DISS. ETH NO. 26627

Clonal selection and immunodominance in the human T cell response to microbial antigens and biologicals

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

ANTONINO CASSOTTA

MSc. Medical Biotechnology
University of Rome "La Sapienza", Italy

Born on 08.02.1989
in Rome, Italy

accepted on the recommendation of

Prof. Dr. Federica Sallusto (examiner)
Prof. Dr. Antonio Lanzavecchia (co-examiner)
Prof. Dr. Annette Oxenius (co-examiner)

Zurich, 2020

A mio zio Giorgio,
della vita innamorato.

Table of contents

1. Acknowledgements

Queste quattro righe sghembe sono per ringraziare tutti coloro che mi hanno accompagnato in questo lungo, tortuoso ma indimenticabile cammino, che a più riprese si è intrecciato con la vita personale fino a rendere difficile distinguere lavoro e piacere.

A Federica e Antonio, che nel mio percorso di dottorato sono stati fari nella notte, fiaccole luminose con cui evitare i precipizi ed esplorare gli spazi del possibile e della conoscenza. È stato per me un onore aver avuto la possibilità di lavorare fianco a fianco con persone e scienziati di questo calibro, ed è stata una magnifica esperienza che mi ha segnato profondamente e porterò sempre con me.

A mamma e papà, a mio fratello Giovanni, ai miei nonni, ai miei cugini, ai miei zii, a tutta la mia famiglia sparsa per l'Italia. A voi che siete stati le mie gambe e le mie braccia, la mia fibra e il mio nervo in quello che ricorderò come il periodo più difficile della mia vita. Ci siete sempre stati, a sostenermi e darmi tenacia anche quando le forze mancavano e lo sguardo si offuscava, senza mai chiedere nulla in cambio: vi voglio bene, e non sarà mai abbastanza ripeterlo.

Ai miei amici, a voi che siete stati musica e risate, rifugio e spensieratezza: Philipp, Eric, Federico, Matteo, Concetta, Sabrina, Mimmo, Paola, Marisa, Serena disasters, Benny, David, Fernando, Gaia, Giada, Francesco, Chiara e tutti i bailadores, i Ciccibelli sparsi per il mondo, Emilio, Vincenzo, il lab gallinaceo (way) too much. Ringrazio ciascuno di voi per aver sempre avuto un sorriso da regalarmi, e per avermi fatto vivere con più leggerezza questo carosello.

A Daniela, che quando iniziai intimorito a muovere i primi passi in lab mi ha mostrato il sentiero sicuro, investendo il suo tempo e le sue energie nella mia formazione sempre con il sorriso in volto e tanta dedizione.

Al resto dei miei compagni di viaggio, che chiamarli solo colleghi sarebbe riduttivo: Luca, Joshua, Dominik, Maria Cristina, Jun, Tobias, Marc, Sara, Samuele, Mengyun, Yiwei, Sandra, Alessia, Corinne, Marco, Laurent, Diego, Francesca, Jeremie, Greta, Daniela P, Alex, Blanca, Chiara, Mathilde, Isa, Simone, Daniela V, Josipa, Michael, Marcus.

I wish to thank also Annette Oxenius, for being part of my thesis committee and positively supporting the completion of my doctoral studies, and Silvia Monticelli for the time invested in organizing the PhD lecture course and for the (lot of) patience she had with me.

Alle lacrime, alle gioie, al sudore, alle risa, alle insicurezze e alle piccole e grandi soddisfazioni, che mi hanno insegnato a mettere a fuoco ciò che è importante nella vita, e cosa la rende così meravigliosa. Agli sbagli che ho fatto e che farò, perché ancor di più mi faranno crescere. Il dottorato è stato il periodo più intenso, difficile e costruttivo della mia vita, professionale e personale. Ho incrociato tante belle persone in questo cammino: molte ancora mi accompagnano e lo faranno a lungo, a prescindere dal tempo e dallo spazio che ci separa.

Qui vi celebro, qui vi ringrazio e vi dedico un pensiero, perché sento di essere la somma delle vostre esperienze, del vostro calore, della vostra impronta.

Bellinzona, Gennaio 2020

Antonino Cassotta

2. General summary

2.1 Summary (English)

CD4$^+$ T cells are necessary for the induction and sustainment of antibody and CD8$^+$ T cell responses, and therefore can be considered as the orchestra conductors of the adaptive immune system. In this thesis, I analyzed the role of CD4$^+$ T cells in the development of neutralizing anti-drug antibodies, in the response against influenza A virus, and in the recognition of commensal and pathogenic bacteria of the Enterobacteriaceae family.

Monoclonal antibodies, and more in general biological drugs, are a very powerful class of therapeutic agents that is increasingly developed and used in clinics. Focusing on the immune response to natalizumab, a humanized monoclonal antibody used in the treatment of multiple sclerosis, we describe an integrated approach to profile the immunogenic landscape of therapeutic antibodies, and we delineate a pathway for the engineering of biobetters endowed with lower immunogenicity and higher efficacy.

Influenza viruses are responsible every year of significant economic loss, and severe morbidity and mortality in certain categories of patients. Moreover, the highly evolving nature of influenza viruses poses serious risks of pandemics and potential transmission of highly pathogenic strains from zoonotic hosts. By analyzing antigen presentation and CD4$^+$ T cell recognition of influenza hemagglutinin, we dissect the impact of antigen processing on T cell clonal selection and immunodominance, and we propose methods to improve in silico predictors of CD4$^+$ T cell responses to vaccination or infection.

Enterobacteriaceae is a large family of bacteria that includes both human commensals and highly pathogenic microbes, the latter causing serious morbidity and mortality especially in underdeveloped countries. Furthermore, the increasing spread of multi-antibiotic resistant strains poses serious threats also in developed countries, making the improvement of therapeutic options a top priority. By interrogating the T cell repertoires of healthy donors and septic patients, here we offer a detailed description of the CD4$^+$ T cell response to Enterobacteriaceae in terms of phenotype, TCR Vβ sequences, antigen specificity and cross-reactivity.

2.2   Riassunto (Italiano)

I linfociti T CD4[+] sono necessari per l'induzione e il sostegno delle risposte anticorpali e CD8[+] citotossiche, e dunque possono essere considerati come i direttori d'orchestra del sistema immunitario adattativo. In questa tesi ho analizzato il ruolo delle cellule T CD4[+] nello sviluppo di anticorpi neutralizzanti contro agenti terapeutici biologici, nella risposta antivirale contro influenza A, e nel riconoscimento di batteri commensali e patogeni appartenenti alla famiglia delle Enterobacteriaceae.

Gli anticorpi monoclonali, e piu' in generale i farmaci biologici, sono una classe di agenti terapeutici in costante sviluppo ed utilizzo in clinica. Focalizzandoci sulla risposta immunitaria al natalizumab, un anticorpo monoclonale umanizzato usato per il trattamento della sclerosi multipla, qui descriviamo un approccio integrato per la caratterizzazione del profilo immunogenico degli anticorpi terapeutici, e delineiamo una strategia per ingegnerizzare versioni migliorate di tali farmaci dotate di minor immunogenicità e dunque aumentata efficacia.

I virus influenzali sono responsabili ogni anno di significative perdite economiche, e di severa morbosità e mortalità in alcune categorie di pazienti. Inoltre, la natura altamente mutevole di tali virus pone seri rischi di pandemie e di potenziale trasmissione di ceppi altamente patogeni da ospiti animali. Analizzando la presentazione dell'emoagglutinina influenzale e il suo riconoscimento da parte delle cellule T CD4[+], in questo studio valutiamo l'impatto della processazione dell'antigene sulla selezione clonale e immunodominanza dei linfociti T, e proponiamo metodi per migliorare i predittori informatici delle risposte T CD4[+] a vaccinazioni e infezioni.

Le Enterobacteriaceae sono una vasta famiglia batterica comprendente sia commensali dell'uomo che microbi altamente patogeni, quest'ultimi causa di severa morbosità e mortalità specialmente nei paesi sottosviluppati. Inoltre, la crescente diffusione di ceppi resistenti a molteplici antibiotici costituisce una grave minaccia anche nei paesi sviluppati, rendendo il miglioramento delle opzioni terapeutiche una grande priorità. Mediante l'analisi dei repertori linfocitari di donatori sani e pazienti con sepsi, con il nostro lavoro offriamo una descrizione dettagliata della risposta immunitaria delle cellule T CD4[+] alle Enterobacteriaceae in termini di fenotipo, sequenze Vβ del TCR, specificità antigenica e cross-reattività.

3. List of abbreviations

APC	Antigen-presenting cell
BCR	B cell receptor
BCL-6	B cell lymphoma 6
BfA	Brefeldin A
CCR	C-C chemokine receptor
CD	Cluster of differentiation
CDR	Complementarity-determining region
CFSE	Carboxyfluorescein succinimidyl ester
CLIP	Class II associated Ii peptide
CTLA-4	Cytotoxic T-Lymphocyte Antigen 4
CXCR	CXC-chemokine receptor
DC	Dendritic cell
EC50	Half maximal effective concentration
ER	Endoplasmic reticulum
FcγR	Fcγ receptor
GATA3	GATA binding protein 3
HLA	Human leukocyte antigen
ICOS	Inducible costimulator
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
ITAM	Immunoreceptor tyrosine-based activation motif
LT	Lymphotoxin
MHC-II	Major histocompatibility complex class II
MIIC	MHC-II compartment
MIC	MHC class I polypeptide-related sequence
MTB	*Mycobacterium tuberculosis*
PAMP	Pathogen-associated molecular pattern
PMA	Phorbol myristate acetate
PBMC	Peripheral blood mononuclear cell
pMHC-II	Peptide-MHC-II complexes
RORγt	RAR-related orphan receptor gamma, isoform 2

SHM   Somatic hypermutation

TAP   Transporter associated with antigen processing

TAPBP   Tapasin

T-bet   T-cell-specific T-box

TNF   Tumor necrosis factor

Tregs   T regulatory cells

TCR   T cell receptor

Tcm   T central memory cells

Tem   T effector memory cells

Tfh   T follicular helper cells

Th   T helper cells

SAP   SLAM-associated protein

V(D)J   Variable, diversity, joining

4. General introduction

The immune system is composed of a highly integrated network of cells and molecules that have evolved to ensure host protection from infectious diseases. Innate immunity provides an immediate line of defense against invading pathogens, and is enforced by many cell types such as phagocytes (neutrophils, macrophages), dendritic cells, and natural killer cells (NK). The cellular components of the innate immune system have generic receptors for structures conserved on different classes of pathogens (known as PAMPs, pathogen-associated molecular patterns), that can trigger inflammatory response to limit pathogen invasion (Cooper and Alder, 2006). Innate immunity can therefore sense the invading pathogens in a broad, yet aspecific fashion.

The adaptive immune system, also known as specific or acquired immunity, instead, ensure a protective response that "adapts" the host to infections to prepare it in case of future challenges, in an highly specific manner for each microbe since their recognition is mediated by an extremely diverse repertoire of somatically rearranged antigen receptors, which are acquired postnatally (Abbas et al., 2012). The hallmark of the adaptive immune system is the ability to learn from experience and remember a previous antigen exposure, a property which is known as immunological memory. In this way upon subsequent encounter with the same pathogen, the immune system will respond more rapidly and vigorously, principles that lays at the basis of vaccinations.

The adaptive immune system comprises B lymphocytes, which mediate humoral immunity through production of antibodies, and T lymphocytes, which mediate cellular immunity. The latter include $CD8^+$ T cells that possess cytotoxic activity towards virus-infected or cancerous cells, and $CD4^+$ T helper cells that are necessary for the induction, promotion and regulation of the whole process by various means such as cytokine production and providing help to B cells. Key features of protective adaptive immune reactions are the ability to specifically recognize the antigen, the priming of a mature immune response against the antigen, and the development of immunological memory in order to confer long lasting protection. The antigen recognition by B and T cells is mediated by specific receptors that are expressed after a process of somatic gene rearrangement that leads to the acquisition of unique antigen receptors on the cell surface, named B-cell (BCR) and T-cell receptor (TCR), respectively.

B and T cells recognize antigens in very different forms: while B cells are able to recognize antigens in their native tridimensional structures, CD8$^+$ and CD4$^+$ T cells classically sense antigens in the form of small linear peptides that are exposed in complex with MHC-I and MHC-II molecules, respectively. Therefore, the activation of T lymphocytes depends on the processing and presentation of the antigen, operated by antigen presenting cells (APCs) in order to mount effective cellular responses.

In the next paragraphs I will introduce some principles at the basis of antigen processing and presentation on MHC-II molecules, of peptide-MHC-II-recognition by CD4$^+$ T cells, and of T-cell and B-cell collaboration.

4.1   Presentation of the antigen to CD4$^+$ T cells

As mentioned above, for antigen recognition CD4$^+$ T cells rely on the interaction with antigen presenting cells (APCs). Classically, for the induction of a primary CD4$^+$ T cell response to a foreign protein antigen, phagocytic APCs such as dendritic cells take up the antigen and process it in the endosomal compartment to small peptides that gets accommodated in the binding cleft of major histocompatibility complex class II (MHC-II) molecules. Once loaded on MHC-II molecules, these small peptides are exposed to the cell surface in order to allow their recognition by specific CD4$^+$ naïve T cells patrolling secondary lymphoid organs. Moreover, presentation of peptides in the context of MHC-II molecules is fundamental also to trigger the effector function of CD4$^+$ T helper cells, to provide for example help to B cells for antibody affinity maturation and class switching, or to activate phagocytic cells in peripheral tissues.

4.1.1   Professional antigen presenting cells

While MHC class I molecules are constitutively expressed on all nucleated cells, MHC class II molecules are constitutively expressed on professional APCs like dendritic cells (DCs), macrophages and B cells (Neefjes et al., 2011), although under inflammatory conditions (such as exposure to IFN-γ) they can be upregulated in many cell types, such as epithelial and vascular endothelial cells (Schroder, 2016). Nevertheless, the MHC-II expression itself is not always sufficient for efficient antigen presentation to CD4$^+$ T cells: indeed, to maximize the probability of antigen

presentation to specific T cells, professional APCs have evolved several mechanisms to efficiently deliver exogenous and endogenous antigens, either self or foreign, into pathways that leads to their presentation on the cell surface in the context of MHC molecules (Roche and Furuta, 2015).

Antigens can enter the endocytic pathway in a multitude of ways, and each APC type has evolved one or more specialized ways to take up, concentrate and present antigens to CD4[+] T cells. Fluid phase macropinocytosis is a nonspecific process of internalization of extracellular material that allow immature DCs to continuously internalize and concentrate large volumes of fluid, in order to efficiently sample antigens in peripheral tissues (Sallusto et al., 1995). Activation of DCs leads to reduction in macropinocytosis activity and accumulation of long-lived peptide-MHC-II (pMHC-II) complexes on the cell surface (Cella et al., 1997), with a concomitant induction of costimulatory molecules (Sallusto et al., 1995) that convert DCs from immature antigen-sampling cells to mature antigen-presenting cells (Roche and Furuta, 2015).

Alongside non-specific antigen uptake by macropinocytosis, APCs can express also several cell surface receptors to facilitate antigen entry, such as lectin receptors, Fcγ receptors (FcγRs), complement receptors and scavenger receptors, that allow the interaction with native or IgG-opsonized microbes and apoptotic cells, therefore mediating receptor-mediated endocytosis or phagocytosis (Roche and Furuta, 2015). Moreover, B cell receptors represent an extremely sensitive way to concentrate antigens by receptor-mediated endocytosis: antigen-specific B cells are indeed 10^4-times more efficient in presenting antigens from the external microenvironment to cognate T cells, compared to B cells with irrelevant specificity (Lanzavecchia, 1985).

Antigen presentation by different type of APCs leads to different outcomes of the CD4[+] T cell response, that is function of the different mechanisms of antigen uptake, the nature of the antigen processing compartments and the lifetime of cell surface peptide-MHC-II complexes (**Fig. 4-1**) (Roche and Furuta, 2015). Dendritic cells represent the quintessential professional antigen-presenting cells for induction of primary CD4[+] T cell responses: DCs can sample antigens by several mechanisms, they can upregulate costimulatory molecules needed for proper naïve T cell priming, while concomitantly

migrating in the T cell zone of secondary lymphoid organs, therefore maximizing the chances of encounter and activation of antigen-specific CD4[+] naïve T cells patrolling the same sites. Macrophages and B cells instead mostly function as APCs for previously activated CD4[+] T helper cells, rather than naïve T cells, to trigger their effector functions. In case of infection, macrophages in peripheral tissues can present peptides derived from phagocyted microbes to CD4[+] helper T cells, leading to their activation and production of effector cytokines, such as IFN-γ that enhance the microbicidal activities of macrophages. For the deployment of protective antibody responses, antigen-specific B cells need to present peptides derived from the internalized antigen to CD4[+] T follicular helper (Tfh) cells, which in turn provide help for proper B cell activation, Ig class-switching and antibody affinity maturation.
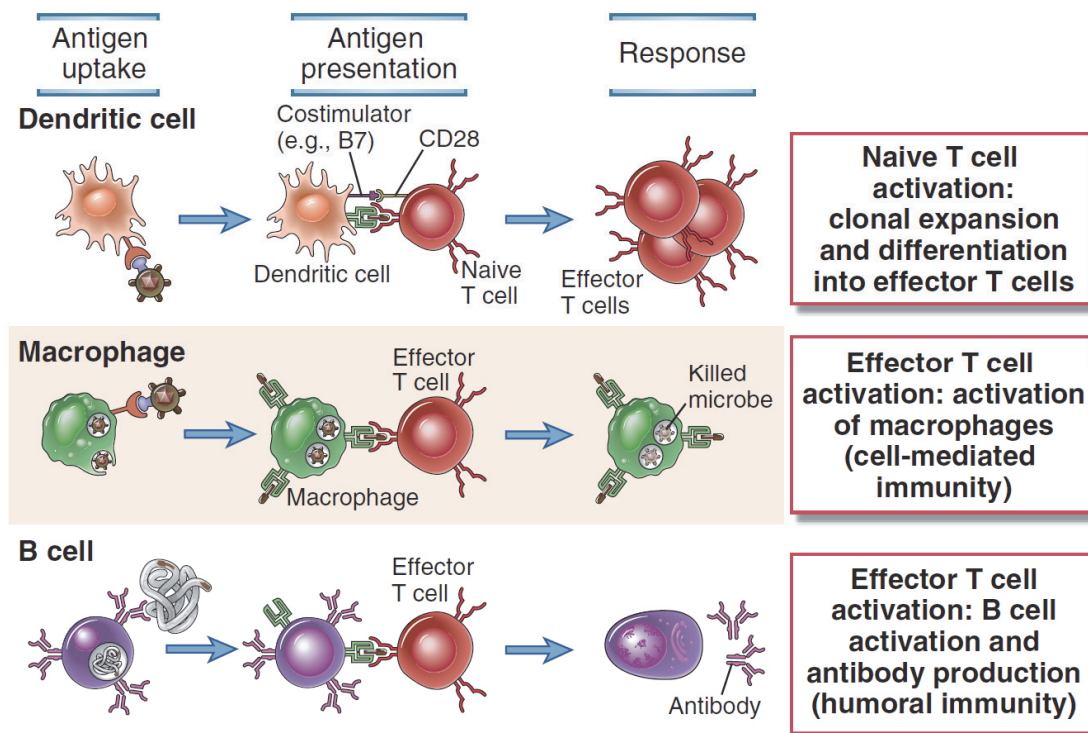
Figure 4-1. Functions of different APCs interacting with CD4[+] T cells by pMHC-II - TCR engagement. From (Abbas et al., 2012).

4.1.2 Genetic diversity of MHC-II locus

Genes of the major histocompatibility complex (MHC) encode surface glycoproteins that play the fundamental role of presentation of short peptides derived from foreign antigens or self-proteins to T lymphocytes. Those molecules are highly polymorphic and are responsible for alloreactivity and graft rejection between non syngeneic individuals (Abbas et al., 2012). In human, MHC genes are named HLA (Human Leukocyte Antigen, so called because the first MHC molecules were identified on the cell surface of leukocytes) and are localized in the short arm of chromosome 6 (region 6p21.3). The HLA locus is organized in 3 sub-regions from centromere to telomere, namely class II, class III and class I, which contains a total of 224 gene loci encoding for an estimate of 128 proteins, of which around 40% have functions closely related to the immune system (Beck et al., 1999).

MHC class II region contains the genes for HLA-DP, -DQ and -DR molecules, whose products are expressed by professional APCs such as DCs, macrophages and B lymphocytes, and can be upregulated by certain cell types such as activated T lymphocytes and endothelial cells (Abbas et al., 2012). Other important genes that map in this regions are HLA-DM and HLA-DO, molecular chaperons involved in the peptide loading and editing of MHC-II-presented peptidome (Mellins and Stern, 2014), as well as genes involved in the MHC-I presentation pathway, such as the transporter associated with antigen processing (TAP), tapasin (TAPBP) and the IFN-γ-inducible LMP2 and LMP7 subunits of the immunoproteasome (Basler et al., 2013). Complete MHC-II molecules are formed by the dimerization of an alpha and a beta chain, whose association allows the constitution of the peptide-binding cleft (**Fig. 4-2a**). In humans the gene for alpha chain of HLA-DR (HLA-DRA) is monomorphic, whereas the beta chains are encoded by HLA-DRB1, HLA-DRB3, HLA-DRB4 or HLA-DRB5 genes. While the highly polymorphic HLA-DRB1 gene is present in all individuals, HLA-DRB3, 4 and 5 alleles are present only in some haplotypes, linked with certain DRB1 alleles (Shiina et al., 2009). As a consequence, some individuals co-express different HLA-DR isoforms formed upon dimerization of alpha chain with gene products encoded by both HLA-DRB1 and HLA-DRB3, 4 or 5. In the case of HLA-DQ and HLA-DP molecules, instead, both the genes encoding alpha and beta chain are highly polymorphic, leading to even higher diversity (**Fig. 4-2b**).

The MHC class III region contains genes encoding components of the complement system (C2, C4 and factor B), and genes encoding the cytokines tumor necrosis factor α (TNF-α) and lymphotoxin (LT) α e β, as well as many genes with no obvious implications with immune function.

The most telomeric region of the HLA locus contains HLA-A, -B and -C genes encoding the classical MHC-I molecules constitutively expressed by all nucleated cells, as well as the genes of non-classical MHC-I molecules, namely HLA-E, -F and -G, which are structurally related to classical MHC-I molecules but have much lower polymorphisms. Moreover, in this region reside also genes encoding MIC-A and MIC-B, stress-induced molecules related to MHC-I that do not associate to β_2-microglobulin nor bind peptides, but can interact with NKG2D receptor expressed by NK cells, γδ T cells or CD8$^+$ αβ T cells (Lanier, 2015).

The genes encoding HLA molecules are the most polymorphic genes of the human genome, with more of 17,000 and 6,700 alleles reported to date for class I and class II, respectively (Robinson et al., 2015), and with most of the polymorphisms leading to nonsynonymous amino acid changes in the peptide-binding groove of the mature HLA molecules (Dendrou et al., 2018). Moreover, usually both parental alleles of each MHC gene are co-dominantly expressed, and after translation polymorphic alpha and beta protein chains of HLA-DP and HLA-DQ can form heterodimers both by cis- and trans-haplotype pairing. Taken together, the polygenic nature of HLA locus, the high polymorphism and the co-dominant expression result in an extremely high inter-individual diversity in terms of mature MHC protein isoforms expressed on the cell surface of APCs. Such high diversity and complexity of the MHC system likely evolved to maximize the probability that at least some individuals within the general population can efficiently mount immune responses to newly encountered infections and survive (Dendrou et al., 2018).

4.1.3   Structure of pMHC-II molecules

Mature MHC-II molecules structurally consist of an extracellular peptide-binding cleft, followed by immunoglobulin (Ig)-like domains, a transmembrane domain and a cytoplasmic tail (**Fig. 4-2a**). The dimerization of alpha and beta chains allows the

correct folding of the extracellular peptide-binding groove at the top of the molecule, that is composed of an eight-stranded β-sheet floor delimited by two α-helices, where the peptide sits between the two helices (Unanue et al., 2016). The majority of HLA polymorphisms (with the exception of HLA-DR alpha gene) are contained in the exon 2 of alpha and beta chains, that encodes the peptide binding cleft (**Fig. 4-2b**), whereas the non-polymorphic Ig-like domains contain binding sites for the CD4 co-receptor of T cells.

Peptides bound to MHC-II cleft assume a stretched conformation in which some amino acid side chains point toward the groove, whereas some others are solvent exposed and therefore potentially recognizable by the T cell receptor (TCR). Binding of the peptide to MHC-II cleft relies on hydrogen bonding of the peptide backbone with the α-helices, and on chemical interactions of certain residues of the peptide, termed anchor residues, with pockets buried deep in the floor of the binding groove that act as acceptor sites. The peptide binding groove of MHC-II molecules have open ends compared to MHC-I molecules: this allows greater flexibility in the length of bound peptides, which are typically 14-20 amino acid long and extend out of the MHC-II structure (Mohan and Unanue, 2012; Unanue et al., 2016). Peptides bound to MHC-II have a central 9-amino-acid core that is stretched within the MHC-II groove, and is known as peptide binding register. The latter contains 3-4 anchor residues, usually at position P1, P4, P6/7 and P9, which interact with residues of the pockets of MHC-II groove. Conversely, the P2, P5 and P8 residues of the core are surface exposed, thus being potential TCR contact sites; the flanking residues that are outside the binding groove can help in stabilizing the peptide-MHC-II interaction and can influence the topology of TCR contact residues (Mohan and Unanue, 2012).

As anticipated, the residues of the MHC-II pockets responsible for binding the peptide within the groove are highly polymorphic: hence, depending on the amino acid composition of the peptide binding cleft, different HLA alleles can bind sets of peptides endowed with anchor residues of slightly different, yet defined, biochemical characteristics. An important implication of the bio-chemical properties that governs peptide binding to HLA groove is that MHC molecules do not discriminate between self and non-self peptides, binding and therefore presenting both foreign peptides (e.g. derived from microbial proteins) and peptides derived from proteins of the individual.

The discrimination between self and non-self is therefore remitted to downstream mechanisms of tolerance that avoid the generation or activation of strongly self-reactive T lymphocytes.
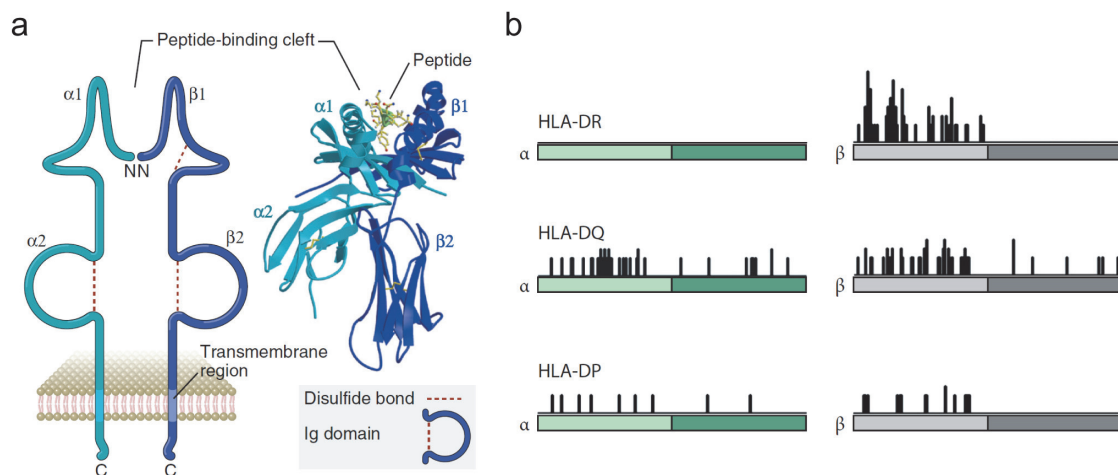
Figure 4-2. Structure and genetic diversity of human MHC-II molecules. (**a**) Schematic diagram of the different domains of MHC-II molecules *(left)* and crystal structure of the extracellular portion of HLA-DR1 molecule with a bound peptide *(right)*. Adapted from (Abbas et al., 2012). (**b**) Location of polymorphic residues in MHC-II molecules. Each α and β chains of HLA-DR, HLA-DQ, and HLA-DP molecules are split into α1, α2 and β1, β2 domains reported in light and dark colors, respectively. Histograms of polymorphic amino acids show that allelic diversity is mainly localized within and around the peptide-binding cleft of MHC-II. Adapted from (Unanue et al., 2016).

4.1.4   MHC-II presentation pathway

Protein antigens taken up by professional APCs can follow a multistep pathway that involves their processing in the endosomal compartments, and subsequent exposure on the cell surface in the form of small peptides bound to the peptide-binding cleft of MHC-II molecules, where they are physically presented for recognition by T cells (**Fig. 4-3**). Alpha and beta chains of MHC-II molecules are synthetized in the ER and associate with the invariant chain CD74, a molecular chaperone that facilitates both the correct folding and the subcellular trafficking of the newly synthetized MHC-II molecules towards an acidic late endosomal compartment named the MHC-II compartment (MIIC) (Neefjes et al., 2011; Schroder, 2016). In the MIIC compartment, invariant chain associated with the newly translated MHC-II is digested by cathepsin proteases except for a small peptide termed class II associated Ii peptide (CLIP) that is protected from proteolysis when embedded in the peptide-binding groove of the MHC-

II alpha-beta dimer. In the meantime, progressive pH acidification of the endosomes ensures unfolding of phagocyted proteins and activation of catalytic activity of proteases such as cathepsins, that start to cut and digest the proteins to small peptides that can potentially be accommodated in the MHC-II binding cleft. Once liberated by proteolytic cut, peptides endowed with sufficient binding affinity for the MHC-II cleft can replace the CLIP fragment, a process that is facilitated by the action of HLA-DM molecules. When the MHC-II alpha beta dimer is loaded with a peptide in the binding cleft, the heterotrimeric pMHC-II complex is transported to the plasma membrane so that the peptide cargo is exposed on the cell surface space and can potentially be recognized by antigen-specific CD4$^+$ T cells.

The HLA-DM molecule is a MHC-II homolog expressed in the MIIC compartment that acts as a peptide exchange factor, thus being able to remove peptides bound to MHC-II cleft with weak affinity or unstable conformations. The HLA-DM editing of MHC-II peptidome is counterbalanced by HLA-DO molecules, that act as substrate mimics that competitively inhibit HLA-DM-mediated catalysis of MHC-II peptide exchange (Mellins and Stern, 2014). By "cleaning up" the MHC-II peptidome from unstable conformers, HLA-DM ensures that only the most stable peptide-MHC-II complexes are exposed on the cell surface for presentation to CD4$^+$ T cells (Mellins and Stern, 2014; Mohan and Unanue, 2012). If peptide acquisition by MHC-II molecules happens in sub-cellular compartments devoid of HLA-DM, it will lead to a broad MHC-II-presented peptidome in which also sub-optimal binders and conformers are allowed to be exposed on the cell surface of APCs. This is the case, for instance, of soluble peptides (such as peptide hormones like insulin) that can get loaded into the cleft of MHC-II directly at the cell surface or on recycling MHC-II molecules in the early endosomes of APCs. The "maturation" of the MHC-II peptidome by HLA-DM has important implications for autoimmunity: since thymic education happens in the presence of HLA-DM editing, it spares from negative selection T cell clones potentially able to recognize self-peptides presented on MHC-II as conformers normally removed by HLA-DM. Those cells are termed "type B" CD4$^+$ T cells, to distinguish them from the conventional "type A" T cells that recognize HLA-DM-matured pMHC-II complexes: if in the peripheral tissues their cognate self-peptides get presented in a HLA-DM-independent manner, self-reactive "type B" T cells can potentially get

activated and therefore lead to autoimmunity (**Fig. 4-4**) (Mohan and Unanue, 2012; Unanue et al., 2016).
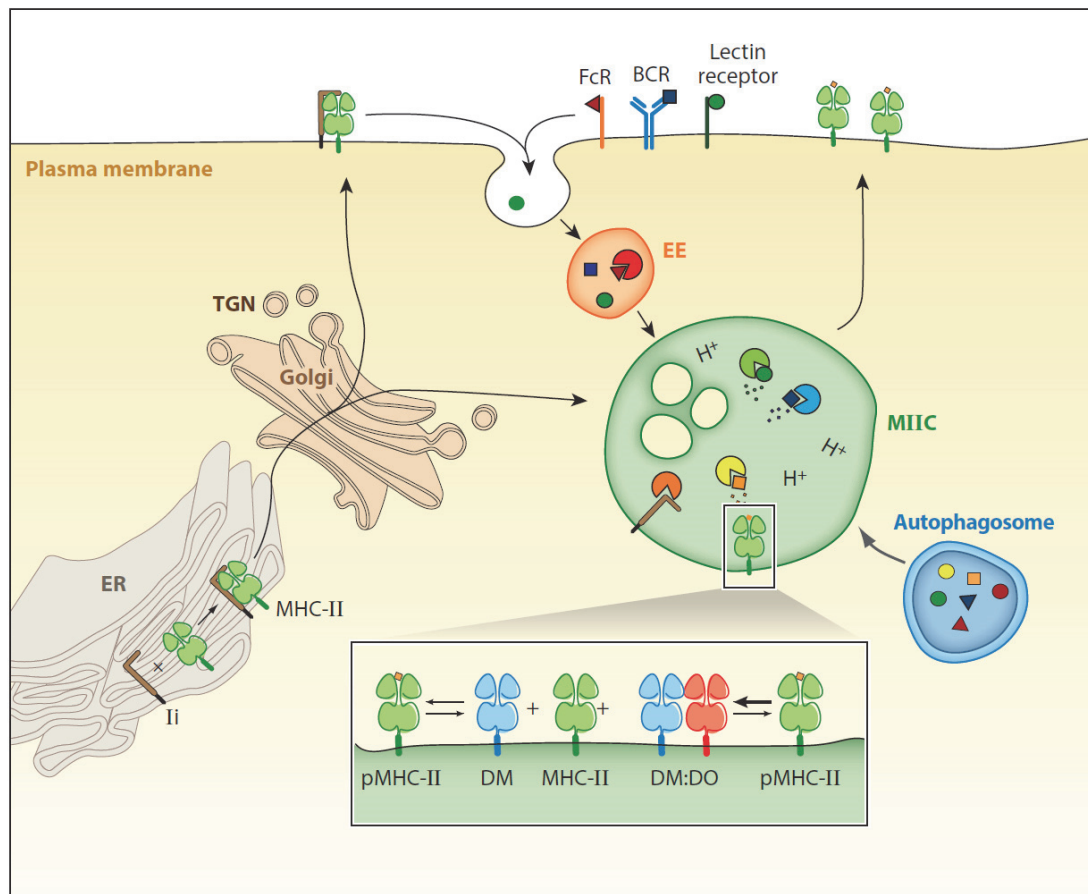
Figure 4-3. General scheme of MHC-II antigen presentation pathway. Abbreviations: BCR, B cell receptor; EE, early endosome; FcR, Fc receptor; PM, plasma membrane; pMHC-II, peptide-loaded MHC-II; TGN, trans-Golgi network. From (Unanue et al., 2016).
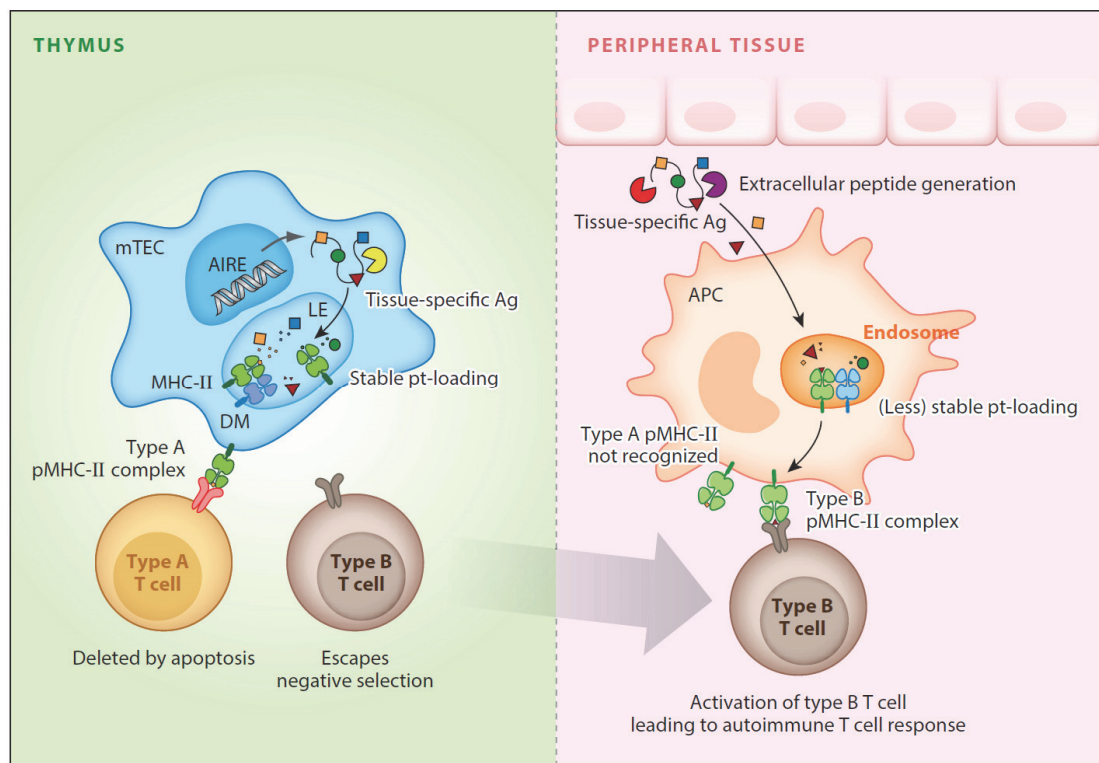
Figure 4-4. Model of evasion from central tolerance by self-reactive "type B" CD4$^+$ T cells *(left)*, that could provoke autoimmunity upon recognition in periphery of pMHC-II complexes not edited by HLA-DM *(right)*. Abbreviations: Ag, antigen; APC, antigen-presenting cell; LE, late endosome; pMHC-II, peptide-loaded MHC-II; pt, peptide. From (Unanue et al., 2016).

4.2   CD4$^+$ T helper cells and B lymphocytes: a tight collaboration

CD4$^+$ T helper lymphocytes play a fundamental role in orchestrating the adaptive immune response to foreign antigens: once antigen-specific CD4$^+$ T cells are primed by DCs in secondary lymphoid organs, they can migrate to follicles to facilitate the induction and sustainment of the antibody response by B cells, and to peripheral tissues where they can deliver effector cytokines that boost immune defense mechanisms (Sallusto and Lanzavecchia, 2009). Moreover, T helper cells are required also for the efficient induction of CD8$^+$ T cytotoxic responses, since they mediate the "licensing" of DCs for the activation of killer CD8$^+$ T cells against poorly immunogenic antigens such as protein antigens. In particular, antigen-stimulated CD4$^+$ T cells can upregulate CD40L and thus deliver a positive signal to CD40-expressing DCs: once activated by the interaction with T helper cells, DCs acquire the ability to prime potent cytotoxic CD8$^+$ T cell responses (Lanzavecchia, 1998; Sallusto et al., 2010).

4.2.1 Generation of a rich and clonally diverse repertoire of receptors for the antigen

T cells are generated early in life in the thymus, a bilobate organ that enlarges during childhood and undergoes atrophy at puberty. In the thymus T cells acquire a unique antigen receptor by genetic recombination of the TCR genes localized in chromosome 7 (TCR beta and gamma loci) and chromosome 14 (TCR alpha and delta loci), a process known as V(D)J recombination. Moreover, T cells undergo a complex process of selection by which lymphocytes that rearranged TCRs able to interact with self MHC molecules are positively selected over those that are unable to do so. At the same time, T cells carrying TCRs that confer strong reactivity to self-derived peptides are negatively selected by deletion or conversion to regulatory T cells (Tregs) in a process called central tolerance. The final outcome of the thymic education is the establishment of an extremely rich and clonally diverse T cell repertoire in which each lymphocyte carries a unique antigen receptor (Cooper and Alder, 2006) that is restricted to self MHC molecules to allow efficient recognition of the MHC-presented peptides, and somehow tolerant to self in order to avoid strong autoimmunity, at least in pre-reproductive age (Davis and Brodin, 2018).

The alpha beta TCR is a heterodimer formed by the association of transmembrane alpha and beta chains covalently linked by a disulfide bridge: the apical domains of the extracellular portion of the TCR are responsible for antigen recognition, and are known as variable regions. The variable region of alpha and beta chains has an Ig-like folding consisting of two adjacent β-sheets held together by an intra-chain disulfide bond, with three protruding peptide loops known as complementarity-determining regions (CDRs). CDRs of alpha and beta chains can make contacts with peptide-MHC complexes, and are the sites where the variability between different TCRs is concentrated. While CDR1 and CDR2 are somatically encoded in each Variable (V) gene, CDR3 sequences originate as new, not germline-encoded sequences during the random and imprecise rearrangement of the V, D and J segments of TCR beta genes, and V and J segments of the TCR alpha. Due to the mechanism of V(D)J recombination, CDR3 loops of TCRs holds the greatest variability: the number of different amino acid sequences in CDR3s that results from somatic recombination is indeed much higher of the number of CDR1 and CDR2 sequences that can be possibly encoded by germline V gene segments,

leading to a theoretical upper limit of 10^{16} possible alpha-beta TCRs (Abbas et al., 2012). Nevertheless, the actual size of the naïve CD4$^+$ T cell compartment of each individual is estimated in 100 million unique TCR Vβ sequences for young adults, repertoire that remains highly diverse during aging despite the occurrence of thymic involution, resulting in a modest two- to fivefold contraction at the seventh decade of life (Qi et al., 2014). The high richness of the naïve repertoires is an important feature to maximize the chances of the acquired immune system to recognize the largest number of antigens as possible, and ensure protection to the myriad of possibly harmful pathogens that the host will encounter during its life-time.

4.2.2 CD4$^+$ T cell recognition of the antigen

The recognition of the antigen involves dramatic phenotypical changes in T lymphocytes, such as the suppression of cell mobility, the organization of a specialized cell-cell junction with the APC, and T cell polarization. TCR engagement require the formation of a supramolecular complex of proteins that stabilize the interaction with the APC across a 15-nm gap space. This interface is known as the immunological synapse, and involves the TCR complex and CD4 co-receptor, adhesion molecules and costimulatory or coinhibitory (checkpoint) receptors (Dustin, 2014).

The TCR interaction with peptide-loaded MHC molecules controls the specificity of the response: once the peptide is exposed on the MHC-II binding cleft, the heterotrimeric pMHC-II complex acts as docking site for the six CDR loops of the alpha and beta chains of the TCR, which allows the scanning of the peptide cargo by CD4$^+$ T cells until the establishment of a cognate interaction. There are multiple contact sites in this interface (**Fig. 4-5**): structural biology has revealed that the germline-encoded CDR1 and CDR2 loops preferentially interact with the conserved α-helices of the MHC molecules, whereas the somatically recombined CDR3s mainly make contacts with the surface exposed residues of the bound peptide (Garcia and Adams, 2005). Interestingly, whereas CDR3 of beta chain is always found in close proximity to the antigenic peptide, in some cases CDR3 of alpha chain makes no contact, suggesting that the former is necessary and the latter sometime can be dispensable for antigen recognition (Glanville et al., 2017). Overall, the TCR interaction with α-helices at the boundaries of the binding cleft is responsible for the MHC restriction of the T cells, whereas the contact

with the bound peptide ensure the specificity of antigen recognition. It is important to notice that despite such fine contacts, the interaction of TCRs with pMHC-II complexes are short-lived and of much lower affinity compared to the binding of antibodies to their cognate antigens. This property might facilitate the serial engagement of a multitude of TCRs by each peptide-MHC-II complex, and the rapid screening by many different TCRs to favor a sensitive detection of the small amount of antigenic determinants presented at each moment on the surface of APCs (Huppa and Davis, 2013; Valitutti et al., 1995).

Whereas the extracellular portion of the TCR ensures specificity, the triggering of the intracellular signaling upon antigen recognition relies on other proteins non-covalently associated with the TCR alpha-beta chains to form the TCR complex. Those comprise CD3 heterodimers (composed of ε chain associated with γ or δ chain) and a homodimer of ζ-chain, which contains particular motifs in their cytoplasmic tails called immunoreceptor tyrosine-based activation motif (ITAM) that can be substrates of tyrosine kinases of the Src family such as Lck. The CD4 co-receptor plays a key role in the ignition of the TCR signaling: the extracellular domain of CD4 can bind to residues of the non-polymorphic β2 domain of MHC-II molecules, while the cytoplasmic tails of CD4 associate with Lck tyrosine kinase. In such way, CD4 co-receptor localizes in close proximity of the TCR complex that contacts the same pMHC-II molecule and recruits Lck next to the ITAMs of CD3 and ζ proteins. Upon phosphorylation by Lck, ITAMs of ζ-chain act as docking sites for the tyrosine kinase ZAP-70, that upon activation can phosphorylate adaptor proteins such as LAT that allow great amplification of the signaling cascade.

While TCR and co-receptor ensure the specificity of the T cell activation, adhesion molecules such as the integrin LFA-1 (CD11a/CD18) and the Ig-superfamily member CD2 provide the energy to pull the T cell and the APC together enough to allow sustained antigen recognition (Dustin, 2014). Moreover, costimulatory and coinhibitory receptors are able to amplify or dampen the net balance of the signaling by the TCR complex, respectively, thus altering the functional outcome of the immunological synapse. Costimulatory receptors involve the archetypal CD28 that can recruit and activate Lck and PKC-θ, and plays a fundamental role in naïve T cell priming: CD28 co-stimulatory activity is dependent on the binding to B7.1 (CD80) and

B7.2 (CD86) ligands upregulated on DCs stimulated by maturation signals. Other important costimulatory receptors includes the inducible ICOS, various members of the TNFR superfamily, such as CD27, GITR, 4-1BB and OX40, and members of the lectin-like receptor superfamily such as NKG2D (Dustin, 2014).

Coinhibitory or checkpoint receptors are the negative counterparts of the costimulators, and include several members, such as CTLA-4, PD-1, LAG-3, TIGIT, TIM-3, Vista and others: in recent years, the blockade of functionality of checkpoints such as CTLA-4 and PD-1 by monoclonal antibodies have been demonstrated being a successful strategy to drastically improve cancer immunotherapy (Larkin et al., 2015). CTLA-4 interacts with CD80 and CD86 with higher affinity and avidity than CD28 does, thus competing with the latter for ligand binding and antagonizing CD28-mediated co-stimulation (Rowshanravan et al., 2018). Whereas CD28 is mainly expressed on the plasma membrane, CTLA-4 has a constitutively active internalization motif in its cytoplasmic domain: therefore CTLA-4 from the cell surface undergo continuous endocytosis, such that more than 90% of the protein is localized in intracellular vesicles. In Tregs, CTLA-4 is constitutively expressed and, due to its highly endocytic nature, can operate a physical removal of CD80 and CD86 ligands from surface of APCs, depriving them of co-stimulatory activity (Dustin, 2014; Rowshanravan et al., 2018).
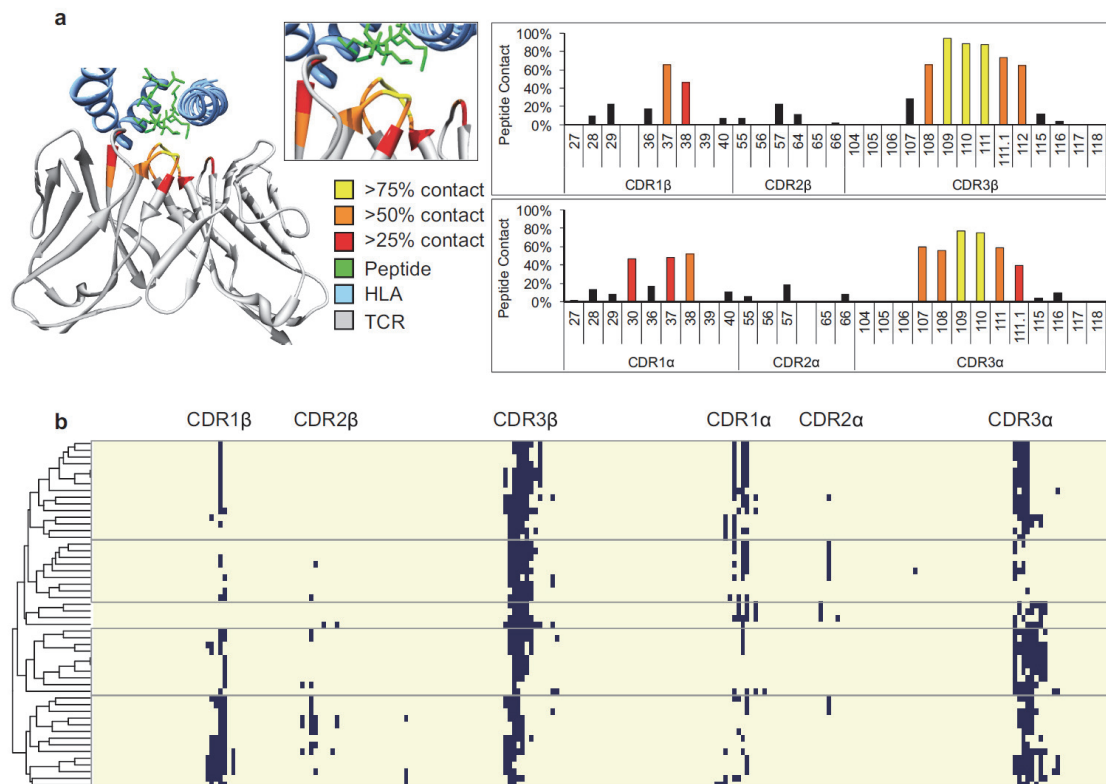
Figure 4-5. Analysis of published crystal structures reveals structural features of TCR-pMHC interaction. CDR residues found within 5 Å of distance from the antigenic peptide were defined as putative epitope contact sites. (**a**) Probabilities of putative contact for each residue of the six CDRs, calculated from 52 published crystal structures of TCR-pMHC interactions. (**b**) Heatmap of putative contact sites from 52 nonredundant TCR sequences from published crystals, clustered into five contact modes according to contact profiles of all six CDRs. From (Glanville et al., 2017).

4.2.3   T cell clonal selection and immunodominance

Each naïve T lymphocyte that arise from the thymus is equipped with a unique, productively-rearranged TCR: in case of recognition of their cognate antigen, reactive naïve T cell clones undergo vigorous clonal expansion and differentiate into effector cells that provide immediate protection, and can survive as memory lymphocytes that can persist for a lifetime (Sallusto et al., 2004). Out of the multitude of potentially immunogenic peptides contained in a complex foreign antigen, only a small fraction is able to induce a measurable T cell response: this phenomenon is known as immunodominance, term that refers to the fact that the T cell response is usually limited to a small proportion of the potential determinants of the protein antigen (Sercarz et al., 1993; Yewdell and Bennink, 1999; Yewdell and Del Val, 2004).

Dominance of T cell epitopes can be defined based on the frequency of response in a population of individuals, or on the magnitude of response in a particular individual following priming with a protein antigen (Landry, 2008). Determinants that elicit the most vigorous immune responses are usually defined as immunodominant, while those inducing weaker responses are referred as subdominant: T cell responses to dominant and subdominant epitopes are both induced, and can be recalled, by the natural processing of the native antigen by APCs. Conversely, the term cryptic epitopes indicates a third group of determinants that are strongly immunogenic only in the form of synthetic peptides or larger proteolytic fragments, but do not induce any T cell response after recall with the native protein, thus indicating inefficient generation of the naturally processed peptides corresponding to such determinants (Adorini, 1998; Sercarz et al., 1993; Yewdell and Bennink, 1999). Many reasons have been proposed to explain crypticity, collectively suggesting that such "hidden" peptides could either not be generated, destroyed or outcompeted by other peptides during MHC-II processing of the native antigen. In this sense, the aforementioned HLA-DM-sensitive pMHC-II conformers that can be recognized by self-reactive "type B" CD4$^+$ T cells represent a particular case of cryptic epitopes, due to the peculiarity of their MHC-II presentation pathway (Mohan and Unanue, 2012; Sercarz et al., 1993).

Understanding the rules that govern immunodominance is of primary importance in order to improve the strategies of vaccine development, by identifying the most immunogenic parts of the antigen. Moreover, the possibility to predict immunogenicity of protein antigens would be very helpful to improve the design of therapeutic proteins such as monoclonal antibodies with reduced immunogenicity, and it could facilitate the monitoring of antigen-specific immune responses in health and disease. T cell responses to foreign antigens are integrated and complex processes, therefore multiple factors can be influencing at different steps the dominance of some epitopes over others, with the establishment of a hierarchy of immunogenicity of the different antigenic determinants. Some of those factors reflect the biochemical rules of the MHC-II processing and presentation pathway in APCs, whereas others hold on the intrinsic architecture of the T cell repertoires and on the mechanisms of antigen recognition and T cell clonal selection (Yewdell, 2006).

MHC-II processing is a cooperative process, so many aspects can affect the composition of the MHC-II peptidome that gets ultimately presented on the surface of APCs. The molecular context in which each peptide is embedded and its structural accessibility might influence the propensity of unfolding during the progressive pH acidification that occurs in the endocytic pathway, therefore affecting the exposition of denatured stretches of the antigen to the proteolytic environment of the late endosomes (Graham et al., 2018; Kim and Sadegh-Nasseri, 2015; Landry, 2008; Mirano-Bascos et al., 2008). Moreover, the set of endocytic proteases with whom different APC types are equipped adds an additional level of complexity: different isoforms of cathepsins have different substrate specificity and their catalytic activity can be tightly regulated by pH and endogenous inhibitors such as cystatins, thus influencing the probabilities of liberation of antigenic peptides from the native protein (Kim and Sadegh-Nasseri, 2015; Unanue et al., 2016). Furthermore, the intrinsic affinity of the liberated peptides for the MHC-II binding cleft and the resistance of newly formed peptide-MHC-II complexes to the HLA-DM editing are thought to be key aspects in the selection of CD4$^+$ T cell epitopes (Kim and Sadegh-Nasseri, 2015; Mellins and Stern, 2014). In a broader view, once stably bound to the MHC-II groove the peptides are protected from further cleavage and potential total destruction by endosomal proteases: some peptides might outperform others in this highly competitive process, therefore leading to higher relative abundance of some peptide-MHC-II complexes that progressively become dominant over others (Kim and Sadegh-Nasseri, 2015). Moreover, the peptide-intrinsic kinetic stability of the mature pMHC-II complexes can be another important parameter in determining immunodominance: peptide-MHC-II complexes endowed with high half-life could persist longer time exposed on the cell surface of the APC, thus increasing their chances to be recognized by the cognate CD4$^+$ T cells (Sant et al., 2005). As mentioned before, receptor-mediated endocytosis, such as BCR- or FcγR-mediated uptake, facilitates efficient sampling of low-abundant antigens by the APCs, but the receptor-antigen interaction is an additional variable that can potentially affect processing by masking some portions of the antigen from the endosomal proteolysis (Unanue et al., 2016). For instance, high affinity antibodies can maintain stable interaction with their cognate antigen even at endosomal pH, thus being able to strongly

suppress the presentation of some T cell epitopes when associated either in the form of BCR or as part of immune complexes (Watts and Lanzavecchia, 1993).

Alongside epitope selection during antigen processing, also the availability of the specific T cells within the naïve CD4[+] T cell repertoire arisen from the thymus can play a role in determining immunodominance to foreign antigens. For instance, the frequency of naïve CD4[+] T cell precursors potentially able to recognize each particular antigen-derived peptide (Jenkins and Moon, 2012; Moon et al., 2007; Nelson et al., 2015), as well as the affinity of interaction of their TCRs with the pMHC-II complexes has been demonstrated to deeply bias the clonal selection and the magnitude of the T cell response (Malherbe et al., 2004). TCR profiling of naïve repertoires of young and elderly individuals has revealed uneven homeostatic proliferation of naïve T cells occurring during aging, with some antigen-inexperienced TCR clonotypes that acquire inequalities in clonal sizes. Age-associated skewing in clonal size distributions of the naïve repertoire might therefore affect the frequency of antigen-specific naïve precursors and potentially lead to "holes" in the repertoire, thus influencing immunodominance to foreign antigens in the elderly (Qi et al., 2014).

4.2.4   Heterogeneity of T helper cell response

Naïve T lymphocytes express surface molecules such as L-selectin (CD62L) and the chemokine receptor CCR7 that allow their homing in the T cell zones of secondary lymphoid organs, where they can interact with antigen-presenting DCs. Recirculating between blood and secondary lymphoid organs, naïve T cells continuously patrol the body in search for their cognate antigen to ensure host protection (Sallusto et al., 1999). In a primary immune response, three distinct signals are needed to be delivered by DCs for the efficient activation and differentiation of naïve CD4[+] T cells: the 1st signal is the presentation of the antigen in the context of MHC-II molecules, which offer a substrate for recognition by the TCR. Moreover, mature DCs provide co-stimulation by upregulating the ligands of costimulatory receptors expressed by naïve CD4[+] T cells (2nd signal), and they produce cytokines (3rd signal) that can direct the differentiation of activated naïve T cells towards multiple fates.

Once activated, naïve T cells undergo robust proliferation and differentiate into effector T cells that can migrate to inflamed tissues, and T follicular helper (Tfh) cells that migrate to B cell follicles to provide help for B cell responses (Sallusto et al., 2018). Upon resolution of the primary immune response, some antigen-experienced T cells can survive long-term as memory cells ensuring faster and stronger response to secondary challenges. Depending on their homing capabilities and effector potential, long-lived memory CD4[+] T cells can be divided into two main subsets: T effector memory (Tem) cells lack the expression of CCR7 and are equipped with sets of chemokine receptors for homing in inflamed peripheral tissues, where they can deliver immediate effector functions and protection against invading microbes (Sallusto et al., 2004; Sallusto et al., 2010; Sallusto et al., 1999). Conversely, T central memory (Tcm) cells maintain expression of CCR7 and other molecules for the re-circulation into T-cell zones of secondary lymphoid organs. Even if Tcm cells lack immediate effector functions, they can vigorously proliferate upon secondary encounter with the antigen and differentiate to effector cells. In such way, Tcm cells are responsible for patrolling the secondary lymphoid organs and mounting recall responses to subsequent antigenic exposures (Sallusto et al., 2004; Sallusto et al., 2010).

To ensure host protection, the immune system has evolved elegant mechanisms by which, from the first sensing of invading microbes by innate immune cells, it triggers a chain reaction series of events that lead ultimately to the differentiation of antigen-specific T helper cells with effector functions useful to the clearance of the initial cause of inflammation. The fine dissection of the repertoire of antigen-experienced CD4[+] memory T cells has revealed that T cell responses have a multiplicity of possible differentiation fates, characterized by a high level of heterogeneity and interconnection between different T helper (Th) subsets (Becattini et al., 2015; Sallusto and Lanzavecchia, 2009). Such great diversity in T helper cell responses is acquired during naïve CD4[+] T cell priming as function of the strength and type of the signals provided by DCs. Depending on the nature of the pathogen and other environmental cues, antigen-presenting DCs have the ability to provide different co-stimulatory signals (2nd signal) and produce a variety of cytokines (3rd signal), therefore instructing with differential polarizing conditions the naïve T cell priming. The final result is the differentiation of T helper cells into distinct functional modules, each tailored for the

particular agent that induced the immune response and able to elicit appropriate defense mechanisms to effectively fight different classes of pathogens. Each T helper cell phenotype is enforced by the expression of lineage-defining transcription factors that orchestrate the gene expression of specific sets of homing receptors and effector cytokines (Sallusto and Lanzavecchia, 2009). The combinatorial expression of homing receptors, such as chemokine receptors, has been instrumental for the study of the phenotype of each T helper subset, as well as for the enumeration of the frequency of T cells reactive for different classes of pathogens in each Th compartment (**Fig. 4-6**) (Sallusto, 2016; Sallusto et al., 2018).

