Inaugural dissertation

for

obtaining the doctoral degree

of the combined Mathematics, Engineering and Natural Sciences

of the Ruprecht-Karls-University

Heidelberg

Presented by

M.Sc. Hannah Barbara van Dijk

Born in: Düsseldorf

Oral examination: 2024-07-12

Single-cell based assessment of the human T cell response against *Plasmodium falciparum* circumsporozoite protein

Referees:

Prof. Dr. Nina Papavasiliou

Prof. Dr. Hedda Wardemann

1. Abstract

The human malaria parasite *Plasmodium falciparum* (*Pf*) is a unicellular organism that infects hepatocytes before reaching the symptom-causing blood stage. The development of a T cell vaccine that could target the liver stage and prevent disease onset remains a challenge. Circumsporozoite protein (CSP) is expressed in infected hepatocytes and elicits a protective antibody response, making it a promising vaccine target. While several CSP T cell epitopes have been described, information of T cell receptor (TCR) gene sequence features, HLA restriction and effects of vaccination on the TCR repertoire are largely unknown. Here, I assessed the human CD8 and CD4 T cell response to CSP after repeated vaccination with recombinant full-length protein compared to pre-vaccination time points at single-cell level. By *TR* gene repertoire sequencing and the generation of large numbers of transgenic T cell lines expressing representative TCRs, I tracked the clonal dynamics of the CD4 and CD8 T cells over time and assessed the target epitopes of CSP-reactive TCRs and their HLA-linkage. I demonstrate that the activated CD8 T cell pool shows limited clonal diversity at baseline. Prior to vaccination, individual clones comprised up to 10% of the sequenced repertoire and exhibit similarity to genes encoding for TCRs with known virus-reactivity. Upon CSP vaccination, activated cells with shared sequence features were recruited into the immune response. However, among these, I identified only a single CSP-reactive TCR with unique functional properties. In contrast, I identified a potent CSP-specific CD4 response against the same vaccine and find numerous TCRs targeting both novel and known CSP epitopes and defined their HLA linkage. My results demonstrate the differential immunogenicity of CSP among CD4 and CD8 T cells. Through *in vitro* stimulation combined with repertoire analysis, I address limitations in quantification of a functional and antigen-specific T cell response. The deep assessment of the TCR properties of CSP-reactive T cells is crucial for the evaluation of potent T cell responses by future vaccine approaches.

2. Zusammenfassung

Plasmodium falciparum (Pf) ist ein einzelliger Parasit und häufigster Verursacher der Malaria. Die Parasiten werden im Sporozoitenstadium von Mücken auf den Menschen übertragen und infizieren innerhalb kurzer Zeit Hepatozyten in der Leber, bevor der Parasit das symptomverursachende Blutstadium erreicht. Für die immunologische Bekämpfung des frühen Parasitenstadiums sind Pf-spezifische Antikörper, CD4 T-Helferzellen und cytotoxische CD8 T-Zellen notwendig, indem sie die Infektion der Hepatozyten verhindern oder infizierte Zellen eliminieren. Die Entwicklung eines Impfstoffs, der starke humorale und zelluläre Immunantworten hervorruft und eine langanhaltenen Schutz bietet, bleibt eine Herausforderung. Das Circumsporozoitenprotein (CSP) wird auf Sporozoiten und in infizierten Hepatozyten exprimiert und löst eine schützende Antikörperreaktion aus, was es zu einem vielversprechenden Impfstoffziel macht. Während viele Studien die Dynamiken der Antikörperantworten charakterisiert haben, sind die Auswirkungen der Impfung auf das T-Zell-Repertoire weitgehend unbekannt. In dieser Studie habe ich die Kinetik menschlicher CD8- und CD4-T-Zellen auf CSP nach wiederholter Impfung mit rekombinantem CSP im Vergleich zu den Zeitpunkten vor der Impfung auf Einzelzellniveau untersucht. Durch die Sequenzierung des T-Zell-Rezeptoren (TCR), konnte ich die klonale Dynamik der CD4- und CD8-T-Zellen im Laufe der Zeit verfolgen und die Zielepitope der CSP-reaktiven TCRs und ihre HLA-Verknüpfung identifizieren. Die Analyse verdeutlicht, dass aktivierte CD8-T-Zellen zu Beginn eine begrenzte klonale Vielfalt aufweisen. Vor der Impfung machten einzelne Klone bis zu 10 % des sequenzierten Repertoires aus und wiesen Ähnlichkeiten mit TCR Genen auf, die für TCRs mit bekannter Virusreaktivität kodieren. Nach der CSP-Impfung wurden aktivierte Zellen mit gemeinsamen Sequenzmerkmalen für die Immunantwort rekrutiert. Unter diesen konnte ich jedoch nur einen CSP-reaktiven TCR mit einzigartigen funktionellen Eigenschaften identifizieren. Im Gegensatz dazu konnten wir eine starke CSP-spezifische CD4-Antwort gegen denselben Impfstoff nachweisen und zahlreiche TCRs finden, die sowohl auf neue als auch auf bekannte CSP-Epitope abzielen und deren HLA-Bindung definieren. Unsere Ergebnisse belegen die unterschiedliche Immunogenität von CSP bei CD4- und CD8-T-Zellen. Durch In-vitro-Stimulation in Kombination mit einer Repertoire-Analyse können wir die Einschränkungen bei der Quantifizierung einer funktionellen und antigenspezifischen T-Zell-Antwort überwinden. Die eingehende Bewertung der TCR-Eigenschaften von CSP-reaktiven T-Zellen ist entscheidend für die Bewertung potenter T-Zell-Antworten durch zukünftige Impfstoffansätze.

3. Table of contents

1. Abstract 5
2. Zusammenfassung 7
3. Table of contents 8
4. Table of figures 12
5. Table of Supplementary figures 14
6. Abbreviations 15
7. Introduction 17
7.1. The adaptive immune system 17
7.2. T cell development 17
7.3. The TCR-pMHC complex – activation, signaling and differentiation 19
7.4. Antigen processing and presentation 20
7.5. The human TCR repertoire 22
7.6. Immunity to malaria-causing *Pf* 23
7.7. T cell response to *Pf*CSP 25
7.8. Methodologies to study antigen-specific T cells 27
8. Objectives 30
9. Methods 31
9.1. Human FMP013 immunization trial 31
9.1.1. Thawing PBMCs 31
9.1.2. *Ex vivo* T cell expansion 31
9.2. T cell receptor repertoire sequencing 32
9.2.1. Fluorescence activated cell analysis and single cell sorting 32
9.2.2. T cell receptor amplification 33
9.2.3. Next generation sequencing 36
9.2.4. Repertoire analysis 37
9.3. T cell receptor cloning 38

9.3.1. Specific amplification of *TR* genes 38
9.3.2. Restriction digestion of vector and linker segment 39
9.3.3. DNA assembly 39
9.3.4. Transformation into competent bacteria 40
9.3.5. Screening of bacterial colonies by PCR and sequence validation 40
9.3.6. Preparation of vector DNA 41
9.4. T cell receptor expression 41
9.4.1. Culture of Phoenix Ampho cells 42
9.4.2. Culture of Jurkat76 cells 42
9.4.3. Transfection of Phoenix Ampho cells 42
9.4.4. Retroviral transduction of Jurkat76 cells 43
9.4.5. Detection of TCR expression in transduced Jurkat76 cells 43
9.5. Functional characterization of T cell receptors 43
9.5.1. B cell Immortalization 43
9.5.2. T cell stimulation 44
9.5.3. Detection of T cell activation 44
9.5.4. *In-vitro* HLA-blocking experiments 45
9.6. MHC peptide binding predictions 45
9.7. *In vitro* peptide MHC binding assay 45
9.8. Peptide *in vitro* expansion and dextramer staining 45
9.9. HLA typing 46
9.10. Statistical analysis 46
10. Results 47
10.1. Human CSP immunization against malaria 47
10.2. Full-length CSP vaccination induces polyclonal CD4 T cell response to C- and N- terminus 47
10.2.1. Identification of TCRs targeting N- and C-terminal epitopes 51
10.2.2. The majority of CSP epitopes are targeted in an HLA-DR context 52

10.2.3. Groups of convergent TCRs target individual CSP epitopes 53

10.2.4. Epitopes outside Th2R/T* region reveal HLA restriction with low prevalence 56

10.2.5. Epitopes outside the Th2R/T* region are highly conserved across different *Pf* isolates 57

10.2.6. The activated CD8 T_{EM} and T_{EMRA} pool is stable over subsequent FMP013 vaccinations 57

10.2.7. The activated CD8 repertoire is dominated by large pre-existing clones..... 58

10.2.8. Activated T cells contain many TCRs with reported reactivities to common viruses ... 60

10.2.9. Convergent clones appear with higher frequency after vaccination..... 61

10.2.10. The post-vaccination repertoire reveals donor-specific enrichment of V segments 62

10.2.11. TCRs that showed signs of antigen-driven selection are not CSP-specific..... 64

10.2.12. CSP peptide expansion induces CD4 dependent bystander activation among CD8 T cells 66

10.2.13. Functional characterization of expanded clones reveals only one CSP-specific TCR 69

10.2.14. TCR B11 targets CS.T3 in a co-receptor independent manner 70

10.2.15. *In silico* peptide binding predictions reveal shared CSP binders across diverse HLA alleles 74

10.2.16. *In vitro* binding shows diverse binding pattern across distinct HLA alleles 75

10.2.17. Strong HLA binders show limited immunogenicity in vaccinated individuals 76

11. Discussion..... 79

11.1. CD4 T cells can target novel N-terminal CSP epitopes..... 79

11.2. CD4 T cell responses are genetically restricted to MHC class II genes with variable prevalence in humans..... 80

11.3. Polymorphic hotspots in T cell epitopes associate with immune pressure by high immunoprevalence..... 81

11.4. N-terminal CSP epitopes were previously associated with an MHC..... 82

class I restricted response..... 82

11.5. Limited immunogenicity of CSP vaccination in CD8 T cell response..... 83

11.6. Antigen-independent activation of CD8 T cells during *in vitro* stimulation..... 85

11.7. CD4-CD8 shared clone can target CSP in a co-receptor independent manner 86
11.8. Limited association of HLA peptide binding and peptide immunogenicity 87
11.9. High requirements for liver stage-specific T cell responses 88
12. Outlook 91
13. References 92
14. Supplement 102
14.1. Supplementary figures 102
14.2. Supplementary tables 106
14.3. Primer sequences 113
14.4. Supplementary Material 121
14.4.1. Antibodies 121
14.4.2. FACS analysis antibodies 122
14.4.3. Peptide pool 123
14.4.4. Bacteria 126
14.4.5. Buffers, solutions and chemicals 126
14.4.6. Cell lines 128
14.4.7. Cell culture media 128
14.4.8. Commercial kits 128
14.4.9. Enzymes and additives 129
14.4.10. RT and PCRs 130
14.4.11. Nucleotides and nucleic acids 130
14.4.12. Instruments and consumables 131
14.4.13. Software 133
14.4.14. Web Resources 134
15. Acknowledgement 135

4. Table of figures

Figure 1: Thymic T cell development. 19
Figure 2: Somatic V(D)J arrangement in the alpha and beta TCR chain. 20
Figure 3: Lifecycle of malaria causing Pf alternating between human and Anopheles mosquito. 25
Figure 4: Schematic representation of CSP. 27
Figure 5: Schematic overview of the human malaria vaccination trial FMP013 (Falciparum Malaria Protein 013). 47
Figure 6: Full-length CSP vaccination induces polyclonal CD4 T cell response to the C- and N- terminus. 50
Figure 7: TCRs target N-terminal, junctional and C-terminal epitopes. 51
Figure 8: N and C-terminal epitopes are predominantly presented in the context of HLA-DR molecules. 53
Figure 9: Donor-specific V segments link to C-terminal epitope specificity. 54
Figure 10: N-terminal epitopes are targeted by polyclonal T cells restricted to one donor. 55
Figure 11: N-terminal epitopes are targeted by low frequent HLA-alleles. 56
Figure 12: Epitopes outside the Th2R region are highly conserved. 57
Figure 13: The activated TEM and TEMRA pool is stable throughout subsequent FMP013 vaccinations. 58
Figure 14: The post-vaccination T cell repertoire is dominated by large preexisting clones. 60
Figure 15: TCRs of large pre-existing clones can be linked to reactivity against common viruses. 61
Figure 16: The post-vaccination repertoire contains higher frequency of convergent clones compared to pre-vaccination. 62
Figure 17: The post-vaccination repertoire reveals donor-specific enrichment of V segments. 64
Figure 18: Enriched sequence features and clonal expansion of post-vaccination T cell clones are independent of CSP reactivity. 66
Figure 19: CSP peptide expansion induces CD4-dependent bystander activation in the CD8 T cell compartment. 68
Figure 20: Low number of CSP-specific TCRs among peptide-expanded CD8 T cell clones. 70
Figure 21: TCR B11 targets CS.T3 in a co-receptor independent manner. 71
Figure 22: The minimal epitope EKCSSVFNVV is targeted in an HLA-DP restricted context. 72
Figure 23: CS.T3 core peptide KMEKCSSVF strongly binds HLA-E*01 molecule. 73
Figure 24: Diverse HLA alleles reveal shared peptide binding along CSP sequence. 75
Figure 25: Th2R and Th3R reveal strongest binding across different HLA supertypes. 76

Figure 26: Th3R epitope induces potent CD8 response with high proliferative capacity in vitro. 78

5. Table of Supplementary figures

Supplementary figure 1: CD4 TCR cloning selection criteria. 102
Supplementary figure 2: Gating strategy of TCR-transgenic Jurkat76 cell lines. 102
Supplementary figure 3: Immune cell frequencies in vaccinated individuals remain stable throughout FMP013 vaccination. 103
Supplementary figure 4: Constant TCR alpha and beta amplification efficiency across timepoints and donors. 103
Supplementary figure 5: Public T cell clones across donors appear at low frequency. 104
Supplementary figure 6: Donor-specific CDR3 length distribution changes in response to FMP013 vaccination. 104
Supplementary figure 7: HLA-E*03:01 dextramer staining of TCR B11 expression Jurkat cell line.... 104
Supplementary figure 8: Broad HLA I haplotype distribution in FMP013 donors. 105

6. Abbreviations

AIM ..... activation induced marker
ALFQ ..... Army liposome formulation containing QS-21
APC ..... antigen presenting cell
CD ..... cluster of differentiation
CDR ..... complementary determining region
CHMI ..... controlled human malaria infection
CIP ..... calf-intestinal-phosphatase
CMV ..... Cytomegalovirus
CSP ..... circumsporozoite protein
CTL ..... cytotoxic T lymphocyte
DN ..... double negative
DNA ..... desoxyribonucleic acid
DP ..... double positive
DTT ..... dithiothreitol
EBV ..... Epstein-Barr-virus
E. coli ..... *Escherichia coli*
ELISA ..... enzyme-linked immunosorbent assay
ELISpot ..... enzyme linked immunosorbent spot assay
FACS ..... fluorescence-activated cell sorting
FBS ..... fetal bovine serum
FMP ..... Falciparum Malaria Protein
Pf ..... *Plasmodium falciparum*
GPI ..... glycosylphosphatidyl-inositol
GLIPH ..... grouping of lymphocyte interaction by paratope hotspots
HBS ..... HEPES buffered saline
HLA ..... human leucocyte antigen
ICS ..... intracellular cytokine staining
IFN-γ ..... interferon-γ
LB ..... lysogeny broth
LTRs ..... long terminal repeats
MFI ..... mean fluorescent intensity
MHC ..... major histocompatibility complex

mRNA ..... messenger RNA
OD ..... optical density
PAC ..... puromycin resistance gene N-acetyltransferase
PBMC ..... peripheral blood mononuclear cell
PCR ..... polymerase chain reaction
PD-1 ..... programmed cell death 1
PEV ..... pre erythrocytic stage vaccines
PTKs ..... protein tyrosine kinases
SS ..... signal sequence
TAP ..... transported associated with antigen processing
T_{CM} ..... central memory T cells
TCR-pMHC ..... T cell receptor-peptideMHC
TCR ..... T cell receptor
TCRdist ..... TCR distance metrics
T_{EM} ..... effector memory
T_{EMRA} ..... terminally differentiated effector memory T cell
TLR ..... toll-like receptor
TRA ..... T cell receptor alpha
TRB ..... T cell receptor beta
VDJ ..... variable diversity joining
RHP ..... random hexamer primers
RNA ..... ribonucleic acid
RT ..... reverse transcription
SARS-CoV-2 ..... severe acute respiratory syndrome coronavirus type 2
WHO ..... World Health Organization
7AAD ..... 7-aminoactinomycin D

7. Introduction

7.1. The adaptive immune system

The human immune system protects the body from infections by pathogens through both an innate and adaptive immune response. While the innate immune system reacts within hours to fight invading pathogens, the adaptive immune reaction is established after several days with high specificity to the encountered antigen. The induction and maintenance of adaptive immune responses against pathogens, as well as homeostasis and memory formation, are exerted by B and T cells. B cells produce antibodies, which can recognize and bind specific antigens, forming the humoral compartment of the adaptive immune system. On the other hand, cellular immunity is mediated by T cells, which directly interact with infected or abnormal cells. The high specificity of T cells is ensured by their surface T cell receptor (TCR), that binds the pathogenic antigen and mediates activation and downstream signaling. After rapid expansion in response to antigenic stimulation, the pathogen-specific T cell response diminishes over time and immunological memory is formed. Memory T cells are antigen-experienced cells, that can respond rapidly to the same antigen upon re-exposure. To exert a variety of immunological functions, T cells comprise two lineages with distinct functional capacities: CD4 T helper cells and CD8 cytotoxic T cells.

CD4 T cells recognize peptides presented by major histocompatibility complex class II (MHC) on the surface of antigen presenting immune cells (APC). This recognition is mediated by the surface TCR and the associated co-receptor CD4, which binds to the MHC molecule. Upon activation, CD4 T cells release signaling molecules, cytokines, in order to assist other immune cells, such as B cells and macrophages. In contrast, CD8 T cells recognize MHC class I presented peptides on all nucleated cells and exert direct cytolytic activity against infected cells. To mediate efficient binding to the MHC class I molecules, CD8 T cells are equipped with MHC class I binding protein CD8.

7.2. T cell development

The thymus is the primary site of T cell development, where both lineages originate from a common bone marrow progenitor lacking both TCR and co-receptor expression, termed double negative (DN) thymocytes [1]. Thymocytes undergo four distinct differentiation stages, characterized by CD44 and CD25 surface expression, which involve *TRA* and *TRB* gene rearrangements (Figure 1). The TCR consists of two chains, alpha and beta, which together form a heterodimer. Both chains are encoded by a gene locus consisting of several segments, that are recombined during T cell development, producing a TCR with defined specificity. During *TRB* recombination one from each of the multiple variable (V), diversity

(D) and, joining (J) gene segments is selected and joined together in two consecutive steps, generating high combinatorial diversity (Figure 2a). The resulting junction region is characterized by a hypervariable region, which is formed through the removal and addition of random nucleotides, further increasing the TCR's diversity [2]. After successful DNA recombination of the TCR beta chain locus, the pre-TCR is assembled, consisting of a pre-TCR alpha chain, which is encoded by a non-rearranging locus, paired with a recombined pre-beta chain. Pre-TCR expressing cells undergo several cell divisions and subsequent recombination of the alpha locus producing the mature alpha-beta antigen receptor. While the mature TCR is displayed at low levels, expression of the pre-TCR is lost at this stage. Simultaneously, CD4 and CD8 co-receptor expression is upregulated to form a large population of double positive (DP) TCR-expressing immature cells. Next, DP thymocytes undergo a selection process, that is dependent on the TCR interaction with self-peptide-MHC-ligands to avoid autoimmunity. During TCR-mediated selection, 90% of thymocytes undergo death by neglect due to poor interaction with available self MHC ligands and associated inadequate signaling [3, 4]. Through further interaction with cortical epithelial cells, TCRs with a strong interaction with a self-ligand are eliminated by apoptotic death, termed negative selection [5]. Only recognition with intermediate intensity result in positive selection and lineage-specific differentiation into either CD4 or CD8 T cells. The commitment of the immature thymocyte to the single positive CD4 or CD8 lineages depends on the self-MHC class specificity of the expressed TCR [6]. Through the processes of positive and negative selection the receptor functionality is ensured and concurrently the recognition of self-antigens that could cause autoimmune pathology is avoided [7].

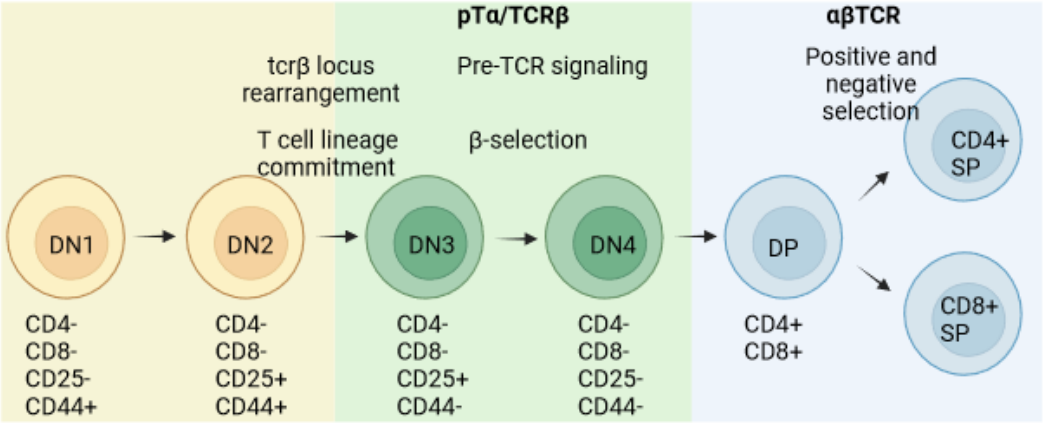


Figure 1: Thymic T cell development.

Early committed T cells lack expression of the T cell receptor, CD4 and CD8 and are termed double negative (DN) thymocytes (yellow). DN undergo four successive differentiation stages, defined by characteristic expression of surface markers. Upon rearrangement of the TCR beta chain, DN cells express the pre-TCR (light green). During the transition to double positive (DP) thymocytes, the pre-TCR alpha chain is replaced with a newly rearranged TCR alpha chain. Through interaction with cortical epithelial cells DP cells are exposed to positive and negative selection based on their interaction strength with self-peptides bound to MHC class I and class I molecules.

7.3. The TCR-pMHC complex – activation, signaling and differentiation

After maturation in the thymus, naïve single positive T cells are released into the periphery, where they continuously screen lymphoid and peripheral tissue for antigens. The TCR ligands are short peptides bound to either MHC class I or MHC class II complexes for CD4 and CD8 T cells, respectively. The TCR peptide MHC (TCR-pMHC) interaction is the main driver of T cell activation, which allows T cell proliferation and differentiation to ensure its effector function. This interaction is structurally mediated by three complementary determining regions (CDR) located in each chain of the TCR (Figure 2, B-C). While the CDR1 and CDR2 regions are involved in TCR-MHC interaction, the CDR3 primarily interacts with the antigenic peptide [8].

The heterodimeric TCR forms a multiprotein complex with four CD3 chains, which mediate signaling after recognition of the cognate antigen. In addition, the receptor associates with its co-receptor in order to facilitate MHC interaction and trigger downstream signaling, mediated through their extracellular and intracellular domain, respectively (Figure 2C). This process involves receptor aggregation and conformational changes, which initiates phosphorylation of tyrosine residues by protein tyrosine kinases (PTKs). This serves as docking site for signaling molecules that translate biochemical signals into transcription factor activation and gene activation [9].

After antigen-mediated activation, naïve CD8 T cells can either differentiate into effector memory (T_{EM}) or central memory (T_{CM}) T cells with unique functional properties. While T_{EM} cells exhibit high proliferative potential and potent cytolytic function, T_{CM} cells are equipped to persist over time following the infection. The surface expression of CCR7 on T_{CM} facilitates homing to secondary lymphoid organs, where cells persist until antigen encounter upon reinfection to maintain long-term immunity. On the other hand, T_{EM} cells are recruited to inflamed tissue through the expression of chemokine receptors. As effector CD8 T cells expand and differentiate, CD45RA expression is upregulated, classifying them as terminally differentiated cells (T_{EMRA}) [10].

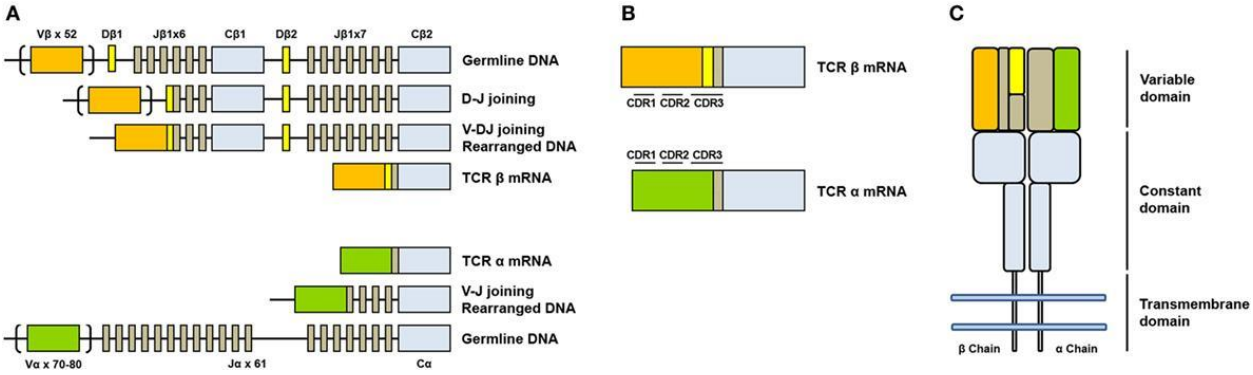


Figure 2: Somatic V(D)J arrangement in the alpha and beta TCR chain.

A Genomic organization and somatic recombination of the TCRB and TCRA loci. **B** Productive arrangements of beta and alpha transcripts **C** Organization of the TCR composed of alpha and beta subunit, each organized in a constant and variable region. Image taken from [11].

7.4. Antigen processing and presentation

TCRs are restricted to recognize peptides presented by MHC class I or class II molecules. Peptides bound to MHC class I molecules are predominantly derived from intracellular proteins, presented on all nucleated cells in order to alert the immune system to ongoing intracellular infection. During homeostasis, only self-derived endogenous antigen is processed within the cytoplasm of the cell and bound to MHC molecules through internal transportation by the transporter associated with antigen processing (TAP). In the absence of infection, CD8 T cells that scan the MHC class I molecules on the cell surface do not receive activation signals and maintain a resting state. Upon infection with intracellular infectious agents, microbe-derived peptides are presented, which can be recognized by antigen-specific CD8 T cells.

MHC class I complexes are formed of a heavy chain and non-covalently associated light chain called ß2-microglobulin. The heavy chain consists of three domains (α1, α2, α3): α1 and α2 form the peptide-binding groove, while α3 anchors the molecules to the cell membrane. To stabilize the structure of

the HLA class I molecule, the ß2-microglobulin associates with the α3 domain of the heavy chain. The peptide-binding groove accommodates only short peptides, typically 8-10 amino acids in length, due to its closed structure. Peptide binding to MHC molecules is mediated through anchor residues located at amino acid positions 2 and 8-10 of the presented peptide [12]. The MHC class I complex is encoded by HLA-A, HLA-B and and HLA-C genes located on chromosome 6 with high polymorphism in humans. The extensive polymorphism of these genes is concentrated in the peptide-binding groove of HLA molecules, which translates into distinct peptide-binding preferences of each allele. The presence of two alleles per locus enables the expression of up to six different HLA-I molecules by one individual, which facilitates the presentation of a diverse range of antigenic peptides. In addition to peptide-binding preferences by a single HLA molecule, intracellular processing pathways including steps like proteasomal cleavage and TAP transporter binding also exhibit preferences along the amino acid sequence of an antigen [13].

In contrast, an MHC class II mediated immune response combats extracellular infection through presentation by antigen presenting cells (APC) including B cells, dendritic cells and macrophages. Through phagocytosis and endocytosis, antigens derived from extracellular pathogens are engulfed by APCs and proteolytically processed. In the endosomal compartment, degraded antigen-derived peptides with high affinity to the MHC molecule are loaded and transported to the cell surface of APCs, where they are displayed for recognition by CD4 T cells. In contrast to MHC class I molecules, MHC class II complexes are alpha-beta chain heterodimers. Both chains consist of two domains with the peptide-binding groove formed by α1 and β1 domain and anchoring to the cell membrane through the α2 and β2 domain. The open structure of the peptide-binding groove facilitates binding of longer peptides of up to 25 amino acids [14]. The difference in structure allows a greater variety of peptides to be bound to MHC class II molecules compared to a narrow peptide repertoire presented by MHC class I molecules. The MHC class II molecules are encoded by four HLA genes: HLA-DR1, HLA-DR3/4/5, HLA-DQ and HLA-DP, linking to eight distinct HLA-II molecules expressed by one individual. All four HLA class II genes exhibit large polymorphisms in the beta chain locus and only a small number of alpha chain alleles among the world's population.

In the human population, thousands of MHC gene variants, alleles, exist that translate into distinct allotypes. The generation and maintenance of the huge diversity provides clear evidence for the role of natural selection. It is believed that the high level of polymorphism results from evolutionary pressures by pathogens, driving the diversification of peptide-binding specificities to evade immune detection. As MHC alleles differ in their peptide-binding profiles, susceptibility to disease is affected and drives selection in natural populations [15]. Furthermore, high levels of heterozygosity in the human

population can be explained by the advantage of heterozygous individuals being capable of presenting a larger variety of peptides, which would enhance TCR diversity and protection against pathogens.

Among the three HLA class I genes, the HLA-B gene locus exhibits the highest degree of polymorphism with nearly 4000 described alleles to date. Accordingly, HLA-B alleles reveal highest number of associations with susceptibility to infection and disease severity [16]. Among the class II gene locus, the highest number of alleles of the HLA-DR locus associates with higher levels of cell surface presentation and increased immunogenicity of HLA-DR presented peptides [17, 18]. Large expression differences of HLA alleles on the surface affect the development and frequency of potentially reactive TCRs [19]. Therefore, immunodominant epitopes are often correlated with a high level of surface expression and a broad loading capacity of several high-frequency HLA alleles [20, 21].

In addition to HLA surface expression levels, the immunogenicity of a peptide is dependent on several factors including protein expression, antigen processing, transport, peptide-MHC affinity and stability of the resulting complex [22]. As the binding of the antigenic peptide to the MHC molecule is the most selective step in the antigen presentation pathway, MHC binding prediction tools can be a potent instrument to predict possible specificity of a T cell response. These algorithms are trained on *in vitro* binding affinity measurements and mass spectrometry eluted ligands and have progressed in predictive performance in recent years [22, 23]. However, as training data is skewed towards high prevalent MHC types, performance diverges with allele frequency and peptide length.

7.5. The human TCR repertoire

The key molecules in adaptive immune responses are antigen receptors, which are found on the surface of B and T cells. The high specificity and sensitivity of T cells is achieved through genetic recombination during thymic development (Figure 1, Figure 2), resulting in the generation of a huge number of receptor variants capable of recognizing any potential self or foreign antigen. The T cell repertoire of a single individual reflects the vast number of T cell receptors present in the body. The estimates of possible combinations produced by VDJ recombination by far exceed the number of distinct antigen receptors in the human body explaining the uniqueness of TCR repertoires of unrelated individuals. Besides initial recombination events, the peripheral repertoire is shaped by positive and negative selection during immune cell maturation driven by the individual's HLA allotype (Figure 1). Throughout a person's life, the T cell repertoire undergoes dynamic changes through constant interactions with the environment. Encounters with infectious agents, commensal microbiota, and food components initiate the expansion of particular T cells bearing the same TCR. Accordingly, the absence of an antigen leads to the contraction of antigen-specific clones.

Despite the high diversity, unrelated individuals share a considerable number of TCR sequences, called public TCRs. Their emergence has been attributed to selection biases due to encountered antigens as well as convergent recombination [24]. The recombination process favors certain sequences, which are thus likely to be found more frequently in individuals [25]. This is further supported by the broad distribution of clone sizes found among the naïve T cell repertoire [26]. In addition, selection and expansion induced by common antigen exposure is needed to explain the quantity of public clones observed in a large cohort of donors [27]. These clones are of particular interest when it comes to the development of vaccines and therapeutics with good population-wide efficacy. Furthermore, also T cell clones with distinct nucleotide sequences, but similar or identical amino acid sequences in their CDRs can bind the same epitope, which are referred to as convergent TCRs. These TCRs target the same or similar antigenic epitope presented by MHC molecules, suggesting convergent evolution of T cell responses to specific antigens. The diversity and sequence patterns among TCRs targeting one epitope are reflected in an epitope-specific TCR repertoire. Several clustering algorithms try to capture and predict TCR-pMHC interaction based on weighted similarity scores for CDR1, 2, and 3 regions and V segment usages. The most commonly used and advanced tools include the grouping of lymphocytes interaction by paratope hotspots (GLIPH) algorithm and TCR distance metrics (TCRdist) tool. [28, 29]. However, the ability to predict antigen recognition and associated T cell activation based on the antigen and TCR sequences is still a key challenge due to the lack of training data and adequate models.

In addition to the identification of convergent TCRs, recent advances in high-throughput TCR sequencing urges a quantitative resolution of the immune repertoire. A large number of computational methods have been developed and applied to quantify diversity, repertoire architecture, and molecular convergence [30]. Different mathematical indices are used to interpret and compare immune repertoires by parameterizing the repertoire space [31]. Popular diversity indices in the immune repertoire field include Shannon index, Simpson index, and species richness with distinct influence of more abundant clones, which are used and compared for diversity discrimination.