The immune response to viruses and intracellular bacteria mostly relies on type 1 helper cells (Th1): those cells express the transcription factor T-bet and produce IFN-γ that activate macrophage effector functions, induce an anti-viral state in non-immune cells and promote B cell isotype switch to IgG subclasses. The expression of the chemokine receptor CXCR3 guides the migration of Th1 cells into inflamed peripheral tissues and can facilitate their interaction with APCs (Groom and Luster, 2011).

Type 2 helper cells (Th2) express the transcription factor GATA3 and ensure protection against helminths and venoms. By production of the cytokines IL-4, IL-5 and IL-13, Th2 cells are responsible for the induction of type 2 responses that involve eosinophils and basophils infiltration, mucus production, smooth muscle contraction and production of IgE that can lead to mast-cell activation. Type 2 responses are necessary to fight parasitic worms and for protection against venoms, but if dysregulated can lead to the onset of allergic diseases.

Protection against fungi and extracellular bacteria is mediated by Th17 cells, that express the transcription factor ROR-γt and the chemokine receptors CCR6 and CCR4. The cytokines IL-17 and IL-22 produced by Th17 cells drive the recruitment and activation of neutrophils at the site of infection and act on epithelial cells, stimulating the production of antimicrobial peptides (such as defensins) and promoting the epithelial barrier function.

A recently emerged category of T helper cells includes the non-classic Th1, or Th1* subset: these cells co-express T-bet and ROR-γt, as well as the chemokine receptors CXCR3 and CCR6, and are particularly important in immunity to

Mycobacterium tuberculosis (MTB) (Acosta-Rodriguez et al., 2007; Lindestam Arlehamn et al., 2013). Despite the fact that both classic Th1 and non-classic Th1* produce large amount of IFN-γ, the analysis of patients with primary immunodeficiency disorders revealed that the absence of functional ROR-γt results in the severe impairment of the ability of *Mycobacterium*-reactive Th1* cells to produce IFN-γ. However, classic Th1 cell response to viral antigens was unaffected in those patients, thus proving the existence of two different pathways for the differentiation of IFN-γ-producing T helper cells (Okada et al., 2015; Sallusto, 2016).

The phenotype of T follicular helper cells (Tfh) is defined by the transcription factor BCL-6 and the expression of the chemokine receptor CXCR5 that allows migration of Tfh cells to the B cell zones of secondary lymphoid organs. Tfh cells are necessary for the generation of affinity-matured, isotype-switched long-lived plasma cells and memory B cells (Sallusto, 2016). Once homed in the B cell follicles, Tfh establish cell-cell interactions with the antigen-presenting B cells to provide them help by different means that include secretion of the B-cell trophic IL-21, IL-4 and IL-10 cytokines, or activation of costimulatory receptors expressed by B cells such as CD40 engaged by Tfh-upregulated CD40L (Crotty, 2015). By directing affinity maturation and Ig isotype switching, Tfh cells represent the limiting factor for germinal center reactions, therefore being in control of the final outcome of antibody responses in terms of affinity and breadth (Sallusto, 2016).

Type	Differentiation signal(s)	Transcription factor(s)	Effector molecule(s)	Homing receptors	Target cells	Function
Th1	IL-12, IFN	T-bet	IFN-γ	CXCR3	Macrophages	Virus, Bacteria
Th2	IL-4	GATA3	IL-4, IL-5, IL-13	CCR4 / CRTh2	Eosinophils	Parasites, Venoms
Th17	IL-6, IL-23	ROR-γt	IL-17, IL-22	CCR6 / CCR4	Neutrophils Epithelial cells	Fungi
Th1*	IL-12, IFN, ?	T-bet, ROR-γt	IFN-γ	CCR6 / CXCR3	Macrophages	Bacteria
Tfh	IL-21	BCL-6	IL-21	CXCR5	B cells	Antibodies

Figure 4-6. CD4[+] T helper subsets involved in tailored immunity to different classes of pathogens by specialized mechanisms. Adapted from (Sallusto and Lanzavecchia, 2009).

4.2.5   B cells and humoral immunity

B lymphocytes are cells specialized in the production of antibodies, very powerful effector molecules in the toolbox of the acquired immune system. Antibodies can provide host protection by physical neutralization of the bound antigen, and by eliciting a cascade of effector functions mediated by their Fc portions, such as promotion of phagocytosis and cytotoxicity by interaction with Fc-receptors and activation of the complement system (Nimmerjahn and Ravetch, 2008).

As a parallelism with the naïve T cell repertoire, also the naïve B cell repertoire is constituted of an extremely rich and diverse collection of receptors for the antigen, named B cell receptors (BCR), which are acquired by genetic recombination in the bone marrow to ensure the broadest coverage as possible against invading pathogens. A fundamental difference lays in the way of antigen recognition: while T cells need the presentation of the antigen, B cells are able to recognize conformational and linear epitopes contained within the native structure of the antigen. Nevertheless, the deployment of an effective B cell response to protein antigens marked by long-lasting titers of high-affinity antibodies involves a complex cross-talk with CD4[+] T helper cells that depends on MHC-II antigen presentation by B cells (Sallusto et al., 2010). In fact, B cells are able to concentrate antigens with high sensitivity by BCR-mediated uptake and present antigen-derived peptides in the context of MHC-II molecules, thereby allowing interaction with cognate Tfh cells that in turn help B cell activation by providing co-stimulatory signals and cytokines.

Once activated, antigen-specific B cells can undergo differentiation into long-lived plasma cells, a cell type that engraft in the bone marrow and is specialized in the secretion of large amounts of antibodies independently of antigen contact, thus being crucial for maintaining the antibody titers over time and ensure protective memory. Alongside, some antigen-experienced B lymphocytes differentiate into class-switched memory B cells, which can persist for decades into secondary lymphoid organs or recirculate in the blood, thus ensuring reactive memory in case of secondary antigenic stimulation (Kurosaki et al., 2015; Sallusto et al., 2004).

4.2.6   Maturation of the antibody response and CD4[+] T cell help

As compared to the primary response, reactive memory B cell response to subsequent antigen challenges is faster and of higher magnitude, leading to the development of plasma cells that produce antibodies with increased affinity and switched isotypes (Kurosaki et al., 2015; Victora and Nussenzweig, 2012). Hallmark of the antibody response is indeed the ability to evolve during time in terms of affinity for the antigen and breadth of reactivity, a phenomenon known as affinity maturation, as well as to acquire new isotypes that can mediate very different effector functions.

Such events rely on the ability of activated B cells to operate an endogenous process of genetic engineering that increases furthermore the diversity of their somatically rearranged B cell receptors. Somatic hypermutation (SHM) of activated B cells is a process that results essentially in the random mutagenesis of the sequence of the antigen-reactive BCR. The outcome of this creative process is a progeny of related B cell clones bearing variants of the original BCR that can potentially have brand new properties in terms of interaction with the cognate antigen. For instance, some of the newly acquired mutations could be detrimental for the antigen binding, but some others mutated BCRs might have higher affinity for the antigen, or recognize a broader spectrum of closely related structures. This newly generated diverse repertoire of B cell clones can therefore be subject to iterative rounds of Darwinian-like selection, ultimately driven by the antigen itself, with the positive selection of high-affinity mutant clones that possess the highest fit in the highly competitive environment of the germinal centers. Affinity maturation is therefore the result of a B cell clonal selection that is function of antigen dose and exposure (primary versus secondary challenge), and availability of help from antigen-specific Tfh cells. Since in the germinal centers B cell clones fight each other for restricted T cell help, the latter should be finely regulated to ensure "high-quality" affinity maturation: if the amount of T cell help stops being rate-limiting, such as in case of the experimental dysregulation of the number of antigen-specific Tfh cells, its selective force is abolished and affinity maturation results compromised (Preite et al., 2015).

The outcome of the selection of the mutant B cell clones generated by SHM is generally the increase, hence the maturation, of the affinity of the serum antibodies over

time (Victora and Nussenzweig, 2012), and furthermore can lead to the production of antibodies endowed with increased breadth of antigen recognition. Indeed, as it has been showed for HIV-1, influenza virus and paramyxoviruses, SHM can extend further beyond the generation of high affinity clones, with additional mutations that can continue to accumulate and can broaden the antibody reactivity (Corti et al., 2013; Corti et al., 2011; Liao et al., 2013; Pappas et al., 2014). In the evolutional race against infectious diseases, the extensive intraclonal diversification obtained by SHM might have evolved to be a factory of antibodies' upgraded versions potentially able to recognize with increased potency slightly different epitopes, to promptly face with rapidly diversifying viruses.

5. Aims

Given the fundamental role of CD4$^+$ T helper cells in induction and sustainment of antibody and CD8$^+$ T cell responses, in my doctoral studies I focused on the characterization of antigen-specific CD4$^+$ T cells isolated from human donors under health or disease conditions. The analysis of antigen specificity offers a privileged point of view to understand the dynamics of adaptive immune responses to foreign antigens, such as pathogens and vaccines. Acquired immune system has evolved to ensure host protection against harmful pathogens but it can be detrimental under certain conditions, for instance if the responses are directed against exogenous proteins used as biological drugs (e.g. therapeutic monoclonal antibodies) or if they are endowed with inappropriate effector phenotypes.

In particular, with my work I aimed to:

- Dissect the principles of T-B cells collaboration at the basis of immunogenicity of therapeutic antibodies (i.e. natalizumab used to treat multiple sclerosis), to delineate new approaches to engineer improved versions of the drugs endowed with higher efficacy due to lower immunogenicity.

- Unravel the relationship between antigen presentation and CD4$^+$ T cell immunodominance to influenza A virus hemagglutinin, with the aim to ameliorate predictors of immune responses elicited by vaccination or viral infections.

- Measure and characterize the CD4$^+$ T cell responses to commensal and pathogenic Enterobacteriaceae species in healthy donors or septic patients, to evaluate the impact of heterologous immunity in immune homeostasis and host protection.

Each of these goals was pursued in three different projects, which I will describe in detail in the next chapters.

6. Results

6.1 A single T cell epitope drives the neutralizing anti-drug antibody response to natalizumab in multiple sclerosis patients

Antonino Cassotta[1,2], Vincent Mikol[3], Thomas Bertrand[3], Stéphanie Pouzieux[3], Josiane Le Parc[3], Paul Ferrari[3], Jacques Dumas[3], Michael Auer[4], Florian Deisenhammer[4], Matteo Gastaldi[5], Diego Franciotta[5], Chiara Silacci-Fregni[1], Blanca Fernandez Rodriguez[1], Isabella Giacchetto-Sasselli[1], Mathilde Foglierini[1,6], David Jarrossay[1], Roger Geiger[1], Federica Sallusto[1,2], Antonio Lanzavecchia[1] & Luca Piccoli[1]

[1] Institute for Research in Biomedicine, Università della Svizzera italiana, Bellinzona, Switzerland.

[2] Institute of Microbiology, ETH Zurich, Zurich, Switzerland.

[3] Research Platform, Sanofi R&D, Vitry-sur-Seine Cedex, France.

[4] Department of Neurology, Innsbruck Medical University, Innsbruck, Austria.

[5] Laboratory of Neuroimmunology, IRCCS Mondino Foundation, Pavia, Italy.

[6] Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland.

Correspondence should be addressed to L.P. (luca.piccoli@irb.usi.ch)

6.1.1 Author contributions

A.C. characterized the T cell response, performed the peptidomics, analyzed the data and wrote the manuscript; V.M. performed structural analyses, modeling, deimmunization and supervision of structural studies; T.B. determined the crystal structures; S.P. performed crystallization and characterization of antibody complexes; J.L.P. cloned the antigen-binding fragments for crystallization; P.F. purified antibodies; J.D. expressed the antigen-binding fragments for crystallization; M.A. collected clinical data and samples; F.D. collected clinical data and provided supervision; M.G. collected clinical data and samples; D.F. collected clinical data and provided supervision; C.S.-F. immortalized memory B cells and performed screenings; B.F.R. sequenced and expressed antibodies; I.G.-S. analyzed antibody sequences; M.F. performed bioinformatics analyses; D.J. performed cell sorting; R.G. analyzed mass-spectrometry data; F.S. provided supervision and wrote the manuscript; A.L. provided supervision, analyzed the data and wrote the manuscript; L.P. provided overall supervision, designed the experiments, characterized the antibodies, analyzed the data and wrote the manuscript.

6.1.2 Abstract

Natalizumab (NZM), a humanized monoclonal IgG4 antibody to α4 integrins, is used to treat patients with relapsing-remitting multiple sclerosis (MS) (Chataway and Miller, 2013; Li et al., 2018), but in about 6% of the cases persistent neutralizing anti-drug antibodies (ADAs) are induced leading to therapy discontinuation (Bachelet et al., 2016; Calabresi et al., 2007). To understand the basis of the ADA response and the mechanism of ADA-mediated neutralization, we performed an in-depth analysis of the B and T cell responses in two patients. By characterizing a large panel of NZM-specific monoclonal antibodies, we found that, in both patients, the response was polyclonal and targeted different epitopes of the NZM idiotype. The neutralizing activity was acquired through somatic mutations and correlated with a slow dissociation rate, a finding that was supported by structural data. Interestingly, in both patients, the analysis of the CD4$^+$ T cell response, combined with mass spectrometry-based peptidomics, revealed a single immunodominant T cell epitope spanning the FR2-CDR2 region of the NZM light chain. Moreover, a CDR2-modified version of NZM was not recognized by T cells, while retaining binding to α4 integrins. Collectively, our integrated analysis identifies the basis of T-B collaboration that leads to ADA-mediated therapeutic resistance and delineates an approach to design novel deimmunized antibodies for autoimmune disease and cancer treatment.

6.1.3   Results and discussion

The therapeutic use of monoclonal antibodies and other biopharmaceutical products can result in an immune response to the drug that, in some cases, affects its efficacy due to the production of neutralizing ADAs (Rup et al., 2015). Several clinical studies have measured ADA levels in sera of selected cohorts of patients and concluded that not all antibody responses lead to drug neutralization (Dunn et al., 2018; Jensen et al., 2019; Link et al., 2017; Murdaca et al., 2016; Quistrebert et al., 2019; Zare et al., 2013). However, an explanation for these heterogeneous responses and an integrated characterization of the B and T cell responses to the drug are still missing. In this study, we isolated NZM-specific B and T cell clones from memory cells of two MS patients, who had a hypersensitivity reaction following drug infusion and developed high titers of ADAs (**Supplementary Table 6.1-1** and **Supplementary Fig. 6.1-1**).

To identify NZM-specific antibodies, we screened supernatants of immortalized B cells (Traggiai et al., 2004) and isolated 30 and 10 anti-NZM monoclonal antibodies from patients A and B, respectively (**Fig. 6.1-1**). Most antibodies from patient A showed high affinity for NZM (KD values 1-6,790 pM, median 6.1 pM), while antibodies from patient B showed lower affinity (KD values 0.4-22.7 nM, median 2.3 nM) (**Supplementary Table 6.1-2**). Sixty percent (18/30) of the antibodies from patient A potently inhibited binding of NZM to α4 integrins on the surface of T cells (IC90 values 17-271 ng/ml, defined as NAbs, neutralizing antibodies), while the remaining showed reduced or no inhibitory capacity (IC90 values > 1,000 ng/ml, defined as BAbs, binding antibodies) (**Fig. 6.1-2a** and **Supplementary Table 6.1-2**). Interestingly, patient B developed only non-neutralizing BAbs, a finding that may be related to the lower number of NZM infusions received.

The antibodies isolated were mostly IgG1-λ, used different V(D)J genes and showed moderate levels of somatic mutations (**Fig. 6.1-1**). NAbs from patient A carried a high load of replacement mutations in the complementary determining regions (CDRs), consistent with an antigen-driven selection (**Fig. 6.1-2b**), while BAbs, from both patients, carried few replacement mutations in the CDRs. The antibody fine specificity was tested using 64 NZM variants generated by swapping the hypervariable CDR loops with the counterparts of the human scaffold antibody used for NZM

humanization (**Extended Data Fig. 6.1-1a-b**). The antibodies recognized epitopes comprising one to six NZM CDRs, with preferential recognition of the heavy chain CDRs (**Fig. 6.1-1** and **Extended Data Fig. 6.1-1c**). These findings demonstrate that NZM induces a neutralizing polyclonal anti-idiotypic antibody response that targets multiple epitopes located primarily in the heavy chain CDRs.

The difference in neutralizing activity of the anti-NZM antibodies isolated may be due to the binding to distinct epitopes or, alternatively, to different affinity or kinetics of binding. When compared for their capacity to bind to the 64 NZM CDR swap variants, BAbs and NAbs did not cluster separately (**Extended Data Fig. 6.1-1c**), indicating that the difference is not related to epitope specificity. However, BAbs and NAbs showed different binding kinetics, as assessed by surface plasmon resonance (SPR). In particular, while the association constant (k_a) was comparable, NAbs showed a lower dissociation constant (k_d) that significantly correlated with NZM neutralization ability (**Fig. 6.1-2c-d** and **Supplementary Table 6.1-2**). The role of somatic mutations was addressed by comparing the antibodies to the unmutated common ancestors (UCAs). In three out of four BAbs tested, the UCAs had low binding affinity for NZM, which was increased by somatic mutations. In contrast, in two out of three NAbs tested, the UCAs showed already high affinity, but required somatic mutations to gain full neutralizing activity (**Fig. 6.1-2e**). Collectively, these findings indicate a critical role for somatic mutations in the generation of high-affinity antibodies and highlight the importance of dissociation rate for ADA neutralizing activity.

We next determined the crystal structure of a BAb (NAA32) and a NAb (NAA84) in complex with NZM. With 2.8 Å resolution, the structures revealed that both antibodies interacted with the same surface area on NZM but engaged the molecule with different orientation (**Fig. 6.1-3a** and **Supplementary Table 6.1-3**). NAA32 and NAA84 recognized 22 and 18 residues of NZM CDRs, respectively, which were mostly located in the heavy chain, a finding consistent with the CDR swap variant specificity (**Fig. 6.1-3b**). In particular, 14 of these residues were recognized by both antibodies, highlighting potential immunodominant B-cell epitopes on NZM. The contact surface area between NAA32 and NZM was 757Å^2, although the main interactions occurred at two separate contact points leaving an empty space in between the two interfaces (**Fig. 6.1-3c** and **Extended Data Fig. 6.1-2a**). In contrast, the contact surface area between

NAA84 and NZM was 617Å^2, but interactions were much tighter, with deeper residue contacts and contiguous surface complementarity that could account for its slower dissociation rate from NZM (**Fig. 6.1-3c** and **Extended Data Fig. 6.1-2a**). Importantly, the NZM surface area engaged by NAA32 and NAA84 overlaps with the area engaged by α4 integrin (Yu et al., 2013) with 13 shared residues (**Fig. 6.1-3b**). While both NAA32 and NAA84 antibodies occluded α4 integrin binding site, neither of them provided molecular mimicry of the integrin, which exhibited a different binding orientation (**Extended Data Fig. 6.1-2b**). Together with the functional data, these structural data suggest that the difference between BAbs and NAbs cannot be explained by the fine epitope specificity, but rather by the strength and the fit of the interaction with NZM.

The generation of neutralizing ADAs through somatic mutations is consistent with an affinity maturation driven by CD4$^+$ T cells targeting non-self regions of the NZM idiotype. To address this hypothesis and to characterize the specificity of the T-cell response, we stimulated CFSE-labeled memory CD4$^+$ T cells from both patients with overlapping peptides covering the variable regions of NZM (**Extended Data Fig. 6.1-3**). Cloning of activated CFSE^{low} T cells resulted in the isolation of several NZM-reactive T cell clones (12 from patient A and 54 from patient B) (**Fig. 6.1-4a**). T cell receptor (TCR) Vβ-gene sequencing revealed the presence of 5 and at least 7 distinct T cell clonotypes in patient A and B, respectively (**Fig. 6.1-4b** and **Supplementary Table 6.1-4**). Strikingly, most of the T cell clones from both patients recognized two overlapping peptides spanning a region comprising the end of framework region 2 (FR2) and the CDR2 of the NZM light chain (GKAPRLLIHYTSALQPGI, named NZM-LC$_{FR2-CDR2}$) (**Fig. 6.1-4c**). Of note, T cell recognition of NZM-LC$_{FR2-CDR2}$ was restricted by HLA-DRB1*14/16 in patient A and by DRB1*07/07 in patient B (**Extended Data Fig. 6.1-4a** and **Supplementary Table 6.1-1** and **6.1-4**). Taken together, these findings suggest that, in these patients, a single T cell epitope in the NZM molecule induced an HLA-DR-restricted CD4$^+$ T-cell response that was sufficient to sustain a strong polyclonal B cell response to the NZM idiotype.

To investigate the mechanism that leads to the presentation of the NZM T cell epitope, we pulsed NZM-specific B cell clones with NZM and identified the naturally processed MHC-II-bound peptides by mass spectrometry-based peptidomics. Three sets

of nested peptides that mapped to the variable regions of NZM were identified (**Fig. 6.1-4d** and **Supplementary Table 6.1-5**). Two of these sets covered the heavy chain FR3 and the FR4-CH1 region, which are in the human germline configuration and therefore are not expected to induce an immune response, a notion consistent with our failure to isolate specific T cell clones. Remarkably, the third set of peptides mapped to the FR2-CDR2 region of the NZM light chain (consensus, TPGKAPRLLIHYTSALQPGIPSR) that spanned the immunodominant NZM-LC$_{FR2-CDR2}$ epitope (GKAPRLLIHYTSALQPGI) recognized by NZM-reactive T cell clones (**Fig. 6.1-4d** and **Supplementary Table 6.1-5**). The NetMHCIIpan algorithm (Jensen et al., 2018b) applied to predict binding to the HLA-DR alleles carried by the two patients identified the same three peptides, as well as several other peptides mapping to CDRs and FRs of NZM. (**Extended Data Fig. 6.1-4b**). In addition, this algorithm predicted the binding of the NZM-LC$_{FR2-CDR2}$ epitope to a reference panel of nine DRB1 and DRB3/4/5 alleles (de la Hera et al., 2014; Paul et al., 2015) (**Extended Data Fig. 6.1-4c**), suggesting its potential immunodominance also in individuals with a diverse HLA background. Collectively, these findings demonstrate that only one of the potentially immunogenic peptides encoded by the six CDRs of NZM was a naturally presented T cell epitope able to generate a polyclonal CD4[+] T-cell response.

The identification of the immunodominant NZM-LC$_{FR2-CDR2}$ epitope prompted us to use a structure-guided design to engineer a "deimmunized" version of NZM. First, we identified residues of NZM light chain CDR2 that were not engaging α4 integrin binding and modelled different mutants with the constraint to preserve the conformation of the CDR2 and the specificity of NZM. Four NZM variants (var1-4) were retained for experimental testing and validation together with a fifth variant (var5) in which the CDR2 of the light chain was reverted to the germline sequence of the human antibody scaffold (**Fig. 6.1-4e**). Two variants, var1 and var3, retained binding to α4 integrins, while var2 and var4 showed partial loss of binding that was considerably reduced in the case of var5 (**Fig. 6.1-4f**). Remarkably, none of the five NZM variants was able to trigger proliferation of T cell clones specific for the naturally presented NZM-LC$_{FR2-CDR2}$ peptide (**Fig. 6.1-4g**), a finding consistent with either absence of TCR cross-reactivity or with a reduced binding to class II molecules. Prediction of binding to a reference set of nine DRB1 and DRB3/4/5 alleles (de la Hera et al., 2014; Paul et al.,

2015), as well as to the DRB1 alleles of both patients, showed a reduction of predicted binding affinity for the var1 and var3 peptides compared to the original NZM peptide (**Fig. 6.1-4h** and **Extended Data Fig. 6.1-4d**). These results provide two deimmunized versions of NZM that can be tested *in vivo*.

This study integrates, for the first time, clonal analysis of B and T cell repertoires and mass spectrometry-based peptidomics to identify the factors that underpin the neutralizing antibody response to a humanized therapeutic antibody. The large number of monoclonal antibodies isolated recognize multiple epitopes spanning different NZM CDRs and therefore represent classical anti-idiotypic antibodies (Pan et al., 1995; van Schie et al., 2017) rather than internal image of the antigen. Interestingly, neutralizing antibodies showed high-level replacement mutations in the CDRs and low dissociation rate, suggesting that B cell selection was driven by decreased k_d rather than increased k_a (Foote and Milstein, 1991).

The highly diverse anti-idiotypic response is consistent with the presence of multiple B cell epitopes recognized by naïve B cells and contrasts with the T cell response that is largely limited to a single epitope that we mapped to the FR2-CDR2 region of NZM light chain. This finding highlights the merit of the humanization technology in limiting the T cell immunogenicity of therapeutic antibodies, since chimeric antibodies were found to elicit T cell responses against multiple epitopes in the FRs and CDRs of both heavy and light chains (Hamze et al., 2017). As expected, the residual immunogenicity is primarily in the CDR regions (Harding et al., 2010), but it is still limited by HLA restriction and processing by antigen-presenting cells.

Previous studies reported a positive association of NZM-related hypersensitivity reactions with DRB1*13 and 14 alleles (de la Hera et al., 2014), but did not investigate any T cell response. The finding that the NZM-LC$_{FR2-CDR2}$ peptide is naturally presented in the context of the HLA-DRB1*07 and DRB1*14/16 alleles of the two patients, together with the prediction of its binding to different alleles, including DRB1*13, suggests that this peptide is a major source of T cell help driving the anti-idiotypic B cell response to NZM. This may explain the high frequency of MS patients producing ADAs to NZM (Bachelet et al., 2016; Calabresi et al., 2007) and offers the possibility for the deimmunization of the drug for a more safe treatment. Structured-guided

mutagenesis and experimental tests suggest that a little amendment of only three residues of the CDR2 of the NZM light chain may be sufficient to remove the immunogenic T cell epitope without interfering with the interaction of NZM with its target.

Collectively, our results demonstrate how the integration of peptidomics, structural data, *in silico* predictions and dissection of the specific B and T cell responses represents a powerful approach to define the immunogenic landscape of therapeutic antibodies and to guide the deimmunization strategies of next-generation biological therapeutics for autoimmune and cancer disease. Currently used fully human antibodies, including checkpoint inhibitors, may also benefit from the deimmunization strategy. Reciprocally, this approach could be used to improve the immunogenicity of vaccines through the engineering of dominant T cell epitopes driving neutralizing antibody responses.

6.1.4 Figures and tables

MS patient	B cell clone	Isotype		Heavy chain VDJ genes (% identity to germline)					Light chain VJ genes (% identity to germline)				H1	H2	H3	L1	L2	L3	(%)
A	NAA2	IgG1	κ	VH1-46	(97.6)	D4-23	JH4	(87.5)	VK3-11	(98.6)	JK3	(91.4)							0
	NAA5	IgG1	λ	VH4-39	(96.6)	D3-22	JH3	(94)	VL2-8	(98.3)	JL2	(100)							10
	NAA6	IgG1	λ	VH3-30-3	(97.2)	D1-26	JH3	(96)	VL3-19	(97.5)	JL3	(100)							20
	NAA7	IgG1	κ	VH4-4	(97.5)	D3-3	JH6	(79)	VK1-5	(97.9)	JK1	(100)							30
	NAA9	IgG3	λ	VH3-23	(99)	D7-27	JH4	(91.7)	VL1-51	(98.6)	JL2	(94.4)							40
	NAA20	IgG1	λ	VH4-34	(97.5)	D6-13	JH4	(87.5)	VL2-8	(97.9)	JL1	(94.7)							50
	NAA32	IgG1	κ	VH1-3	(97.9)	D2-15	JH2	(98.1)	VK3-15	(100)	JK4	(97.4)							60
	NAA33	IgG1	λ	VH5-51	(97.2)	D6-6	JH5	(90.2)	VL3-1	(95.3)	JL2	(94.7)							70
	NAA34	IgG3	κ	VH3-23	(98.6)	D2-15	JH4	(100)	VK1-9	(98.6)	JK3	(100)							80
	NAA36	IgG1	λ	VH3-30-3	(95.5)	D1-1	JH4	(85.4)	VL2-11	(97.9)	JL2	(94.3)							90
	NAA40	IgG1	λ	VH4-61	(99.7)	D6-13	JH5	(90.2)	VL1-44	(99)	JL2	(94.7)							100
	NAA44	IgG1	λ	VH4-39	(96.9)	D5-12	JH5	(82.4)	VL2-8	(97.5)	JL2	(92.1)							
	NAA45	IgG1	λ	VH4-61	(97.9)	D3-16	JH4	(93.8)	VL1-44	(99)	JL3	(100)							
	NAA49	IgG1	λ	VH3-33	(98.6)	D5-24	JH4	(89.6)	VL3-21	(98.9)	JL2	(100)							
	NAA59	IgG1	λ	VH1-69	(97.9)	D1-1	JH3	(96)	VL3-19	(98.6)	JL2	(94.7)							
	NAA62	IgG3	κ	VH3-11	(97.6)	D4-17	JH4	(91.7)	VK3-11	(97.5)	JK5	(100)							
	NAA80	IgG1	λ	VH3-48	(96.9)	D2-15	JH6	(85.5)	VL3-19	(98.9)	JL3	(97.1)							
	NAA84	IgG1	λ	VH3-21	(97.6)	D3-22	JH3	(94)	VL3-19	(97.5)	JL2	(92.1)							
	NAA92	IgG3	λ	VH3-30	(99.3)	D3-22	JH4	(91.7)	VL4-69	(97.3)	JL2	(97.4)							
	NAA94	IgG2	λ	VH1-3	(96.9)	D3-22	JH5	(96.1)	VL1-40	(100)	JL3	(100)							
	NAA96	IgG3	λ	VH4-61	(98.3)	D3-3	JH4	(85.4)	VL1-44	(98.3)	JL3	(100)							
	NAA104	IgG1	κ	VH2-70D	(100)	D6-19	JH3	(98)	VK3-11	(98.2)	JK2	(97.4)							
	NAA105	IgG1	λ	VH1-46	(96.5)	D1-20	JH3	(94)	VL3-19	(98.9)	JL2	(97.1)							
	NAA110	IgG1	κ	VH3-30-3	(96.9)	D2-15	JH6	(83.9)	VK1-5	(98.2)	JK1	(100)							
	NAA113	IgG1	κ	VH3-33	(96.2)	D1-14	JH4	(89.6)	VK3-11	(100)	JK4	(100)							
	NAA114	IgG1	λ	VH4-39	(98.5)	D4-17	JH4	(87.5)	VL1-51	(99.3)	JL3	(100)							
	NAA116	IgG1	κ	VH3-30-3	(97.6)	D2-15	JH6	(90.3)	VK3-20	(100)	JK2	(100)							
	NAA119	IgG1	κ	VH1-3	(96.5)	D6-19	JH1	(94.2)	VK3-20	(98.6)	JK3	(94.7)							
	NAA125	IgG1	λ	VH4-39	(95.2)	D2-21	JH6	(85.5)	VL1-40	(98.6)	JL1	(100)							
	NAA128	IgG1	κ	VH1-69	(96.9)	D4-11	JH4	(97.9)	VK3-11	(100)	JK4	(100)							
B	NAE125	IgG1	λ	VH4-39	(97.6)	D1-7	JH4	(81.3)	VL1-47	(98.3)	JL3	(97.3)							
	NAE194	IgG1	κ	VH3-11	(94.1)	D6-19	JH4	(79.2)	VK1-12	(94.6)	JK4	(91.7)							
	NAE197	IgG1	κ	VH3-30	(100)	D3-16	JH4	(100)	VK1-39	(95.3)	JK1	(97.4)							
	NAE199	IgG1	λ	VH3-7	(97.2)	D7-27	JH3	(94)	VL3-19	(100)	JL1	(97.4)							
	NAE203	IgG3	λ	VH2-70	(97.6)	D3-16	JH3	(100)	VL3-9	(98.2)	JL2	(100)							
	NAE205	IgG3	λ	VH5-51	(94.8)	D4-11	JH6	(87.1)	VL3-10	(96.4)	JL1	(86.8)							
	NAE206	IgG1	λ	VH1-18	(94.8)	D3-22	JH5	(96.1)	VK3-20	(93.6)	JK2	(100)							
	NAE207	IgG1	λ	VH5-10	(100)	D2-2	JH4	(75)	VL4-60	(99.7)	JL3	(97.4)							
	NAE208	IgG1	κ	VH3-30	(89.9)	D3-9	JH3	(90)	VK3-15	(93.9)	JK1	(94.7)							
	NAE210	IgG1	κ	VH3-33	(95.1)	D3-22	JH5	(84.3)	VK1-5	(96.8)	JK1	(100)							

Figure 6.1-1. V(D)J gene usage and epitope mapping of 40 anti-natalizumab monoclonal antibodies. The colored cells show the binding of the antibodies to six selected NZM variants in which individual CDRs were swapped with the counterparts of the human scaffold antibody used for NZM humanization (H, heavy chain; L, light chain; 1, CDR1; 2, CDR2; 3, CDR3). Shown is the percentage of binding of the antibodies to the NZM CDR swap variants relative to NZM, as tested by ELISA, with a three-color gradation scale from minimum (0%, blue) to maximum (100%, white). OD values and exact % binding values are available as Source Data.
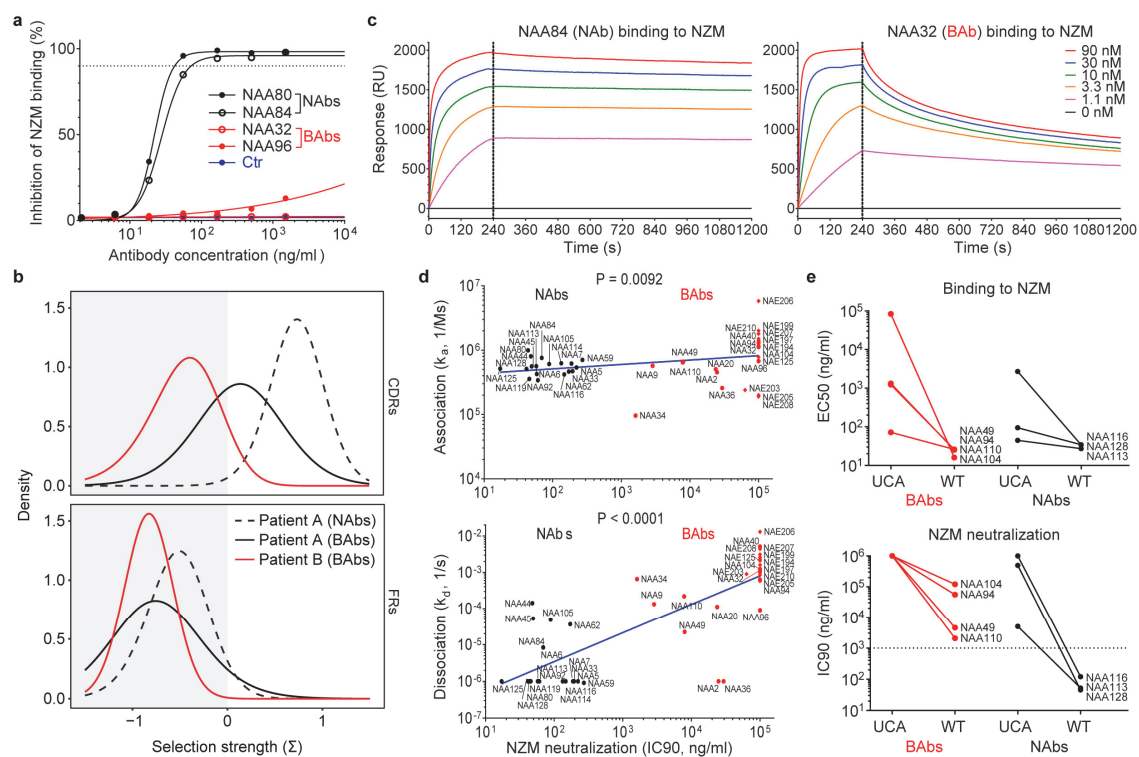
Figure 6.1-2. The neutralizing activity of ADAs is acquired through somatic mutations and correlates with a slow dissociation rate. a, Dose-dependent inhibition of NZM binding to T cells of four representative antibodies (two NAbs in black and two BAbs in red) compared to an unrelated antibody used as a control (representative of n = 2 independent experiments). The dotted line indicates the 90% of inhibition. **b**, Positive and negative selection strengths (Σ) in CDRs and FRs of 18 NAb and 12 BAb heavy chain sequences from patient A (black and dotted black lines) and 10 BAb heavy chain sequences from patient B (red line) estimated using the Bayesian estimation of Antigen-driven SELectIoN (BASELINe), which compares the observed frequencies of replacement and silent mutations with the expected mutations. **c**, Kinetics of binding of two representative antibodies (NAA84 and NAA32) to different doses of NZM as measured by SPR (representative of n = 2 independent experiments). The dotted line separates the association and the dissociation phase. RU, resonance units. **d**, Correlation of NZM neutralization (expressed as IC90, i.e. the antibody concentration required for 90% inhibition of NZM binding to T cells) with association constant (k_a, upper panel) and dissociation constant (k_d, lower panel) of n = 40 antibodies measured through SPR (representative of n = 2 experiments). BAbs and NAbs are shown as red and black symbols (circle for patient A and diamond for patient B), respectively. A two-tailed Spearman's correlation was performed; r coefficient, 0.4070 (k_a) and 0.7673 (k_d); 95% confidence interval, 0.09993 to 0.6433 (k_a) and 0.5928 to 0.8730 (k_d). **e**, Comparison of binding to NZM expressed as half-maximal effective concentration, EC50 (upper panel) and NZM neutralization expressed as IC90 (lower panel) between the unmutated common ancestor (UCA) and the wild type (WT) versions of seven representative antibodies (four BAbs and three NAbs) (representative of n = 2 independent experiments). The dotted line indicates the threshold of neutralization set at 1,000 ng/ml.

```
a




              NAA32              NZM (H+L)




   NAA84




b
Heavy chain    CDR1          CDR2                    CDR3
NZM   QV[…]ASGFNIKDTYIHWV[…]MGRIDPANGYTKYDPKFQ[…]CAREGYYGNYGVYAMDYW [...]
NAA32          KDTY          D AN Y K        [...]    Y GNYGVY        [...]
NAA84          KDTY          R D AN          [...]  E Y     VY        [...]
α4int          KDTY          R DPAN          [...]  E YYGNYGVY        [...]

Light chain    CDR1          CDR2            CDR3
NZM   DI[…]]TCKTSQDINKYMAWY[…]IHYTSALQPGI[…]YCLQYDNLWTFGQ[…]
NAA32          Q N Y         Y                  DN
NAA84            N Y                          YDN W
α4int            NKY         Y      P         YD  W

c
                                NZM (H+L)




              NAA32                          NAA84
```
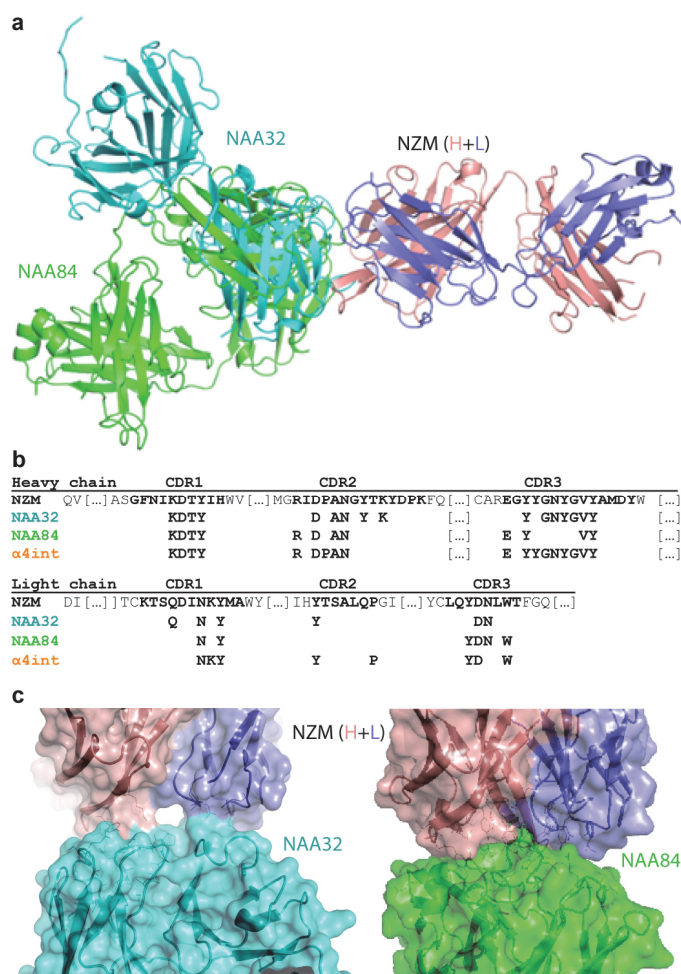
Figure 6.1-3. Structural features of the interaction of NZM with a NAb and a BAb. a, Superimposition of the antigen-binding fragment of NZM in complex with NAA84 (NAb, green) and NAA32 (BAb, cyan). NZM heavy and light chains are shown in salmon and slate blue, respectively. Proteins are displayed in ribbon diagram. **b,** Alignment of the NZM residues that are recognized by NAA32, NAA84 and α4 integrin. **c,** Detailed visualization of the interacting interfaces of NZM and NAA32 (left) and NAA84 (right). The antibodies are shown as ribbon diagrams with overlapping surfaces. The Sc values were 0.696 and 0.707 for NAA32/NZM and NAA84/NZM complexes, respectively.
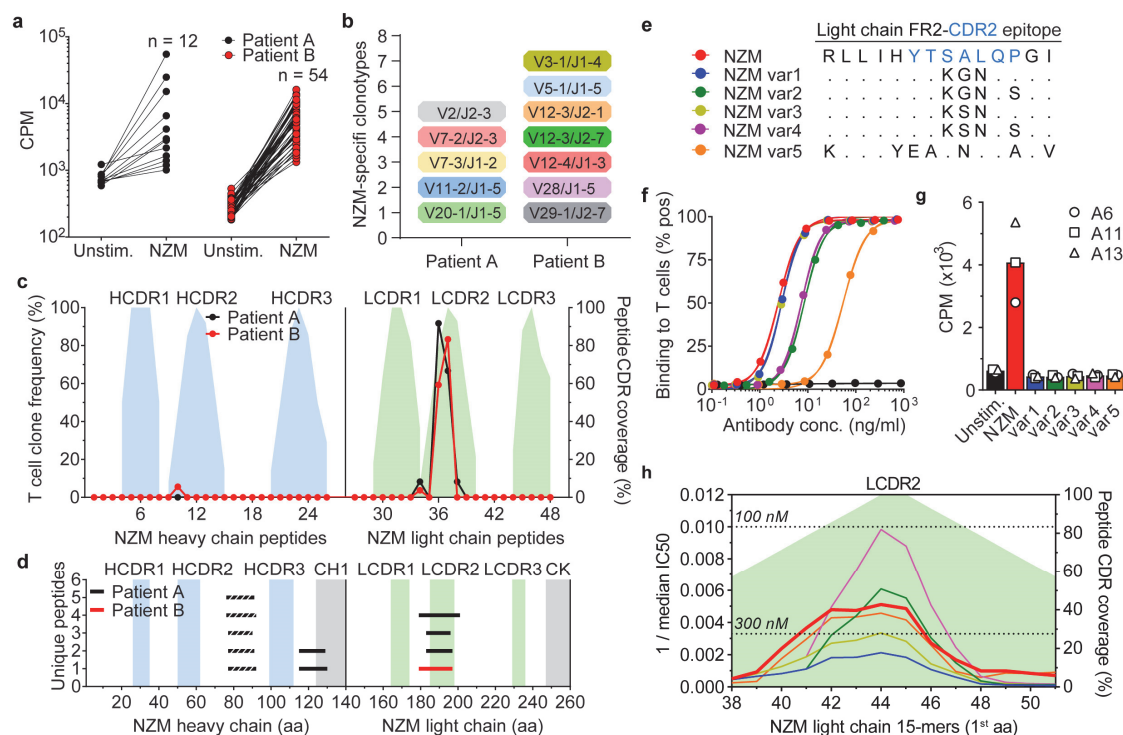
Figure 6.1-4. Identification of a single immunodominant T cell epitope that can be engineered to deimmunize NZM. a, Proliferative response of NZM-reactive T cell clones isolated after *ex-vivo* stimulation of memory $CD4^+$ T cells, upon re-stimulation with the overlapping peptide pool covering the entire sequences of the variable regions of the NZM heavy and light chains. Proliferation was assessed on day 3 after 16 h pulse with [³H]-thymidine. Data are expressed as counts per minute (CPM). Unstim., unstimulated T cells. **b,** TCR Vβ gene repertoire of NZM-reactive $CD4^+$ T cell clones isolated from MS patients A and B. The y-axis indicates the number of unique clonotypes identified. **c,** Epitope mapping of NZM-specific $CD4^+$ T cell clones isolated from patients A and B. The epitopes were identified by screening the T cell clones with overlapping peptides spanning the NZM heavy and light chain variable region. The left y-axis indicates the frequency of T cell clones reactive to each peptide. The right y-axis indicates the percentage of CDR residues within each peptide shown as filled areas in the background (representative of n = 2 experiments). **d,** Mass spectrometry-based MHC peptidomics of NZM-specific B cell clones pulsed with NZM. Each line represents a unique MHC-II-bound peptide identified in n = 2 independent experiments. Dashed lines indicates peptides belonging to both NZM and B cell receptor of the antigen-presenting cells. The y-axis indicates the number of unique peptides overlapping in the same NZM region. **e,** Comparison of the amino acid sequences of the $LC_{FR2-CDR2}$ epitope of NZM and the five variants engineered for the drug deimmunization. **f,** Binding of NZM variants to α4 integrin expressed on T cells analyzed by flow cytometry. The NZM variants are color coded as depicted in panel e. An antibody with different specificity was used a negative control (black curve). **g,** Proliferation of three NZM-$LC_{FR2-CDR2}$-reactive T cell clones (A6, A11 and A13) after stimulation with autologous B cells pulsed with NZM and the five engineered variants (representative of n = 2 independent experiments). The bars show the mean proliferation. **h,** Predicted binding affinities of 15mer peptides spanning the light chain CDR2 region of NZM variants to a reference set of nine DRB1 and DRB3/4/5 alleles (DRB1*0301, DRB1*0701, DRB1*1301, DRB1*1401, DRB1*1501, DRB3*0101, DRB3*0202, DRB4*0101, DRB5*0101) (de la Hera et al., 2014; Paul et al., 2015). The NZM variants are color coded as depicted in panel e. The affinities are shown as reciprocal median IC50 (nM) values. The dotted lines define the thresholds of high-affinity binding set at 100 nM and low-affinity binding set at 300 nM.
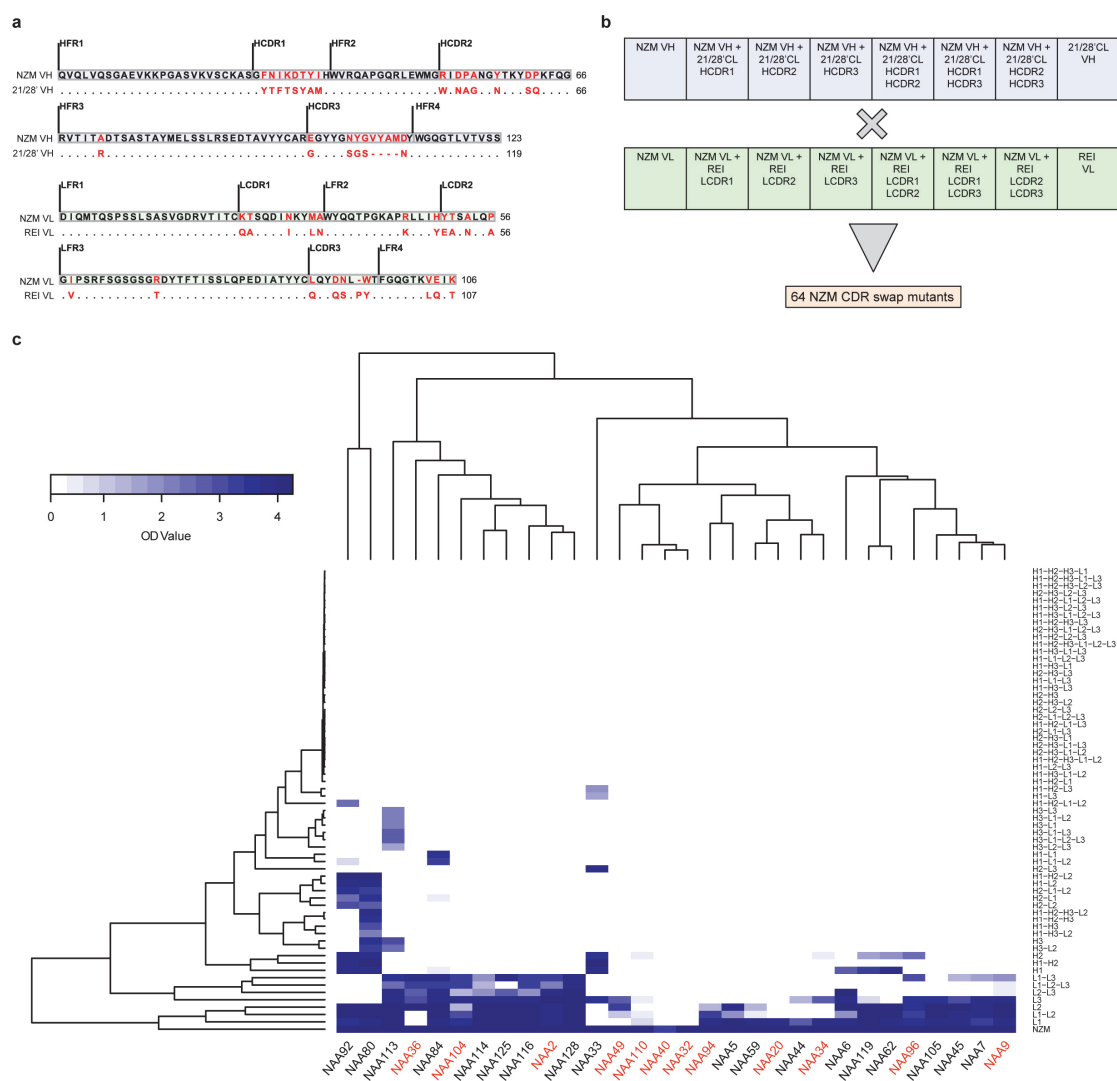
Extended Data Figure 6.1-1. Epitope mapping of NZM-specific antibodies. a, Alignment of NZM heavy and light chain variable regions (NZM VH and NZM VL) to the human scaffold antibody counterparts (21/28'CL and REI) used for NZM humanization. Mutated residues are shown in red. Dots indicate the same residue. **b,** Scheme of the 8 heavy and 8 light chains variants of NZM that were combined in an 8x8 matrix to express 64 different NZM CDR swap variants. **c,** Cluster analysis of binding of 30 antibodies isolated from patient A to the 64 NZM swap variants by ELISA. BAbs and NAbs are indicated on the x-axis in red and black, respectively. The NZM swap variants are shown on the right y-axis (H, heavy chain; L, light chain; 1, CDR1; 2, CDR2; 3, CDR3). Optical density (OD) values are shown with a two-color gradation scale from minimum (white) to maximum (blue).
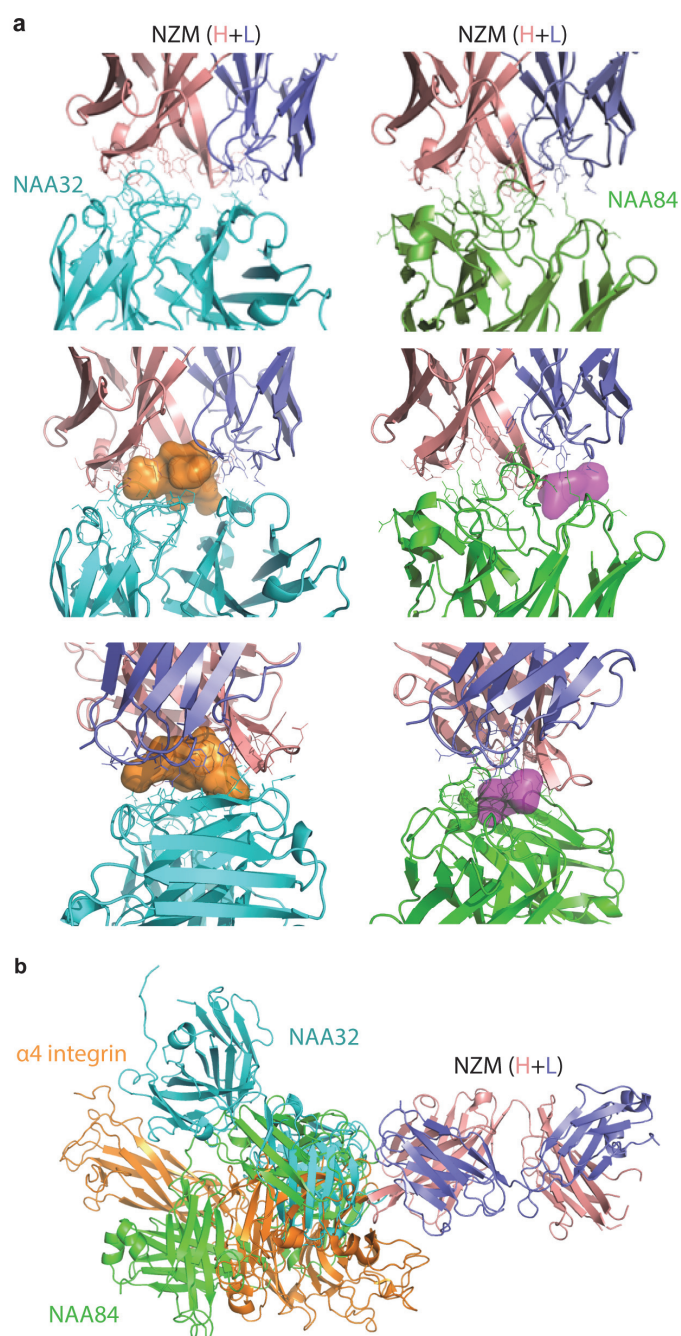
Extended Data Figure 6.1-2. Structural details of the interaction of NZM with a NAb, a BAb and α4-integrin. a, Closer view of the interaction interface between NZM and NAA32 (left) and NAA84 (right). Epitope and paratope residues are shown in solid sticks. Proteins are displayed in ribbon diagram. The empty space in the interface between the NZM and NAA32 or NAA84 is represented as orange or purple surface, respectively, in two different orientations **b,** Superimposition of the antigen-binding fragment of NZM in complex with NAA84 (NAb, green), NAA32 (BAb, cyan) and α4-integrin (orange). NZM heavy and light chains are shown in salmon and slate blue, respectively. Proteins are displayed in ribbon diagram.
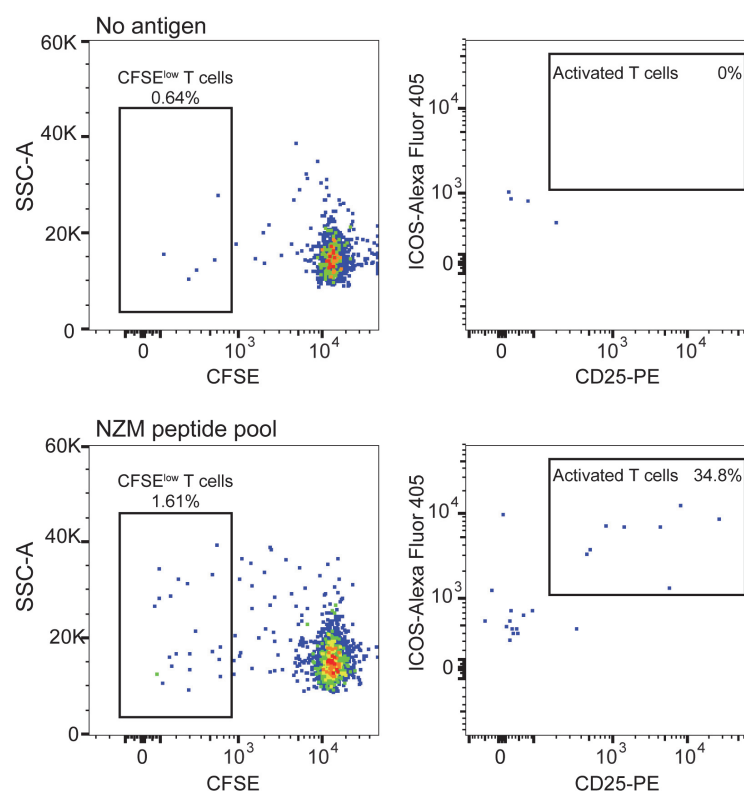
Extended Data Figure 6.1-3. Sorting of NZM-activated memory CD4+ T cells from MS patients.
Flow cytometry analysis of memory CD4+ T cells at day 12 after *ex-vivo* stimulation with irradiated autologous monocytes untreated (upper panels) or pre-pulsed with NZM peptide pool (lower panels). CFSE^{low}CD25+ICOS+ T cells reactive to NZM peptide pool were FACS-sorted and cloned by limiting dilution (representative of n = 2 biologically independent samples).
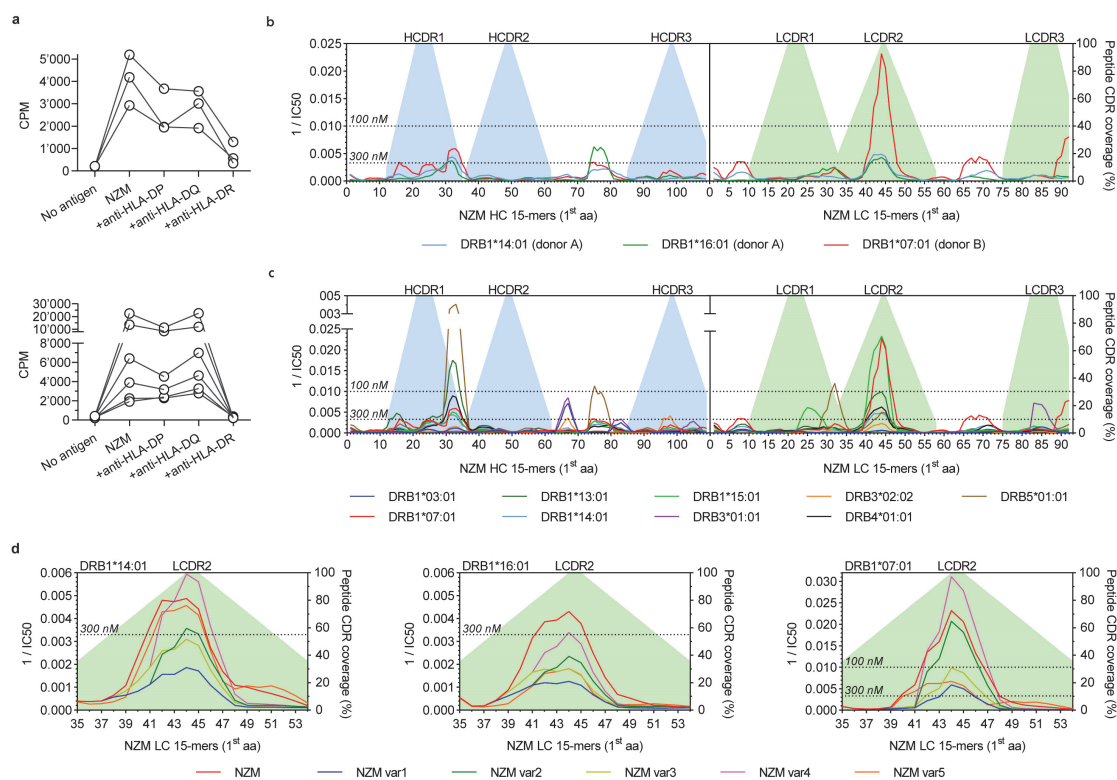
Extended Data Figure 6.1-4. MHC restriction of NZM-reactive CD4[+] T cell clones and peptide-MHC-II binding affinity predictions of NZM and deimmunized variants. a, MHC restriction of NZM-reactive T cell clones. NZM-specific CD4[+] T cell clones isolated from patient A (upper panel) and patient B (lower panel) were stimulated with antigen-pulsed autologous APCs in the absence or presence of blocking anti-MHC-II antibody (anti-HLA-DR, clone L243; anti-HLA-DQ, clone SPVL3; anti-HLA-DP, clone B7/21). Proliferation was measured on day 3 after a 16-h pulse with [3H]-thymidine, and is expressed as counts per minute (cpm). Inhibition of T cell proliferation was >80% only in the presence of the anti-HLA-DR antibody. **b** and **c,** Predicted binding affinities of all theoretical 15mer peptides derived from NZM heavy chain (HC) and light chain (LC) to HLA-DRB1 alleles carried by the two patients (b), or to a reference set of nine HLA-DRB1 and HLA-DRB3/4/5 alleles (c). The affinities are shown as reciprocal IC50 (nM) values. The dotted lines define the thresholds of high-affinity binding set at 100 nM and low-affinity binding set at 300 nM. **d,** Predicted binding affinities of 15mer peptides spanning the light chain CDR2 region of NZM variants to HLA-DRB1 alleles carried by patient A (DRB1*14:01 and DRB1*16:01) and patient B (DRB1*0701). The affinities are shown as reciprocal median IC50 (nM) values. The dotted lines define the thresholds of high-affinity binding set at 100 nM and low-affinity binding set at 300 nM.
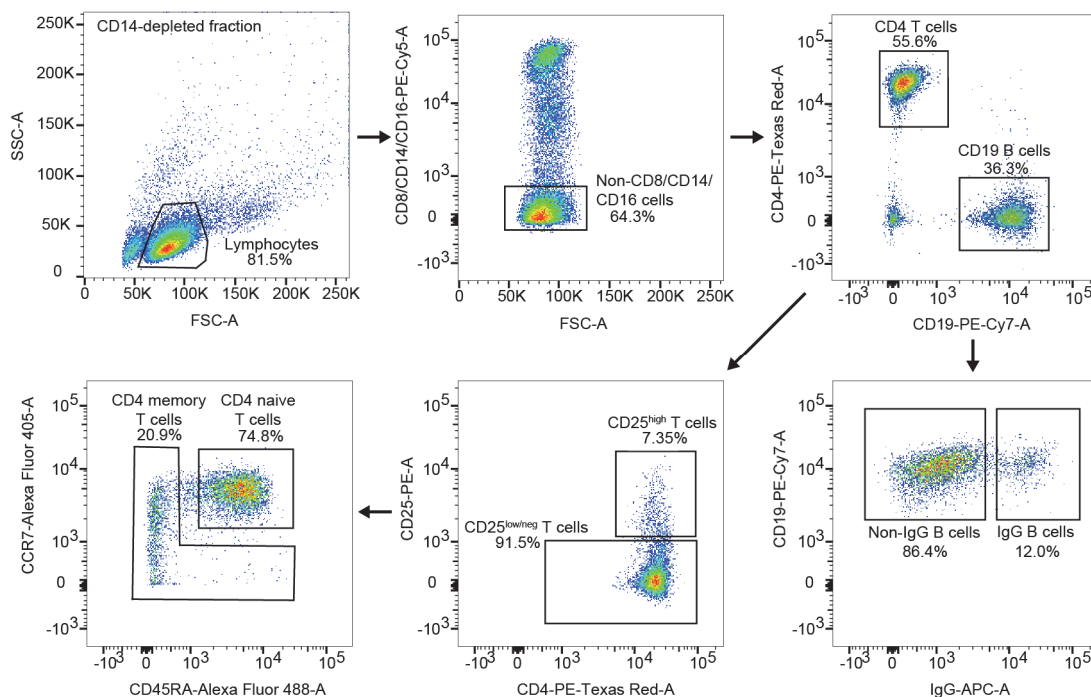
Supplementary Figure 6.1-1. Sorting strategy of memory IgG$^+$ B cells and memory CD4$^+$ T cells from MS patients. Representative flow cytometry analysis of CD14 depleted PBMCs stained with fluorochrome-labeled mouse monoclonal antibodies. Cells were FACS-sorted after gating on CD8$^-$, CD14$^-$, CD16$^-$ population. IgG$^+$ memory B cells were sorted from CD19$^+$ B cells. Memory CD4$^+$ T cells were sorted from the CD4$^+$CD25$^{low/neg}$ gate (to exclude CD4$^+$CD25^{bright} T$_{reg}$ cells), after removal of naïve CD4$^+$ T cells (CD45RA$^+$CCR7$^+$).