7.6. Immunity to malaria-causing *Pf*

The malaria-causing agent *Pf* was responsible for 249 million cases and 608 000 deaths in 2022 and remains a major global health burden. Despite disease control and elimination efforts, efficient vaccine development is difficult, and anti-malarial drug resistance has evolved against nearly all available therapeutics [32]. The adaptability to changing environments, as well as parasite surface antigenic variation, are major reasons for the persistence of the parasite in many geographic regions [33]. The complex life cycle of *Pf* includes alternation between the mosquito and the human host. All

Plasmodium parasites enter the human host in the sporozoite stage through an infective bite by the Anopheles mosquito, which initiates the development of a pre-erythrocytic stage (Figure 3A). After sporozoites migrate from the skin to blood vessel, they invade hepatocytes in the human liver (Figure 3B), where tens of thousands merozoites are produced through a process called schizogony. After 6-7 days merozoites are released into the bloodstream through hepatocyte rupture. Upon invasion into red blood cells, merozoites initiate an asexual development phase including ring-, trophozoite-and schizont stages. Rupturing of the infected cell after ~48 h releases 16-22 merozoites, which are able to invade another un-infected erythrocyte restarting the asexual replication cycle (Figure 3C) [34]. Iterative rounds of replication ensure rapid expansion. During infection a subset of asexual replicating parasites of approximately 0.54 % develops into gametocytes [35], the sexual form of the parasite. The development into mature gametocytes includes five stages over the course of 8-12 days. Only mature stages are found in circulation, while immature stages sequester in host tissue, especially the bone marrow, to avoid clearance by the spleen [36, 37]. In a subsequent mosquito bite the female Anopheles mosquito ingests the gametocytes into its midgut (Figure 3D), where a temperature drop to ~32°C induces further maturation to gametes. Male gametes exflagellate and form a zygote by fusing with the female counterpart. Further development into ookinete and oocyst allows the parasite to form sporozoites that travel to the salivary glands and can then be transmitted via subsequent mosquito bites [38]. Constant switching between invasive and replicative forms ensures an effective parasitic life cycle in which Plasmodium is well-adapted to its two hosts.

All malaria symptoms are caused by asexual blood stage parasites, ranging from mild fever to more severe forms of the disease like cerebral malaria, severe anemia, or metabolic acidosis [39]. The characteristic periodic fever episodes occur with each rupture of the erythrocytes and release of merozoites, inducing proinflammatory immune responses. Different host-derived factors, the parasite strain, and the social environment determine the clinical outcome of a *Pf* infection. A key feature in parasite pathology is the expression of adhesins on infected red blood cells mediating adhesion endothelial and non-infected red blood cells to avoid splenic clearance. This causes obstruction compromising blood flow and tissue perfusion [40]. Moreover, a combination of different parasite factors impairs the oxygen delivery to tissues, which is the most prominent reason for severe disease manifestation. Additionally, clinical disease prognosis depends on host properties, such as age and pre-existing immunity.

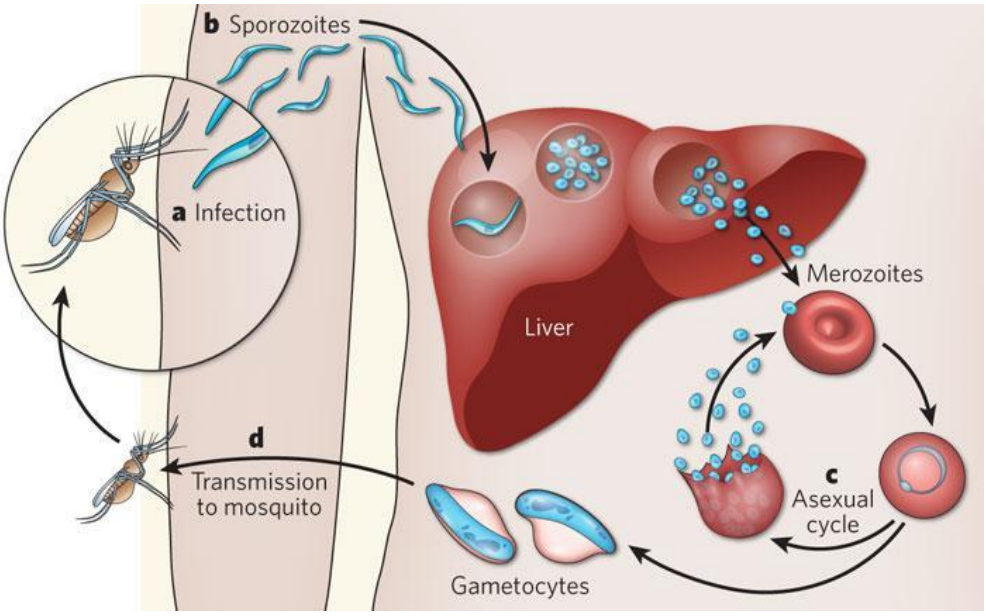


Figure 3: Lifecycle of malaria causing Pf alternating between human and Anopheles mosquito.
Injections of sporozoites into the skin (a) initiates the infection, followed by migration to the liver to induce the pre-erythrocytic stage (b). Merozoites are released into circulation upon hepatocyte rupture, commencing the asexual replication cycle (c). Production of gametocytes, the sexual form of the parasite, ensures transmission to the mosquito by a subsequent bite (d). Image taken from Michalakis & Renaud, 2009 [41].

7.7. T cell response to *Pf*CSP

The control of malaria infection is a complex task for the immune system, which involves different immune cell populations. Throughout an infection, individual parasite stages (sporozoite, merozoite and gametocyte) must be targeted at distinct sites of the human body (skin, liver, blood circulation), which requires co-operation of various immune cells. Many studies that investigated immune responses elicited by *Pf* exposure support the protective role of B cells, CD4 T cells, and CD8 T cells specific for pre-erythrocytic stages of the parasite [42]. While antibodies are essential for neutralization of the sporozoite in the skin and in circulation [43, 44], CD4 T cells provide help to produce potent Plasmodium-specific antibodies as well as mediate activation of macrophages and NK cells [45]. In addition, Plasmodium-specific CD8 T cells can exert cytolytic effector functions, once sporozoites reach the liver and infect hepatocytes [46]. In addition to HLA class I restricted cytolytic activity to eliminate infected hepatocytes, CD8 T cells engage in pathways of protective immunity that depend on the release of cytokines. T cell-derived IFN-γ has been found to be crucial for both direct parasite-killing as well as potentiating the CD8 T cell response against infected hepatocytes [47, 48].

Pre-erythrocytic stage vaccines (PEV) aim to provide protection against the initial sporozoite stage of the infection, before the blood stage and associated symptomatic manifestation of the disease is

initiated [49]. PEVs are designed to induce antibodies, that clear sporozoites from the skin and bloodstream and T cell responses, that eliminate infected hepatocytes. Already early malaria vaccine approaches revealed, that repeated immunization with irradiated sporozoites could confer sterile immunity both in animal models and in humans, which was dependent on memory CD8 T cells targeting the liver stage [50-53]. During controlled human malaria infection (CHMI) studies as well as in naturally exposed individuals, a cytotoxic T lymphocyte (CTL) response to pre-erythrocytic stage antigens could be detected and several epitopes in sporozoite-associated proteins were identified [20, 54]. The circumsporozoite protein (CSP), which is prominently displayed on the sporozoite surface, is known to elicit both, a strong cellular and humoral immune response. Together with its high abundance leading to immunodominance of the protein, CSP is a promising vaccine target. The role of CD8 T cells in controlling infection was underlined in a study, in which depletion of CD8 T cells abrogated immunity in mice, while transfer of CSP-specific CD8 T cells conferred protection [55-57].

CSP consists of a tandem repeat region flanked by an N-terminal and C-terminal domain and is embedded in the plasma membrane via a glycosylphosphatidyl-inositol (GPI) anchor at its C-terminus (Figure 4). The major B cell epitope of the protein forms a four amino acid NANP motif within the non-structured repeat region. Both the N- and C-terminal domains contain several T cell epitopes with different degrees of immunodominance and associated HLA immunoprevalence. Epitopes that are presented by numerous HLA haplotypes are associated with high frequency of responding individuals to a specific epitope, termed immunoprevalence. In contrast, immunodominance relates to high frequency of epitope-specific T cells in a single individual. For CD4 T cells four immunodominant epitopes have been reported that are associated with presentation by numerous frequent HLA-DR haplotypes. Th2R, T*, Th3R, CS.T3 are all located in the C-terminal domain of the protein. Specifically, the Th2R epitope has been previously found to be targeted by a potent T cell response in the context of several high and low frequent HLA-DR alleles [58, 59]. In contrast, many CD8 T cell epitopes are restricted to only one HLA haplotype due to the narrow peptide repertoire and associated low overlap across HLA-I alleles. Most immunogenic epitope regions, in particular Th2R and Th3R, along CSP are associated with a high degree of sequence variability among circulating parasite strains. These naturally occurring variations strongly affect residues that are critical for T cell recognition, which prevents T cell-mediated cross protection against other circulating strains [60].

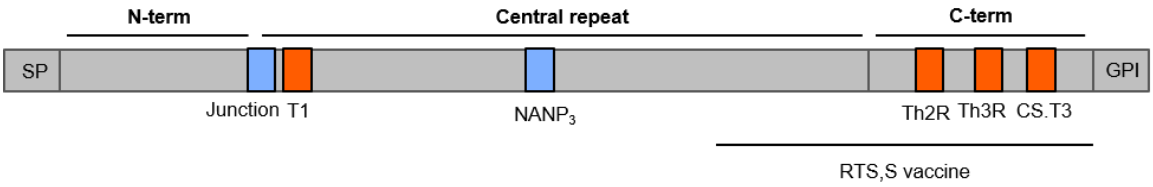


Figure 4: Schematic representation of CSP.

CSP contains three domains, a central repeat region flanked by a N-terminal and C-terminal domain. It is embedded in the plasma membrane through a GPI anchor at the C-terminus. While protective B cell epitopes are located in the central regions (blue), T cell epitopes are located in the N- and C-terminus of the protein. As indicated, the RTS,S vaccine contains part of the repeat region and the C-terminal domain.

To date, two malaria vaccines, RTS,S and R21, received recommendation by the World Health Organization (WHO) for use in regions with moderate to high *P.falciparum* transmission. Both vaccines are CSP-based and comprise part of the repeat region as well as the C-terminal domain of the CSP fused to hepatitis B surface antigen (Figure 4) [61]. Rapid decline in antibody titers and limited efficacy underlines the need to improve next-generation vaccine candidates [62]. As CSP is also expressed during parasite development in the liver, it might serve as a potent CD4 T cell target to promote antibody responses and mediate parasite killing via the production of IFN-γ, or CD8 T cell target for direct killing of infected hepatocytes. The insufficient induction of T cell-mediated cross-protection by available vaccines guided current approaches to investigate the potential of including non-malaria universal epitopes originating from, e.g., tetanus [63, 64]. The use of non-malaria protein-originated universal T cell epitopes circumvents the loss of specificity with subsequent exposure with another strain. However, it also prevents natural boosting by infection to maintain high frequency of antigen-specific T cells. Therefore, the identification of protective, conserved epitopes provides the opportunity to incorporate them into an epitope-based vaccine with increased efficacy.

7.8. Methodologies to study antigen-specific T cells

To evaluate the potency of vaccine or infection induced T cell responses, the detection and quantification of antigen-specific T cells is required. Various approaches aim at the identification of antigen-specific T cells including *in vitro* activation assay, *in vitro* proliferation assays and multimer staining being the most common ones [65]. In recent years, TCR sequencing and repertoire analysis revolutionized the field, allowing researchers to investigate the dynamics of antigen-specific T cell clones over time [66].

Multimer staining techniques allow the direct detection of TCR-pMHC interaction but lack the detection of functionality and measure of biological relevance [67]. Although the TCR-pMHC interaction

is highly specific, it is of low affinity and, hence, produces a weak and transient signal, which is challenging to detect without amplification [68]. Multimer staining techniques can increase the avidity by simultaneous binding of multiple cell surface TCRs. Especially MHC multimers comprising ten MHC monomers revealed superior sensitivity and reduced signal to noise ratio compared to conventionally used tetramers. However, prior knowledge of epitope and MHC restriction is essential for multimer analysis and only a small fraction of the immune response to whole protein is captured by this approach.

In contrast, *in vitro* assays using antigen-induced activation and proliferation as functional readouts, enable sensitive detection without the restriction to certain HLA haplotypes and epitopes. Most commonly, peripheral blood mononuclear cells (PBMCs) are stimulated with peptide libraries and consecutive functional readout include up-regulation of surface activation markers, cytokine release and proliferation detected by activation induced marker (AIM) assay, IFN-γ release assay, enzyme linked immunosorbent spot assay (ELISpot) or intracellular cytokine staining (ICS). *In vitro* peptide stimulation enables the expansion of antigen-specific T cells and the detection of less abundant cell types, but might induce alterations in cell phenotype and bystander expansion of T cell clones independent of antigen-specificity. All *in vitro* approaches differ in their sensitivity, robustness, throughput and affordability, which influences the method choice based on biological question and sample availability (Table 1).

Table 1: Advantages and disadvantages of methods for the detection of antigen-specific T cells.

	Advantage	Disadvantage
Multimer staining	Sorting of antigen-specific T cells Combination with phenotyping possible Sensitive and specific	Limited coverage of HLA alleles and peptides Detection of non-functional antigen-specific cells High costs and error prone
ELISpot assay	Higher sample throughput possible Sensitive and cost effective Low number of cells required for analysis	No sorting of cytokine-secreting cells possible No information on *ex vivo* T cell phenotype
AIM assay	Sorting of activated cells possible	High background dependent on activation marker Less sensitive Cell and work intensive assay
ICS	Assessment of multiple cytokines at single-cell level	No sorting of cytokine secreting cells possible

The possibility of TCR sequencing and high-throughput next generation sequencing enabled detailed insights into TCR sequence diversity. The fact that T cells become activated and proliferate upon

antigen encounter allows the identification of cells contributing to the ongoing immune response by tracking clonal dynamics over time. In contrast to direct analysis of antigen-specific cells by HLA-restricted multimer staining, detection of changes in the repertoire composition allows an unbiased approach by characterizing the immune response and track cell frequencies and phenotypes over time comprehensively. Longitudinal characterization of the TCR repertoire was first restricted to bulk sequencing methods, which concede sequencing of extremely high number of TCRs but prevented single-cell pairing information of alpha and beta TCR chains [69]. Single-cell sequencing methods provided new analytical opportunities such as capturing paired information of both chain sequences and linking the cell phenotype by including single-cell transcriptome or index sorting flow cytometric analysis. The vast diversity of TCR sequence combinations poses one of the major challenges in T cell detection, making it impossible to capture the entire TCR repertoire for a given epitope. The combination of TCR sequence feature information with clustering-based algorithms enabled the search for sequence similarities within a given TCR repertoire and identification of epitope-specific TCR patterns. Recently, efforts have been made to extend TCR-pMHC databases and improve clustering algorithms based on sequence similarity. Nonetheless, the amount of experimentally validated data is still limited and significantly skewed towards a few HLA haplotypes and epitopes.

8. Objectives

The overall goal of this study was to investigate limitations of currently available CSP-targeting vaccines to induce potent T cell responses by characterizing CSP-specific T cells after recombinant CSP vaccination on a single-cell level.

The specific aims were:

1. Repertoire analysis to characterize the T cell response dynamics in recombinant CSP vaccinations
2. Identification and functional characterization of CSP-reactive TCRs
3. Investigation of the role of HLA genotype in protein immunogenicity

9. Methods

9.1. Human FMP013 immunization trial

Whole blood was obtained from participants of the *Falciparum* Malaria protein (FMP) immunization trail FMP013 Part A, which has been conducted by Dr. Sheetij Dutta and colleagues at the Walter Reed Army Institute, Silver Spring, Maryland. Within this study ten volunteers received three doses of either 20 µg or 40 µg FMP013 in ALFQ adjuvant on day 1, 29 and 57 as part of the high dose or low dose arm of the study, respectively. Blood samples were taken 7 days before the first immunization, 28 days after the second and 14 days after the third immunization. PBMC isolation has been performed at the study site and samples were shipped to Germany after completion of the study.

Ethics approval was granted by the ethics committee of the medical faculty of the University Heidelberg in 2020.

9.1.1. Thawing PBMCs

PBMCs were thawed at 37°C for one minute and quickly transferred to a 50 ml falcon containing 30 ml RPMI. After centrifugation for 8 minutes at 1400 rpm at 4°C, the supernatant was discarded and the cell pellet resuspended in appropriate medium. To determine the total viable cell number, tryptophan blue 1:1 dilution and Neubauer counting chamber was used.

9.1.2. *Ex vivo* T cell expansion

For the *ex vivo* T cell expansion assay, PBMCs were resuspended in 5 ml RPMI containing 10 % fetal calve serum (FCS, US origin), 2 mM Glutamine, 1.2 % Penicillin/Streptomycin and 1.5 % 1M HEPES (expansion medium) and divided based on the number of stimulatory conditions. One third of cells for each stimulatory condition were incubated with the CSP peptide pool or an equivalent amount of DMSO for 1 h at 37 °C, while remaining cells were kept at 37°C without any peptides added. After 1 h incubation, stimulated cells were washed by addition of 5 ml RPMI and centrifugation at 500 g for 10 min. Subsequently, unstimulated cells of each condition were combined with the stimulated fraction, spun at 500 g for 10 min and seeded into a 48-well cell culture plate at densities of $1.5x10^6$ cells per well. For co-stimulatory signals, 20 U/ml recombinant IL-2 (Stemcell) and 0.5 µg/ml CD28 antibody (BD Biosciences) were added to the culture. On day 4 and 7, the expansion medium was exchanged and cells were transferred to a 24-well cell culture plate depending on the cell density and medium color. To avoid high levels of unspecific activation, cells were starved on day 9, 24h before restimulation, by omitting IL-2 supplementation.

9.2. T cell receptor repertoire sequencing

For T cell repertoire sequencing, single T cells from PBMC samples were sorted into 384-well plates, the *TR* genes were amplified from the cDNA and the amplicon library was sequenced by next generation sequencing followed by repertoire analysis using an automated analysis pipeline.

9.2.1. Fluorescence activated cell analysis and single cell sorting

For flow cytometric analysis, freshly thawed PBMCs or 10-day stimulated and expanded cells were washed with FACS buffer (4% fetal bovine serum (FBS) in PBS) and subsequently stained in 50 µl (< $1x10^6$ cells) or 100 µl (1-$5x10^6$ cells) antibodies diluted in FACS buffer. Freshly thawed PBMCs were stained with the activation panel and 10-day expanded PBMCs with the T cell expansion panel listed in Table 2. After antibody staining for 30 min on ice in the dark, cells were washed in FACS buffer and the cell pellet was resuspended in 50 or 100 µl live-dead marker 7-aminoactinomycin D (7AAD) and incubated for 10 minutes at 4°C in the dark. Finally, cells were resuspended in 200-800 µl FACS buffer based on the initial cell number and filtered shortly before analysis and index cell sorting at a FACS Aria III (BD). Using freshly thawed PBMCs, activated CD8 T cells were sorted based on 7AAD-CD3+CD8+CCR7- phenotype and positivity in any of the four activation markers 4-1BB (CD137), PD-1, ICOS or CXCR5. In contrast, for the 10-day expansion assay, activated T cells were sorted based on a 7AAD-CD3+CD4+ or 7AAD-CD3+CD8+ phenotype and high co-expression of CD25 and OX40 or CD69 and CD137, respectively. During sorting, index data were acquired, which links the single cell phenotype to the position on the 384-well plate. Single cells were sorted into individual wells of a 384-well plate, that is cooled to 4°C during the sorting process and contained 2 µl of random hexamer primer (RHP) sorting mix per well. This mix contains the RNAse inhibitor RNAsin, the cell lysing agent NP-40, the reducing agent DTT and RHPs for priming reverse transcription (Table 3). After a maximum of 70 minutes of sort time, each plate was immediately frozen on dry ice and stored at -80°C until further processing.

Table 2: Flow cytometry panel for the single cell sorting, I: Activation panel, II: T cell expansion panel

Antigen	Fluorochrome	Clone	Dilution	Panel
CD3	FITC	OKT3	1:250	I, II
CD4	APC-Cy7	A161A1	1:100	I, II
CD8	PE-Cy7	SK1	1:200	I, II
CD45RA	BV510	HI100	1:200	I
CCR7	BV710	G043H7	1:10	I
CXCR5	Alexa647	RF8B2	1:100	I
PD-1	BV605	EH12.2H7	1:20	I
ICOS	PE-Cy7	C398.4A	1:100	I
CD137	BV785	4B4-1	1:20	I, II
CD69	Alexa647	FN50	1:50	II
OX40	BV421	ACT-35	1:50	II
CD25	PE	BC96	1:100	II
7AAD			1:400	I, II

Table 3: Composition of the sort RHP mix.

Reagent	Stock concentration	Volume (ul)
Nuclease-free water	-	1.48
PBS	10x	0.05
DTT	100 mM	0.10
NP-40	10%	0.14
RHP	300 ng/ul	0.14
RNAsin	40 U/ul	0.09
Total		2.00

9.2.2. T cell receptor amplification

For cDNA synthesis, plates were thawed shortly on ice, incubated at 68°C for 60 sec as initial heat step and spun down at 850 g for 1 min at 4°C. 2 µl of reverse transcription (RT) mix containing DTT, RNAsin, dNTPs, RT buffer and reverse transcriptase SuperScript IV (Life Technologies) were added to each well and the plate was incubated on an Eppendorf Mastercycler using the cDNA synthesis program (Table 5). *TR* gene amplification was subsequently performed by semi-nested PCR. Initially, cDNA was diluted by addition of 9 µl nuclease-free water per well and 2.5 µl of diluted cDNA was used as template for the first PCR. Separate primer sets for *TRA* and *TRB* gene amplification were used, covering all V gene

segments respectively and attaching a common linker sequence at the transcript start. The amplified sequence spans across large parts of the V segments until the start of the constant region, where the reverse primer binds. To increase amplification efficiency, the first PCR was performed in duplicates for *TRA* and *TRB* gene amplification separately and 0.5 µl of each duplicate reaction product was used as template for the second PCR. The second PCR forward primer bound the common linker sequence attached to the V segment and the reverse primer bound to the constant segment in a nested way. Both forward and reverse primers contained 16 pb long barcodes unique per row and columns, adding spatial information to the amplified *TR* genes. The detailed composition of the first and second PCR reaction are listed in Table 6 and Table 8, and master cycler conditions in Table 7 and Table 9. All primer sequences used for *TRA* and *TRB* amplification are catalogued in Supplementary Table 5 and Supplementary Table 6. The pipetting was carried out on a Tecan Evo200 automation platform. To determine the amplification efficiency and verify successful amplification, PCR products of selected rows of several 384-plates were run on a 2 % agarose gel by gel electrophoresis and the presence of a band at the correct size was quantified.

Table 4: Composition of the RT mix.

Reagent	Stock concentration	Volume (ul)
Water	-	0.63
RT-Buffer	5x	0.80
DTT	100 mM	0.30
dNTPs	25 mM each	0.14
RNAsin	40 U/ul	0.06
SuperScript IV	200 U/ul	0.07
Total		2.00

Table 5: PCR cycler program for the cDNA synthesis.

Step	Temperature	Duration	Cycles
RNA denaturation	42 °C	5 min	1x
Primer annealing	25 °C	10 min	1x
Reverse transcription	50 °C	60 min	1x
Termination	94 °C	5 min	1x
Cooling	4 °C	Hold	1x

Table 6: Composition of the first PCR reaction.

Reagent	Stock concentration	Final concentration	Volume (ul)
Nuclease-free water			10 – x
PCR Buffer	[10x]	1x	1.00
5’ primer	[50 µM]	0.06 µM each	0.49/0.34
3’ primer	[50 µM]	0.3 µM	0.06
dNTPs	[25 mM each]	0.25 mM	0.10
HotStarTaq	[5 U/ul]	0.025 U/ul	0.05
Template (diluted with 9 µl water)	-		2.50
Total			10.00

Table 7: PCR cycler program for the first PCR reaction.

Step	Temperature	Duration	Cycles
HotStarTaq activation	94 °C	15 min	1x
Denaturation	94 °C	30 sec	50x
Primer annealing	62 °C	1 min	50x
Elongation	72 °C	1 min	50x
Final elongation	72 °C	5 min	1x
Cooling	4 °C	Hold	1x

Table 8: Composition of the second PCR reaction.

Reagent	Stock concentration	Final concentration	Volume (ul)
Nuclease-free water			6.78
PCR Buffer	[10x]	1x	1.00
MgCl2	[25 mM]	4 mM	1.00
5' primer	[50um]	0.167 µM	0.03
3' primer	[50um]	0.167 µM	0.03
dNTPs	[25mM each]	0.25 mM	0.10
HotStarTaq	[5 U/ul]	0.025 U/ul	0.05
Template	-	-	1.00
Total			10.00

Table 9: PCR cycler program for the second PCR reaction.

Step	Temperature	Duration	Cycles
HotStarTaq activation	94 °C	15 min	1x
Denaturation	94 °C	30 sec	50x
Primer annealing	62 °C	30 sec	50x
Elongation	72 °C	1 min	50x
Final elongation	72 °C	5 min	1x
Cooling	4 °C	Hold	1x

9.2.3. Next generation sequencing

For Illumina next generation sequencing the TruSeq Nano library preparation kit (Illumina) was used to ligate adaptors to the amplicons. One sequencing matrix consists of twelve 384-well plates with unique row- and column-specific barcodes. 1 µl DNA of each well corresponding to one matrix and one locus were pooled and 1:3 diluted with nuclease- free water. The pooled DNA was purified using the NucleoSpin Gel and PCR CleanUp kit (Macherey-Nagel) to remove primers. The purification was performed based on manufacturer's instructions with a DNA to NTI ratio of 3:2, based on recommendation for an efficient purification of the given amplicon size. The DNA concentration and purity were determined using a NanoQuant plate in a Tecan M1000 Pro plate reader measuring the optical density (OD) ratio of OD260/OD280 and OD260/OD230. For further purification and size exclusion, the sample was separated by gel electrophoresis on a 4% agarose gel, run at 80V for 6h at room temperature. Afterwards the gel was stained for 30 min in 1:2000 SybrSafe (Invitrogen) diluted in water at room temperature. Using UV light, bands at the expected size were cut and again purified

with the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) using a gel to NTI ratio of 1:4, to avoid clogging of the column due to large gel amounts. For more accurate DNA quantification, picogreen measurement has been performed using the Quant-iT PicoGreen ds DNA kit (Invitrogen) following manufacturer's instructions. The purified DNA samples were used for library preparation using the TruSeq Nano DNA LT kit (Illumina). 1 µg of DNA was complemented up to 27 µl with nuclease free water and 3 µl NEB 2.1 buffer was added. The adaptor ligation reaction was performed according to the kit's protocol, except the the usage of nuclease-free water instead of resuspension buffer and reduction of resuspension buffer volume to resolve DNA amplicons after bead purification to 25 µl. Additionally, only one round of washing has been performed to reduce further DNA losses. To control for the library quality, the amount of adaptor dimers and the amount of ligated amplicons, both a Bioanalyzer profile as well as qPCR quantification was performed. For DNA fragment size analysis the High Sensitivity DNA kit (Agilent) was used following the provided instructions and running the loaded chip at a Bioanalyzer (Agilent) for 40 min. For the quantification of ligated amplicons, the KAPA Library Quant kit (Roche) was utilized using 1:10,000 – 1:40,000 sample dilutions and following the manufacturer's instructions. Based on these quantifications, both libraries of the alpha and beta loci were multiplexed in equimolar ratios and send for sequencing at the MiSeq2 (Illumina) at the Functional Genomics Facility in Zürich using the paired end 300 bp read configuration.

9.2.4. Repertoire analysis

TCR sequence analysis and segment annotation was performed by an TCR adapted version of the automated analysis pipeline SciReptor [70, 71]. In brief, paired sequencing reads were assembled using the PandaSeq tool setting the maximal and minimal length thresholds to 550 and 300 bp, respectively, to capture all reads of the expected length [72]. A minimal overlap of 50 bp is required for successful assembly and the read quality score must by above 0.8. Subsequently, assembled sequences are aligned to bacteriophage PhiX genome to exclude PhiX sequences, which were added during the sequencing process to achieve a balanced base composition. As a next step, reads are assigned to their plate position according to the column- and row-specific barcodes and V, D and J segments as well as CDR and framework regions (FWR) are identified by Ig BLAST and annotated [73]. Afterwards, information about the cellular phenotype from index sorting fcs files, and the metadata table are linked to the sequence information and stored in a relational MySQL database.

Further analysis of the clonal composition and sequence features were performed in R studio using the following packages: ggplot2, ggalluvial, pheatmap, vegan, stringr, dplyr. Clustering of TCR according to sequence similarity was performed using the BLscore package and further visualization was executed in Cytoscape.

9.3. T cell receptor cloning

For further functional characterization of the TCRs of interest, the TCRs were cloned into a retroviral expression vector aiming for stable expression in Jurkat76 cells. To achieve this, alpha and beta first PCR product was used as template for V gene specific amplification. Both alpha and beta TCR amplicon were cloned into a retroviral expression vector, joined by a 2A self-cleaving element using Gibson assembly.

9.3.1. Specific amplification of *TR* genes

To enable cloning of *TRA* and *TRB* genes into the expression vector, regions of sequence homology, required for Gibson assembly were attached to the PCR products. To this end, 0.5 µl of each duplicated primary PCR product was pooled and used as template in a V gene segment specific PCR reaction. For alpha, all forward primers contain 32 bp of the TRAV13-1 signal sequence, whereas all beta forward primers contained 19 bp of the TRBV5-1 signal sequence. The reverse primers bound to the constant region of the *TRA* and *TRB* genes. All V segment specific primers for *TRA* and *TRB* are listed in Supplementary Table 7 and Supplementary Table 8. The composition of the PCR reaction mix and mastercycler program conditions are listed in Table 10 and Table 11. Successful amplification was verified by gel electrophoresis on a 2% agarose gel and PCR products were purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). DNA concentration and purity were determined using a NanoQuant plate in a Tecan M1000 Pro plate reader.

Table 10: Composition of the specific PCR reaction for TCR alpha/beta cloning.

Reagent	Stock concentration	Final concentration	Volume (ul)
Nuclease-free water	-		10.6
PCR Buffer	[5x]	1x	4.00
5’ primer	[5 µM]	0.5 µM	2.00
3’ primer	[5 µM]	0.5 µM	2.00
dNTPs	[25 mM each]	0.25 mM	0.20
Phusion	[2 U/ul]	0.02 U/ul	0.20
Template	-	-	1.00
Total			20.0

Table 11: PCR program for the specific PCR reaction.

Step	Temperature	Duration	Cycles
Initial denaturation	98 °C	1 min	1x
Denaturation	98 °C	20 sec	5x
Primer annealing	65 °C	30 sec	5x
Elongation	72 °C	1 min	5x
Denaturation	98 °C	20 sec	35x
Primer annealing	72 °C	30 sec	35x
Elongation	72 °C	1 min	35x
Final elongation	72 °C	10 min	1x
Cooling	4 °C	Hold	1x

9.3.2. Restriction digestion of vector and linker segment

The retroviral expression vectors pMSCV-PImC-SSa-TRBC1 or 2, containing the constant region 1 or 2, were digested overnight at 37°C using the MfeI-HF restriction enzyme in 10x Cut-Smart Buffer diluted with nuclease free water (Table 12). To avoid re-ligation, the linearized vector was incubated with 1µl calf-intestinal-phosphatase (CIP) for 5 min at 37°C. Subsequently, the vector was purified by gel electrophoresis on a 1 % agarose gel followed by gel excision and purification with the NucleoSpin Gel and PCR Clean-Up kit (Macherey-Nagel). The pEX-A128-TRAC-2A-SSb vector containing the linker fragment for joining the *TRA* and *TRB* genes, was digested with BsrGI-HF at 37°C overnight (Table 12). The linker fragment of 549 bp size was excised from a 2 % agarose gel and the DNA concentration was measured after Nucleospin Gel and PCR Clean Up (Macherey-Nagel).

Table 12: Restriction endonuclease digests of the TCR cloning vectors

Vector	Restriction endonuclease	Size	Fragment size
pMSCV-PImC-SSa-TRBC1	MfeI-HF	8206	-
pMSCV-PImC-SSa-TRBC2	MfeI-HF	8206	-
pEX-A128-TRAC-2A-SSb	BsrGI-HF	3140 bp	549 bp

9.3.3. DNA assembly

To assemble the 2A-Linker fragment and the *TRA* and *TRB* specific PCR products, into the pMSCV-PlmC backbone, Gibson assembly was performed. Equimolar amounts (0.0125 pmols) of each fragment and backbone were used in a 5 µl reaction with the 2x NEBuilder Hifi DNA Assembly Master Mix (NEB). The assembly reaction was performed at 50° C for 60 min.

9.3.4. Transformation into competent bacteria

For transformation into chemically competent DH10b *E. coli* cells 3 µl of assembly product was added to 5 µl of competent bacteria and incubated on ice for 30 minutes. Subsequently, a heat shock was performed at 42 °C for 45 sec and the samples were immediately cooled on ice. To facilitate initial bacterial growth, 100 µl of LB medium were added and cells were incubated at 37°C for 1 h while shaking at 650 rpm. Next, the bacteria were plated on LB-agar plates containing 100 µg/ml ampicillin and grown overnight at 37°C.

9.3.5. Screening of bacterial colonies by PCR and sequence validation

Successful transformation was first validated by PCR amplification using bacterial colonies as template to check for presence of complete *TR* genes. For each individual TCR three single colonies were picked, resuspended in 3 µl water each and 1 µl of the suspension was utilized as template in the colony PCR amplification. The forward primer (colP_Psi_in_F) and reverse primer (TRBC115-135) bind upstream and downstream of the insertion site, respectively (Supplementary Table 9), ensuring an amplification of the entire TCR construct. The detailed PCR composition and PCR cycler conditions can be found in Table 13 and Table 14. To validate the correct length of the insert, the PCR product was analyzed by gel electrophoresis using a 2 % agarose gel, where a 1500 bp large band indicated the presence of the complete TCR, while no insert is detected by a band of 300 bp size. To exclude sequences with PCR-induced point mutations or short deletions, amplicons of correct size were Sanger sequenced (Eurofins Genomics) using the colony PCR primers. Sequences were aligned to the reference sequence from the database using an automated R script and correct sequences, allowing silent mutations, were taken forward to vector DNA preparation.

Table 13: Composition of the colony PCR for TCR cloning validation.

Reagent	Stock concentration	Final concentration	Volume (ul)
Nuclease-free water	-		16.20
PCR Buffer	[10x]	1x	2.00
5' primer	[50 µM]	0.5 µM	0.20
3' primer	[50 µM]	0.5 µM	0.20
dNTPs	[25 mM each]	0.25 mM	0.20
Taq	[5U/ul]	0.05 U/ul	0.20
Template	-	-	1.00
Total			20.00

Table 14: PCR Program of the colony PCR for TCR cloning validation.

Step	Temperature	Duration	Cycles
Initial denaturation	98 °C	10 min	1x
Denaturation	98 °C	30 sec	30x
Primer annealing	52 °C	40 sec	30x
Elongation	72 °C	3 min	30x
Final elongation	72 °C	10 min	1x
Cooling	4 °C	Hold	1x

9.3.6. Preparation of vector DNA

Single bacterial colonies containing TCR constructs with the correct sequence were inoculated in 5 ml LB medium containing 75 µg/ml ampicillin. Bacterial cultures were incubated overnight at 37°C while shaking (180rpm) and harvested the next day by centrifugation. The bacterial DNA was purified using the NucleoSpin Plasmid kit (Macherey-Nagel) according to manufacturer's instructions. DNA purity and concentrations were measured with a NanoQuant plate using a Tecan M1000 Pro plate reader.