	Patient A		Patient B	
Gender	Female		Female	
Age at MS diagnosis	29		19	
Age at NZM first infusion	40		24	
Number of NZM infusions	3*		2*	
Type of hypersensitivity reaction to NZM	Anaphylactoid-like		Anaphylactoid-like	
▪ Symptoms	Urticaria, flushing, shivering and light dyspnea		Diffuse urticaria and flushing	
▪ Post-infusion therapy	Diphenhydramine and prednisolone		Chlorpheniramine and methylprednisolone	
ADA serum titers (EC50):				
▪ IgG	1/700		1/1000	
▪ IgM	nd		nd	
▪ IgA	1/30		1/170	
HLA typing:				
▪ HLA-DRB1	14:01P	16:01	07:01	07:01
▪ HLA-DQA1	01:04	01:02	02:01	02:01
▪ HLA-DQB1	05:03	05:02	02:02	02:02
▪ HLA-DPA1	01:03	01:03	02:01	01:03
▪ HLA-DPB1	04:01P	04:02P	01:01P	04:01P
▪ HLA-A	02:01	03:01	03:01	26:01
▪ HLA-B	27:02	35:03	14:01	57:02
▪ HLA-C	02:02	12:03	04:01	08:02

Supplementary Table 6.1-1. Features of the MS patients treated with NZM. nd = not detected; *last infusion was incomplete due to hypersensitivity reaction.

Patient	Antibody	Type	IC90 (ng/ml)	EC50 (ng/ml)	EC50$_{diss}$ (ng/ml)	k$_a$ (1/Ms)	k$_d$ (1/s)	KD (M)
	NAA104	BAb	100,000.0	35.1	16,704.0	1.1E+06	1.2E-03	1.1E-09
	NAA32	BAb	100,000.0	68.1	18,663.0	1.2E+06	1.0E-03	9.0E-10
	NAA40	BAb	100,000.0	103.0	37,878.0	1.3E+06	5.1E-03	3.9E-09
	NAA94	BAb	100,000.0	69.8	17,478.0	1.3E+06	6.3E-04	4.7E-10
	NAA96	BAb	100,000.0	32.6	29,967.0	6.8E+05	9.0E-05	1.3E-10
	NAA36	BAb	29,485.0	83.6	100,000.0	2.6E+05	1.0E-06	1.0E-12
	NAA2	BAb	24,849.0	59.2	100,000.0	4.5E+05	1.0E-06	1.0E-12
	NAA20	BAb	23,763.0	99.0	4,227.0	5.1E+05	1.1E-04	2.2E-10
	NAA49	BAb	7,965.0	51.8	100,000.0	6.6E+05	2.2E-05	3.4E-11
	NAA110	BAb	7,843.0	65.9	100,000.0	6.5E+05	2.2E-04	3.4E-10
	NAA9	BAb	2,859.0	78.6	100,000.0	5.7E+05	1.3E-04	2.3E-10
	NAA34	BAb	1,611.0	51.1	2,173.0	9.6E+04	6.5E-04	6.8E-09
	NAA59	NAb	271.1	31.0	220.0	7.1E+05	9.1E-07	1.3E-12
	NAA5	NAb	222.3	51.0	392.0	5.4E+05	1.0E-06	1.0E-12
	NAA33	NAb	194.6	58.3	332.0	4.7E+05	1.0E-06	1.0E-12
A	NAA7	NAb	187.9	69.5	100,000.0	6.2E+05	1.0E-06	1.0E-12
	NAA62	NAb	171.2	47.4	10,339.0	4.6E+05	3.6E-05	7.8E-11
	NAA116	NAb	147.3	44.0	594.2	4.2E+05	1.0E-06	1.0E-12
	NAA114	NAb	133.3	48.0	32,688.0	6.3E+05	1.0E-06	1.0E-12
	NAA105	NAb	88.8	50.5	889.6	6.1E+05	4.9E-05	8.0E-11
	NAA84	NAb	69.1	36.4	434.7	7.6E+05	8.4E-06	1.1E-11
	NAA92	NAb	60.8	33.2	37.6	3.4E+05	1.0E-06	1.0E-12
	NAA6	NAb	58.6	51.2	87.8	4.3E+05	1.0E-06	1.0E-12
	NAA113	NAb	57.7	30.2	589.8	5.6E+05	1.0E-06	1.0E-12
	NAA45	NAb	49.4	31.4	706.0	5.6E+05	5.2E-05	9.4E-11
	NAA44	NAb	48.0	43.8	141.1	8.0E+05	1.4E-04	1.8E-10
	NAA119	NAb	45.5	52.7	262.7	3.6E+05	1.0E-06	1.0E-12
	NAA80	NAb	43.7	46.0	43.0	9.9E+05	1.0E-06	1.0E-12
	NAA128	NAb	41.4	39.8	64.6	5.1E+05	1.0E-06	1.0E-12
	NAA125	NAb	17.1	48.4	531.3	5.2E+05	1.0E-06	1.0E-12
	NAE125	BAb	100,000.0	24.5	100,000.0	7.8E+05	2.3E-03	2.9E-09
	NAE194	BAb	100,000.0	25.6	100,000.0	1.3E+06	2.1E-03	1.6E-09
	NAE197	BAb	100,000.0	823.3	100,000.0	1.5E+06	1.6E-03	1.1E-09
	NAE199	BAb	100,000.0	23.3	15,761.0	1.8E+06	2.6E-03	1.5E-09
B	NAE205	BAb	100,000.0	736.3	100,000.0	2.0E+05	5.8E-04	2.9E-09
	NAE206	BAb	100,000.0	710.0	100,000.0	5.8E+06	1.3E-02	2.3E-09
	NAE207	BAb	100,000.0	972.6	100,000.0	1.4E+06	3.1E-03	2.2E-09
	NAE210	BAb	100,000.0	17.5	100,000.0	2.0E+06	8.5E-04	4.2E-10
	NAE208	BAb	100,000.0	6,433.0	100,000.0	1.9E+05	4.4E-03	2.3E-08
	NAE203	BAb	63,640.0	18.5	5,924.0	2.4E+05	9.0E-04	3.7E-09

Supplementary Table 6.1-2. Comparison of neutralizing and binding abilities of anti-natalizumab antibodies. IC90 indicates the antibody concentration required for 90% inhibition of NZM binding to T cells. EC50 indicates half-maximal antibody concentration required to bind to NZM by ELISA. EC50$_{diss}$ values were calculated by NZM ELISA in presence of high-pH dissociation buffer (maximum value 100,000 for no binding). Kinetics parameters measured by SPR: k$_a$, association constant, k$_d$, dissociation constant and KD, equilibrium dissociation constant (when no dissociation is measured, k$_d$ and KD are set at minimum values of 10^{-6} and 10^{-12}, respectively). NAb, neutralizing antibody (IC90 < 1,000 ng/ml). BAb, binding (non-neutralizing) antibody (IC90 > 1,000 ng/ml, maximum 100,000). Values are shown with a two-color gradation scale from minimum (green) to maximum (red).

	NZM-NAA32	NZM-NAA84
Data collection		
Space group	P4212	P21
Cell dimensions		
a, b, c (Å)	189.1, 189.1, 87.26	105.4, 66.6, 199.5
α, β, γ (°)	90, 90, 90	90, 98.9, 90
Resolution (Å)	1.96 (2.22-1.96)	2.74 (3.14-2.79)
R_{merge}	0.132 (1.529)	0.143 (0.583)
$I / \sigma I$	11.9 (1.8)	6.3 (1.8)
Completeness (%)	99.8 (100)	98.6 (99.2)
Redundancy	8 (12.5)	3 (3.1)
Refinement		
Resolution (Å)	2.0	2.8
No. reflections	63528	40000
R_{work} / R_{free}	0.266 / 0.285	0.238 / 0.263
No. atoms		
Protein	6714	13096
Ligand/ion	18	0
Water	253	0
B-factors		
Protein	39.3	56.8
Ligand/ion	48.5	na
Water	33.5	na
R.m.s. deviations		
Bond lengths (Å)	0.008	0.007
Bond angles (°)	1.08	1.03

Supplementary Table 6.1-3. Data collection and refinement statistics (molecular replacement). One crystal for each complex was used to solve the structures. Values in parentheses are for highest-resolution shell. na, not applicable.

Patient	Nr. of clones	Vβ gene	Jβ gene	Vβ CDR3	Specificity	Epitope	Start	End	MHC restriction
	6	V2	J2-3	CASGFTDTQYF	NZM LC	GKAPRLLIHYTSALQPGI	41	58	HLA-DR
	1	V7-2	J2-3	CASSLYSGANTDTQYF	NZM LC	GKAPRLLIHYTSALQPGI	41	58	HLA-DR
A	3	V7-3	J1-2	CASSSQRDSPYTF	NZM LC	GKAPRLLIHYTSALQ	41	55	nd
	1	V11-2	J1-5	CASSLDNGAGYNQPHHF	NZM LC	KYMAWYQQTPGKAPR	31	45	nd
	1	V20-1	J1-5	CSASTRQGVGRNQPQHF	NZM LC	GKAPRLLIHYTSALQPGIPS	41	60	HLA-DR
	3	V28	J1-5	CASSPAGAYSNQPQHF	NZM HC	PGQRLEWMGRIDPAN	41	55	nd
	1	V3-1	J1-4	CASSPPGVANEKLFF	NZM LC	KYMAWYQQTPGKAPR	31	45	HLA-DR
	11	V12-3	J2-1	CASSPTGTGLNEQFF	NZM LC	GKAPRLLIHYTSALQPGI	41	58	HLA-DR
B	2	V12-3	J2-7	CASSPTGTGLTEQYF	NZM LC	GKAPRLLIHYTSALQPGI	41	58	HLA-DR
	2	V12-4	J1-3	CASSXPPGRVGGNTIYF*	NZM LC	GKAPRLLIHYTSALQPGI	41	58	nd
	1	V5-1	J1-5	CASSYLPDSNQPQHF	NZM LC	GKAPRLLIHYTSALQPGI	41	58	HLA-DR
	7	V29-1	J2-7	CSVAQGGGSYEQYF	NZM LC	PRLLIHYTSALQPGI	44	58	HLA-DR

Supplementary Table 6.1-4. TCR Vβ sequence analysis and MHC restriction of NZM-reactive T cell clones. LC, light chain; HC, heavy chain. Start and end position of the epitope residues in the NZM chains are indicated. nd, not determined. *X can be either S or Y.

Protein	Peptide sequence	Start	End	Length	Identification (patient-replicate)				Uniquely present In NZM
					A-rep1	A-rep2	B-rep1	B-rep2	
NZM heavy chain	DTYIHWVRQAPGQR	31	44	14			+		*
NZM heavy chain	DTSASTAYMELSSLRSEDT	73	91	19	+				
NZM heavy chain	ASTAYMELSSLRSEDT	76	91	16	+	+			
NZM heavy chain	ASTAYMELSSLRSEDTA	76	92	17	+	+			
NZM heavy chain	STAYMELSSLRSED	77	90	14	+	+			
NZM heavy chain	STAYMELSSLRSEDT	77	91	15	+	+			
NZM heavy chain	STAYMELSSLRSEDTA	77	92	16	+	+			
NZM heavy chain	TAYMELSSLRSEDT	78	91	14	+				
NZM heavy chain	GQGTLVTVSSASTKGP	114	129	16	+				*
NZM heavy chain	GQGTLVTVSSASTKGPS	114	130	17	+				*
NZM heavy chain	QGTLVTVSSASTKGP	115	129	15	+	+			*
NZM heavy chain	QGTLVTVSSASTKGPS	115	130	16	+	+			*
NZM light chain	TPGKAPRLLIHYTSALQP	39	56	18	+				*
NZM light chain	TPGKAPRLLIHYTSALQPG	39	57	19	+		+	+	*
NZM light chain	TPGKAPRLLIHYTSALQPGIPS	39	60	22	+				*
NZM light chain	TPGKAPRLLIHYTSALQPGIPSR	39	61	23	+	+			*
NZM light chain	KAPRLLIHYTSALQPG	42	57	16		+			*
NZM light chain	APRLLIHYTSALQP	43	56	14	+	+			*
NZM light chain	APRLLIHYTSALQPG	43	57	15	+	+			*
NZM light chain	APRLLIHYTSALQPGIPS	43	60	18	+				*
NZM light chain	TFTISSLQPEDIAT	72	85	14	+				*

Supplementary Table 6.1-5. NZM peptides identified by mass spectrometry-based peptidomics. Start/end position and length of the peptide residues in the NZM chains are indicated. + indicates identification of a peptide in an experiment replicate. * indicates a peptide belonging uniquely to NZM and not to the B cell receptors of the antigen-presenting cells.

	OD values									% of binding to NZM							Range
	NZM	H1	H2	H3	L1	L2	L3	Neg ctr		H1	H2	H3	L1	L2	L3		
NAA2	4.048	0.124	0.124	0.134	3.907	3.926	4.1	0.123		0.03	0.03	0.28	96.41	96.89	101.32		0
NAA5	4.1	0.122	0.155	0.129	3.985	3.716	0.132	0.123		-0.03	0.80	0.15	97.11	90.34	0.23		10
NAA6	4.1	2.743	0.149	0.123	3.899	4.1	3.897	0.123		65.88	0.65	0.00	94.95	100.00	94.90		20
NAA7	4.046	0.124	0.134	0.127	3.174	3.921	3.971	0.126		-0.05	0.20	0.03	77.76	96.81	98.09		30
NAA9	4.1	0.125	0.129	0.117	3.918	4.006	3.714	0.128		-0.08	0.03	-0.28	95.42	97.63	90.28		40
NAA20	4.1	0.125	0.346	0.129	3.96	0.131	0.325	0.121		0.10	5.65	0.20	96.48	0.25	5.13		50
NAA32	4.1	0.123	0.128	0.128	0.12	0.124	0.124	0.123		0.00	0.13	0.13	-0.08	0.03	0.03		60
NAA33	4.1	3.745	3.621	0.12	0.13	0.134	3.653	0.125		91.07	87.95	-0.13	0.13	0.23	88.75		70
NAA34	3.986	0.13	0.643	0.128	3.749	0.136	2.716	0.137		-0.18	13.15	-0.23	93.84	-0.03	67.00		80
NAA36	4.1	0.117	0.121	0.124	0.112	4.015	2.997	0.119		-0.05	0.05	0.13	-0.18	97.86	72.29		90
NAA40	3.376	0.122	0.119	0.123	0.138	0.126	0.314	0.119		0.09	0.00	0.12	0.58	0.21	5.99		100
NAA44	4.1	0.121	0.127	0.124	2.741	0.126	1.212	0.121		0.00	0.15	0.08	65.85	0.13	27.42		
NAA45	4.027	0.121	0.122	0.124	3.938	4.013	2.916	0.122		-0.03	0.00	0.05	97.72	99.64	71.55		
NAA49	4.241	0.094	0.107	0.101	0.264	3.212	2.65	0.104		-0.24	0.07	-0.07	3.87	75.13	61.54		
NAA59	4.238	0.103	0.104	0.105	3.959	1.321	0.105	0.113		-0.24	0.02	-0.19	93.24	29.28	-0.19		
NAA62	4.1	3.882	1.979	0.124	3.859	4.1	0.123	0.128		94.51	46.60	-0.10	93.93	100.00	-0.13		
NAA80	4.067	3.901	3.869	3.565	3.974	4.1	0.139	0.137		95.78	94.96	87.23	97.63	100.84	0.05		
NAA84	4.1	0.613	0.152	0.129	3.828	4.075	3.755	0.137		12.01	0.38	-0.20	93.14	99.37	91.29		
NAA92	4.1	3.79	3.714	0.13	3.551	4.1	0.126	0.134		92.18	90.27	-0.10	86.16	100.00	-0.20		
NAA94	4.23	0.098	0.102	0.095	3.875	1.331	0.093	0.123		-0.61	-0.51	-0.68	91.36	29.41	-0.73		
NAA96	4.1	0.12	2.589	0.125	3.622	3.623	3.592	0.133		-0.33	61.91	-0.20	87.95	87.98	87.19		
NAA104	4.1	0.118	0.126	0.126	3.593	1.282	3.618	0.129		-0.28	-0.08	-0.08	87.23	29.04	90.08		
NAA105	4.1	0.118	0.134	0.122	3.88	4.1	3.618	0.124		-0.15	0.25	-0.05	94.47	100.00	87.88		
NAA110	4.1	0.125	0.367	0.126	0.912	0.297	0.501	0.131		-0.15	5.95	-0.13	19.68	4.18	9.32		
NAA113	4.209	0.167	0.131	2.915	4.059	4.154	3.934	0.174		-0.17	-1.07	67.93	96.28	98.64	93.18		
NAA114	4.197	0.108	0.163	0.107	3.984	4.089	3.854	0.14		-0.79	0.57	-0.81	94.75	97.34	91.55		
NAA116	4.224	0.138	0.139	0.137	4.007	4.192	3.926	0.164		-0.64	-0.62	-0.67	94.66	99.21	92.66		
NAA119	4.244	3.219	1.575	0.113	4.123	4.145	0.676	0.136		75.05	35.03	-0.56	97.05	97.59	13.15		
NAA125	4.202	0.132	0.118	0.122	3.984	4.157	3.948	0.138		-0.15	-0.49	-0.39	94.64	98.89	93.75		
NAA128	4.26	0.116	0.133	0.107	4.073	4.222	4.067	0.126		-0.24	0.17	-0.46	95.48	99.08	95.33		
NAE125	2.375	0.078	0.123	0.076	0.079	0.093	0.079	0.078		0.00	1.96	-0.09	0.04	0.65	0.04		
NAE194	2.651	0.525	0.069	0.065	0.466	2.186	0.215	0.068		17.69	0.04	-0.12	15.41	82.00	5.69		
NAE197	0.856	0.08	0.081	0.081	0.083	0.081	0.088	0.085		-0.65	-0.52	-0.52	-0.26	-0.52	0.39		
NAE199	2.776	0.067	0.068	0.067	1.398	0.493	0.559	0.07		-0.11	-0.07	-0.11	49.08	15.63	18.07		
NAE203	2.865	0.062	0.062	0.062	1.375	2.275	0.064	0.063		-0.04	-0.04	-0.04	46.82	78.94	0.04		
NAE205	2.152	0.064	0.065	0.063	1.271	1.238	0.295	0.065		-0.05	0.00	-0.10	57.79	56.21	11.02		
NAE206	0.637	0.066	0.067	0.068	0.069	0.069	0.07	0.07		-0.71	-0.53	-0.35	-0.18	-0.18	0.00		
NAE207	1.504	0.064	0.065	0.065	0.107	0.065	0.067	0.066		-0.14	-0.07	-0.07	2.85	-0.07	0.07		
NAE208	0.641	0.063	0.064	0.063	0.069	0.066	0.124	0.065		-0.35	-0.17	-0.35	0.69	0.17	10.24		
NAE210	2.857	0.063	0.063	0.078	1.93	1.445	2.683	0.065		-0.07	-0.07	0.47	66.80	49.43	93.77		

Source data of Figure 6.1-1. OD values and exact % binding values.

6.1.5 Methods

Patients and sample collection. Blood samples were collected from two MS patients (A and B) who had a hypersensitivity reaction following infusion of NZM and developed high titers of ADAs. Patient A was treated with NZM at the University Hospital for Neurology in Innsbruck, Austria, in 2014, while patient B was treated at the Mondino Foundation in Pavia, Italy, in 2018. Both patients provided written informed consent for this study. The study was approved by the Ethical committees of Innsbruck (UN2013-0040_LEK) and Pavia (P-20170027756). Blood samples were processed to obtain serum and peripheral blood mononuclear cells (PBMCs). For serum preparation, whole blood was collected in Vacutainer tubes (BD Biosciences) containing clot activators and kept at room temperature until a clot was formed. The tube was centrifuged at 2,000g for 10 min at 22°C, and the serum fraction was stored at −80°C. PBMCs were isolated from whole blood through Ficoll density gradient centrifugation and were resuspended in freezing medium for long-term storage in liquid nitrogen.

Production of NZM Fc-variants and CDR swap variants. A variant of NZM in which the human Fc was replaced with the murine counterpart (NZM-mFc) was produced by molecular cloning and used in ELISA and FACS assays to avoid reactivity with secondary anti-human Fcγ-specific antibodies. Briefly, synthetic genes expressing the NZM heavy chain and light chain variable regions (KEGG DRUG Database entry: D06886) were produced by Genscript and subcloned into vectors for expression of chimeric human CH1-murine IgG2a heavy chain (mFc) and human Igκ, respectively. The chains were expressed following transient transfection of these vectors into Expi293F cells (ThermoFisher Scientific) using polyethylenimine. Cell lines were routinely tested for mycoplasma contamination. NZM CDR swap variants were designed by aligning the sequences of NZM with those of the human antibodies (21/28'CL and REI) used for humanization (GenBank accession numbers AAA52825 and 751419A), synthetized by Genscript and subcloned into the mFc vector. NZM deimmunized variants were synthetized by Genscript and subcloned into vectors for expression of full human IgG4. The antibodies were purified by protein A or protein G chromatography (GE Healthcare) and concentrated by Amicon Ultra filter units (100K, Millipore). For SPR, mass-spectrometry and T cell experiments, NZM IgG was purified

from TYSABRI® drug solution for infusion (Biogen). Total IgGs were quantified by Pierce BCA protein assay (ThermoFischer).

B cell and T cell sorting. Monocytes were isolated from PBMCs by positive selection using CD14 magnetic microbeads (Miltenyi Biotech). CD14-depleted fractions were stained on ice for 15–20 min with the following fluorochrome-labeled mouse monoclonal antibodies: CD8-PE–Cy5 (clone B9.11; cat. no. A07758), CD14-PE–Cy5 (clone RMO52; cat. no. A07765), CD16-PE–Cy5 (clone 3G8; cat. no. A07767), CD45RA-FITC (clone ALB11; cat. no. A07786) from Beckman Coulter, CD19-PE–Cy7 (clone SJ25C1; cat. no. 341113), CD25–PE (clone M-A251; cat. no. 555432) from BD Biosciences, CD4–PE–Texas Red (clone S3.5; cat. no. MHCD0417) from ThermoFisher Scientific, CCR7–BV421 (clone G043H7; cat. no. 353208) from BioLegend, Alexa Fluor 647-conjugated goat anti-human IgG (cat. no. 109-606-170) from Jackson ImmunoResearch. IgG$^+$ memory B cells were sorted from CD19$^+$ B cells to over 98% purity on a FACSAria III (BD). Memory CD4$^+$ T cells were sorted to over 98% purity after exclusion of naïve CD4$^+$ T cells (CD45RA$^+$CCR7$^+$) and CD8$^+$, CD14$^+$, CD16$^+$, CD19$^+$, CD25^{bright} cells.

B cell immortalization and isolation of monoclonal antibodies. B cells were cultured in RPMI 1640 medium supplemented with 2 mM glutamine, 1% (v/v) nonessential amino acids, 1% (v/v) sodium pyruvate, penicillin (50 U/ml), streptomycin (50 μg/ml) (all from Invitrogen) and 10% fetal bovine serum (HyClone, characterized, GE Healthcare Life Science). Sorted IgG$^+$ memory B cells were immortalized with Epstein–Barr virus (EBV) and plated in single-cell cultures in the presence of CpG-DNA (2.5 μg/ml) and irradiated PBMC-feeder cells, as previously described (Traggiai et al., 2004). Two weeks post immortalization, the culture supernatants were tested by ELISA for binding to NZM-mFc as well as to a control antibody of an irrelevant specificity. B cell cultures that tested positive only for NZM-mFc were isolated and expanded.

Sequence analysis of antibody cDNAs and production of recombinant antibodies. cDNA was synthesized from selected B cell cultures, and both the VH and VL genes were sequenced as previously described (Tiller et al., 2008). The genes that encoded the VH and VL and the number of somatic mutations were determined by analyzing the homology between the genes encoding the VH and VL sequences of the monoclonal

antibodies and the known genes encoding human V, D and J regions that are present in the international immunogenetics information system (IMGT) database (version 3.4.17) (Lefranc et al., 2009). Antibody-coding sequences were amplified and sequenced with primers specific for the V and J regions of the given antibody. Sequences of the unmutated common ancestor (UCA) of the VH- and VL-coding genes were constructed using IMGT/V-QUEST (Lefranc et al., 2009) and synthetized by Genscript. To calculate the frequency of mutations, the entire sequence of each heavy chain variable region was compared to the germline sequence to identify replacement (R) and silent (S) mutations. The selection strength was estimated using BASELINe framework which compares the observed frequencies of replacement and silent mutations with the expected ones (Yaari et al., 2012). All calculations were done using the "calcBaseline" and "groupBaseline" functions from SHazaM (version 0.1.11) (Gupta et al., 2015). Sequences encoding antibody heavy and light chains were cloned into vectors for expression of human IgG1, Igκ and Igγ, and the chains were expressed following transient transfection of these vectors into Expi293F cells, as described above. Selected antibodies were also expressed as antigen-binding fragments for X-ray crystallography analysis.

ELISA assays for screening and characterization of anti-NZM antibodies. Total IgGs were quantified using 96-well MaxiSorp plates (Nunc) coated with 10 μg/ml goat anti–human IgG (SouthernBiotech, cat. no. 2040-01). Plates were then blocked with PBS with 1% BSA and incubated with titrated monoclonal antibodies, using Certified Reference Material 470 (ERMs-DA470, Sigma-Aldrich) as a standard. Plates were then washed and incubated with 1/500 alkaline phosphatase (AP)-conjugated goat anti–human IgG (Southern Biotech, cat. no. 2040-04). Substrate (para-nitrophenyl phosphate (p-NPP), Sigma) was added and plates were read at wavelength of 405 nm to determine optical density (OD) values. To test specific antibody binding, ELISA plates were coated with 1 μg/ml of NZM-mFc or a control antibody to test for nonspecific binding. Plates were blocked with PBS with 1% BSA and incubated with titrated sera or monoclonal antibodies, followed by 1/2,500 AP-conjugated goat anti-human IgG, Fcγ fragment specific antibody (Jackson ImmunoResearch, cat. no. 109-056-098), goat anti-human serum IgA, α chain specific (Jackson ImmunoResearch, cat. no. 109-055-011), or goat anti-human IgM, Fc5μ fragment specific (Jackson ImmunoResearch, cat. no.

109-055-129). In some dissociation assays, the samples were supplemented with 25% (vol/vol) of an alkaline dissociation buffer (2.5% Triton X-100, 2M ethanolamine, 0.15M NaCl, pH 11.6) one minute before the end of incubation. To test antibody binding to NZM CDR swap variants, ELISA plates were coated with 2 µg/ml goat anti–mouse IgG, human adsorbed (Southern Biotech, cat. no. 1030-01) and blocked with PBS with 1% BSA. After washing, the plates were incubated with 2 µg/ml of NZM CDR swap variants, followed by monoclonal antibodies at 2 µg/ml concentration and 1/2,500 AP-conjugated goat anti-human IgG, Fcγ fragment specific antibody. A heat map of the patterns of antibody binding to the 64 NZM CDR swap variants was computed by using the default clustering parameters of the heatmap.2 function from 'gplots' package in R in which the Euclidean metric and complete aggregation method were selected.

Surface plasmon resonance (SPR) assay. To study the kinetics of interaction of NZM-specific monoclonal ADAs, NZM IgG (50 nM) was stabilized in 10 mM acetate buffer, pH 4.5, and immobilized onto a EDC/NHS pre-activated ProteOn sensor chip (Biorad) through amine coupling; unreacted groups were blocked by injection of ethanolamine HCl (1 M). HEPES buffered saline (HBS) (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% surfactant Tween-20) was used as running buffer. All injections were made at flow rate of 100 µl/min. Monoclonal ADAs were diluted to 30 nM and injected onto the NZM coated chip; one channel of the chip was injected with HBS and used as reference for the analysis. Injection time and dissociation time were 240 s and 600 s, respectively. The binding interaction of each antibodies with NZM was assessed using a ProteON XPR36 instrument (BioRad). SPR data were processed with ProteOn Manager Software and k_a (1/Ms), k_d (1/s) and KD (M) parameters were calculated applying the Langmuir fit model.

NZM binding and inhibition of NZM binding assays. Serial dilutions of NZM and NZM variants IgG4 were prepared in MACS buffer (PBS 1% FBS, 2mM EDTA). T cells isolated from healthy donors were used as source of the cell adhesion molecule α4-integrin and added (50,000 cell/well) to the plates for 30 min, 4°C. T cells were washed and stained with 3.75 µg/ml Alexa Fluor 647–conjugated goat anti-human IgG (Jackson ImmunoResearch, cat. no. 109-606-170) for 30 min, 4°C. Cells were washed and analyzed by FACS. NZM binding was calculated as percentage of IgG$^+$ stained cells. To

study NZM neutralization, NZM-mFc was diluted to 5 ng/ml (final concentration) in MACS buffer (PBS 1% FBS, 2mM EDTA) and incubated with titrated monoclonal antibodies for 1 h, 37°C. T cells were added to the plates for 30 min, 4°C, then washed and stained with secondary goat anti-mouse-IgG-PE (SouthernBiotech, cat. no. 1030-09) at 1 μg/ml for 30 min, 4°C. Cells were washed and analyzed by FACS. NZM neutralization was calculated for each well as percentage of inhibition of binding of NZM-mFc to T cells with the following formula: 1 – % of cells stained by NZM-mFc. Gates were defined based on negative and positive controls.

X-ray crystallography and structure-guided deimmunization. NZM, NAA32 and NAA84 antibodies were produced as antigen-binding fragments and purified on IMAC (HitrapTM IMAC HP, GE Healthcare). Stable complexes of NZM-NAA32 and NZM-NAA84 were prepared in D-PBS buffer (PBS Dulbecco's Gibco 14190-094) at 1.5:1 molar ratio, purified by size-exclusion chromatography (Superdex 200 , GE healthcare) and concentrated to 10 mg/ml. Crystal were grown by vapor diffusion with a well solution containing 1.7 M sodium malonate pH 6 at 292 K (NZM-NAA32) or 35 % PEG400, 200 mM NaCl, 4 % MPD, 100 mM MES pH 6 (NZM-NAA84), and they typically appeared within a week. Crystals were flash frozen and x-ray data were collected at the European Synchrotron Radiation Facility (Beamline ID30A-1, Grenoble) using MxCube2 software. Data collection and processing statistics are outlined in **Supplementary Table 6.1-3**. Structure solution was performed by molecular replacement using MOLREP (version 11.6.03) (Vagin and Teplyakov, 1997) through the CCP4 suite (version 7.0.058) (Winn et al., 2011) and crystallographic refinement was carried out using BUSTER (version 2.11.7) (Bricogne et al., 2017). The shape correlation statistic Sc of protein interface was calculated with CCP4. Superimposition was performed using the structure of the NZM-α4β7 integrin complex as a reference (4IRZ, https://www.rcsb.org/). Surface area is calculated based on Van der Waals surfaces of atoms of NZM lying at 4Å distance from any atom of NAA32 or NAA84. For deimmunization of NZM, the CDR2 of the light chain of the antibody was modelled by mutating residues that were positioned at more than 4.0 Å distance from α4-integrin as observed in the 4IRZ structure and did not affect the conformation of the antibody CDR region. A non-exhaustive search was performed and a number of potential mutations were selected for mutagenesis.

***Ex-vivo* stimulation of memory CD4[+] T cells.** T cells were cultured in RPMI 1640 medium supplemented with 2 mM glutamine, 1% (v/v) nonessential amino acids, 1% (v/v) sodium pyruvate, penicillin (50 U/ml), streptomycin (50 μg/ml) (all from Invitrogen) and 5% heat-inactivated human serum (Swiss Red Cross). Sorted memory CD4[+] T cells were labeled with 5-(and 6)-carboxyfluorescein diacetate succinimidyl ester (CFSE, ThermoFisher) and cultured at a ratio of 2:1 with irradiated autologous monocytes untreated or pre-pulsed for 2-3 h with a peptide pool (15mers overlapping of 10) covering the entire sequence of the variable region of the NZM heavy and light chains (NZM peptide pool, 3 μM per peptide, produced by A&A Labs). After 12 days, cells were stained with antibodies to CD25–PE (BD Biosciences, clone M-A251, cat. no. 555432) and ICOS–Pacific Blue (BioLegend, clone C398.4A, cat. no. 313522).

Isolation of NZM-specific T cell clones. Proliferating activated T cells from *ex-vivo* stimulated cultures were sorted as CFSE[low]CD25[+]ICOS[+] and cloned by limiting dilution. T cell clones reactivity was determined by stimulation with irradiated autologous monocytes or B cells, untreated or pre-pulsed for 2-3 h with NZM peptide pool (3 μM per peptide) or, in some experiments, with recombinant NZM (5 μg/ml). T cell clones proliferation was measured on day 3 after 16 h incubation with 1 μCi/ml [methyl-^{3}H]thymidine (Perkin Elmer). Positive T cell clones were selected based on a cut-off value of (i) counts per minute (cpm) with antigen and antigen-presenting cells (APCs) ≥ 1000, and (ii) stimulation index ≥ 1.5 (cpm with antigen and APCs / cpm with APCs only). To determine MHC restriction, stimulation assay was performed in the absence or presence of blocking anti-MHC-II monoclonal antibodies produced in house from hybridoma cell lines (anti-HLA-DR, clone L243 from ATCC, cat. no. HB-55; anti-HLA-DQ, clone SPVL3 (Spits et al., 1983); anti-HLA-DP, clone B7/21 (Watson et al., 1983)). Epitope mapping was performed by stimulation of T cell clones with irradiated EBV-immortalized B-cell (EBV-B) clones, untreated or pre-pulsed for 2-3 h with individual peptides (15mers overlapping of 10) covering the entire sequence of the variable region of the NZM heavy and light chains (3 μM per peptide).

Sequence analysis of TCR Vβ genes. Sequence analysis of rearranged TCR Vβ genes of NZM-specific T cell clones was performed as previously described (Latorre et al., 2018). Briefly, cDNA from individual T cell clones was obtained by reverse transcription of total RNA from 10^3-10^4 cells per reaction. Rearranged TCR Vβ genes

were PCR amplified using forward primer pool targeting Vβ genes, and reverse primer pairing to C1–C2 β-chain constant region. Sequence amplification was assessed through agarose gel electrophoresis; successfully amplified fragments were sequenced by Sanger method, and TCR sequence annotation was carried out by using IMGT/V-QUEST algorithm (Lefranc et al., 2009).

HLA typing and peptide-MHC-II binding affinity predictions. HLA genotype of the patients was determined by reverse sequence-specific oligonucleotides probes (revPCR-SSO) DNA typing (LABType, One Lambda Inc.) performed at the IRCCS San Matteo Hospital Foundation, Pavia, Italy. Predicted IC50 (nM) binding values of all theoretical NZM-derived peptides (15mers overlapping of 14) to HLA-DRB1 alleles carried by the two patients and to a reference set of nine HLA-DRB1 and HLA-DRB3/4/5 alleles including DRB1*0301, DRB1*0701, DRB1*1501, DRB3*0101, DRB3*0202, DRB4*0101, DRB5*0101 (Paul et al., 2015) and DRB1*13:01/14:01 (de la Hera et al., 2014) were calculated using NetMHCIIpan 3.2 server (http://www.cbs.dtu.dk/services/NetMHCIIpan-3.2/) (Jensen et al., 2018b).

Purification of MHC-II presented peptides. NZM-specific EBV-B clones isolated from the two patients were pulsed overnight with 5 µg/ml NZM at a cellular density of $5x10^6$ cells/ml. MHC-II complexes were purified from about 10^9 NZM-pulsed EBV-B cells with a protocol adapted from Bassani-Sternberg M. et al (Bassani-Sternberg et al., 2015). Briefly, the B cells were lysed with 0.25% sodium deoxycholate, 1% octyl-β-D glucopyranoside (Sigma), 0.2 mM iodoacetamide, 1 mM EDTA, and Complete Protease Inhibitor Cocktail (Roche) in PBS at 4°C for 1 h. The lysates were cleared by 20 min centrifugation at 18,000 g at 4°C, and MHC-II complexes were purified by immunoaffinity chromatography with the anti-HLA-DR/DP/DQ HB-145 monoclonal antibody produced in house from hybridoma cell line IVA12 (ATCC, cat. no. HB-145) and covalently bound to Protein-A Sepharose beads (Thermo Fisher Scientific). In detail, the cleared lysates were loaded 3 times into the affinity columns at 4°C, and subsequently washed at 4°C with 10 column volumes of 150 mM NaCl, 20 mM Tris•HCl, pH 8 (buffer A); 10 column volumes of 400 mM NaCl, 20 mM Tris•HCl, pH 8; 10 column volumes of buffer A; and finally 10 column volumes of 20 mM Tris•HCl, pH 8. The HLA-II complexes were eluted at room temperature by addition of 500 µl of 0.1 M acetic acid, in total five elutions for each sample. Small aliquots of each eluted

fraction were analyzed by 12% SDS-PAGE to evaluate yield and purity of MHC-II complexes. Sep-Pak tC18 (Waters, Milford, MA) cartridges were used for further separation of peptides from MHC-II subunits. The cartridges were prewashed with 80% acetonitrile (AcN) in 0.5% formic acid, followed by 0.2% trifluoroacetic acid (TFA), and subsequently loaded 3 times with each fraction eluted from the immunoaffinity column. After loading, the cartridges were washed with 0.2% TFA, and the peptides were separated from the more hydrophobic MHC-II chains by elution with 30% AcN in 0.2% TFA. The peptides were further purified using a Silica C18 column tip (Harvard Apparatus, Holliston, MA) and eluted again with 30% AcN in 0.2% TFA. Finally, the peptides were concentrated by vacuum centrifugation, and resuspended in 2% AcN, 0.1% TFA, 0.5% formic acid for MS analysis.

Liquid chromatography–mass spectrometry (LC-MS/MS) and data analysis. MHC-II peptides were separated on an EASY-nLC 1200 HPLC system coupled online to a Q Exactive mass HF spectrometer via a nanoelectrospray source (Thermo Fisher Scientific). Peptides were loaded in buffer A (0.1% formic acid) on in-house packed columns (75 μm inner diameter, 50 cm length, and 1.9 μm C18 particles from Dr. Maisch GmbH) and eluted with a non-linear 120 min gradient of 5%–60% buffer B (80% ACN, 0.1% formic acid) at a flow rate of 250 nl/min and a column temperature of 50°C. The Q Exactive was operated in a data dependent mode with a survey scan range of 300-1650 m/z and a resolution of 60,000 at m/z 200. Up to 10 most abundant isotope patterns with a charge ≥ 1 were isolated with a 1.8 Th wide isolation window and subjected to higher-energy C-trap dissociation (HCD) fragmentation at a normalized collision energy of 27. Fragmentation spectra were acquired with a resolution of 15,000 at m/z 200. Dynamic exclusion of sequenced peptides was set to 30 s to reduce the number of repeated sequences. Thresholds for the ion injection time and ion target values were set to 80 ms and 3E6 for the survey scans and 120 ms and 1E5 for the MS/MS scans, respectively. Data were acquired using the Xcalibur software (Thermo Scientific). MaxQuant software was used to analyze mass spectrometry raw files. MS/MS spectra were searched against the full-length NZM heavy and light chains sequences, the human Uniprot FASTA database, and a common contaminants database (247 entries) by the Andromeda search engine (Cox et al., 2011). N-terminal acetylation and methionine oxidation were set as variable modifications; no fixed modifications

were selected; the enzyme specificity was set to "unspecific" with a minimum peptide length of 8 amino acids. A false discovery rate (FDR) of 1% was required for peptides. Peptide identification was performed with an allowed precursor mass deviation of up to 4.5 ppm and an allowed fragment mass deviation of 20 ppm; "match between runs" option was disabled.

Statistical analysis. GraphPad Prism 7 software was used for to perform all the statistical analyses. EC50 (ng/ml) and IC90 (ng/ml) values were calculated for every antibody tested with the different ELISA assays and the inhibition of NZM binding assay, respectively, by nonlinear regression analysis using the GraphPad Prism 7 software. A two-tailed Spearman's correlation was performed to correlate NZM neutralization (IC90) with association constant (k_a) and dissociation constant (k_d). "n" indicates the number of antibodies tested. With 17 or more pairs, GraphPad Prism 7 software computes an approximate P value for nonparametric correlation. An extremely significant P value is indicated as $P < 0.0001$.

Data availability. All requests for raw and analyzed data and materials will be promptly reviewed by the Institute for Research in Biomedicine to verify if the request is subject to any intellectual property or confidentiality obligations. Patient-related data not included in the paper may be subject to patient confidentiality. Any data and materials that can be shared will be released via a Material Transfer Agreement. Source data of Fig. 6.1-1 are provided. Sequence data of the monoclonal antibodies isolated in this study have been deposited in GenBank (MN044260-MN044339). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD013599. The X-ray structure factors and coordinates have been deposited in the Protein Data Bank (access numbers are 6FG1 and 6FG2).

6.2 Deciphering and predicting CD4[+] T cell immunodominance of Influenza virus hemagglutinin

A manuscript by:

Antonino Cassotta[1,2], Philipp Paparoditis[1], Roger Geiger[1], Ramgopal R. Mettu[4], Samuel J. Landry[4], Alessia Donati[1], Marco Benevento[1], Mathilde Foglierini[1], David JM Lewis[3], Antonio Lanzavecchia[1] & Federica Sallusto[1,2]

[1] Università della Svizzera italiana (USI), Faculty of Biomedical Sciences, Institute for Research in Biomedicine, Bellinzona, Switzerland.

[2] Institute of Microbiology, ETH Zürich, Zürich, Switzerland.

[3] NIHR/Wellcome Trust Imperial Clinical Research Facility, Hammersmith Hospital, London, United Kingdom.

[4] Tulane University, New Orleans, Louisiana, USA.

6.2.1   Author contributions

A.C. characterized the T cell response by T cell library method, isolated T cell clones and sequenced their TCR Vβ, performed the sample preparation of MHC-II peptidomics, performed MHC binding predictions, performed bioinformatics analyses of TCR Vβ deep sequencing, analyzed the data and wrote the manuscript; P.P. immortalized memory B cells, performed screenings and isolated B cell clones; R.G. and M.B. acquired mass-spectrometry data; R.R.M. and S.J.L. performed bioinformatics analyses of epitope likelihood and MHC binding; A.D. provided assistance for the characterization of the T cell response; M.F. provided assistance for bioinformatics analyses of TCR Vβ deep sequencing; D.J.M.L. collected biological samples; A.L. provided supervision, designed the experiments and wrote the manuscript; F.S. provided overall supervision, designed the experiments, analyzed the data and wrote the manuscript.

6.2.2 Abstract

Influenza viruses represent a public health concern due to their pandemic potential and to the sporadic spread of highly pathogenic strains from zoonotic hosts to humans (Webster and Govorkova, 2014). In this study, we used T cell libraries to identify the influenza HA peptides recognized by naïve and memory CD4$^+$ T cells in individuals with different HLA haplotypes. While naïve T cells recognized a variety of peptides spanning the whole HA sequence, memory T cells were highly focused on just a few peptides. Interestingly, these immunodominant peptides were readily identified by mass spectrometry (MS)-based analysis of peptides eluted from MHC class II (MHC-II) molecules isolated from dendritic cells or antigen-specific B cell clones. When compared to MS-based MHC-II peptidomics, prediction of immunodominant epitopes using algorithms based uniquely on peptide binding was far less efficient, but could be improved by considering the structural accessibility of the antigen to proteolytic cleavage. Collectively, these findings indicate that processing of native proteins represents a major constraint determining the dominance of certain peptides and delineate new methods to determine T cell immunodominance.

6.2.3   Introduction

CD4[+] T lymphocytes orchestrate adaptive immune responses by secreting cytokines that promote multiple types of inflammatory responses in tissues and by providing help to B cells and CD8[+] T cells (Sallusto et al., 2010). For antigen recognition, CD4[+] T cells rely on the interaction with antigen-presenting cells (APCs) that take up, process and present antigen in the form of short linear peptides bound to MHC-II molecules (Roche and Furuta, 2015; Unanue et al., 2016). Typically, only a small fraction of the multitude of potentially immunogenic peptides contained in a complex foreign antigen is able to induce a measurable T cell response, with some peptides that are recognized with higher magnitude and thus arise as immunodominant, and others that remains subdominant or even cryptic (Sercarz et al., 1993; Yewdell and Bennink, 1999; Yewdell and Del Val, 2004).

Given the complexity and tight connection between antigen presentation and recognition, many factors may concur to peptide and T cell immunodominance. Some of those reflect the biochemical rules of antigen processing and MHC presentation, such as the molecular context in which the peptides are embedded (Graham et al., 2018; Kim and Sadegh-Nasseri, 2015; Landry, 2008; Mirano-Bascos et al., 2008), the affinity of the generated peptides for MHC-II binding, the resistance to HLA-DM-mediated editing of newly formed peptide MHC-II (pMHC-II) complexes (Kim and Sadegh-Nasseri, 2015; Mellins and Stern, 2014) or their kinetic stability on the cell surface of APCs (Sant et al., 2005). Furthermore, the heterogeneous set of proteolytic enzymes and endogenous inhibitors that different kinds of APCs are equipped with (Unanue et al., 2016), as well as the interactions with molecular partners that facilitate antigen uptake, such as B cell receptors (BCRs) or soluble antibodies (Watts and Lanzavecchia, 1993), can affect the antigen processing and the composition of the MHC-II-presented peptidome. Other variables influencing T cell immunodominance depend on the architecture of the T cell repertoires and the mechanisms of antigen recognition (Yewdell, 2006), such as the availability of antigen-specific naïve precursors (Jenkins and Moon, 2012; Moon et al., 2007), the interaction affinity of their TCRs with pMHC-II complexes (Malherbe et al., 2004) or the occurrence of TCR cross-reactivity to similar antigenic peptides (Campion et al., 2014; Nelson et al., 2015; Su et al., 2013).

In this study, we chose influenza A virus as a model infectious agent that triggers complex adaptive immune reactions comprising both humoral and cellular responses. Influenza viruses infect every year over a billion people worldwide, being the cause of prominent economic loss as well as significant morbidity and mortality, especially in children under five and adults over 65 (Krammer et al., 2018; Lee et al., 2019; Zens and Farber, 2015). Despite great efforts in research, vaccines are only moderately effective against seasonal strains and are challenged by the rapidly evolving nature of influenza viruses that occasionally emerge as new strains causative of serious epidemics or pandemics (Angeletti and Yewdell, 2018; Krammer et al., 2018; Webster and Govorkova, 2014; Zens and Farber, 2015). We focused our attention on hemagglutinin (HA) that represents the main target of antibody response to influenza virus upon vaccination or infection (Angeletti and Yewdell, 2018; Corti et al., 2011; Krammer et al., 2018; Lee et al., 2019; Pappas et al., 2014). The detailed characterization of HA-reactive memory and, for the first time, naïve CD4$^+$ T cell repertoires, paralleled by a deep analysis of HA processing and presentation on MHC-II molecules, allowed us to shed new light on the factors governing CD4$^+$ T cell clonal selection and immunodominance to influenza HA in humans.

6.2.4   Results

Memory T cells target an immunodominant region of influenza H1 hemagglutinin

To capture the entire repertoire of memory T cells specific for H1 hemagglutinin, we obtained multiple and large blood samples from a donor (HD1) after vaccination with the 2013/14 seasonal Influenza vaccine Inflexal containing the A/California/7/2009 H1N1 HA (H1-HA). Central memory (Tcm), effector memory (Tem) and circulating follicular helper (cTfh) CD4$^+$ T cells were isolated by cell sorting, labelled with CFSE and stimulated with Inflexal. When analyzed on day 6, proliferating Influenza-specific T cells were detected in all three memory subsets from blood samples obtained 6 and 12 months after vaccination (**Fig. 6.2-1 A**). To select H1-HA-reactive T cells, the CFSE-negative T cells were sorted, relabeled with CFSE and stimulated with H1-HA (**Fig. 6.2-1 B**). T cells proliferating in the secondary stimulation were cloned and 456 H1-HA-specific clones were isolated (**Fig. 6.2-1 C** and **Supplementary Table 6.2-1**) and characterized for peptide specificity, MHC restriction and TCR Vβ usage. Strikingly, more than 85% of the clones isolated (393 out of 456) recognized two overlapping peptides (H1-HA$_{401-420}$ or H1-HA$_{411-430}$), thus defining, in this individual, a highly immunodominant region (**Fig. 6.2-1 D**). T cells specific for the immunodominant H1-HA$_{401-430}$ region were found in all 3 memory subsets (94 clones in Tcm, 112 clones in Tem and 187 clones in Tfh) and were HLA-DR restricted, as shown by antibody blocking experiments (**Supplementary Fig. 6.2-1, A**). Several T cell clones specific for subdominant H1-HA regions were also HLA-DR restricted, with a minority being HLA-DQ or HLA-DP restricted (**Supplementary Fig. 6.2-1, A-C**).

TCR-Vβ Sanger sequencing performed on 274 T cell clones showed that the response to H1-HA was highly polyclonal, comprising 88 distinct clonotypes, even when directed against the immunodominant region (**Supplementary Table 6.2-2**). For instance, 62 T cell clones specific for the immunodominant peptide H1-HA$_{401-420}$ included 16 different clonotypes and 157 T cell clones specific for the immunodominant peptide H1-HA$_{411-430}$ included 39 different clonotypes (**Fig. 6.2-1, E** and **Supplementary Table 6.2-2**). Of note, 26 out of the 39 H1-HA$_{411-430}$-specific clonotypes used the TRBV19 gene, suggesting a preferential TCR CDR1 and CDR2 usage that might facilitate cognate interaction with the peptide-MHC complex.

Tracking of HA-reactive T cell clonotypes within the CFSE-negative T cell population responding to Inflexal (in Fig. 6.2-1, A) showed that those against the immunodominant H1-HA region were among the most represented and that some were also found in the Tcm, Tem or Tfh repertoire *ex vivo* (**Fig. 6.2-1, F**). Strikingly, several of these clonotypes were still detected in memory T cell subsets isolated from donor HD1 48 months later (**Fig. 6.2-1, G**).

Collectively, these findings indicate that in an Inflexal-immunized donor a polyclonal memory CD4$^+$ T cell response to influenza HA is highly focused on a small immunodominant region.

Memory T cells are focused against immunodominant regions while naïve T cells recognize multiple peptides spanning the entire H1-HA sequence

We next investigated whether the immunodominance observed in the memory repertoire is a general phenomenon and whether it is reflected in the naïve repertoire of a given individual. To address this question, we employed the highly sensitive T cell library method (Geiger et al., 2009) to screen naïve and total memory CD4$^+$ T cells from HD1 and from 3 other immune donors with a diverse HLA background (**Supplementary Table 6.2-3**). For each donor, naïve and memory T cells were polyclonally expanded in multiple cultures (each containing 2,000 cells) in the presence of PHA, IL-2 and feeder cells. For a broad and unbiased screening of T cell reactivity against H1-HA, the T cell libraries were then screened using overlapping 15mer peptides covering the entire A/California/7/2009 H1N1 HA sequence. In all four donors tested, H1-HA-peptide-specific T cell clones were readily detected in naïve and memory libraries although, as expected, their frequencies measured in the naïve libraries were lower compared to that measured in the memory libraries (**Fig. 6.2-2, A and B**). Epitope mapping of memory T cell clones confirmed in all 4 donors a skewing towards one or two immunodominant regions, which for donor HD1 coincided with those detected by antigen-driven proliferation of memory T cell subsets (**Fig. 6.2-2, C**). Strikingly, however, epitope mapping of naïve T cell clones showed that these cells covered a broad range of peptide specificities (**Fig. 6.2-2, C and D**). This pattern was particularly evident for donor HD1 that was analyzed at high depth. In this donor, naïve

T cells with diverse TCR-Vβs recognized peptides spanning virtually all the HA sequence (**Fig. 6.2-2, C** and **Supplementary Table 6.2-2**).

Collectively, these findings demonstrate that the naïve T cell repertoire has a very broad coverage of the HA sequence and that only a fraction of this repertoire is selected in the memory repertoire.

Naïve and memory T cell clones show different functional avidities for peptide and naturally processed H1-HA

A plausible explanation for the selection of immunodominant peptides is their binding affinity to MHC molecules and/or TCR. We therefore measured the binding affinity of dominant and subdominant peptides to the recombinant HLA-DR molecules of donor HD1 (alleles *DRB1*01:01* or **08:01*) (**Supplementary Table 6.2-3**). *In vitro* refolding assay carried out in the presence of titrated peptides showed that the immunodominant H1-HA$_{401-420}$ and H1-HA$_{411-430}$ peptides bound to HLA-DRB1*08:01 molecules with an affinity that was comparable to that of subdominant peptides binding to the same HLA-DR molecules (**Supplementary Fig. 6.2-1, C**), indicating that immunodominance of these peptides is not explained by preferential binding to MHC class II molecules.

We then selected a large number of naïve and memory T cell clones from donor HD1 and determined their functional avidity by measuring the proliferative response to autologous monocytes pulsed with different concentrations of the H1-HA peptides or of H1-HA protein, that need processing for presentation on MHC-II molecules. Functional avidity, measured in response to peptide stimulation and expressed as EC50 value, was spread over almost 3 logs for both memory and naïve T cell clones, with memory clones being enriched for high avidity cells (**Fig. 6.2-3, A** and **Supplementary Fig. 6.2-2, A and C**), consistent with previous observations on the response to tetanus toxoid (Geiger et al., 2009). When the same clones were tested for their response to protein, several observations were made. First, all memory clones responded to H1-HA protein and there was an overall correlation between EC50 values for protein and peptide, with the immunodominant clones, such as M1 specific for the H1-HA$_{411-430}$, showing the highest avidity for both peptide and protein (red dots in **Fig. 6.2-3, B**). However, there were a

few notable exceptions. For instance, the subdominant clones M2 and M3, specific for H1-HA$_{396-410}$ and H1-HA$_{526-540}$, had high functional avidity for peptide, comparable to clone M1, but showed 1000-fold lower avidity for the H1-HA protein (**Fig. 6.2-3, B**). Second, memory T cell clones restricted by HLA-DP (M4 specific for H1-HA$_{121-140}$) or by HLA-DQ molecules (M5 and M6 specific for H1-HA$_{186-200}$ and H1-HA$_{446-460}$, respectively) recognized peptides and protein with functional avidity lower than the median of the distribution (**Fig. 6.2-3, A and B**). Third, several naïve T cell clones, including some with intermediate avidity for peptides, did not proliferate in response to H1-HA protein (**Fig. 6.2-3, C**). Finally, and most importantly, the differential recognition of dominant and subdominant epitopes in response to H1-HA was also observed for naive T cell clones. For instance, clone N1, derived from naïve cells and specific for the immunodominant H1-HA$_{411-430}$ peptide, had high functional avidity for both peptide and naturally processed H1-HA, comparable to that of memory clone M1, while the naïve clone N2 and N3 (specific for the subdominant H1-HA$_{521-540}$ and H1-HA$_{536-555}$ peptides) had high avidity for peptide but 100-fold lower avidity for protein, similar to what observed for memory clones M2 and M3 (**Fig. 6.2-3, C**).