9.4. T cell receptor expression

To stably express TCRs in TCR-deficient T cells, retroviral transduction was performed. Phoenix Ampho cells were transfected with retroviral expression vector, which contained retroviral packaging sequence Psi and 5' and 3' long terminal repeats (LTRs) in addition to the *TRA* and *TRB* genes. The combination of retroviral components originating from the transfected vector and endogenous components of the Ampho Phoenix cells allowed the production of amphotropic retroviral particles

containing the TR gene of interest. Subsequently, retroviral particles were used to stably integrate the TR gene into to target T cell line. To enable the quantification of transduction efficiency and the selection of transduced cells, an mCherry marker and a puromycin resistance gene were additionally included in the retroviral expression vector.

9.4.1. Culture of Phoenix Ampho cells

Phoenix Ampho cells were cultured in DMEM GlutaMAX medium supplemented with 10 % heat-inactivated FBS and 1 % Penicillin/Streptomycin (Phoenix medium). Cells were incubated at 37°C and 5 % CO_2 and split every 2-3 days. Passaging was performed by addition of trypsin and incubation at room temperature for 2 minutes. Detached cells were washed and seeded at a density of $1x10^5$ cells per ml.

9.4.2. Culture of Jurkat76 cells

The TCR-deficient T cell line Jurkat76 [74] and its derivates containing the CD3 and CD4 or CD8 co-receptor were cultured in RPMI 1640 supplemented with 10 % heat-inactivated FBS, 1 % Penicillin/Streptomycin and 2 mM Glutamine (Jurkat medium). Cells were incubated at 37°C and 5 % CO_2 and passaged every 2-3 days by seeding concentrations of $8x10^5$ cells/ml.

9.4.3. Transfection of Phoenix Ampho cells

For transfection, $5x10^5$ Phoenix Ampho cells were seeded in 2 ml Phoenix medium into each well of a 6-well cell culture plate. After incubation for ~ 36 h to reach cell confluency of 60-70 %, cells were transfected with the retroviral TCR expression vectors using calcium phosphate. 1 µg of plasmid DNA (0.1 µg/ul DNA in water) were incubated at 50°C for 30 minutes and subsequently 102.5 µl water and 12.5 µl $CaCl_2$ were added. To form DNA complexes, 125 µl of 2x HEPES buffered saline (HBS) were added dropwise, while vortexing. During incubation at room temperature for 10-20 minutes, medium of Phoenix Ampho cells was changed to Phoenix medium without antibiotics. Then, 250 µl of the DNA-HBS mix were added slowly and dropwise to the cells and cells were incubated for 6 h at 37°C. Subsequently, medium was aspirated and 2 ml of Jurkat76 medium without antibiotics was added to each well, followed by overnight incubation at 37°C. The next day, the supernatant containing the retroviral particles was harvested and centrifuged at 1000g for 5 min to remove cellular debris and avoid contamination with Phoenix Ampho cells in the subsequent transduction of Jurkat76 cells.

9.4.4. Retroviral transduction of Jurkat76 cells

For retroviral transduction, $5x10^5$ Jurkat76 cells expressing the CD3 and CD4 or CD8 co-receptor, were resuspended in 1 ml retroviral supernatant (see section 9.4.3) containing 10 µg/ml protamine sulfate and were plated in 24-well tissue culture plates. Spin infection was performed by centrifugation at 2000 g and 32°C for 1.5 h without brakes. Afterwards, 1 ml of Jurkat medium without antibiotics was added to each well and cells were incubated for 2 days at 37°C and 5% CO_2. The puromycin resistance gene N-acetyltransferase (PAC), contained in the TCR expression vectors allows selection for transduced cells by addition of 0.8 µg/ml puromycin dihydrochloride to the culture.

9.4.5. Detection of TCR expression in transduced Jurkat76 cells

To assess protein expression of genes introduced by retroviral transduction, Jurkat76 cells were analyzed for TCR and mCherry expression by flow cytometric analysis. $1x10^6$ Jurkat76 cells were harvested on day 10 after start of puromycin treatment and transferred to individual wells of a 96-well plate for all subsequent washing and staining steps. After an initial wash with FACS buffer and centrifugation at 1400 rpm for 5 minutes, cells were resuspended in 50 µl FACS buffer containing anti-CD3, anti-CD8 and anti-TCR antibody (see Supplementary Material 14.4.2) and incubated for 30 minutes at 4°C in the dark. After subsequent washing, cells were stained in 50 µl FACS buffer containing live-dead 7AAD marker for 10 minutes at 4°C. Afterwards, cells were washed and analyzed at the FACS Aria III (BD).

The transduction efficiency was assessed by flow cytometric analysis of mCherry and TCR expression (Supplementary figure 2). For subsequent functional characterization, a TCR expression threshold cutoff of 7 % of TCR positive cells among mCherry expressing cells was defined based on minimal TCR expression of previously identified CSP-reactive TCRs.

9.5. Functional characterization of T cell receptors

To characterize T cell receptors with regard to their CSP-specificities, TCR transgenic T cell lines were co-cultured with autologous B cells pulsed with CSP peptides. The degree of activation was assessed by measuring the IL-2 concentration by ELISA.

9.5.1. B cell Immortalization

The Epstein-Barr virus (EBV), used for B cell immortalization, was obtained from the B95-8 cell line, which are described to release high titers of EBV. EBV stocks were previously generated by Ilka Wahl.

In brief, B95-8 cells were cultured for seven days without changing the medium, and the supernatant was subsequently harvested and stored in cryotubes in liquid nitrogen until further use.

To generate autologous B cell lines of each donor, human B cells were isolated from PBMCs using fluorescence-activated cell sorting. PBMC were thawed as described previously and stained with anti-CD19-PE-Cy7 antibody (clone SJ25C1) diluted 1:20 in 100 µl FACS buffer for 30 minutes at 4 °C in the dark. Cells were subsequently washed and stained with live-dead marker 7AAD staining diluted 1:400 in 100 µl FACS buffer. After 10 min incubation at 4°C in the dark, cells were washed and resuspended in FACS buffer and analyzed at the FACS Aria III (BD). All live B cells, defined as 7AAD-CD19+, were bulk sorted, yielding 1-4x10^5 B cells per donor. B cells were then washed by addition of 5 ml RPMI and after centrifugation resuspended in 250 µl EBV containing supernatant and 250 µl RPMI containing 10 % FBS, 2 mM Glutamine, 1 % Penicillin/Streptomycin, 1% non-essential amino acids and 1% pyruvate (B cell medium) In order to improve transformation efficiency by activating B cells, 1.25 µg/ml resiquimod (Sigma Aldrich) was supplemented to the culture. On day 5 after sort, 500 µl of fresh B cell medium containing resiquimod (Sigma Aldrich) were added to the cells. On day 7, the culture medium was removed after centrifugation and exchanged with B cell medium without toll-like receptor (TLR) agonist. Then, B cells were expanded for additional 20-30 days, while cells were maintained at cell densities of 0.8x10^6 cells/ml. Expanded cells were frozen in FBS containing 10% DMSO and stored in liquid nitrogen.

9.5.2. T cell stimulation

TCR transgenic T cell lines were stimulated in a 96-well U bottom tissue culture plate with a T:B cell ratio of 1:2 and a total number of 6.4x10^5 cells per well. Each T cell line was tested for reactivity in two biological replicates and two technical replicates. After seeding autologous B cells in 100 µl AIM-V medium, respective CSP peptides were added in 50 µl AIM-V at a concentration of 2.5 µg/ml for each peptide and incubated for 2 h at 37°C and 5 % CO_2. As negative control same volume of DMSO was added to one control well. Afterwards, TCR transgenic Jurkat76 cells were resupended in 50 µl AIM-V and added to the B cells. After 24h incubation at 37°C and 5% CO_2, plates were centrifuged at 1400 rpm for 5 min and supernatant was harvested for IL-2 quantification.

9.5.3. Detection of T cell activation

The IL-2 concentration in the cell culture supernatant was determined by sandwich ELISA using the human IL-2 ELISA MAX Deluxe kit (Biolegend). The ELISA was performed according to the manufacturers protocol with the adaptation to 384-well plate by reducing all volumes per well by a factor of 4. All supernatants were analyzed undiluted, 1:3, 1:9 and 1:27 diluted to ensure that the

quantified IL-2 concentration is within the standard range of the kit. All washing steps were performed using an Aquamax 2000 384 plate washer and the OD450 was measured at a Tecan M1000 Pro Plate reader.

9.5.4. *In-vitro* HLA-blocking experiments

To identify the peptide presenting HLA locus, TCR-transgenic Jurkat76 cell lines were stimulated in the co-culture assay as described in section 9.5.2 in presence or absence of HLA locus-specific blocking antibodies listed in section 14.4.1. Each antibody was added to a separate well at a final concentration of 10 µg/ml together with Jurkat76 cell lines after peptide loading of autologous B cells lines. Blocking of the target peptide presenting HLA molecules is seen in a strong decrease or absence of IL-2 detection in the co-culture supernatant.

9.6. MHC peptide binding predictions

MHC peptide binding predictions were performed using the NetMHCpan4.1 algorithm [23]. All 8mer, 9mer and 10mer peptides covering the full-length 3D7 CSP sequence were included in the analysis and evaluated for binding to HLA-A, HLA-B and HLA-C alleles expressed by at least one donor enrolled in the FMP013 study. All weak binders (rank <2) and strong binders (rank <0.5) were plotted in the respective HLA-allele context.

9.7. *In vitro* pMHC binding assay

CSP peptides covering the major predicted binders identified by NetMHCpan4.1 were loaded onto five MHC monomers (Immudex) and subsequently coupled to streptavidin beads (Spherotech) following the instruction provided by Immudex. Since potentially immunogenic epitopes located within the GPI and signal peptide sequence are not included in the FMP013 vaccine sequence, the respective peptides were excluded from *in vitro* binding assays as well. The level of peptide binding was quantified using a fluorescently labelled antibody staining, which targets a conformational epitope within the ß2-microglobulin domain. Efficient binding of the antibody is dependent on the peptide bound state of the MHC molecule and therefore the MFI of antibody stained beads correlates with peptide binding. For better comparison across alleles, the percent binding was calculated relative to the allele-specific positive control peptide.

9.8. Peptide *in vitro* expansion and dextramer staining

For CSP peptide-loaded dextramer stainings of HLA-matched PBMC samples, peptide-loaded HLA monomers were loaded onto a fluorescently labelled dextramer backbone following instructions

provided by Immudex. Based on *in vitro* HLA binding, one to three peptides with strong binding phenotype were selected for dextramer staining per HLA allele. Freshly thawed PBMCs isolated 14 days after the third vaccination time point of four donors (F1-F4) were stained with the T cell activation panel (Table 2) in addition to the PE-labelled dextramer.

For peptide specific expansion, a 15mer peptide covering the respective dextramer bound peptide was used for 10-day in vitro PBMC expansion described in section 9.1.2. On day ten, without additional restimulation, expanded cells were stained with the T cell expansion panel (Table 2) and the PE-labelled dextramer.

9.9. HLA typing

For HLA typing, gDNA was extracted from PBMCs using the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's instructions. HLA typing was performed at the Institute for Immunology and Genetics, Kaiserslautern or at the DKMS facility.

9.10. Statistical analysis

For running statistical tests, GraphPad Prism (Version 9.3.1) and R (Version 1.3.1093) were used. The corresponding statistical test used for each comparison is stated in the figure legend. Significance levels are indicated as stars corresponding to following p-value ranges: * $p < 0.05$, ** $p < 0.01$, *** 0.001, **** $p< 0.0001$.

10. Results

10.1. Human CSP immunization against malaria

To characterize the T cell response against CSP in depth, I utilized PBMC samples of a vaccination trial using nearly full-length recombinant CSP. The *Falciparum* Malaria protein immunogen (FMP013) includes the non-immunogenic CSP N-terminus distinguishing it from currently licensed vaccines, RTS,S and R21, that solely contain the C-terminal part of the protein.

Within the part A of the FMP013 trial performed at the Walter Reed Army Institute of Research, Silver Spring, Maryland, ten volunteers received three doses monthly distanced apart of either 20 ug (low dose, donors F1-F5) or 40 ug (high dose, donors F6-F10) recombinant CS protein (Figure 5) [75]. The protein was administered in 1 ml Army liposome formulation containing QS-21 (ALFQ). The liposome-based adjuvant contains phospholipids, cholesterol and monophosphoryl lipid A as immunostimulants and was improved by inclusion of QS-21 saponin, which was well tolerated and induced high antibody titers leading to enhanced protection [76-78]. The blood collections were performed 7 days before the first dose, 28 days after the second, or 14 days after the third dose, and PBMCs were isolated subsequently (Figure 5). Two donors did not receive the full vaccination regimen, with F1 only receiving the first dose and F5 receiving the first two doses.

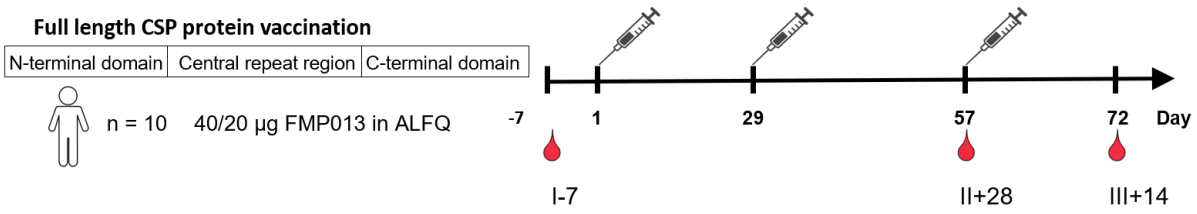


Figure 5: Schematic overview of the human malaria vaccination trial FMP013 (Falciparum Malaria Protein 013).

Ten volunteers received three monthly doses of 40 or 20 ug recombinant full length CSP in 1 ml ALFQ as part of the high and low dose regimen of the study, respectively. Blood collections were performed 7 days prior to the first vaccination, 28 days post second- and 14 days post third dose. The study was conducted by Dr. Sheetij Dutta and colleagues at the Walther Reed Army Institute of Research, Silver Spring, Maryland.

10.2. Full-length CSP vaccination induces polyclonal CD4 T cell response to C- and N-terminus

To enrich for potentially low frequent CD4 T cell populations targeting weakly immunogenic epitopes, I performed a 10-day *in vitro* expansion of PBMC samples isolated 28 days after the second FMP013

vaccine dose, in the presence of CSP peptides and performed activation marker expression readout on day 11. Optimal culture condition to minimize cell death and promote T cell proliferation were established by IL-2 cytokine addition, anti-CD28 antibody concentration and optimal cell density. For peptide loading onto APC surface, PBMCs was pulsed with a peptide library of 52 overlapping 15mer peptides covering the complete CS protein contained in the FMP013 vaccine. Cells were kept in culture for ten days to expand all CSP-reactive T cells along with a donor-matched non-stimulated control culture. To distinguish CSP-reactive cells specific for individual subregions of the protein, cells were split on day 10 and restimulated with sub pools covering the N-terminus/junction region or C-terminus of the protein (Figure 6A). Upon 18 h restimulation, CD4 T cells were stained for activation markers CD25 and OX40 to allow activation phenotype specific single-cell sorting and subsequent sequencing of amplified paired *TR* genes (Figure 6B).

To examine whether the activation induced by initial exposure is downregulated until restimulation on day 10, I compared the clonal composition in C-term or N-term/junction stimulated cultures in two donors (F7 and F9). In both donors, using the 28 days post 2nd vaccination PBMC samples, I detected a positive activation fold change after stimulation with C-term or N-term/junction peptides compared to unstimulated control with stronger activation induced by C-term stimulation (Figure 6C). After sequencing of *TR* genes from 116 and 138 activated CD4 T cells from donor F7 and F9, respectively, clonal composition analysis of both cultures revealed on average 51% clonal expansion in both cultures and donors, however with strong clonal overlap across C-term and N-term antigen stimulation (Figure 6D). The detection of shared clones between N-term/junction and C-term stimulation, but absence among unstimulated culture, suggests that the initial peptide exposure induced an activation, which is maintained over 10-day *in vitro* culture.

To determine the magnitude of the C- and N-terminus specific T cell response, I adapted the assay by splitting the expansion culture for separate C- and N-term specific expansion in samples from donor F1-F5 and kept the same activation readout on day 11 after restimulation (Figure 6A). The level of activation among stimulated CD4 T cells compared to the unstimulated control, determined by flow cytometry staining, varied from fold changes of 0.6 to 4.3 among individual donors (Figure 6E). This suggests variations in the strength of T cell response among donors. All five donors responded to C-term peptide stimulation with positive activation fold changes from 1.5 to 3.4 compared to unstimulated control, in line with various studies describing the C-terminus as immunodominant for T cell responses [79-81]. The less immunogenic N-terminus induced positive activation fold changes compared to unstimulated culture in only two donors (F1, F3) (Figure 6E). Based on reactivity to C- and N-terminus or C-term reactivity only, single activated CD4 T cells from two donors (F3 and F4)

were isolated using an indexed flow cytometric cell sorting and amplified and sequenced their paired *TRA* and *TRB* genes (Supplementary Table 1).

All cultures showed a diverse clonal composition with 67 and 58 unique clones detected upon N-term stimulation and 100 and 124 detected upon C-term stimulation in both donors (F3, F4), respectively (Figure 6E). Corresponding to differences in activation phenotype, a lower degree of clonal expansion was detected in N-term/junction stimulated cultures (36% in F3, 7% in F4) compared to the C-term stimulation in both donors (61% in F3, 27% in F4) (Figure 6F). The absence of clonal overlap between two expansion cultures confirmed the peptide-specific expansion of distinct clones. In summary, 10-day in vitro expansion allowed efficient amplification of the CSP-specific CD4 T cell response.

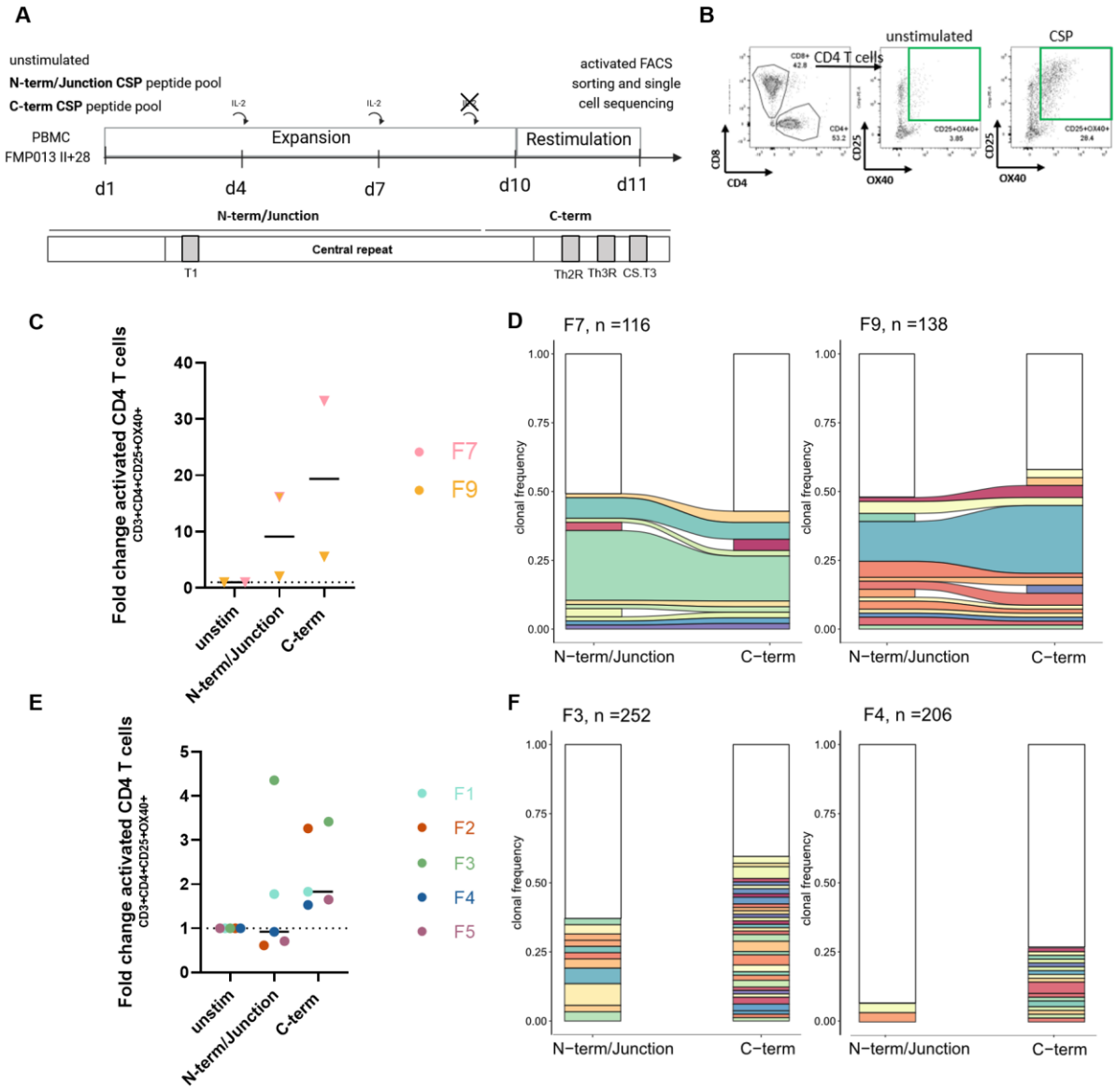


Figure 6: Full-length CSP vaccination induces polyclonal CD4 T cell response to the C- and N- terminus.

A Experimental setup of 10-day T cell expansion assay. PBMCs were incubated 1 h with CSP peptides covering the whole protein (donor F7 and F9) or C- or N-term peptide pool (donor F1-F5) of the protein. On day 10, cells are re-stimulated with N-terminal or C-terminal sub pools and activation marker expression was assessed on day 11 by flow cytometric analysis. **B** Gating strategy after restimulation on day 11. Activated CD3+CD4+OX40+CD25+ T cells (green square) were sorted for subsequent single-cell TCR sequencing. **C + E** Fold change in frequency of activated CD4 T cells compared to unstimulated control after full length CSP expansion (C) or N- and C-term specific expansion (E). **D + F** Clonal composition of activated CD4 T cells in N-term and C-term expansion cultures after full length CSP expansion (D) or N- and C-term specific expansion (F). Expanded clones are shown in color, while non-expanded clones are combined in white compartment. Clonal overlap between N-term/junction and C-term expansion cultures are indicated by connecting bands.

10.2.1. Identification of TCRs targeting N- and C-terminal epitopes

To investigate, whether the expanded clones were CSP-reactive, I next examined whether enriched sequence features in individual cultures can be linked to CSP reactivity of single TCRs. In total, 86 TCRs originating from all four donors (F3, F4, F7, F9) were selected based on enriched V segment usage, V segment pairing, clonal expansion and TCR convergence for further functional characterization (Supplementary figure 1; Supplementary Table 2).

Through gene cloning of paired *TRA* and *TRB* genes and subsequent retroviral transduction, I stably expressed the selected clones in CD3+CD4+ TCR-deficient Jurkat76 T cell line [71]. Based on mCherry expression in 67 transgenic cell lines, the transduction efficiency was assessed (Figure 7A, Supplementary figure 2). In total, 63 out of 67 generated cell lines revealed robust TCR surface expression, qualifying them for functional characterization (Figure 7A). To screen for CSP reactivity, the T cell lines were co-cultured with EBV immortalized autologous B cells pulsed with pools of overlapping 15mer peptides covering the complete N-term, junction/repeat or C-term region of the protein (Figure 7, B-D). Using IL-2 secretion as readout for T cell activation, I identified in total 29 CSP-specific T cell clones. While the majority of clones (69%; 20 of 29) showed specificity to C-terminal peptides (Figure 7B), a relatively small number of TCRs were found to be specific for the N-terminal domain (17%; 5 of 29) and junction/repeat region (14%; 4 of 29) (Figure 7, C-D).

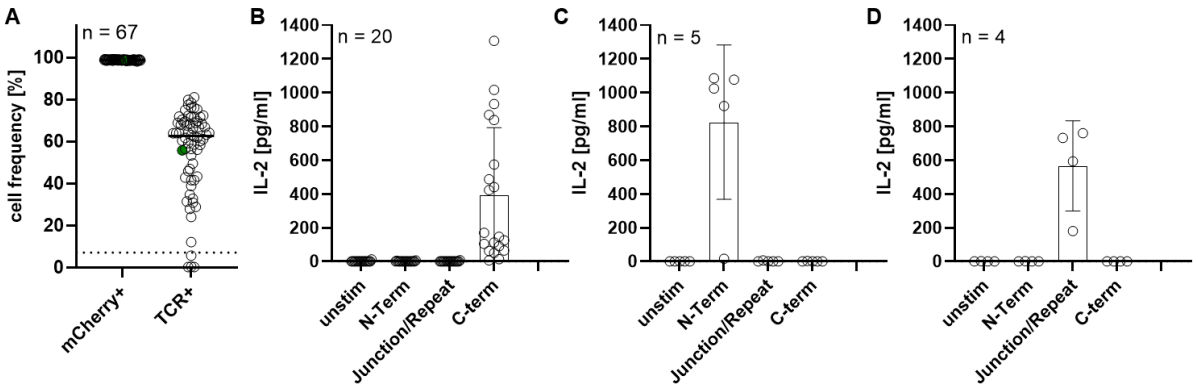


Figure 7: TCRs target N-terminal, junctional and C-terminal epitopes.

A Transduction efficiency of TCR-transgenic Jurkat76 T cell lines based on mCherry (mCherry+) and TCR (TCR+) expression quantified by flow cytometric analysis. 67 transgenic T cell lines were generated by retroviral transduction of TCRs isolated after *in vitro* expansion and were compared to a T cell line expressing a control TCR (green dot). 63 T cell lines with TCR expression above 7% (dotted line) were subjected to subsequent CSP reactivity testing. **B – D** Activation of 29 CSP-specific T cell lines. IL-2 was quantified by ELISA in the supernatant after 24h co-culture with autologous B cell lines loaded with CSP peptide pools. Each T cell line was stimulated with the N-term, junction/repeat and C-term peptide pool or in absence of peptide (unstimulated control). Transgenic T cell lines were categorized by reactivity against the N-terminal (B), junction/repeat (C) or C-terminal (D) peptide pool.

To identify the fine specificity of each CSP-reactive T cell line, subsequent stimulations with peptide sub pools and individual peptides were performed. The majority of TCRs (16/29, 55%) isolated from three of the donors (F3, F4, F7) targeted the immunodominant C-terminal Th2R/T* region (CSP315 HIKEYYLNKIQNSLST; Figure 8A). TCRs (4/29, 14 %) from two other donors (F4, F9) targeted the CS.T3 epitope (CSP367 KICKMEKCSSVFNVV) located downstream of the Th2R region. Further, I identified four TCRs targeting the T1 epitope at the N-terminal junction, all originating from one donor (F3). In addtion, two novel N-terminal epitopes spanning CSP position 61-75 (ENWYSLKKNSRSLGE) and 81-95 (NEDNEKLRKPKHKKL) were identified. All four CSP61 targeting TCRs originated from donor F3 (4/29, 14%), whereas CSP81 was targeted by a single TCR from donor F4 (1/29, 3%) (Figure 8A). Thus, a large fraction of in vitro expanded clones (29/63, 46%) were CSP-specific and targeted several epitopes located in the N-term, junction and C-term of the protein.

10.2.2. The majority of CSP epitopes are targeted in an HLA-DR context

To determine the presenting HLA molecule of each T cell-epitope recognition, I used HLA-DR, HLA-DQ and HLA-DP specific antibodies to block TCR-pMHC interaction specifically for each locus. A complete absence or strong decrease in detected IL-2 concentration in locus-specific HLA blocking condition confirms the presentation of target epitope by the blocked HLA molecule. By specific TCR-pMHC blocking, an HLA-DR restriction of CS.T3 and Th2R epitope was detected for all generated T cell lines confirming the reported strong HLA-DR dominance of both C-terminal epitopes. In contrast, the T* targeting TCR from donor F7 was linked to HLA-DQ peptide presentation and the N-junctional T1 epitope was recognized in the context of HLA-DQ as well. Both N-terminal epitopes revealed HLA-DR restricted peptide presentation, seen in a strong decrease or complete absence of activation in presence of HLA-DR specific antibody (Figure 8A).

Interestingly, IL-2 levels after stimulation with target peptide varied strongly across TCRs and target epitopes. To evaluate whether differences in IL-2 levels originate from differences in TCR-pMHC binding or are caused by technical artefacts as differences in cell number or TCR expression, both cell numbers and TCR expression was monitored during co-culture assays. Although decreased TCR expression upon increased numbers of cell passages correlates with a decrease in IL-2 secretion, the differences between individual T cell lines remained the same, suggesting TCR dependent differences in activation (Figure 7, B-D, Figure 8A). Also, insignificant differences in TCR expression do not correlate with IL-2 secretion (Figure 8B) and the most potent TCRs with highest IL-2 levels were found among Th2R and CSP61 targeting T cell lines (Figure 8A). Together, the data show that the majority of TCRs (24/29, 83%) targeted a HLA-DR presented CSP epitope with variable TCR affinity.

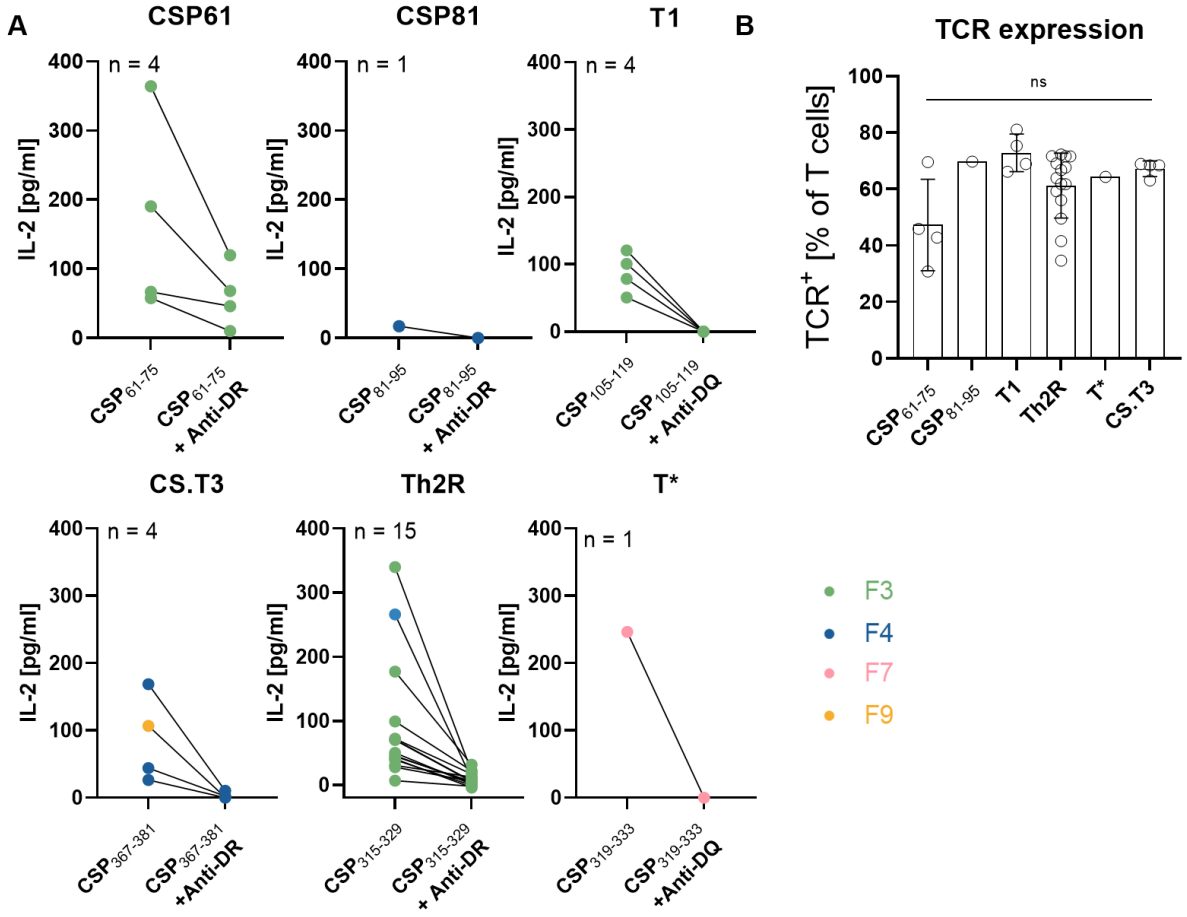


Figure 8: N and C-terminal epitopes are predominantly presented in the context of HLA-DR molecules.

A Activation of TCR-transgenic T cell lines in co-culture with autologous B cell lines loaded with the target peptide in presence or absence of HLA locus-specific blocking antibodies. IL-2 quantification in the supernatant after 24h co-culture is shown color-coded by donor. **B** TCR expression of CSP reactive T cell lines stratified by target epitope. Statistical significance calculated by One-way ANOVA (B).

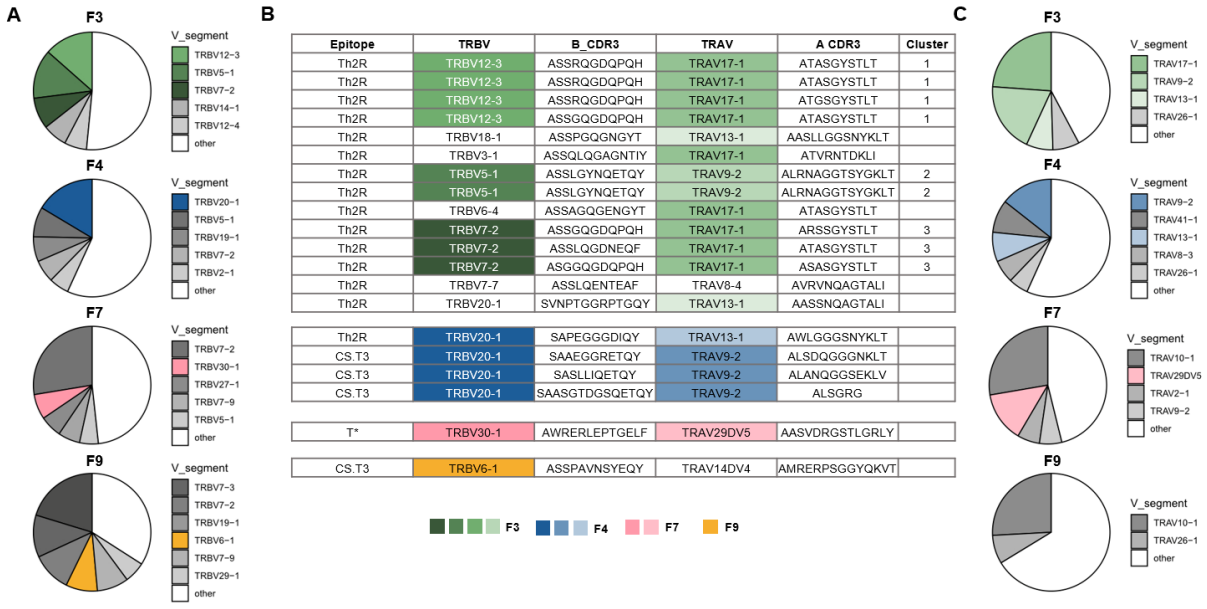
10.2.3. Groups of convergent TCRs target individual CSP epitopes

To define and characterize TCR properties associated with individual epitopes and correlate with enriched sequence features in the expansion cultures, I investigated the TCR sequence features of the CSP-specific TCRs and calculated their abundance among the peptide expanded repertoire. Specifically, I performed similarity clustering analysis and investigated patterns in V segment usage and CDR3 regions within groups of TCRs with shared epitope specificity.