Collectively, these findings indicate that in both the naïve and memory repertoires immunodominant peptides are recognized with high avidity by T cells. They also reveal the presence of a hidden broad repertoire of naïve T cells for cryptic T cell epitopes. Finally, they show that the functional avidity of peptide recognition is not predictive of the response to the naturally processed antigen, suggesting a major role for antigen processing in determining the abundance of the processed peptide generated.

The MHC class II peptidome defines immunodominant regions and reveals modulation for antigen processing by antibodies

Based on the above results, we hypothesize that CD4$^+$ T cell immunodominance could be primarily related to the yield of peptides generated by antigen processing by APCs. We therefore used mass spectrometry (MS)-based immunopeptidomics to identify peptides naturally presented on MHC class II molecules by different types of APCs. Briefly, monocyte-derived DCs from donor HD1 and EBV-immortalized B cell clones carrying surface BCR specific for HA from all 4 donors were pulsed overnight

with recombinant H1-HA. MHC-II molecules were isolated from lysed cells using a pan-anti-MHC-II antibody and peptides were purified by reversed phase chromatography. Using this approach, we identified thousands of total MHC-II bound individual peptides derived from self-molecules as well as 3-35 HA-derived peptides (median 9) with the expected length peaking at 12-15 amino acids (**Fig. 6.2-4, A and B**). Remarkably, almost 60% of the H1-HA-derived peptides presented by DCs of donor HD1 corresponded to a rich set of nested peptides overlapping the immunodominant regions $HA_{401-430}$ targeted by memory T cells, suggesting that peptides in this region are presented in high abundance and are recognized by T cells with high functional avidity (**Fig. 6.2-4, C** and **Supplementary Table 6.2-4**). A similar correspondence between immunodominant peptides recognized by memory T cells and peptides presented on MHC-II molecules was found in the other 3 donors analyzed (**Fig. 6.2-4, C**).

The antibody response to Influenza virus is directed against several regions of HA, such as the highly variable globular head or the conserved stem region, that can be target of neutralizing antibodies with higher breadth to multiple viral subtypes (Corti et al., 2011). The MHC-II peptidome of H1-HA-pulsed B cell clones with distinct epitope specificity isolated from donor HD1 revealed that a clone specific for the HA globular head presented the H1-HA immunodominant region in a similar manner as DCs, whereas a B cell clone specific for the HA stalk region was a poorer presenter of H1-HA-derived peptides (**Supplementary Fig. 6.2-3, A and B**). Indeed, even if we resolved a comparable total number of MHC-II presented peptides in the two kind of B cells (**Supplementary Fig. 6.2-3, A**), from the anti-stem clone we did not detect any H1-HA-derived peptides corresponding to the dominant region $HA_{401-420}$, neither any subdominant epitopes from the HA1 domain (**Supplementary Fig. 6.2-3, B**). Consistently, presentation of the recombinant H1-HA by anti-stem B cells resulted in a much lower activation and proliferation of T cell clones specific for $HA_{401-420}$ or $HA_{241-260}$ peptides (both HLA-DR-restricted), resulting in a 10-fold reduction of their functional avidity as compared to antigen presentation by anti-head B cells (**Supplementary Fig. 6.2-3, C and D**). Coherently, HLA-DR-restricted T cell clones specific for the $HA_{416-430}$ or $HA_{386-400}$ peptides detected by MHC-II immunopeptidomics both on anti-head and anti-stem B cells clones, showed comparable functional avidity when stimulated with either kind of APCs (**Supplementary Fig. 6.2-3, C and D**).

Taken together, these data show that the MHC-II peptidome presented by professional APCs defines the immunodominant regions of HA recognized by memory CD4[+] T cells. Moreover, the spectrum of peptides naturally presented by B cells can be modulated by antibody binding to HA, thus potentially being an additional variable affecting B cell clonal selection during T cell-dependent immune responses.

Epitope prediction can be improved by combining binding algorithm with epitope likelihood based on structural constraints

The binding of processed peptides to the cleft of MHC-II molecules follows precise rules which have been instrumental to generate sophisticated algorithms able to predict with high accuracy the affinity of binding (Jensen et al., 2018b; Unanue et al., 2016; Wang et al., 2010). Accordingly, virtually all the HA 15mer peptides *in silico* predicted as good binders for MHC-II alleles carried by the four donors analyzed were recognized by T cell clones, isolated either from the naïve or the memory compartment (**Supplementary Fig. 6.2-4**). Nevertheless, it was surprising to notice that the immunodominant HA peptides did not show stronger binding for the MHC-II cleft as compared to subdominant or cryptic peptides, neither by *in silico* prediction nor, as demonstrated before, by *in vitro* measurement of MHC-II binding. These data suggest that the binding to MHC-II is a necessary, but not sufficient feature for T cell immunodominance. We therefore set out to explore further parameters that might improve the *in silico* predictions of HA-derived T cell epitopes.

The molecular context in which a peptide is embedded and its structural accessibility might influence the propensity of unfolding during the progressive pH acidification that occurs in the endocytic pathway, therefore affecting the exposition of denatured stretches of the antigen to the proteolytic environment of the late endosomes (Graham et al., 2018; Kim and Sadegh-Nasseri, 2015; Landry, 2008). To evaluate the role of structural constraints of HA in influencing the immunodominance observed in the memory repertoires, we adopted a recently developed algorithm (Mettu et al., 2016). In brief, an aggregate z-score of conformational stability was determined for each HA residue by integrating four structural parameters obtained from the 3D structure of post-fusion HA resolved by X-ray diffraction (PDB codes: 3LZG for HA1 domain (Xu et al.,

2010), 1HTM for HA2 domain in the post-fusion conformation (Bullough et al., 1994)). The z-score statistic was then used to calculate an epitope likelihood for each theoretical HA 15mer peptide, following the rationale that the liberation of antigenic peptides might be facilitated by surrounding unstable regions that are readily unfolded and targeted by endosomal proteases (**Fig. 6.2-5, A**). We then performed peptide binding affinity predictions of each theoretical 15mer HA peptide for the MHC-II alleles carried by each donor (**Supplementary Table 6.2-3**), considered for each peptide the best scoring affinity within each group of MHC-II alleles and built a combined predictor by iteratively weighting the contributions of epitope likelihood based on structural accessibility and peptide binding affinity to MHC-II, until we could maximize the Area under the Receiver Operating Characteristic Curve (AUROC) of the predictor using as reference the set of epitopes recognized by memory T cells of each donor (**Fig. 6.2-5, A**). With this approach we found 7-41 weighting combinations that maximize the AUROC for the combined prediction up to 0.79, outperforming the single predictors (**Fig. 6.2-5, B**). Interestingly, depending on the HLA background of each donor, we alternatively observed a preponderant role of peptide binding or epitope liberation likelihood, with the latter alone explaining up to 60% of the memory T cell epitopes in some donors (**Fig. 6.2-5, C**). Of note, all the immunodominant peptides identified in the memory repertoire of the four donors analyzed localized within a peak of high predicted epitope likelihood (as defined by a threshold of 0.6) (**Fig. 6.2-5, D**). On average, the top 20 predicted peptides accounted for more than 50% of the HA-specific memory T cell clones, with a sensitivity comparable to the set of peptides measured by MHC-II immunopeptidomics (**Fig. 6.2-5, E**). Thus, a combined analysis of structural accessibility and peptide binding affinity to MHC-II cleft may be helpful to improve prediction of the immunogenicity for CD4$^+$ T cell responses.

6.2.5 Discussion

In this study, we report the identification of influenza HA peptides recognized by naïve and memory CD4$^+$ T cells in individuals with different HLA haplotypes. While naïve T cells recognized a variety of peptides spanning the whole HA sequence, memory T cells were highly focused on just a few peptides. Interestingly, these immunodominant peptides were readily identified by MS-based analysis of peptides eluted from MHC class II molecules isolated from DCs or HA-specific B cell clones. Collectively, these findings indicate that processing of native proteins represents a major constraint determining the dominance of certain peptides and delineate new methods to identify cryptic and immunodominant T cell epitopes.

The identification and characterization of antigen-specific T cells in the naive and memory repertoire is of both fundamental and practical relevance. The high throughput T cell library method used in this study can rapidly identify, in different individuals, the range of peptides that can be recognized by T cells, thus determining, with a simple assay, both peptide binding to MHC-II and the presence of specific TCRs (Campion et al., 2014; Geiger et al., 2009; Latorre et al., 2018). We show that naïve T cells from donors with different HLA haplotypes recognize multiple peptides spanning the whole HA sequence, while memory T cells recognized only a few regions, of which one or two behaved as immunodominant. Considering the small size of the naïve T cell libraries analyzed (3-7x10^5 T cells) compared to the total naïve T cell pool, we can estimate that the naive repertoire contains a large number of HA-specific T cells, with frequencies ranging from 10^{-5} to 10^{-6} for each epitope and with a range of functional avidities. This indicates that, at least in the case of the response to influenza HA, immunodominance cannot be explained by holes in the repertoire, but rather by the selective generation of certain epitopes following antigen processing. The finding that the functional avidity of naive T cell clones was on average 10-fold lower as compared to that of memory T cell clones could be explained by clonal selection of high avidity T cells in the memory pool, as originally reported for mouse T cells (Busch and Pamer, 1999; McHeyzer-Williams et al., 1999; Savage et al., 1999). However, it is important to consider that many naïve T cell clones with intermediate avidities for peptides, comparable to that of memory clones, proliferated in response to peptide but not to processed HA, thus defining a series of cryptic epitopes. These T cell clones have been

defined a type-B T cells and may be involved in immune responses driven by non-conventional antigen processing, as suggested in some cases of autoimmunity (Mohan and Unanue, 2012; Sadegh-Nasseri and Kim, 2015). It should be noted however that HA-specific type-B T cell clones recognizing cryptic epitopes were not found in the memory pool, since all the memory clones isolated in response to HA peptides were found to recognize the endogenously processed antigen as well.

The diversity of epitope recognition measured in the naïve compartment indicate that a multiplicity of HA-derived peptides can potentially trigger T cell activation, underlining the redundancy of MHC-II system in accommodating and presenting a large variety of different peptides. Nevertheless, many T cell clones isolated from the naïve compartment showed high functional avidity when stimulated by peptides but not by naturally processed protein. This finding suggests that the naïve repertoire retains T cell precursors recognizing peptides that fail to be generated and/or presented by professional APCs (cryptic peptides) or that are produced in amount insufficient to trigger priming of cognate naïve T cells. It remains to be established whether non-professional APCs, such as epithelial cells in the respiratory tract that are main targets of Influenza viruses, may generate a different set of peptides compared to professional APCs. Epithelial cells readily upregulate MHC class II molecules in response to inflammatory cytokines or viral infection (Gao et al., 1999; Reith and Mach, 2001) and, although they have limited endocytic potential, they can generate peptide ligands for MHC-II presentation through endogenous degradation pathways, such as autophagy and macroautophagy (Dengjel et al., 2005; Schmid et al., 2007).

The analysis performed on memory T cell libraries and on *ex vivo* stimulated memory T cells revealed the presence, in each individual, of one or two immunodominant sites targeted by polyclonal and clonally expanded T cells as well as a few subdominant sites. Importantly, our study shows that the immunodominant sites correspond to those that are found most abundantly in the MHC-II peptidome isolated from HA-pulsed DCs or HA-specific B cells. These findings point to a simple model whereby immunodominance is determined by the abundance of a given peptide-MHC complex generated by processing followed by selection and clonal expansion of high avidity T cells. It is noteworthy that some subdominant memory T cell clones showed very high functional avidity when stimulated by peptides, but not by naturally processed

HA protein, albeit their cognate peptides showed similar binding affinity for MHC-II as compared to immunodominant peptides. This finding suggests that sub-dominance may be simply due to a lower abundance of the naturally processed peptide. Collectively our data reveal that only a small fraction of the peptides that bind to MHC-II molecules and can be recognized by T cells is generated by antigen processing and even in this case the yield of processed peptide can vary at least 100-fold between immunodominant and subdominant peptides.

Previous studies using tetanus toxoid as a model antigen showed that high-affinity antibodies can modulate antigen processing by enhancing or suppressing the generation of different T cell epitopes (Watts and Lanzavecchia, 1993). These findings are extended by the analysis of donor HD1, where a B cell clone with a BCR specific for the HA globular head generates the same sets of immunodominant peptides as DCs, while a B cell clone specific for the HA stem generates a lower yield and a different set of peptides. This difference, which is reflected in a different efficiency in presentation of the HA antigen, suggests a mechanism whereby BCR targeting the stem region of HA in its pre-fusion conformation can perturb antigen presentation of the immunodominant T cell epitope $HA_{401-420}$ localized in the loop between helix A and helix B of HA stem, that undergo a major conformational change at acidic pH (Bullough et al., 1994). As a consequence, in individuals with a particular HLA background such as HD1, B cell specific for the HA stem may be less capable to present antigen to immunodominant T cell clones and consequently to receive T-cell help.

The suppressive effect of antibodies might be mediated by several non-exclusive mechanisms: for instance, the antibody could impair the proteolytic liberation of certain T cell epitopes by directly masking cathepsin cleavage sites, or indirectly by limiting or delaying conformational changes and unfolding of HA at endosomal pH, therefore potentially altering the portions of the antigen exposed to proteases. Modulation of antigen processing by BCR specificity can have profound implications in the B cell clonal selection that happens in conditions of limited availability of antigen and T cell help. Indeed, in the highly competitive environment of the germinal center, little variations in abundance of MHC-II presented peptides might largely influence the probability of interaction with cognate Tfh cells, and ultimately affect the amount of T cell help received by individual B cell clones. In such sense, B cells bearing BCRs that

cause decreased presentation of immunodominant T cell epitopes (i.e. the most frequent specificity of cognate T helper cells), might be disadvantaged over other B cell clones whom BCR do not affect antigen processing.

The high-resolution epitope mapping of naïve and memory T cells from donors with a diverse HLA background offered us the possibility to benchmark currently available in silico predictors of CD4[+] T cell immunogenicity. Indeed, we found that, although being a pre-requisite for recognition by T cells, peptide-MHC-II binding affinity, either predicted in silico or measured *in vitro*, is only a weak correlate of T cell recognition and in particular of immunodominance. During antigen processing, peptides are liberated by enzymatic activity of cathepsins and other proteases activated by the acidification of pH in late endosomes and lysosomes, and assembled within the peptide-binding cleft of MHC-II molecules prior to exposure on the cell surface of APCs (Roche and Furuta, 2015; Unanue et al., 2016). It has been proposed that, along with the protein antigen expression level and subcellular localization, also the position within the tridimensional structure of the native antigen may profoundly influence the amount of processed peptides (Abelin et al., 2019; Graham et al., 2018). Recent reports have shown that T cell epitopes from viral antigens tend to localize in highly flexible, surface-exposed regions of the protein that could act as sites of initial proteolytic cleavage (Koblischke et al., 2017; Landry, 2008; Mirano-Bascos et al., 2008), suggesting that the physical accessibility within the tertiary structure is a requirement for efficient peptide release by proteases. By analyzing the tridimensional conformation of HA, we found that immunodominant epitopes are embedded in regions predicted as readily accessible targets of endosomal proteases. Furthermore, combined in silico analysis of both peptide-MHC-II binding affinity and structural accessibility yielded higher predictive values for the set of HA memory epitopes here described, thus indicating a possible strategy to develop more accurate predictive algorithms for T cell immunogenicity of protein antigens. The combined model of accessibility and peptide-MHC-II binding would be consistent with the MHC-guided antigen processing proposed by the "bind first, cut later" model (Sercarz and Maverakis, 2003): readily unfolded / accessible regions of the antigen might be bound by MHC-II cleft more promptly, and therefore protected from destruction by endosomal proteases.

In conclusion, the findings here reported suggest a model for epitope selection by antigen processing based on a trade-off between multiple factors, including the structural accessibility to proteases and the binding affinity of liberated peptides for the groove of MHC-II molecules. Structural constrains might define regions prone to be liberated at higher rate by protease cut, thus being a property intrinsic of the antigen tridimensional structure, whereas the MHC-II allelic background of each individual would select the fine sequences of high-affinity binder peptides. The implications of this model is that highly accessible epitopes could be presented at high abundance on the APC surface even if they are not strong MHC-II binders, and vice versa potentially strong binder peptides could never be presented at relevant amount if they are not accessible to proteolytic liberation or if they are destroyed by cathepsins. The net result of such complex processes would be the differential abundance of some MHC-II-presented peptides, which might drive more prominent clonal expansion of cognate naïve T cells leading to immunodominance. Perturbation of the antigen structure, for instance by bound immunoglobulins, might further alter the substrate for endosomal proteolysis and therefore influence the antigen presentation and interaction with cognate T cells.

6.2.6 Figures and tables
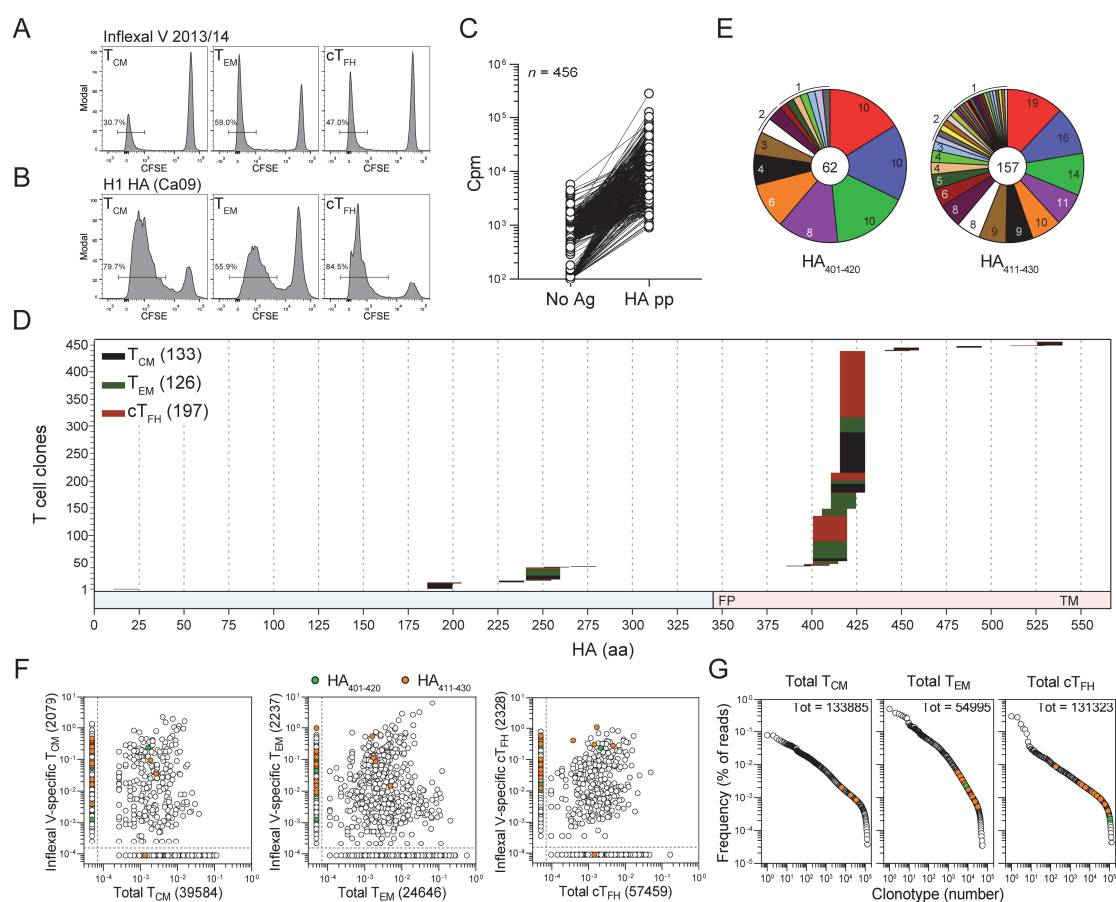
Figure 6.2-1. Clonally expanded memory CD4⁺ T cells target an immunodominant region of influenza H1-HA. (**A**) Memory CD4⁺ Tcm, Tem and cTfh cell subsets were isolated from blood samples of donor HD1 six and twelve months after Inflexal V vaccination. Cells were labeled with CFSE and stimulated with the Inflexal V vaccine in the presence of autologous monocytes. Shown is the CFSE profile and the percentage of proliferating CFSE^{low} cells on day 6 in the twelve month-sample. Percentage of CFSE^{low} in the six month-samples was 38% (Tcm), 65% Tem and 57% cTfh. (**B**) Inflexal V-reactive CFSE^{low} T cells were sorted, relabeled with CFSE and stimulated with recombinant H1-HA in the presence of autologous monocytes. After 5 days, CFSE^{low} proliferating T cells from month 6 and month 12 samples were sorted and cloned by limiting dilution. (**C**) H1-HA-specific CD4⁺ T cell clones were identified based on the proliferative response (stimulation index, SI, >3) to a pool of overlapping peptides spanning the entire H1-HA sequence. Proliferation was assessed on day 3 after a 16-h pulse with [³H]-thymidine and expressed as the counts per minute (Cpm). (**D**) Epitope mapping of H1-HA-specific T cell clones isolated from memory CD4⁺ T cell subsets of donor HD1. The epitopes were identified by screening the T cell clones with the individual H1-HA peptides. The *x* axis indicates HA amino acid sequence; each color-coded segment represents the sequence recognized by individual clones isolated from Tcm (black), Tem (green) or cTfh (red) cultures. The numbers of H1-HA-reactive T cell clones isolated from each subset are reported. (**E**) Rearranged TCR Vβ sequences of H1-HA-specific T cell clones were determined by RT-PCR followed by Sanger sequencing. Pie chart of the repertoire of rearranged TCR Vβ sequences of T cell clones recognizing immunodominant HA$_{401-420}$ or HA$_{411-430}$ epitopes. Each slice of the chart indicates a different TCR Vβ clonotype (HA$_{401-420}$, *n*=16 and HA$_{411-430}$, *n*=39); the number of sister clones bearing the same TCR Vβ sequence are reported for each slice. The total number of clones sequenced is reported at the center. (**F**, **G**) To evaluate the clonal expansion of H1-

HA-reactive memory T cell clones, TCR Vβ deep sequencing was performed on Inflexal V-reactive CFSE^{low} Tcm, Tem and cTfh cells and on *ex vivo* total Tcm, Tem and Tfh cells isolated from the same blood samples and immediately sequenced. Comparison of TCR Vβ clonotype frequency distribution of total (*x* axis) and Inflexal V-reactive cells (*y* axis) is shown in panel F. Circles outside the dotted lines represent clonotypes that were found in only one of the two samples and that were assigned an arbitrary frequency value for graphical purposes. Panel G reports the frequency distribution of TCR Vβ clonotypes from *ex vivo* total Tcm, Tem and Tfh cells isolated from a blood sample of donor HD1 obtained 48 months after the 2013/14 Inflexal vaccination. TCR Vβ clonotypes recognizing the immunodominant HA$_{401-420}$ or HA$_{411-430}$ peptides found in any of the samples analyzed are colored in green and orange, respectively.
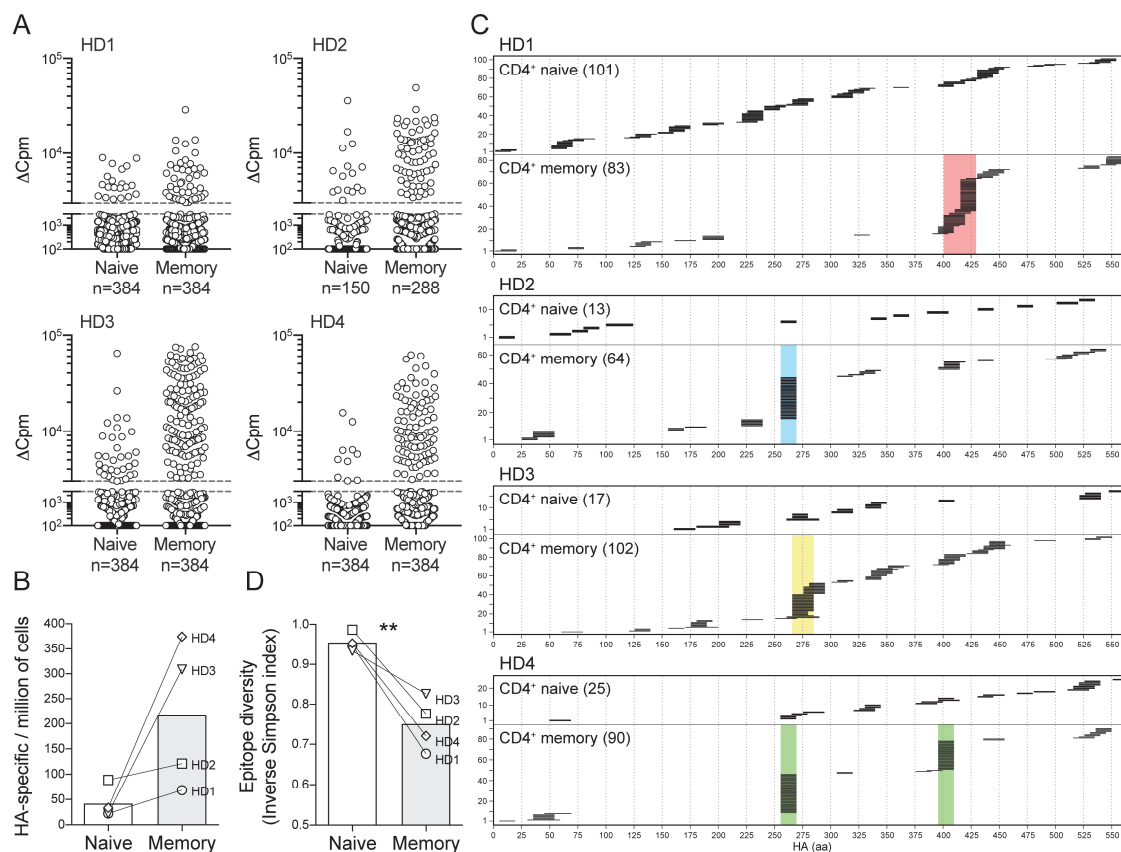
Figure 6.2-2. Isolation and epitope mapping of HA-reactive CD4+ T cells from naïve and memory compartments of influenza vaccinated donors. (**A**) Naïve and memory CD4+ T cells were FACS-sorted from PBMCs of four donors and polyclonally expanded in multiple wells, each containing 1000 to 2000 cells. The number of wells ranged from 150 to 384, depending on the number of cells isolated (see *Methods*). After 14-21 days, the amplified naïve and memory T cell libraries were screened against a pool of overlapping peptides spanning the entire H1-HA sequence in the presence of autologous APCs. Proliferation was assessed on day 4 after a 16-h pulse with [3H]-thymidine. Data are expressed as the counts per minute, after subtraction of background proliferation (ΔCpm). The proliferative response of each T cell clone in the library is represented by a single dot. The specificity of positive cultures was confirmed in 3 independent experiments. (**B**) Frequencies of H1-HA-specific T cells within naïve or memory CD4+ T cells was calculated based on number of negative wells according to the Poisson distribution. Data are expressed as frequency per million of naïve or memory CD4+ T cells. Each symbol indicates a different donor. (**C**) Epitope mapping of H1-HA-specific T cell clones from naïve or memory libraries. The epitopes were identified by screening the T cell cultures with overlapping peptides spanning the entire H1-HA sequence. The *x* axis indicates HA amino acid sequence; each segment represents the sequence recognized by individual T cell lines. The numbers of HA-reactive T cell clones in the naïve or memory library of each donor is reported. The immunodominant regions identified in the memory compartment of each donor are highlighted with color-coded shadows. (**D**) Richness and evenness of the pool of HA-specific T cell clones isolated from naïve or memory libraries were evaluated as Inverse Simpson index of diversity (1-D), calculated based on the number of T cell clones recognizing each particular H1-HA peptide. Inverse Simpson index (1-D) ranges between 0 and 1, and reflects the probability that two HA-specific T cell clones randomly selected from a repertoire will recognize different HA epitopes. ** *p*-value = 0.0097, as determined by two-tailed paired t-test.
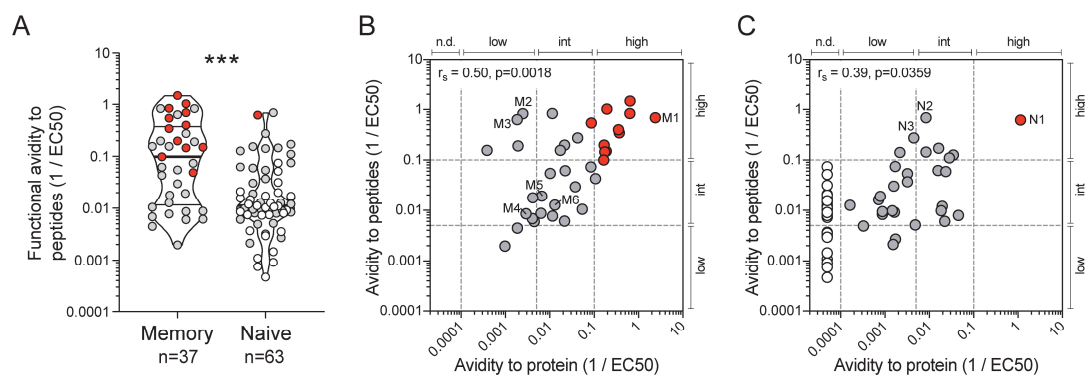
Figure 6.2-3. H1-HA specific naïve and memory T cell clones show different functional avidities. Functional avidity of H1-HA-reactive T cell clones isolated from donor HD1 was determined by stimulation with titrated doses of synthetic H1-HA peptides or recombinant H1-HA in the presence of autologous monocytes. Proliferation was assessed on day 3 after a 16-h pulse with [^{3}H]-thymidine; EC50 values were calculated by nonlinear regression curve fit. (**A**) Violin plots of the frequency distribution of reciprocal EC50 values of T cell clones from the memory or naïve compartment stimulated with titrated doses of H1-HA peptides. T cell clones specific for the immunodominant HA$_{411-430}$ epitope are reported as red dots. Lines represent the median and quartiles. *** *p*-value < 0.001, as determined by two-tailed Mann–Whitney *U*-test. (**B** and **C**) Scatter plots of reciprocal EC50 values of T cell clones from the memory (**B**) or naïve (**C**) compartment, stimulated in parallel with recombinant H1-HA (*x* axis) and synthetic peptides (*y* axis). EC50 values below the detection limit for stimulations with recombinant H1-HA were set arbitrarily to 20 µg/ml; the corresponding T cell clones are reported as white dots. Spearman correlation was calculated based on EC50 pairs from T cell clones responding to both peptides and recombinant H1 HA (**B**, *n*=36 and **C**, *n*=29). Thresholds of functional avidity were set arbitrarly at EC50 values of 10 µg/ml, 200 ng/ml and 10 ng/ml of antigen.
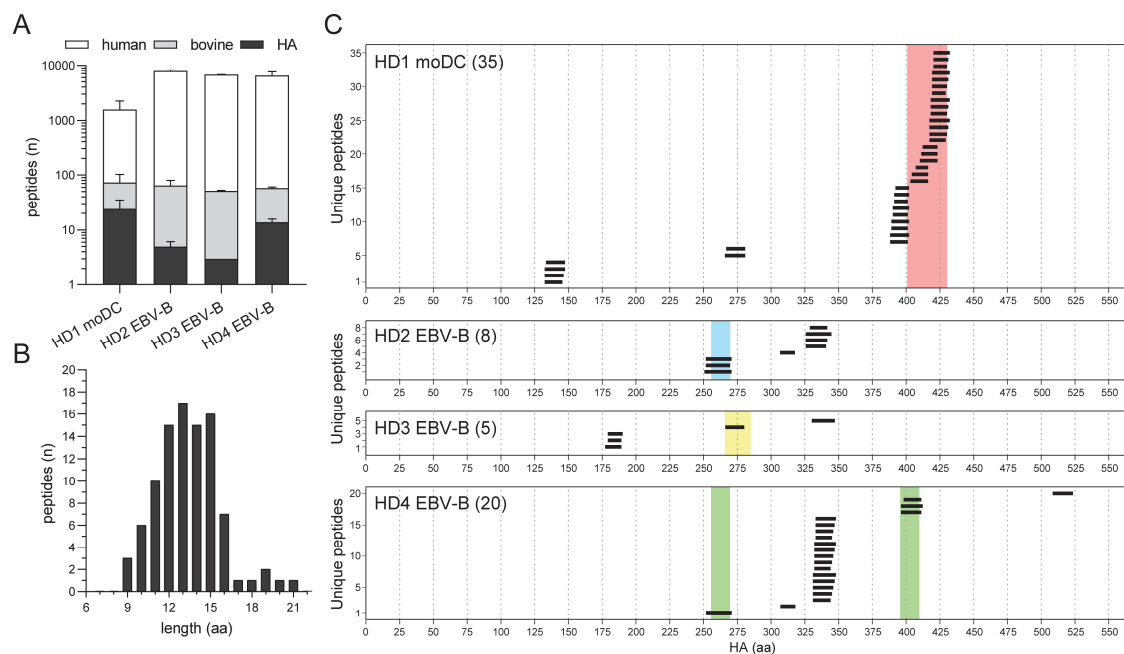
Figure 6.2-4. The MHC class II peptidome defines H1-HA immunodominant regions targeted by memory CD4+ T cells. Monocyte-derived DCs (moDC) from donor HD1 and HA-specific EBV-B cell clones from donors HD2-HD4 were pulsed with recombinant H1-HA and MHC-II presented peptides were measured by MS-based immunopeptidomics. (**A**) Number of MHC-II eluted peptides measured by MS in the different donors (mean ± SD of *n*=2 independent experiments for HD1 and HD3, *n*=3 for HD2, *n*=4 for HD4). Color code indicates the different source organism of the measured peptides. (**B**) Histogram of the lengths of H1-HA-derived peptides eluted from MHC-II molecules reported in (A). (**C**) Sets of H1-HA-derived peptides eluted from MHC-II molecules from each donor. The *x* axis indicates HA amino acid sequence. Each segment represents a unique HA-derived peptide identified by MS (union of *n*=2 independent experiments for HD1 and HD3, *n*=3 for HD2, *n*=4 for HD4); the total numbers of HA-derived peptides are reported. The immunodominant regions targeted by memory CD4+ T cells of each donor are reported with color-coded shadows.
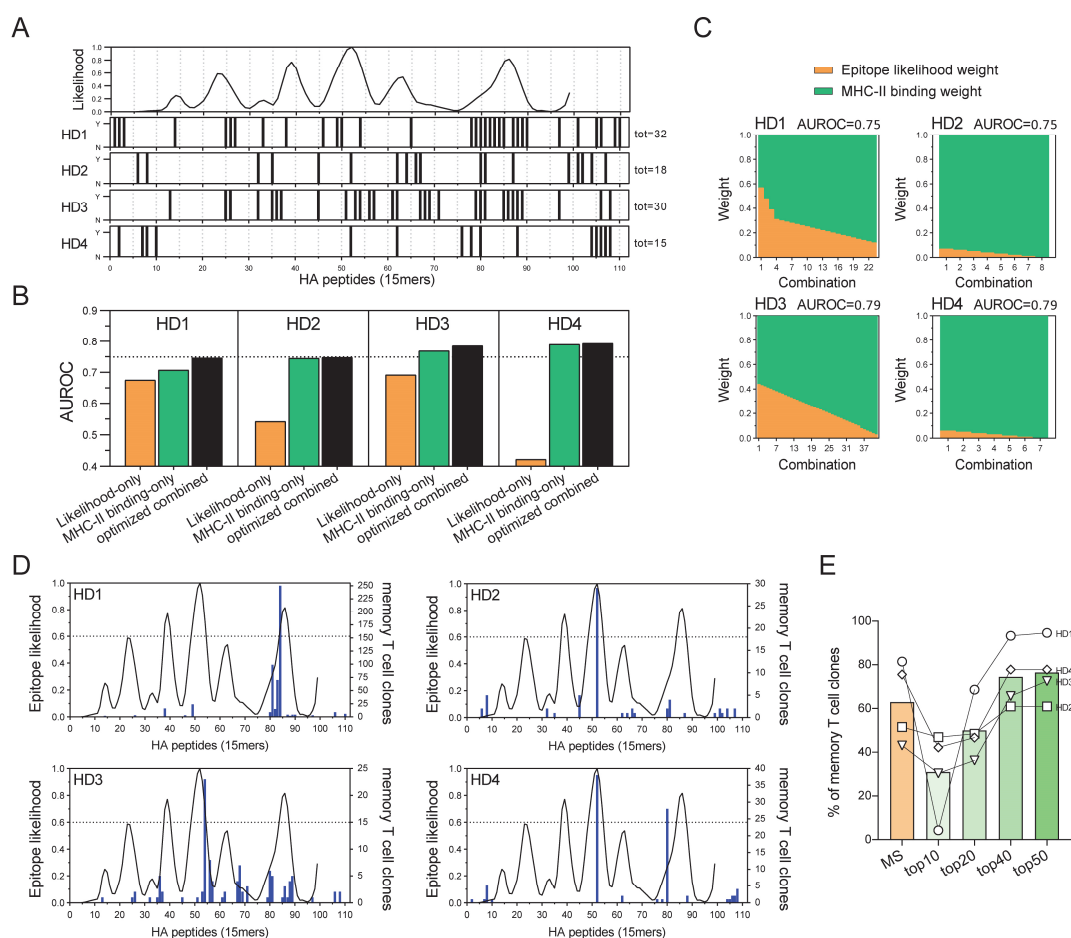
Figure 6.2-5. Immunodominant epitopes localize in HA regions predicted as promptly liberated by endosomal proteases. (**A**) Epitope likelihood based on structural accessibility was calculated for each theoretical HA 15mer peptide (*upper panel*). The sets of HA epitopes mapped in the memory CD4[+] T cell repertoire of each donor is reported (*lower panels*); each segment indicate a peptide recognized by at least one T cell clone isolated from memory CD4[+] T cells. (**B**) Performance of in silico predictors was benchmarked to the sets of HA memory epitopes reported in (**A**) used as references, and measured by Area under the Receiver Operating Characteristic Curve (AUROC). Shown are the maximal AUROC values achieved by single predictors based uniquely on structural accessibility (orange bars) or MHC-II binding calculated using NetMHCIIpan 3.2 (green bars), or by the combined predictors (black bars). (**C**) Relative contributions of epitope likelihood based on structural accessibility (orange bars) and MHC-II predicted binding (green bars) in the combined predictor for each donor. Shown are all the different weight combinations that yield the best performances (AUROC values within 0.05 of the maximum achieved). (**D**) Epitope likelihood based on structural accessibility was compared to the epitope mapping of HA-specific cells isolated from the memory CD4[+] T cell compartment of each donor. The black curve indicates predicted epitope likelihood; blue bars of the histogram indicate the number of individual T cell clones found specific for each HA 15mer peptide. (**E**) Sensitivity was evaluated in terms of percentage of memory T cell clones for which the cognate peptide was identified by MS-based MHC-II peptidomics (orange bar) or by epitope prediction based on structural accessibility at different thresholds (green bars). Each symbol represents a different donor.
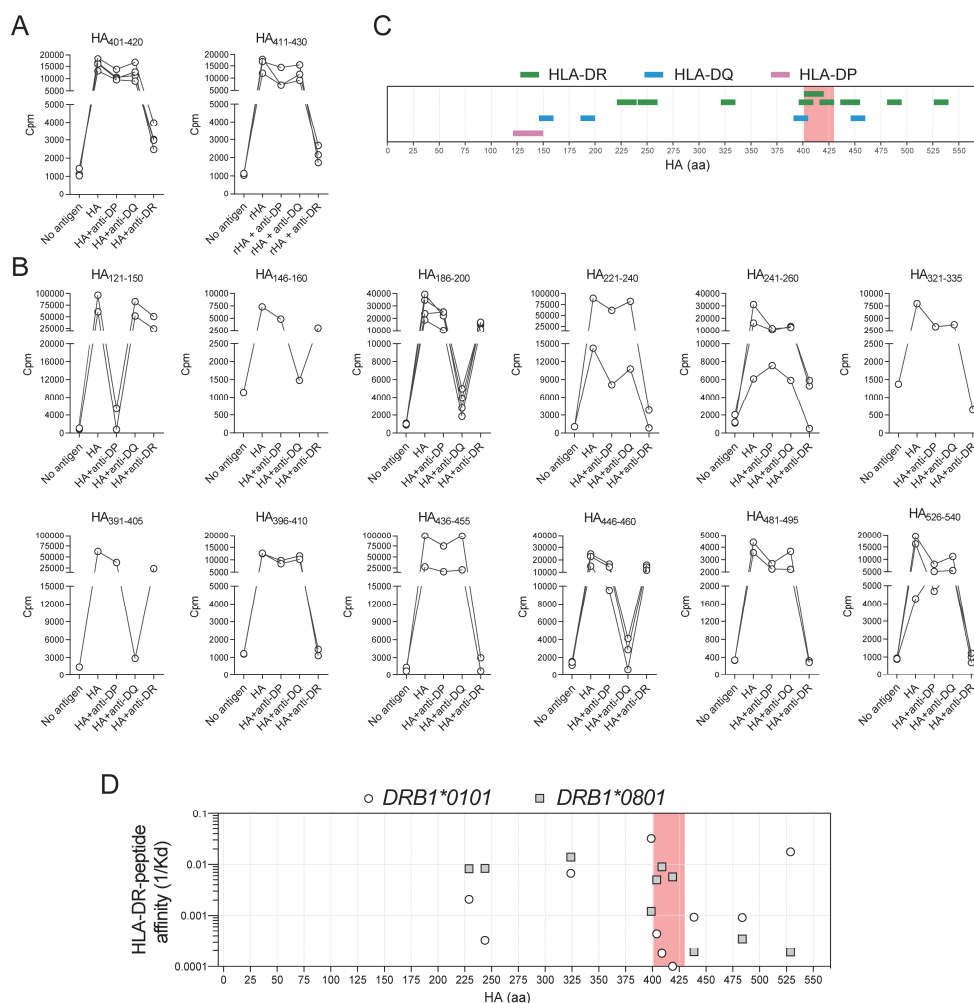
Supplementary Figure 6.2-1. MHC restriction of H1-HA-specific T cell clones and measurement of peptide binding affinity to HLA-DR. (**A** and **B**) MHC-II restriction of T cell clones from donor HD1 specific for different H1-HA peptides was determined by stimulation with autologous APCs pulsed with recombinant H1-HA, in the absence or presence of anti-MHC-II blocking antibodies (anti-HLA-DR, clone L243; anti-HLA-DQ, clone SPVL3; anti-HLA-DP, clone B7/21). Proliferation was assessed on day 3 after a 16-h pulse with [^{3}H]-thymidine and expressed as counts per minute (Cpm). MHC-II restriction was defined based on inhibition of T cell proliferation >80%. A panel of T cell clones specific for immunodominant (**A**) or subdominant (**B**) HA epitopes are shown. Data are grouped based on the epitope specificity of the T cell clones tested, reported on top of each plot. (**C**) Summary of MHC-II restriction of H1-HA-specific T cell clones isolated from donor HD1. The *x* axis indicates HA amino acid sequence; each color-coded segment represents the peptide recognized by T cell clones restricted by HLA-DP (pink), HLA-DQ (blue) or HLA-DR (green). The immunodominant HA$_{401-430}$ region identified in the memory compartment of donor HD1 is highlighted with a red shadow. (**D**) MHC-II binding affinity of H1-HA peptides recognized by HLA-DR-restricted T cell clones from donor HD1 was measured in vitro. Briefly, recombinant HLA-DRB1 isoforms were refolded in the presence of recombinant HLA-DRA and increasing concentration of peptides, at room temperature and pH 7. Kd values were calculated by nonlinear regression fitting of pMHC-II refolding curves. The plot reports the inverse Kd values of each HA peptide tested with either *HLA-DRB1*01:01* (white dots) or *HLA-DRB1*08:01* (grey squares) alleles. The immunodominant HA$_{401-430}$ region identified in the memory compartment is highlighted with a red shadow.
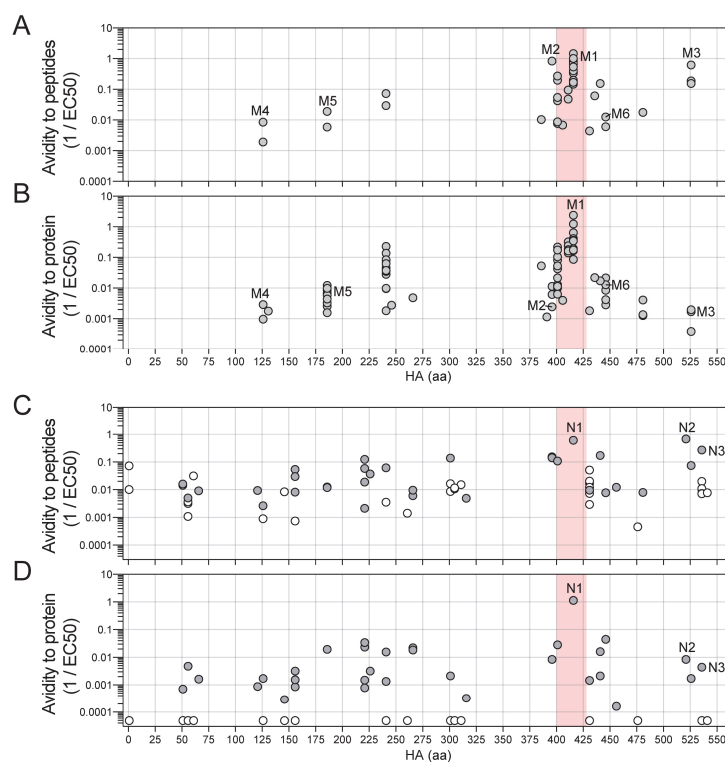
Supplementary Figure 6.2-2. Functional avidities for peptide and naturally processed H1-HA of T cell clones specific for different epitopes. Functional avidity of H1-HA-reactive T cell clones isolated from the memory (**A** and **B**) or the naïve (**C** and **D**) compartment of donor HD1 was determined by stimulation with titrated doses of synthetic peptides (**A**, *n*=37 clones and **C**, *n*=63) or recombinant H1-HA (**B**, *n*=79 and **D**, *n*=63). Data are expressed as reciprocal EC50 values. Each dot represents an individual T cell clone; the position of the dots on the *x* axis indicates the starting residue of the cognate peptide. EC50 values below the detection limit for stimulations with recombinant H1-HA were set arbitrarily to 20 µg/ml; the corresponding T cell clones are reported as white dots. The immunodominant H1-HA$_{401-430}$ region identified in the memory compartment is highlighted with a red shadow.
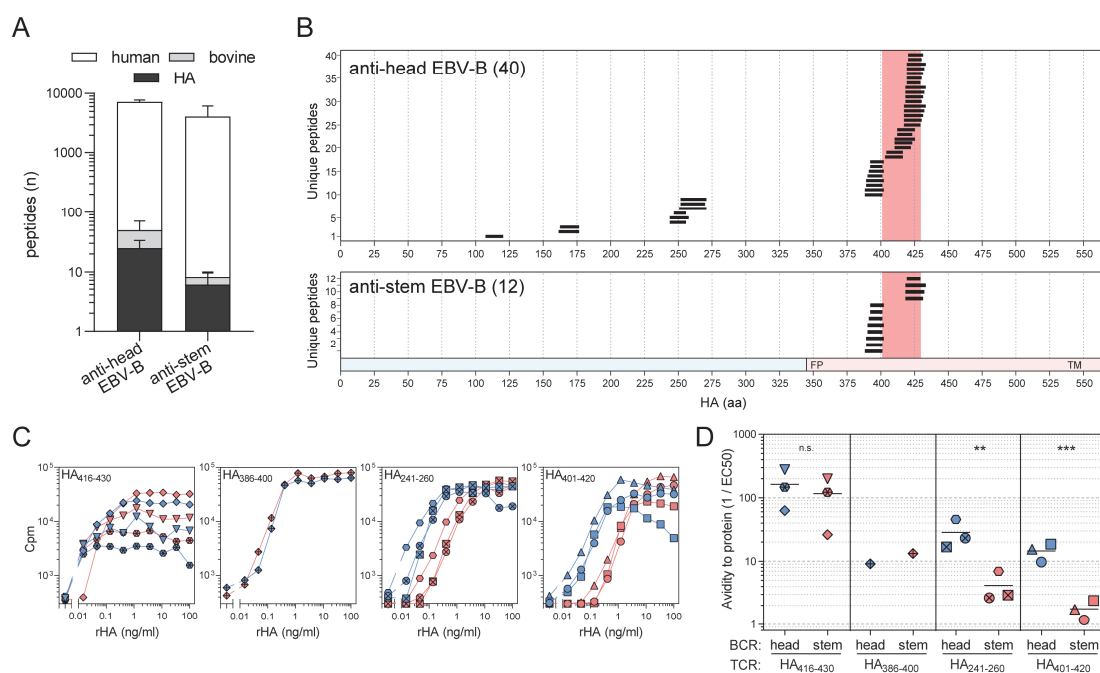
Supplementary Figure 6.2-3. The MHC class II peptidome reveals modulation for antigen processing by antibodies. EBV-B cell clones from donor HD1 specific for HA head or HA stem were pulsed with recombinant H1-HA, and MHC-II presented peptides were measured by MS-based immunopeptidomics. (**A**) Number of MHC-II eluted peptides measured by MS in anti-head or anti-stem EBV-B cell clones (mean ± SD of *n*=3 independent experiments). Color code indicates the different source organism of the measured peptides. (**B**) Sets of H1-HA-derived peptides eluted from MHC-II molecules of anti-head or anti-stem EBV-B cell clones. The *x* axis indicates HA amino acid sequence. Each segment represents a unique H1-HA-derived peptide identified by MS (union of *n*=3 independent experiments); the total numbers of HA-derived peptides are reported. The immunodominant H1-HA$_{401-430}$ region targeted by memory CD4^{+} T cells is highlighted with a red shadow. The chart at the bottom indicates HA1 and HA2 domains colored in blue and red, respectively (FP, fusion peptide; TM, transmembrane). (**C** and **D**) The effect of BCR specificity in modulating antigen processing was tested by co-culture of HA-specific T and EBV-B cell clones in the presence of titrated antigen. Briefly, functional avidity of memory T cell clones specific for different HA epitopes was determined by stimulation with titrated doses of recombinant H1-HA in the presence of the indicated EBV-B cell clone as APCs. Each individual T cell clone tested was restricted by HLA-DR and carried a different rearranged TCR Vβ sequence. (**C**) Proliferation was assessed on day 3 after a 16-h pulse with [^{3}H]-thymidine and expressed as counts per minute (Cpm). Proliferation curves are grouped based on the epitope specificity of the T cell clones, reported in the top-left corner of each plot. Each symbol refers to a distinct T cell clone stimulated in the presence of anti-head (colored in blue) or anti-stem (in red) EBV-B cell clone as APCs. (**D**) EC50 values were calculated by nonlinear regression curve fit. Each symbol indicates the reciprocal EC50 value of T cell clones co-cultured with anti-head (in blue) or anti-stem (in red) EBV-B cell clone as APCs. Lines represent mean values. Data are grouped based on the epitope specificity of the T cell clones. n.s. not significant, ** *p*-value = 0.0041, *** *p*-value = 0.0004, as determined by two-tailed ratio paired t-test.
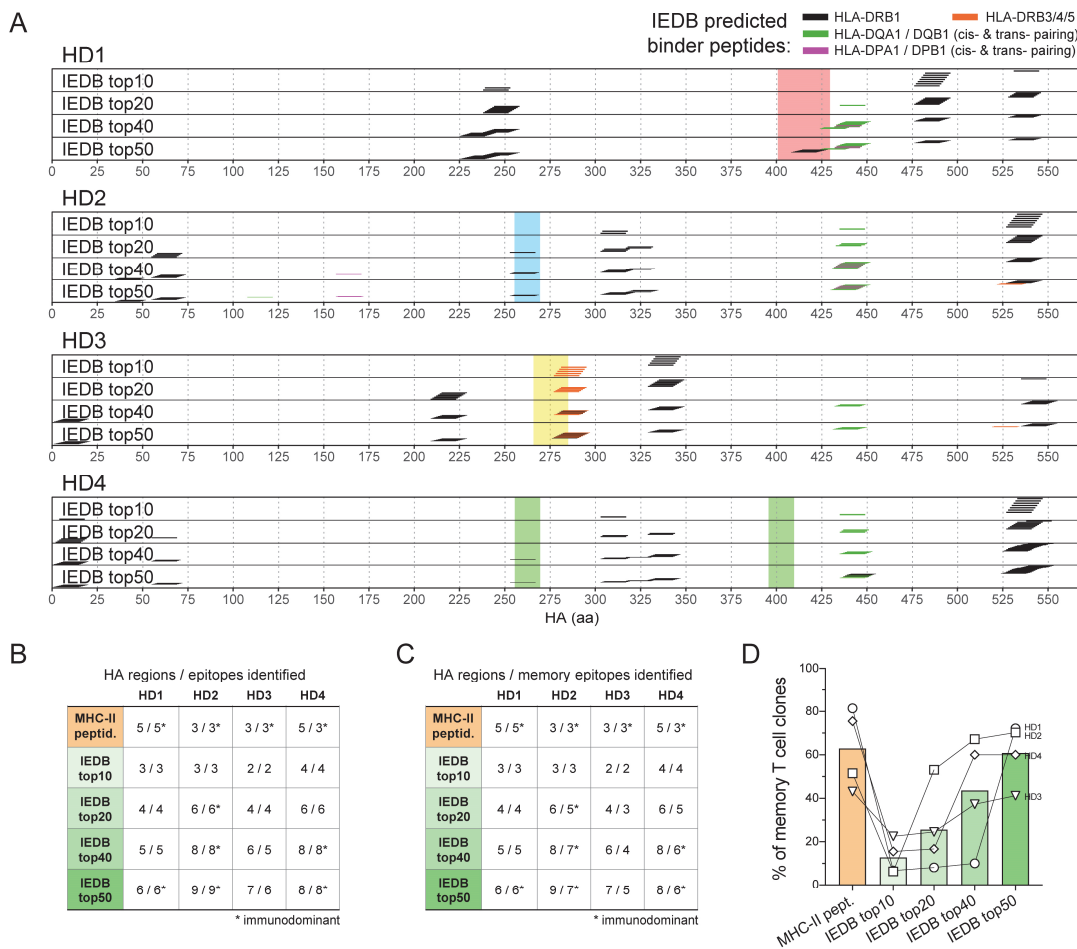
Supplementary Figure 6.2-4. Peptide binding to MHC-II is a necessary but not sufficient condition to define immunodominance. MHC-II binding affinity of each theoretical H1-HA 15mer peptide was calculated using IEDB tool for MHC-II binding prediction (http://tools.iedb.org/mhcii/). Personalized analyses were performed by considering the MHC-II alleles carried by each donor (HLA-DRB1, HLA-DRB3/4/5, HLA-DQA1/DQB1 in cis- or trans-pairing, HLA-DPA1/DPB1 in cis- or trans-pairing). Top scoring H1-HA 15mer peptides for each donor were selected based on percentile rank calculated by comparison to a large set of random natural peptides. (**A**) Sets of H1-HA peptides predicted as MHC-II binders at different thresholds for each donor. The *x* axis indicates HA amino acid sequence. The sets of top predicted MHC-II binder peptides are reported as color-coded segments. The immunodominant regions targeted by memory CD4[+] T cells of each donor are reported with color-coded shadows. (**B** and **C**) IEDB-predicted peptides and MHC-II eluted peptides measured by MS-based peptidomics defined discrete H1-HA regions. The tables summarize the number of HA regions found presented on MHC-II by MS-based peptidomics, or predicted as MHC-II binders in different donors. The corresponding number of HA epitopes recognized by at least one T cell clone regardless of the subset of origin (**B**) or isolated from the memory compartment (**C**) are reported. Identification of immunodominant epitopes is marked by an asterisk. (**D**) Sensitivity was evaluated in terms of percentage of memory T cell clones for which the cognate peptide was identified by MS-based MHC-II peptidomics (orange bar) or by MHC-II binding predictions at different thresholds (green bars). Each symbol represents a different donor.

Supplementary Table 6.2-1. Epitope specificity of H1-HA-reactive T cell clones isolated from memory CD4+ T cell subsets. Epitope mapping of 456 H1-HA-reactive T cell clones from donor HD1 was performed by screening with overlapping peptides spanning the entire H1-HA sequence in the presence of autologous APCs. Proliferation was assessed on day 3 after a 16-h pulse with [³H]-thymidine and expressed as counts per minute (cpm). For each T cell clone the subset of origin, the start and end position of the epitope residues in H1-HA, and the cpm values after stimulation with autologous APCs untreated or pulsed with the indicated HA peptide are reported.

Donor	Subset	Clone ID	Peptide	Start	Stop	Length	No Ag (cpm)	Peptide (cpm)
HD1-t1	Tem	D2	TFATANADTLCIGYH	11	25	15	96	2780
HD1-t1	Tcm	C9	KGKEVLVLWGIHHPS	186	200	15	35	16800
HD1-t1	Tcm	F9	KGKEVLVLWGIHHPS	186	200	15	122	37220
HD1-t1	Tcm	G3	KGKEVLVLWGIHHPS	186	200	15	56	22389
HD1-t1	Tcm	G8	KGKEVLVLWGIHHPS	186	200	15	80	1087
HD1-t1	Tcm	E5	KGKEVLVLWGIHHPS	186	200	15	57	34909
HD1-t1	Tcm	M8	KGKEVLVLWGIHHPS	186	200	15	70	38751
HD1-t1	Tcm	N5	KGKEVLVLWGIHHPS	186	200	15	9	7557
HD1-t1	Tcm	G5	KGKEVLVLWGIHHPS	186	200	15	74	3209
HD1-t1	Tcm	1_C10	KGKEVLVLWGIHHPS	186	200	15	112	31556
HD1-t1	Tem	C10	KGKEVLVLWGIHHPS	186	200	15	23	1921
HD1-t1	cTfh	1_E2	KGKEVLVLWGIHHPSTSADQ	186	205	20	152	5961
HD1-t1	Tcm	H11	KGKEVLVLWGIHHPSTSADQ	186	205	20	9	7607
HD1-t1	Tcm	B10	KFKPEIAIRPKVRDQ	226	240	15	37	52927
HD1-t1	Tcm	E10	KFKPEIAIRPKVRDQ	226	240	15	28	13504
HD1-t1	Tcm	H10	KFKPEIAIRPKVRDQ	226	240	15	17	14525
HD1-t1	cTfh	B10	EGRMNYYWTLVEPGD	241	255	15	30	1384
HD1-t1	Tcm	E11	EGRMNYYWTLVEPGD	241	255	15	2	2194
HD1-t1	cTfh	B5	EGRMNYYWTLVEPGDKITFE	241	260	20	133	2276
HD1-t1	cTfh	C3	EGRMNYYWTLVEPGDKITFE	241	260	20	26	2806
HD1-t1	cTfh	D5	EGRMNYYWTLVEPGDKITFE	241	260	20	26	1411
HD1-t1	Tcm	C2	EGRMNYYWTLVEPGDKITFE	241	260	20	17	1545
HD1-t1	Tcm	C4	EGRMNYYWTLVEPGDKITFE	241	260	20	87	4061
HD1-t1	Tcm	C8	EGRMNYYWTLVEPGDKITFE	241	260	20	6	2473
HD1-t1	Tcm	H2	EGRMNYYWTLVEPGDKITFE	241	260	20	30	4149
HD1-t1	Tcm	H7	EGRMNYYWTLVEPGDKITFE	241	260	20	56	8047
HD1-t1	Tcm	G9	EGRMNYYWTLVEPGDKITFE	241	260	20	15	6585
HD1-t2	Tcm	P2_H1	EGRMNYYWTLVEPGDKITFE	241	260	20	1057	13480
HD1-t2	Tem	P1_B3	EGRMNYYWTLVEPGDKITFE	241	260	20	779	6019
HD1-t2	Tem	P1_B6	EGRMNYYWTLVEPGDKITFE	241	260	20	999	3372
HD1-t2	Tem	P1_D11	EGRMNYYWTLVEPGDKITFE	241	260	20	831	15987
HD1-t2	Tem	P1_F12	EGRMNYYWTLVEPGDKITFE	241	260	20	1079	25246
HD1-t2	Tem	P1_H2	EGRMNYYWTLVEPGDKITFE	241	260	20	860	4606
HD1-t2	Tem	P1_H3	EGRMNYYWTLVEPGDKITFE	241	260	20	711	2737
HD1-t2	Tem	P2_A5	EGRMNYYWTLVEPGDKITFE	241	260	20	960	9814
HD1-t2	Tem	P2_B2	EGRMNYYWTLVEPGDKITFE	241	260	20	1471	24848
HD1-t2	Tem	P2_B10	EGRMNYYWTLVEPGDKITFE	241	260	20	1898	15899
HD1-t2	Tem	P2_D9	EGRMNYYWTLVEPGDKITFE	241	260	20	610	5304
HD1-t2	Tem	P2_F3	EGRMNYYWTLVEPGDKITFE	241	260	20	746	7698
HD1-t2	Tem	P2_H1	EGRMNYYWTLVEPGDKITFE	241	260	20	934	3470
HD1-t1	Tcm	N2	VEPGDKITFEATGNL	251	265	15	40	6888
HD1-t1	Tcm	I11	VVPRYAFAMERNAGS	266	280	15	20	43000
HD1-t1	Tcm	B11	QNAIDEITNKVNSVI	386	400	15	11	3949
HD1-t1	cTfh	D10	VNSVIEKMNTQFTAV	396	410	15	2	24126
HD1-t1	cTfh	D4	VNSVIEKMNTQFTAV	396	410	15	27	8367
HD1-t1	Tcm	C7	VNSVIEKMNTQFTAV	396	410	15	111	31432
HD1-t1	cTfh	1_E6	EKMNTQFTAVGKEFN	401	415	15	50	1516
HD1-t2	Tem	P1_E9	EKMNTQFTAVGKEFN	401	415	15	705	3439
HD1-t2	Tem	P2_C4	EKMNTQFTAVGKEFN	401	415	15	607	4359
HD1-t2	Tem	P2_C5	EKMNTQFTAVGKEFN	401	415	15	1693	13277
HD1-t2	Tem	P2_E4	EKMNTQFTAVGKEFN	401	415	15	546	3244
HD1-t2	Tem	P2_H6	EKMNTQFTAVGKEFN	401	415	15	513	2298
HD1-t2	Tem	P2_H9	EKMNTQFTAVGKEFN	401	415	15	557	2334
HD1-t1	cTfh	B11	EKMNTQFTAVGKEFNHLEKR	401	420	20	41	10207
HD1-t1	cTfh	B2	EKMNTQFTAVGKEFNHLEKR	401	420	20	28	8679
HD1-t1	cTfh	B4	EKMNTQFTAVGKEFNHLEKR	401	420	20	11	9663
HD1-t1	cTfh	C4	EKMNTQFTAVGKEFNHLEKR	401	420	20	14	10216
HD1-t1	cTfh	C7	EKMNTQFTAVGKEFNHLEKR	401	420	20	15	6012
HD1-t1	cTfh	C8	EKMNTQFTAVGKEFNHLEKR	401	420	20	37	4298
HD1-t1	cTfh	D11	EKMNTQFTAVGKEFNHLEKR	401	420	20	35	6440
HD1-t2	cTfh	P2_B8	EKMNTQFTAVGKEFNHLEKR	401	420	20	744	3175

HD1-t2	cTfh	P2_B9	EKMNTQFTAVGKEFNHLEKR	401	420	20	670	5584
HD1-t2	cTfh	P2_C1	EKMNTQFTAVGKEFNHLEKR	401	420	20	1049	10934
HD1-t2	cTfh	P2_C5	EKMNTQFTAVGKEFNHLEKR	401	420	20	636	6484
HD1-t2	cTfh	P2_C7	EKMNTQFTAVGKEFNHLEKR	401	420	20	636	5305
HD1-t2	cTfh	P2_E3	EKMNTQFTAVGKEFNHLEKR	401	420	20	653	2090
HD1-t2	cTfh	P2_E12	EKMNTQFTAVGKEFNHLEKR	401	420	20	714	29826
HD1-t2	cTfh	P2_F7	EKMNTQFTAVGKEFNHLEKR	401	420	20	666	18331
HD1-t2	cTfh	P2_G1	EKMNTQFTAVGKEFNHLEKR	401	420	20	558	82246
HD1-t2	cTfh	P3_A8	EKMNTQFTAVGKEFNHLEKR	401	420	20	536	5856
HD1-t2	cTfh	P3_B8	EKMNTQFTAVGKEFNHLEKR	401	420	20	548	15338
HD1-t2	cTfh	P3_B11	EKMNTQFTAVGKEFNHLEKR	401	420	20	710	9439
HD1-t2	cTfh	P3_C6	EKMNTQFTAVGKEFNHLEKR	401	420	20	615	1859
HD1-t2	cTfh	P3_C10	EKMNTQFTAVGKEFNHLEKR	401	420	20	518	35719
HD1-t2	cTfh	P3_D11	EKMNTQFTAVGKEFNHLEKR	401	420	20	555	26321
HD1-t2	cTfh	P3_E11	EKMNTQFTAVGKEFNHLEKR	401	420	20	522	41415
HD1-t2	cTfh	P3_F5	EKMNTQFTAVGKEFNHLEKR	401	420	20	805	2621
HD1-t2	cTfh	P3_G2	EKMNTQFTAVGKEFNHLEKR	401	420	20	580	2138
HD1-t2	cTfh	P3_G9	EKMNTQFTAVGKEFNHLEKR	401	420	20	843	16556
HD1-t2	cTfh	P3_H9	EKMNTQFTAVGKEFNHLEKR	401	420	20	783	9039
HD1-t2	cTfh	P1_B6	EKMNTQFTAVGKEFNHLEKR	401	420	20	757	21726
HD1-t2	cTfh	P1_D1	EKMNTQFTAVGKEFNHLEKR	401	420	20	762	31463
HD1-t2	cTfh	P1_D6	EKMNTQFTAVGKEFNHLEKR	401	420	20	813	8569
HD1-t2	cTfh	P1_E12	EKMNTQFTAVGKEFNHLEKR	401	420	20	636	40637
HD1-t2	cTfh	P1_F6	EKMNTQFTAVGKEFNHLEKR	401	420	20	695	5182
HD1-t2	cTfh	P1_F8	EKMNTQFTAVGKEFNHLEKR	401	420	20	717	38826
HD1-t2	cTfh	P1_G8	EKMNTQFTAVGKEFNHLEKR	401	420	20	589	1880
HD1-t2	cTfh	P1_G11	EKMNTQFTAVGKEFNHLEKR	401	420	20	545	48497
HD1-t2	cTfh	P2_A2	EKMNTQFTAVGKEFNHLEKR	401	420	20	544	19086
HD1-t2	cTfh	P2_B1	EKMNTQFTAVGKEFNHLEKR	401	420	20	558	42241
HD1-t2	cTfh	P2_B2	EKMNTQFTAVGKEFNHLEKR	401	420	20	588	22689
HD1-t2	cTfh	P2_B3	EKMNTQFTAVGKEFNHLEKR	401	420	20	439	22995
HD1-t2	cTfh	P2_D9	EKMNTQFTAVGKEFNHLEKR	401	420	20	822	55302
HD1-t2	cTfh	P2_G8	EKMNTQFTAVGKEFNHLEKR	401	420	20	896	33606
HD1-t2	cTfh	P3_A2	EKMNTQFTAVGKEFNHLEKR	401	420	20	1058	31877
HD1-t2	cTfh	P3_C6	EKMNTQFTAVGKEFNHLEKR	401	420	20	860	18204
HD1-t2	cTfh	P3_C10	EKMNTQFTAVGKEFNHLEKR	401	420	20	839	76180
HD1-t2	cTfh	P3_G11	EKMNTQFTAVGKEFNHLEKR	401	420	20	653	38717
HD1-t2	cTfh	P1_B7	EKMNTQFTAVGKEFNHLEKR	401	420	20	1114	15735
HD1-t1	Tcm	M4	EKMNTQFTAVGKEFNHLEKR	401	420	20	59	18915
HD1-t1	Tcm	G5	EKMNTQFTAVGKEFNHLEKR	401	420	20	4	11291
HD1-t2	Tcm	P1_F10	EKMNTQFTAVGKEFNHLEKR	401	420	20	527	9991
HD1-t2	Tcm	P2_B8	EKMNTQFTAVGKEFNHLEKR	401	420	20	500	9565
HD1-t1	Tcm	1_F3	EKMNTQFTAVGKEFNHLEKR	401	420	20	182	4927
HD1-t2	Tem	P1_C2	EKMNTQFTAVGKEFNHLEKR	401	420	20	695	11262
HD1-t2	Tem	P1_C5	EKMNTQFTAVGKEFNHLEKR	401	420	20	829	7612
HD1-t2	Tem	P1_C8	EKMNTQFTAVGKEFNHLEKR	401	420	20	583	12065
HD1-t2	Tem	P1_C10	EKMNTQFTAVGKEFNHLEKR	401	420	20	566	125856
HD1-t2	Tem	P1_D6	EKMNTQFTAVGKEFNHLEKR	401	420	20	562	4678
HD1-t2	Tem	P1_D7	EKMNTQFTAVGKEFNHLEKR	401	420	20	727	2219
HD1-t2	Tem	P1_D12	EKMNTQFTAVGKEFNHLEKR	401	420	20	804	85168
HD1-t2	Tem	P1_E1	EKMNTQFTAVGKEFNHLEKR	401	420	20	701	32527
HD1-t2	Tem	P1_F2	EKMNTQFTAVGKEFNHLEKR	401	420	20	636	12684
HD1-t2	Tem	P1_H8	EKMNTQFTAVGKEFNHLEKR	401	420	20	506	13317
HD1-t2	Tem	P2_A10	EKMNTQFTAVGKEFNHLEKR	401	420	20	749	2090
HD1-t2	Tem	P2_C6	EKMNTQFTAVGKEFNHLEKR	401	420	20	549	5318
HD1-t2	Tem	P2_C10	EKMNTQFTAVGKEFNHLEKR	401	420	20	705	3678
HD1-t2	Tem	P2_C12	EKMNTQFTAVGKEFNHLEKR	401	420	20	596	5462
HD1-t2	Tem	P2_E5	EKMNTQFTAVGKEFNHLEKR	401	420	20	504	1975
HD1-t2	Tem	P2_E11	EKMNTQFTAVGKEFNHLEKR	401	420	20	614	6112
HD1-t2	Tem	P2_F2	EKMNTQFTAVGKEFNHLEKR	401	420	20	732	53825
HD1-t2	Tem	P2_F6	EKMNTQFTAVGKEFNHLEKR	401	420	20	732	9841
HD1-t2	Tem	P2_F12	EKMNTQFTAVGKEFNHLEKR	401	420	20	452	21091
HD1-t2	Tem	P2_G9	EKMNTQFTAVGKEFNHLEKR	401	420	20	363	2675
HD1-t2	Tem	P3_B5	EKMNTQFTAVGKEFNHLEKR	401	420	20	452	6693
HD1-t2	Tem	P3_B8	EKMNTQFTAVGKEFNHLEKR	401	420	20	418	15288
HD1-t2	Tem	P1_B12	EKMNTQFTAVGKEFNHLEKR	401	420	20	749	2173
HD1-t2	Tem	P1_B5	EKMNTQFTAVGKEFNHLEKR	401	420	20	822	3397
HD1-t2	Tem	P1_F3	EKMNTQFTAVGKEFNHLEKR	401	420	20	614	4137
HD1-t2	Tem	P2_C4	EKMNTQFTAVGKEFNHLEKR	401	420	20	708	3788
HD1-t2	Tem	P2_C5	EKMNTQFTAVGKEFNHLEKR	401	420	20	468	5469
HD1-t2	Tem	P2_E12	EKMNTQFTAVGKEFNHLEKR	401	420	20	555	3705
HD1-t2	Tem	P2_H6	EKMNTQFTAVGKEFNHLEKR	401	420	20	546	7255
HD1-t2	Tem	P2_H9	EKMNTQFTAVGKEFNHLEKR	401	420	20	404	2090
HD1-t2	Tem	P2_H11	EKMNTQFTAVGKEFNHLEKR	401	420	20	1128	36506

HD1-t2	cTfh	P3_D8	QFTAVGKEFNHLEKR	406	420	15	568	2965
HD1-t2	Tem	P1_B1	QFTAVGKEFNHLEKR	406	420	15	749	63510
HD1-t2	Tem	P1_B5	QFTAVGKEFNHLEKR	406	420	15	688	16281
HD1-t2	Tem	P1_B12	QFTAVGKEFNHLEKR	406	420	15	862	7668
HD1-t2	Tem	P1_D4	QFTAVGKEFNHLEKR	406	420	15	558	24127
HD1-t2	Tem	P1_F10	QFTAVGKEFNHLEKR	406	420	15	697	8586
HD1-t2	Tem	P2_B8	QFTAVGKEFNHLEKR	406	420	15	649	4252
HD1-t2	Tem	P2_E12	QFTAVGKEFNHLEKR	406	420	15	1189	11381
HD1-t2	Tem	P2_G6	QFTAVGKEFNHLEKR	406	420	15	976	9414
HD1-t2	Tem	P1_B1	QFTAVGKEFNHLEKR	406	420	15	873	59881
HD1-t2	Tem	P1_D4	QFTAVGKEFNHLEKR	406	420	15	658	25169
HD1-t2	Tem	P1_F10	QFTAVGKEFNHLEKR	406	420	15	1237	25675
HD1-t2	Tem	P2_G6	QFTAVGKEFNHLEKR	406	420	15	683	23579
HD1-t2	cTfh	P1_C8	GKEFNHLEKRIENLN	411	425	15	1297	6809
HD1-t2	cTfh	P1_D11	GKEFNHLEKRIENLN	411	425	15	609	9181
HD1-t2	cTfh	P1_B2	GKEFNHLEKRIENLN	411	425	15	949	16047
HD1-t1	Tem	C3	GKEFNHLEKRIENLN	411	425	15	39	3202
HD1-t1	Tem	C6	GKEFNHLEKRIENLN	411	425	15	48	4812
HD1-t1	Tem	C8	GKEFNHLEKRIENLN	411	425	15	75	1054
HD1-t1	Tem	D9	GKEFNHLEKRIENLN	411	425	15	9	908
HD1-t1	Tem	E3	GKEFNHLEKRIENLN	411	425	15	24	10746
HD1-t2	Tem	P1_A10	GKEFNHLEKRIENLN	411	425	15	645	11100
HD1-t2	Tem	P1_C4	GKEFNHLEKRIENLN	411	425	15	671	19960
HD1-t2	Tem	P1_C7	GKEFNHLEKRIENLN	411	425	15	737	5220
HD1-t2	Tem	P1_G9	GKEFNHLEKRIENLN	411	425	15	505	10000
HD1-t2	Tem	P1_G11	GKEFNHLEKRIENLN	411	425	15	575	4441
HD1-t2	Tem	P2_A3	GKEFNHLEKRIENLN	411	425	15	705	2983
HD1-t2	Tem	P2_A6	GKEFNHLEKRIENLN	411	425	15	735	15644
HD1-t2	Tem	P2_A11	GKEFNHLEKRIENLN	411	425	15	557	14299
HD1-t2	Tem	P2_B1	GKEFNHLEKRIENLN	411	425	15	827	7762
HD1-t2	Tem	P2_B11	GKEFNHLEKRIENLN	411	425	15	666	11052
HD1-t2	Tem	P2_D8	GKEFNHLEKRIENLN	411	425	15	486	3339
HD1-t2	Tem	P2_E3	GKEFNHLEKRIENLN	411	425	15	589	13167
HD1-t2	Tem	P2_E6	GKEFNHLEKRIENLN	411	425	15	572	7822
HD1-t2	Tem	P2_E7	GKEFNHLEKRIENLN	411	425	15	513	4700
HD1-t2	Tem	P2_G5	GKEFNHLEKRIENLN	411	425	15	482	2349
HD1-t2	Tem	P2_G10	GKEFNHLEKRIENLN	411	425	15	436	9767
HD1-t2	Tem	P3_B3	GKEFNHLEKRIENLN	411	425	15	572	3352
HD1-t2	Tem	P3_B7	GKEFNHLEKRIENLN	411	425	15	590	2656
HD1-t2	Tem	P1_E3	GKEFNHLEKRIENLN	411	425	15	826	8149
HD1-t2	Tem	P2_C2	GKEFNHLEKRIENLN	411	425	15	1167	4046
HD1-t2	Tem	P2_D5	GKEFNHLEKRIENLN	411	425	15	839	35993
HD1-t2	Tem	P2_D6	GKEFNHLEKRIENLN	411	425	15	839	21265
HD1-t2	Tem	P2_D7	GKEFNHLEKRIENLN	411	425	15	873	9431
HD1-t2	cTfh	P1_B8	GKEFNHLEKRIENLNKKVDD	411	430	20	1557	5049
HD1-t2	cTfh	P1_B9	GKEFNHLEKRIENLNKKVDD	411	430	20	1335	10451
HD1-t2	cTfh	P1_C10	GKEFNHLEKRIENLNKKVDD	411	430	20	2712	11114
HD1-t2	cTfh	P1_D3	GKEFNHLEKRIENLNKKVDD	411	430	20	3428	12258
HD1-t2	cTfh	P1_D8	GKEFNHLEKRIENLNKKVDD	411	430	20	1607	15770
HD1-t2	cTfh	P1_E6	GKEFNHLEKRIENLNKKVDD	411	430	20	1541	15125
HD1-t2	cTfh	P2_G3	GKEFNHLEKRIENLNKKVDD	411	430	20	1389	7065
HD1-t2	cTfh	P3_B10	GKEFNHLEKRIENLNKKVDD	411	430	20	1135	5140
HD1-t2	cTfh	P3_D5	GKEFNHLEKRIENLNKKVDD	411	430	20	1730	6265
HD1-t2	cTfh	P3_F3	GKEFNHLEKRIENLNKKVDD	411	430	20	1126	11392
HD1-t2	cTfh	P3_G4	GKEFNHLEKRIENLNKKVDD	411	430	20	3916	19325
HD1-t2	cTfh	P3_G7	GKEFNHLEKRIENLNKKVDD	411	430	20	961	5573
HD1-t2	cTfh	P3_H7	GKEFNHLEKRIENLNKKVDD	411	430	20	1635	22621
HD1-t2	cTfh	P1_G2	GKEFNHLEKRIENLNKKVDD	411	430	20	1385	12365
HD1-t2	Tcm	P1_B10	GKEFNHLEKRIENLNKKVDD	411	430	20	1376	17180
HD1-t2	Tcm	P1_D9	GKEFNHLEKRIENLNKKVDD	411	430	20	1643	10950
HD1-t2	Tcm	P1_E4	GKEFNHLEKRIENLNKKVDD	411	430	20	1422	14131
HD1-t2	Tcm	P1_F5	GKEFNHLEKRIENLNKKVDD	411	430	20	3510	10997
HD1-t2	Tcm	P1_G8	GKEFNHLEKRIENLNKKVDD	411	430	20	1668	29311
HD1-t2	Tcm	P1_G9	GKEFNHLEKRIENLNKKVDD	411	430	20	636	2660
HD1-t2	Tcm	P2_C8	GKEFNHLEKRIENLNKKVDD	411	430	20	1439	4397
HD1-t2	Tcm	P2_E8	GKEFNHLEKRIENLNKKVDD	411	430	20	1259	22230
HD1-t2	Tcm	P2_E11	GKEFNHLEKRIENLNKKVDD	411	430	20	2595	8928
HD1-t2	Tcm	P2_F10	GKEFNHLEKRIENLNKKVDD	411	430	20	3821	22210
HD1-t2	Tcm	P2_H7	GKEFNHLEKRIENLNKKVDD	411	430	20	1602	5226
HD1-t2	Tcm	P3_B6	GKEFNHLEKRIENLNKKVDD	411	430	20	1607	36033
HD1-t2	Tcm	P3_C4	GKEFNHLEKRIENLNKKVDD	411	430	20	1494	4869
HD1-t2	Tcm	P1_C5	GKEFNHLEKRIENLNKKVDD	411	430	20	1567	5604
HD1-t2	Tcm	P1_G4	GKEFNHLEKRIENLNKKVDD	411	430	20	4132	13350
HD1-t2	Tcm	P2_E9	GKEFNHLEKRIENLNKKVDD	411	430	20	1771	14502

107

HD1-t2	Tem	P2_D5	GKEFNHLEKRIENLNKKVDD	411	430	20	623	19462
HD1-t2	Tem	P1_C12	GKEFNHLEKRIENLNKKVDD	411	430	20	1394	7942
HD1-t2	Tem	P2_C11	GKEFNHLEKRIENLNKKVDD	411	430	20	2523	21745
HD1-t2	Tem	P2_D4	GKEFNHLEKRIENLNKKVDD	411	430	20	1153	11388
HD1-t2	Tem	P2_G1	GKEFNHLEKRIENLNKKVDD	411	430	20	1825	28035
HD1-t2	Tem	P2_H7	GKEFNHLEKRIENLNKKVDD	411	430	20	1588	8809
HD1-t1	cTfh	B3	HLEKRIENLNKKVDD	416	430	15	2	8812
HD1-t1	cTfh	C10	HLEKRIENLNKKVDD	416	430	15	145	2412
HD1-t1	cTfh	C11	HLEKRIENLNKKVDD	416	430	15	22	19297
HD1-t1	cTfh	C6	HLEKRIENLNKKVDD	416	430	15	9	28233
HD1-t1	cTfh	E2	HLEKRIENLNKKVDD	416	430	15	17	16733
HD1-t2	cTfh	P2_B11	HLEKRIENLNKKVDD	416	430	15	636	9331
HD1-t2	cTfh	P2_C6	HLEKRIENLNKKVDD	416	430	15	783	42993
HD1-t2	cTfh	P2_C8	HLEKRIENLNKKVDD	416	430	15	714	16421
HD1-t2	cTfh	P2_C9	HLEKRIENLNKKVDD	416	430	15	1068	10181
HD1-t2	cTfh	P2_C11	HLEKRIENLNKKVDD	416	430	15	649	7886
HD1-t2	cTfh	P2_D1	HLEKRIENLNKKVDD	416	430	15	458	2312
HD1-t2	cTfh	P2_D4	HLEKRIENLNKKVDD	416	430	15	1045	8908
HD1-t2	cTfh	P2_D5	HLEKRIENLNKKVDD	416	430	15	506	3719
HD1-t2	cTfh	P2_D8	HLEKRIENLNKKVDD	416	430	15	575	2233
HD1-t2	cTfh	P2_D12	HLEKRIENLNKKVDD	416	430	15	602	6771
HD1-t2	cTfh	P2_E2	HLEKRIENLNKKVDD	416	430	15	548	4468
HD1-t2	cTfh	P2_E5	HLEKRIENLNKKVDD	416	430	15	522	2094
HD1-t2	cTfh	P2_E8	HLEKRIENLNKKVDD	416	430	15	584	7408
HD1-t2	cTfh	P2_F1	HLEKRIENLNKKVDD	416	430	15	546	1667
HD1-t2	cTfh	P2_F2	HLEKRIENLNKKVDD	416	430	15	748	2565
HD1-t2	cTfh	P2_F4	HLEKRIENLNKKVDD	416	430	15	571	9583
HD1-t2	cTfh	P2_F6	HLEKRIENLNKKVDD	416	430	15	714	16973
HD1-t2	cTfh	P2_F9	HLEKRIENLNKKVDD	416	430	15	500	1637
HD1-t2	cTfh	P2_F12	HLEKRIENLNKKVDD	416	430	15	996	3867
HD1-t2	cTfh	P2_G4	HLEKRIENLNKKVDD	416	430	15	701	18411
HD1-t2	cTfh	P2_G6	HLEKRIENLNKKVDD	416	430	15	554	2664
HD1-t2	cTfh	P2_H4	HLEKRIENLNKKVDD	416	430	15	454	6014
HD1-t2	cTfh	P2_H7	HLEKRIENLNKKVDD	416	430	15	477	8635
HD1-t2	cTfh	P2_H11	HLEKRIENLNKKVDD	416	430	15	684	2588
HD1-t2	cTfh	P3_A4	HLEKRIENLNKKVDD	416	430	15	551	4145
HD1-t2	cTfh	P3_B2	HLEKRIENLNKKVDD	416	430	15	395	5750
HD1-t2	cTfh	P3_B3	HLEKRIENLNKKVDD	416	430	15	568	2775
HD1-t2	cTfh	P3_B6	HLEKRIENLNKKVDD	416	430	15	601	10210
HD1-t2	cTfh	P3_B9	HLEKRIENLNKKVDD	416	430	15	465	14841
HD1-t2	cTfh	P3_C3	HLEKRIENLNKKVDD	416	430	15	422	21295
HD1-t2	cTfh	P3_C4	HLEKRIENLNKKVDD	416	430	15	645	1738
HD1-t2	cTfh	P3_C7	HLEKRIENLNKKVDD	416	430	15	736	12330
HD1-t2	cTfh	P3_C8	HLEKRIENLNKKVDD	416	430	15	468	20087
HD1-t2	cTfh	P3_C9	HLEKRIENLNKKVDD	416	430	15	454	25808
HD1-t2	cTfh	P3_C11	HLEKRIENLNKKVDD	416	430	15	430	22544
HD1-t2	cTfh	P3_D6	HLEKRIENLNKKVDD	416	430	15	526	1597
HD1-t2	cTfh	P3_E2	HLEKRIENLNKKVDD	416	430	15	447	1515
HD1-t2	cTfh	P3_E3	HLEKRIENLNKKVDD	416	430	15	667	49977
HD1-t2	cTfh	P3_E5	HLEKRIENLNKKVDD	416	430	15	490	5587
HD1-t2	cTfh	P3_E9	HLEKRIENLNKKVDD	416	430	15	421	5007
HD1-t2	cTfh	P3_F4	HLEKRIENLNKKVDD	416	430	15	517	4743
HD1-t2	cTfh	P3_F6	HLEKRIENLNKKVDD	416	430	15	670	5717
HD1-t2	cTfh	P3_F8	HLEKRIENLNKKVDD	416	430	15	496	18268
HD1-t2	cTfh	P3_F9	HLEKRIENLNKKVDD	416	430	15	644	2795
HD1-t2	cTfh	P3_F10	HLEKRIENLNKKVDD	416	430	15	567	1748
HD1-t2	cTfh	P3_G5	HLEKRIENLNKKVDD	416	430	15	697	17982
HD1-t2	cTfh	P3_G6	HLEKRIENLNKKVDD	416	430	15	662	8158
HD1-t2	cTfh	P3_G8	HLEKRIENLNKKVDD	416	430	15	796	4640
HD1-t2	cTfh	P3_H2	HLEKRIENLNKKVDD	416	430	15	662	21611
HD1-t2	cTfh	P3_H4	HLEKRIENLNKKVDD	416	430	15	1153	14160
HD1-t2	cTfh	P3_H6	HLEKRIENLNKKVDD	416	430	15	432	1743
HD1-t2	cTfh	P3_H8	HLEKRIENLNKKVDD	416	430	15	1035	6409
HD1-t2	cTfh	P1_A4	HLEKRIENLNKKVDD	416	430	15	679	25835
HD1-t2	cTfh	P1_A6	HLEKRIENLNKKVDD	416	430	15	636	47018
HD1-t2	cTfh	P1_C1	HLEKRIENLNKKVDD	416	430	15	554	22117
HD1-t2	cTfh	P1_C3	HLEKRIENLNKKVDD	416	430	15	675	2371
HD1-t2	cTfh	P1_C11	HLEKRIENLNKKVDD	416	430	15	779	37658
HD1-t2	cTfh	P1_D9	HLEKRIENLNKKVDD	416	430	15	770	8599
HD1-t2	cTfh	P1_D12	HLEKRIENLNKKVDD	416	430	15	531	3470
HD1-t2	cTfh	P1_E2	HLEKRIENLNKKVDD	416	430	15	566	1872
HD1-t2	cTfh	P1_E3	HLEKRIENLNKKVDD	416	430	15	514	4496
HD1-t2	cTfh	P1_E7	HLEKRIENLNKKVDD	416	430	15	792	24727
HD1-t2	cTfh	P1_E9	HLEKRIENLNKKVDD	416	430	15	692	7909

HD1-t2	cTfh	P1_F1	HLEKRIENLNKKVDD	416	430	15	1356	19303
HD1-t2	cTfh	P1_F2	HLEKRIENLNKKVDD	416	430	15	521	1988
HD1-t2	cTfh	P1_F4	HLEKRIENLNKKVDD	416	430	15	632	5014
HD1-t2	cTfh	P1_F7	HLEKRIENLNKKVDD	416	430	15	1344	10526
HD1-t2	cTfh	P1_F10	HLEKRIENLNKKVDD	416	430	15	649	20758
HD1-t2	cTfh	P1_F11	HLEKRIENLNKKVDD	416	430	15	680	2873
HD1-t2	cTfh	P1_F12	HLEKRIENLNKKVDD	416	430	15	724	13223
HD1-t2	cTfh	P1_G1	HLEKRIENLNKKVDD	416	430	15	772	2277
HD1-t2	cTfh	P1_G6	HLEKRIENLNKKVDD	416	430	15	589	21166
HD1-t2	cTfh	P1_G9	HLEKRIENLNKKVDD	416	430	15	690	59730
HD1-t2	cTfh	P1_G10	HLEKRIENLNKKVDD	416	430	15	626	62687
HD1-t2	cTfh	P1_H2	HLEKRIENLNKKVDD	416	430	15	509	21180
HD1-t2	cTfh	P1_H4	HLEKRIENLNKKVDD	416	430	15	364	12485
HD1-t2	cTfh	P1_H5	HLEKRIENLNKKVDD	416	430	15	697	4826
HD1-t2	cTfh	P1_H6	HLEKRIENLNKKVDD	416	430	15	723	42107
HD1-t2	cTfh	P1_H8	HLEKRIENLNKKVDD	416	430	15	641	4145
HD1-t2	cTfh	P1_H9	HLEKRIENLNKKVDD	416	430	15	434	10398
HD1-t2	cTfh	P1_H10	HLEKRIENLNKKVDD	416	430	15	754	6232
HD1-t2	cTfh	P2_A4	HLEKRIENLNKKVDD	416	430	15	535	27800
HD1-t2	cTfh	P2_A8	HLEKRIENLNKKVDD	416	430	15	529	54851
HD1-t2	cTfh	P2_A9	HLEKRIENLNKKVDD	416	430	15	800	28103
HD1-t2	cTfh	P2_A10	HLEKRIENLNKKVDD	416	430	15	468	19765
HD1-t2	cTfh	P2_B4	HLEKRIENLNKKVDD	416	430	15	553	31324
HD1-t2	cTfh	P1_B3	HLEKRIENLNKKVDD	416	430	15	1593	6093
HD1-t2	cTfh	P1_B4	HLEKRIENLNKKVDD	416	430	15	1284	4691
HD1-t2	cTfh	P1_B5	HLEKRIENLNKKVDD	416	430	15	996	3010
HD1-t2	cTfh	P1_B10	HLEKRIENLNKKVDD	416	430	15	1027	19946
HD1-t2	cTfh	P1_B11	HLEKRIENLNKKVDD	416	430	15	1214	8141
HD1-t2	cTfh	P1_C2	HLEKRIENLNKKVDD	416	430	15	1302	5954
HD1-t2	cTfh	P1_C3	HLEKRIENLNKKVDD	416	430	15	886	2569
HD1-t2	cTfh	P1_C7	HLEKRIENLNKKVDD	416	430	15	1426	7808
HD1-t2	cTfh	P1_D2	HLEKRIENLNKKVDD	416	430	15	1428	5030
HD1-t2	cTfh	P1_D4	HLEKRIENLNKKVDD	416	430	15	696	2299
HD1-t2	cTfh	P1_D5	HLEKRIENLNKKVDD	416	430	15	1448	6387
HD1-t2	cTfh	P1_D7	HLEKRIENLNKKVDD	416	430	15	1330	16393
HD1-t2	cTfh	P1_E4	HLEKRIENLNKKVDD	416	430	15	1295	4201
HD1-t2	cTfh	P1_F3	HLEKRIENLNKKVDD	416	430	15	1299	22039
HD1-t2	cTfh	P1_G5	HLEKRIENLNKKVDD	416	430	15	999	72858
HD1-t2	cTfh	P2_A7	HLEKRIENLNKKVDD	416	430	15	892	2546
HD1-t2	cTfh	P2_B5	HLEKRIENLNKKVDD	416	430	15	870	45752
HD1-t2	cTfh	P2_D4	HLEKRIENLNKKVDD	416	430	15	727	2470
HD1-t2	cTfh	P2_D10	HLEKRIENLNKKVDD	416	430	15	723	3966
HD1-t2	cTfh	P2_F1	HLEKRIENLNKKVDD	416	430	15	731	3178
HD1-t2	cTfh	P2_F9	HLEKRIENLNKKVDD	416	430	15	1531	17466
HD1-t2	cTfh	P3_C11	HLEKRIENLNKKVDD	416	430	15	809	103047
HD1-t2	cTfh	P3_D6	HLEKRIENLNKKVDD	416	430	15	886	34388
HD1-t2	cTfh	P3_E2	HLEKRIENLNKKVDD	416	430	15	809	2500
HD1-t2	cTfh	P3_E4	HLEKRIENLNKKVDD	416	430	15	1239	21131
HD1-t2	cTfh	P3_E6	HLEKRIENLNKKVDD	416	430	15	685	23651
HD1-t2	cTfh	P3_G10	HLEKRIENLNKKVDD	416	430	15	1158	7250
HD1-t2	cTfh	P3_H4	HLEKRIENLNKKVDD	416	430	15	732	59705
HD1-t2	cTfh	P3_H5	HLEKRIENLNKKVDD	416	430	15	1085	6219
HD1-t2	cTfh	P3_H6	HLEKRIENLNKKVDD	416	430	15	916	19507
HD1-t2	cTfh	P3_H8	HLEKRIENLNKKVDD	416	430	15	723	13909
HD1-t1	Tcm	B8	HLEKRIENLNKKVDD	416	430	15	54	20572
HD1-t1	Tcm	C5	HLEKRIENLNKKVDD	416	430	15	9	971
HD1-t1	Tcm	C5	HLEKRIENLNKKVDD	416	430	15	22	3102
HD1-t1	Tcm	C3	HLEKRIENLNKKVDD	416	430	15	2130	7311
HD1-t2	Tcm	P1_A5	HLEKRIENLNKKVDD	416	430	15	696	22273
HD1-t2	Tcm	P1_A7	HLEKRIENLNKKVDD	416	430	15	592	17321
HD1-t2	Tcm	P1_A9	HLEKRIENLNKKVDD	416	430	15	479	25765
HD1-t2	Tcm	P1_A10	HLEKRIENLNKKVDD	416	430	15	736	4320
HD1-t2	Tcm	P1_B6	HLEKRIENLNKKVDD	416	430	15	814	17457
HD1-t2	Tcm	P1_B7	HLEKRIENLNKKVDD	416	430	15	510	11746
HD1-t2	Tcm	P1_B8	HLEKRIENLNKKVDD	416	430	15	549	9538
HD1-t2	Tcm	P1_B9	HLEKRIENLNKKVDD	416	430	15	705	9536
HD1-t2	Tcm	P1_C1	HLEKRIENLNKKVDD	416	430	15	610	12153
HD1-t2	Tcm	P1_C3	HLEKRIENLNKKVDD	416	430	15	692	16949
HD1-t2	Tcm	P1_C6	HLEKRIENLNKKVDD	416	430	15	684	7714
HD1-t2	Tcm	P1_C7	HLEKRIENLNKKVDD	416	430	15	503	10704
HD1-t2	Tcm	P1_C9	HLEKRIENLNKKVDD	416	430	15	602	9964
HD1-t2	Tcm	P1_C11	HLEKRIENLNKKVDD	416	430	15	701	31107
HD1-t2	Tcm	P1_D6	HLEKRIENLNKKVDD	416	430	15	513	46150
HD1-t2	Tcm	P1_D10	HLEKRIENLNKKVDD	416	430	15	592	5779

HD1-t2	Tcm	P1_E5	HLEKRIENLNKKVDD	416	430	15	561	6183
HD1-t2	Tcm	P1_E6	HLEKRIENLNKKVDD	416	430	15	474	3685
HD1-t2	Tcm	P1_E7	HLEKRIENLNKKVDD	416	430	15	478	4291
HD1-t2	Tcm	P1_E9	HLEKRIENLNKKVDD	416	430	15	503	13853
HD1-t2	Tcm	P1_F1	HLEKRIENLNKKVDD	416	430	15	632	8111
HD1-t2	Tcm	P1_F2	HLEKRIENLNKKVDD	416	430	15	538	37319
HD1-t2	Tcm	P1_F4	HLEKRIENLNKKVDD	416	430	15	772	13290
HD1-t2	Tcm	P1_F8	HLEKRIENLNKKVDD	416	430	15	486	6378
HD1-t2	Tcm	P1_F9	HLEKRIENLNKKVDD	416	430	15	658	6128
HD1-t2	Tcm	P1_G1	HLEKRIENLNKKVDD	416	430	15	447	14854
HD1-t2	Tcm	P1_G6	HLEKRIENLNKKVDD	416	430	15	561	22736
HD1-t2	Tcm	P1_G7	HLEKRIENLNKKVDD	416	430	15	533	10665
HD1-t2	Tcm	P1_G11	HLEKRIENLNKKVDD	416	430	15	749	3215
HD1-t2	Tcm	P1_G12	HLEKRIENLNKKVDD	416	430	15	525	17067
HD1-t2	Tcm	P1_H2	HLEKRIENLNKKVDD	416	430	15	529	14523
HD1-t2	Tcm	P1_H3	HLEKRIENLNKKVDD	416	430	15	649	43914
HD1-t2	Tcm	P1_H7	HLEKRIENLNKKVDD	416	430	15	518	2471
HD1-t2	Tcm	P1_H9	HLEKRIENLNKKVDD	416	430	15	613	21731
HD1-t2	Tcm	P2_A2	HLEKRIENLNKKVDD	416	430	15	566	4115
HD1-t2	Tcm	P2_A3	HLEKRIENLNKKVDD	416	430	15	610	3712
HD1-t2	Tcm	P2_A4	HLEKRIENLNKKVDD	416	430	15	702	4304
HD1-t2	Tcm	P2_A9	HLEKRIENLNKKVDD	416	430	15	469	10350
HD1-t2	Tcm	P2_B1	HLEKRIENLNKKVDD	416	430	15	568	8437
HD1-t2	Tcm	P2_B6	HLEKRIENLNKKVDD	416	430	15	535	20170
HD1-t2	Tcm	P2_B7	HLEKRIENLNKKVDD	416	430	15	535	3549
HD1-t2	Tcm	P2_B9	HLEKRIENLNKKVDD	416	430	15	816	7662
HD1-t2	Tcm	P2_D10	HLEKRIENLNKKVDD	416	430	15	1063	3651
HD1-t2	Tcm	P2_E6	HLEKRIENLNKKVDD	416	430	15	705	3001
HD1-t2	Tcm	P2_E10	HLEKRIENLNKKVDD	416	430	15	830	6882
HD1-t2	Tcm	P2_F2	HLEKRIENLNKKVDD	416	430	15	1179	22353
HD1-t2	Tcm	P2_F4	HLEKRIENLNKKVDD	416	430	15	1248	25969
HD1-t2	Tcm	P2_F5	HLEKRIENLNKKVDD	416	430	15	936	4872
HD1-t2	Tcm	P2_F12	HLEKRIENLNKKVDD	416	430	15	836	9370
HD1-t2	Tcm	P2_G1	HLEKRIENLNKKVDD	416	430	15	784	6850
HD1-t2	Tcm	P2_G7	HLEKRIENLNKKVDD	416	430	15	662	3717
HD1-t2	Tcm	P2_G10	HLEKRIENLNKKVDD	416	430	15	843	3189
HD1-t2	Tcm	P2_H6	HLEKRIENLNKKVDD	416	430	15	845	16522
HD1-t2	Tcm	P3_B7	HLEKRIENLNKKVDD	416	430	15	910	9052
HD1-t2	Tcm	P3_B8	HLEKRIENLNKKVDD	416	430	15	940	3402
HD1-t2	Tcm	P3_B9	HLEKRIENLNKKVDD	416	430	15	597	9496
HD1-t2	Tcm	P3_C5	HLEKRIENLNKKVDD	416	430	15	572	2782
HD1-t2	Tcm	P1_C4	HLEKRIENLNKKVDD	416	430	15	783	2272
HD1-t2	Tcm	P2_B4	HLEKRIENLNKKVDD	416	430	15	737	8482
HD1-t2	Tcm	P2_B10	HLEKRIENLNKKVDD	416	430	15	605	7337
HD1-t2	Tcm	P2_B12	HLEKRIENLNKKVDD	416	430	15	792	8661
HD1-t2	Tcm	P2_C6	HLEKRIENLNKKVDD	416	430	15	869	4114
HD1-t2	Tcm	P2_C11	HLEKRIENLNKKVDD	416	430	15	525	1958
HD1-t2	Tcm	P2_C12	HLEKRIENLNKKVDD	416	430	15	895	37463
HD1-t2	Tcm	P2_D9	HLEKRIENLNKKVDD	416	430	15	590	3942
HD1-t2	Tcm	P2_F1	HLEKRIENLNKKVDD	416	430	15	926	21726
HD1-t2	Tcm	P2_F7	HLEKRIENLNKKVDD	416	430	15	631	2259
HD1-t2	Tcm	P2_G6	HLEKRIENLNKKVDD	416	430	15	1480	5311
HD1-t2	Tcm	P3_B8	HLEKRIENLNKKVDD	416	430	15	770	12547
HD1-t1	Tem	B3	HLEKRIENLNKKVDD	416	430	15	57	3913
HD1-t1	Tem	F4	HLEKRIENLNKKVDD	416	430	15	48	6015
HD1-t2	Tem	P1_A9	HLEKRIENLNKKVDD	416	430	15	757	6386
HD1-t2	Tem	P1_A11	HLEKRIENLNKKVDD	416	430	15	705	36772
HD1-t2	Tem	P1_A12	HLEKRIENLNKKVDD	416	430	15	853	2501
HD1-t2	Tem	P1_C1	HLEKRIENLNKKVDD	416	430	15	684	4556
HD1-t2	Tem	P1_C3	HLEKRIENLNKKVDD	416	430	15	858	2466
HD1-t2	Tem	P1_C6	HLEKRIENLNKKVDD	416	430	15	1292	5707
HD1-t2	Tem	P1_C9	HLEKRIENLNKKVDD	416	430	15	614	2460
HD1-t2	Tem	P1_E12	HLEKRIENLNKKVDD	416	430	15	770	8641
HD1-t2	Tem	P1_F4	HLEKRIENLNKKVDD	416	430	15	588	2316
HD1-t2	Tem	P1_G1	HLEKRIENLNKKVDD	416	430	15	645	7054
HD1-t2	Tem	P1_G2	HLEKRIENLNKKVDD	416	430	15	921	4779
HD1-t2	Tem	P1_G3	HLEKRIENLNKKVDD	416	430	15	671	17904
HD1-t2	Tem	P1_H6	HLEKRIENLNKKVDD	416	430	15	653	4661
HD1-t2	Tem	P2_A1	HLEKRIENLNKKVDD	416	430	15	653	3303
HD1-t2	Tem	P2_A2	HLEKRIENLNKKVDD	416	430	15	894	5115
HD1-t2	Tem	P2_C3	HLEKRIENLNKKVDD	416	430	15	675	7089
HD1-t2	Tem	P2_D11	HLEKRIENLNKKVDD	416	430	15	542	3240
HD1-t2	Tem	P2_E10	HLEKRIENLNKKVDD	416	430	15	599	8805
HD1-t2	Tem	P2_F8	HLEKRIENLNKKVDD	416	430	15	564	3133

110

HD1-t2	Tem	P2_H3	HLEKRIENLNKKVDD	416	430	15	672	8814
HD1-t2	Tem	P2_H8	HLEKRIENLNKKVDD	416	430	15	631	2878
HD1-t2	Tem	P3_B4	HLEKRIENLNKKVDD	416	430	15	555	6973
HD1-t2	Tem	P1_E2	HLEKRIENLNKKVDD	416	430	15	626	4187
HD1-t2	Tem	P1_F7	HLEKRIENLNKKVDD	416	430	15	658	66991
HD1-t2	Tem	P1_G10	HLEKRIENLNKKVDD	416	430	15	766	3856
HD1-t2	Tem	P2_B4	HLEKRIENLNKKVDD	416	430	15	731	4028
HD1-t2	Tem	P2_B8	HLEKRIENLNKKVDD	416	430	15	1105	278741
HD1-t2	Tcm	P1_H6	ELLVLLENERTLDYH	441	455	15	5720	76062
HD1-t2	cTfh	P1_C5	ELLVLLENERTLDYHDSNVK	441	460	20	2630	12579
HD1-t1	Tcm	B5	LENERTLDYHDSNVK	446	460	15	67	38816
HD1-t1	Tcm	I4	LENERTLDYHDSNVK	446	460	15	17	13057
HD1-t1	Tcm	L3	LENERTLDYHDSNVK	446	460	15	9	6635
HD1-t1	Tcm	M11	LENERTLDYHDSNVK	446	460	15	23	8795
HD1-t1	Tcm	F8	CFEFYHKCDNTCMES	481	495	15	1576	10408
HD1-t1	Tcm	C6	CFEFYHKCDNTCMES	481	495	15	291	1631
HD1-t1	Tcm	M2	CFEFYHKCDNTCMES	481	495	15	323	3155
HD1-t2	cTfh	P2_B6	KLNREEIDGVKLESTRIYQI	511	530	20	1576	7823
HD1-t2	cTfh	P1_E5	RIYQILAIYSTVASS	526	540	15	4464	20392
HD1-t1	Tcm	B3	RIYQILAIYSTVASS	526	540	15	54	2331
HD1-t1	Tcm	B7	RIYQILAIYSTVASS	526	540	15	22	34735
HD1-t1	Tcm	E3	RIYQILAIYSTVASS	526	540	15	50	16289
HD1-t1	Tcm	H4	RIYQILAIYSTVASS	526	540	15	19	29613
HD1-t1	Tcm	I6	RIYQILAIYSTVASS	526	540	15	4	12824
HD1-t1	Tcm	M9	RIYQILAIYSTVASS	526	540	15	75	42415

Supplementary Table 6.2-2. TCR-Vβ sequence and epitope specificity of H1-HA-reactive T cell clones isolated from memory or naïve compartments. Start and end position of the epitope residues in H1-HA A/California/07/2009(H1N1) recognized by CD4[+] Tcm, Tem and cTfh subsets are indicated. The number of sister clones isolated from each subset is reported.

Subset	ID	Vβ gene	Jβ gene	Vβ CDR3	Epitope	Start	Stop	Length	N sister clones			
									tot	from Tcm	from Tem	from cTfh
Memory	mem_1	TRBV20-1	TRBJ2-5	CSASPGRGETQYF	TFATANADTLCIGYH	11	25	15	1		1	
Memory	mem_2	TRBV12-5	TRBJ1-5	CASGINSNHPQHF	PLHLGKCNIAGWILG	66	80	15	1		1	
Memory	mem_3	TRBV11-3	TRBJ2-5	CASSPGGETQYF	KGKEVLVLWGIHHPS	186	200	15	2	1		1
Memory	mem_4	TRBV6-1	TRBJ2-5	CASSISGTGFQETQYF	KGKEVLVLWGIHHPS	186	200	15	5	5		
Memory	mem_5	TRBV18	TRBJ2-6	CASSPSAAGANVLTF	KGKEVLVLWGIHHPS	186	200	15	6	6		
Memory	mem_6	TRBV6-5	TRBJ1-5	CASSYSSFLREPQHF	KGKEVLVLWGIHHPS	186	200	15	1	1		
Memory	mem_7	TRBV12-4	TRBJ1-3	CASTRESWRRNTIYF	KGKEVLVLWGIHHPS	186	200	15	2	1	1	
Memory	mem_8	TRBV11-3	TRBJ2-3	CASRTGADTQYF	KGKEVLVLWGIHHPSTSADQ	186	205	20	1	1		
Memory	mem_9	TRBV12-4	TRBJ2-6	CASRRDSGTGANVLTF	KFKPEIAIRPKVRDQ	226	240	15	2	2		
Memory	mem_10	TRBV11-2	TRBJ2-7	CASSLAPGQGLYEQYF	KFKPEIAIRPKVRDQ	226	240	15	1	1		
Memory	mem_11	TRBV5-1	TRBJ2-2	CASSLGPQSKNTGELFF	EGRMNYYWTLVEPGD	241	255	15	1		1	
Memory	mem_12	TRBV11-2	TRBJ1-6	CASSLQGAGSSPLHF	EGRMNYYWTLVEPGD	241	255	15	2		1	1
Memory	mem_13	TRBV7-2	TRBJ2-1	CASSLWRRGSYNEQFF	EGRMNYYWTLVEPGD	241	255	15	1	1		
Memory	mem_14	TRBV6-4	TRBJ2-1	CSSGEQGAGGNEQFF	EGRMNYYWTLVEPGD	241	255	15	1			1
Memory	mem_15	TRBV7-9	TRBJ1-3	CASSPGTESSGNTIYF	EGRMNYYWTLVEPGDKITFE	241	260	20	2		2	
Memory	mem_16	TRBV6-4	TRBJ2-6	CASSPQGCGANVLTF	EGRMNYYWTLVEPGDKITFE	241	260	20	1			1
Memory	mem_17	TRBV6-4	TRBJ2-2	CASSQAGNSGELFF	EGRMNYYWTLVEPGDKITFE	241	260	20	1		1	
Memory	mem_18	TRBV6-4	TRBJ2-2	CASSQAGNTGELFF	EGRMNYYWTLVEPGDKITFE	241	260	20	1		1	
Memory	mem_19	TRBV12-3	TRBJ2-6	CASSQQGSGANVLTF	EGRMNYYWTLVEPGDKITFE	241	260	20	2	2		
Memory	mem_20	TRBV6-4	TRBJ2-3	CASSQQGTGADTQYF	EGRMNYYWTLVEPGDKITFE	241	260	20	1	1		
Memory	mem_21	TRBV5-5	TRBJ2-7	CASSSTVPSYEQYF	EGRMNYYWTLVEPGDKITFE	241	260	20	2	2		
Memory	mem_22	TRBV19	TRBJ2-7	CASSTNLGDYEQYF	YYWTLVEPGDKITFE	246	260	15	1	1		
Memory	mem_23	TRBV2	TRBJ1-5	CASSVGQGVSQPQHF	VVPRYAFAMERNAGS	266	280	15	1	1		
Memory	mem_24	TRBV6-1	TRBJ2-1	CASSESTSGTFNEQFF	QNAIDEITNKVNSVI	386	400	15	1	1		
Memory	mem_25	TRBV18	TRBJ1-3	CASSPLQGISGRPRSF	VNSVIEKMNTQFTAV	396	410	15	1	1		
Memory	mem_26	TRBV6-5	TRBJ2-5	CASSTGDTGGRQETQYF	VNSVIEKMNTQFTAV	396	410	15	1			1
Memory	mem_27	TRBV28	TRBJ1-5	CASRRKSVGQVGQPQHF	EKMNTQFTAVGKEFN	401	415	15	1			1
Memory	mem_28	TRBV10-2	TRBJ2-3	CASSELGQATDTQYF	EKMNTQFTAVGKEFNHLEKR	401	420	20	10		10	
Memory	mem_29	TRBV11-1	TRBJ2-7	CASSQVRGAQTYEQYF	EKMNTQFTAVGKEFNHLEKR	401	420	20	1			1
Memory	mem_30	TRBV12-4	TRBJ1-6	CASSFPPPIHF	EKMNTQFTAVGKEFNHLEKR	401	420	20	2			2
Memory	mem_31	TRBV14	TRBJ1-2	CASSQVVSGYGYTF	EKMNTQFTAVGKEFNHLEKR	401	420	20	8			8
Memory	mem_32	TRBV14	TRBJ1-2	CASSQLSGAYGHFF	EKMNTQFTAVGKEFNHLEKR	401	420	20	1			1
Memory	mem_33	TRBV14	TRBJ1-2	CASSQVNSAYGHTF	EKMNTQFTAVGKEFNHLEKR	401	420	20	1			1
Memory	mem_34	TRBV20-1	TRBJ1-2	CSASTPGGIYGYTF	EKMNTQFTAVGKEFNHLEKR	401	420	20	10			10
Memory	mem_35	TRBV20-1	TRBJ1-5	CSLDRPNQPQHF	EKMNTQFTAVGKEFNHLEKR	401	420	20	6			6
Memory	mem_36	TRBV20-1	TRBJ2-3	CSARVGGSSDTQYF	EKMNTQFTAVGKEFNHLEKR	401	420	20	1		1	
Memory	mem_37	TRBV24-1	TRBJ2-2	CATLGMGTGELFF	EKMNTQFTAVGKEFNHLEKR	401	420	20	3	3		
Memory	mem_38	TRBV24-1	TRBJ2-3	CATQGMGTDTQYF	EKMNTQFTAVGKEFNHLEKR	401	420	20	2			2
Memory	mem_39	TRBV27	TRBJ2-6	CASRSGGSGANVLTF	EKMNTQFTAVGKEFNHLEKR	401	420	20	4		4	
Memory	mem_40	TRBV3-1	TRBJ2-1	CASSPSTSGGSQFF	EKMNTQFTAVGKEFNHLEKR	401	420	20	10		10	
Memory	mem_41	TRBV7-8	TRBJ2-2	CASTQAAGLNTGELFF	EKMNTQFTAVGKEFNHLEKR	401	420	20	1	1		

Memory	mem_42	TRBV6-6	TRBJ1-4	CASTPAANEKLFF	QFTAVGKEFNHLEKR	406	420	15	1			1
Memory	mem_43	TRBV7-9	TRBJ2-4	CASSVTSGGEVKNIQYF	GKEFNHLEKRIENLN	411	425	15	14		14	
Memory	mem_44	TRBV7-9	TRBJ2-4	CASSVTSGGEVKNVQYF	GKEFNHLEKRIENLN	411	425	15	2			2
Memory	mem_45	TRBV7-9	TRBJ1-1	CGAAAYPEGGEIKNKQFF	GKEFNHLEKRIENLN	411	425	15	1		1	
Memory	mem_46	TRBV12-4	TRBJ1-1	CASSFTMNTEAFF	HLEKRIENLNKKVDD	416	430	15	5	5		
Memory	mem_47	TRBV12-4	TRBJ1-1	CASSFTMYTKAFF	HLEKRIENLNKKVDD	416	430	15	1	1		
Memory	mem_48	TRBV19	TRBJ2-7	CASSTASGRSSYEQYF	HLEKRIENLNKKVDD	416	430	15	19			19
Memory	mem_49	TRBV19	TRBJ1-2	CASSSEGGGRANGYTF	HLEKRIENLNKKVDD	416	430	15	16			16
Memory	mem_50	TRBV19	TRBJ1-5	CAIGQGGSNQPQHF	HLEKRIENLNKKVDD	416	430	15	11			11
Memory	mem_51	TRBV19	TRBJ2-3	CASSPSSGSRDTQYF	HLEKRIENLNKKVDD	416	430	15	9		4	5
Memory	mem_52	TRBV19	TRBJ1-5	CASSSEGGGRQPQHF	HLEKRIENLNKKVDD	416	430	15	9		7	2
Memory	mem_53	TRBV19	TRBJ2-5	CASQGTSGGRETQYF	HLEKRIENLNKKVDD	416	430	15	8	1	1	6
Memory	mem_54	TRBV19	TRBJ1-1	CASSEVRGGRTEAFF	HLEKRIENLNKKVDD	416	430	15	8	5	3	
Memory	mem_55	TRBV19	TRBJ1-5	CATGQGGSNQPQHF	HLEKRIENLNKKVDD	416	430	15	6		1	5
Memory	mem_56	TRBV19	TRBJ1-5	CASGQGGSNQPQHF	HLEKRIENLNKKVDD	416	430	15	4			4
Memory	mem_57	TRBV19	TRBJ1-5	CASTSETGGRQPQHF	HLEKRIENLNKKVDD	416	430	15	4			4
Memory	mem_58	TRBV19	TRBJ1-2	CASKDVAGTAKDGYTF	HLEKRIENLNKKVDD	416	430	15	2		1	1
Memory	mem_59	TRBV19	TRBJ1-5	CASSAEFQGRQPQHF	HLEKRIENLNKKVDD	416	430	15	2		1	1
Memory	mem_60	TRBV19	TRBJ2-5	CASSPQSGKETQYF	HLEKRIENLNKKVDD	416	430	15	2	1	1	
Memory	mem_61	TRBV19	TRBJ1-2	CASSSEGGGRVNGYTF	HLEKRIENLNKKVDD	416	430	15	2			2
Memory	mem_62	TRBV19	TRBJ1-1	CASNEVRGGNHQPLF	HLEKRIENLNKKVDD	416	430	15	1	1		
Memory	mem_63	TRBV19	TRBJ1-2	CASRSTGGGLEGYTF	HLEKRIENLNKKVDD	416	430	15	1			1
Memory	mem_64	TRBV19	TRBJ1-5	CASSADTGGRQPQHF	HLEKRIENLNKKVDD	416	430	15	1			1
Memory	mem_65	TRBV19	TRBJ1-1	CASSEVRGGLTEAFF	HLEKRIENLNKKVDD	416	430	15	1	1		
Memory	mem_66	TRBV19	TRBJ2-5	CASSSAGKETQYF	HLEKRIENLNKKVDD	416	430	15	1			1
Memory	mem_67	TRBV19	TRBJ1-2	CASSSEGQGRGGYTF	HLEKRIENLNKKVDD	416	430	15	1		1	
Memory	mem_68	TRBV19	TRBJ2-7	CASSTASGRFSYEQYF	HLEKRIENLNKKVDD	416	430	15	1			1
Memory	mem_69	TRBV19	TRBJ2-7	CASSTASGRSFFEQYF	HLEKRIENLNKKVDD	416	430	15	1			1
Memory	mem_70	TRBV19	TRBJ2-7	CASSTASGRSFYEQYF	HLEKRIENLNKKVDD	416	430	15	1			1
Memory	mem_71	TRBV19	TRBJ2-7	CASSTASGRSLYEQYF	HLEKRIENLNKKVDD	416	430	15	1			1
Memory	mem_72	TRBV19	TRBJ2-7	CASVPPAGGHSYEQYF	HLEKRIENLNKKVDD	416	430	15	1			1
Memory	mem_73	TRBV19	TRBJ1-2	CHSSSEGQGRGGHTF	HLEKRIENLNKKVDD	416	430	15	1		1	
Memory	mem_74	TRBV20-1	TRBJ2-2	CSARDGSWTGELFF	HLEKRIENLNKKVDD	416	430	15	10	10		
Memory	mem_75	TRBV20-1	TRBJ2-2	CSARDGSWSGELFF	HLEKRIENLNKKVDD	416	430	15	1	1		
Memory	mem_76	TRBV20-1	TRBJ2-2	CSGRDGSWNGELFF	HLEKRIENLNKKVDD	416	430	15	1	1		
Memory	mem_77	TRBV3-1	TRBJ1-6	CASSQGTPGSPLHF	HLEKRIENLNKKVDD	416	430	15	2		2	
Memory	mem_78	TRBV3-1	TRBJ2-5	CASSESSGRETQYF	HLEKRIENLNKKVDD	416	430	15	1			1
Memory	mem_79	TRBV3-1	TRBJ2-3	CASSSQGSRTDTQYF	HLEKRIENLNKKVDD	416	430	15	1		1	
Memory	mem_80	TRBV6-6	TRBJ1-5	CASSYSSANTGGNQPQHF	HLEKRIENLNKKVDD	416	430	15	1			1
Memory	mem_81	TRBV7-2	TRBJ1-4	CAAGTAKEKLFF	HLEKRIENLNKKVDD	416	430	15	3	2		1
Memory	mem_82	TRBV10-2	TRBJ2-1	CSSRPGEGGNNEQFF	ELLVLLENERTLDYH	441	455	15	1	1		
Memory	mem_83	TRBV11-2	TRBJ2-7	CASSPRTDYEQYF	LENERTLDYHDSNVK	446	460	15	3	3		
Memory	mem_84	TRBV30	TRBJ2-2	CAWSGDSYTGELFF	LENERTLDYHDSNVK	446	460	15	2	2		
Memory	mem_85	TRBV11-3	TRBJ2-1	CASSLDPTNNEQFF	CFEFYHKCDNTCMES	481	495	15	2	2		
Memory	mem_86	TRBV6-2	TRBJ2-2	CASSYREVGDRGPVF	CFEFYHKCDNTCMES	481	495	15	1	1		
Memory	mem_87	TRBV5-5	TRBJ2-7	CASSLNGLSYEQYF	RIYQILAIYSTVASS	526	540	15	1	1		
Memory	mem_88	TRBV9	TRBJ2-7	CASSVAAQAYEQYF	RIYQILAIYSTVASS	526	540	15	3	3		

Clonotypes (tot): 88   Clones (tot): 274

113

Subset	ID	Vβ gene	Jβ gene	Vβ CDR3	Epitope	Start	Stop	Length	N sister clones
Naive	nai_1	TRBV9	TRBJ2-5	CASSLQGQETQYF	MKAILVVLLYTFATANADTL	1	20	20	1
Naive	nai_2	TRBV11-2	TRBJ2-3	CASSTGTGPDTQYF	EDKHNGKLCKLRGVAPLHLG	51	70	20	1
Naive	nai_3	TRBV30	TRBJ1-1	CAWTPPQSTGDTEAFF	EDKHNGKLCKLRGVAPLHLG	51	70	20	1
Naive	nai_4	TRBV11-1	TRBJ2-7	CASSLKDIYEQYF	GKLCKLRGVAPLHLG	56	70	15	1
Naive	nai_5	TRBV11-2	TRBJ1-5	CASSSDSNQPQHF	GKLCKLRGVAPLHLG	56	70	15	1
Naive	nai_6	TRBV9	TRBJ2-1	CASSVGDYLGYNEQFF	GKLCKLRGVAPLHLG	56	70	15	1
Naive	nai_7	TRBV20-1	TRBJ1-5	CSATQGINQPQHF	PLHLGKCNIAGWILG	66	80	15	1
Naive	nai_8	TRBV6-6	TRBJ1-1	CASSYQMRGTEAFF	SSFERFEIFPKTSSW	126	140	15	1
Naive	nai_9	TRBV14	not resolved	not resolved	NKGVTAACPHAGAKS	146	160	15	1
Naive	nai_10	TRBV20-1	TRBJ2-6	CSATQSCGANVQTF	AGAKSFYKNLIWLVK	156	170	15	1
Naive	nai_11	TRBV11-2	TRBJ2-1	CASSLDRGIEQFF	AGAKSFYKNLIWLVKKGNSY	156	175	20	1
Naive	nai_12	TRBV5-1	TRBJ2-3	CASSLENNRPGTMDTQYF	AGAKSFYKNLIWLVKKGNSY	156	175	20	1
Naive	nai_13	TRBV11-3	TRBJ2-1	CASSLSIWQGAYNEQFF	SRYSKKFKPEIAIRP	221	235	15	1
Naive	nai_14	TRBV20-1	TRBJ2-7	CSAPTGTSPYEQYF	SRYSKKFKPEIAIRPKVRDQ	221	240	20	1
Naive	nai_15	TRBV20-1	TRBJ2-1	CSARDPSGSVNEQLF	SRYSKKFKPEIAIRPKVRDQ	221	240	20	1
Naive	nai_16	TRBV12-3	TRBJ2-2	CASSYSGGGNTGELFF	KFKPEIAIRPKVRDQ	226	240	15	1
Naive	nai_17	TRBV5-1	TRBJ1-1	CASSLQGENTEAFF	EGRMNYYWTLVEPGD	241	255	15	1
Naive	nai_18	TRBV5-1	TRBJ1-1	CASRQGMNTEAFF	EGRMNYYWTLVEPGDKITFE	241	260	20	1
Naive	nai_19	TRBV20-1	TRBJ2-1	CSARDPSGSVNEQFF	ATGNLVVPRYAFAMERNAGS	261	280	20	1
Naive	nai_20	TRBV3-1	TRBJ2-2	CATKPGGTGELFF	VVPRYAFAMERNAGS	266	280	15	1
Naive	nai_21	TRBV10-3	TRBJ1-4	CAISEMATNEKLFF	VVPRYAFAMERNAGSGIIIS	266	285	20	2
Naive	nai_22	TRBV6-6	TRBJ1-3	CASSRQGGSGNTIYF	GAINTSLPFQNIHPITIGKC	301	320	20	2
Naive	nai_23	TRBV7-8	TRBJ1-1	not resolved	GAINTSLPFQNIHPITIGKC	301	320	20	1
Naive	nai_24	TRBV10-1	TRBJ2-1	CASSTPGVGNEQFF	SLPFQNIHPITIGKC	306	320	15	2
Naive	nai_25	TRBV7-8	TRBJ2-7	CASSLLTSGGNGQYF	EKMNTQFTAVGKEFNHLEKR	401	420	20	1
Naive	nai_26	TRBV11-2	TRBJ1-3	CASSYGAPTSDTIYF	GFLDIWTYNAELLVL	431	445	15	1
Naive	nai_27	TRBV11-2	TRBJ1-1	CASSNRVNTEAFF	GFLDIWTYNAELLVLLENER	431	450	20	1
Naive	nai_28	TRBV12-4	TRBJ2-2	not resolved	GFLDIWTYNAELLVLLENER	431	450	20	1
Naive	nai_29	TRBV5-1	TRBJ1-2	CASTPRYEGAKYGYTF	GFLDIWTYNAELLVLLENER	431	450	20	1
Naive	nai_30	TRBV6-5	TRBJ1-6	CASSYSGRGESPLHF	LENERTLDYHDSNVK	446	460	15	1
Naive	nai_31	TRBV6-5	TRBJ2-5	CASSYLGETQYF	EIGNGCFEFYHKCDN	476	490	15	1
Naive	nai_32	TRBV5-5	TRBJ1-5	CASSFLALGQPQHF	KLESTRIYQILAIYSTVASS	521	540	20	1
Naive	nai_33	TRBV20-1	TRBJ1-1	CSARKTGEVWTEAFF	TVASSLVLVVSLGAI	536	550	15	1
Naive	nai_34	TRBV14	TRBJ1-6	CASSQGTGSPLHF	TVASSLVLVVSLGAISFWMC	536	555	20	1
Naive	nai_35	TRBV20-1	TRBJ2-7	CSALTSGSGEQYF	TVASSLVLVVSLGAISFWMC	536	555	20	1
Naive	nai_36	TRBV6-2	TRBJ2-2	CASSSTSGRPGELFF	LVLVVSLGAISFWMC	541	555	15	1

Clonotypes (tot): 36   Clones (tot): 39

Supplementary Table 6.2-3. HLA typing of the four healthy donors included in this study.