With the exception of V segment TRBV20-1 found in CSP reactive TCRs across two epitopes (CS.T3, Th2R), TCRs targeting C-terminal epitopes revealed unique non-shared V segments associated with the CSP target epitope (Figure 9). The majority of TCRs targeting C-terminal epitopes (19/20) used a TRBV gene segment, that was enriched (>5%) in peptide specific expansion culture (Figure 9). Th2R

specific TCRs mostly used the TRAV17-1 gene (9/15) paired with TRBV3-1, TRBV12-3 or TRBV7-2, indicating a strong contribution of the alpha chain in TCR-pMHC interaction (Figure 9B). In addition to shared V segment usage among Th2R targeting TCRs, three groups of TCRs revealed high CDR3 amino acid sequence similarity. All three clusters of similarity identified by BLscore clustering algorithm were identified among TCRs originating from the same donor (F3). In contrast, TCRs targeting the CS.T3 and T* epitope lack high amino acid sequence similarity in their CDR3 region (Figure 9B).

In line with detection of diverse V gene segment enriched in expansion cultures from individual donors, the majority of V segments used by C-term-reactive TCRs (15/17, 88%) are linked to only one donor (Figure 9). The sharing of V segment genes was limited to TRAV13-1 and TRBV20-1 in Th2R targeting TCRs isolated from donor F3 and F4. The fact that TRBV20-1 was found across different target epitopes and presenting HLA alleles implies a differential involvement in either HLA or peptide binding. The diverse Th2R specific TCR repertoire supports the establishment of a potent T cell response against this supertope in different genetic backgrounds.



Epitope	TRBV	B_CDR3	TRAV	A CDR3	Cluster
Th2R	TRBV12-3	ASSRQGDQPQH	TRAV17-1	ATASGYSTLT	1
Th2R	TRBV12-3	ASSRQGDQPQH	TRAV17-1	ATASGYSTLT	1
Th2R	TRBV12-3	ASSRQGDQPQH	TRAV17-1	ATGSGYSTLT	1
Th2R	TRBV12-3	ASSGQGDQPQH	TRAV17-1	ATASGYSTLT	1
Th2R	TRBV18-1	ASSPGQGNGYT	TRAV13-1	AASLLGGSNYKLT	
Th2R	TRBV3-1	ASSQLQGAGNTIY	TRAV17-1	ATVRNTDKLI	
Th2R	TRBV5-1	ASSLGYNQETQY	TRAV9-2	ALRNAGGTSYGKLT	2
Th2R	TRBV5-1	ASSLGYNQETQY	TRAV9-2	ALRNAGGTSYGKLT	2
Th2R	TRBV6-4	ASSAGQGENGYT	TRAV17-1	ATASGYSTLT	
Th2R	TRBV7-2	ASSGQGDQPQH	TRAV17-1	ARSSGYSTLT	3
Th2R	TRBV7-2	ASSLQGDNEQF	TRAV17-1	ATASGYSTLT	3
Th2R	TRBV7-2	ASGGQGDQPQH	TRAV17-1	ASASGYSTLT	3
Th2R	TRBV7-7	ASSLQENTEAF	TRAV8-4	AVRVNQAGTALI	
Th2R	TRBV20-1	SVNPTGGRPTGQY	TRAV13-1	AASSNQAGTALI	
Th2R	TRBV20-1	SAPEGGGDIQY	TRAV13-1	AWLGGGSNYKLT	
CS.T3	TRBV20-1	SAAEGGRETQY	TRAV9-2	ALSDQGGGNKLT	
CS.T3	TRBV20-1	SASLLIQETQY	TRAV9-2	ALANQGGSEKLV	
CS.T3	TRBV20-1	SAASGTDGSQETQY	TRAV9-2	ALSGRG	
T*	TRBV30-1	AWRERLEPTGELF	TRAV29DV5	AASVDRGSTLGRLY	
CS.T3	TRBV6-1	ASSPAVNSYEQY	TRAV14DV4	AMRERPSGGYQKVT	

Figure 9: Donor-specific V segments link to C-terminal epitope specificity.

A + C V segment usage in C-terminal expansion cultures. V segments with frequencies above 5% are shown individually while low frequent V segments (<5%) are combined among "other". V segments with a link to CSP reactivity are color coded by donor. Alpha (A) and beta (C) V segment usage are shown separately. **B** Sequence data of C-term specific TCRs. Alpha and beta V segment usage and CDR3 sequence are linked to C-terminal epitopes. V segments are color coded by originating donor and frequency among expansion culture (pie charts).

Subsequently, V segment enrichment analysis after N-term specific expansion combined with *TR* gene sequence feature analysis of N-term specific TCRs was performed. N-term specific expansion induced

an increase in TRBV30-1, TRBV5-1 and TRBV3-1 V segments that could be linked to epitope specificity in donor F3 (Figure 10). T1-targeting TCRs exclusively harbor the TRBV30-1 V segment paired with diverse alpha chains, indicating an essential contribution of TRBV30-1 to the TCR-pMHC interaction. Likewise, TRBV5-1 was enriched in both N- and C-term cultures, paired with epitope-specific non-shared alpha chains, and could be linked to CSP61 and Th2R targeting (Figure 10). Since both epitopes are linked to HLA-DR presentation, TRBV5-1 is potentially involved in HLA allele specific binding. The TCR targeting the N-terminal CSP81 epitope by donor F4 is encoded by non-enriched alpha and beta chains, potentially associated with its weak avidity seen during *in vitro* stimulation (Figure 10, Figure 8).

TRBV7 was found enriched in all expansion cultures across C- and N-term stimulation and all donors, but could only be linked to CSP reactivity and Th2R specificity in donor F3. Thus, the higher frequency of TRBV7 V segment usage in CD8 T cells is most likely independent of a CSP-specific response. Accordingly, TRBV7 subgroups were identified as among the favored V segments in human T cells through antigen independent T cell repertoire sequencing [82, 83].

Taken together, I identified TCR sequence features associated with epitope specificity and presenting HLA alleles for both C- and N-term specific CD4 T cell responses, which could guide the identification of additional CSP-specific TCRs. Gene segment enrichment analysis in the peptide expanded repertoire is a potent indicator of antigen specificity. However, the detection of common sequence features irrespective of peptide specificity is a confounding factor and urges functional validation of antigen-specificity.

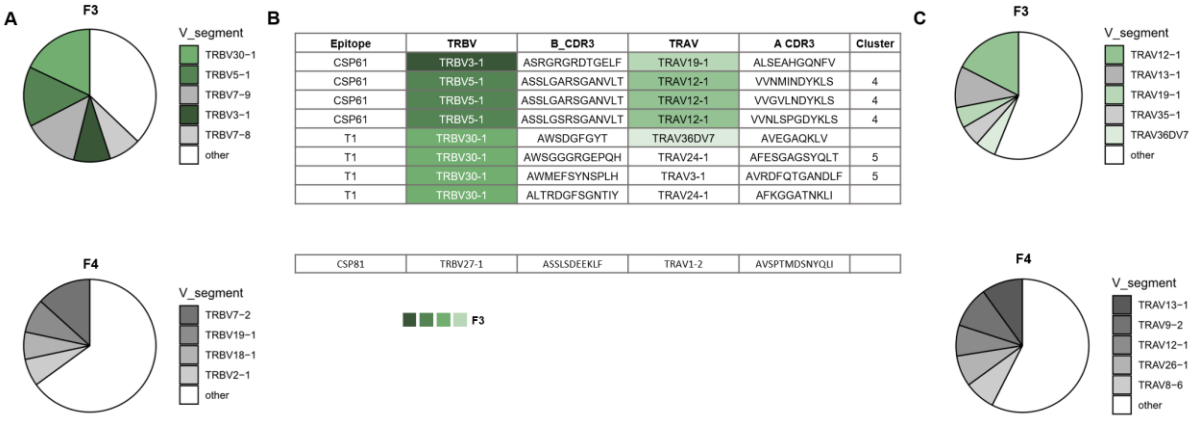


Figure 10: N-terminal epitopes are targeted by polyclonal T cells restricted to one donor.

A + C V segment usage in N-terminal expansion cultures of two donors (F3, F4). V segments with frequencies above 5% are shown individually while low frequent V segments (<5%) are combined among "other". V segments with a link to CSP reactivity are color-coded by donor. Alpha (A) and beta (C) V segment usage is shown separately. **B** Sequence data of N-term specific TCRs. Alpha and beta V segment usage and CDR3

sequence are linked to N-terminal epitopes. V segments are color-coded by originating donor and frequency among expansion culture (pie charts).

10.2.4. Epitopes outside Th2R/T* region reveal HLA restriction with low prevalence

To investigate the breath of the presenting HLAs for each CSP epitope, I assessed HLA allele sharing and associated epitope targeting across donors. I found shared DRB alleles, DRB*03:01 and DRB*07:01, associated with a Th2R specific response (Figure 11A). In addition, low frequent non-shared alleles DRB*13:01, DRB*13:02, DRB*10:01 and DRB*04:01 associate with ability to target the Th2R epitope highlighting that numerous HLA alleles can present the Th2R region. In contrast, targeting of both novel N-terminal epitopes is linked to unique non-shared HLA-DR alleles (DRB1* 13:01/13:02 of donor F3 and DRB1* 04:01/10:01 of donor F4) (Figure 11A). Due to the exclusive expression of both potential alleles by a single donor, the exact peptide presenting HLA allele remains unknown. This analysis reveals that the majority of donors harbor an HLA allele with capacity to induce a Th2R response while targeting of N-terminal and junctional epitopes are linked to a single non-shared HLA allele (Figure 11B). Thus, the breath of presenting HLAs is high for the immunodominant Th2R epitope, while the recognition of additional CSP epitopes is limited to single or very few presenting HLA alleles.

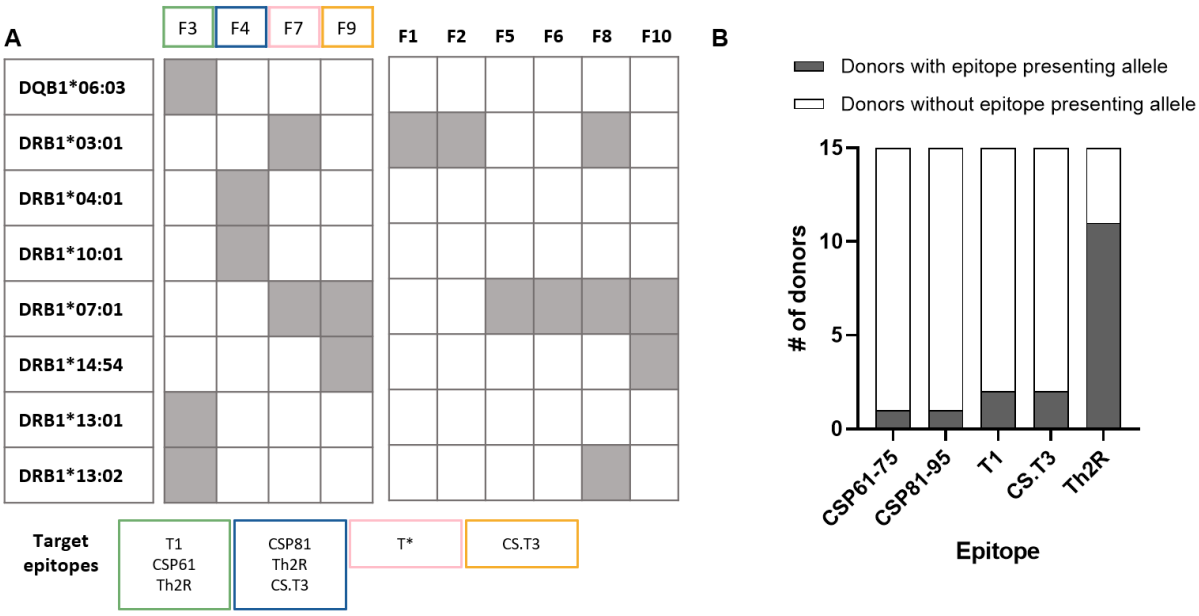


Figure 11: N-terminal epitopes are targeted by low frequent HLA-alleles.

A List of HLA alleles of ten donors (F1-F10). Grey boxes mark the presence of respective allele in individual donor. Donors with CSP-specific TCRs (F3, F4, F7, F9) and respective CSP target epitopes are highlighted. **B** Number of donors with HLA alleles known to present the respective CSP epitope.

10.2.5. Epitopes outside the Th2R/T* region are highly conserved across different *Pf* isolates

To define the degree of sequence conservation in CSP regions targeted by CD4 T cells identified in this study, I analyzed a publicly accessible data set of 481 CSP amino acid sequences from *Pf* field isolates originating from seven geographic regions for sequence conservation [84]. While the 15mer peptide sequence covering the CS.T3 and CSP61 epitopes were conserved in all 481 sequences, T1 and CSP81 were identical in 90% (435/481) and 99% (478/481) of the analyzed sequences, respectively (Figure 12). Unlike the high sequence conservation observed for epitopes outside the Th2R region, only six identical matches (1%) with vaccine sequence were identified for the highly polymorphic Th2R sequence (Figure 12) [81, 85]. Taken together, all non-Th2R peptides targeted by CD4 T cells are highly conserved across geographical regions.

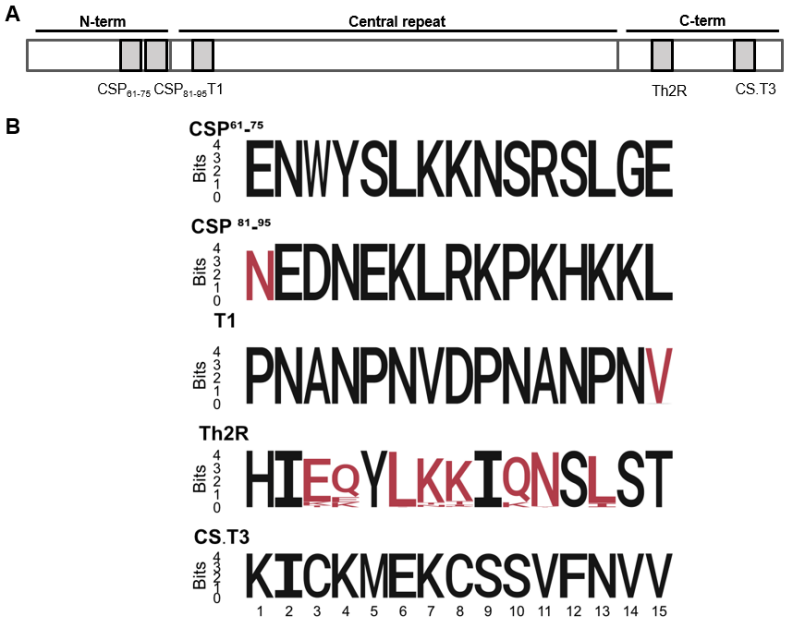


Figure 12: Epitopes outside the Th2R region are highly conserved.

A Schematic representation of the CS protein indicating the position of epitopes identified within this study. **B** Sequence diversity among 481 CSP sequences isolated from 7 geographic regions published by Tanabe, Mita et al. 2013 [84]. Sequence logos for each epitope are shown, polymorphic amino acids are highlighted in red.

10.2.6. The activated CD8 T_{EM} and T_{EMRA} pool is stable over subsequent FMP013 vaccinations

Next, I aimed to characterize the CD8 T cell response to subsequent FMP013 vaccinations using longitudinal peripheral blood samples of five donors (F1-F5) (Figure 5). To track changes in T cell composition longitudinally and allow for a subsequent targeted single-cell based analysis of the T cell

repertoire that is involved in the ongoing anti-vaccine response, I designed a staining flow cytometric panel including various immune cell and activation markers. CCR7 and CD45RA served as markers to distinguish T_{EM}, T_{EMRA} and T_{CM} from naive T cells. To further characterize the activation status of different T cell subsets, the T cell activation markers PD-1, ICOS, CXCR5 and CD137 were included [58, 86-88]. While large inter-individual differences in immune cell frequencies were detected, the activated CCR7- memory pool remained rather constant over the time of vaccination (Figure 13, B-C, Supplementary figure 3). In summary, the activation phenotype of the effector memory CD8 compartment is unaffected by subsequent FMP013 vaccinations.

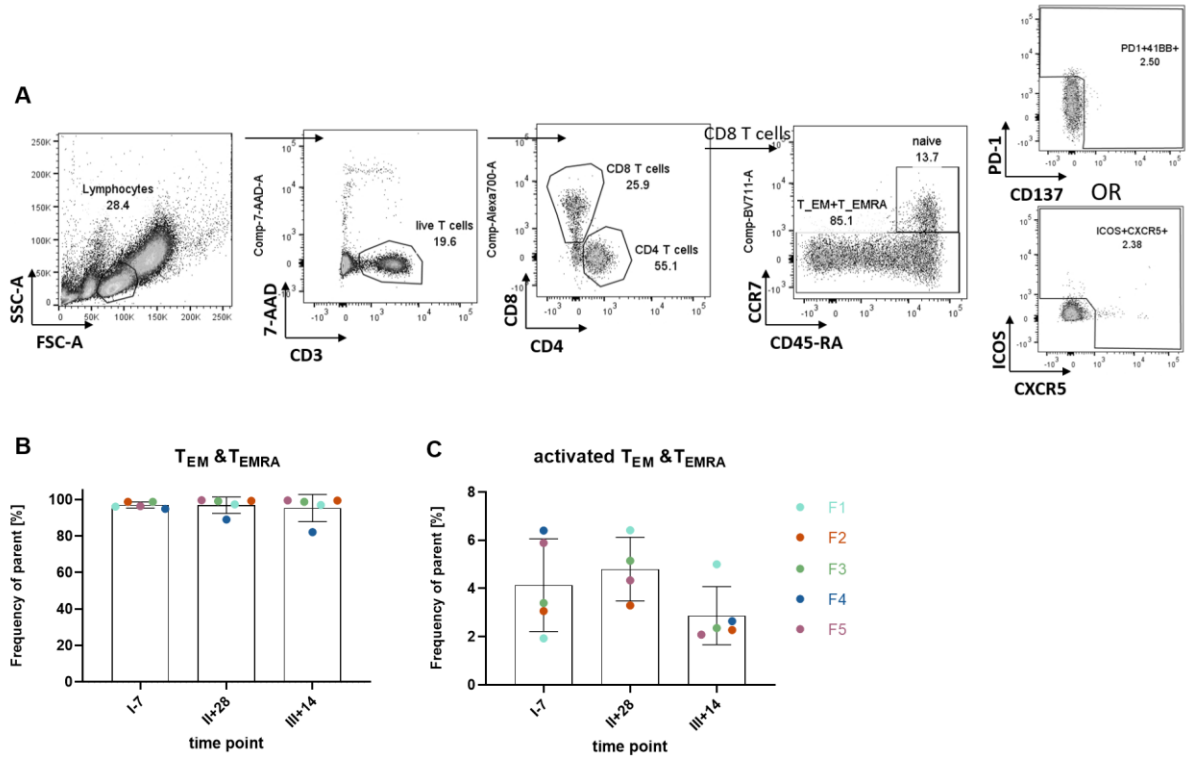


Figure 13: The activated TEM and TEMRA pool is stable throughout subsequent FMP013 vaccinations.

A Representative gating strategy for flow cytometric analysis of PBMC samples from individuals receiving FMP013 vaccination. **B-C** Frequency of effector memory (T_{EM}) or terminally differentiated effector memory (T_{EMRA}) (B) and frequency of activated EM and EMRA cells (C) are shown as frequency of respective gating parent for five donors receiving the high dose vaccination regimen. Bar plots of individual timepoints before first (I-7), post second (II+28) and post third (III+14) vaccination dose show mean and standard deviation with individual dots representing individual donors.

10.2.7. The activated CD8 repertoire is dominated by large pre-existing clones

To characterize the antigen-experienced repertoire after vaccination and study the impact of subsequent CSP vaccinations on the CD8 repertoire, cell sorting was focused on the T_{EM} (CCR7-

CD45RA-) and T_{EMRA} (CCR7- CD45RA+) memory compartment (Figure 14A), which have been previously reported to contain the majority of vaccine-specific T cells 2-4 weeks after vaccination [89]. Due to the complexity in T cell activation by timing and level of expression, I sorted activated T_{EM} and T_{EMRA}, that express at least one of the four activation markers (Figure 13C). In total, I isolated 5528 activated CD8 T cells using index single-cell sorting from five donors and three individual time points each. On average 250 cells per donor and time point were sorted and their *TRA* and *TRB* genes amplified and sequenced [71] (Supplementary figure 4; Supplementary Table 1). The sampling depth was increased for post 2nd and 3rd dose time point to enable the detection of low-frequent CSP-specific T cell clones.

Clonal composition analysis based on the TCR beta chain showed that the CD8 T cell repertoire is dominated by large clones, that are persistent throughout vaccination (Figure 14A). On average 60-70% of the repertoire was expanded before vaccination. The expanded fraction is decreased upon vaccination to 50-60% in all donors, suggesting the recruitment of small clones into the ongoing immune response (Figure 14, A-B). This finding is supported by the increase in the mean Shannon diversity from 3.1 to 3.9 (Figure 14C). Together this indicates the emergence of small clones, rather than rapid and strong expansion of pre-existing clones, in these malaria naïve individuals. Following individual clones over the course of subsequent vaccinations, several large clones shrank in size with subsequent immunizations. In contrast, several small clones were present at both timepoints after vaccination, but absent before the first vaccination dose, implying an active contribution to the vaccine response and potential CSP-reactivity.

To assess and quantify the independent emergence of several clones with distinct nucleotide but shared amino acid sequence in several donors, I conducted a search for public T cell clones. In this analysis a public clone was defined as identical CDR3 beta amino acid sequence, that was detected in more than one donor. The analysis revealed two public clones shared by donor F2 and F3 and donor F4 and F5, respectively. In support of CSP-specificity, one public clone was only detected after vaccination and contained several T cell clonotypes with distinct CDR3 beta nucleotide sequence in each donor. In contrast, the shared clone by donor F4 and F5 was already detected before vaccination, contradicting CSP-specificity (Supplementary figure 5). Apart from these two shared clones, the vast majority of clones (937/943, 99.4 %) are detected in only one donor. Taken together, the activated CD8 effector memory repertoire is dominated by large pre-existing clones and diversifies upon vaccination through the emergence of small clones unique to individual donors.

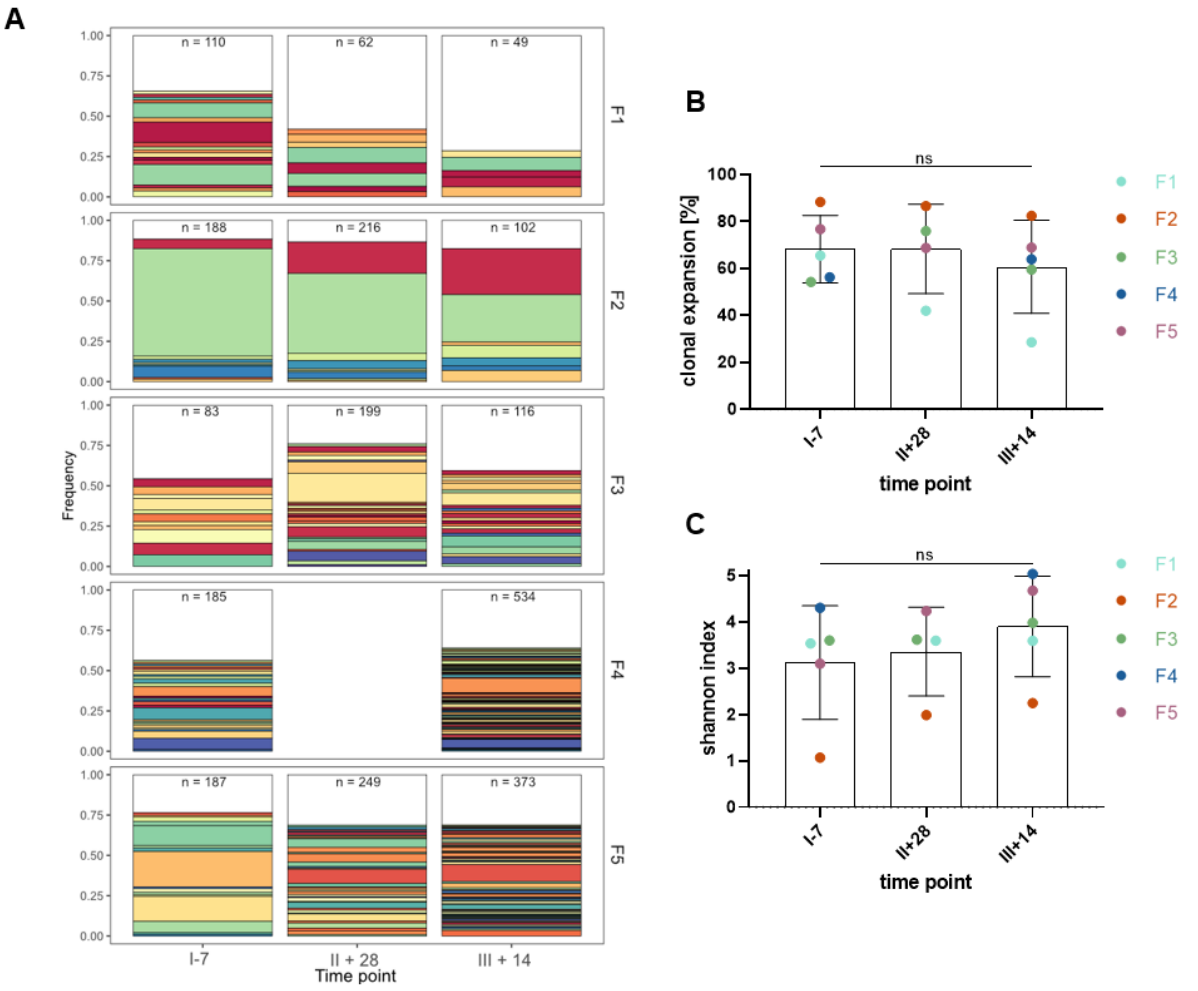


Figure 14: The post-vaccination T cell repertoire is dominated by large preexisting clones.

Activated CD8+ $T_{EM+EMRA}$ cells were index single-cell sorted, *TR* genes amplified and sequenced for subsequent repertoire analysis. **A** Clonal composition of the TCR repertoire in individual donors across multiple blood collection time points. Expanded clones are shown in color while unique TCR sequences are combined in the white compartment. Same clonal color within each donor represent the same clone, while color sharing across different donors does not. **B** Percent clonal expansion and **C** Shannon diversity index before and after FMP013 vaccination are shown for each donor individually. II+28 time point was not available for donor F4. Statistical significance was calculated by Wilcoxon test (B-C).

10.2.8. Activated T cells contain many TCRs with reported reactivities to common viruses

The detection of pre-existing clones after vaccination likely reflect the response to preceding or chronic infections of common viruses [90, 91]. To determine their potential specificity, I devised the TCRmatch tool for an in-depth data base search for TCRs with reported specificities as well as specificity prediction by sequence similarity to previously reported receptors [92]. Among the activated clones, in total 209 TCRs match with TCRs against common viruses inducing chronic (EBV, Cytomegalovirus (CMV)) and non-chronic (Influenza) infection (Figure 15). Especially viruses with high

prevalence and high number of published TCRs like CMV, EBV and influenza mapped to at least two TCRs in every donor. The fact that the FMP013 study was conducted before the emergence of SARS-CoV-2 pandemic explains the low number of TCRs mapping to the SARS coronavirus. Taken together, the activated CD8 repertoire throughout subsequent CSP vaccinations comprises a high number of pre-existing clones with specificities to common viruses.

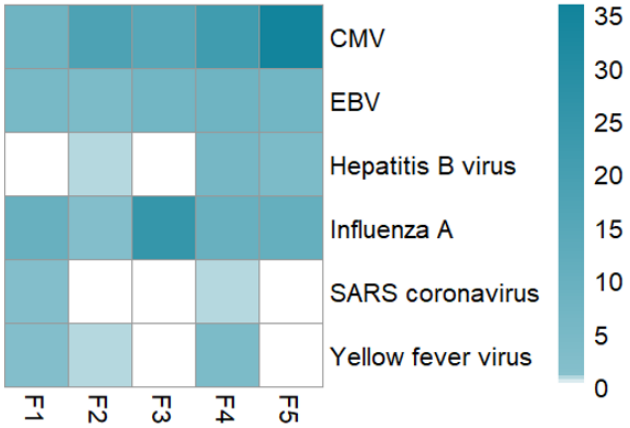


Figure 15: TCRs of large pre-existing clones can be linked to reactivity against common viruses.
Number of TCRs with beta chain match in IEDB found identified by TCRmatch tool. Identified TCRs are stratified by donor (F1-F5) and pathogenic origin of the TCR epitope. CMV: Cytomegalovirus, EBV: Epstein-Barr virus

10.2.9. Convergent clones appear with higher frequency after vaccination

To detect clonal convergence among emerging clones after vaccination, a clustering algorithm, BLscore, was employed to identify clusters of convergent TCRs by alignment of CDR1-3 sequences [58]. Within all detected sequences, 55 clones clustered in 19 similarity clusters based on high similarity in TCR alpha and beta chain (Figure 16). While only one or two clusters were identified in donors F1, F2 and F3, eight and seven clusters were identified in donor F4 and F5, respectively. Seven and six of them only comprised clones emerging after vaccination, respectively, whereas all clusters identified in donors F1, F2, and F3 included T cell clones detected pre-vaccination, potentially indicating a stronger vaccine-induced response in donors F4 and F5 (Figure 16). The two largest clusters with seven and nine clonal members, respectively, did not include any clone present pre-vaccination (Figure 17). Overall, the higher frequency of convergent clones in the post-vaccination repertoire suggests a vaccine-induced emergence of similar TCRs with high potential of same epitope specificity.

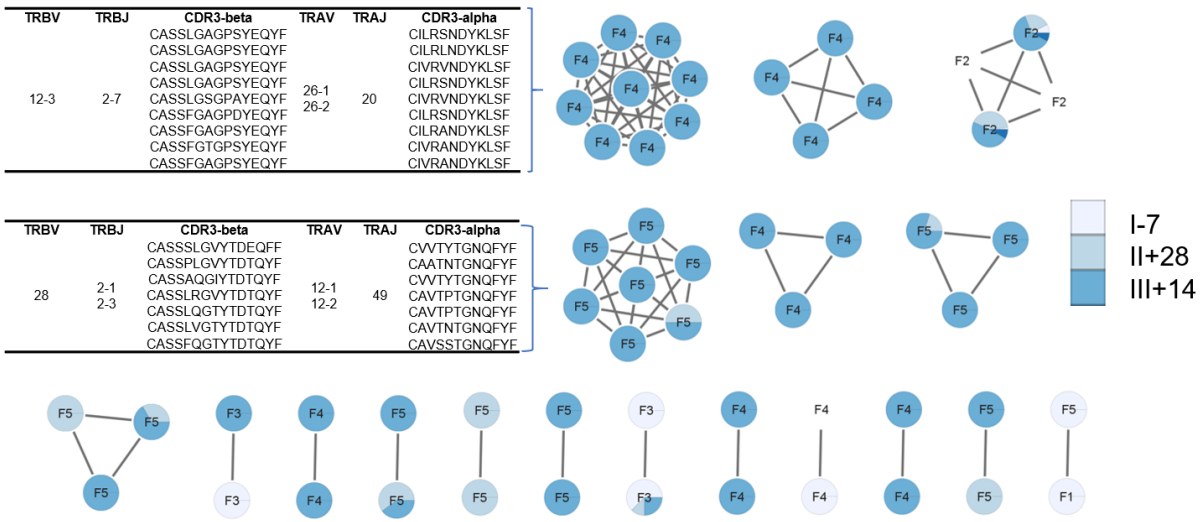


Figure 16: The post-vaccination repertoire contains higher frequency of convergent clones compared to pre-vaccination.

Clusters of similarity of CCR7- effector memory CD8 T cells based on BLscore algorithm. Each circle indicates a TCR clone, labelled with the ID of the originating donor, and pie charts indicating clonal frequency at individual time points. Sequence features of the two largest clusters of similarity are shown exemplarily.

10.2.10. The post-vaccination repertoire reveals donor-specific enrichment of V segments

To identify TCR sequence features associated with antigen-driven selection without the requirement of high amino acid CDR3 sequence similarity, enrichment of TCR sequence features was assessed longitudinally and compared across donors. Analysis of the donor-specific V segment usage revealed highly heterogenous *TR* gene segment usage with detection of 35-40 individual beta V genes and 35-40 alpha V genes per donor, respectively. The frequency of individual V genes stratified by collection time point discloses both preferential enrichment as well as contraction of individual genes with subsequent vaccination. To correct for differences in sampling depth between individual time points and donors, the relative abundance of each V segment was calculated relative to pre-vaccination level and plotted against the corresponding p-value of each comparison. Defining a p-value threshold of 0.05 and fold change of above 2, several donor-specific beta V segments were significantly enriched in donor F1-F4 at both post-vaccination time points (TRBV27-1 in donor F1, TRBV28-1 in donor F2, TRBV19-1 in donor F3, TRBV12-3 in donor F4), indicating preferential use upon vaccination (Figure 17A). In donor F5, six V segments revealed significant increase after vaccination, of which four were detected above the defined thresholds at both post-vaccination timepoints (Figure 17A). The strong overlap in significantly enriched donor-specific V segments across post-immunization time points indicates a constant recruitment of clones with these beta V segments into the ongoing immune

response. Using the same approach for alpha V segment usage, identified only TRAV29DV5 V segment as preferentially used in TCRs from donor F1 (Figure 17B).