	Gender	Age	HLA-DRB1*		HLA-DQA1*		HLA-DQB1*		HLA-DPA1*		HLA-DPB1*	
HD1	Male	50	01:01	08:01	01:01	04:01	04:02	05:01	01:03	02:01	04:02	13:01
HD2	Female	60	07:01		02:01	02:01	02:02	02:02	02:01	02:01	11:01	17:01
HD3	Male	19	03:01	04:08	03:03	05:01	02:01	03:01	01:03	02:02/ 02:06	04:01	05:01
HD4	Female	41	04:04	07:01	02:01	03:01	03:02	03:03	01:03	01:03	04:01	03:01/ 104:01

	HLA-A*		HLA-B*		HLA-C*	
HD1	02	11	35	51	04	14
HD2	02	29	44		16	
HD3	01:01	25:01	08:01	14:02	07:01	08:02
HD4	01:01	03:01	40:01	57:01	03:04	06:02

Supplementary Table 6.2-4. H1-HA peptides identified by MS-based MHC-II peptidomics. Start/end position and length of the peptide residues in H1-HA are indicated. + indicates identification of a peptide in an experiment replicate.

				Identification (donor-cell type-replicate) — HD1							
Peptide sequence	Start	End	Length	HD1-moDC-rep1	HD1-moDC-rep2	HD1-EBV-B(anti-head)-rep1	HD1-EBV-B(anti-head)-rep2	HD1-EBV-B(anti-head)-rep3	HD1-EBV-B(anti-stem)-rep1	HD1-EBV-B(anti-stem)-rep2	HD1-EBV-B(anti-stem)-rep3
YPGDFIDYEELREQ	108	121	14					+			
IFPKTSSWPNHDSN	133	146	14		+						
IFPKTSSWPNHDSNK	133	147	15	+	+						
IFPKTSSWPNHDSNKG	133	148	16		+						
FPKTSSWPNHDSNKG	134	148	15		+						
YKNLIWLVKKGNSYPK	162	177	16			+		+			
KNLIWLVKKGNSYPK	163	177	15			+		+			
LSKSYINDKGKEV	178	190	13								
KSYINDKGKEV	180	190	11								
KSYINDKGKEVL	180	191	12								
MNYYWTLVEPGDK	244	256	13			+	+	+	+		
MNYYWTLVEPGDKIT	244	258	15			+			+		
YWTLVEPGDK	247	256	10								
VEPGDKITFEATGNLVVPRYA	251	271	21				+				
EPGDKITFEATGNLVVPRY	252	270	19				+				
EPGDKITFEATGNLVVPRYA	252	271	20				+	+			
VVPRYAFAMERNAGSG	266	281	16		+						
VPRYAFAMERNAGSG	267	281	15		+						
LPFQNIHPITIG	307	318	12								
STKLRLATGLRNIPSI	326	341	16								
STKLRLATGLRNIPSIQ	326	342	17								
STKLRLATGLRNIPSIQSRG	326	345	20								
LRLATGLRNIPSIQ	329	342	14								
LATGLRNIPSIQSR	331	344	14								
LATGLRNIPSIQSRG	331	345	15								
LATGLRNIPSIQSRGL	331	346	16								
LATGLRNIPSIQSRGLF	331	347	17								
LATGLRNIPSIQSRGLFG	331	348	18								
ATGLRNIPSIQSR	332	344	13								
ATGLRNIPSIQSRG	332	345	14								
ATGLRNIPSIQSRGL	332	346	15								
ATGLRNIPSIQSRGLF	332	347	16								
ATGLRNIPSIQSRGLFG	332	348	17								
TGLRNIPSIQSRG	333	345	13								
TGLRNIPSIQSRGL	333	346	14								
TGLRNIPSIQSRGLF	333	347	15								
TGLRNIPSIQSRGLFG	333	348	16								
AIDEITNKVNSVIE	388	401	14		+			+		+	
AIDEITNKVNSVIEK	388	402	15		+			+			

Sequence	Start	End	Length								
IDEITNKVNSVIE	389	401	13			+	+	+			
IDEITNKVNSVIEK	389	402	14	+	+	+	+	+	+	+	+
DEITNKVNSVIE	390	401	12			+	+	+		+	+
DEITNKVNSVIEK	390	402	13	+		+		+		+	+
EITNKVNSVIE	391	401	11			+	+			+	
EITNKVNSVIEK	391	402	12	+		+	+				
ITNKVNSVIE	392	401	10			+	+	+		+	+
ITNKVNSVIEK	392	402	11			+	+	+		+	
VNSVIEKMNTQFTAVG	396	411	16								
VNSVIEKMNTQFTAVGK	396	412	17								
SVIEKMNTQFTAVG	398	411	14								
MNTQFTAVGKEFNH	403	416	14	+	+	+	+	+			
NTQFTAVGKEFNH	404	416	13	+	+	+	+	+			
FTAVGKEFNH	407	416	10		+		+				
VGKEFNHLEKRIE	410	422	13			+	+				
VGKEFNHLEKRIEN	410	423	14		+	+	+				
VGKEFNHLEKRIENLN	410	425	16			+	+				
GKEFNHLEKRIEN	411	423	13		+						
KEFNHLEKRIEN	412	423	12		+		+				
KEFNHLEKRIENLN	412	425	14		+		+				
LEKRIENLNKKVD	417	429	13		+		+				
LEKRIENLNKKVDD	417	430	14		+	+	+				
LEKRIENLNKKVDDG	417	431	15	+	+	+	+				
LEKRIENLNKKVDDGF	417	432	16		+	+	+				
LEKRIENLNKKVDDGFL	417	433	17		+	+	+				
EKRIENLNKKVDD	418	430	13		+		+				
EKRIENLNKKVDDG	418	431	14	+	+	+	+				
EKRIENLNKKVDDGF	418	432	15		+	+	+	+	+		
EKRIENLNKKVDDGFL	418	433	16	+	+	+	+	+			
KRIENLNKKVD	419	429	11		+		+	+			
KRIENLNKKVDD	419	430	12	+	+		+				
KRIENLNKKVDDG	419	431	13	+	+	+	+				
KRIENLNKKVDDGF	419	432	14	+		+	+				
KRIENLNKKVDDGFL	419	433	15	+	+	+	+				
RIENLNKKVDD	420	430	11	+	+	+	+				
RIENLNKKVDDG	420	431	12	+	+	+	+				
RIENLNKKVDDGF	420	432	13	+							
EEAKLNREEIDGVKLE	508	523	16								

117

Peptide sequence	Start	End	Length	HD2			HD3		HD4			
				HD2-EBV-B-rep1	HD2-EBV-B-rep2	HD2-EBV-B-rep3	HD3-EBV-B-rep1	HD3-EBV-B-rep2	HD4-EBV-B-rep1	HD4-EBV-B-rep2	HD4-EBV-B-rep3	HD4-EBV-B-rep4
YPGDFIDYEELREQ	108	121	14									
IFPKTSSWPNHDSN	133	146	14									
IFPKTSSWPNHDSNK	133	147	15									
IFPKTSSWPNHDSNKG	133	148	16									
FPKTSSWPNHDSNKG	134	148	15									
YKNLIWLVKKGNSYPK	162	177	16									
KNLIWLVKKGNSYPK	163	177	15									
LSKSYINDKGKGKEV	178	190	13					+				
KSYINDKGKEV	180	190	11				+					
KSYINDKGKEVL	180	191	12					+				
MNYYWTLVEPGDK	244	256	13									
MNYYWTLVEPGDKIT	244	258	15									
YWTLVEPGDK	247	256	10									
VEPGDKITFEATGNLVVPRYA	251	271	21	+	+							
EPGDKITFEATGNLVVPRY	252	270	19	+	+							
EPGDKITFEATGNLVVPRYA	252	271	20	+	+	+				+		+
VVPRYAFAMERNAGSG	266	281	16					+		+		
VPRYAFAMERNAGSG	267	281	15	+	+		+					
LPFQNIHPITIG	307	318	12	+	+				+			
STKLRLATGLRNIPSI	326	341	16	+								
STKLRLATGLRNIPSIQ	326	342	17	+	+	+		+				
STKLRLATGLRNIPSIQSRG	326	345	20	+								
LRLATGLRNIPSIQ	329	342	14			+						
LATGLRNIPSIQSR	331	344	14						+			
LATGLRNIPSIQSRG	331	345	15						+	+		+
LATGLRNIPSIQSRGL	331	346	16						+	+	+	+
LATGLRNIPSIQSRGLF	331	347	17						+	+	+	+
LATGLRNIPSIQSRGLFG	331	348	18						+	+	+	+
ATGLRNIPSIQSR	332	344	13				+		+	+		
ATGLRNIPSIQSRG	332	345	14					+	+	+	+	+
ATGLRNIPSIQSRGL	332	346	15						+	+	+	+
ATGLRNIPSIQSRGLF	332	347	16						+	+	+	+
ATGLRNIPSIQSRGLFG	332	348	17						+	+	+	+
TGLRNIPSIQSRG	333	345	13						+	+	+	+
TGLRNIPSIQSRGL	333	346	14									
TGLRNIPSIQSRGLF	333	347	15						+	+	+	+
TGLRNIPSIQSRGLFG	333	348	16						+	+	+	+
AIDEITNKVNSVIE	388	401	14									
AIDEITNKVNSVIEK	388	402	15									
IDEITNKVNSVIE	389	401	13									
IDEITNKVNSVIEK	389	402	14									
DEITNKVNSVIE	390	401	12									
DEITNKVNSVIEK	390	402	13									
EITNKVNSVIE	391	401	11									
EITNKVNSVIEK	391	402	12									
ITNKVNSVIE	392	401	10									

Peptide	Start	End	Length					
ITNKVNSVIEK	392	402	11					
VNSVIEKMNTQFTAVG	396	411	16		+	+	+	+
VNSVIEKMNTQFTAVGK	396	412	17		+	+	+	
SVIEKMNTQFTAVG	398	411	14		+	+	+	+
MNTQFTAVGKEFNH	403	416	14					
NTQFTAVGKEFNH	404	416	13					
FTAVGKEFNH	407	416	10					
VGKEFNHLEKRIE	410	422	13					
VGKEFNHLEKRIEN	410	423	14					
VGKEFNHLEKRIENLN	410	425	16					
GKEFNHLEKRIEN	411	423	13					
KEFNHLEKRIEN	412	423	12					
KEFNHLEKRIENLN	412	425	14					
LEKRIENLNKKVD	417	429	13					
LEKRIENLNKKVDD	417	430	14					
LEKRIENLNKKVDDG	417	431	15					
LEKRIENLNKKVDDGF	417	432	16					
LEKRIENLNKKVDDGFL	417	433	17					
EKRIENLNKKVDD	418	430	13					
EKRIENLNKKVDDG	418	431	14					
EKRIENLNKKVDDGF	418	432	15					
EKRIENLNKKVDDGFL	418	433	16					
KRIENLNKKVD	419	429	11					
KRIENLNKKVDD	419	430	12					
KRIENLNKKVDDG	419	431	13					
KRIENLNKKVDDGF	419	432	14					
KRIENLNKKVDDGFL	419	433	15					
RIENLNKKVDD	420	430	11					
RIENLNKKVDDG	420	431	12					
RIENLNKKVDDGF	420	432	13					
EEAKLNREEIDGVKLE	508	523	16	+				

119

6.2.7   Methods

Cell purification and sorting. Serial blood samples from healthy donor HD1 vaccinated with Inflexal V 2013/14 (Crucell) was obtained and used in compliance with the Federal Office of Public Health (authorization no. A000197/2 to F.S.). Blood from healthy donors HD2,HD3 and HD4 vaccinated with Fluarix Tetra 2015/16 (GlaxoSmithKline) was obtained from the Surrey Clinical Research Centre (University of Surrey, UK). All participants provided written informed consent for participation in the study. Peripheral Blood Mononuclear Cells (PBMCs) were isolated with Ficoll-Paque Plus (GE Healthcare). Monocytes were isolated from PBMCs by positive selection using CD14 magnetic microbeads (Miltenyi Biotech). CD14-depleted fractions were stained at 37°C for 15 min with a primary anti-human-CXCR5 (clone 51505; cat. no. MAB190) from Bio-Techne, followed by staining with a biotinylated secondary goat anti-mouse IgG$_{2B}$ (cat. no. 1090-08) from Southern Biotech. After washing, cells were stained on ice for 30 min with PE/Cy7-Streptavidin (cat. no. 405206) from BioLegend, and with the following fluorochrome-labeled mouse monoclonal antibodies: CD8-PE–Cy5 (clone B9.11; cat. no. A07758), CD25-PE–Cy5 (clone B1.49.9; cat. no. IM2646) from Beckman Coulter, CD22-FITC (clone HIB22; cat. no. 555424) from BD Biosciences, CD4–PE–Texas Red (clone S3.5; cat. no. MHCD0417), CD45RA-Qdot 655 (clone MEM-56; cat. no. Q10069), CD95-PerCP-eFluor 710 (clone DX2; cat. no. 46-0959-42) from ThermoFisher Scientific, CCR7–BV421 (clone G043H7; cat. no. 353208) from BioLegend, Alexa Fluor 647-conjugated goat anti-human IgG (cat. no. 109-606-170) from Jackson ImmunoResearch. Naïve and memory CD4$^+$ T cells were sorted to over 98% purity on a FACSAria III (BD) after exclusion of CD8$^+$, CD22$^+$ and CD25^{bright} cells. Naïve T cells were sorted as CD4$^+$CD45RA$^+$CCR7$^+$CD95$^-$; the remaining CD4$^+$ T cells were sorted as total memory cells. In some experiments with donor HD1, total memory CD4$^+$ T cells were divided in cTfh (sorted as CXCR5$^+$ cells), Tcm (sorted as CCR7$^+$CXCR5$^-$ cells) and Tem (sorted as CCR7$^-$CXCR5$^-$ cells). IgG$^+$ memory B cells and IgG$^-$ B cells were sorted to over 98% purity after gating on CD22$^+$CD4$^-$CD8$^-$CD25$^-$ cells.

Antigens and peptides. Peptides were synthesized as crude material on a small scale (1 mg) by A&A Labs (San Diego). Peptides used in the study included 15mers

overlapping of 10 (112 peptides) or 20mers overlapping of 10 (56 peptides) covering the entire sequence of A/California/07/2009 (H1N1) hemagglutinin (H1-HA). The numbering of HA T cell epitopes reported in the text refers to the 566-amino acid long sequence of A/California/07/2009 HA (UniProtKB: A0A075EXW1). Recombinant A/California/07/2009 (H1-HA) and A/Vietnam/1203/2004 (H5-HA) hemagglutinins were purchased from Protein Sciences Corporation.

Cell culture. T cells were cultured in RPMI 1640 medium supplemented with 2 mM glutamine, 1% (vol/vol) nonessential amino acids, 1% (vol/vol) sodium pyruvate, penicillin (50 U/ml), streptomycin (50 µg/ml) (all from Invitrogen) and 5% human serum (Swiss Red Cross). For some experiments, medium was supplemented with IL-2 (500 IU/ml). B cells were cultured in RPMI 1640 medium supplemented with 2 mM glutamine, 1% (v/v) nonessential amino acids, 1% (v/v) sodium pyruvate, penicillin (50 U/ml), streptomycin (50 µg/ml) (all from Invitrogen) and 10% fetal bovine serum (HyClone, characterized, GE Healthcare Life Science). Sorted IgG$^+$ memory B cells were immortalized with Epstein–Barr virus (EBV) and plated in single-cell cultures in the presence of CpG-DNA (2.5 µg/ml) and irradiated PBMC-feeder cells, as previously described (Traggiai et al., 2004). Two weeks post immortalization, the culture supernatants were screened by high throughput ELISA for binding to H1-HA or H5-HA as described (Pappas et al., 2014). EBV-B cell clones that resulted positive for binding to H1-HA and/or H5-HA were isolated and expanded. IgG$^-$ B cells to be used as antigen-presenting cells for T cell libraries were expanded with CD40L according to an established protocol (Zand et al., 2005). Autologous monocyte-derived dendritic cells (moDCs) were generated by culture in complete medium containing 10% fetal bovine serum (HyClone) supplemented with recombinant GM-CSF (Gentaur) and IL-4 (ImmunoTools), as previously described (Sallusto and Lanzavecchia, 1994).

T cell library. Sorted naïve or memory CD4$^+$ T cells were polyclonally stimulated with 1 µg/ml PHA (Remel) in the presence of irradiated (45Gy) allogeneic feeder cells (5×10^4 per well) and IL-2 (500 IU/ml) in a 96-well plate. The size of the library (number of wells and initial input of cells seeded per well) depended on the number of cells isolated from each donor. T cell lines were expanded as previously described (Geiger et al., 2009). Library screening was performed 14-21 days after initial stimulation, by culturing thoroughly washed T cells (2.5×10^5 per well) with autologous

irradiated B cells (2.5×10^4 per well), untreated or pulsed with a pool of HA overlapping peptides (2 µM per peptide) composed of 15mers (112 peptides, 15mers overlapping of 10) and 20mers (56 peptides, 20mers overlapping of 10) covering the H1-HA sequence. Proliferation was assessed on day 4, after incubation for 16 h with 1 µCi/ml [methyl-^{3}H]-thymidine (Perkin Elmer). Data were expressed as counts per min (cpm). Stringent criteria were used to score positive T cell lines based on a cut-off value of (i) a ΔCpm value $\geq 3 \times 10^3$ (Cpm with antigen and APCs – Cpm with APCs only) and (ii) a stimulation index (SI) ≥ 3 (Cpm with antigen and APCs / Cpm with APCs only). This threshold was chosen based upon previous observations made across multiple negative and positive samples assessed by the T cell library technique and with a variety of donors and antigens (Campion et al., 2014; Geiger et al., 2009; Latorre et al., 2018; Lindestam Arlehamn et al., 2013). The specificity of positive cultures was confirmed in subsequent independent experiments of epitope mapping. Precursor frequencies were calculated based on numbers of negative wells, assuming a Poisson distribution (Geiger et al., 2009), and are expressed per million cells within each subset. Inverse Simpson index of diversity (1-D) was calculated for each donor by considering the number of individual T cell clones recognizing each particular H1-HA peptide (Simpson, 1949). Inverse Simpson index (1-D) quantify the richness and evenness of populations, it ranges between 0 and 1 and can be interpreted as the probability that two H1-HA-specific T cell clones randomly selected from a repertoire will recognize different epitopes.

Isolation of H1-HA-specific T cell clones. Sorted memory CD4[+] T cell subsets from donor HD1 were labelled with carboxyfluorescein succinimidyl ester (CFSE) and cultured at a ratio of 2:1 with irradiated autologous monocytes untreated or pulsed with Inflexal V 2013/2014 (3 µg/ml). After 6 days, cells were stained with antibodies to CD25-PE (clone M-A251; cat. no. 555432) from BD Biosciences and ICOS-APC (clone C398.4A; cat. no. 313510) from BioLegend. Proliferating activated T cells were FACS-sorted as CFSE[low]CD25[+]ICOS[+] and expanded *in vitro* in the presence of IL-2 (500 IU/ml). To select HA-specific T cells, Inflexal V-reactive CFSE[low] cultures were relabeled with CFSE and stimulated with irradiated autologous monocytes untreated or pulsed with recombinant H1-HA (5 µg/ml). After 5 days, proliferating activated T cells were sorted as CFSE[low]CD25[+]ICOS[+] and cloned by limiting dilution. In some

experiments, positive cultures from T cell libraries were labelled with CFSE and cultured at a ratio of 2:1 with irradiated autologous monocytes untreated or pulsed with H1-HA peptide pool (2 μM per peptide). After 5 days, proliferating activated T cells were sorted as CFSE^{low}CD25$^+$ICOS$^+$ and cloned by limiting dilution. T cell clone reactivity was determined by stimulation with H1-HA peptide pool (2 μM per peptide) or recombinant H1-HA (5 μg/ml) in the presence of irradiated autologous monocytes or B cells as APCs. In some experiments, H1-HA peptides or recombinant H1-HA were titrated by serial dilution. Epitope mapping was performed by stimulation of T cell clones with irradiated autologous EBV-immortalized B-cell (EBV-B) clones, untreated or pre-pulsed for 2-3 h with individual peptides (15mers overlapping of 10 or 20mers overlapping of 10) covering the entire sequence of H1-HA (2 μM per peptide). To determine MHC restriction, T cell clones were stimulated with autologous APCs pulsed with recombinant H1-HA, in the absence or presence of blocking anti-MHC-II monoclonal antibodies produced in house from hybridoma cell lines (anti-HLA-DR, clone L243 from ATCC, cat. no. HB-55; anti-HLA-DQ, clone SPVL3 (Spits et al., 1983); anti-HLA-DP, clone B7/21 (Watson et al., 1983)). In all experiments proliferation was assessed on day 3, after incubation for 16 h with 1 μCi/ml [methyl-^{3}H]-thymidine (Perkin Elmer). Data were expressed as counts per min (Cpm).

TCR Vβ deep sequencing. *Ex vivo*-sorted memory CD4$^+$ T cell subsets and CFSE^{low} fractions of Inflexal V-stimulated memory CD4$^+$ T cell subsets from donor HD1 were analyzed by deep sequencing. In brief, 2.5-5×10^5 T cells were centrifuged and washed in PBS, and genomic DNA was extracted from the pellet using QIAamp DNA Micro Kit (Qiagen), according to manufacturer's instructions. Genomic DNA quantity and purity were assessed through spectrophotometric analysis. Sequencing of TCR Vβ CDR3 was performed by Adaptive Biotechnologies using the ImmunoSEQ platform (http://www.immunoseq.com). In brief, following multiplex PCR reaction designed to target any CDR3 Vβ fragments; amplicons were sequenced using the Illumina HiSeq platform. Raw data consisting of all retrieved sequences of 87 nucleotides or corresponding amino acid sequences and containing the CDR3 region were exported and further processed. The assay was performed at deep level for *ex vivo*-sorted total memory CD4$^+$ (detection sensitivity, 1 cell in 200,000) and at survey level for CFSE^{low} Inflexal V-reactive cultures (detection sensitivity, 1 cell in 40,000). Each clonotype was

defined as a unique productively rearranged TCR Vβ nucleotide sequence; data processing was done using the productive frequency of reads provided by ImmunoSEQ Analyzer V.3.0 (http://www.immunoseq.com).

Sequence analysis of TCR Vβ genes. Sequence analysis of rearranged TCR Vβ genes of HA-specific T cell clones from donor HD1 was performed as previously described (Latorre et al., 2018). Briefly, cDNA from individual T cell clones was obtained by reverse transcription of total RNA from 10^3-10^4 cells per reaction. Rearranged TCR Vβ genes were PCR amplified using forward primer pool targeting Vβ genes, and reverse primer pairing to C1-C2 β-chain constant region. Sequence amplification was assessed through agarose gel electrophoresis; successfully amplified fragments were sequenced by Sanger method, and TCR sequence annotation was carried out by using IMGT/V-QUEST algorithm (Lefranc et al., 2009).

HLA typing and peptide-MHC-II binding affinity measurement. HLA genotype of the patients was determined by reverse sequence-specific oligonucleotides probes (revPCR-SSO) DNA typing (LABType, One Lambda Inc.) performed at the IRCCS San Matteo Hospital Foundation (Pavia, Italy). Affinity measurements of H1-HA 15mer peptides recognized by HLA-DR-restricted T cell clones from donor HD1 to recombinant *HLA-DRB1*01:01* or *HLA-DRB1*08:01* alleles was performed by Immunitrack (Copenhagen, Denmark), as previously described (Justesen et al., 2009). Briefly, recombinant HLA-DRB1 isoforms were refolded *in vitro* in the presence of recombinant HLA-DRA and increasing concentration of H1-HA 15mer peptides. Titrated pan-HLA-DR-binding epitope (PADRE) was used as positive control. After 24 hrs incubation at room temperature and pH 7, correctly folded heterotrimeric pMHC-II complexes were detected by ELISA; data were analyzed using GraphPad Prism 8 software.

Purification of MHC-II presented peptides. MoDCs generated from donor HD1 were pulsed 2 hrs with 10 µg/ml recombinant H1-HA at a cellular density of 3×10^6 cells/ml, and matured overnight with 100 ng/ml LPS (Enzo Life Sciences) at a cellular density of 1×10^6 cells/ml. HA-specific EBV-B cell clones isolated from IgG$^+$ memory B cells of each of the four donors were pulsed overnight with 200 ng/ml recombinant H1-HA at a cellular density of 5×10^6 cells/ml. MHC-II complexes were purified from about 3×10^7

HA-pulsed moDCs or 10^9 HA-pulsed EBV-B cells with a protocol adapted from Bassani-Sternberg M. et al (Bassani-Sternberg et al., 2015). Briefly, the B cells were lysed with 0.25% sodium deoxycholate, 1% octyl-β-D glucopyranoside (Sigma), 0.2 mM iodoacetamide, 1 mM EDTA, and Complete Protease Inhibitor Cocktail (Roche) in PBS at 4°C for 1 h. The lysates were cleared by 20 min centrifugation at 18,000 g at 4°C, and MHC-II complexes were purified by immunoaffinity chromatography with the anti-HLA-DR/DP/DQ HB-145 monoclonal antibody produced in house from hybridoma cell line IVA12 (ATCC, cat. no. HB-145) and covalently bound to Protein-A Sepharose beads (Thermo Fisher Scientific). The cleared lysates were loaded 3 times into the affinity columns at 4°C, and subsequently washed at 4°C with 10 column volumes of 150 mM NaCl, 20 mM Tris•HCl, pH 8 (buffer A); 10 column volumes of 400 mM NaCl, 20 mM Tris•HCl, pH 8; 10 column volumes of buffer A; and finally 10 column volumes of 20 mM Tris•HCl, pH 8. The MHC-II complexes were eluted at room temperature by addition of 500 µl of 0.1 M acetic acid, in total five elutions for each sample. Small aliquots of each eluted fraction were analyzed by 12% SDS-PAGE to evaluate yield and purity of MHC-II complexes. Sep-Pak tC18 (Waters, Milford, MA) cartridges were used for further separation of peptides from MHC-II subunits. The cartridges were prewashed with 80% acetonitrile (AcN) in 0.5% formic acid, followed by 0.2% trifluoroacetic acid (TFA), and subsequently loaded 3 times with each fraction eluted from the immunoaffinity column. After loading, the cartridges were washed with 0.2% TFA, and the peptides were separated from the more hydrophobic MHC-II chains by elution with 30% AcN in 0.2% TFA. The peptides were further purified using a Silica C18 column tip (Harvard Apparatus, Holliston, MA) and eluted again with 30% AcN in 0.2% TFA. Finally, the peptides were concentrated by vacuum centrifugation, and resuspended in 2% AcN, 0.1% TFA, 0.5% formic acid for MS analysis.

Liquid chromatography-mass spectrometry (LC-MS/MS) and data analysis. MHC-II peptides were separated on an EASY-nLC 1200 HPLC system coupled online to a Q Exactive mass HF spectrometer via a nanoelectrospray source (Thermo Fisher Scientific). Peptides were loaded in buffer A (0.1% formic acid) on in-house packed columns (75 µm inner diameter, 50 cm length, and 1.9 µm C18 particles from Dr. Maisch GmbH) and eluted with a non-linear 120 min gradient of 5% – 60% buffer B (80% ACN, 0.1% formic acid) at a flow rate of 250 nl/min and a column temperature of

50°C. The Q Exactive was operated in a data dependent mode with a survey scan range of 300 – 1650 m/z and a resolution of 60,000 at m/z 200. Up to 10 most abundant isotope patterns with a charge ≥ 1 were isolated with a 1.8 Th wide isolation window and subjected to higher-energy C-trap dissociation (HCD) fragmentation at a normalized collision energy of 27. Fragmentation spectra were acquired with a resolution of 15,000 at m/z 200. Dynamic exclusion of sequenced peptides was set to 30 s to reduce the number of repeated sequences. Thresholds for the ion injection time and ion target values were set to 80 ms and 3E6 for the survey scans and 120 ms and 1E5 for the MS/MS scans, respectively. Data were acquired using the Xcalibur software (Thermo Scientific). MaxQuant software was used to analyze mass spectrometry raw files. MS/MS spectra were searched against the A/California/07/2009 (H1N1) HA sequence (UniProtKB: A0A075EXW1), the bovine Uniprot FASTA database, the human Uniprot FASTA database, and a common contaminants database (247 entries) by the Andromeda search engine (Cox et al., 2011). N-terminal acetylation and methionine oxidation were set as variable modifications; no fixed modifications were selected; the enzyme specificity was set to "unspecific" with a minimum peptide length of 8 amino acids. A false discovery rate (FDR) of 1% was required for peptides. Peptide identification was performed with an allowed precursor mass deviation of up to 4.5 ppm and an allowed fragment mass deviation of 20 ppm; "match between runs" option was disabled.

In silico analysis. MHC-II binding affinity of each theoretical H1-HA-derived 15mer peptide was calculated using IEDB tool for MHC-II binding prediction (http://tools.iedb.org/mhcii/, (Paul et al., 2015)). Donor-tailored analyses were performed using IEDB recommended method and considering the set of MHC-II alleles carried by each donor at the following loci: HLA-DRB1, HLA-DRB3/4/5 (if associated), HLA-DQA1/DQB1 in cis- or trans-pairing, HLA-DPA1/DPB1 in cis- or trans-pairing. Top scoring H1-HA 15mer peptides for each donor were selected based on percentile rank calculated by comparison to a large set of random natural peptides. Epitope prediction based on structural accessibility was performed as previously described (Mettu et al., 2016). Briefly, an aggregate z-score of conformational stability was determined for each H1-HA residue by integrating four structural parameters obtained from the 3D structure of post-fusion HA resolved by X-ray diffraction (PDB

codes: 3LZG for HA1 domain (Xu et al., 2010), 1HTM for HA2 domain in the post-fusion conformation (Bullough et al., 1994)). The z-score statistic was then used to calculate an epitope likelihood for each theoretical H1-HA 15mer peptide, following the rationale that the liberation of antigenic peptides might be facilitated by surrounding unstable regions that are readily unfolded and targeted by endosomal proteases. For the optimization of combined predictors, peptide-MHC-II binding affinity was calculated using NetMHCIIpan 3.2 (Jensen et al., 2018b). In particular, we systematically performed peptide binding affinity predictions of each theoretical H1-HA 15mer peptide for the MHC-II alleles carried by each donor, considering for each peptide the best scoring affinity within each group of MHC-II alleles. Combined predictors for each donor were then built by iteratively weighting the contributions of epitope likelihood based on structural accessibility and peptide binding affinity to MHC-II, until we could maximize the Area under the Receiver Operating Characteristic Curve (AUROC) of the predictor using as reference the set of epitopes recognized by memory T cells of each donor.

Statistical analysis. Statistical analyses were performed using GraphPad Prism 8 software or R software version 3.5.1. EC50 (ng/ml) and Kd (ng/ml) values were calculated by nonlinear regression curve fit (4PL with automatic outlier elimination) using GraphPad Prism 8 software. Significance was assigned at P value < 0.05, unless stated otherwise. Specific tests are indicated in the figure legends for each comparison.

6.3 Broadly reactive CD4[+] T cells against Enterobacteriaceae are present in the naïve and enriched in the memory human immune repertoires

A manuscript by:

Antonino Cassotta[1,2*], Jérémie Goldstein[1,8*], Greta Durini[1], David Jarrossay[1], Franca Baggi Menozzi[3], Mario Venditti[4], Alessandro Russo[5], Marco Falcone[6], Antonio Lanzavecchia[1], Maria Cristina Gagliardi[7], Daniela Latorre[1,2#], and Federica Sallusto[1,2#].

[1] Institute for Research in Biomedicine, Università della Svizzera italiana, Bellinzona, Switzerland.

[2] Institute of Microbiology, ETH Zurich, Switzerland.

[3] Servizio di Microbiologia EOLAB, Ente Ospedaliero Cantonale, Bellinzona, Switzerland.

[4] Department of Public Health and Infectious Diseases, Sapienza University of Rome, Rome, Italy.

[5] Infectious Diseases Clinic, Department of Medicine, University of Udine, Udine, Italy.

[6] Department of Clinical and Experimental Medicine, University of Pisa, Italy.

[7] Center for Gender Specific Medicine, Istituto Superiore di Sanità, Rome, Italy.

[8] Present address: Department of Pathology-Immunology, Faculty of Medicine, University of Geneva, Switzerland.

[*,#] These authors contributed equally to this work.

6.3.1 Author contributions

A.C. characterized the T cell response in healthy donors and septic patients, isolated T cell clones, performed bioinformatics analyses of TCR Vβ deep sequencing, analyzed the data and wrote the manuscript; J.G. characterized the T cell response in healthy donors, isolated T cell clones, analyzed the data and wrote the manuscript; G.D. provided assistance for the characterization of the T cell response in healthy donors; D.J. performed cell sorting; F.B.M. provided microbes and assistance for antigen preparation; M.V., A.R. and M.F. collected biological samples; A.L. provided supervision and wrote the manuscript; M.C.G. provided supervision and designed the experiments; D.L. designed the experiments, characterized the T cell response in healthy donors and septic patients, isolated T cell clones, analyzed the data and wrote the manuscript; F.S. provided overall supervision, designed the experiments and wrote the manuscript.

6.3.2 Abstract

Enterobacteriaceae are a large family of Gram-negative bacteria that includes both commensals and mutualistic or opportunistic pathogens. The latter can cause severe nosocomial infections, with outbreaks of multi-antibiotics resistant strains, thus being a major public health threat. A better understanding of immune defense mechanisms against Enterobacteriaceae and the identification of relevant microbial antigens may yield new strategies for prevention and treatment of severe infections. In this study, we combined several experimental approaches to analyze the CD4$^+$ T helper (Th) cell response to Enterobacteriaceae in the blood of healthy donors and of patients with sepsis. This approach led to the isolation of Enterobacteriaceae-reactive memory Th cells that were highly enriched in the CCR6$^+$CXCR3$^+$ Th1*/17 cell subset, expressed gut homing markers and produced IFN-γ, IL-17A and IL-22. This T cell subset was severely reduced in septic patients with *K. pneumoniae* bloodstream infection who also selective lacked circulating *K. pneumonie*-reactive T cells. By combining heterologous antigenic stimulation, single cell cloning and TCR Vβ sequencing, we found that a large fraction of memory Th cell clones was broadly cross-reactive to several Enterobacteriaceae species. The cross-reactive Th cell clones were expanded *in vivo* and a large fraction of them recognized the conserved outer membrane protein A (OmpA) antigen. Th cell clone cross-reactivity to different Enterobacteriaceae species was also detected in the naïve CD4$^+$ T cell compartment, although to a lesser extent than in the memory compartment. These data point to the existence of T cell determinants conserved among different Enterobacteriaceae species and targeted by broadly cross-reactive T cells, which are present in the pre-immune repertoire and become dominant in the memory repertoire.

6.3.3 Introduction

Enterobacteriaceae are a vast family of Gram-negative bacteria that includes harmless symbionts and potentially harmful pathogens (McSorley, 2014; Zeng et al., 2017). Among pathogenic Enterobacteriaceae, thyphoidal *Salmonella enterica* serovars, such as *S. typhi* and *S. paratyphi*, are responsible for life-threatening enteric fever with more than 20 million people affected worldwide annually (Bhan et al., 2005; McSorley, 2014; Napolitani et al., 2018), whereas non-typhoidal *S. enterica* serovars, such as *S. typhimurium* and *S. enteriditis,* are common causes of gastroenteritis due to food poisoning and can induce serious systemic infections in immunocompromised individuals (Feasey et al., 2012; Godinez et al., 2011). *Shigella* species, such as *S. flexneri*, *S. sonnei* and *S. dysenteriae*, can cause a highly contagious and severe inflammatory diarrhea, which affects around 164 million people per year, predominantly children under the age of 5 years (Kotloff et al., 1999), while *Klebsiella pneumoniae*, which is part of the normal microbiome colonizing several mucosal sites (Atarashi et al., 2017; Calfee, 2017), can become, under certain conditions (i.e. immunocompromised patients), an opportunistic pathogen and cause healthcare-associated infections, including bloodstream, catheter-associated urinary tract or respiratory tract infections (Calfee, 2017; Weiner et al., 2016). Given the lack of effective vaccines to prevent Enterobacteriaceae infections (Mani et al., 2016; Milligan et al., 2018) as well as the constant increase of multidrug resistant strains (Nordmann et al., 2011; Sheu et al., 2019; van Duin et al., 2013; Weiner et al., 2016), Enterobacteriaceae can cause severe deadly infections, thus representing a global health threat. Recently, carbapenem-resistant Enterobacteriaceae bacteria have been included in the list of antibiotic-resistant "priority pathogens" by the WHO.

Increasing evidence suggests that CD4[+] T cells can recognize commensals-derived antigens under homeostatic conditions and during gut dysbiosis or inflammation (Belkaid and Harrison, 2017; Sorini et al., 2018). Although many studies have investigated the human T cell response to commensals and pathogens (Duchmann et al., 1999; Duhen and Campbell, 2014; Hegazy et al., 2017; Napolitani et al., 2018; Sheikh et al., 2011), studies of Enterobacteriaceae-reactive CD4[+] T cells and their characterization in terms of phenotype, antigen specificity and TCR repertoire composition in physiological and pathological conditions remain elusive. Here, we

combined several experimental approaches, including heterologous stimulation of bacteria-specific memory T cells, high-throughput TCR-Vβ sequencing and antigen-specific priming of naïve T cells, to perform a systematic characterization of the CD4$^+$ Th cell response to multiple Enterobacteriaceae species in heathy individuals as well as in patients suffering from *K. pneumoniae* infections. We report that the CXCR3$^+$CCR6$^+$ Th1*/17 subset contains most Enterobacteriaceae-reactive memory CD4$^+$ T cells and that this subset is severely reduced and *K. pneumonie*-reactive T cells are absent in septic patients with *K. pneumoniae* bloodstream infection. Interestingly, most Enterobacteriaceae-reactive memory CD4$^+$ T cells cross-react to several Enterobacteriaceae species and some target OmpA, which is a highly conserved antigen among Enterobacteriaceae species. Importantly, broadly cross-reactive CD4$^+$ T cells are already present in the naïve repertoire and become highly expanded in the memory repertoire. Collectively, these data provide a better understanding of the human immune responses against Enterobacteriaceae and provide insights for the development of new strategies for prevention of severe Enterobacteriaceae-infections.

6.3.4   Results

Enterobacteriaceae-reactive memory CD4[+] T cells are present in the blood of healthy donors and have a Th1*/17 phenotype

To investigate the human CD4[+] T cell response to Enterobacteriaceae, we isolated by FACS-sorting memory CD4[+] Th cells from peripheral blood mononuclear cells (PBMCs) of a large number of healthy donors. Cells were labeled with carboxyfluorescein succinimidyl ester (CFSE) and stimulated *in vitro* with autologous monocytes either untreated or pulsed with different bacteria (**Fig. 6.3-1A, B**). In most of the donors tested, a clear proliferative T cell response to the Enterobacteriaceae species *E. coli, K. pneumoniae, E. aerogenes, Shigella* and *S. typhimurium* was detected, as shown by the CFSE profiles, although the magnitude of the response was significantly lower (p value < 0.0001) compared to the response elicited by *S. aureus* or *M. tuberculosis* in control cultures (**Fig. 6.3-1B**). Enterobacteriaceae-induced T cell proliferation was inhibited by anti-HLA-DP, -DQ and -DR blocking antibodies, indicating that Th cells respond to bacterial antigens in the context of MHC class II molecules (**Supplementary Fig. 6.3-1A, B**). The proliferating Th cells produced IFN-γ in combination with IL-17A and IL-22 (**Fig. 6.3-1C, D**) and, consistent with this phenotype, were significantly enriched in the CXCR3[+]CCR6[+] Th1*/17 cell subset (**Supplementary Fig. 6.3-2A**). *S. aureus*-reactive CD4[+] T cells produced also IFN-γ in combination with IL-17A and IL-22 while *M. tuberculosis*-reactive cells produced mainly IFN-γ, as previously reported (Acosta-Rodriguez et al., 2007; Becattini et al., 2015).

Because of their main localization at mucosal surfaces, we then analyzed in CXCR3[+]CCR6[+] Th1*/17 cells expression of the integrin α4β7 and of CD161, which have been previously linked to a subset of gut-homing IL-17-producing T cells in humans (Berlin et al., 1993; Cosmi et al., 2008; Kleinschek et al., 2009; Wagner et al., 1996). As shown in **Supplementary Fig. 6.3-2B**, a large fraction of CXCR3[+]CCR6[+] Th1*/17 cells expressed CD161 and α4β7, alone or in combination. When the distribution of Enterobacteriacae-reactive Th cells was analyzed in CD161[+] and CD161[−] Th1*/17, we found that *E. coli-* and *K. pneumoniae*-reactive T cells were enriched in the CD161[+] subset, whereas *S. typhimurium-* and *M. tuberculosis*-specific T cells were

present in both CD161$^+$ and CD161$^-$ subsets (**Supplementary Fig. 6.3-2C**). All bacteria-reactive CD161$^+$ Th cells produced both IL-17A and IFN-γ (thus displaying a Th1*/17 phenotype), while *S. typhimurium*- and *M. tuberculosis*-specific Th cells produced mainly IFN-γ only (thus displaying a Th1* phenotype) (data not shown).

Collectively, these results indicate that Enterobacteriaceae-reactive memory CD4$^+$ T cells are present in the blood of healthy individuals and are mainly confined in the CCR6$^+$CXCR3$^+$ Th1*/17 cell subset expressing the α4β7 integrin and CD161, and producing IFN-γ, IL-17A and IL-22.

Patients with systemic *K. pneumoniae* infection show reduced frequency of CXCR3$^+$ Th cells and selectively lack *K. pneumoniae*-reactive Th cells in blood

Enterobacteriaceae are among the major causes of systemic infections (Alhashem et al., 2017; Neuner et al., 2011; Sheu et al., 2019; van Duin et al., 2013; Weiner et al., 2016; Wilson et al., 2011). Reduction in the overall frequency of blood-circulating CD4$^+$ T cells and impaired T cell functions have been described in animal models of sepsis and in patients, as reviewed recently (Cabrera-Perez et al., 2014; Jensen et al., 2018a). To evaluate the presence and function of Enterobacteriaceae-reactive T cells in pathological conditions, we analyzed memory CD4$^+$ Th cells in the blood of patients with severe sepsis caused by systemic carbapenem-resistant *K. pneumoniae* infections. In these patients, frequency of circulating CXCR3$^+$ cells (CCR6$^-$ Th1 and CCR6$^+$ Th1*/17) was significantly reduced compared to healthy controls, while frequency of CCR4$^+$ cells (CCR6$^-$ Th2 and CCR6$^+$ Th17) was comparable (**Supplementary Fig. 6.3-3A**). In addition, CXCR3$^+$CCR6$^+$ memory T cells from septic patients did not proliferate in response to *K. pneumoniae,* while they proliferated in response to *M. tuberculosis*, although to a variable extent (**Supplementary Fig. 6.3-3B**). Of note, none of the other Th subsets isolated from septic patients proliferated in response to *K. pneumoniae* stimulation (data not shown), thus indicating a lack of a compensatory T cell response.

Memory CD4+ T cells are broadly cross-reactive to Enterobacteriaceae species

Previous studies provided evidence of CD4+ T cells able to cross-recognize multiple intestinal microbes in physiological and pathological conditions (Duchmann et al., 1999; Hegazy et al., 2017). To thoroughly define the extent of T cell cross-reactivity in the response to Enterobacteriaceae, we used two different approaches. First, we isolated CFSE^low memory CD4+ T cells proliferating in response to Enterobacteriaceae, re-labeled the cells with CFSE and performed heterologous re-stimulation. Th cell lines obtained from the primary stimulation with *E. coli, K. pneumoniae, E. aerogenes, S. typhimurium* or *S. aureus* showed robust proliferation upon secondary stimulation with the same bacteria, as expected, but also to heterologous stimulation with all the other Enterobacteriaceae species tested, but not to *S. aureus* (**Fig. 6.3-2A, B**). In contrast, Th cell lines obtained from the primary stimulation with *S. aureus* proliferated upon secondary stimulation with *S. aureus* but not with Enterobacteriaceae species.

In the second approach, we isolated from 8 donors a large number of T cell clones from CFSE^low memory CD4+ T cells proliferating in response to Enterobacteriaceae. The T cell clones (*n* = 685) were screened for their reactivity against the panel of Enterobacteriaceae species or *S. aureus* as control (**Fig. 6.3-2C, D**). Specific Th cell clones (i.e. reactive only to the original bacteria used in the primary stimulation) were very rare among clones obtained from *E. coli*-, *K. pneumoniae*-, *E. aerogenes*- and *Shigella*-cultures, while, strikingly, around 60% of them cross-reacted with 3 or more Enterobacteriaceae species (**Fig. 6.3-2C, D**). As an exception, the fraction of *S. typhimurium*-specific Th cell clones was higher (around 65%) and less than 30% cross-reacted with 3 other Enterobacteriaceae species (**Fig. 6.3-2C, D**). As control, *S. aureus*-specific Th cell clones did not cross-react to any of the Enterobacteriaceae species tested (**Fig. 6.3-2C, D**).

Collectively, these results reveal an extensive cross-reactivity in the human memory T cell repertoire against Enterobacteriaceae antigens.

Extensive clonotype sharing among Enterobacteriaceae-reactive memory CD4+ T cell repertoires

To define the TCR repertoire composition of Enterobacteriaceae-reactive CD4+ T cells, we sequenced TCR Vβ genes in CFSE^low cells obtained after primary stimulation, as previously described (Becattini et al., 2015). In all donors analyzed, the TCR repertoire of Enterobacteriaceae-reactive Th cells was polyclonal and comprised a variable number of clonotypes (mean ± s.d.: *E.coli* 379 ± 370; *K. pneumoniae* 318 ± 182; *E. aerogenes* 525 ± 337; *Shigella ssp* 261 ± 99; *S. typhimurium* 907 ± 543), which were fewer compared to the number of clonotypes in control *S. aureus*- or *M. tuberculosis*-reactive T cell cultures (mean ± s.d.: *S. aureus* 1512 ± 386; *M. tuberculosis* 1190 ± 658) (**Fig. 6.3-3A**). A non-metric multidimensional scaling analysis showed that Enterobacteriaceae-reactive TCR Vβ repertoires co-clustered together and separately from *S. aureus*- or *M. tuberculosis*-reactive repertoires, suggesting a high degree of overlap (**Fig. 6.3-3B**). Indeed, as shown in **Fig. 6.3-3C** for one representative donor, many *E. coli*-, *K. pneumoniae*-, *E. aerogenes*- and *S. typhimurium*-reactive TCR Vβ clonotypes were found shared with 1 (green), 2 (yellow) and even 3 (red) additional Enterobacteriaceae-reactive TCR Vβ clonotypes. Of note, the most expanded TCR Vβ clonotypes in each Enterobacteriaceae-reactive repertoire (top 5%) were mainly broadly cross-reactive, being found in all 4 Enterobacteriaceae-reactive T cell populations (**Fig. 6.3-3C**). In addition, although the Enterobacteriaceae cross-reactive clonotypes comprised 26%-49% of the unique TCR Vβ productive rearrangements of each Enterobacteriaceae-responding repertoire, the cumulative frequency accounted for most of the total T cell response (range 66-93% cumulative percentage of templates) (**Fig. 6.3-3D**). Consistent with the lack of cross-reactivity observed in polyclonal cultures and T cell clones, very rare TCR Vβ clonotypes were shared between *S. aureus*- or *M. tuberculosis*-reactive and Enterobacteriaceae-reactive T cells.

To evaluate the level of clonal expansion of Enterobacteriaceae-reactive TCR Vβ clonotypes *in vivo*, we compared in 2 blood donors the Enterobacteriaceae-reactive TCR Vβ clonotypes of CFSE^low memory CD4+ T cells with that of total memory CD4+ T cells sequenced directly *ex vivo* (**Fig. 6.3-3E**). Several Enterobacteriaceae-reactive clonotypes could be identified in circulating T cells. Interestingly, broadly cross-reactive TCR clonotypes (red circles) showed higher clonal expansion compared to

specific-TCR clonotypes, with 11 out of 19 and 21 out of 34 broadly cross-reactive TCR Vβ clonotypes being within the top 5% expanded clonotypes in circulating memory CD4$^+$ T cells of donor A and B, respectively.

Collectively, these findings demonstrate that the human memory CD4$^+$ T cell response to Enterobacteriaceae comprises specific and broadly cross-reactive T cell clones, with cross-reactive clones being the most expanded *in vivo*.

Cross-reactivity to Enterobacteriaceae is imprinted in the naïve repertoire

The high extent of cross-reactivity of the Enterobacteriaceae-reactive memory CD4$^+$ T cell repertoire raised the question as whether cross-reactivity is already set in the naïve repertoire. To address this question, we performed *in vitro* priming experiments following a protocol previously adopted to study *S. aureus* and *C. albicans* naïve T cells (Becattini et al., 2015; Zielinski et al., 2012). Briefly, highly pure naïve CD4$^+$ T cells were FACS-sorted from the blood of healthy donors, labelled with CFSE and cultured with autologous monocytes pulsed with the different Enterobacteriaceae species or *S. aureus*, as control. Proliferating CFSE^{low} T cells were detected in all stimulated cultures but not in control cultures or in cultures performed in the presence of MHC-II blocking antibodies (**Fig. 6.3-4A**), consistent with a high frequency of naïve T cell precursors for these complex antigens. A large panel of Th cell clones (*n* = 585) were then isolated from primed CFSE^{low} cells and tested for their reactivity against different Enterobacteriaceae species. As shown in **Fig. 6.3-4B**, a sizable fraction of the Enterobacteriaceae-reactive Th cell clones isolated from *in vitro* primed naïve cells were cross-reactive to at least another Enterobacteriaceae species. Analysis of several donors revealed that up to 14% of the T cell clones cross-reacted to 3 additional Enterobacteriaceae species (**Fig. 6.3-4C**). However, a higher proportion of naïve T cell-derived clones than memory T cell-derived clones was specific for the bacteria eliciting the primary response (12% *vs* 3.5% in *E. coli*, 43% *vs* 6.9% in *K. pneumoniae*), with the exception of *S. typhimurium* (43% vs 65%) (**Fig. 6.3-4C** compared to **Fig. 6.3-2D**). Of note, up to 15% of the Th cell clones derived from *in vitro S. aureus*-primed naïve T cells were able to cross-react with phylogenetically distant Enterobacteriaceae species (**Fig. 6.3-4B, C**).

Collectively, these findings show that Enterobacteriaceae cross-reactive T cells are already present in the naïve CD4[+] T cell compartment, which concurrently holds a sizable fraction of precursors specific for individual Enterobacteriaceae species.

Outer membrane protein A is an important antigen determinant of cross-reactivity to Enterobacteriaceae

The T cell cross-reactivity observed for Enterobacteriaceae may be associated with the presence of conserved antigenic epitopes in different bacteria species or due to an intrinsic TCR binding degeneracy (Huppa and Davis, 2013; Mason, 1998; Sewell, 2012). To address this point, we interrogated the literature for potentially immunogenic proteins conserved within the Enterobacteriaceae family. Outer membrane protein A (OmpA) is a porin widely expressed by gram-negative bacteria, highly conserved among the Enterobacteriaceae family (**Supplementary Fig. 6.3-4**) and throughout evolution (Delcour, 2002; Krishnan and Prasadarao, 2012), and an important target of both humoral and cellular protective immune responses *in vivo* (Lee et al., 2012; Pennini et al., 2017; Pore and Chakrabarti, 2013). To evaluate the immunogenicity of OmpA for CD4[+] T cells, Enterobacteriaceae-reactive T cell clones isolated from memory and naïve compartments were stimulated *in vitro* with recombinant OmpA from *K. pneumoniae* in presence of autologous monocytes. We observed reactivity to OmpA in 12% to 33% (*n* = 397) of clones isolated from memory CD4[+] T cells, and in 7% to 30% (*n* = 262) of clones from *in vitro* primed naïve CD4[+] T cells (**Fig. 6.3-5A, B**). Importantly, most OmpA-specific T cell clones were broadly cross-reactive to three or more Enterobacteriacaeae.