While the V segments of both TCR chains can contribute independently to the mode of TCR-pMHC interaction, epitope specificity can be linked to distinct TCR V segment pairing [93]. The closer analysis of alpha and beta V segment pairing in all five donors and at three individual timepoints revealed distinct pairing patterns. Upon vaccination, one to four TRBV-TRAV pairings were enriched per donor, with limited overlap of gene segments shared across donors. Among the enriched TRBV- pairings, four of the eleven beta segments were identified in V segment enrichment analysis, whereas only one of the ten identified segments were included. TCRs emerging after vaccination showed preferential use of TRAV26-1+TRBV27-1, TRAV8-6+TRBV28-1, TRAV26-2+TRBV7-2 and TRAV19-1+TRBV9-1 pairings in donors F1, F2, F4 and F5, respectively (Figure 17C). Likewise, most enriched V segment pairings were found across both post-vaccination timepoint.

In addition to preferential V segment usage, changes in CDR3 sequence length distribution can also imply a vaccine-induced recruitment of similar TCRs. The analysis revealed donor-specific CDR3 beta length distribution from 10-17 amino acids in all five donors with distinct dominating length before vaccination (Supplementary figure 6). Upon vaccination, the dominant CDR3 beta length of 10 amino acids in donor F2 shifted to a dominant length of 11 amino acids after second vaccine dose and 13 amino acids after third vaccine dose. In contrast, TCRs from donor F5 showed dominant CDR3 length of 14 amino acids before vaccination and 12 and 13 amino acids after vaccination. In summary, the detailed longitudinal analysis of TCR sequence features revealed a significant enrichment of donor-specific beta chain V segment usage and alpha-beta pairing in all donors, whereas preferred alpha chain V segment usage and CDR3 length was only detected in one or two donors.

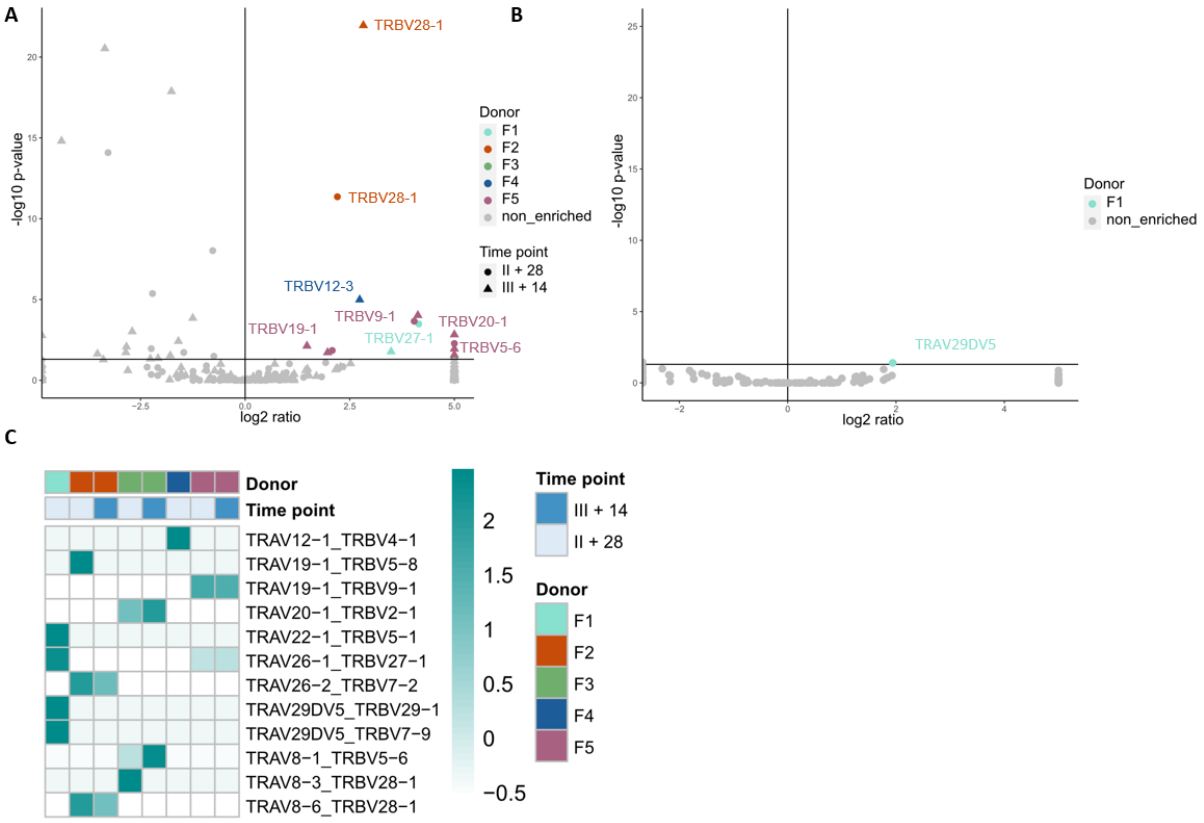


Figure 17: The post-vaccination repertoire reveals donor-specific enrichment of V segments.

A-B V segment frequencies of individual donors (F1-F5) normalized to the number of detected clones per time point. Frequencies of beta (A) and alpha (B) V segments relative to pre-vaccination time point are plotted against the respective p-value of the comparison. **C** Heatmap of alpha-beta V segment pairing enriched after vaccination in at least one of the donors and time points (> 5% cumulative clonal frequency in at least one post vaccine timepoint, >1 ratio compared to pre-vaccination time point).

10.2.11. TCRs that showed signs of antigen-driven selection are not CSP-specific

To investigate whether the donor-specific changes in sequence features or emergence of convergent clones after vaccination indeed relate to CSP-specificity, a selection of TCRs were cloned and stably expressed in Jurkat76-CD8 T cells using retroviral transduction for further functional characterization. In total, 97 TCRs covering all five donors were selected based on preferential beta V segment usage, clustering by BLscore algorithm and clonal expansion after vaccination (Figure 18, A-B; Supplementary Table 3). Despite identical selection criteria across donors, the highest number of selected TCRs (71%, 69 of 97) originated from donor F5 with high prevalence of clones expanded after the vaccination that were not detected at the pre-vaccination time point (Figure 18A). To increase the likelihood of identifying CSP-reactive clones, 71% of the selected TCRs exhibited V segment usage enriched after vaccination and detection of convergent clones and/or clonal enrichment after vaccination in 32% of the selected TCRs (Figure 18B).

Subsequently, 77 cloned TCRs were stably introduced into TCR-deficient Jurkat cells and transduction efficiency confirmed by mCherry positivity and TCR surface expression (Figure 18C, Supplementary figure 2). 62 TCR with robust TCR surface expression were screened for CSP reactivity by stimulation with autologous B cells presenting three pools of overlapping peptides covering the full-length protein in addition to a negative control of autologous B cell culture without external peptide loading. IL-2 detection after 24h co-culture in the culture supernatant by ELISA did not reveal any CSP-specific T cell activation. However, two TCRs exhibited positive IL-2 signal in all stimulatory conditions including the negative control, indicating reactivity against an EBV-derived peptide, originating from EBV-induced B cell immortalization. Thus, although I could detect co-culture induced activation of TCR-transgenic T cell line, none of the selected clones were specific to CSP.

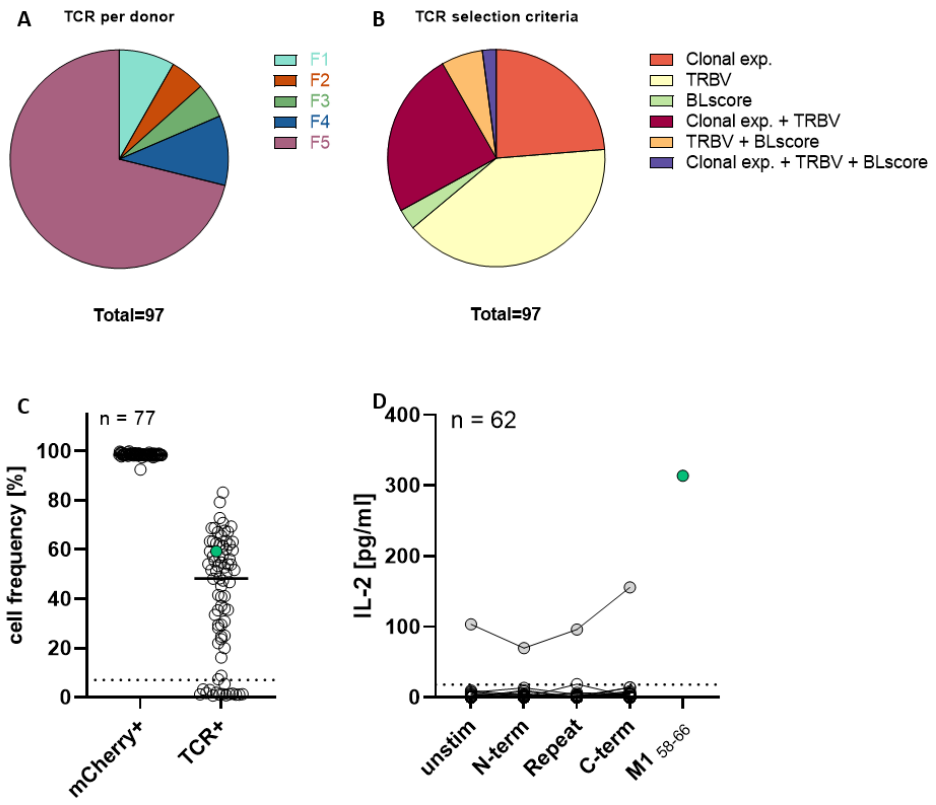


Figure 18: Enriched sequence features and clonal expansion of post-vaccination T cell clones are independent of CSP reactivity.

A-B 97 TCRs were selected for functional characterization covering all five donors (A) and based on fulfillment of one or several selection criteria (B). **C** Transduction efficiency of TCR-transgenic Jurkat76 cell lines based on mCherry (mCherry+) and TCR (TCR+) expression quantified by flow cytometric analysis. 77 transgenic T cell lines were generated by retroviral transduction in comparison to a control TCR (green dot). 62 T cell lines with TCR expression above 7% (dotted line) were subjected to subsequent CSP reactivity testing. **D** Activation of 62 T cell lines stimulated with CSP peptide loaded autologous B cell lines and influenza M1 specific control TCR (green dot). IL-2 was quantified by ELISA in the supernatant after 24h co-culture. Each T cell line was stimulated with a N-term, junction/repeat and C-term peptide pool in comparison to a non-stimulated control. Data were generated together with Sara Kraker.

10.2.12. CSP peptide expansion induces CD4 dependent bystander activation among CD8 T cells

To increase the likelihood of detecting rare CSP-specific CD8 T cells and to enrich for CSP-specificity, I stimulated peripheral blood samples from II+28 time point of five donors (F1-F5) for ten days *in vitro* with CSP peptide pools covering the N-term/junction or C-term and measured activation by flow cytometric analysis using CD69 and CD137 as markers (Figure 19, A-B).

Cultures from 3 out of 4 donors (F1, F3 and F4) exhibited activation upon stimulation with C-term peptides with activation fold changes between 1.9 and 2.8 compared to unstimulated controls,

whereas only donor F3 showed strong activation fold change (4.6) in response to N-terminal peptide expansion compared to the unstimulated control (Figure 19B). Interestingly, the level of CD8 T cell activation upon stimulation strongly correlated with the level of CD4 T cell activation in individual donors (Figure 19C). To closely investigate the clonal composition of CD8 T cells and quantify non-specific bystander activation, activated CD8 T cells from three donors (F1, F3 and F4) were singe-cell sorted and their *TR* genes amplified and sequenced (Supplementary Table 1). I detected a diverse clonal composition in both N- and C-term cultures with comparable level of expansion, exhibiting strong clonal overlap across stimulation cultures in donors F1 and F3 indicating antigen-independent activation and proliferation (Figure 19D). Taken together, a substantial fraction of activated CD8 T cell clones expanded after *in vitro* stimulation was bystander-activated and non-specific to CSP peptides.

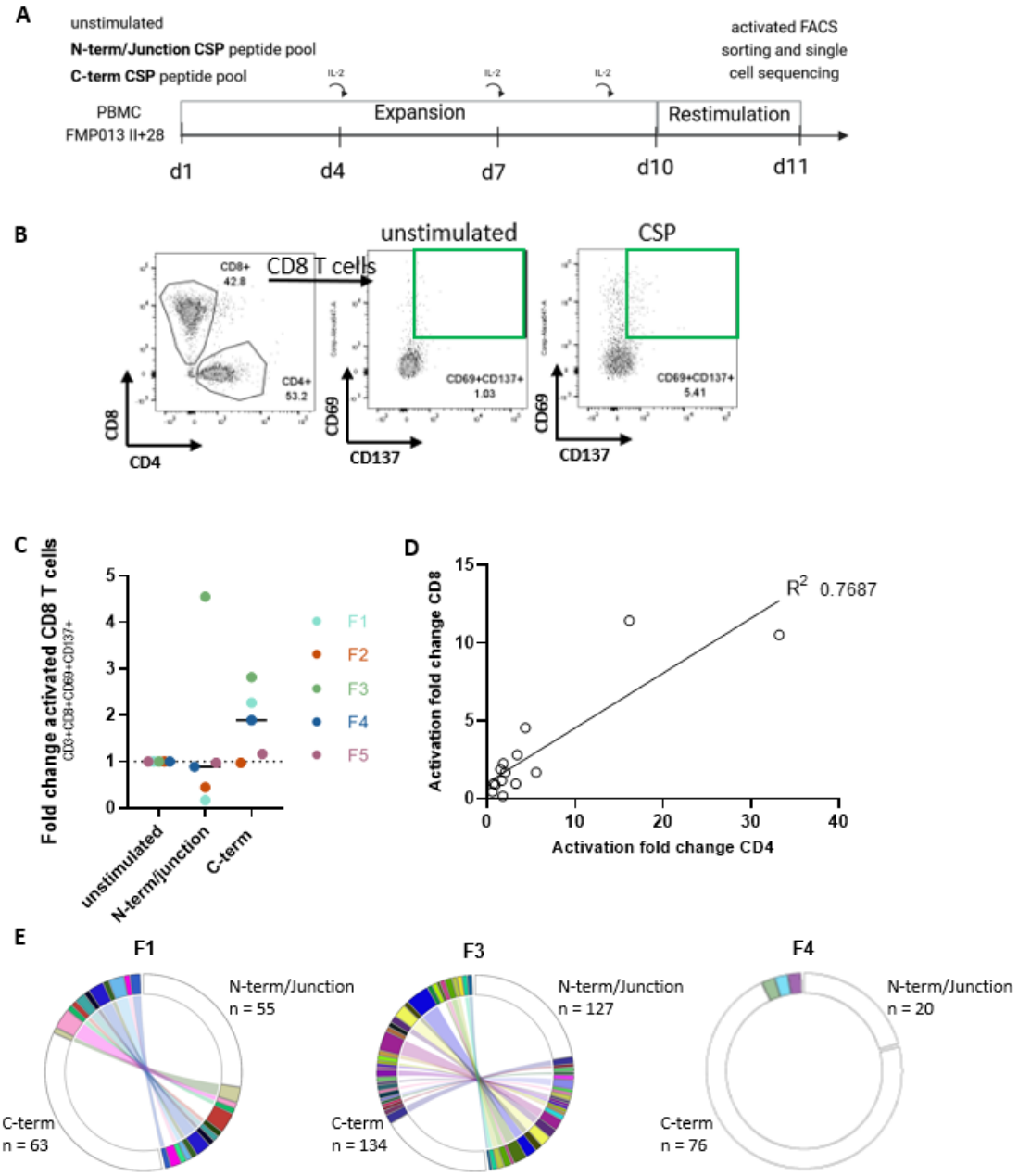


Figure 19: CSP peptide expansion induces CD4-dependent bystander activation in the CD8 T cell compartment.

A Experimental setup of 10-day T cell expansion assay. PBMCs are stimulated with CSP peptides covering the C- or N-terminus of the protein. On day 10 cells are re-stimulated for 18h with N-terminal or C-terminal subpools. Activation marker expression is assessed on day 11 by flow cytometric analysis. **B** Gating strategy after restimulation on day 11. **C** Fold change in frequency of activated CD8 T cells after N-term or C-term specific expansion compared to the unstimulated control. **D** Correlation of activation fold change among CD4 and CD8 T cell in individual donors. **E** Clonal composition and overlap of CSP expanded PBMC cultures. Expanded clones are shown in color and nonexpanded clones are combined in white. Shared clones across separate expansion cultures are connected by colored ribbon.

10.2.13. Functional characterization of expanded clones reveals only one CSP-specific TCR

To investigate whether clones that expanded in response to only one peptide pool might be CSP-specific, I performed a repertoire analysis including V segment usage, V segment pairing, clone size and TCR convergence. The TRBV segment usage was diverse in cultures from all donors. Differences in V segment usage were only observed in C-term and N-term expanded cultures from donor F3. TRBV4-1 and TRBV19-1 showed preferred usage in C-term expanded CD8 cultures, whereas TRBV5-1 and TRBV12-4 was used by several clones expanded after N-term-specific stimulation. To identify TCRs with high sequence similarity, similarity clustering analysis identified three clusters of convergent TCRs all containing overlapping clones detected in separate expansion cultures rather pointing towards bystander activation induced expansion. Similarity analysis across the non-expanded repertoire identified nine TCRs clustering with clones that showed clonal expansion after vaccination. To investigate whether TCRs are specific to CSP, I generated 31 cell lines with robust TCR expression from five donors (F1, F3, F4, F7, F9) for further functional characterization (Figure 20C, Supplementary figure 2). 40 TCRs were initially selected based on a scoring system including criteria like peptide culture specific clonal expansion, TRBV segment usage, clustering by BLscore algorithm and clonal overlap non-expanded repertoire (Figure 20, A-B; Supplementary Table 4). The majority of TCRs (57%; 23/40) were selected based on V segment usage, of which eight of them (20%; 8/40) also met criteria in one of the other selection categories including clonal expansion and overlap with non-expanded repertoire (Figure 20B).

To assess CSP reactivity of TCR-transgenic T cell lines, co-culture stimulation assay with major CSP peptide pools (N-term, junction/Repeat, C-term) were performed. Only one T cell clone, derived from donor F9, revealed CSP reactivity (3%, 1 of 31) (Figure 20D). Subsequent downstream stimulation with sub-pools and individual peptides allowed epitope identification of the TCR expressing cell line (Figure 20, E-F). TCR B11 targets the C-terminal epitope located at CSP position 367-381, previously identified as MHC class II restricted epitope termed CS.T3. In addition to CSP-specific T cell clone, two T cell clones with shared TRBV7-9 and 12-2 V segment usage responded to peptide-loaded as well as unloaded autologous B cell co-culture, indicating EBV-specificity. Thus, also among activated CD8 T cell clones that expanded exclusively in response to C-term or N-term stimulation but not both, the frequency of CSP-specific T cells was low (1/31, 3%).

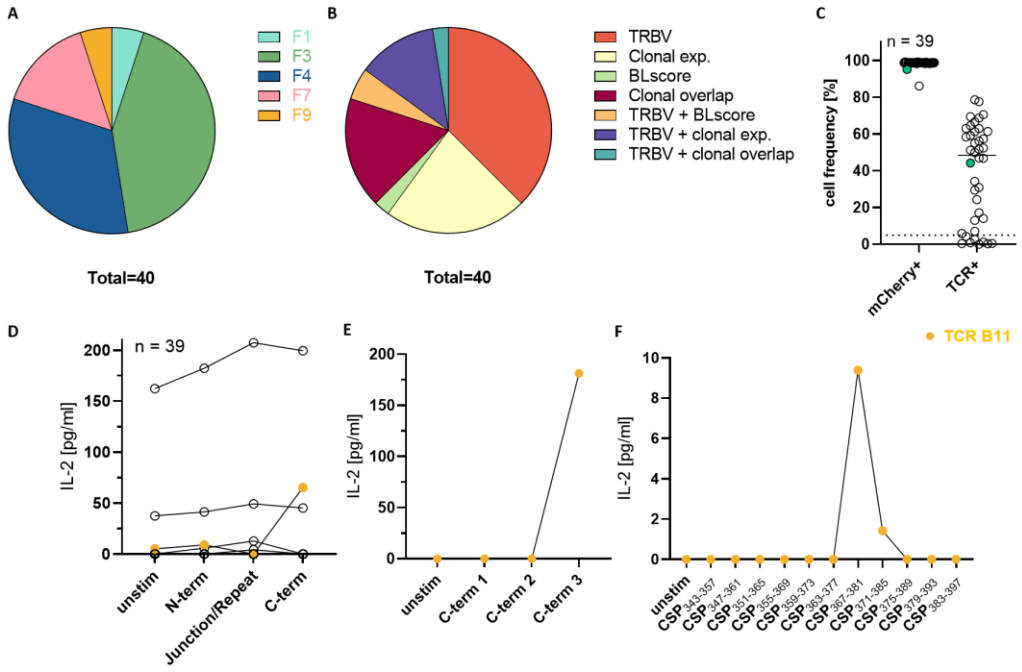


Figure 20: Low number of CSP-specific TCRs among peptide-expanded CD8 T cell clones.

A-B 40 TCRs were selected for functional characterization covering all five donors (A) and based on fulfillment of one or several selection criteria (B). **C** Transduction efficiency of TCR-transgenic Jurkat76 cell lines based on mCherry (mCherry+) and TCR (TCR+) expression quantified by flow cytometric analysis. 39 transgenic T cell lines were generated by retroviral transduction in comparison to control TCR (green dot). 31 T cell lines with TCR expression above 7% (dotted line) were subjected to subsequent CSP reactivity testing. **D** Activation of 31 T cell lines after stimulation with CSP peptide pools loaded onto autologous B cell lines quantified by IL-2 ELISA in the supernatant after 24h co-culture. C-term reactive TCR B11 is shown in yellow. **E-F** Activation of TCR B11 expressing T cell line after stimulation with C-term sub pools 1-3 (E) and single 15mer peptides contained in peptide pool C-term 3 (F) in comparison to unstimulated control.

10.2.14. TCR B11 targets CS.T3 in a co-receptor independent manner

The CSP reactive TCR B11 exhibited a shared CD4 and CD8 phenotype in PBMC expansion culture: two cells displayed a clear CD8 phenotype in flow cytometric analysis and five cells stained positive for CD4 expression (Figure 21A). To characterize TCR B11 activation in both co-receptor contexts, both co-receptor expressing TCR-transgenic Jurkat cell lines were generated and showed 68 % and 47 % TCR expression in the context of the CD4 or CD8 co-receptor, respectively. Peptide titration of the target epitope revealed similar concentration dependent activation kinetics for both cell lines (Figure 21B) suggesting a TCR-pMHC interaction induced activation independent of co-receptor association.

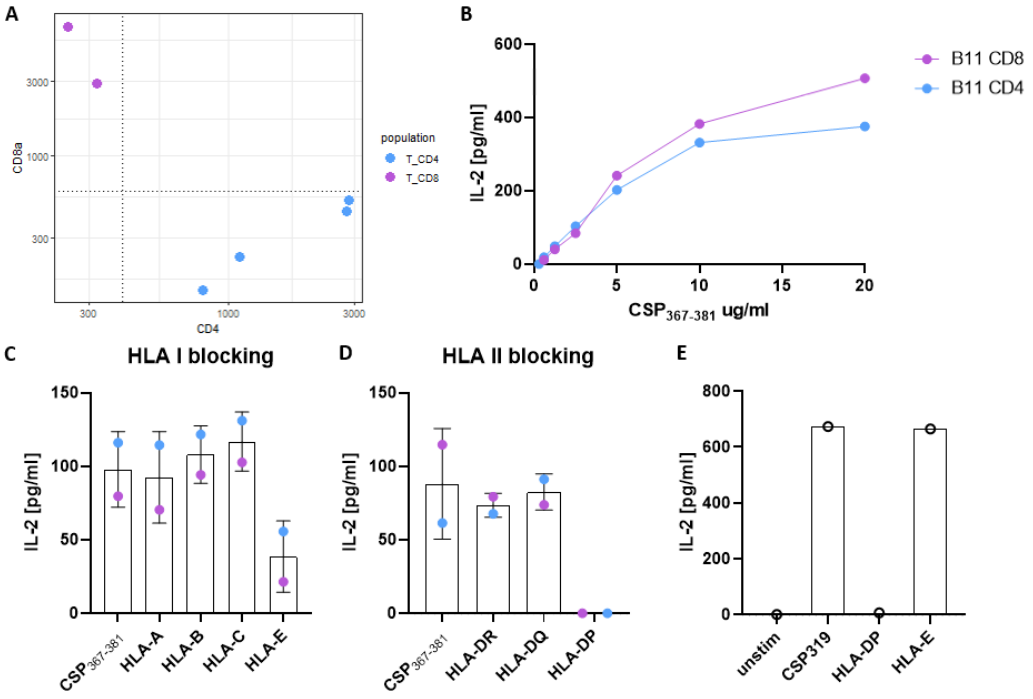


Figure 21: TCR B11 targets CS.T3 in a co-receptor independent manner.

A CD4/CD8 phenotype of B11 T cell clone in C-terminal expansion culture detected by index single-cell sorting. **B** Peptide titration of 15mer peptide covering the CS.T3 epitope ($CSP_{367-381}$) in context of CD4 or CD8 coreceptor expressing Jurkat76 cell line. **C** Activation after 24h co-culture stimulation with $CSP_{367-381}$ peptide-loaded autologous B cells in presence of locus specific HLA blocking antibodies. **D** Activation level (IL-2 expression) of HLA-DP restricted control T cell line in presence of target peptide and HLA-DP and HLA-E specific blocking antibody.

Next, I aimed to identify the presenting HLA context required for T cell activation. In the same co-culture assay, TCR B11 transgenic cell lines were stimulated in the presence of individual HLA class I and class II blocking antibodies. During screening of locus-specific TCR-pMHC blockings, both HLA-E and HLA-DP blocking resulted in a substantial decrease in IL-2 secretion, suggesting the presentation of the target peptide in both HLA contexts (Figure 21, C-D). To eliminate the possibility of potential cross-reactivity of the blocking antibodies, I performed blocking of another HLA-DP restricted CSP-specific T cell line, which showed absence of an HLA-E blocking effect (Figure 21E). The absence of T cell activation in presence of HLA-DP and HLA-E antibody in both co-receptor contexts further supports the co-receptor independence of TCR B11.

To determine the minimal epitope within the CS.T3 epitope region in both presenting HLA contexts, I performed stimulation with overlapping 9mer, 10mer and 11mer pools covering the 15mer amino acid sequence in the presence and absence of HLA-DP and HLA-E blocking antibodies. The quantification of IL-2 in the supernatant revealed a minimal length of ten amino acids to induce weak

activation of B11 TCR and strongly increased activation by the 11mer pool (Figure 22, A-B). Further stimulation with individual peptides revealed MEKCSSVFNVV as the strongest activator, while KCSSVFNVV was identified as minimal epitope essential for weak activation of the B11 TCR (Figure 22, A-B). Strikingly, HLA-DP blocking resulted in complete absence of activation, while HLA-E blocking slightly enhanced detected activation levels (Figure 22, C-D).

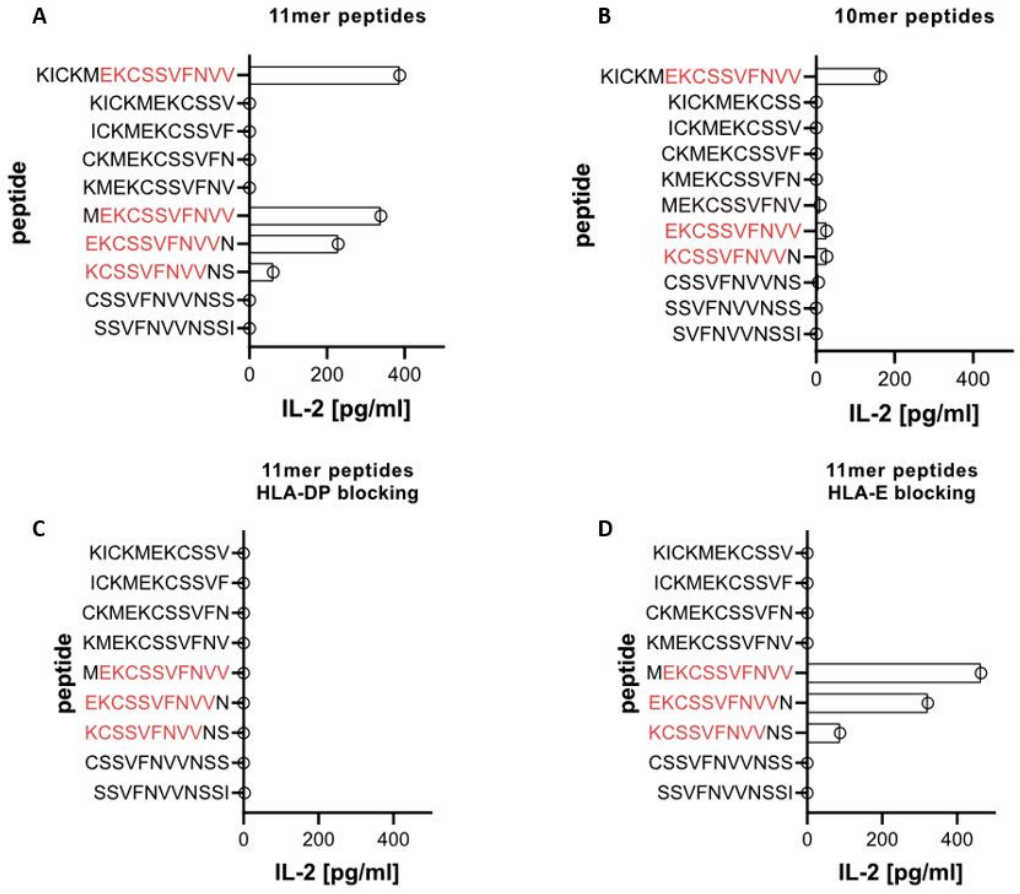


Figure 22: The minimal epitope EKCSSVFNVV is targeted in an HLA-DP restricted context.

A-B Activation level quantified by IL-2 in the supernatant after stimulation with single 10mer (A) and 11mer (B) peptides covering the CS.T3 epitope in comparison to 15mer peptide stimulation. **C-D** Activation level quantified by IL-2 in the supernatant after stimulation with single 11mer in presence of HLA-DP (C) or HLA-E (D) specific blocking antibody. Minimal epitope which leads to activation phenotype is highlighted in red.

To uncouple potential pMHC interaction and TCR binding from functional activation, I employed both MHC binding predictions as well as TCR-pMHC model predictions to asses binding of CS.T3 epitope region to the HLA-E molecule. Within the human population, the HLA-E gene exhibits strong conservation, with only two alleles differing by a single amino acid. Peptide binding prediction analysis revealed KMEKCSSVF as predicted strong binder for HLA-E*01:01 by NetMHCpan4.1. Strikingly this peptide is the only predicted strong binder within the full CSP sequence. Further modelling of the TCR

B11 interaction with the peptide HLA-E complex by ImmuneScape supported the high potential of the KMEKCSSVF peptide presented by the HLA-E*01:01 allele. Lysin, methionine and phenylalanine at position 1,2 and 9 serve as anchor residues by building several hydrogen bonds and pi stacking with the HLA-E molecule (Figure 23). In line with an absence of HLA-E dependent activation by KMEKCSSVF peptide, our TCR-pMHC model predicts limited binding capacity of TCR B11 to the pMHC complex. The CDR1 region of the alpha chain can potentially form hydrogen bonds with the HLA molecule to stabilize the interaction. The CDR3 beta chain forms the main contact point with pMHC complex based on model predictions. Given the high potential of the CS.T3 epitope being presented by HLA-E*01 alleles, I aimed to express a HLA-E*01 dextramer in order to assess successful peptide binding and TCR binding subsequently. In collaboration with Immudex, a stable dextramer loaded with KMEKCSSVF peptide could be produced implying successful peptide binding. Nonetheless, subsequent staining of the TCR B11 expressing T cell line did not show TCR dependent binding of the dextramer (Supplementary figure 7).

Even though, peptide prediction and successful dextramer production indicate efficient binding of the CSP-derived peptide to HLA-E molecules and HLA-E blocking can affect B11 activation levels, dextramer staining and minimal epitope identification revealed a strong preference in binding to the HLA-DP presented peptide in the interaction with TCR B11.

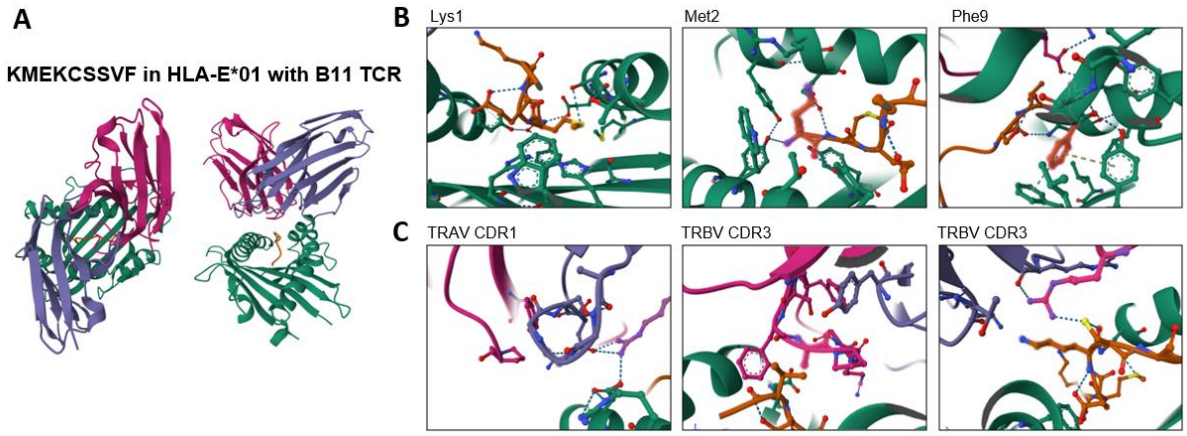


Figure 23: CS.T3 core peptide KMEKCSSVF strongly binds HLA-E*01 molecule.

A Structural model of B11 TCR with KMEKCSSVF presenting HLA-E*01:01 molecule modelled by ImmuneScape. Top view (left) and front view (right) of complex formed by peptide (orange), HLA molecule (green), TCR alpha chain (purple) and TCR beta chain (pink) are shown. **B** Hydrogen bonds and pi stacking interaction between peptide anchors and MHC molecule. **C** Interaction sites of TCR alpha and beta chain with peptide.

10.2.15. *In silico* peptide binding predictions reveal shared CSP binders across diverse HLA alleles

To explore potential reasons for the absence of an T cell response restricted to classical MHC class I molecules, I investigated MHC peptide binding across individual alleles and donors. An MHC binding prediction algorithm, NetMHCpan4.1, was devised to assess the potential binding capacity of CSP peptides in individual HLA contexts across ten donors (Supplementary figure 8A). Although the study's participant size of ten was relatively small with respect to the vast diversity of HLA haplotypes within the population, this study encompasses the majority of HLA supertypes described for both HLA-A and HLA-B alleles with shared peptide binding characteristics (Supplementary figure 8B) [94]. Numerous peptides exhibited predicted binding in many HLA haplotype contexts and only limited number of peptides are predicted per HLA allele (Figure 24). Upon further investigation of previously identified CSP epitopes and their corresponding HLA contexts, it became apparent that the same peptide regions along the amino acid sequence were implicated. In line with previous studies, the C-terminus contained six predicted epitope regions, while the N-terminus only exhibited two peptides with predicted binding across several HLA alleles (Figure 24). Furthermore, the signal peptide contained two peptide regions, that exhibited predicted binding by most HLA alleles included in this analysis. Thus, based on MHC peptide binding predictions, several CSP peptides can be presented by different HLA molecules and do not limit the immunogenic potential of an MHC class I restricted CSP response.

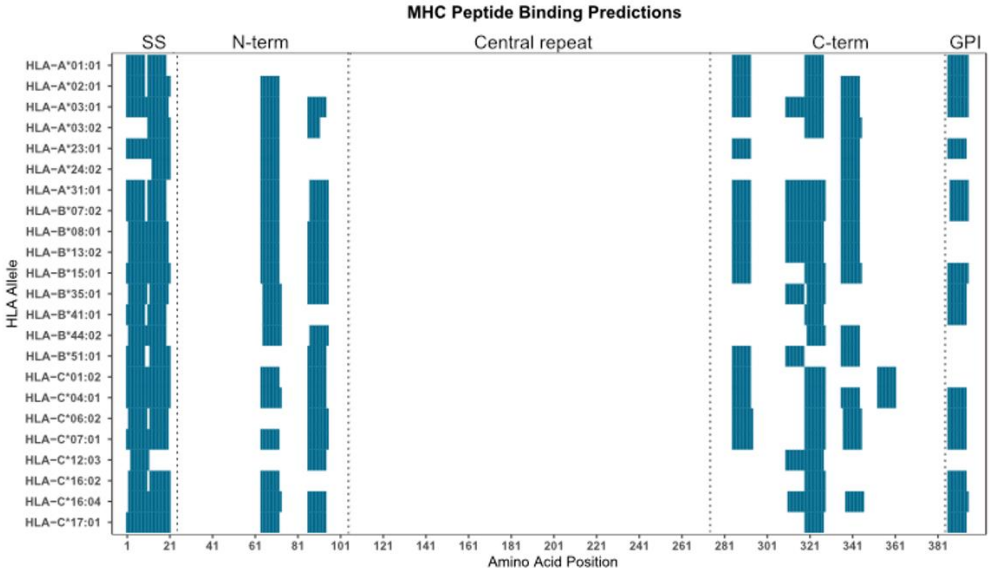


Figure 24: Diverse HLA alleles reveal shared peptide binding along CSP sequence.

Peptide binding prediction by HLA alleles of ten FMP013 donors along CSP sequence. Predictions were performed using by NetMHCpan4.1 tool and predicted binders are highlighted in blue.

10.2.16. *In vitro* binding shows diverse binding pattern across distinct HLA alleles

To investigate the predictive performance of NetMHCpan4.1, I performed an *in vitro* binding assay for a selection of CSP peptides and HLA alleles [22]. Five HLA alleles were selected based on their prevalence among the FMP013 donors. Here, the most common HLA-A alleles, HLA-A*02:01 and HLA-A*01:01 were included to cover the HLA context of previously published CSP epitopes with commercially available MHC monomers. In addition, a small number of less common alleles, HLA-A*11:01, HLA-B*07:02, HLA-B*08:01, and known association with at least one CSP epitope were selected. In total, seven peptides covering the major predicted sequences located in the N- and C-terminal domain of the CSP were evaluated by MHC loading and flow cytometric analysis through ß2-miroglobulin antibody staining (Figure 25). HLA-A molecules bound only one or two peptides, whereas HLA-B molecules revealed binding of six out of seven tested peptides (Figure 25A). Peptides CSP319 and CSP336, corresponding to T cell epitopes in Th2R and Th3R regions, showed the strongest binding for three out of the five tested alleles (Figure 25B). The comparison of *in vitro* peptide binding with *in silico* prediction revealed a correlation between predicted strong binders and higher *in vitro* binding compared to weak binders (Figure 25C). However, it was observed, that several peptides predicted as non-binders exhibited strong binding *in vitro*. These unpredicted peptide binders were identified across four out of five alleles, suggesting that the limited predictive performance is not specific to

particular alleles (Figure 25C). In summary, only a subset of predicted binders by NetMHCpan4.1 revealed peptide binding to matched HLA monomer during *in vitro* MHC binding assessment.

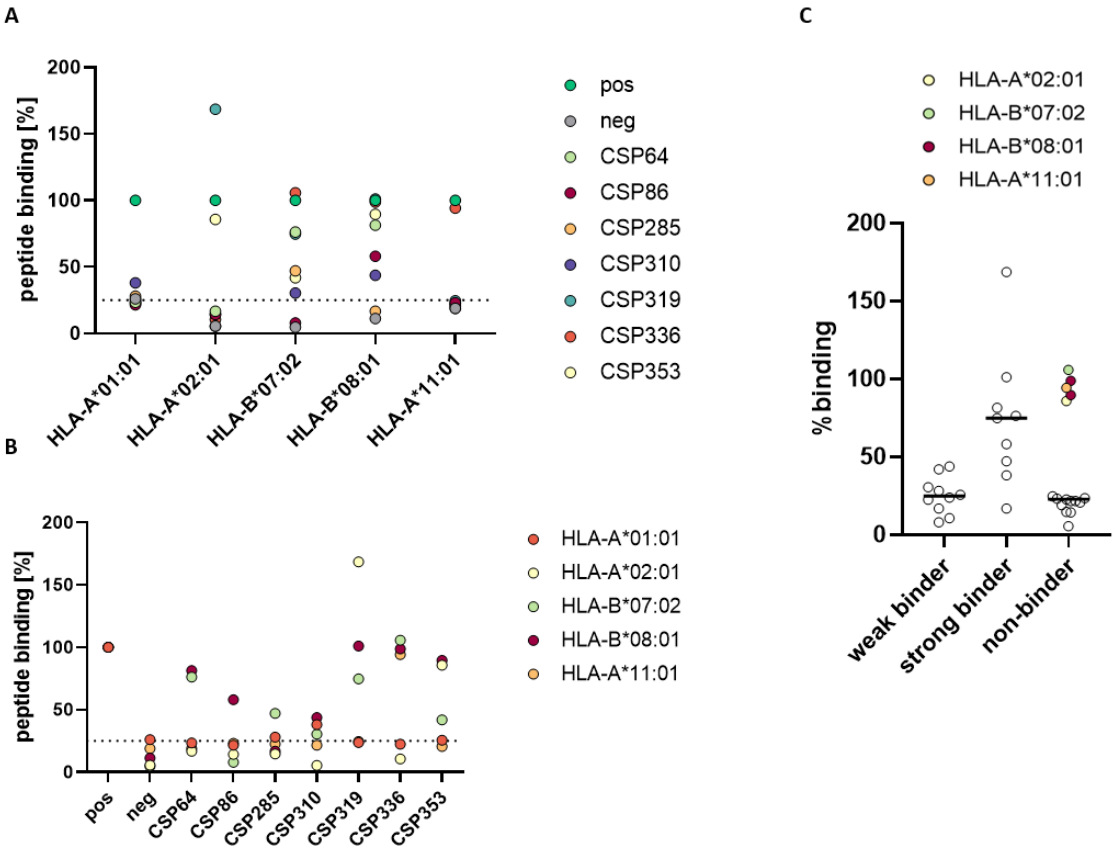


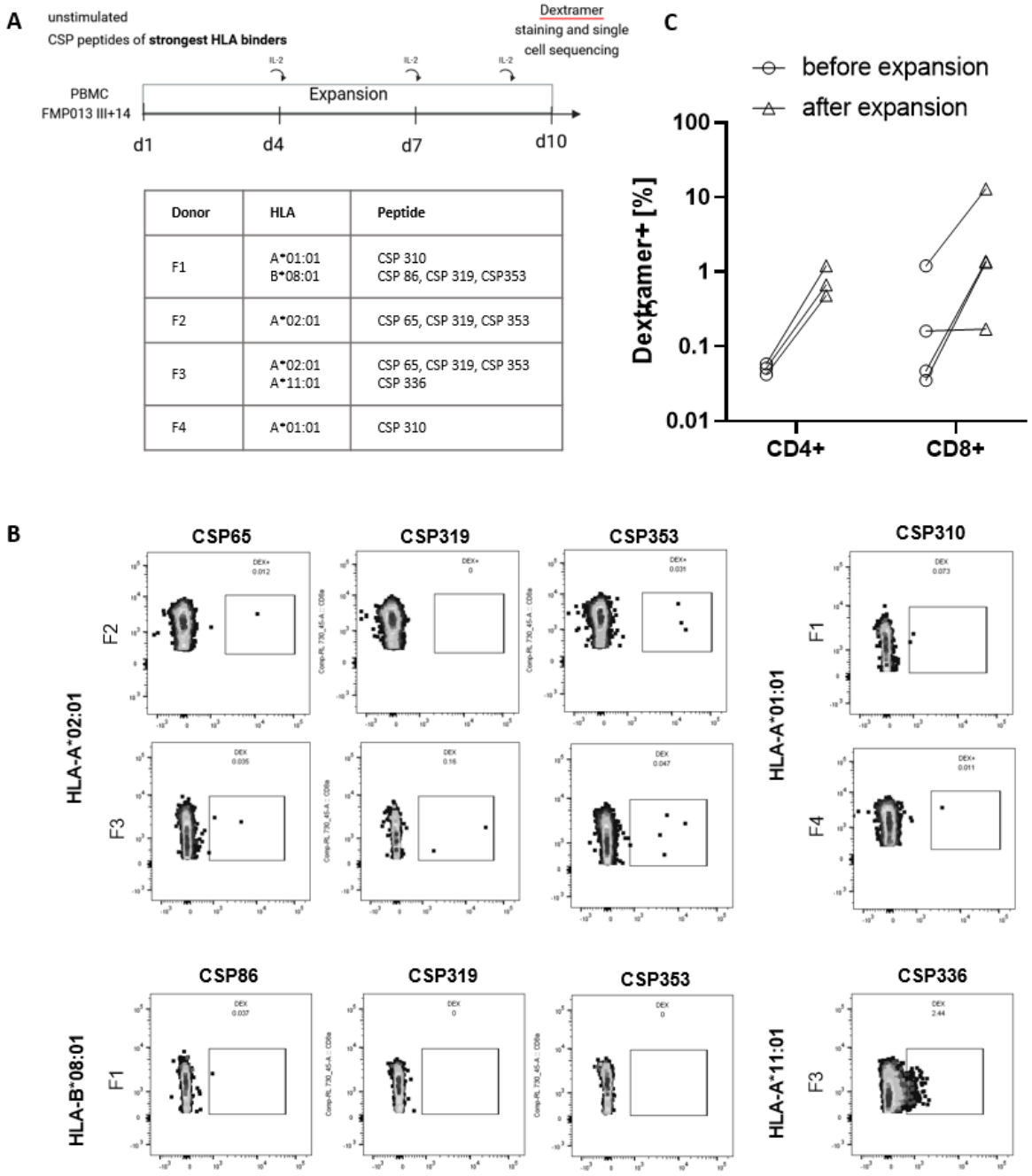
Figure 25: Th2R and Th3R reveal strongest binding across different HLA supertypes.

A-B Peptide binding quantified by *in vitro* binding assay relative to allele specific positive control (light green). Threshold for binding (dotted line) was set by negative control (grey). Percent binding stratified by loaded peptide (A) and HLA type (B) are shown. **C** *In vitro* peptide binding grouped by binding prediction of NetMHCpan4.1.

10.2.17. Strong HLA binders show limited immunogenicity in vaccinated individuals

To further investigate the correlation of *in vitro* pMHC binding and peptide immunogenicity *in vivo* in the context of an anti-CSP T cell response, I performed multimer staining's in HLA-matched vaccinated donors combined with peptide *in vitro* expansion. For four donors, one to four peptides were selected based on matching HLA haplotype and strong peptide binding detected in the *in vitro* binding assay (Figure 26A). To evaluate dextramer-binding T cell frequency and their functional potential in response to peptide expansion, I compared *ex vivo* frequencies before and after 10-day expansion. *Ex vivo* frequencies of CSP-binding CD8 T cells induced by vaccination were assessed in PBMC samples from III+14 time point. At baseline, HLA-B*08:01 dextramer staining loaded with three different CSP peptides did not detect a CD8 population in non-stimulated PBMC samples from the III+14 time point (Figure 26B). HLA-A*02:01 loaded with CSP 353 peptide dextramer staining showed 0.03 and 0.05 %

specific cells among all CD8 cells in two HLA-matched donors. In contrast, the Th2R peptide (CSP 319) loaded HLA-A*02:01 dextramer revealed a positive population in only one of two HLA-matched donors (Figure 26B). The Th3R overlapping peptide CSP 336 presented in the context of an HLA-A*11:01 allele associated with 1.3 % dextramer binding cells among CD8 expressing T cells in one HLA-matched donor. Also, the assessment of dextramer binding among CD4 T cells revealed non-specific background staining of below 0.06 % in all donors and peptides. After 10-day expansion in the presence of the respective stimulating peptide, I could detect an increase in frequency of several pMHC specific T cell populations in several donors (Figure 26C). However, also among CD4 expressing T cells an increase in dextramer positive population was detected. Only the expansion culture stimulated with Th3R peptide revealed strong CD8-specific enrichment of the dextramer positive population. Taken together, I found highest Th3R-HLA-A*11:01 dextramer binding T cell frequencies and strongest CD8-specific expansion upon *in vitro* stimulation. Also, the peptide *in vitro* expansion induced an amplification of both CD4 and CD8 dextramer binding T cells, indicating higher non-specific binding to the dextramer after *in vitro* expansion.



Donor	HLA	Peptide
F1	A*01:01 B*08:01	CSP 310 CSP 86, CSP 319, CSP353
F2	A*02:01	CSP 65, CSP 319, CSP 353
F3	A*02:01 A*11:01	CSP 65, CSP 319, CSP 353 CSP 336
F4	A*01:01	CSP 310

Figure 26: Th3R epitope induces potent CD8 response with high proliferative capacity *in vitro*.

A Experimental setup of 10-day T cell expansion assay. PBMCs isolated 14 days post 3rd vaccination are stimulated with CSP peptides the strongest HLA binders in HLA matched donor. Table specifies which peptides **B** Dextramer staining's loaded with CSP peptides in the context of HLA matched donors (F1, F2, F3, F4) before 10-day expansion. **C** Dextramer frequency among CD4 and CD8 positive T cells before and after 10-day expansion.

11. Discussion

The single-cell based assessment of the T cell response induced by the CSP-based FMP013 vaccine revealed good immunogenicity among CD4 and limited immunogenicity among CD8 T cells. Through the combined approach of *in vitro* peptide stimulation and repertoire analysis, I addressed limitations in quantification of functional and antigen-specific T cell responses. The deep assessment of the TCR properties of CSP-reactive T cells is crucial for the evaluation of potent T cell responses by future vaccine approaches.

11.1. CD4 T cells can target novel N-terminal CSP epitopes

Despite the compelling evidence demonstrating the importance of potent CD4 T cell responses in protection against malaria [95], current CSP-based vaccines fail to induce cross-protective T cell responses due to the strong polymorphism in immunodominant epitopes [58, 96, 97]. This study provides insight into the CD4 T cell response against CSP protein vaccination and identifies novel conserved epitopes by the use of *in vitro* antigen-specific expansion combined with single-cell based TCR sequencing. The combined approach allows enrichment for low frequent CSP-specific T cell clones and determination of their fine epitope specificity through an expression cloning platform. I discovered individual TCRs targeting previously reported immunodominant C-terminal epitopes CS.T3, Th2R and T*. As reported before, despite the overlap between the Th2R and T* epitope regions, identified TCRs in this study targeting one of them did not exhibit cross-reactivity to the other [98]. Historically, the same set of four immunodominant epitopes has been characterized in the context of infectious sporozoites, irradiated sporozoites and subunit vaccines [79, 80, 98-101]. However, the fourth epitope, Th3R, was not identified in this study or in response to irradiated sporozoites [58] or RTS,S subunit vaccine [102]. This contradicts the finding of Good and colleagues, who identified Th3R as most immuno-prevalent epitope [80]. However, in support of an absent CD4 restricted response, more recent studies suggest a stronger MHC class I association of the Th3R epitope [20, 103-105]. In support of previous studies, I find that the C-terminal anti-CSP CD4 T cell response targets a few highly immunogenic epitopes and is dominated by a response targeting the Th2R/T* “supertope” region [80, 98].

In addition to a C-term targeting T cell response, two of four donors revealed an N-term specific response, where T cells target CSP 61-75 and CSP 81-95 in addition to previously reported junctional T1 epitope. Transgenic T cell lines targeting the N-terminal CSP 61 epitope revealed potent activation in response to peptide-pulsed co-culture, suggesting high avidity of CSP61 targeting TCRs. Furthermore, the CSP61 epitope can be targeted by a diverse polyclonal T cell pool, as single TCRs revealed heterogenous V segment usage and CDR3 amino acid sequence. In contrast, the CSP81

epitope was targeted by a less potent T cell response as only a single low reactive TCR was identified in one donor. However, the CSP81 epitope as well as the T1 epitope overlap with the R1 region and N-terminal junction, which is known to be essential in hepatocyte invasion and highly conserved across different species. Several antibodies targeting this region have shown good inhibitory activity during hepatocyte invasion and high potency in protection [43, 106-108], presumably due to the essential function of this region in sporozoite biology. Targeting of this vulnerable site with high sequence conservation supports the high potential as protective T cell epitopes as well.

The identification of CSP-specific CD4 T cell responses in all four donors analyzed in this study underscores CSP as an immunogenic target for CD4 T cells. The identification of novel T cell epitopes within the N-terminus emphasizes the immunogenic potential of a so far mostly neglected region of the protein. Additional studies in larger cohorts are needed to investigate the existence of additional unknown epitopes within this domain and their importance for protection.

11.2. CD4 T cell responses are genetically restricted to MHC class II genes with variable prevalence in humans

A major challenge in inducing potent vaccine-specific responses with broad population coverage are the limitations in HLA presentation and great variability in HLA alleles. Historic data have established the Th2R/T* epitope region as immunodominant CD4 T cell epitope that can be universally presented on diverse HLA-DR molecules [58, 109]. This is supported by the HLA-DR restriction of all TCRs targeting the Th2R epitope identified in this study. In addition to confirmation of previously reported HLA alleles (DRB3, DRB4, DRB7 and DRB10), I identified DRB13 as potent Th2R presenting molecule.

Although the CS.T3 epitope was demonstrated as universal epitope targeted by diverse HLA-DR molecules [79], targeting of the C-terminal epitope located downstream of the TSR region in this study was limited to two donors and linked to non-shared DR4, DR10 and DR14 alleles. However, in support of the promiscuous potential of the CS.T3 epitope, I find targeting of the CS.T3 epitope in the context of HLA-DP presentation by a shared CD4 and CD8 T cell clone.

The ability to target the T1 epitope was restricted to only one donor and a non-shared HLA-DQ allele. This is in support of the previous finding by our lab in the context of irradiated sporozoite vaccination, that the activation of TCRs targeting the T1 epitope was limited to autologous B cells originating from the same donor [58]. The fact that both donors with T1-specific TCRs, F3 in this study and D3 in the irradiated sporozoite vaccination trial, present the epitope on the same DQ allele DQA1*01:02P-DQB1*06:03P, further supports the strong HLA restriction of this epitope as reported by others [110-112]. Furthermore, T1-specific TCRs isolated after irradiated sporozoite vaccination and during this study

showed strong enrichment of TRBV30-1 V segment, suggesting essential involvement in recognition of the T1-HLA-DQ complex.

The targeting of N-terminal epitopes, identified in this study, was restricted to non-shared HLA alleles, which explains the relatively low abundance of TCRs against these epitopes. However, based on allele frequency studies, around 15 % of the population in the United States, where the study has been conducted, carry the CSP61-presenting DR13 allele. Also, within the African population, the prevalence of the DR13 allele ranges from 6 and 25 %, varying across different regions of the continent.

The small number of ten donors included in this study does not reflect the human HLA diversity and aggravates general correlations of MHC genotype with immune response phenotype. Also, the prevalence of HLA haplotypes differs in African donors, and it will be critical to assess HLA-allele frequency linked to CSP epitopes in a larger cohort. Nonetheless, I here demonstrate the potential existence of additional low prevalent and less immunodominant CD4 T cell epitopes that remained undiscovered to date. The development of peptide-based subunit vaccines targeting the pre-erythrocytic stage has been evolved and include immunogenic epitopes of several sporozoite-associated proteins to broaden the response [113]. These approaches aim at the population wide coverage through combination of several conserved epitopes restricted to low prevalent HLA alleles. Since their population coverage and ability to induce strong T cell responses is still limited, they would likely benefit from the inclusion of additional CSP-derived conserved epitopes. Simultaneously, epitope-based vaccines enable the exclusion of immunodominant polymorphic epitopes which can amplify the response to less immunogenic, but conserved epitopes.

11.3. Polymorphic hotspots in T cell epitopes associate with immune pressure by high immunoprevalence

By decades of co-evolution, the *Pf* parasite developed many immune evasion strategies, including intracellular parasitism and antigenic polymorphism, which challenges the induction of potent immune responses [114]. As the antigen presentation process is sensitive to single amino acid changes, point mutations allows the pathogen with an altered epitope to escape detection and elimination by the immune system. Sequence polymorphism in major T cell epitope regions manifests one of the major challenges in vaccine design. Even though the mechanisms behind CSP polymorphisms are poorly understood, polymorphic hotspots, like the Th2R region, are a strong indicator for a role in protective immunity. Also, the absence of functionally significant role in sporozoite biology within the vertebrate host emphasizes the aspects of natural immune selection [115]. The data presented here

support the role of immune pressure by broad immunogenicity in many high frequent HLA alleles. Several studies have identified the same polymorphic regions across geographic regions, but with distinct haplotype distribution, emphasizing the role of local factors in generating polymorphism [85]. These local factors could be linked to differences in HLA haplotype distributions. In addition, differences in susceptibility of anopheline species are proposed as selective force and could relate to differential species distributions as local factor as well [116]. In contrast, the CS.T3 epitope is discussed as universal T cell epitope capable of presentation by many HLA-DR alleles and reveals high sequence conservation across field isolates [79, 84, 117]. However, functional characterization of single CSP-specific TCR isolated from irradiated sporozoite vaccination [58] or recombinant CSP protein vaccination analyzed within this study rather suggests a low frequent T cell response targeting this epitope. The fact that I identified a co-receptor independent TCR against this epitope indicates promiscuous presentation by MHC class I and class II molecules and suggests immunogenic limitations originating from low frequent naïve T cell precursors, which links to the high sequence conservation seen in field isolates. In contrast, the high sequence conservation seen in novel N-terminal epitopes is likely linked to limited promiscuous presentation in the human population. However, the potent T cell response in presence of a CSP61-presenting allele highlights its relevance for vaccine design to induce cross-protective responses [115].

11.4. N-terminal CSP epitopes were previously associated with an MHC class I restricted response

Both N-terminal epitopes targeted by several CD4 associated TCRs identified in this study have been associated with an CD8-restricted response in naturally exposed individuals using ELISpot assay [118]. Historically, most CSP T cell epitopes have been identified by chromium release and ELISpot assay [59] and detected levels of CSP-specific CD8 responses are very low [105]. ELISpot has shown good sensitivity in assessing CD8 T lymphocyte responses to viral infections in humans and allows the detection of IFN-γ production at single cell level and enumeration of antigen-specific cells [119-121]. But, antigen reactivity cannot be linked to cell phenotype nor TCR sequence. Various studies link antigen reactivity to MHC class I restriction on the basis of stimulating peptide length, CD4 T cell depletion or CD8 T cell enrichment [20, 118]. A substantial IFN-γ response in ELISpot assay independent of CD8 T cells has been reported by others and can falsely elevate the enumeration of antigen-specific cells or even lead to a wrong HLA linkage [122]. The detected IFN-γ production after stimulation with CSP 61 and CSP 81 peptides in ELISpot assay by Gonzalez and colleagues may arise from 80% purity of CD8 enriched culture, wherein 20% CD4 T cells could contribute to the detected signal [118]. Also, both peptides associated with undetectable *in vitro* binding to HLA-A donor-matched complexes, further supporting

a potential MHC class II linkage of both epitopes. Through TCR expression cloning and locus-specific TCR-pMHC blocking, the data presented in this study reveal exclusive HLA-DR restriction of all TCRs targeting N-terminal epitope. Whether additional published class I restricted epitopes might reveal HLA class II linkage needs to be elucidated by *in vitro* stimulation assays with a flow cytometric readout that integrates the cell phenotype of the responding cell.

11.5. Limited immunogenicity of CSP vaccination in CD8 T cell response

While we and others have shown good immunogenicity of CSP for CD4 T cells, the induction of a potent anti-CSP response by CD8 T cells remains a challenge [123]. Through flow cytometric analysis of peripheral blood samples, I detect constant frequencies of activated CD8 effector memory cells over subsequent CSP vaccinations. In contrast to the assumption of vaccine-induced activation and proliferation, the activated CD8 effector memory pool is unaffected by subsequent vaccinations. Among CD4 T cells, in particular Tfh cells, the detection of vaccine-induced upregulation of activation markers could be detected during irradiated sporozoite vaccination and SARS-CoV-2 vaccination [124, 125]. The overall lower expression of PD-1 and ICOS on CD8 T cells compared to CD4 T cells aggravates the detection of vaccination-induced changes above the variability of flow cytometric analysis across independent analysis days [126]. Longitudinal phenotypic analysis of PD-1+ CD8 memory T cells during SARS-CoV-2 vaccination also implied overall constant activation level that only increased after the third vaccine dose [127]. In addition, the activation phenotype of the CD8 memory compartment is highly affected by antigen exposure to chronic infection (EBV, CMV) and acute infection (influenza), which have not been monitored throughout the FMP013 vaccination trial. Also, differences in sampling time after last dose will affect the detected level of activation induced by the vaccine due to timely limited expression of surface activation markers. Several activation markers including PD-1 show strongest antigen-induced upregulation seven days after vaccination, while analyzed blood samples in this study were taken 14 or 28 days after the preceding vaccine dose [128].

Subsequent longitudinal repertoire analysis of the activated memory compartment further revealed a high degree of clonal expansion with single clones strongly dominating the repertoire. In comparison to the pre-vaccination timepoint, single largely expanded clones were maintained, although their frequency decreased upon successive immunizations, implying that they do not contribute to a CSP-specific immune response. The stable composition of the CD8 TCR repertoire over months and broad distribution of clone sizes have been overserved, both in healthy individuals [129] and during SARS-CoV-2 vaccination [127] or irradiated sporozoite immunizations [124]. In support of involvement to previous infection, many TCR beta chains of pre-existing clones exhibit reported reactivities against common viruses, primarily EBV and CMV. The detection of expanded clones involved in preceding infections is

a general confounding factor in human repertoire studies over subsequent vaccinations [66], specifically among CD8 T cells with a more constant memory repertoire composition [127]. The phenomenon of increasingly dominating CD8 T cell clones, called memory inflation, has been attributed to chronic antigen exposure in CMV seropositive individuals [91, 130]. As reported for influenza virus vaccination, CMV seropositivity and associated memory inflation can have a negative effect on the generation of a vaccine-specific response [131, 132]. The low number of analyzed individuals and absence of clinical information on CMV seropositivity for the donors enrolled in this study prevent a correlation analysis with vaccine efficacy.

In contrast to large pre-existing clones, subsequent vaccinations led to the recruitment of small clones diversifying the repertoire. As the malaria naïve individuals enrolled in this study have not been exposed to malaria infection an expansion of pre-existing immunity is highly unlikely. Moreover, vaccine studies for SARS-CoV-2 virus in individuals without pre-existing immunity have also shown that antigen-specific CD8 T cells evolve at very low frequency increasing with repeated immunizations [125]. Among small clones recruited after vaccinations, I found donor-specific enrichment of TCR sequence features. The detection of TCR convergence in the activated repertoire implies antigen-specificity of the recruited clones. In line with diverse HLA background of individuals donors and importance of V gene segment for TCR interaction with the HLA molecule, I found donor-specific unique TRBV segments and TRBV-TRAV pairings with increased usage upon vaccination. The more constant TRAV segment usage over time of vaccination potentially originates from lower contribution of the alpha chain to pMHC recognition, resulting in weaker convergent selection of alpha chain V segments [133]. The very low abundance of public TCRs found in analyzed individuals is supported by their diverse HLA background and associated low repertoire overlap in unrelated individuals [24]. Despite common exposure to subsequent vaccinations, the overall high diversity and privacy of the activated CD8 T cell repertoire across donors can be linked to diverse antigen-exposure before and throughout vaccinations as well as their HLA background.

Despite identifying alterations in the repertoire induced by the vaccine, the activated CD8 effector T cell population does not exhibit a potent CSP-specific CD8 T cell response. The lack of CSP-specificity among examined TCR expressing cell lines selected among clones with shared sequence features could have various explanations. Through limited sampling depth of TCR sequences from individual time points and donors, infrequent clones remain undetected and overlap to pre-vaccination might be underestimated. Various studies have shown the ability to identify vaccine-specific T cell responses by repertoire analysis without antigen-specific *in vitro* enrichment, however, either focusing on CD4 T cell responses, tissue-resident CD8 T cell responses or CD8 T cell responses to highly T cell immunogenic yellow fever vaccination [58, 90, 134], which are all associated with a high frequency of

antigen-specific cells. The lower immunogenicity of the CSP protein, originating from very low naïve precursor frequency aggravates their detection through repertoire analysis without additional amplification [135].

In contrast, a lack of CSP-specificity detection could originate from low TCR affinity associated with inefficient T cell activation and low cytokine release during co-culture assay. Nevertheless, the assay used in this study was optimized for sensitive detection of low affinity TCRs through extended co-culture time, optimized cell counts, effector:target cell ratio, high peptide concentrations and IL-2 quantification, that accumulates in the supernatant over time [71]. In contrast to commonly used platforms for functional screening of tumor-infiltrating T cell, which use lower number of T cells and flow cytometric activation readout [136, 137], I detected higher sensitivity through increased T cell numbers and IL-2 quantification by ELISA after extended culture time. The detection of two reactive clones from donor F7 independent of CSP peptide stimulation supports ability for sensitive TCR reactivity detection in autologous B cell stimulation assay. The shared TRAV12-2+TRBV7-9 V segment pairing of both TCRs suggests specificity against a shared EBV-derived epitope, as EBV containing supernatant was used for B cell immortalization. Moreover, the identification of EBV-reactive clones by *in vitro* stimulation in addition to clones identified by sequence similarity underlines the significant fraction of CSP non-specific bystander activation and suggests the presence of other common virus-associated epitope specificities among tested T cell lines. To further entangle whether clones with shared sequence features actively contribute to a vaccine-specific response, an in-depth repertoire analysis and identification of their antigen specificity is essential. These orphan TCRs can be deorphanized by multimer screenings of common immunodominant viral epitopes and their specificity linked back to shared sequence features and activation phenotype over subsequent vaccinations.

11.6. Antigen-independent activation of CD8 T cells during *in vitro* stimulation

I detect evidence for a high fraction of bystander activation among peptide-expanded CD8 T cells, based on strong clonal overlap across N-term and C-term peptide-specific *in vitro* expansion. While *in vitro* stimulation assays have been successfully used to detect a vaccine-specific T cell responses against SARS-CoV-2 or Influenza vaccination [65, 138], functional validation and quantification of non-specific bystander activation showed variable outcomes with respect to the ratio of peptide-specific and non-specific activation [139]. The high correlation of the activation status and clonal expansion among CD4 and CD8 T cells and lower activation phenotype in unstimulated control culture suggests a CD4 T cell-dependent activation of CD8 T cells. A cytokine-dependent mechanism behind bystander activation has already been proposed by Martin et al., who quantified CD8 bystander activation in

PBMC stimulation assay with an intracellular cytokine readout. The authors detected a T cell response comprised of both antigen-specific and bystander component, and improved the assay accuracy through addition of cytokine-specific antibodies [140]. On the other hand, cytokine blocking and the absence of costimulatory signals limits potent antigen-specific activation and proliferation and decreased assay sensitivity [139]. The detection of CD8 bystander activation induced by cytokines produced by antigen-specific CD4 T cells highlights the importance of functional validation of individual T cell clones responding to peptide stimulation *in vitro* and suggests the overestimation of antigen-specific T cell frequencies solely based on activation phenotype in bulk stimulations.

Moreover, the antigen-independent activation of T cells can still play an important immunopathological role *in vivo*. In mice and humans, a cytokine response by antigen-inexperienced CD8 T cells in the absence of cognate T cell receptor stimulation has been observed [141]. Especially during bacterial infection, the contribution of bystander T cell activation plays an important immunopathological role [142, 143]. These innate like responses by CD8 T cells can limit pathogen growth through the production of effector molecules like IFN-γ, which could contribute to parasite elimination in the liver during malaria infection. On the basis of non CSP-specific T cell recruitment upon vaccination and high frequency of bystander-activated CD8 T cells during PBMC stimulation, a cytokine-induced bystander CD8 T cell response to CSP vaccinations can be proposed. However, whether bystander-activated CD8 T cells can reach the liver or can be induced during infection in the liver and can play a protective role in liver stage-specific immunity through production of IFN-γ remains to be elucidated. A high production of IFN-γ by bystander-activated CD8 T cells irrespective of TCR antigen specificity could compensate for the sparsity of CSP-specific cells induced by vaccination.

11.7. CD4-CD8 shared clone can target CSP in a co-receptor independent manner

In addition to the identification of CD4 T cell associated TCRs targeting the CS.T3 epitope, I find recognition of the C-terminal epitope by a shared CD4 and CD8 clonotype. Shared clonotypes across the CD4 and CD8 compartments are very rare and have been associated with dual recognition of a single TCR of both MHC class I and MHC class II complexes [144]. Dual HLA specificity is essential for the emergence of identical clones with CD4 or CD8 phenotype during thymic selection and lineage commitment. In line with variable affinity in TCR recognition of both HLA complexes reported by Heemskerk et al., the T cell clone B11 reveals strong activation in response to HLA-DP specific stimulation. The identification of a CSP-reactive T cell clone with shared CD4 and CD8 phenotype suggests that these can be amplified by CSP vaccination.