These results indicate OmpA as one of the immunogenic antigens targeted by a large percentage of Enterobacteriaceae broadly cross-reactive CD4[+] T cells, which likely recognize conserved epitopes with similar amino acid sequences encoded by different bacterial species.

6.3.5 Discussion

In this study we performed a thoroughly analysis of the human CD4$^+$ T cell responses to commensals and pathogens belonging to the Enterobacteriaceae family. In line with recent studies (Duhen and Campbell, 2014; Hegazy et al., 2017; Zimmermann et al., 2015), these cells show a Th1*/17 phenotype characterized by the expression of CCR6 and CXCR3 as well as production of IFN-γ, IL-17A and IL-22. Further dissection of the CCR6$^+$CXCR3$^+$ Th cell compartment established that Enterobacteriaceae-reactive cells are enriched in a T cell subpopulation expressing the C-type lectin-like receptor CD161 associated with the gut homing marker α4β7 integrin (Cosmi et al., 2008; Kleinschek et al., 2009), thus remarking their distinction from the Th1* cells induced by Mycobacteria (Acosta-Rodriguez et al., 2007; Lindestam Arlehamn et al., 2013; Okada et al., 2015; Sallusto et al., 2018).

By combining multiple techniques that allow different levels of sensitivity and throughput, we report a high level of cross-reactivity at the monoclonal (T cell clones), oligoclonal (T cell lines) and polyclonal (TCR Vβ sequencing) level. In particular, we showed that Enterobacteriaceae broadly cross-reactive memory CD4$^+$ T cells are clonally expanded *in vivo*, with some TCR Vβ clonotypes found at relatively high frequency in the circulating antigen-experienced repertoire. Surprisingly, contrarily to previous reports pointing to an almost total absence of bacteria cross-reactive naïve T cells (Bacher et al., 2019; Hegazy et al., 2017), circulating naïve CD4$^+$ T cells comprised a high frequency of Enterobacteriaceae cross-reactive T cells. This finding indicates that the high level of CD4$^+$ T cell cross-reactivity to Enterobacteriaceae is already imprinted in the antigen-inexperienced naïve repertoire and may provide a selective advantage for protective host immunity. Our results also indicate that Enterobacteriaceae cross-reactive Th cells become dominant in the memory Th cell compartment, which may be the manifestation of an "original antigenic sin". This term refers to the phenomenon observed in T and B cell responses by which pre-existing immunity acquired by earlier antigen encounter biases the subsequent immune responses to closely related antigens, that therefore boosts the expansion of cross-reactive memory cells arose from prior challenges rather than inducing de novo priming of antigen-specific naïve precursors (Lee et al., 2019; Sewell, 2012). The high clonal expansion of Enterobacteriaceae broadly cross-reactive memory Th cells measured in

healthy donors might therefore be the result of the preferential expansion of cross-reactive over specific clones within the pool of early primed T cells, driven by multiple encounters with closely related members of the human microbiota.

The selection of a pool of clonally expanded cross-reactive memory T cells might be on the one hand advantageous for host protection, since it could potentially confer heterologous immunity to an extended range of pathogens prior to antigen exposure (Campion et al., 2014; Su et al., 2013). On the other hand, pre-existing cross-reactive immunity could also be detrimental if the pool of expanded T cells holds an effector phenotype not suitable to deliver sterilizing immunity to the newly encountered pathogens. Development of antigen-specific Th responses with inappropriate phenotypes can result in failure of host protection, as shown in humans for *Mycobacterium leprae* (Sieling and Modlin, 1994; Yamamura et al., 1991), *M. tuberculosis* (Lindestam Arlehamn and Sette, 2014) and *Candida albicans* (Okada et al., 2015). Cross-reactive T cells can be a double-edged sword, as shown recently in the case of *Candida*-induced Th17 cells that, while ensuring intestinal homeostasis, can contribute to lung inflammation and immunopathology upon cross-recognition of airborne *Aspergillus fumigatus* (Bacher et al., 2019).

We identified Outer membrane protein A (OmpA) as one of the antigenic determinants of Th cells broadly-cross-reactive with Enterobacteriaceae. Sequence analysis of OmpA homologs in different bacterial species revealed high sequence identity, thus suggesting that cross-reactivity to Enterobacteriaceae is mainly due to T cell recognition of conserved epitopes in different, although phylogenetically closely-related bacteria. Given the evidence of OmpA as an important target of antibody response to Enterobacteriaceae (Lee et al., 2012; Pennini et al., 2017; Pore and Chakrabarti, 2013), our findings are consistent with a tight connection existing between T- and B-cell immune responses, often converging to the same antigenic targets. T cell cross-reactivity to multiple antigens can be also due to promiscuous TCR engagement of many MHC-II molecules loaded with highly different peptides. The latter phenomenon is known as TCR degeneracy and relies to the intrinsic weak affinity of TCR-pMHC-II interactions (Huppa and Davis, 2013; Mason, 1998; Sewell, 2012) that can allow unexpected patterns of cross-recognition (Campion et al., 2014; Su et al., 2013). TCR degeneracy might be at the basis of the observed cross-reactivity of rare T

cell clones that respond to *S. aureus* and *M. tuberculosis* that are present in both memory and naïve CD4[+] T cell compartments.

An interesting observation of this study was that septic patients suffering from *K. pneumoniae* systemic infection have a selective reduction of circulating CXCR3[+] Th cells (both Th1 and Th1*/17) and lack *K. pneumoniae*-reactive memory Th cells, while *M. tuberculosis*-reactive T cells are still detectable. The selective defect in *K. pneumoniae*-reactive T in these patients suggests that sepsis-induced immunosuppression can be elicited in an antigen-specific, TCR-mediated fashion, for instance as the result of impaired T cell function or of activation-induced cell death (AICD) caused by massive antigen exposure (Cabrera-Perez et al., 2014; Jensen et al., 2018a). Interestingly, in the blood of two patients that survived *K. pneumoniae* sepsis, the Th1*/17 response to *K. pneumoniae* measured after recovery was unaffected and comparable to the one observed in healthy controls (data not shown). These data suggest that the presence of a circulating pool of *K. pneumoniae*-reactive T cells may correlate with protection from sepsis, although further studies are needed to definitively dissect this point.

In spite of extensive research and clinical trials, for many pathogens of the Enterobacteriaceae family such as *Shigella* spp, *K. pneumoniae* or *Yersinia pestis* there are no safe and effective vaccines (Levine et al., 2007; Mani et al., 2016; Verma and Tuteja, 2016), and currently available vaccines for other species, such as thyphoidal *Salmonella enterica* serovars, showed mild efficacy and conferred short-lasting immunity (Milligan et al., 2018). In light of our findings, we propose that the large level of Enterobacteriaceae cross-reactivity observed in the memory CD4[+] T cell repertoire can have a profound impact on immunization outcome and development of protective immunity to closely related pathogens, and therefore should be a parameter to take carefully in consideration to improve vaccines efficacy. The identification of antigenic determinants of cross-reactive and specific T cell responses to commensals and pathogens might be extremely relevant for the rational design of subunit vaccines against newly emerging multi-drug resistant Enterobacteriaceae. Directing the immune response to common or unique antigenic targets might allow to hijack, or alternatively overcome, the pre-existing immunity to ensure host protection.

6.3.6   Figures
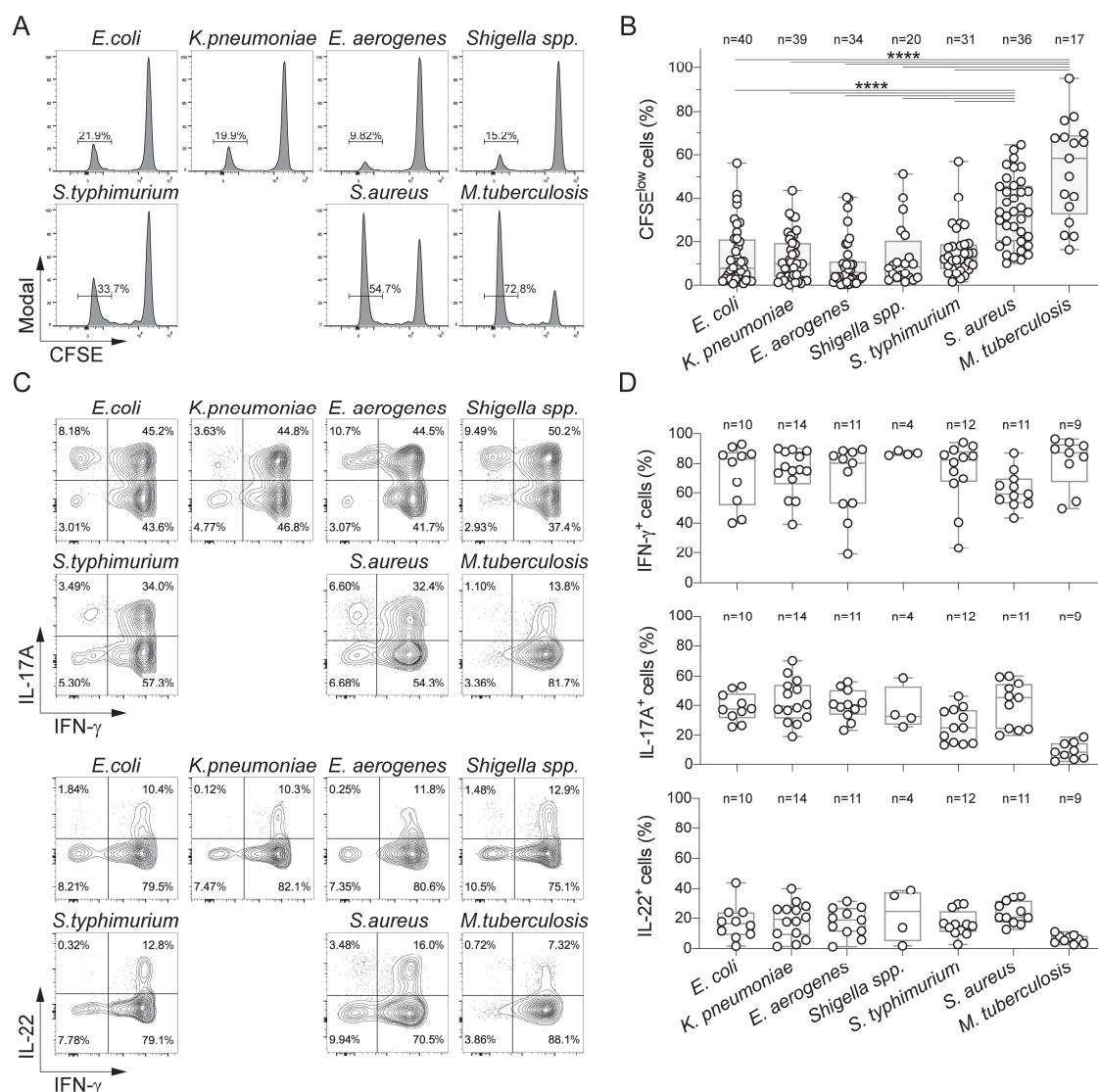
Figure 6.3-1. Enterobacteriaceae-reactive memory CD4+ T cells are present in the blood of healthy donors and show a Th1*/17 phenotype. Human memory CD4+ T cells were isolated from PBMCs, labeled with CFSE and stimulated with the indicated heat-inactivated bacteria in the presence of autologous monocytes. (**A**) CFSE profiles on day 6 after stimulation and percentage of CFSE^{low} proliferating cells from a representative donor. (**B**) Pooled data of percentage of CFSE^{low} cells from multiple donors (n = 17-40). ******** p-value < 0.0001, as determined by two-tailed unpaired t-test. (**C** and **D**) IFN-γ, IL-17A and IL-22 production by CFSE^{low} fractions was measured by intracellular staining after stimulation with phorbol myristate acetate (PMA) and ionomycin in the presence of Brefeldin A (BfA) in a representative donor (**C**) and in several donors (n = 4-14) (**D**).
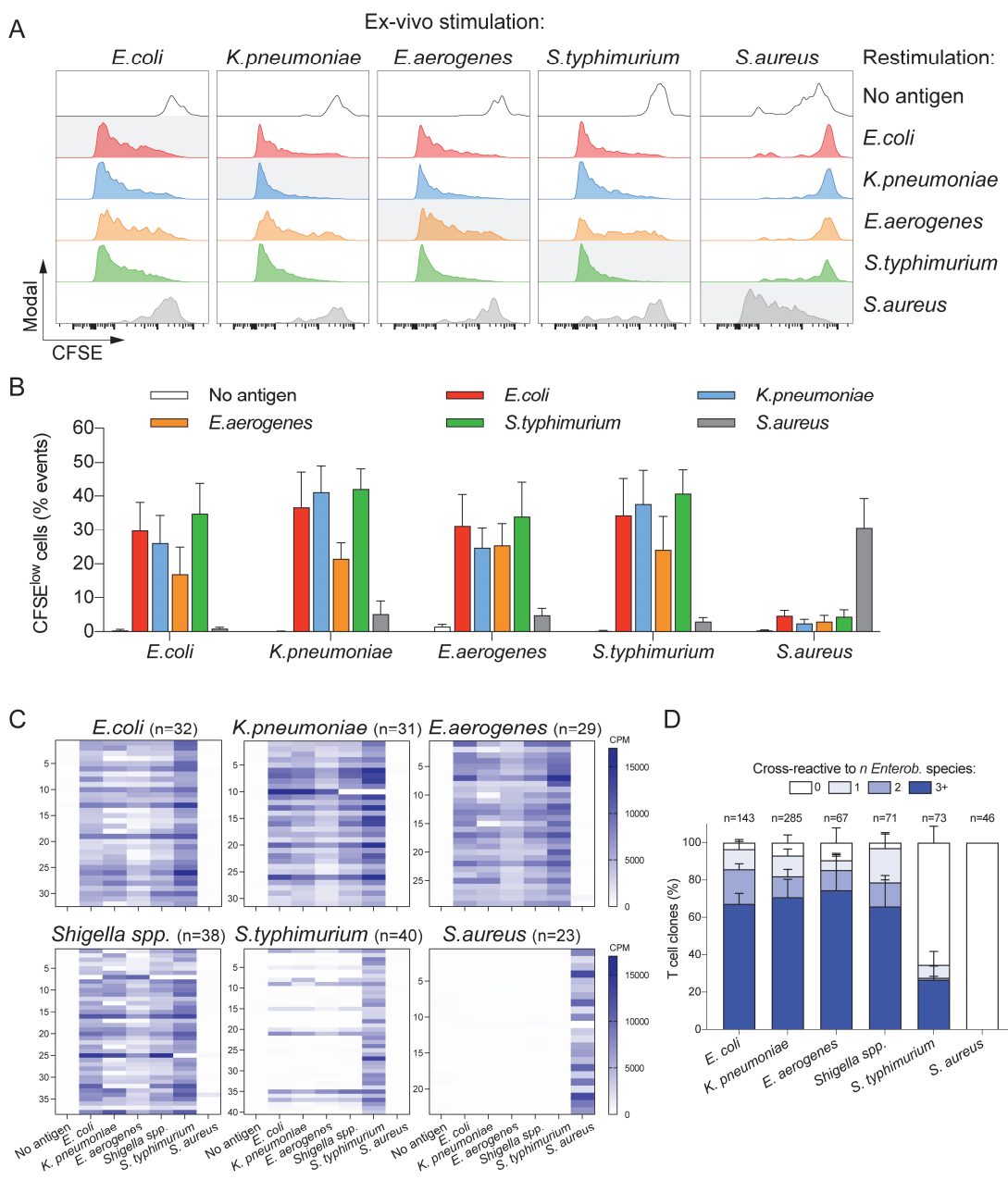
Figure 6.3-2. Memory CD4⁺ T cells are broadly cross-reactive to Enterobacteriaceae. Bacteria-reactive polyclonal T cell lines were obtained by FACS-sorting of CFSE^{low} fractions from primary cultures of memory CD4⁺ T cells. Following expansion, CFSE^{low} cultures were re-labeled with CFSE and stimulated with the same bacteria (homologous re-stimulation) or with different bacteria (heterologous re-stimulation) and autologous monocytes. (**A**) CFSE profiles of polyclonal T cell lines reactive for the antigens indicated on top, restimulated for 5 days with the bacterial antigens reported on the right, from a representative donor. Homologous re-stimulation for each culture is highlighted with a grey shadow. (**B**) Pooled data of percentage of CFSE^{low} cells after secondary stimulation from multiple donors ($n = 5$). CFSE^{low} cells are reported as percentage of total events after normalization on live lymphocytes. (**C** and **D**) A large panel of T cell clones were isolated from CFSE^{low} cultures by limiting dilution. T cell clones isolated from each CFSE^{low} fraction were screened with a panel of bacteria in the presence of autologous monocytes. Proliferation was assessed on day 3 after a 16-h pulse with [³H]-thymidine and expressed as counts per min (CPM). The heatmaps report the CPM of T cell clones isolated from CFSE^{low} fractions of

a representative donor (**C**). The antigen used in the primary stimulation and the number of clones tested are reported on top of each heatmap. Each row of the heatmaps refers to an individual T cell clone, stimulated with the antigens reported at the bottom. (**D**) Pooled data of the patterns of reactivity of T cell clones isolated from the CFSE^{low} fractions of memory CD4$^+$ T cells from multiple donors ($n = 8$). The barplot represents the frequency of non-cross-reactive T cell clones (0) or of T cell clones cross-reactive to other 1, 2 and 3 or more enterobacteriaceae species. The total number of clones analyzed is reported on top.
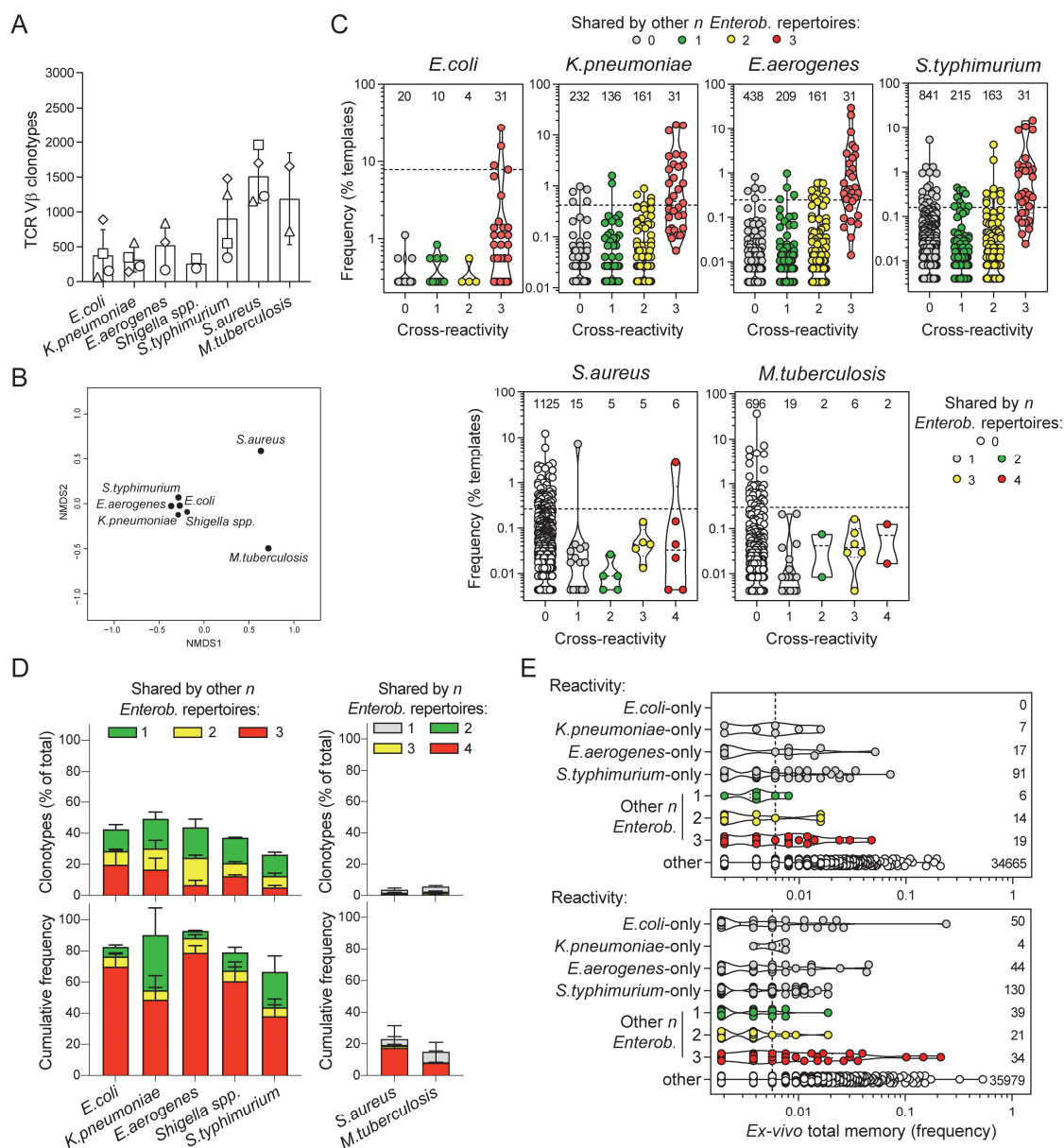
Figure 6.3-3. Extensive clonotype sharing among Enterobacteriaceae-reactive memory CD4[+] T cell repertoires. Bacteria-reactive CFSE[low] fractions were FACS-sorted from stimulated memory CD4[+] T cells, and their TCR Vβ repertoire was determined by deep sequencing. (**A**) Number of unique productive TCR Vβ nucleotide rearrangements resolved from each bacteria-responsive memory repertoire. Each symbol represents a different donor (*n* = 2-4). (**B**) Pairwise similarity of each antigen-reactive TCR Vβ repertoire was calculated by Chao-Jaccard overlap. The reciprocal distance of each TCR Vβ repertoire was plotted using Kruskal's non-metric multidimensional scaling. Each point of the graph represents a distinct bacteria-reactive memory CD4[+] repertoire (average from *n* = 2-4 donors): the closer the points, the more similar the repertoires in terms of sharing of TCR Vβ clonotypes. (**C**) Frequency distribution of TCR Vβ clonotypes reactive to Enterobacteriaceae or control bacteria from a representative donor. Each circle represents a unique TCR Vβ nucleotide rearrangement, and the color code indicates the number of distinct Enterobacteriaceae-reactive repertoires in which the clonotype was found. The number of clonotypes for each class of reactivity is reported on top. Dotted lines indicate the frequency of the top 5% expanded clonotypes. Because of limited cell number, *Shigella* was not tested in this donor. (**D**)

Pooled data of percentage of unique TCR Vβ rearrangements (*upper panels*) and corresponding cumulative frequency of templates (*lower panels*) of Enterobacteriaceae cross-reactive TCR Vβ clonotypes measured from multiple donors (*n* = 2-4). Color code indicates cross-reactive clonotypes found shared between the indicated Enterobacteriaceae-reactive repertoires or control repertoires, and other 1, 2 and 3 or more enterobacteriaceae-reactive repertoires. (**E**) Frequency distribution of Enterobacteriaceae-reactive TCR Vβ clonotypes from two different donors measured by deep sequencing of total memory CD4$^+$ T cells directly *ex vivo* after FACS-sorting (*upper panel*: donor A; *lower panel*: donor B). The reactivity of each clonotype was determined by comparison with the collection of TCR Vβ nucleotide sequences measured in CFSE^{low} cultures obtained by stimulation of memory CD4$^+$ T cells from the same donors. The color code indicates clonotypes specific (grey circles) or cross-reactive with 1 (green circles), 2 (yellow circles) or 3 Enterobacteriaceae species (red circles). The number of clonotypes for each class of reactivity is reported on the right. The number and frequency distribution of all other clonotypes in the memory CD4$^+$ T cell repertoire is also reported (white circles). Dotted lines indicate the frequency of the top 5% expanded clonotypes.
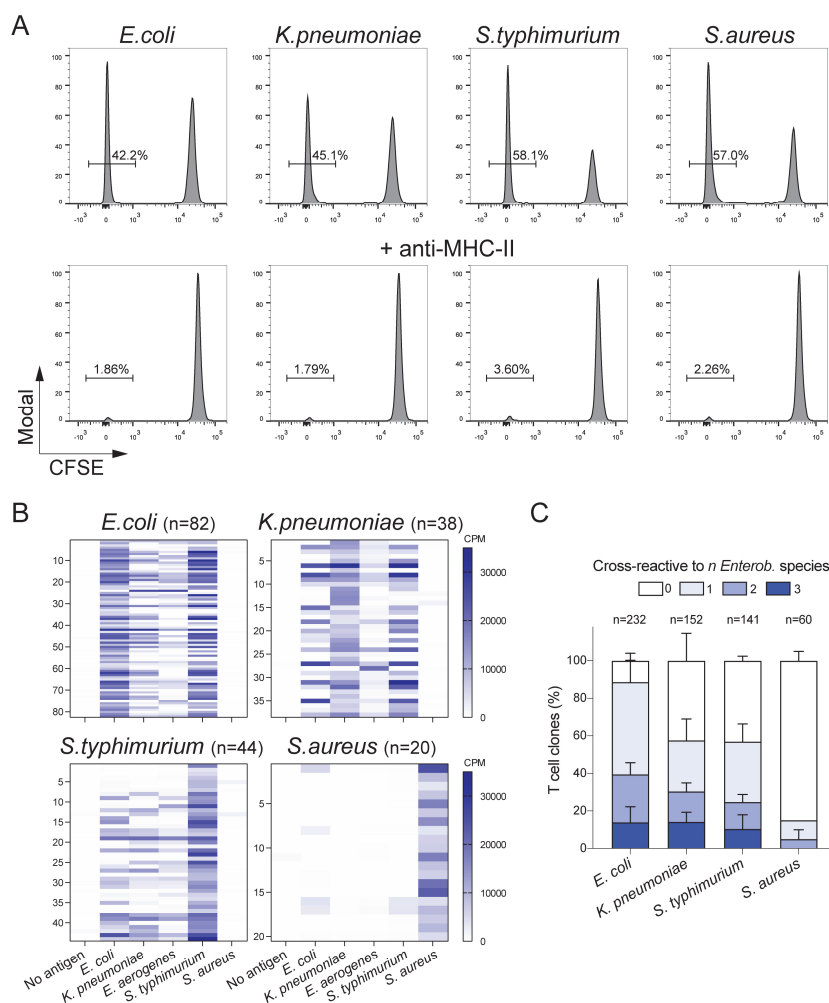
Figure 6.3-4. Cross-reactive T cells to Enterobacteriaceae are already detected in the naïve repertoire. Human naïve CD4⁺CD95⁻ T cells were FACS-sorted at high purity, labeled with CFSE and primed *in vitro* for 10 days in the presence of the indicated heat-inactivated bacteria and autologous monocytes. (**A**) CFSE profiles on day 10 after priming and percentage of CFSE^{low} proliferating cells from a representative donor (*upper panel*). Priming was performed in the presence of anti-MHC-II blocking antibodies (anti-HLA-DR, clone L243; anti-HLA-DQ, clone SPVL3; anti-HLA-DP, clone B7/21; *lower panel*). Complete suppression of CFSE^{low} responses was observed in all donors tested (*n*=3). (**B**) Antigen-primed CFSE^{low} T cells were cloned by limiting dilution. T cell clones isolated from each CFSE^{low} fraction were screened with a panel of heat-inactivated bacteria in the presence of autologous monocytes. Proliferation was assessed on day 3 after a 16-h pulse with [³H]-thymidine and expressed as counts per min (CPM). The heatmaps report the CPM of T cell clones isolated from CFSE^{low} fractions of a representative donor. The antigen used for the initial *in vitro* priming and the number of clones tested are reported on top of each heatmap. Each row of the heatmaps refers to an individual T cell clone, stimulated with the antigens reported at the bottom. (**C**) Pooled data of the patterns of reactivity of T cell clones isolated from each CFSE^{low} fraction of *in vitro* primed naïve CD4⁺ T cells from multiple donors (*n* = 6). The barplot represents the frequency of T cell clones cross-reactive to other 1, 2 and 3 Enterobacteriaceae species. The total number of clones analyzed is reported on top.
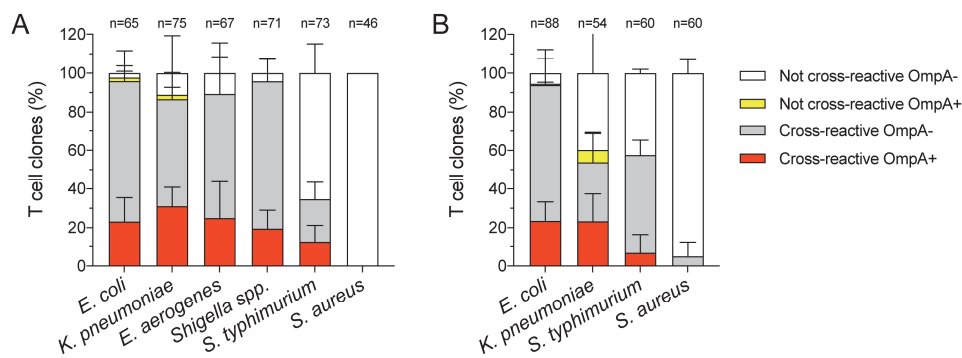
Figure 6.3-5. Outer membrane protein A is an important antigen determinant of cross-reactivity to Enterobacteriaceae. Enterobacteriaceae-reactive T cell clones isolated from multiple donors (*n* = 3) were stimulated with recombinant Outer membrane protein A (OmpA) from *K. pneumoniae* in the presence of autologous monocytes. Pooled data of reactivity to *K. pneumoniae*-OmpA of T cell clones isolated from the CFSE^{low} fractions of memory (**A**) or naïve (**B**) CD4$^+$ T cells stimulated with the indicated antigens. The barplots represent the percentage of cross-reactive (to 3 or more Enterobacteriaceae species) and specific T cell clones that react to OmpA (colored in red and yellow, respectively), isolated from each CFSE^{low} fraction. The total number of clones analyzed is reported on top.
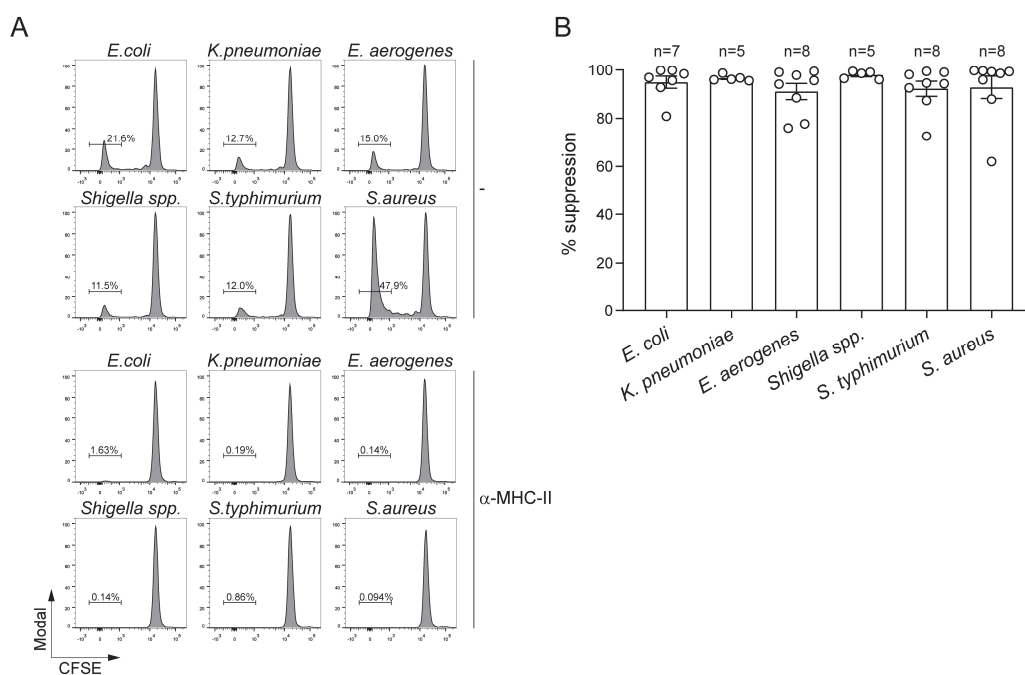
Supplementary figure 6.3-1. CD4$^+$ T cell proliferation to Enterobacteriaceae is inhibited by anti-MHC-II blocking antibodies. (**A**) Human memory CD4$^+$ T cells were isolated from PBMCs, labeled with CFSE and stimulated with the indicated heat-inactivated bacteria and autologous monocytes in the absence or presence of anti-MHC-II blocking antibodies (anti-HLA-DR, clone L243; anti-HLA-DQ, clone SPVL3; anti-HLA-DP, clone B7/21). Shown are the CFSE profiles from a representative donor (**A**) and pooled data of percentage of inhibition by anti-MHC-II blocking antibodies in *n* = 5-8 donors (**B**).
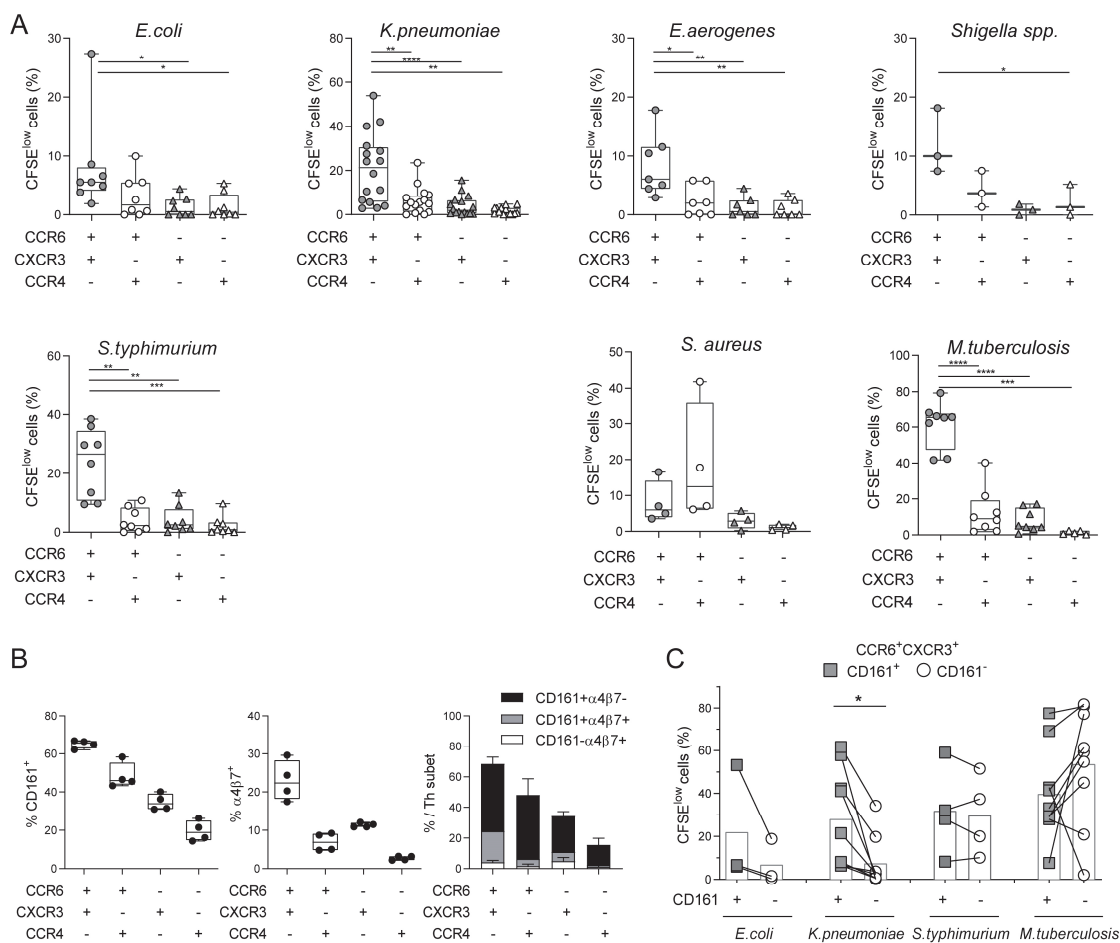
Supplementary figure 6.3-2. Enterobacteriaceae-reactive T cells are contained in the CXCR3⁺CCR6⁺ Th subset that is enriched in CD161⁺α4β7⁺ cells. (**A**) Human memory CD4⁺ T cell subsets were isolated according to the expression of CCR6, CXCR3 and CCR4, labeled with CFSE, and stimulated with the indicated heat-inactivated bacteria in the presence of autologous monocytes. Pooled data of percentage of CFSE^low cells on day 6 from multiple donors (*n* = 3-16). (**B**) Surface expression of CD161 and α4β7 integrin on human memory CD4⁺ T cell subsets measured by flow cytometry; pooled data from *n* = 4 donors. (**C**) Human memory CXCR3⁺CCR6⁺ Th cells were sorted based on CD161 expression, labeled with CFSE, and stimulated with the indicated antigens in the presence of autologous monocytes. Pooled data of percentage of CFSE^low cells on day 6 from *n* = 3-9 donors. * *p*-value < 0.05, ** *p* < 0.01, *** *p* < 0.001, **** *p* < 0.0001, as determined by two-tailed paired t-test.
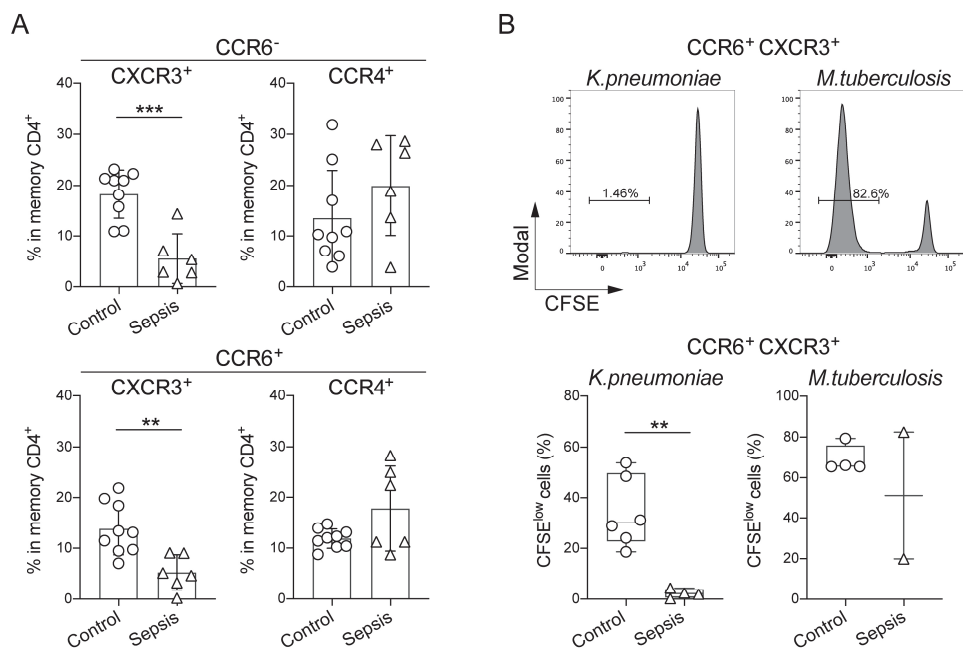
Supplementary figure 6.3-3. Septic patients with systemic *K. pneumoniae* infection show reduced frequency of circulating CXCR3⁺ T cells and selectively lack *K. pneumoniae*-reactive T cells. (A) Surface expression of chemokine receptors by memory CD4$^+$ T cells from PBMCs of healthy controls (*n* = 9) or septic patients with systemic *K. pneumoniae* infection (*n* = 6) was analzed by flow cytometry. Shown are pooled data of percentage of memory CD4$^+$ T cells positive for combinations of chemokine receptors. (**B**) Human memory CXCR3$^+$CCR6$^+$CCR4$^-$ Th cells were sorted, labeled with CFSE, and stimulated with the indicated antigens in the presence of autologous monocytes. Shown are CFSE profiles on day 6 in a representative septic patient (*upper panel*), and pooled data of percentage of CFSE^{low} cells in healthy controls or septic patients (*lower panel*). ** *p*-value < 0.01, *** *p* < 0.001, as determined by two-tailed unpaired t-test.
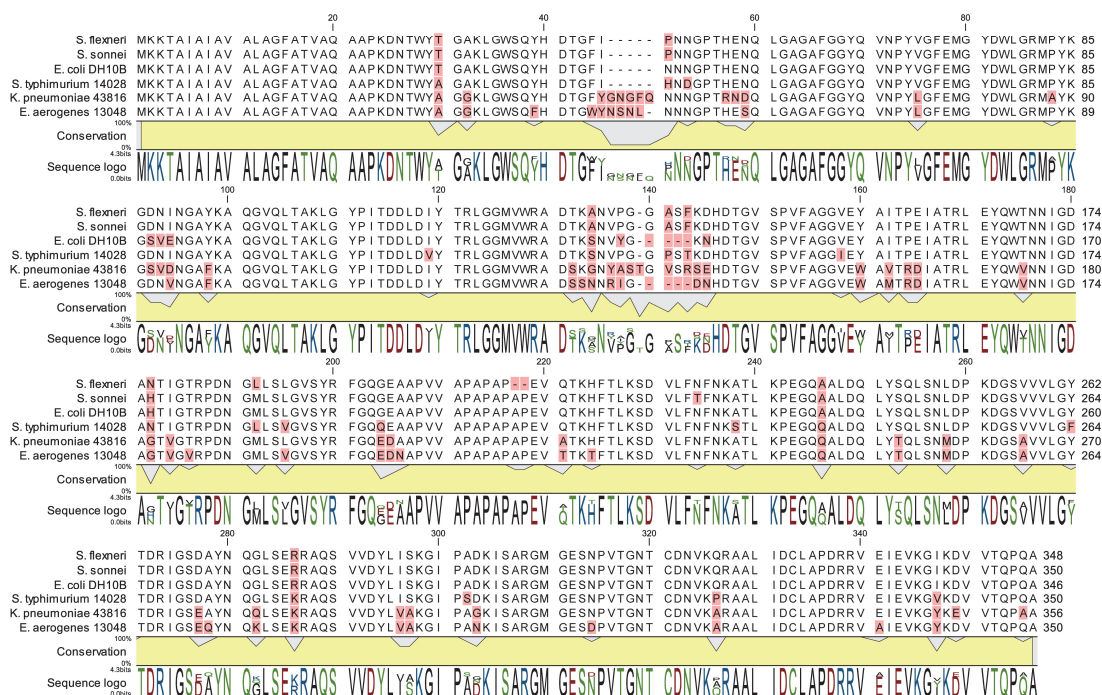
Supplementary figure 6.3-4. Multiple sequence alignment of OmpA homologs in different Enterobacteriaceae species. Variable residues are highlighted in red. A histogram of conservation for each residue and a sequence LOGO are reported below the alignment.

6.3.7 Methods

Cells and cell sorting. Blood from healthy donors was obtained from the Swiss Blood Donation Center of Basel and Lugano. Blood from septic patients was obtained from Policlinico Umberto I, Department of Public Health and Infectious Diseases (University of Rome "La Sapienza", Italy). All patients provided written informed consent for participation in the study. Human primary cell protocols were approved by the Federal Office of Public Health (no. A000197/2 to F.S.). Peripheral Blood Mononuclear Cells (PBMCs) were isolated with Ficoll-Paque Plus (GE Healthcare). Monocytes and total CD4$^+$ T cells were isolated by positive selection using CD14 and CD4 magnetic microbeads, respectively (Miltenyi Biotech). Total CD4$^+$ cells obtained by positive selection were stained on ice for 15-20 min with the following fluorochrome-labeled mouse monoclonal antibodies: CD8-PE–Cy5 (clone B9.11; cat. no. A07758), CD14-PE–Cy5 (clone RMO52; cat. no. A07765), CD16-PE–Cy5 (clone 3G8; cat. no. A07767), CD56-PE/Cy5 (clone N901; cat. no. A07789), CD19-PE/Cy5 (clone J3-119; cat. no. A07771), CD25-PE–Cy5 (clone B1.49.9; cat. no. IM2646) from Beckman Coulter, CD4–PE–Texas Red (clone S3.5; cat. no. MHCD0417), CD45RA-Qdot 655 (clone MEM-56; cat. no. Q10069), CD95-PerCP-eFluor 710 (clone DX2; cat. no. 46-0959-42) from ThermoFisher Scientific, CCR7–BV421 (clone G043H7; cat. no. 353208), CCR6-BV605 (clone G034E3; cat. no. 353420), CXCR3-Alexa Fluor 647 (clone G025H7; cat. no. 353712) from BioLegend, CCR4-PE-Cy7 (clone 1G1; cat. no. 557864) from BD Biosciences. In some experiments CD161-PE/Cy7 (clone HP-3G10; cat. no. 339918) from BioLegend and anti-α4β7 integrin (Vedolizumab) from Humabs BioMed (Bellinzona, Switzerland) were used. Naïve and memory CD4$^+$ T cells were sorted to over 98% purity on a FACSAria III (BD) after exclusion of CD8$^+$, CD14$^+$, CD16$^+$, CD56$^+$, CD19$^+$ and CD25^{bright} cells. Naïve cells were sorted as CD4$^+$CD45RA$^+$CCR7$^+$CD95$^-$ cells; the remaining CD4$^+$ T cells were sorted as total memory. Memory Th cell subsets were sorted as follows: CXCR3$^+$CCR4$^-$CCR6$^-$ (enriched in Th1), CCR4$^+$CXCR3$^-$CCR6$^-$ (enriched in Th2), CCR6$^+$CXCR3$^+$CCR4$^-$ (enriched in Th1*/17), CCR6$^+$CCR4$^+$CXCR3$^-$ (enriched in Th17). In some experiments, CCR6$^+$CXCR3$^+$ memory Th cells were further divided based the expression of CD161.

Microbes and antigens. The following microbial strains were kindly provided by Servizio di Microbiologia EOLAB, Ente Ospedaliero Cantonale (Bellinzona,

Switzerland): *Escherichia coli* (ATCC DH10B), *Klebsiella pneumoniae* (ATCC 43816), *Enterobacter aerogenes* (ATCC 13048), *Shigella flexneri* (ATCC 12021), *Shigella sonnei* (ATCC 9290), *Salmonella typhimurium* (ATCC 14028), *Staphylococcus aureus* (ATCC 29213) and *Streptococcus pyogenes* (ATCC 19615). Bacteria were cultured in aerobic condition at 37°C in Luria-Bertani broth. After expansion, bacteria were extensively washed in PBS and heat inactivated at 60°C for 2 hrs. Bacteria count of each isolate was determined by staining with 4', 6-diamidino-2'-phenylindole, dihydrochloride (DAPI, Fisher Scientific) and counting at the fluorescence microscope. Ratio used for stimulation assays was 2 bacteria particles per monocyte. Whole cell lysate of *Mycobacterium tuberculosis* (strain CDC1551, cat. no. NR-14823) was obtained through BEI Resources (NIAID, NIH), and was used at a final concentration of 3 µg/ml. Recombinant Outer membrane protein A (OmpA) from *Klebsiella pneumoniae* (ATCC 43816) was kindly provided by Humabs BioMed (Bellinzona, Switzerland), and used at 5 µg/ml for stimulation assays.

T cell stimulation. T cells were cultured in RPMI 1640 medium supplemented with 2 mM glutamine, 1% (vol/vol) nonessential amino acids, 1% (vol/vol) sodium pyruvate, penicillin (50 U/ml), streptomycin (50 µg/ml) (all from Invitrogen) and 5% human serum (Swiss Red Cross). Sorted memory CD4$^+$ T cells or naïve CD4$^+$CD95$^-$ T cells were labelled with carboxyfluorescein succinimidyl ester (CFSE) and cultured at a ratio of 2:1 with untreated or antigen-pulsed irradiated autologous monocytes. Depending of the antigen, monocytes were pulsed 3-5 hrs with heat-inactivated bacteria (2 particles per monocytes) or *M. tuberculosis* whole cell lysate (3 µg/ml). To determine MHC restriction, the assay was performed in the absence or presence of blocking anti-MHC-II monoclonal antibodies produced in house from hybridoma cell lines (anti-HLA-DR, clone L243 from ATCC, cat. no. HB-55; anti-HLA-DQ, clone SPVL3 (Spits et al., 1983); anti-HLA-DP, clone B7/21 (Watson et al., 1983)). Stimulated memory and naïve T cell cultures were collected at day 6-7 and 8-10, respectively, and stained with antibodies to CD25–PE (clone M-A251; cat. no. 555432) from BD Biosciences and ICOS–APC (clone C398.4A; cat. no. 313510) from BioLegend. Proliferating activated T cells were FACS-sorted as CFSE^{low}CD25$^+$ICOS$^+$ and expanded in vitro in the presence of IL-2 (500 IU/ml). In some experiments, CFSE^{low} cultures were labelled again with CFSE and challenged in secondary stimulations with antigen-pulsed

irradiated autologous monocytes. Readout of T cell proliferation was determined at day 4-5 after secondary stimulation. Percentage of CFSE^{low} cells were normalized on live lymphocytes as follows: (% lymphocytes) × (% live cells) × (% CFSE^{low} cells).

Intracellular cytokine staining. CFSE^{low} cultures from ex vivo stimulated memory cells were stimulated with PMA and ionomycin for 5h in the presence of brefeldin A for the last 2h (all reagents from Sigma-Aldrich). Cell viability was determined by staining with LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (ThermoFisher), according to the manufacturer's instructions. Subsequently cells were fixed and permeabilized with Cytofix/Cytoperm (BD Biosciences), and then stained with the following antibodies: IL-22-PerCP-eFluor710 (22URTI, cat. no. 46-7229-42) from eBioscience, IL-17A-BV605 (BL168, cat. no. 512326) and IFN-γ-APC-Cy7 (4S.B3, cat. no. 502530) from BioLegend. Stained cells were analyzed using a BD LSRFortessa (BD Biosciences), and flow cytometry data were analyzed with FlowJo software (Tree Star).

Isolation of T cell clones. CFSE^{low} cultures FACS-sorted from ex vivo stimulated memory CD4$^+$ T cells or from in vitro primed naïve CD4$^+$CD95$^-$ T cells were cloned by limiting dilution, as previously described (Latorre et al., 2018). T cell clone reactivity was determined by stimulation with untreated or antigen-pulsed irradiated autologous monocytes. Depending of the antigen, monocytes were pulsed 3-5 hrs with heat-inactivated bacteria (2 particles per monocytes), *M. tuberculosis* whole cell lysate (3 μg/ml) or recombinant OmpA (5 μg/ml). To determine MHC restriction, stimulation of T cell clones was performed in the absence or presence of blocking anti-MHC-II monoclonal antibodies produced in house from hybridoma cell lines (anti-HLA-DR, clone L243 from ATCC, cat. no. HB-55; anti-HLA-DQ, clone SPVL3 (Spits et al., 1983); anti-HLA-DP, clone B7/21 (Watson et al., 1983)). In all experiments proliferation was assessed on day 3, after incubation for 16 h with 1 μCi/ml [methyl-^{3}H]thymidine (Perkin Elmer). Data were expressed as counts per min (CPM).

TCR Vβ deep sequencing. Ex vivo-sorted total memory CD4$^+$ and CFSE^{low} fractions from antigen-stimulated memory CD4$^+$ T cell cultures (2.5-5×10^5 cells) were analyzed by deep sequencing. In brief, cells were centrifuged and washed in PBS, and genomic DNA was extracted from the pellet using QIAamp DNA Micro Kit (Qiagen), according to manufacturer's instructions. Genomic DNA quantity and purity were assessed

through spectrophotometric analysis. Sequencing of TCR Vβ CDR3 was performed by Adaptive Biotechnologies using the ImmunoSEQ assay (http://www.immunoseq.com). In brief, following multiplex PCR reaction designed to target any CDR3 Vβ fragments; amplicons were sequenced using the Illumina HiSeq platform. Raw data consisting of all retrieved sequences of 87 nucleotides or corresponding amino acid sequences and containing the CDR3 region were exported and further processed. The assay was performed at deep level for ex vivo-sorted total memory CD4$^+$ (detection sensitivity, 1 cell in 200,000) and at survey level for CFSE^{low} antigen-reactive cultures (detection sensitivity, 1 cell in 40,000). Each clonotype was defined as a unique productively rearranged TCR Vβ nucleotide sequence; data processing was done using the productive frequency of templates provided by ImmunoSEQ Analyzer V.3.0 (http://www.immunoseq.com). For each repertoire, a frequency corresponding to the top 5th percentile in the frequency-ranked list of unique clonotypes was chosen as threshold (top 5%). Chao-Jaccard overlap between pairs of TCR repertoires was calculated using R package "fossil" (Vavrek, 2011); Kruskal's Non-metric Multidimensional Scaling (NMDS) of average Chao-Jaccard overlaps was performed using R package "MASS" (Venables and Ripley, 2002).

Statistical analysis. Statistical analyses were performed using GraphPad Prism 8 software. Significance was assigned at P value < 0.05 unless stated otherwise. Specific tests are indicated in the figure legends for each comparison. Analysis of TCR Vβ repertoires was performed using R software version 3.5.1. Multiple sequence alignment was performed using CLC Genomics Workbench version 8 (QIAGEN).

7. Concluding remarks

During my doctoral studies I analyzed the human CD4$^+$ T cell responses directed against antigens of different nature, such as therapeutic antibodies, viruses and bacteria. Overall, some general considerations can be drawn from the multiple projects here presented: some are of conceptual interest, others are of practical importance.

Given the fundamental role of CD4$^+$ T helper cells in directing acquired immune responses, the enumeration of antigen-specific T cells and the identification of their cognate epitopes is of primary importance to assess immunogenicity of vaccines, pathogens or biological drugs. The integration of cellular, biochemical and informatics techniques allowed us to describe with an unprecedented level of detail the architecture of antigen-reactive human CD4$^+$ T cell repertoires. For instance, we remarked that the naïve repertoire is a "gold mine" of antigen specificities: such high richness is necessary to cope with the large variety of pathogens that the host will encounter during life. However, studying the reactivity to influenza hemagglutinin, we showed that human donors hold a wide repertoire of naïve precursors recognizing cryptic peptides that remains underexploited. Further studies are needed to evaluate if this hidden repertoire might be targeted under certain conditions, for instance in case of antigen presentation by different cell types or virus-infected cells, or if the antigen is provided in a different form or formulation. The priming of a broader repertoire of naïve precursors has the potential to increase the magnitude of the first wave of T cell help upon vaccination, and therefore could be beneficial to improve immunization strategies against viral infections.

The huge diversity of the naïve repertoire implies also the possibility that unwanted CD4$^+$ T cell responses might be elicited against exogenous proteins used as therapeutic drugs, driving anti-drug antibody responses that might neutralize their efficacy. Biological drugs such as proteins and peptides constitute a powerful class of therapeutic agents that is largely, and increasingly, used in clinics for the treatment of various diseases, as is the case of monoclonal antibodies for cell depletion or immune checkpoint blockade, or of recombinant proteins and peptides for replacement therapy of endocrine disorders (Anselmo et al., 2019; Leader et al., 2008). A problematic aspect

of biological drugs is that they might be sensed as non-self antigens by the immune system, especially in therapeutic regimens involving chronic and massive exposure, with the triggering of anti-drug immune responses that can lower their effectiveness. The integrated approach that we presented here, involving the detailed profiling of antigen presentation and CD4$^+$ T cell response, delineates a generalizable pathway to identify, and possibly defuse, the immunogenic potential of biological drugs. The optimization of biobetter drugs bearing minimal immunogenicity is important to improve their efficacy and it may benefit a large variety of clinical fields, including the treatment of autoimmunity, cancer or endocrine disorders. Besides, the comprehensive measurement of the immunogenic landscape of infectious agents, such as viruses and bacteria, can be beneficial also for the design of improved and safer subunit vaccines. By focusing on the most immunogenic antigens targeted by naturally developed immune responses, rationally designed vaccines might confer host protection while minimizing the risks of infection or toxic reactions associated with heat-killed or attenuated vaccines (Leader et al., 2008).

The data here presented offer insights about the "rules of engagement" and the mechanisms of clonal selection of CD4$^+$ T cell responses in humans. In particular, we showed that the selection of naïve precursors upon first encounter with foreign antigens is strongly influenced by antigen processing. Depending on the relative abundance of cognate peptides presented by APCs on MHC-II complexes, some naïve precursors might gain selective advantage during priming, and therefore undergo preferential clonal expansion leading to the immunodominance observed in the memory CD4$^+$ T cell compartment. In addition, we showed that the composition of the memory repertoire can be strongly skewed by T cell cross-reactivity, which has the potential of profoundly bias future responses to closely related antigens. Heterologous immunity and original antigenic sin of CD4$^+$ T cell responses are key factors that could impact immunization outcome and host protection against pathogens, and therefore should be carefully considered for the design of improved vaccination strategies.

Using MHC-II peptidomics and stimulation of T cell clones with different specificities, we showed that antibody binding can tune antigen processing, for instance by suppressing the presentation of some T cell epitopes. An implication of such finding is that the amount of Tfh help received by B cells during germinal center reactions

might be personalized for each B cell clone depending on their BCR specificity, and it could be influenced also by pre-existing serum antibodies able to form immune complexes with the antigen. Under limiting conditions, little variations in the amount of Tfh help received might result in largely different outcomes in the quality of B cell and antibody responses. This phenomenon might explain why some classes of antibodies, such as those targeting the stem region of HA, are often subdominant in the antibody response to influenza. One reason could be ascribed to the limited accessibility of the stem domain to BCR recognition (Corti and Lanzavecchia, 2013); alongside, here we showed that an anti-stem B cell clone presents a narrower HA-derived MHC-II peptidome compared to an anti-head B cell clone, therefore suggesting that B cells specific for HA stem might receive less cognate Tfh help and thus result subdominant in the overall B cell response. Further studies are needed to determine to what extent the tuning of antigen processing by antibodies impacts B cell clonal selection and affinity maturation. The identification of T cell epitopes selectively suppressed or enhanced by antigen presenting B cells depending on their BCR paratope could have profound implications for vaccinology, with the ultimate goal of directing the antibody response towards desired portions of the antigen by rational induction of Tfh responses selectively targeting certain epitopes. The development of immunization protocols to boost broadly neutralizing antibodies directed to the most tempting antigenic targets, such as conserved structures, might be extremely beneficial in the challenging quest for universal vaccines against rapidly evolving viruses such as influenza or HIV.

Overall, our studies highlight the tight link existing between the humoral and cellular arms of adaptive immunity. In particular, the data here presented remark the importance of cognate T helper cells for the induction of affinity matured, class-switched antibodies, and the intertwinement between T and B cell responses often converging on the same antigenic targets. In other words, the presence of antigen-specific CD4$^+$ T cell responses can be considered as a proxy for matured antibody responses. Likewise, serological studies aiming to identify antigenic targets of high-affinity antibodies could also be predictive of the antigens against which CD4$^+$ T cell responses are directed, with implications for vaccines design. The detailed characterization of the ground rules of T-B collaboration is important in order to actively impact, and possibly ameliorate, immune responses by rational intervention. In

the era of big data, the integrated study of antigen specificity of T and B cells, combined with multi-omics approaches, offers the potentiality to boost up next generation immunology both in the fields of infectious diseases and autoimmunity, and in many emerging areas such as cancer immunotherapy, neuroimmunology and behavioral medicine.

8. References

Abbas, A.K., A.H. Lichtman, and S. Pillai. 2012. Cellular and molecular immunology. Elsevier/Saunders, Philadelphia. x, 545 p. pp.

Abelin, J.G., D. Harjanto, M. Malloy, P. Suri, T. Colson, S.P. Goulding, A.L. Creech, L.R. Serrano, G. Nasir, Y. Nasrullah, C.D. McGann, D. Velez, Y.S. Ting, A. Poran, D.A. Rothenberg, S. Chhangawala, A. Rubinsteyn, J. Hammerbacher, R.B. Gaynor, E.F. Fritsch, J. Greshock, R.C. Oslund, D. Barthelme, T.A. Addona, C.M. Arieta, and M.S. Rooney. 2019. Defining HLA-II Ligand Processing and Binding Rules with Mass Spectrometry Enhances Cancer Epitope Prediction. *Immunity* 51:766-779 e717.

Acosta-Rodriguez, E.V., L. Rivino, J. Geginat, D. Jarrossay, M. Gattorno, A. Lanzavecchia, F. Sallusto, and G. Napolitani. 2007. Surface phenotype and antigenic specificity of human interleukin 17-producing T helper memory cells. *Nat Immunol* 8:639-646.

Adorini, L. 1998. Immunodominance. In Encyclopedia of Immunology. Elsevier, 1290-1292.

Alhashem, F., N.L. Tiren-Verbeet, E. Alp, and M. Doganay. 2017. Treatment of sepsis: What is the antibiotic choice in bacteremia due to carbapenem resistant Enterobacteriaceae? *World J Clin Cases* 5:324-332.

Angeletti, D., and J.W. Yewdell. 2018. Is It Possible to Develop a "Universal" Influenza Virus Vaccine? Outflanking Antibody Immunodominance on the Road to Universal Influenza Vaccination. *Cold Spring Harb Perspect Biol* 10:

Anselmo, A.C., Y. Gokarn, and S. Mitragotri. 2019. Non-invasive delivery strategies for biologics. *Nat Rev Drug Discov* 18:19-40.

Atarashi, K., W. Suda, C. Luo, T. Kawaguchi, I. Motoo, S. Narushima, Y. Kiguchi, K. Yasuma, E. Watanabe, T. Tanoue, C.A. Thaiss, M. Sato, K. Toyooka, H.S. Said, H. Yamagami, S.A. Rice, D. Gevers, R.C. Johnson, J.A. Segre, K. Chen, J.K. Kolls, E. Elinav, H. Morita, R.J. Xavier, M. Hattori, and K. Honda. 2017. Ectopic

colonization of oral bacteria in the intestine drives TH1 cell induction and inflammation. *Science* 358:359-365.

Bachelet, D., S. Hassler, C. Mbogning, J. Link, M. Ryner, R. Ramanujam, M. Auer, P.E. Hyldgaard Jensen, N. Koch-Henriksen, C. Warnke, K. Ingenhoven, D. Buck, V. Grummel, A. Lawton, N. Donnellan, A. Hincelin-Mery, D. Sikkema, M. Pallardy, B. Kieseier, B. Hemmer, H.P. Hartung, P. Soelberg Sorensen, F. Deisenhammer, P. Donnes, J. Davidson, A. Fogdell-Hahn, P. Broet, and A. Consortium. 2016. Occurrence of Anti-Drug Antibodies against Interferon-Beta and Natalizumab in Multiple Sclerosis: A Collaborative Cohort Analysis. *PLoS One* 11:e0162752.

Bacher, P., T. Hohnstein, E. Beerbaum, M. Rocker, M.G. Blango, S. Kaufmann, J. Rohmel, P. Eschenhagen, C. Grehn, K. Seidel, V. Rickerts, L. Lozza, U. Stervbo, M. Nienen, N. Babel, J. Milleck, M. Assenmacher, O.A. Cornely, M. Ziegler, H. Wisplinghoff, G. Heine, M. Worm, B. Siegmund, J. Maul, P. Creutz, C. Tabeling, C. Ruwwe-Glosenkamp, L.E. Sander, C. Knosalla, S. Brunke, B. Hube, O. Kniemeyer, A.A. Brakhage, C. Schwarz, and A. Scheffold. 2019. Human Anti-fungal Th17 Immunity and Pathology Rely on Cross-Reactivity against Candida albicans. *Cell* 176:1340-1355 e1315.

Basler, M., C.J. Kirk, and M. Groettrup. 2013. The immunoproteasome in antigen processing and other immunological functions. *Curr Opin Immunol* 25:74-80.

Bassani-Sternberg, M., S. Pletscher-Frankild, L.J. Jensen, and M. Mann. 2015. Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Mol Cell Proteomics* 14:658-673.

Becattini, S., D. Latorre, F. Mele, M. Foglierini, C. De Gregorio, A. Cassotta, B. Fernandez, S. Kelderman, T.N. Schumacher, D. Corti, A. Lanzavecchia, and F. Sallusto. 2015. T cell immunity. Functional heterogeneity of human memory CD4(+) T cell clones primed by pathogens or vaccines. *Science* 347:400-406.

Beck, S., D.E. Geraghty, H. Inoko, and L. Rowen. 1999. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. *Nature* 401:921-923.

Belkaid, Y., and O.J. Harrison. 2017. Homeostatic Immunity and the Microbiota. *Immunity* 46:562-576.

Berlin, C., E.L. Berg, M.J. Briskin, D.P. Andrew, P.J. Kilshaw, B. Holzmann, I.L. Weissman, A. Hamann, and E.C. Butcher. 1993. Alpha 4 beta 7 integrin mediates lymphocyte binding to the mucosal vascular addressin MAdCAM-1. *Cell* 74:185-195.

Bhan, M.K., R. Bahl, and S. Bhatnagar. 2005. Typhoid and paratyphoid fever. *Lancet* 366:749-762.

Bricogne, G., E. Blanc, M. Brandl, C. Flensburg, P. Keller, W. Paciorek, P. Roversi, A. Sharff, O.S. Smart, C. Vonrhein, and T.O. Womack. 2017. BUSTER version 2.11.7. *Cambridge, United Kingdom: Global Phasing Ltd*

Bullough, P.A., F.M. Hughson, J.J. Skehel, and D.C. Wiley. 1994. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* 371:37-43.

Busch, D.H., and E.G. Pamer. 1999. T cell affinity maturation by selective expansion during infection. *J Exp Med* 189:701-710.

Cabrera-Perez, J., S.A. Condotta, V.P. Badovinac, and T.S. Griffith. 2014. Impact of sepsis on CD4 T cell immunity. *J Leukoc Biol* 96:767-777.

Calabresi, P.A., G. Giovannoni, C. Confavreux, S.L. Galetta, E. Havrdova, M. Hutchinson, L. Kappos, D.H. Miller, P.W. O'Connor, J.T. Phillips, C.H. Polman, E.W. Radue, R.A. Rudick, W.H. Stuart, F.D. Lublin, A. Wajgt, B. Weinstock-Guttman, D.R. Wynn, F. Lynn, M.A. Panzara, Affirm, and S. Investigators. 2007. The incidence and significance of anti-natalizumab antibodies: results from AFFIRM and SENTINEL. *Neurology* 69:1391-1403.

Calfee, D.P. 2017. Recent advances in the understanding and management of Klebsiella pneumoniae. *F1000Res* 6:1760.

Campion, S.L., T.M. Brodie, W. Fischer, B.T. Korber, A. Rossetti, N. Goonetilleke, A.J. McMichael, and F. Sallusto. 2014. Proteome-wide analysis of HIV-specific naive and memory CD4(+) T cells in unexposed blood donors. *J Exp Med* 211:1273-1280.

Cella, M., A. Engering, V. Pinet, J. Pieters, and A. Lanzavecchia. 1997. Inflammatory stimuli induce accumulation of MHC class II complexes on dendritic cells. *Nature* 388:782-787.

Chataway, J., and D.H. Miller. 2013. Natalizumab therapy for multiple sclerosis. *Neurotherapeutics* 10:19-28.

Cooper, M.D., and M.N. Alder. 2006. The evolution of adaptive immune systems. *Cell* 124:815-822.

Corti, D., S. Bianchi, F. Vanzetta, A. Minola, L. Perez, G. Agatic, B. Guarino, C. Silacci, J. Marcandalli, B.J. Marsland, A. Piralla, E. Percivalle, F. Sallusto, F. Baldanti, and A. Lanzavecchia. 2013. Cross-neutralization of four paramyxoviruses by a human monoclonal antibody. *Nature* 501:439-443.

Corti, D., and A. Lanzavecchia. 2013. Broadly neutralizing antiviral antibodies. *Annu Rev Immunol* 31:705-742.

Corti, D., J. Voss, S.J. Gamblin, G. Codoni, A. Macagno, D. Jarrossay, S.G. Vachieri, D. Pinna, A. Minola, F. Vanzetta, C. Silacci, B.M. Fernandez-Rodriguez, G. Agatic, S. Bianchi, I. Giacchetto-Sasselli, L. Calder, F. Sallusto, P. Collins, L.F. Haire, N. Temperton, J.P. Langedijk, J.J. Skehel, and A. Lanzavecchia. 2011. A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. *Science* 333:850-856.

Cosmi, L., R. De Palma, V. Santarlasci, L. Maggi, M. Capone, F. Frosali, G. Rodolico, V. Querci, G. Abbate, R. Angeli, L. Berrino, M. Fambrini, M. Caproni, F. Tonelli, E. Lazzeri, P. Parronchi, F. Liotta, E. Maggi, S. Romagnani, and F. Annunziato.

2008. Human interleukin 17-producing cells originate from a CD161+CD4+ T cell precursor. *J Exp Med* 205:1903-1916.

Cox, J., N. Neuhauser, A. Michalski, R.A. Scheltema, J.V. Olsen, and M. Mann. 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 10:1794-1805.

Crotty, S. 2015. A brief history of T cell help to B cells. *Nat Rev Immunol* 15:185-189.

Davis, M.M., and P. Brodin. 2018. Rebooting Human Immunology. *Annu Rev Immunol* 36:843-864.

de la Hera, B., E. Urcelay, D. Brassat, A. Chan, A. Vidal-Jordana, A. Salmen, L.M. Villar, J.C. Alvarez-Cermeno, G. Izquierdo, O. Fernandez, B. Oliver, A. Saiz, J.R. Ara, A.G. Vigo, R. Arroyo, V. Meca, S. Malhotra, N. Fissolo, A. Horga, X. Montalban, and M. Comabella. 2014. Natalizumab-related anaphylactoid reactions in MS patients are associated with HLA class II alleles. *Neurol Neuroimmunol Neuroinflamm* 1:e47.

Delcour, A.H. 2002. Structure and function of pore-forming beta-barrels from bacteria. *J Mol Microbiol Biotechnol* 4:1-10.

Dendrou, C.A., J. Petersen, J. Rossjohn, and L. Fugger. 2018. HLA variation and disease. *Nat Rev Immunol* 18:325-339.

Dengjel, J., O. Schoor, R. Fischer, M. Reich, M. Kraus, M. Muller, K. Kreymborg, F. Altenberend, J. Brandenburg, H. Kalbacher, R. Brock, C. Driessen, H.G. Rammensee, and S. Stevanovic. 2005. Autophagy promotes MHC class II presentation of peptides from intracellular source proteins. *Proc Natl Acad Sci U S A* 102:7922-7927.

Duchmann, R., E. May, M. Heike, P. Knolle, M. Neurath, and K.H. Meyer zum Buschenfelde. 1999. T cell specificity and cross reactivity towards enterobacteria, bacteroides, bifidobacterium, and antigens from resident intestinal flora in humans. *Gut* 44:812-818.

Duhen, T., and D.J. Campbell. 2014. IL-1beta promotes the differentiation of polyfunctional human CCR6+CXCR3+ Th1/17 cells that are specific for pathogenic and commensal microbes. *J Immunol* 193:120-129.

Dunn, N., A. Juto, M. Ryner, A. Manouchehrinia, L. Piccoli, K. Fink, F. Piehl, and A. Fogdell-Hahn. 2018. Rituximab in multiple sclerosis: Frequency and clinical relevance of anti-drug antibodies. *Mult Scler* 24:1224-1233.

Dustin, M.L. 2014. The immunological synapse. *Cancer Immunol Res* 2:1023-1033.

Feasey, N.A., G. Dougan, R.A. Kingsley, R.S. Heyderman, and M.A. Gordon. 2012. Invasive non-typhoidal salmonella disease: an emerging and neglected tropical disease in Africa. *Lancet* 379:2489-2499.

Foote, J., and C. Milstein. 1991. Kinetic maturation of an immune response. *Nature* 352:530-532.

Gao, J., B.P. De, and A.K. Banerjee. 1999. Human parainfluenza virus type 3 up-regulates major histocompatibility complex class I and II expression on respiratory epithelial cells: involvement of a STAT1- and CIITA-independent pathway. *J Virol* 73:1411-1418.

Garcia, K.C., and E.J. Adams. 2005. How the T cell receptor sees antigen--a structural view. *Cell* 122:333-336.

Geiger, R., T. Duhen, A. Lanzavecchia, and F. Sallusto. 2009. Human naive and memory CD4+ T cell repertoires specific for naturally processed antigens analyzed using libraries of amplified T cells. *J Exp Med* 206:1525-1534.

Glanville, J., H. Huang, A. Nau, O. Hatton, L.E. Wagar, F. Rubelt, X. Ji, A. Han, S.M. Krams, C. Pettus, N. Haas, C.S.L. Arlehamn, A. Sette, S.D. Boyd, T.J. Scriba, O.M. Martinez, and M.M. Davis. 2017. Identifying specificity groups in the T cell receptor repertoire. *Nature* 547:94-98.

Godinez, I., A.M. Keestra, A. Spees, and A.J. Baumler. 2011. The IL-23 axis in Salmonella gastroenteritis. *Cell Microbiol* 13:1639-1647.

Graham, D.B., C. Luo, D.J. O'Connell, A. Lefkovith, E.M. Brown, M. Yassour, M. Varma, J.G. Abelin, K.L. Conway, G.J. Jasso, C.G. Matar, S.A. Carr, and R.J. Xavier. 2018. Antigen discovery and specification of immunodominance hierarchies for MHCII-restricted epitopes. *Nat Med* 24:1762-1772.

Groom, J.R., and A.D. Luster. 2011. CXCR3 in T cell function. *Exp Cell Res* 317:620-631.

Gupta, N.T., J.A. Vander Heiden, M. Uduman, D. Gadala-Maria, G. Yaari, and S.H. Kleinstein. 2015. Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* 31:3356-3358.

Hamze, M., S. Meunier, A. Karle, A. Gdoura, A. Goudet, N. Szely, M. Pallardy, F. Carbonnel, S. Spindeldreher, X. Mariette, C. Miceli-Richard, and B. Maillere. 2017. Characterization of CD4 T Cell Epitopes of Infliximab and Rituximab Identified from Healthy Donors. *Front Immunol* 8:500.

Harding, F.A., M.M. Stickler, J. Razo, and R.B. DuBridge. 2010. The immunogenicity of humanized and fully human antibodies: residual immunogenicity resides in the CDR regions. *MAbs* 2:256-265.

Hegazy, A.N., N.R. West, M.J.T. Stubbington, E. Wendt, K.I.M. Suijker, A. Datsi, S. This, C. Danne, S. Campion, S.H. Duncan, B.M.J. Owens, H.H. Uhlig, A. McMichael, I.B.D.C.I. Oxford, A. Bergthaler, S.A. Teichmann, S. Keshav, and F. Powrie. 2017. Circulating and Tissue-Resident CD4(+) T Cells With Reactivity to Intestinal Microbiota Are Abundant in Healthy Individuals and Function Is Altered During Inflammation. *Gastroenterology* 153:1320-1337 e1316.

Huppa, J.B., and M.M. Davis. 2013. The interdisciplinary science of T-cell recognition. *Adv Immunol* 119:1-50.

Jenkins, M.K., and J.J. Moon. 2012. The role of naive T cell precursor frequency and recruitment in dictating immune response magnitude. *J Immunol* 188:4135-4140.

Jensen, I.J., F.V. Sjaastad, T.S. Griffith, and V.P. Badovinac. 2018a. Sepsis-Induced T Cell Immunoparalysis: The Ins and Outs of Impaired T Cell Immunity. *J Immunol* 200:1543-1553.

Jensen, K.K., M. Andreatta, P. Marcatili, S. Buus, J.A. Greenbaum, Z. Yan, A. Sette, B. Peters, and M. Nielsen. 2018b. Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* 154:394-406.

Jensen, P.E.H., C. Warnke, K. Ingenhoven, L. Piccoli, M. Gasis, C. Hermanrud, B.M. Fernandez-Rodriguez, M. Ryner, D. Kramer, J. Link, R. Ramanujam, M. Auer, D. Buck, V. Grummel, E. Bertotti, N. Fissolo, B. Oliver-Martos, P. Nytrova, M. Khalil, M. Guger, S. Rathmaier, C. Sievers-Stober, R.L.P. Lindberg, S. Hassler, D. Bachelet, O. Aktas, N. Donnellan, A. Lawton, B. Hemmer, E.K. Havrdova, B. Kieseier, H.P. Hartung, M. Comabella, X. Montalban, T. Derfuss, F. Sellebjerg, P. Donnes, M. Pallardy, S. Spindeldreher, P. Broet, F. Deisenhammer, A. Fogdell-Hahn, P.S. Sorensen, and A. Consortium. 2019. Detection and kinetics of persistent neutralizing anti-interferon-beta antibodies in patients with multiple sclerosis. Results from the ABIRISK prospective cohort study. *J Neuroimmunol* 326:19-27.

Justesen, S., M. Harndahl, K. Lamberth, L.L. Nielsen, and S. Buus. 2009. Functional recombinant MHC class II molecules and high-throughput peptide-binding assays. *Immunome Res* 5:2.

Kim, A., and S. Sadegh-Nasseri. 2015. Determinants of immunodominance for CD4 T cells. *Curr Opin Immunol* 34:9-15.

Kleinschek, M.A., K. Boniface, S. Sadekova, J. Grein, E.E. Murphy, S.P. Turner, L. Raskin, B. Desai, W.A. Faubion, R. de Waal Malefyt, R.H. Pierce, T. McClanahan, and R.A. Kastelein. 2009. Circulating and gut-resident human Th17 cells express CD161 and promote intestinal inflammation. *J Exp Med* 206:525-534.

Koblischke, M., M.S. Mackroth, J. Schwaiger, I. Fae, G. Fischer, K. Stiasny, F.X. Heinz, and J.H. Aberle. 2017. Protein structure shapes immunodominance in the CD4 T cell response to yellow fever vaccination. *Sci Rep* 7:8907.

Kotloff, K.L., J.P. Winickoff, B. Ivanoff, J.D. Clemens, D.L. Swerdlow, P.J. Sansonetti, G.K. Adak, and M.M. Levine. 1999. Global burden of Shigella infections: implications for vaccine development and implementation of control strategies. *Bull World Health Organ* 77:651-666.

Krammer, F., A. Garcia-Sastre, and P. Palese. 2018. Is It Possible to Develop a "Universal" Influenza Virus Vaccine? Potential Target Antigens and Critical Aspects for a Universal Influenza Vaccine. *Cold Spring Harb Perspect Biol* 10:

Krishnan, S., and N.V. Prasadarao. 2012. Outer membrane protein A and OprF: versatile roles in Gram-negative bacterial infections. *FEBS J* 279:919-931.

Kurosaki, T., K. Kometani, and W. Ise. 2015. Memory B cells. *Nat Rev Immunol* 15:149-159.

Landry, S.J. 2008. Three-dimensional structure determines the pattern of CD4+ T-cell epitope dominance in influenza virus hemagglutinin. *J Virol* 82:1238-1248.

Lanier, L.L. 2015. NKG2D Receptor and Its Ligands in Host Defense. *Cancer Immunol Res* 3:575-582.

Lanzavecchia, A. 1985. Antigen-specific interaction between T and B cells. *Nature* 314:537-539.

Lanzavecchia, A. 1998. Immunology. Licence to kill. *Nature* 393:413-414.

Larkin, J., V. Chiarion-Sileni, R. Gonzalez, J.J. Grob, C.L. Cowey, C.D. Lao, D. Schadendorf, R. Dummer, M. Smylie, P. Rutkowski, P.F. Ferrucci, A. Hill, J. Wagstaff, M.S. Carlino, J.B. Haanen, M. Maio, I. Marquez-Rodas, G.A. McArthur, P.A. Ascierto, G.V. Long, M.K. Callahan, M.A. Postow, K. Grossmann, M. Sznol, B. Dreno, L. Bastholt, A. Yang, L.M. Rollin, C. Horak,

F.S. Hodi, and J.D. Wolchok. 2015. Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *N Engl J Med* 373:23-34.

Latorre, D., U. Kallweit, E. Armentani, M. Foglierini, F. Mele, A. Cassotta, S. Jovic, D. Jarrossay, J. Mathis, F. Zellini, B. Becher, A. Lanzavecchia, R. Khatami, M. Manconi, M. Tafti, C.L. Bassetti, and F. Sallusto. 2018. T cells in patients with narcolepsy target self-antigens of hypocretin neurons. *Nature* 562:63-68.

Leader, B., Q.J. Baca, and D.E. Golan. 2008. Protein therapeutics: a summary and pharmacological classification. *Nat Rev Drug Discov* 7:21-39.

Lee, J., P. Paparoditis, A.P. Horton, A. Fruhwirth, J.R. McDaniel, J. Jung, D.R. Boutz, D.A. Hussein, Y. Tanno, L. Pappas, G.C. Ippolito, D. Corti, A. Lanzavecchia, and G. Georgiou. 2019. Persistent Antibody Clonotypes Dominate the Serum Response to Influenza over Multiple Years and Repeated Vaccinations. *Cell Host Microbe* 25:367-376 e365.

Lee, S.J., L. Liang, S. Juarez, M.R. Nanton, E.N. Gondwe, C.L. Msefula, M.A. Kayala, F. Necchi, J.N. Heath, P. Hart, R.M. Tsolis, R.S. Heyderman, C.A. MacLennan, P.L. Felgner, D.H. Davies, and S.J. McSorley. 2012. Identification of a common immune signature in murine and human systemic Salmonellosis. *Proc Natl Acad Sci U S A* 109:4998-5003.

Lefranc, M.P., V. Giudicelli, C. Ginestoux, J. Jabado-Michaloud, G. Folch, F. Bellahcene, Y. Wu, E. Gemrot, X. Brochet, J. Lane, L. Regnier, F. Ehrenmann, G. Lefranc, and P. Duroux. 2009. IMGT, the international ImMunoGeneTics information system. *Nucleic Acids Res* 37:D1006-1012.

Levine, M.M., K.L. Kotloff, E.M. Barry, M.F. Pasetti, and M.B. Sztein. 2007. Clinical trials of Shigella vaccines: two steps forward and one step back on a long, hard road. *Nat Rev Microbiol* 5:540-553.

Li, H., F.H. Shi, S.Y. Huang, S.G. Zhang, and M.L. Chen. 2018. A Review on Clinical Pharmacokinetics, Pharmacodynamics, and Pharmacogenomics of Natalizumab:

A Humanized Anti-alpha4 Integrin Monoclonal Antibody. *Curr Drug Metab* 19:1213-1223.

Liao, H.X., R. Lynch, T. Zhou, F. Gao, S.M. Alam, S.D. Boyd, A.Z. Fire, K.M. Roskin, C.A. Schramm, Z. Zhang, J. Zhu, L. Shapiro, N.C.S. Program, J.C. Mullikin, S. Gnanakaran, P. Hraber, K. Wiehe, G. Kelsoe, G. Yang, S.M. Xia, D.C. Montefiori, R. Parks, K.E. Lloyd, R.M. Scearce, K.A. Soderberg, M. Cohen, G. Kamanga, M.K. Louder, L.M. Tran, Y. Chen, F. Cai, S. Chen, S. Moquin, X. Du, M.G. Joyce, S. Srivatsan, B. Zhang, A. Zheng, G.M. Shaw, B.H. Hahn, T.B. Kepler, B.T. Korber, P.D. Kwong, J.R. Mascola, and B.F. Haynes. 2013. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* 496:469-476.

Lindestam Arlehamn, C.S., A. Gerasimova, F. Mele, R. Henderson, J. Swann, J.A. Greenbaum, Y. Kim, J. Sidney, E.A. James, R. Taplitz, D.M. McKinney, W.W. Kwok, H. Grey, F. Sallusto, B. Peters, and A. Sette. 2013. Memory T cells in latent Mycobacterium tuberculosis infection are directed against three antigenic islands and largely contained in a CXCR3+CCR6+ Th1 subset. *PLoS Pathog* 9:e1003130.

Lindestam Arlehamn, C.S., and A. Sette. 2014. Definition of CD4 Immunosignatures Associated with MTB. *Front Immunol* 5:124.

Link, J., R. Ramanujam, M. Auer, M. Ryner, S. Hassler, D. Bachelet, C. Mbogning, C. Warnke, D. Buck, P.E. Hyldgaard Jensen, C. Sievers, K. Ingenhoven, N. Fissolo, R. Lindberg, V. Grummel, N. Donnellan, M. Comabella, X. Montalban, B. Kieseier, P. Soelberg Sorensen, H.P. Hartung, T. Derfuss, A. Lawton, D. Sikkema, M. Pallardy, B. Hemmer, F. Deisenhammer, P. Broet, P. Donnes, J. Davidson, A. Fogdell-Hahn, and A. Consortium. 2017. Clinical practice of analysis of anti-drug antibodies against interferon beta and natalizumab in multiple sclerosis patients in Europe: A descriptive study of test results. *PLoS One* 12:e0170395.

Malherbe, L., C. Hausl, L. Teyton, and M.G. McHeyzer-Williams. 2004. Clonal selection of helper T cells is determined by an affinity threshold with no further skewing of TCR binding properties. *Immunity* 21:669-679.

Mani, S., T. Wierzba, and R.I. Walker. 2016. Status of vaccine research and development for Shigella. *Vaccine* 34:2887-2894.

Mason, D. 1998. A very high level of crossreactivity is an essential feature of the T-cell receptor. *Immunol Today* 19:395-404.

McHeyzer-Williams, L.J., J.F. Panus, J.A. Mikszta, and M.G. McHeyzer-Williams. 1999. Evolution of antigen-specific T cell receptors in vivo: preimmune and antigen-driven selection of preferred complementarity-determining region 3 (CDR3) motifs. *J Exp Med* 189:1823-1838.

McSorley, S.J. 2014. Immunity to intestinal pathogens: lessons learned from Salmonella. *Immunol Rev* 260:168-182.

Mellins, E.D., and L.J. Stern. 2014. HLA-DM and HLA-DO, key regulators of MHC-II processing and presentation. *Curr Opin Immunol* 26:115-122.

Mettu, R.R., T. Charles, and S.J. Landry. 2016. CD4+ T-cell epitope prediction using antigen processing constraints. *J Immunol Methods* 432:72-81.

Milligan, R., M. Paul, M. Richardson, and A. Neuberger. 2018. Vaccines for preventing typhoid fever. *Cochrane Database Syst Rev* 5:CD001261.

Mirano-Bascos, D., M. Tary-Lehmann, and S.J. Landry. 2008. Antigen structure influences helper T-cell epitope dominance in the human immune response to HIV envelope glycoprotein gp120. *Eur J Immunol* 38:1231-1237.

Mohan, J.F., and E.R. Unanue. 2012. Unconventional recognition of peptides by T cells and the implications for autoimmunity. *Nat Rev Immunol* 12:721-728.

Moon, J.J., H.H. Chu, M. Pepper, S.J. McSorley, S.C. Jameson, R.M. Kedl, and M.K. Jenkins. 2007. Naive CD4(+) T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. *Immunity* 27:203-213.

Murdaca, G., F. Spano, M. Contatore, A. Guastalla, E. Penza, O. Magnani, and F. Puppo. 2016. Immunogenicity of infliximab and adalimumab: what is its role in hypersensitivity and modulation of therapeutic efficacy and safety? *Expert Opin Drug Saf* 15:43-52.

Napolitani, G., P. Kurupati, K.W.W. Teng, M.M. Gibani, M. Rei, A. Aulicino, L. Preciado-Llanes, M.T. Wong, E. Becht, L. Howson, P. de Haas, M. Salio, C.J. Blohmke, L.R. Olsen, D.M.S. Pinto, L. Scifo, C. Jones, H. Dobinson, D. Campbell, H.B. Juel, H. Thomaides-Brears, D. Pickard, D. Bumann, S. Baker, G. Dougan, A. Simmons, M.A. Gordon, E.W. Newell, A.J. Pollard, and V. Cerundolo. 2018. Clonal analysis of Salmonella-specific effector T cells reveals serovar-specific and cross-reactive T cell responses. *Nat Immunol* 19:742-754.

Neefjes, J., M.L. Jongsma, P. Paul, and O. Bakke. 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol* 11:823-836.

Nelson, R.W., D. Beisang, N.J. Tubo, T. Dileepan, D.L. Wiesner, K. Nielsen, M. Wuthrich, B.S. Klein, D.I. Kotov, J.A. Spanier, B.T. Fife, J.J. Moon, and M.K. Jenkins. 2015. T cell receptor cross-reactivity between similar foreign and self peptides influences naive cell population size and autoimmunity. *Immunity* 42:95-107.

Neuner, E.A., J.Y. Yeh, G.S. Hall, J. Sekeres, A. Endimiani, R.A. Bonomo, N.K. Shrestha, T.G. Fraser, and D. van Duin. 2011. Treatment and outcomes in carbapenem-resistant Klebsiella pneumoniae bloodstream infections. *Diagn Microbiol Infect Dis* 69:357-362.

Nimmerjahn, F., and J.V. Ravetch. 2008. Fcgamma receptors as regulators of immune responses. *Nat Rev Immunol* 8:34-47.

Nordmann, P., T. Naas, and L. Poirel. 2011. Global spread of Carbapenemase-producing Enterobacteriaceae. *Emerg Infect Dis* 17:1791-1798.

Okada, S., J.G. Markle, E.K. Deenick, F. Mele, D. Averbuch, M. Lagos, M. Alzahrani, S. Al-Muhsen, R. Halwani, C.S. Ma, N. Wong, C. Soudais, L.A. Henderson, H. Marzouqa, J. Shamma, M. Gonzalez, R. Martinez-Barricarte, C. Okada, D.T. Avery, D. Latorre, C. Deswarte, F. Jabot-Hanin, E. Torrado, J. Fountain, A. Belkadi, Y. Itan, B. Boisson, M. Migaud, C.S.L. Arlehamn, A. Sette, S. Breton, J. McCluskey, J. Rossjohn, J.P. de Villartay, D. Moshous, S. Hambleton, S. Latour, P.D. Arkwright, C. Picard, O. Lantz, D. Engelhard, M. Kobayashi, L. Abel, A.M. Cooper, L.D. Notarangelo, S. Boisson-Dupuis, A. Puel, F. Sallusto, J. Bustamante, S.G. Tangye, and J.L. Casanova. 2015. IMMUNODEFICIENCIES. Impairment of immunity to Candida and Mycobacterium in humans with bi-allelic RORC mutations. *Science* 349:606-613.

Pan, Y., S.C. Yuhasz, and L.M. Amzel. 1995. Anti-idiotypic antibodies: biological function and structural studies. *FASEB J* 9:43-49.

Pappas, L., M. Foglierini, L. Piccoli, N.L. Kallewaard, F. Turrini, C. Silacci, B. Fernandez-Rodriguez, G. Agatic, I. Giacchetto-Sasselli, G. Pellicciotta, F. Sallusto, Q. Zhu, E. Vicenzi, D. Corti, and A. Lanzavecchia. 2014. Rapid development of broadly influenza neutralizing antibodies through redundant mutations. *Nature* 516:418-422.

Paul, S., C.S. Lindestam Arlehamn, T.J. Scriba, M.B. Dillon, C. Oseroff, D. Hinz, D.M. McKinney, S. Carrasco Pro, J. Sidney, B. Peters, and A. Sette. 2015. Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes. *J Immunol Methods* 422:28-34.

Pennini, M.E., A. De Marco, M. Pelletier, J. Bonnell, R. Cvitkovic, M. Beltramello, E. Cameroni, S. Bianchi, F. Zatta, W. Zhao, X. Xiao, M.M. Camara, A. DiGiandomenico, E. Semenova, A. Lanzavecchia, P. Warrener, J. Suzich, Q. Wang, D. Corti, and C.K. Stover. 2017. Immune stealth-driven O2 serotype

prevalence and potential for therapeutic antibodies against multidrug resistant Klebsiella pneumoniae. *Nat Commun* 8:1991.

Perez-Riverol, Y., A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D.J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. Perez, J. Uszkoreit, J. Pfeuffer, T. Sachsenberg, S. Yilmaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A.F. Jarnuczak, T. Ternent, A. Brazma, and J.A. Vizcaino. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* 47:D442-D450.

Pore, D., and M.K. Chakrabarti. 2013. Outer membrane protein A (OmpA) from Shigella flexneri 2a: a promising subunit vaccine candidate. *Vaccine* 31:3644-3650.

Preite, S., D. Baumjohann, M. Foglierini, C. Basso, F. Ronchi, B.M. Fernandez Rodriguez, D. Corti, A. Lanzavecchia, and F. Sallusto. 2015. Somatic mutations and affinity maturation are impaired by excessive numbers of T follicular helper cells and restored by Treg cells or memory T cells. *Eur J Immunol* 45:3010-3021.

Qi, Q., Y. Liu, Y. Cheng, J. Glanville, D. Zhang, J.Y. Lee, R.A. Olshen, C.M. Weyand, S.D. Boyd, and J.J. Goronzy. 2014. Diversity and clonal selection in the human T-cell repertoire. *Proc Natl Acad Sci U S A* 111:13139-13144.

Quistrebert, J., S. Hassler, D. Bachelet, C. Mbogning, A. Musters, P.P. Tak, C.A. Wijbrandts, M. Herenius, S.A. Bergstra, G. Akdemir, M. Johannesson, B. Combe, B. Fautrel, S. Chollet-Martin, A. Gleizes, N. Donnellan, F. Deisenhammer, J. Davidson, A. Hincelin-Mery, P. Donnes, A. Fogdell-Hahn, N. De Vries, T. Huizinga, I. Abugessaisa, S. Saevarsdottir, S. Hacein-Bey-Abina, M. Pallardy, P. Broet, X. Mariette, and A. Consortium. 2019. Incidence and risk factors for adalimumab and infliximab anti-drug antibodies in rheumatoid arthritis: A European retrospective multicohort analysis. *Semin Arthritis Rheum* 48:967-975.

Reith, W., and B. Mach. 2001. The bare lymphocyte syndrome and the regulation of MHC expression. *Annu Rev Immunol* 19:331-373.

Robinson, J., J.A. Halliwell, J.D. Hayhurst, P. Flicek, P. Parham, and S.G. Marsh. 2015. The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res* 43:D423-431.

Roche, P.A., and K. Furuta. 2015. The ins and outs of MHC class II-mediated antigen processing and presentation. *Nat Rev Immunol* 15:203-216.

Rowshanravan, B., N. Halliday, and D.M. Sansom. 2018. CTLA-4: a moving target in immunotherapy. *Blood* 131:58-67.

Rup, B., M. Pallardy, D. Sikkema, T. Albert, M. Allez, P. Broet, C. Carini, P. Creeke, J. Davidson, N. De Vries, D. Finco, A. Fogdell-Hahn, E. Havrdova, A. Hincelin-Mery, C.H. M, H.J. PE, E.C. Jury, H. Kirby, D. Kramer, S. Lacroix-Desmazes, J. Legrand, E. Maggi, B. Maillere, X. Mariette, C. Mauri, V. Mikol, D. Mulleman, J. Oldenburg, G. Paintaud, R.P. C, N. Ruperto, R. Seitz, S. Spindeldreher, F. Deisenhammer, and A. Consortium. 2015. Standardizing terms, definitions and concepts for describing and interpreting unwanted immunogenicity of biopharmaceuticals: recommendations of the Innovative Medicines Initiative ABIRISK consortium. *Clin Exp Immunol* 181:385-400.

Sadegh-Nasseri, S., and A. Kim. 2015. MHC Class II Auto-Antigen Presentation is Unconventional. *Front Immunol* 6:372.

Sallusto, F. 2016. Heterogeneity of Human CD4(+) T Cells Against Microbes. *Annu Rev Immunol* 34:317-334.

Sallusto, F., A. Cassotta, D. Hoces, M. Foglierini, and A. Lanzavecchia. 2018. Do Memory CD4 T Cells Keep Their Cell-Type Programming: Plasticity versus Fate Commitment? T-Cell Heterogeneity, Plasticity, and Selection in Humans. *Cold Spring Harb Perspect Biol* 10:

Sallusto, F., M. Cella, C. Danieli, and A. Lanzavecchia. 1995. Dendritic cells use macropinocytosis and the mannose receptor to concentrate macromolecules in the major histocompatibility complex class II compartment: downregulation by cytokines and bacterial products. *J Exp Med* 182:389-400.

Sallusto, F., J. Geginat, and A. Lanzavecchia. 2004. Central memory and effector memory T cell subsets: function, generation, and maintenance. *Annu Rev Immunol* 22:745-763.

Sallusto, F., and A. Lanzavecchia. 1994. Efficient presentation of soluble antigen by cultured human dendritic cells is maintained by granulocyte/macrophage colony-stimulating factor plus interleukin 4 and downregulated by tumor necrosis factor alpha. *J Exp Med* 179:1109-1118.

Sallusto, F., and A. Lanzavecchia. 2009. Heterogeneity of CD4+ memory T cells: functional modules for tailored immunity. *Eur J Immunol* 39:2076-2082.

Sallusto, F., A. Lanzavecchia, K. Araki, and R. Ahmed. 2010. From vaccines to memory and back. *Immunity* 33:451-463.

Sallusto, F., D. Lenig, R. Forster, M. Lipp, and A. Lanzavecchia. 1999. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature* 401:708-712.

Sant, A.J., F.A. Chaves, S.A. Jenks, K.A. Richards, P. Menges, J.M. Weaver, and C.A. Lazarski. 2005. The relationship between immunodominance, DM editing, and the kinetic stability of MHC class II:peptide complexes. *Immunol Rev* 207:261-278.

Savage, P.A., J.J. Boniface, and M.M. Davis. 1999. A kinetic basis for T cell receptor repertoire selection during an immune response. *Immunity* 10:485-492.

Schmid, D., M. Pypaert, and C. Munz. 2007. Antigen-loading compartments for major histocompatibility complex class II molecules continuously receive input from autophagosomes. *Immunity* 26:79-92.

Schroder, B. 2016. The multifaceted roles of the invariant chain CD74--More than just a chaperone. *Biochim Biophys Acta* 1863:1269-1281.

Sercarz, E.E., P.V. Lehmann, A. Ametani, G. Benichou, A. Miller, and K. Moudgil. 1993. Dominance and crypticity of T cell antigenic determinants. *Annu Rev Immunol* 11:729-766.

Sercarz, E.E., and E. Maverakis. 2003. Mhc-guided processing: binding of large antigen fragments. *Nat Rev Immunol* 3:621-629.

Sewell, A.K. 2012. Why must T cells be cross-reactive? *Nat Rev Immunol* 12:669-677.

Sheikh, A., F. Khanam, M.A. Sayeed, T. Rahman, M. Pacek, Y. Hu, A. Rollins, M.S. Bhuiyan, S. Rollins, A. Kalsy, M. Arifuzzaman, D.T. Leung, D.A. Sarracino, B. Krastins, R.C. Charles, R.C. Larocque, A. Cravioto, S.B. Calderwood, W.A. Brooks, J.B. Harris, J. Labaer, F. Qadri, and E.T. Ryan. 2011. Interferon-gamma and proliferation responses to Salmonella enterica Serotype Typhi proteins in patients with S. Typhi Bacteremia in Dhaka, Bangladesh. *PLoS Negl Trop Dis* 5:e1193.

Sheu, C.C., Y.T. Chang, S.Y. Lin, Y.H. Chen, and P.R. Hsueh. 2019. Infections Caused by Carbapenem-Resistant Enterobacteriaceae: An Update on Therapeutic Options. *Front Microbiol* 10:80.

Shiina, T., K. Hosomichi, H. Inoko, and J.K. Kulski. 2009. The HLA genomic loci map: expression, interaction, diversity and disease. *J Hum Genet* 54:15-39.

Sieling, P.A., and R.L. Modlin. 1994. Cytokine patterns at the site of mycobacterial infection. *Immunobiology* 191:378-387.

Simpson, E.H. 1949. Measurement of Diversity. *Nature* 163:688-688.

Sorini, C., R.F. Cardoso, N. Gagliani, and E.J. Villablanca. 2018. Commensal Bacteria-Specific CD4(+) T Cell Responses in Health and Disease. *Front Immunol* 9:2667.

Spits, H., G. Keizer, J. Borst, C. Terhorst, A. Hekman, and J.E. de Vries. 1983. Characterization of monoclonal antibodies against cell surface molecules associated with cytotoxic activity of natural and activated killer cells and cloned CTL lines. *Hybridoma* 2:423-437.

Su, L.F., B.A. Kidd, A. Han, J.J. Kotzin, and M.M. Davis. 2013. Virus-specific CD4(+) memory-phenotype T cells are abundant in unexposed adults. *Immunity* 38:373-383.

Tiller, T., E. Meffre, S. Yurasov, M. Tsuiji, M.C. Nussenzweig, and H. Wardemann. 2008. Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J Immunol Methods* 329:112-124.

Traggiai, E., S. Becker, K. Subbarao, L. Kolesnikova, Y. Uematsu, M.R. Gismondo, B.R. Murphy, R. Rappuoli, and A. Lanzavecchia. 2004. An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. *Nat Med* 10:871-875.

Unanue, E.R., V. Turk, and J. Neefjes. 2016. Variations in MHC Class II Antigen Processing and Presentation in Health and Disease. *Annu Rev Immunol* 34:265-297.

Vagin, A., and A. Teplyakov. 1997. MOLREP: an Automated Program for Molecular Replacement. *Journal of Applied Crystallography* 30:1022-1025.

Valitutti, S., S. Muller, M. Cella, E. Padovan, and A. Lanzavecchia. 1995. Serial triggering of many T-cell receptors by a few peptide-MHC complexes. *Nature* 375:148-151.

van Duin, D., K.S. Kaye, E.A. Neuner, and R.A. Bonomo. 2013. Carbapenem-resistant Enterobacteriaceae: a review of treatment and outcomes. *Diagn Microbiol Infect Dis* 75:115-120.

van Schie, K.A., S. Kruithof, P.A. van Schouwenburg, A. Vennegoor, J. Killestein, G. Wolbink, and T. Rispens. 2017. Neutralizing capacity of monoclonal and polyclonal anti-natalizumab antibodies: The immune response to antibody therapeutics preferentially targets the antigen-binding site. *J Allergy Clin Immunol* 139:1035-1037 e1036.

Vavrek, M.J. 2011. Fossil: palaeoecological and palaeogeographical analysis tools. *Palaeontologia Electronica* 14:16.

Venables, W., and B. Ripley. 2002. Modern applied statistics (Fourth S., editor) New York. In Springer.

Verma, S.K., and U. Tuteja. 2016. Plague Vaccine Development: Current Research and Future Trends. *Front Immunol* 7:602.

Victora, G.D., and M.C. Nussenzweig. 2012. Germinal centers. *Annu Rev Immunol* 30:429-457.

Wagner, N., J. Lohler, E.J. Kunkel, K. Ley, E. Leung, G. Krissansen, K. Rajewsky, and W. Muller. 1996. Critical role for beta7 integrins in formation of the gut-associated lymphoid tissue. *Nature* 382:366-370.

Wang, P., J. Sidney, Y. Kim, A. Sette, O. Lund, M. Nielsen, and B. Peters. 2010. Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinformatics* 11:568.

Watson, A.J., R. DeMars, I.S. Trowbridge, and F.H. Bach. 1983. Detection of a novel human class II HLA antigen. *Nature* 304:358-361.

Watts, C., and A. Lanzavecchia. 1993. Suppressive effect of antibody on processing of T cell epitopes. *J Exp Med* 178:1459-1463.

Webster, R.G., and E.A. Govorkova. 2014. Continuing challenges in influenza. *Ann N Y Acad Sci* 1323:115-139.

Weiner, L.M., A.K. Webb, B. Limbago, M.A. Dudeck, J. Patel, A.J. Kallen, J.R. Edwards, and D.M. Sievert. 2016. Antimicrobial-Resistant Pathogens Associated With Healthcare-Associated Infections: Summary of Data Reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2011-2014. *Infect Control Hosp Epidemiol* 37:1288-1301.

Wilson, J., S. Elgohari, D.M. Livermore, B. Cookson, A. Johnson, T. Lamagni, A. Chronias, and E. Sheridan. 2011. Trends among pathogens reported as causing bacteraemia in England, 2004-2008. *Clin Microbiol Infect* 17:451-458.

Winn, M.D., C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, and

K.S. Wilson. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235-242.

Xu, R., D.C. Ekiert, J.C. Krause, R. Hai, J.E. Crowe, Jr., and I.A. Wilson. 2010. Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science* 328:357-360.

Yaari, G., M. Uduman, and S.H. Kleinstein. 2012. Quantifying selection in high-throughput Immunoglobulin sequencing data sets. *Nucleic Acids Res* 40:e134.

Yamamura, M., K. Uyemura, R.J. Deans, K. Weinberg, T.H. Rea, B.R. Bloom, and R.L. Modlin. 1991. Defining protective responses to pathogens: cytokine profiles in leprosy lesions. *Science* 254:277-279.

Yewdell, J.W. 2006. Confronting complexity: real-world immunodominance in antiviral CD8+ T cell responses. *Immunity* 25:533-543.

Yewdell, J.W., and J.R. Bennink. 1999. Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses. *Annu Rev Immunol* 17:51-88.

Yewdell, J.W., and M. Del Val. 2004. Immunodominance in TCD8+ responses to viruses: cell biology, cellular immunology, and mathematical models. *Immunity* 21:149-153.

Yu, Y., T. Schurpf, and T.A. Springer. 2013. How natalizumab binds and antagonizes alpha4 integrins. *J Biol Chem* 288:32314-32325.

Zand, M.S., A. Bose, T. Vo, M. Coppage, T. Pellegrin, L. Arend, F.E. Lee, A. Bozorgzadeh, and N. Leong. 2005. A renewable source of donor cells for repetitive monitoring of T- and B-cell alloreactivity. *Am J Transplant* 5:76-86.

Zare, N., S.H. Zarkesh-Esfahani, M. Gharagozloo, and V. Shaygannejad. 2013. Antibodies to interferon beta in patients with multiple sclerosis receiving CinnoVex, rebif, and betaferon. *J Korean Med Sci* 28:1801-1806.

Zeng, M.Y., N. Inohara, and G. Nunez. 2017. Mechanisms of inflammation-driven bacterial dysbiosis in the gut. *Mucosal Immunol* 10:18-26.

Zens, K.D., and D.L. Farber. 2015. Memory CD4 T cells in influenza. *Curr Top Microbiol Immunol* 386:399-421.

Zielinski, C.E., F. Mele, D. Aschenbrenner, D. Jarrossay, F. Ronchi, M. Gattorno, S. Monticelli, A. Lanzavecchia, and F. Sallusto. 2012. Pathogen-induced human TH17 cells produce IFN-gamma or IL-10 and are regulated by IL-1beta. *Nature* 484:514-518.

Zimmermann, K., S. Bastidas, L. Knecht, H. Kuster, S.R. Vavricka, H.F. Gunthard, and A. Oxenius. 2015. Gut commensal microbes do not represent a dominant antigenic source for continuous CD4+ T-cell activation during HIV-1 infection. *Eur J Immunol* 45:3107-3113.

9. Curriculum vitae

Personal details

Name	Antonino Cassotta
Nationality	Italian
Current address	Via Canonico Ghiringhelli 41, 6500 Bellinzona, Switzerland
Date of birth	February 8, 1989
Place of birth	Rome, Italy
Driving license	B
Email	antonino.cassotta@irb.usi.ch
Skype	ntoohm
LinkedIn	www.linkedin.com/in/antonino-cassotta
ORCID	https://orcid.org/0000-0001-8674-4294

Education and training

Mar 2014 to date

Doctoral studies under the supervision of Prof. Dr. Federica Sallusto and Prof. Dr. Antonio Lanzavecchia.

Department of Microbiology, ETH Zurich (Switzerland). Experimental work at the Institute for Research in Biomedicine. Via Vincenzo Vela 6, CH 6500 Bellinzona, Switzerland.

Feb 2011 – Feb 2014

Student Internship at Laboratory of Experimental Medicine directed by **Professor Rosa Sorrentino**

Dept. of Biology and Biotechnology "C. Darwin", University of Rome "La Sapienza". Piazzale Aldo Moro 5, 00185 Rome, Italy.

• Genetic and functional studies of spondyloarthropathies

• Differentiation of human monocyte and related human cell lines

• Cross-talk between monocytes/macrophages and NK cells

• MHC antigen presentation pathways

Nov 2011 – Oct 2013

Master's degree in Medical Biotechnology. 110 / 110 cum laude.

University of Rome "La Sapienza". Piazzale Aldo Moro 5, 00185 Rome, Italy.

• Principal subjects: Immunology, Molecular biology, Molecular oncology, Human genetics, Virology, Parasitology, Biotechnology of human reproduction, Bioinformatics, Molecular medicine,

Hematology, Embriology, Endocrinology, Molecular diagnostic and imaging

• Research thesis: "***Study of myeloid cells differentiation: modulation of molecules involved in classical and non-classical pathways of antigen presentation***". Training coordinator: Professor Rosa Sorrentino.

My Master's degree thesis focused on the differentiation of human monocyte and related human cell lines. In particular I studied the modulation of molecules of the endogenous pathway of antigen presentation in human monocyte-derived macrophages and in human cell lines.

Techniques used: Cell cultures. PBMC isolation and depletion of cellular subpopulations. Differentiation of dendritic cells and macrophages. RNA isolation and Reverse Transcription. RealTime PCR. High-Fidelity PCR. Protein extraction. Western blot. Immunoprecipitation. Flow cytometry. Cytotoxicity Test.

Oct 2008 – *Nov 2011*	**Bachelor degree in Biotechnology**. 110 / 110 cum laude.

University of Rome "La Sapienza". Piazzale Aldo Moro 5, 00185 Rome, Italy.

• Principal subjects: Genetics, Cellular biology, Biochemistry, Molecular biology, Chemistry, Physics, Immunology, General pathology, Bioinformatics

• Research thesis: "***Study of genetic association between ERAP1 gene and Ankylosing Spondylitis in Sardinia***". Training coordinator: Professor Rosa Sorrentino.

My Bachelor's degree thesis was a case-control genetic study of Ankylosing Spondylitis conducted on Sardinian population, to point out genetic association between Endoplasmic Reticulum Aminopeptidase 1 (ERAP1) SNPs and the disease.

Techniques used: Isolation of genomic DNA from cells. SNP genotyping by RFLPs and Allelic Discrimination with real-time PCR.

Sep 2003 – *Jul 2008*	**High school diploma**. 91/100.

Liceo Scientifico Statale "A.Pacinotti". Via XV Giugno, 19123 La Spezia, Italy.

Courses

- Introduction to Statistics for Biologists. February 2-4, 2016; Swiss Institute of Bioinformatics, Bellinzona (Switzerland).
- The Data Scientist's Toolbox. Completed March 30, 2015 (4 weeks course). Johns Hopkins University @ Coursera.org.

Languages

Italian   mother tongue

English   proficient user

Teaching

- Student Course (ETHZ autumn semesters 2015, 2016, 2017): *Analysis of Human T and B Cell Responses to Infectious Agents* (IRB).

- General laboratory supervision of undergraduate students (IRB).

Technical skills

- Human T and B cell culture
- PBMC isolation and sorting of cellular subpopulations
- T cell cloning by limiting dilution
- T cell library
- Differentiation of monocytes-derived dendritic cells
- Isolation of genomic DNA from cells
- SNP genotyping by RFLPs and Allelic Discrimination with real-time PCR
- RNA isolation and Reverse Transcription
- PCR
- RT-qPCR
- Western blotting
- Immunoprecipitation
- Multicolor Flow Cytometry
- TCR-sequencing (NGS and Sanger)
- Luminex
- Nanostring
- MHC-I / -II immunopeptidomics
- Quantitative proteomics of human T cells
- ATAC-seq
- Transfection of primary human T cells
- CRISPR-Cas9 genome editing of human T cells

Computer skills

- E.C.D.L. Core certification (European Computer Driving Licence)

Excellent command:
- Microsoft Office™ suite (Word™, Excel™, Powerpoint™)
- Adobe™ suite (Illustrator, Photoshop, Lightroom)
- GraphPad Prism
- Bioinformatics (Databanks usage, ClustalW, PyMol, Molegro, BioEdit, Integrative Genomics Viewer)
- Data analysis (ABI™ ImageLab, ABI™ Sequence Detection Software, ImageJ, BD CellQuest™ Pro, FlowJo)

Intermediate:
- R, R studio

Basic:
- Web development (HTML, PHP)
- Database management (MySQL)

10.   List of Publications, Presentations and Awards

Publications

- De Simone G, Mazza EMC, **Cassotta A**, Davydov AN, Kuka M, Zanon V, De Paoli F, Scamardella E, Metsger M, Roberto A, Pilipow K, Colombo FS, Tenedini E, Tagliafico E, Gattinoni L, Mavilio D, Peano C, Price DA, Singh SP, Farber JM, Serra V, Cucca F, Ferrari F, Orrù V, Fiorillo E, Iannacone M, Chudakov DM, Sallusto F, Lugli E. "CXCR3 Identifies Human Naive CD8[+] T Cells with Enhanced Effector Differentiation Potential". *J Immunol*. 2019 Dec 15;203(12):3179-3189. doi: 10.4049/jimmunol.1901072.

- **Cassotta A**, Mikol V, Bertrand T, Pouzieux S, Le Parc J, Ferrari P, Dumas J, Auer M, Deisenhammer F, Gastaldi M, Franciotta D, Silacci-Fregni C, Fernandez Rodriguez B, Giacchetto-Sasselli I, Foglierini M, Jarrossay D, Geiger R, Sallusto F, Lanzavecchia A, Piccoli L. "A single T cell epitope drives the neutralizing anti-drug antibody response to natalizumab in multiple sclerosis patients." *Nature Medicine*. 2019 Sep;25(9):1402-1407. doi: 10.1038/s41591-019-0568-2.

- Latorre D, Kallweit U, Armentani E, Foglierini M, Mele F, **Cassotta A**, Jovic S, Jarrossay D, Mathis J, Zellini F, Becher B, Lanzavecchia A, Khatami R, Manconi M, Tafti M, Bassetti CL, Sallusto F. "T cells in patients with narcolepsy target self-antigens of hypocretin neurons". *Nature*. 2018 Oct;562(7725):63-68. doi: 10.1038/s41586-018-0540-1.

- Sallusto F, **Cassotta A**, Hoces D, Foglierini M, Lanzavecchia A. "T-Cell Heterogeneity, Plasticity, and Selection in Humans". *Cold Spring Harb Perspect Biol*. 2018 Mar 1;10(3). doi: 10.1101/cshperspect.a029421.

- Stettler K, Beltramello M, Espinosa DA, Graham V, **Cassotta A**, Bianchi S, Vanzetta F, Minola A, Jaconi S, Mele F, Foglierini M, Pedotti M, Simonelli L, Dowall S, Atkinson B, Percivalle E, Simmons CP, Varani L, Blum J, Baldanti F, Cameroni E, Hewson R, Harris E, Lanzavecchia A, Sallusto F, Corti D.K. "Specificity, cross-reactivity and function of antibodies elicited by Zika virus infection". *Science*. 2016 Aug 19;353(6301):823-6. doi: 10.1126/science.aaf8505.

- Camilli G, **Cassotta A**, Battella S, Palmieri G, Santoni A, Paladini F, Fiorillo MT, Sorrentino R. "Regulation and trafficking of the HLA-E molecules during monocyte-macrophage differentiation". *J Leukoc Biol*. 2016 Jan;99(1):121-30. doi: 10.1189/jlb.1A0415-172R.

- Becattini S, Latorre D, Mele F, Foglierini M, De Gregorio C, **Cassotta A**, Fernandez B, Kelderman S, Schumacher TN, Corti D, Lanzavecchia A, Sallusto F. "Functional heterogeneity of human memory CD4[+] T cell clones primed by pathogens or vaccines". *Science*. 2015 Jan 23;347(6220):400-6. doi: 10.1126/science.1260668.

Oral presentations

- **<u>Cassotta A</u>**, Paparoditis P, Geiger R, Mettu RR, Landry SJ, Donati A, Benevento M, Foglierini M, Lewis DJM, Lanzavecchia A, Sallusto F. "Deciphering CD4$^+$ T cell immunodominance to Influenza virus". 17th International Congress of Immunology, IUIS 2019. October 19-23, 2019; Beijing (China).

- **<u>Cassotta A</u>**. "Deciphering CD4$^+$ T cell immunodominance to Influenza virus". SSAI annual congress 2019: viral immunity symposium. September 7, 2019; Universita' della Svizzera italiana, USI (Lugano, Switzerland).

- **<u>Cassotta A</u>**, Paparoditis P. "Heterologous prime-boost in human immunization". ADITEC Annual Meeting. September 18-20, 2017; Siena (Italy).

- **<u>Cassotta A</u>**, Foglierini M, Mele F, Lanzavecchia A, Sallusto F. "In depth immune profiling of human CD4$^+$ T cell response to Influenza virus". XXVIII Meeting of the Swiss Immunology PhD students in Immunology. February 8-10, 2016; Ermatingen, TG (Switzerland).

Poster presentations

- **<u>Cassotta A</u>**, Geiger R, Paparoditis P, Donati A, Benevento M, Foglierini M, Lewis DJM, Lanzavecchia A, Sallusto F. "Deciphering CD4$^+$ T cell immunodominance to Influenza virus". Gordon Research Conference in Immunochemistry & Immunobiology. June 10-15, 2018; Mount Snow in West Dover (VT, United States).

- **<u>Cassotta A</u>**, Geiger R, Paparoditis P, Donati A, Benevento M, Foglierini M, Lewis DJM, Lanzavecchia A, Sallusto F. "Dominant and cryptic epitopes in Influenza HA identified by combination of T cell libraries and peptidomics". EMBO workshop on Antigen Processing and Presentation. May 28-31, 2017; Salamanca (Spain).

- **<u>Cassotta A</u>**, Goldstein J, Mele F, Baggi Menozzi F, Sette A, Peters B, Lanzavecchia A, Sallusto F. "Heterologous immunity: Investigation of pathogen-cross-reactive TCRs at the clonal level". Gordon Research Conference in Immunochemistry & Immunobiology. June 19-24, 2016; Barga (Lucca, Italy).

- **<u>Cassotta A</u>**, Foglierini M, Mele F, Lanzavecchia A, Sallusto F. "Dissecting the human T helper cell repertoire to Influenza virus". EMBO Conference Next Gen Immunology 2016. February 14-16, 2016; Weizmann Institute of Science, Rehovot (Israel).

- **<u>Cassotta A</u>**, Foglierini M, Mele F, Lanzavecchia A, Sallusto F. "Dissecting the human T helper cell repertoire to Influenza virus". 18th International Summer School on Immunology. September 12-19, 2015; Rabac (Croatia).

Honors and awards

- Poster Award at the Gordon Research Conference in Immunochemistry & Immunobiology. June 10-15, 2018; Mount Snow in West Dover (VT, United States). Poster title: "Deciphering CD4$^+$ T cell immunodominance to Influenza virus".

- Travel grant for the EMBO Conference Next Gen Immunology 2016. February 14-16, 2016; Weizmann Institute of Science, Rehovot (Israel).

- Working scholarship at the Faculty of Pharmacy and Medicine. April 2013-December 2013; University of Rome "La Sapienza". Piazzale Aldo Moro 5, 00185 Rome, Italy.

- Working scholarship at the library of the Department of Biochemical Sciences "Rossi Fanelli". October 2010-October 2011; University of Rome "La Sapienza". Piazzale Aldo Moro 5, 00185 Rome, Italy.

- Working scholarship at the Presidency of the Faculty of Mathematical, Physical and Natural Sciences. October 2009-October 2010; University of Rome "La Sapienza". Piazzale Aldo Moro 5, 00185 Rome, Italy.