Alongside the HLA-DP-restricted CS.T3 presentation, I detect evidence of peptide presentation linked to HLA-E molecule through peptide binding prediction, *in silico* modelling and *in vitro* binding to HLA-E dextramer. The virtual genetic monomorphism of the HLA-E locus with only two alleles with a single amino acid difference emphasizes the potential of HLA-E targeting T cell responses in vaccine design [145]. HLA-E restricted CD8 T cell epitopes have been identified in response to SARS-CoV-2 vaccines [146], their significance in controlling EBV infection [147] or HIV infection [148] described, yet their involvement in *Pf*-specific T cell responses has not been elucidated. The CSP-derived KMEKCSSVF peptide reveals promiscuous binding potential to both alleles through preferred binding motifs of methionine and phenylalanine at amino acid positions P2 and P9, respectively [149]. Despite the high pMHC binding potential, the evidence presented in this study of recognition and activation by TCR B11 mediated by interaction with the HLA-E presented peptide is limited. The detection of HLA-E and HLA-DP restricted activation independently is constrained by autologous B cell stimulation assay, which could be circumvented by the use of a multimer-based stimulation system solely restricted to one allele and peptide [150]. Also, the lack of HLA-E typing data of individuals enrolled in this study further aggravates the linkage to one of the two HLA-E alleles. Through the multimer-based platform both existing HLA-E alleles could be compared in their ability to bind the peptide and induce T cell activation, separately.

Evidence for co-receptor independent T cell activation, similar to the observed absence of a co-receptor enhancing effect in TCR B11 expressing cell line, has been reported before. The essence of co-receptor interaction for potent TCR triggering is under current debate [151]. While several groups have observed TCR triggering even if co-receptors are blocked [152], the recruitment of p56 to phosphorylated TCRs by the co-receptors amplifies TCR signals. Laugel et al. demonstrates that CD8 co-receptor dependence is inversely related with TCR-pMHC affinity through impaired CD8 binding of tested TCR-pMHC complexes [153]. The authors show that TCR activation is independent of any aggregation-based triggering mechanism and that the difference in signal amplification through TCR affinity can compensate for co-receptor enhancing effects. The ability of the co-receptor-independent TCR to activate CD4 and CD8 T cells could potentially enhance cellular responses through the simultaneous activation of functionally diverse T cell subsets [154].

11.8. Limited association of HLA peptide binding and peptide immunogenicity

Understanding the correlations of HLA binding and peptide immunogenicity is highly important for targeted vaccine and immunotherapy design and several algorithms aim at resolving the complexity of peptide immunogenicity [23]. While several CSP peptides are predicted to bind a variety of HLA alleles covering several HLA supertypes [155], *in vitro* MHC binding assays revealed limited binding potential of CSP peptides. The good correlation of prediction as strong binder and *in vitro* binding, but weak

correlation among predicted weak binders, have been reported by others and relate to limited predictive performance of current MHC binding algorithms [22, 23]. The identification of strong *in vitro* binders among non-predicted binders underlines the limitations in the use of *in silico* predictions for immunogenicity predictions. In support of the reported high immunogenic potential of Th2R and Th3R epitope region, I detect broad binding capacity of both peptides across diverse MHC class I alleles.

The data presented in this study revealed limited immunogenicity of strong HLA binders in vaccinated individuals, in contrast to reported strong correlation in the context of SARS-CoV-2 infection and yellow fever vaccination [156, 157]. The combined approach of dextramer staining and *in vitro* expansion used in this study allows the quantification of *ex vivo* epitope-specific T cell frequencies and capacity of antigen binding T cells to proliferate in response to peptide stimulation. In support of the high CD8 T cell immunogenicity of the Th3R region, I found CD8 T cell specific expansion in response to stimulation with the CSP peptide overlapping with the Th3R region. However, the high sequence polymorphism within this region restricts the ability to induce a long-term CD8 memory response by vaccination that can recognize and rapidly respond upon re-infection with a circulating strain [97].

11.9. High requirements for liver stage-specific T cell responses

Both, the tissue residence of a liver stage-specific T cell response and huge variety of potential immunogenic sporozoite-associated proteins complicate the detection of circulating CSP-specific CD8 T cells during malaria infection or sporozoite vaccination [158, 159]. Despite the circumvented liver stage by recombinant protein vaccination and focusing the immune response to a single protein, subsequent full-length recombinant CSP vaccinations fail to induce highly frequent CSP-specific CD8 T cells in circulation. A full-length CSP vaccination trail including the immunogenic C-terminus containing major T cell epitopes as well as the less immunogenic N-terminus offers an optimal setting to study the T cell response against CSP in detail. The absence of a detected MHC class I restricted T cell response both through repertoire analysis and *in vitro* expansion supports the limited immunogenic potential of CSP to induce CD8 T cell responses reported previously [109, 160]. In contrast to a protective anti-CSP response induced by irradiated and live sporozoites, subunit-based vaccines are inefficient in reaching the same protective correlation. Live replicating pathogens often mount robust CD8 T cell responses, but sterile and subunit immunizations have so far been unsuccessful in inducing optimal CD8 T cell responses. Successful vaccines as the yellow fever and small pox vaccine based on live replicating pathogens, which elicited strong CD8 immunity in addition to humoral responses associated with long-term protection underline the protective potential of vaccine-induced CD8 T cell responses [161]. One crucial bottleneck in effective T cell immunity induced by subunit vaccines originates from limitations in antigen presentation. External antigen that is taken up by professional

APCs needs to be cross-presented on MHC class I molecules for CD8 T cell priming. To modulate cross-presentation efficacy, adjuvants can engage in the cross-presentation pathway [162]. The ALFQ adjuvant utilized in the FMP013 study is suboptimal at inducing effective MHC class I restricted antigen presentation [77]. The liposome-based adjuvant only contains saturated lipids, which, in contrast to unsaturated lipids, only induces significant antigen-specific CD4 T cell responses and often lack CD8 T cell response [163]. Also, during RTS,S vaccination in liposome-based adjuvant AS01, a limited potential to induce antigen-specific CD8 T cell responses has been observed, whereas the adjuvant could be attributed to a strong production of Th1 type CD4 immune response [164].

Our understanding of the protective requirements for vaccine-elicited CD8 T cell responses is still very limited. For malaria, where solely antibody-focused vaccines fail to induce long-term protection, a multi-stage vaccine that induces potent T cell responses as well can enhance protection [165]. In addition to their high potential in targeting infectious disease stages like the liver stage, where neutralizing antibodies cannot access, T cells can recognize a broader antigenic repertoire and are not limited to outer surface proteins [166]. Irradiated sporozoite vaccines induce sterilizing protection against malaria in both humans and rodents, but in mice the protective immunity is linked to extremely high frequency of memory CD8 T cells specific for sporozoite-derived proteins [51]. In order to generate higher frequency of vaccine specific memory CD8 T cells, DNA and mRNA-based vaccines that induce strong MHC class I antigen presentation associated with effectiveness in inducing potent CD8 CTL responses could represent a promising alternative vaccination approach [167].

Another requirement for liver stage-specific T cell immunity confers efficient presentation of *Pf* antigens on infected hepatocytes. In mice, CSP-derived peptides from *P. yoelii* and *P. berghei* are expressed on the surface of infected hepatocytes and potent cytotoxic T cells can recognize presented sporozoite antigens [168-171]. In humans the level of presented CSP peptides by infected hepatocytes and time of expression during liver stage are still unknown and resemble a major limitation in understanding the generation of live stage specific protective T cell response. Unlike interference with antigen presentation processes, which is a favored immune evasion strategy, MHC presentation is unaffected by sporozoite infection. This further promotes the protective potential of liver stage-specific T cell immunity [172].

More recently, CD4 T cells were also found to recognize parasite-derived peptides on infected hepatocytes in association with class II MHC molecules and eliminate *P.yoelii*-infected hepatocytes [173, 174]. Through their reported cytolytic activity, CD4 T cells could participate in effector mechanisms against malaria liver stages and could compensate for limited CD8 T cell responses [98, 175]. Also, in

malaria endemic regions clonally expanded CSP-specific cytolytic memory CD4 Th1 cells associate with clinical immunity, supporting their protective potential [176].

12. Outlook

The identification of N-terminal CD4 epitopes that remained undiscovered to date imply the presence of additional low immunogenic epitopes. To screen for additional CSP epitopes, PBMC samples from a larger cohort are required and low frequent T cells need to be amplified by *in vitro* expansion. A linkage to responding cell phenotype is hereby required to allow the identification of presenting HLA context. Also, to assess the immunoprevalence of identified epitopes, HLA binding and allele frequency in different geographic regions should be monitored and quantified. Initially CSP peptides can be screened for *in vitro* binding to diverse HLA monomers to narrow down the number of HLA alleles with potential to elicit a T cell response. In a second step, the frequency of epitope-specific T cells and functional capacity should be assessed by multimer staining and *in vitro* stimulation in HLA-matched donors that received a CSP-based vaccination or have been naturally exposed to *Pf* infection. The comparison of vaccine-induced T cell response and natural exposure is crucial to evaluate the protective potential of vaccine-elicited responses.

The detection of a highly expanded effector memory CD8 repertoire and emerging clones with shared sequence features upon vaccination without CSP-specificity raises new questions. What is the role of bystander-activated CD8 T cells induced by vaccination? Can they contribute to liver-stage specific T cell response upon infection? How can CSP immunogenicity among CD8 T cells be increased? These questions could be addressed by an in-depth characterization of bystander-activated clones and further *in vitro* hepatocyte killing experiments. Through multimer stainings against common viral epitopes (CMV, EBV, Influenza) and tracking the activation phenotype over time of vaccination, the magnitude of bystander and antigen-specific CD4 and CD8 T cell responses can be quantified. To evaluate their pathogenic role, the correlation with protection and/or clinical symptoms should be addressed. In addition, in-depth repertoire analysis by bulk sequencing might assist the more sensitive detection of expanding and contracting clones and more accurate quantification of bystander-activation induced expansion. During *in vitro Pf*-infected hepatocyte killing experiments, the production and effect of IFN-γ produced by bystander-activated CD8 T cells with non-*Pf*-specificity can be monitored and can give insight on the immunopathological role of bystander-activated cells.

The data presented in this study support the essence of functional validation of *in vitro* assays and multimer stainings [177]. By *TR* gene sequencing of peptide-expanded and dextramer-stained Th3R binding cells, I will assess the clonal composition of dextramer binding and activated CD8 T cells before and after peptide-specific expansion. Through *TR* gene cloning and expression, I will subsequently characterize the functional capacities of a Th3R targeting CD8 T cell response.

13. References

1. Lugo, J.P., et al., *Early precursor thymocytes can produce interleukin 2 upon stimulation with calcium ionophore and phorbol ester.* Proc Natl Acad Sci U S A, 1986. **83**(6): p. 1862-6.
2. Born, W., et al., *Rearrangement of T-cell receptor beta-chain genes during T-cell development.* Proc Natl Acad Sci U S A, 1985. **82**(9): p. 2925-9.
3. Alam, S.M., et al., *Qualitative and quantitative differences in T cell receptor binding of agonist and antagonist ligands.* Immunity, 1999. **10**(2): p. 227-37.
4. Alam, S.M., et al., *T-cell-receptor affinity and thymocyte positive selection.* Nature, 1996. **381**(6583): p. 616-20.
5. Palmer, E., *Negative selection — clearing out the bad apples from the T-cell repertoire.* Nature Reviews Immunology, 2003. **3**(5): p. 383-391.
6. Germain, R.N., *T-cell development and the CD4-CD8 lineage decision.* Nature Reviews. Immunology, 2002. **2**(5): p. 309-322.
7. Starr, T.K., S.C. Jameson, and K.A. Hogquist, *Positive and negative selection of T cells.* Annual Review of Immunology, 2003. **21**: p. 139-176.
8. Williams, W.V., et al., *The antigen-major histocompatibility complex-T cell receptor interaction. A structural analysis.* Immunol Res, 1988. **7**(4): p. 339-50.
9. Shah, K., et al., *T cell receptor (TCR) signaling in health and disease.* Signal Transduct Target Ther, 2021. **6**(1): p. 412.
10. Martin, M.D. and V.P. Badovinac, *Defining Memory CD8 T Cell.* Frontiers in Immunology, 2018. **9**: p. 2692.
11. De Simone, M., G. Rossetti, and M. Pagani, *Single Cell T Cell Receptor Sequencing: Techniques and Future Challenges.* Frontiers in Immunology, 2018. **9**: p. 1638.
12. Falk, K., et al., *Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules.* Nature, 1991. **351**(6324): p. 290-296.
13. Kloetzel, P.-M. and F. Ossendorp, *Proteasome and peptidase function in MHC-class-I-mediated antigen presentation.* Current Opinion in Immunology, 2004. **16**(1): p. 76-81.
14. Zhang, C., A. Anderson, and C. DeLisi, *Structural principles that govern the peptide-binding motifs of class I MHC molecules.* J Mol Biol, 1998. **281**(5): p. 929-47.
15. Radwan, J., et al., *Advances in the Evolutionary Understanding of MHC Polymorphism.* Trends Genet, 2020. **36**(4): p. 298-311.
16. Sanchez-Mazas, A., *A review of HLA allele and SNP associations with highly prevalent infectious diseases in human populations.* Swiss Med Wkly, 2020. **150**: p. w20214.
17. Glimcher, L.H. and C.J. Kara, *Sequences and factors: a guide to MHC class-II transcription.* Annu Rev Immunol, 1992. **10**: p. 13-49.
18. Guardiola, J. and A. Maffei, *Control of MHC class II gene expression in autoimmune, infectious, and neoplastic diseases.* Crit Rev Immunol, 1993. **13**(3-4): p. 247-68.
19. Bendukidze, N., et al., *Identification of HLA alleles with low or no cell surface expression in the Czech population.* Folia Biologica, 2003. **49**(6): p. 227-229.
20. Doolan, D.L., et al., *Degenerate cytotoxic T cell epitopes from P. falciparum restricted by multiple HLA-A and HLA-B supertype alleles.* Immunity, 1997. **7**(1): p. 97-112.

21. Tarke, A., et al., *Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases.* Cell Rep Med, 2021. **2**(2): p. 100204.
22. Bonsack, M., et al., *Performance Evaluation of MHC Class-I Binding Prediction Tools Based on an Experimentally Validated MHC–Peptide Binding Data Set.* Cancer Immunology Research, 2019. **7**(5): p. 719-736.
23. Reynisson, B., et al., *NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data.* Nucleic Acids Research, 2020. **48**(W1): p. W449-W454.
24. Venturi, V., et al., *The molecular basis for public T-cell responses?* Nat Rev Immunol, 2008. **8**(3): p. 231-8.
25. Russell, M.L., et al., *Combining genotypes and T cell receptor distributions to infer genetic loci determining V(D)J recombination probabilities*. 2021, Immunology.
26. de Greef, P.C., et al., *The naive T-cell receptor repertoire has an extremely broad distribution of clone sizes.* eLife, 2020. **9**: p. e49900.
27. Ruiz Ortega, M., et al., *Modeling and predicting the overlap of B- and T-cell receptor repertoires in healthy and SARS-CoV-2 infected individuals.* PLoS genetics, 2023. **19**(2): p. e1010652.
28. Davis, M.M. and S.D. Boyd, *Recent progress in the analysis of αβT cell and B cell receptor repertoires.* Current Opinion in Immunology, 2019. **59**: p. 109-114.
29. Fischer, D.S., et al., *Predicting antigen specificity of single T cells based on TCR CDR 3 regions.* Molecular Systems Biology, 2020. **16**(8).
30. Chiffelle, J., et al., *T-cell repertoire analysis and metrics of diversity and clonality.* Curr Opin Biotechnol, 2020. **65**: p. 284-295.
31. Miho, E., et al., *Computational Strategies for Dissecting the High-Dimensional Complexity of Adaptive Immune Repertoires.* Frontiers in Immunology, 2018. **9**: p. 224.
32. Venkatesan, P., *The 2023 WHO World malaria report.* The Lancet. Microbe, 2024: p. S2666-5247(24)00016-8.
33. Weiss, D.J., et al., *Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000-17: a spatial and temporal modelling study.* Lancet (London, England), 2019. **394**(10195): p. 322-331.
34. Silvie, O., et al., *Interactions of the malaria parasite and its mammalian host.* Current Opinion in Microbiology, 2008. **11**(4): p. 352-359.
35. Cao, P., et al., *Modeling the dynamics of Plasmodium falciparum gametocytes in humans during malaria infection.* Elife, 2019. **8**.
36. Farfour, E., et al., *The extravascular compartment of the bone marrow: a niche for Plasmodium falciparum gametocyte maturation?* Malaria Journal, 2012. **11**: p. 285.
37. Tibúrcio, M., et al., *Early gametocytes of the malaria parasite Plasmodium falciparum specifically remodel the adhesive properties of infected erythrocyte surface.* Cellular Microbiology, 2013. **15**(4): p. 647-659.
38. Angrisano, F., et al., *Malaria parasite colonisation of the mosquito midgut--placing the Plasmodium ookinete centre stage.* International Journal for Parasitology, 2012. **42**(6): p. 519-527.
39. Miller, L.H., et al., *The pathogenic basis of malaria.* Nature, 2002. **415**(6872): p. 673-679.
40. English, M., et al., *Acidosis in severe childhood malaria.* QJM: monthly journal of the Association of Physicians, 1997. **90**(4): p. 263-270.

41. Michalakis, Y. and F. Renaud, *Malaria: Evolution in vector control.* Nature, 2009. **462**(7271): p. 298-300.
42. Cowman, A.F., et al., *Malaria: Biology and Disease.* Cell, 2016. **167**(3): p. 610-624.
43. Murugan, R., et al., *Evolution of protective human antibodies against Plasmodium falciparum circumsporozoite protein repeat motifs.* Nat Med, 2020. **26**(7): p. 1135-1145.
44. Flores-Garcia, Y., et al., *Antibody-Mediated Protection against Plasmodium Sporozoites Begins at the Dermal Inoculation Site.* mBio, 2018. **9**(6).
45. Weiss, W.R., et al., *The role of CD4+ T cells in immunity to malaria sporozoites.* J Immunol, 1993. **151**(5): p. 2690-8.
46. Hassert, M., S. Arumugam, and J.T. Harty, *Memory CD8+ T cell-mediated protection against liver-stage malaria.* Immunol Rev, 2023. **316**(1): p. 84-103.
47. Mellouk, S., et al., *Inhibitory activity of interferons and interleukin 1 on the development of Plasmodium falciparum in human hepatocyte cultures.* J Immunol, 1987. **139**(12): p. 4192-5.
48. Seguin, M.C., et al., *Induction of nitric oxide synthase protects against malaria in mice exposed to irradiated Plasmodium berghei infected mosquitoes: involvement of interferon gamma and CD8+ T cells.* J Exp Med, 1994. **180**(1): p. 353-8.
49. Druilhe, P. and J.W. Barnwell, *Pre-erythrocytic stage malaria vaccines: time for a change in path.* Curr Opin Microbiol, 2007. **10**(4): p. 371-8.
50. Nussenzweig, R.S., et al., *Protective Immunity produced by the Injection of X-irradiated Sporozoites of Plasmodium berghei.* Nature, 1967. **216**(5111): p. 160-162.
51. Schmidt, N.W., et al., *Extreme CD8 T cell requirements for anti-malarial liver-stage immunity following immunization with radiation attenuated sporozoites.* PLoS pathogens, 2010. **6**(7): p. e1000998.
52. Weiss, W.R. and C.G. Jiang, *Protective CD8+ T lymphocytes in primates immunized with malaria sporozoites.* PloS One, 2012. **7**(2): p. e31247.
53. Van Braeckel-Budimir, N. and J.T. Harty, *CD8 T-cell-mediated protection against liver-stage malaria: lessons from a mouse model.* Frontiers in Microbiology, 2014. **5**.
54. Epstein, J.E., et al., *Live attenuated malaria vaccine designed to protect through hepatic CD8$^+$ T cell immunity.* Science (New York, N.Y.), 2011. **334**(6055): p. 475-480.
55. Schofield, L., et al., *γ Interferon, CD8+ T cells and antibodies required for immunity to malaria sporozoites.* Nature, 1987. **330**(6149): p. 664-666.
56. Weiss, W.R., et al., *CD8+ T cells (cytotoxic/suppressors) are required for protection in mice immunized with malaria sporozoites.* Proceedings of the National Academy of Sciences, 1988. **85**(2): p. 573-576.
57. Rodrigues, M.M., et al., *CD8 $^+$ cytolytic T cell clones derived against the Plasmodium yoelii circumsporozoite protein protect against malaria.* International Immunology, 1991. **3**(6): p. 579-585.
58. Wahl, I., et al., *Clonal evolution and TCR specificity of the human T $_{FH}$ cell response to Plasmodium falciparum CSP.* Science Immunology, 2022. **7**(72): p. eabm9644.
59. Heide, J., et al., *Comprehensive Review of Human Plasmodium falciparum-Specific CD8+ T Cell Epitopes.* Frontiers in Immunology, 2019. **10**: p. 397.
60. Bonelo, A., et al., *Generation and characterization of malaria-specific human CD8(+) lymphocyte clones: effect of natural polymorphism on T cell recognition and endogenous cognate antigen presentationby liver cells.* Eur J Immunol, 2000. **30**(11): p. 3079-88.

61. Laurens, M.B., *RTS,S/AS01 vaccine (Mosquirix™): an overview.* Human Vaccines & Immunotherapeutics, 2020. **16**(3): p. 480-489.
62. Duffy, P.E. and J. Patrick Gorres, *Malaria vaccines since 2000: progress, priorities, products.* npj Vaccines, 2020. **5**(1): p. 48.
63. Ludwig, J., et al., *Glycosylated nanoparticle-based PfCSP vaccine confers long-lasting antibody responses and sterile protection in mouse malaria model.* NPJ Vaccines, 2023. **8**(1): p. 52.
64. Fraser, C.C., et al., *Generation of a universal CD4 memory T cell recall peptide effective in humans, mice and non-human primates.* Vaccine, 2014. **32**(24): p. 2896-903.
65. Gondré-Lewis, T.A., et al., *NIAID workshop on T cell technologies.* Nature Immunology, 2023. **24**(1): p. 14-18.
66. Minervina, A., M. Pogorelyy, and I. Mamedov, *T-cell receptor and B-cell receptor repertoire profiling in adaptive immunity.* Transplant International, 2019. **32**(11): p. 1111-1123.
67. Bacher, P. and A. Scheffold, *Flow-cytometric analysis of rare antigen-specific T cells.* Cytometry. Part A: The Journal of the International Society for Analytical Cytology, 2013. **83**(8): p. 692-701.
68. Dolton, G., et al., *Optimized Peptide-MHC Multimer Protocols for Detection and Isolation of Autoimmune T-Cells.* Frontiers in Immunology, 2018. **9**: p. 1378.
69. Rosati, E., et al., *Overview of methodologies for T-cell receptor repertoire analysis.* BMC Biotechnology, 2017. **17**(1): p. 61.
70. Imkeller, K., et al., *sciReptor: analysis of single-cell level immunoglobulin repertoires.* BMC Bioinformatics, 2016. **17**: p. 67.
71. Wahl, I., et al., *An efficient single-cell based method for linking human T cell phenotype to T cell receptor sequence and specificity.* European Journal of Immunology, 2022. **52**(2): p. 237-246.
72. Masella, A.P., et al., *PANDAseq: paired-end assembler for illumina sequences.* BMC Bioinformatics, 2012. **13**: p. 31.
73. Ye, J., et al., *IgBLAST: an immunoglobulin variable domain sequence analysis tool.* Nucleic Acids Res, 2013. **41**(Web Server issue): p. W34-40.
74. Heemskerk, M.H., et al., *Redirection of antileukemic reactivity of peripheral T lymphocytes using gene transfer of minor histocompatibility antigen HA-2-specific T-cell receptor complexes expressing a conserved alpha joining region.* Blood, 2003. **102**(10): p. 3530-40.
75. Hutter, J.N., et al., *First-in-human assessment of safety and immunogenicity of low and high doses of Plasmodium falciparum malaria protein 013 (FMP013) administered intramuscularly with ALFQ adjuvant in healthy malaria-naïve adults.* Vaccine, 2022. **40**(40): p. 5781-5790.
76. Genito, C.J., et al., *Liposomes containing monophosphoryl lipid A and QS-21 serve as an effective adjuvant for soluble circumsporozoite protein malaria vaccine FMP013.* Vaccine, 2017. **35**(31): p. 3865-3874.
77. Alving, C.R., et al., *Army Liposome Formulation (ALF) family of vaccine adjuvants.* Expert Review of Vaccines, 2020. **19**(3): p. 279-292.
78. Cawlfield, A., et al., *Safety, toxicity and immunogenicity of a malaria vaccine based on the circumsporozoite protein (FMP013) with the adjuvant army liposome formulation containing QS21 (ALFQ).* Vaccine, 2019. **37**(29): p. 3793-3803.

79. Sinigaglia, F., et al., *A malaria T-cell epitope recognized in association with most mouse and human MHC class II molecules.* Nature, 1988. **336**(6201): p. 778-80.
80. Good, M.F., et al., *Human T-cell recognition of the circumsporozoite protein of Plasmodium falciparum: immunodominant T-cell domains map to the polymorphic regions of the molecule.* Proc Natl Acad Sci U S A, 1988. **85**(4): p. 1199-203.
81. Guttinger, M., et al., *Human T cells recognize polymorphic and non-polymorphic regions of the Plasmodium falciparum circumsporozoite protein.* Embo j, 1988. **7**(8): p. 2555-8.
82. Emerson, R., et al., *Estimating the ratio of CD4+ to CD8+ T cells using high-throughput sequence data.* J Immunol Methods, 2013. **391**(1-2): p. 14-21.
83. Kitaura, K., et al., *A new high-throughput sequencing method for determining diversity and similarity of T cell receptor (TCR) α and β repertoires and identifying potential new invariant TCR α chains.* BMC Immunol, 2016. **17**(1): p. 38.
84. Tanabe, K., et al., *Within-population genetic diversity of Plasmodium falciparum vaccine candidate antigens reveals geographic distance from a Central sub-Saharan African origin.* Vaccine, 2013. **31**(9): p. 1334-1339.
85. Jalloh, A., M. Jalloh, and H. Matsuoka, *T-cell epitope polymorphisms of the Plasmodium falciparum circumsporozoite protein among field isolates from Sierra Leone: age-dependent haplotype distribution?* Malaria Journal, 2009. **8**(1): p. 120.
86. Valentine, K.M. and K.K. Hoyer, *CXCR5+ CD8 T Cells: Protective or Pathogenic?* Frontiers in Immunology, 2019. **10**: p. 1322.
87. Wen, T., J. Bukczynski, and T.H. Watts, *4-1BB ligand-mediated costimulation of human T cells induces CD4 and CD8 T cell expansion, cytokine production, and the development of cytolytic effector function.* Journal of Immunology (Baltimore, Md.: 1950), 2002. **168**(10): p. 4897-4906.
88. Park, S.L., S.N. Christo, and L.K. Mackay, *ICOS-play: dressing T cells for residency.* Trends in Immunology, 2022. **43**(4): p. 280-282.
89. Lyke, K.E., et al., *Attenuated PfSPZ Vaccine induces strain-transcending T cells and durable protection against heterologous controlled human malaria infection.* Proc Natl Acad Sci U S A, 2017. **114**(10): p. 2711-2716.
90. Sant, S., et al., *Single-Cell Approach to Influenza-Specific CD8+ T Cell Receptor Repertoires Across Different Age Groups, Tissues, and Following Influenza Virus Infection.* Frontiers in Immunology, 2018. **9**: p. 1453.
91. Kim, J., A.-R. Kim, and E.-C. Shin, *Cytomegalovirus Infection and Memory T Cell Inflation.* Immune Network, 2015. **15**(4): p. 186.
92. Chronister, W.D., et al., *TCRMatch: Predicting T-Cell Receptor Specificity Based on Sequence Similarity to Previously Characterized Receptors.* Frontiers in Immunology, 2021. **12**: p. 640725.
93. Dash, P., et al., *Quantifiable predictive features define epitope-specific T cell receptor repertoires.* Nature, 2017. **547**(7661): p. 89-93.
94. Sidney, J., et al., *HLA class I supertypes: a revised and updated classification.* BMC Immunology, 2008. **9**(1): p. 1.
95. Molano, A., et al., *Cutting edge: the IgG response to the circumsporozoite protein is MHC class II-dependent and CD1d-independent: exploring the role of GPIs in NK T cell activation and antimalarial responses.* J Immunol, 2000. **164**(10): p. 5005-9.
96. de la Cruz, V.F., et al., *Lack of cross-reactivity between variant T cell determinants from malaria circumsporozoite protein.* J Immunol, 1988. **141**(7): p. 2456-60.

97. Zeeshan, M., et al., *Genetic variation in the Plasmodium falciparum circumsporozoite protein in India and its relevance to RTS,S malaria vaccine.* PLoS One, 2012. **7**(8): p. e43430.
98. Moreno, A., et al., *Cytotoxic CD4+ T cells from a sporozoite-immunized volunteer recognize the Plasmodium falciparum CS protein.* Int Immunol, 1991. **3**(10): p. 997-1003.
99. Good, M.F., et al., *Construction of synthetic immunogen: use of new T-helper epitope on malaria circumsporozoite protein.* Science, 1987. **235**(4792): p. 1059-62.
100. Moreno, A., et al., *CD4+ T cell clones obtained from Plasmodium falciparum sporozoite-immunized volunteers recognize polymorphic sequences of the circumsporozoite protein.* J Immunol, 1993. **151**(1): p. 489-99.
101. Sinigaglia, F., et al., *Epitopes recognized by human T lymphocytes on malaria circumsporozoite protein.* Eur J Immunol, 1988. **18**(4): p. 633-6.
102. Schwenk, R., et al., *Immunization with the RTS,S/AS malaria vaccine induces IFN-γ(+)CD4 T cells that recognize only discrete regions of the circumsporozoite protein and these specificities are maintained following booster immunizations and challenge.* Vaccine, 2011. **29**(48): p. 8847-54.
103. Wang, R., et al., *Induction of antigen-specific cytotoxic T lymphocytes in humans by a malaria DNA vaccine.* Science, 1998. **282**(5388): p. 476-80.
104. Wang, R., et al., *Induction in humans of CD8+ and CD4+ T cell and antibody responses by sequential immunization with malaria DNA and recombinant protein.* J Immunol, 2004. **172**(9): p. 5561-9.
105. Dodoo, D., et al., *Measuring naturally acquired immune responses to candidate malaria vaccine antigens in Ghanaian adults.* Malar J, 2011. **10**: p. 168.
106. Kisalu, N.K., et al., *A human monoclonal antibody prevents malaria infection by targeting a new site of vulnerability on the parasite.* Nat Med, 2018. **24**(4): p. 408-416.
107. Murugan, R., et al., *Clonal selection drives protective memory B cell responses in controlled human malaria infection.* Sci Immunol, 2018. **3**(20).
108. Tan, J., et al., *A public antibody lineage that potently inhibits malaria infection through dual binding to the circumsporozoite protein.* Nat Med, 2018. **24**(4): p. 401-407.
109. Lalvani, A., et al., *Potent induction of focused Th1-type cellular and humoral immune responses by RTS,S/SBAS2, a recombinant Plasmodium falciparum malaria vaccine.* J Infect Dis, 1999. **180**(5): p. 1656-64.
110. Calvo-Calle, J.M., et al., *Binding of malaria T cell epitopes to DR and DQ molecules in vitro correlates with immunogenicity in vivo: identification of a universal T cell epitope in the Plasmodium falciparum circumsporozoite protein.* Journal of Immunology (Baltimore, Md.: 1950), 1997. **159**(3): p. 1362-1373.
111. Nardin, E.H., et al., *Conserved repetitive epitope recognized by CD4+ clones from a malaria-immunized volunteer.* Science, 1989. **246**(4937): p. 1603-6.
112. Nardin, E.H., et al., *Synthetic malaria peptide vaccine elicits high levels of antibodies in vaccinees of defined HLA genotypes.* J Infect Dis, 2000. **182**(5): p. 1486-96.
113. Tsuji, M. and F. Zavala, *Peptide-based subunit vaccines against pre-erythrocytic stages of malaria parasites.* Mol Immunol, 2001. **38**(6): p. 433-42.
114. Hisaeda, H., K. Yasutomo, and K. Himeno, *Malaria: immune evasion by parasites.* The International Journal of Biochemistry & Cell Biology, 2005. **37**(4): p. 700-706.
115. Gandhi, K., et al., *Variation in the circumsporozoite protein of Plasmodium falciparum: vaccine development implications.* PloS One, 2014. **9**(7): p. e101783.

116. Kumkhaek, C., et al., *Are extensive T cell epitope polymorphisms in the Plasmodium falciparum circumsporozoite antigen, a leading sporozoite vaccine candidate, selected by immune pressure?* J Immunol, 2005. **175**(6): p. 3935-9.
117. Kumar, P., S. Biswas, and D.N. Rao, *Potentiation of immune response against the RESA peptides of Plasmodium falciparum by incorporating a universal T-cell epitope (CS.T3) and an immunomodulator (polytuftsin), and delivery through liposomes.* Microbiol Immunol, 1999. **43**(6): p. 567-76.
118. González, J.M., et al., *HLA-A*0201 restricted CD8+ T-lymphocyte responses to malaria: identification of new Plasmodium falciparum epitopes by IFN-gamma ELISPOT.* Parasite Immunol, 2000. **22**(10): p. 501-14.
119. Woldemeskel, B.A., C.C. Garliss, and J.N. Blankson, *SARS-CoV-2 mRNA vaccines induce broad CD4+ T cell responses that recognize SARS-CoV-2 variants and HCoV-NL63.* The Journal of clinical investigation, 2021. **131**(10).
120. Sahin, U., et al., *COVID-19 vaccine BNT162b1 elicits human antibody and TH1 T cell responses.* Nature, 2020. **586**(7830): p. 594-599.
121. Freen-van Heeren, J.J., et al., *Assessing Antigen-Specific T Cell Responses Through IFN-γ Enzyme-Linked Immune Absorbent Spot (ELISpot).* Methods Mol Biol, 2024. **2782**: p. 209-226.
122. Fuchs, Y.F., et al., *Vagaries of the ELISpot assay: specific detection of antigen responsive cells requires purified CD8(+) T cells and MHC class I expressing antigen presenting cell lines.* Clin Immunol, 2015. **157**(2): p. 216-25.
123. Ockenhouse, C.F., et al., *Ad35.CS.01-RTS,S/AS01 Heterologous Prime Boost Vaccine Efficacy against Sporozoite Challenge in Healthy Malaria-Naïve Adults.* PLoS One, 2015. **10**(7): p. e0131571.
124. Bartsch, I.M. and H.R. Rodewald, *Dynamics of the Human T Cell Response to Repeated Immunization with Irradiated Malaria Parasites*. 2020: Ruprecht-Karls-Universität Heidelberg.
125. Zhang, Z., et al., *Humoral and cellular immune memory to four COVID-19 vaccines.* Cell, 2022. **185**(14): p. 2434-2451.e17.
126. Bayiyana, A., et al., *Longitudinal Changes in Cd4(+), Cd8(+) T Cell Phenotype and Activation Marker Expression Following Antiretroviral Therapy Initiation among Patients with Cryptococcal Meningitis.* J Fungi (Basel), 2019. **5**(3).
127. Galeota, E., et al., *Tracking the immune response profiles elicited by the BNT162b2 vaccine in COVID-19 unexperienced and experienced individuals.* Clin Immunol, 2024. **261**: p. 110164.
128. Ewer, K.J., et al., *T cell and antibody responses induced by a single dose of ChAdOx1 nCoV-19 (AZD1222) vaccine in a phase 1/2 clinical trial.* Nat Med, 2021. **27**(2): p. 270-278.
129. Baron, V., et al., *The repertoires of circulating human CD8(+) central and effector memory T cell subsets are largely distinct.* Immunity, 2003. **18**(2): p. 193-204.
130. Schober, K., et al., *Reverse TCR repertoire evolution toward dominant low-affinity clones during chronic CMV infection.* Nat Immunol, 2020. **21**(4): p. 434-441.
131. Goronzy, J.J., et al., *Value of immunological markers in predicting responsiveness to influenza vaccination in elderly individuals.* J Virol, 2001. **75**(24): p. 12182-7.
132. Derhovanessian, E., et al., *Cytomegalovirus-associated accumulation of late-differentiated CD4 T-cells correlates with poor humoral response to influenza vaccination.* Vaccine, 2013. **31**(4): p. 685-90.

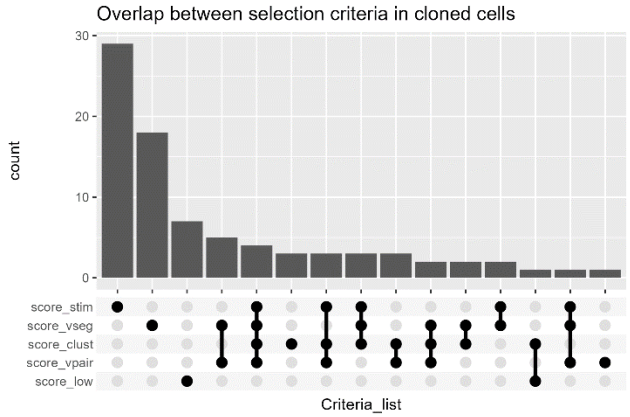
133. Springer, I., N. Tickotsky, and Y. Louzoun, *Contribution of T Cell Receptor Alpha and Beta CDR3, MHC Typing, V and J Genes to Peptide Binding Prediction.* Frontiers in Immunology, 2021. **12**: p. 664514.
134. Minervina, A.A., et al., *Primary and secondary anti-viral response captured by the dynamics and phenotype of individual T cell clones.* Elife, 2020. **9**.
135. Straub, A., et al., *Recruitment of epitope-specific T cell clones with a low-avidity threshold supports efficacy against mutational escape upon re-infection.* Immunity, 2023. **56**(6): p. 1269-1284.e6.
136. Ehrenfried, A.R., et al., *T-Cell-Based Platform for Functional Screening of T-Cell Receptors Identified in Single-Cell RNA Sequencing Data Sets of Tumor-Infiltrating T-Cells.* Bio Protoc, 2024. **14**(8): p. e4972.
137. Boschert, T., et al., *H3K27M neoepitope vaccination in diffuse midline glioma induces B and T cell responses across diverse HLA loci of a recovered patient.* Sci Adv, 2024. **10**(5): p. eadi9091.
138. Oberhardt, V., et al., *Rapid and stable mobilization of CD8+ T cells by SARS-CoV-2 mRNA vaccine.* Nature, 2021. **597**(7875): p. 268-273.
139. Montes, M., et al., *Optimum in vitro expansion of human antigen-specific CD8 T cells for adoptive transfer therapy.* Clin Exp Immunol, 2005. **142**(2): p. 292-302.
140. Martin, M.D., et al., *Bystander responses impact accurate detection of murine and human antigen-specific CD8 T cells.* J Clin Invest, 2019. **129**(9): p. 3894-3908.
141. Watson, N.B., et al., *The gene regulatory basis of bystander activation in CD8(+) T cells.* Sci Immunol, 2024. **9**(92): p. eadf8776.
142. Lertmemongkolchai, G., et al., *Bystander activation of CD8+ T cells contributes to the rapid production of IFN-gamma in response to bacterial pathogens.* J Immunol, 2001. **166**(2): p. 1097-105.
143. Lee, H., S. Jeong, and E.C. Shin, *Significance of bystander T cell activation in microbial infection.* Nat Immunol, 2022. **23**(1): p. 13-22.
144. Heemskerk, M.H.M., et al., *Dual HLA class I and class II restricted recognition of alloreactive T lymphocytes mediated by a single T cell receptor complex.* Proceedings of the National Academy of Sciences, 2001. **98**(12): p. 6806.
145. Grant, E.J., et al., *The unconventional role of HLA-E: The road less traveled.* Mol Immunol, 2020. **120**: p. 101-112.
146. Wang, Y., et al., *Ad5-nCoV Vaccination Could Induce HLA-E Restricted CD8(+) T Cell Responses Specific for Epitopes on Severe Acute Respiratory Syndrome Coronavirus 2 Spike Protein.* Viruses, 2023. **16**(1).
147. Vietzen, H., et al., *HLA-E-restricted immune responses are crucial for the control of EBV infections and the prevention of PTLD.* Blood, 2023. **141**(13): p. 1560-1573.
148. Yang, H., et al., *HLA-E-restricted, Gag-specific CD8(+) T cells can suppress HIV-1 infection, offering vaccine opportunities.* Sci Immunol, 2021. **6**(57).
149. Ruibal, P., et al., *Peptide Binding to HLA-E Molecules in Humans, Nonhuman Primates, and Mice Reveals Unique Binding Peptides but Remarkably Conserved Anchor Residues.* J Immunol, 2020. **205**(10): p. 2861-2872.
150. Meyer, M., et al., *MediMer: a versatile do-it-yourself peptide-receptive MHC class I multimer platform for tumor neoantigen-specific T cell detection.* Front Immunol, 2023. **14**: p. 1294565.
151. Davis, S.J. and P.A. van der Merwe, *TCR triggering: co-receptor-dependent or -independent?* Trends Immunol, 2003. **24**(12): p. 624-6; author reply 626-7.

152. Irvine, D.J., et al., *Direct observation of ligand recognition by T cells.* Nature, 2002. **419**(6909): p. 845-9.
153. Laugel, B., et al., *Different T cell receptor affinity thresholds and CD8 coreceptor dependence govern cytotoxic T lymphocyte activation and tetramer binding properties.* J Biol Chem, 2007. **282**(33): p. 23799-810.
154. Davari, K., et al., *Development of a CD8 co-receptor independent T-cell receptor specific for tumor-associated antigen MAGE-A4 for next generation T-cell-based immunotherapy.* J Immunother Cancer, 2021. **9**(3).
155. Dieng, C.C., et al., *Genetic variations of Plasmodium falciparum circumsporozoite protein and the impact on interactions with human immunoproteins and malaria vaccine efficacy.* Infect Genet Evol, 2023. **110**: p. 105418.
156. Lund, O., et al., *Human leukocyte antigen (HLA) class I restricted epitope discovery in yellow fewer and dengue viruses: importance of HLA binding strength.* PloS One, 2011. **6**(10): p. e26494.
157. Low, J.S., et al., *Clonal analysis of immunodominance and cross-reactivity of the CD4 T cell response to SARS-CoV-2.* Science (New York, N.Y.), 2021. **372**(6548): p. 1336-1341.
158. Ishizuka, A.S., et al., *Protection against malaria at 1 year and immune correlates following PfSPZ vaccination.* Nat Med, 2016. **22**(6): p. 614-23.
159. Vaughan, K., et al., *Meta-analysis of immune epitope data for all Plasmodia: overview and applications for malarial immunobiology and vaccine-related issues.* Parasite Immunol, 2009. **31**(2): p. 78-97.
160. Malik, A., et al., *Human cytotoxic T lymphocytes against the Plasmodium falciparum circumsporozoite protein.* Proc Natl Acad Sci U S A, 1991. **88**(8): p. 3300-4.
161. Slifka, M.K., *Immunological memory to viral infection.* Curr Opin Immunol, 2004. **16**(4): p. 443-50.
162. Lee, W. and M. Suresh, *Vaccine adjuvants to engage the cross-presentation pathway.* Front Immunol, 2022. **13**: p. 940047.
163. Taneichi, M., et al., *Antigen chemically coupled to the surface of liposomes are cross-presented to CD8+ T cells and induce potent antitumor immunity.* J Immunol, 2006. **177**(4): p. 2324-30.
164. Rapaka, R.R., A.S. Cross, and M.A. McArthur, *Using Adjuvants to Drive T Cell Responses for Next-Generation Infectious Disease Vaccines.* Vaccines (Basel), 2021. **9**(8).
165. Beura, L.K., S.C. Jameson, and D. Masopust, *Is a Human CD8 T-Cell Vaccine Possible, and if So, What Would It Take? CD8 T-Cell Vaccines: To B or Not to B?* Cold Spring Harbor Perspectives in Biology, 2018. **10**(9): p. a028910.
166. Moutaftsi, M., et al., *Uncovering the interplay between CD8, CD4 and antibody responses to complex pathogens.* Future Microbiol, 2010. **5**(2): p. 221-39.
167. Hoffman, S.L. and D.L. Doolan, *Can malaria DNA vaccines on their own be as immunogenic and protective as prime-boost approaches to immunization?* Developments in Biologicals, 2000. **104**: p. 121-132.
168. Hoffman, S.L., et al., *Irradiated sporozoite vaccine induces cytotoxic T lymphocytes that recognize malaria antigens on the surface of infected hepatocytes.* Immunology Letters, 1990. **25**(1-3): p. 33-38.
169. Romero, P., et al., *Cloned cytotoxic T cells recognize an epitope in the circumsporozoite protein and protect against malaria.* Nature, 1989. **341**(6240): p. 323-6.
170. Hoffman, S.L., et al., *Sporozoite vaccine induces genetically restricted T cell elimination of malaria from hepatocytes.* Science, 1989. **244**(4908): p. 1078-81.

171. Bongfen, S.E., et al., *Plasmodium berghei-infected primary hepatocytes process and present the circumsporozoite protein to specific CD8+ T cells in vitro.* J Immunol, 2007. **178**(11): p. 7054-63.
172. Ma, J., et al., *Dynamics of the Major Histocompatibility Complex Class I Processing and Presentation Pathway in the Course of Malaria Parasite Development in Human Hepatocytes: Implications for Vaccine Development.* PLoS ONE, 2013. **8**(9): p. e75321.
173. Rénia, L., et al., *Effector functions of circumsporozoite peptide-primed CD4+ T cell clones against Plasmodium yoelii liver stages.* J Immunol, 1993. **150**(4): p. 1471-8.
174. Rénia, L., et al., *In vitro activity of CD4+ and CD8+ T lymphocytes from mice immunized with a synthetic malaria peptide.* Proc Natl Acad Sci U S A, 1991. **88**(18): p. 7963-7.
175. Oliveira, G.A., et al., *Class II-restricted protective immunity induced by malaria sporozoites.* Infect Immun, 2008. **76**(3): p. 1200-6.
176. Furtado, R., et al., *Cytolytic circumsporozoite-specific memory CD4(+) T cell clones are expanded during Plasmodium falciparum infection.* Nat Commun, 2023. **14**(1): p. 7726.
177. Braendstrup, P., et al., *Identification and HLA-Tetramer-Validation of Human CD4+ and CD8+ T Cell Responses against HCMV Proteins IE1 and IE2.* PLoS ONE, 2014. **9**(4): p. e94892.

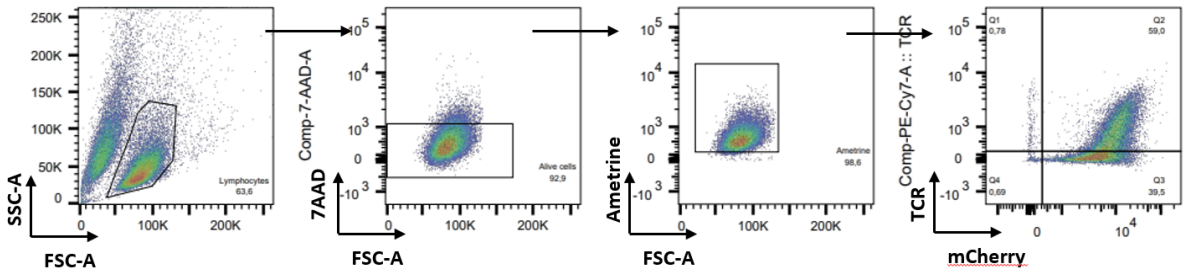
14. Supplement

14.1. Supplementary figures



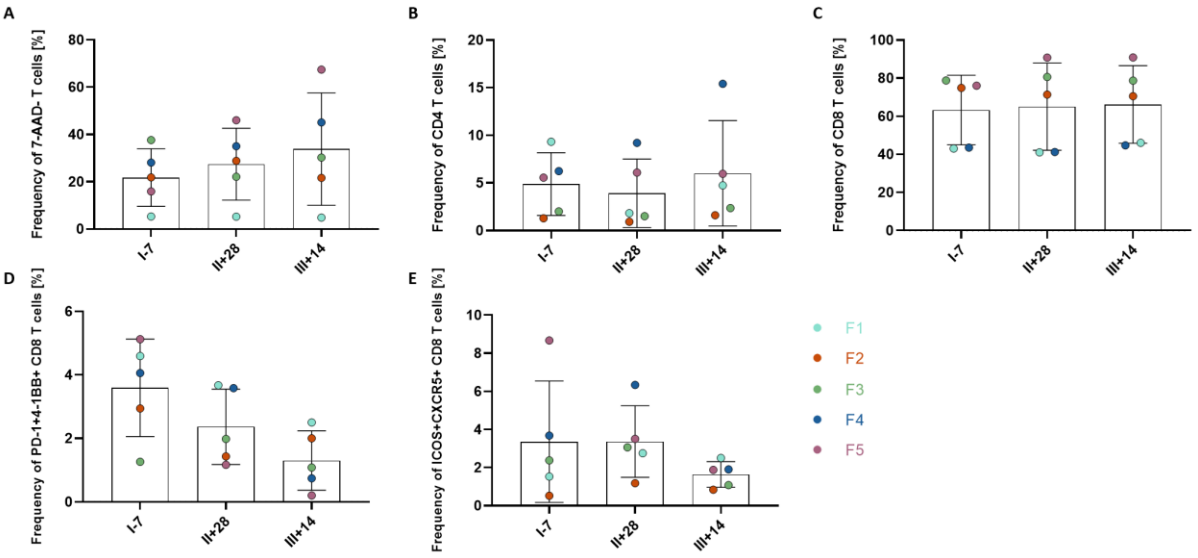
Supplementary figure 1: CD4 TCR cloning selection criteria.

86 TCRs were selected for functional characterization based on clonal expansion in stimulation culture (score stim and score_low), V segment usage (score_vseg), V segment pairing (score_vpair), similarity clustering (score_clust).



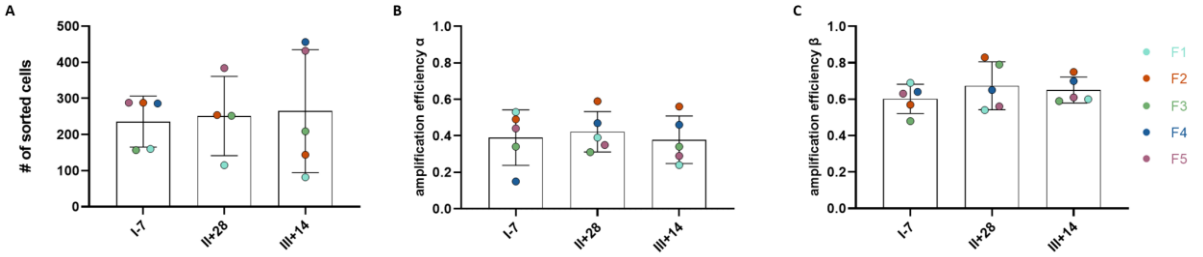
Supplementary figure 2: Gating strategy of TCR-transgenic Jurkat76 cell lines.

TCR constructs were expressed by retroviral transduction and transduction efficiency was assessed by flow cytometric analysis of TCR surface expression and mCherry expression.



Supplementary figure 3: Immune cell frequencies in vaccinated individuals remain stable throughout FMP013 vaccination.

Frequency of live T cells (CD3+7AAD-), CD4 T cells (CD3+7AAD-CD4+), CD8 T cells (CD3+7AAD-CD8+), activated CD8 T cells by PD-1-4-1BB (CD137) or ICOS-CXCR5 expression among gating parent. Frequencies are shown for five individual donors (F1-F5) at three timepoints before (I-7) and after (II+28, III+14) FMP013 vaccination.



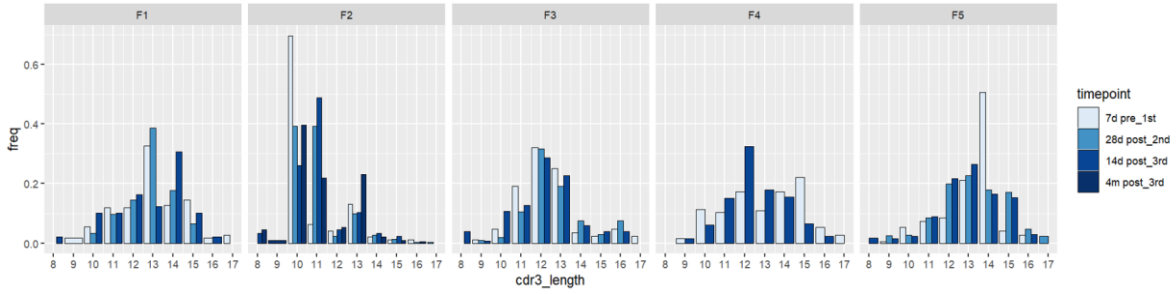
Supplementary figure 4: Constant TCR alpha and beta amplification efficiency across timepoints and donors.

A Number of sorted activated T_{EM} and T_{EMRA} cells per donor and time point. **B-C** Amplification efficiency of TCR alpha (B) and beta (C) genes during PCR amplification and sequencing. Amplification efficiency was calculated based on obtained sequences.

Public clone ID	Clone ID	Donor	B_cdr3	Time point
1	559	F2	ASSSANYGYT	proliferation
1	956	F2	ASSSANYGYT	III+14
1	931	F3	ASSSANYGYT	III+14
1	944	F3	ASSSANYGYT	II+28, III+14
2	605	F4	ASSLDYNEQF	III+14
2	431	F5	ASSLDYNEQF	I-7

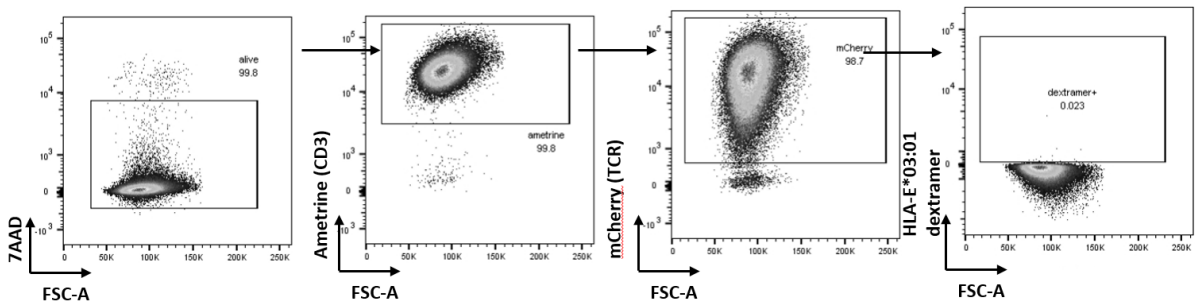
Supplementary figure 5: Public T cell clones across donors appear at low frequency.

Table of two public T cell clones identified among activated T_{EM} and T_{EMRA} CD8 repertoire.



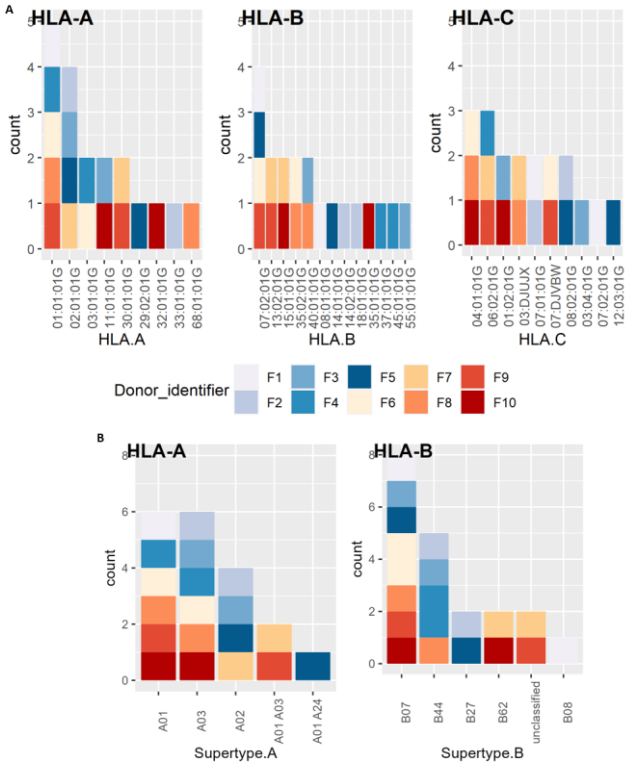
Supplementary figure 6: Donor-specific CDR3 length distribution changes in response to FMP013 vaccination.

Supplementary Figure 6: **Donor-specific CDR3 length distribution changes in response to FMP013 vaccination.** TCR beta CDR3 amino acid length distribution after clonal reduction was quantified for each donor at 4 individual time points before and after subsequent FMP013 vaccinations.



Supplementary figure 7: HLA-E*03:01 dextramer staining of TCR B11 expression Jurkat cell line

Cells were pre-gated on alive (7AAD-), ametrine and TCR expression. Dextramer staining revealed 0.02 % stained cells.



Supplementary figure 8: Broad HLA I haplotype distribution in FMP013 donors.

A HLA-A, HLA-B and HLA-C alleles of ten FMP013 donors (F1-F10) identified by HLA typing. **B** HLA-A and HLA-B supertypes of ten donors according to classification by Sidney et al. [94].

14.2. Supplementary tables

Supplementary Table 1: Number of sorted T cell population and *TR* gene amplification efficiencies.

Sorted population	Sorted cells	Beta sequences	Alpha sequences
Activated CD8 T cells non-expanded	4305	2653	1657
Expanded CD8 T cells	1662	739	454
Expanded CD4 T cells	1038	761	501

Supplementary Table 2: TCR sequence features and reactivity of TCRs cloned from *in vitro* expanded, activated CD4 T cells.

TCR ID	Donor	Time point	Population	TRBV	TRBJ	TRBC	TRAV	TRAJ	TCR exp.	Reactivity
T21F7x1604	F7	II+28	CD4	7-2	2-7	2	10-1	18	32.7	
T21F3x2797	F3	II+28	CD4	5-1	2-6	2	12-1	20	42.8	CSP61
T21F7x1777	F7	II+28	CD4	7-2	2-1	2	10-1	18	56.4	
T21F7x1643	F7	II+28	CD4	7-2	2-7	2	10-1	18	66.8	
T21F9x2172	F9	II+28	CD4	7-3	1-2	1	10-1	43		
T21F4x3722	F4	II+28	CD4	27-1	1-4	1	1-2	33	69.7	CSP81
T21F3x2896	F3	II+28	CD4	12-4	1-5	1	12-1	40		
T21F3x2855	F3	II+28	CD4	5-1	2-6	2	12-1	20	30.8	CSP61
T21F3x2742	F3	II+28	CD4	5-1	2-6	2	12-1	20	69.5	CSP61
T21F4x3804	F4	II+28	CD4	20-1	2-1	2	12-3	45	68	
T21F4x3818	F4	II+28	CD4	19-1	2-7	2	13-1	47	27.8	
T21F9x2086	F9	II+28	CD4	6-1	2-7	2	14DV4	13	63	CS.T3
T21F3x2835	F3	II+28	CD4	12-3	1-5	1	17-1	11	71.6	Th2R
T21F4x3756	F4	II+28	CD4	5-1	2-3	2	12-3	26		
T21F3x2801	F3	II+28	CD4	20-1	2-7	2	13-1	15	61.7	Th2R
T21F3x2911	F3	II+28	CD4	18-1	1-2	1	13-1	53	72.2	Th2R
T21F4x3795	F4	II+28	CD4	20-1	1-4	1	13-1	29		
T21F4x3682	F4	II+28	CD4	20-1	2-4	2	13-1	53	61.8	Th2R
T21F3x2931	F3	II+28	CD4	3-1	1-3	1	17-1	34	59.2	Th2R
T21F4x3646	F4	II+28	CD4	20-1	2-5	2	17-1	50	41.5	
T21F3x3023	F3	II+28	CD4	14-1	1-2	1	19-1	43	46.9	
T21F3x2838	F3	II+28	CD4	12-3	1-5	1	17-1	11	66.7	Th2R
T21F3x2862	F3	II+28	CD4	6-4	1-2	1	17-1	11	67.6	Th2R
T21F3x2915	F3	II+28	CD4	7-2	1-5	1	17-1	11	41.5	Th2R
T21F3x2903	F3	II+28	CD4	7-2	2-1	2	17-1	11	49.5	Th2R
T21F3x2945	F3	II+28	CD4	12-3	1-5	1	17-1	11	63.9	Th2R
T21F3x2960	F3	II+28	CD4	12-3	1-5	1	17-1	11	71.5	Th2R
T21F3x2968	F3	II+28	CD4	12-4	1-2	1	17-1	33	28.8	
T21F3x2863	F3	II+28	CD4	7-2	1-5	1	17-1	11	56	Th2R
T21F3x2893	F3	II+28	CD4	3-1	2-2	2	19-1	26	45.9	CSP61
T21F3x2713	F3	II+28	CD4	7-8	1-5	1	25-1	28		

T21F3x2939	F3	II+28	CD4	12-3	1-3	1	26-1	48	76.1	
T21F3x3026	F3	II+28	CD4	7-2	2-3	2	26-1	23	53.3	
T21F4x3761	F4	II+28	CD4	20-1	1-5	1	20-1	4	0.15	
T21F4x3675	F4	II+28	CD4	20-1	1-1	1	2-1	37		
T21F4x3798	F4	II+28	CD4	10-2	2-7	2	21-1	16	5.62	
T21F3x2949	F3	II+28	CD4	12-3	1-2	1	23DV6	23	43.3	
T21F3x2731	F3	II+28	CD4	30-1	1-5	1	23DV6	36		
T21F3x2714	F3	II+28	CD4	30-1	1-5	1	24-1	28	75.3	T1
T21F3x2705	F3	II+28	CD4	30-1	1-3	1	24-1	32	68.9	T1
T21F3x2802	F3	II+28	CD4	30-1	1-3	1	26-1	52		
T21F3x2996	F3	II+28	CD4	5-1	1-5	1	29DV5	40	71.8	
T21F3x2851	F3	II+28	CD4	30-1	1-6	1	3-1	36	81	T1
T21F3x2865	F3	II+28	CD4	7-2	2-3	2	35-1	17	38.8	
T21F3x2807	F3	II+28	CD4	30-1	2-3	2	26-1	27	62.3	
T21F3x3008	F3	II+28	CD4	27-1	1-6	1	26-2	40	60.1	
T21F4x3826	F4	II+28	CD4	20-1	2-5	2	27-1	34		
T21F7x1635	F7	II+28	CD4	5-8	1-1	1	29DV5	45	58.5	
T21F7x1698	F7	II+28	CD4	30-1	2-2	2	29DV5	18	64.3	T*
T21F9x2073	F9	II+28	CD4	6-1	1-5	1	29DV5	53	64	
T21F9x2068	F9	II+28	CD4	19-1	1-2	1	29DV5	53		
T21F3x2708	F3	II+28	CD4	7-9	2-1	2	35-1	40	0.33	
T21F9x2018	F9	II+28	CD4	12-4	2-7	2	39-1	48	61.1	
T21F4x3758	F4	II+28	CD4	28-1	2-1	2	41-1	57	24	
T21F4x3700	F4	II+28	CD4	7-2	2-7	2	41-1	48		
T21F4x3724	F4	II+28	CD4	20-1	2-2	2	35-1	49	56.4	
T21F4x3647	F4	II+28	CD4	2-1	1-2	1	35-1	48		
T21F4x3653	F4	II+28	CD4	28-1	2-1	2	35-1	28	60.3	
T21F3x2777	F3	II+28	CD4	30-1	1-2	1	36DV7	54	66.1	T1
T21F3x2717	F3	II+28	CD4	30-1	2-7	2	36DV7	54	64	
T21F3x2907	F3	II+28	CD4	5-1	2-1	2	38-2-DV8	52	75	
T21F4x3719	F4	II+28	CD4	12-4	1-5	1	38-2-DV8	53		
T21F9x2033	F9	II+28	CD4	7-9	1-1	1	6-1	37		
T21F3x2989	F3	II+28	CD4	5-1	2-5	2	9-2	52	64.1	
T21F3x3019	F3	II+28	CD4	5-1	2-5	2	9-2	52	69	Th2R
T21F3x2898	F3	II+28	CD4	5-1	2-6	2	9-2	49	72.3	
T21F3x2795	F3	II+28	CD4	3-1	2-4	2	8-3	45	12	
T21F4x3806	F4	II+28	CD4	20-1	2-2	2	8-3	15	31.4	
T21F3x2917	F3	II+28	CD4	7-7	1-1	1	8-4	15	34.6	Th2R
T21F3x3020	F3	II+28	CD4	12-3	1-1	1	8-4	50	76	
T21F4x3712	F4	II+28	CD4	5-4	2-7	2	8-4	49	63.2	
T21F3x3018	F3	II+28	CD4	14-1	1-2	1	9-2	20	70.3	
T21F3x2923	F3	II+28	CD4	14-1	1-5	1	9-2	30		
T21F3x2936	F3	II+28	CD4	14-1	1-1	1	9-2	34		
T21F3x2975	F3	II+28	CD4	5-1	2-5	2	9-2	52	71.6	Th2R

T21F4x3828	F4	II+28	CD4	20-1	2-3	2	9-2	39		
T21F3x2859	F3	II+28	CD4	7-3	2-5	2	9-2	4	0.16	
T21F4x3825	F4	II+28	CD4	20-1	2-5	2	9-2	10	68.9	CS.T3
T21F4x3702	F4	II+28	CD4	20-1	2-7	2	9-2	58	58.3	
T21F4x3791	F4	II+28	CD4	20-1	2-1	2	9-2	32	79.8	
T21F4x3706	F4	II+28	CD4	20-1	2-1	2	9-2	10	78.4	
T21F4x3701	F4	II+28	CD4	20-1	2-5	2	9-2	57	68.4	CS.T3
T21F4x3685	F4	II+28	CD4	20-1	2-2	2	9-2	49	77.5	
T21F4x3762	F4	II+28	CD4	20-1	2-5	2	9-2	35	68.4	CS.T3

Supplementary Table 3: TCR sequence features and reactivity of TCRs cloned from activated CD8 T_{EM} and T_{EMRA} cells.

TCR ID	Donor	Time point	Population	TRBV	TRBJ	TRBC	TRAV	TRAJ	TCR exp.	Reactivity
T19F1x1257	F1	II+28	T_{EM}+T_{EMRA}	27	2-1	2	2	30	16	Non-reactive
T19F1x1620	F1	II+28	T_{EM}+T_{EMRA}	7-9	2-5	2	38-2-DV8	42	7	Non-reactive
T19F1x1662	F1	II+28	T_{EM}+T_{EMRA}	15-1	1-1	1	12-2	30		
T19F1x2356	F1	III+14	T_{EM}+T_{EMRA}	12-3	1-1	1	29/DV5	43	58	Non-reactive
T19F1x2379	F1	III+14	T_{EM}+T_{EMRA}	7-9	2-4	2	29/DV5	43	3	
T19F1x2391	F1	III+14	T_{EM}+T_{EMRA}	11-2	2-1	2	13-2	47	24	Non-reactive
T19F1x2423	F1	III+14	T_{EM}+T_{EMRA}	27	2-7	2	26-1	20	1	
T19F1x2716	F1	III+14	T_{EM}+T_{EMRA}	15-1	1-1	1	24-1	32		
T19F1x2762	F1	III+14	T_{EM}+T_{EMRA}	7-9	2-3	2	29DV5	49		
T19F2x1302	F2	II+28	T_{EM}+T_{EMRA}	27	2-7	2	12-3	13	71	EBV reactive
T19F3x1281	F3	III+14	T_{EM}+T_{EMRA}	5-1	2-3	2	12-2	39		
T19F3x2811	F3	III+14	T_{EM}+T_{EMRA}	2-1	1-2	1	19-1	40		
T19F3x874	F3	III+14	T_{EM}+T_{EMRA}	7-9	1-6	1	41	57	41	Non-reactive
T19F3x906	F3	III+14	T_{EM}+T_{EMRA}	27	2-5	2	29/DV5	47	45	Non-reactive
T19F4x3235	F4	III+14	T_{EM}+T_{EMRA}	4-3	2-2	2	12-2	49	29	Non-reactive
T19F4x3299	F4	III+14	T_{EM}+T_{EMRA}	12-3	2-1	2	12-3	49	5	
T19F4x3308	F4	III+14	T_{EM}+T_{EMRA}	12-3	1-2	1	12-2	49	69	Non-reactive
T19F4x3316	F4	III+14	T_{EM}+T_{EMRA}	3-1	2-3	2	19-1	34		
T19F4x3320	F4	III+14	T_{EM}+T_{EMRA}	12-3	1-5	1	12-3	49	47	Non-reactive
T19F4x3383	F4	III+14	T_{EM}+T_{EMRA}	4-3	2-7	2	12-3	49	2	
T19F5x1657	F5	III+14	T_{EM}+T_{EMRA}	18	1-1	1	41	48	61	Non-reactive
T19F5x1793	F5	III+14	Tet+	28	1-6	1	26-2	48	83	Non-reactive
T19F5x1850	F5	III+14	Tet+	28	2-3	2	12-2	45	59	Non-reactive
T19F5x2037	F5	III+14	T_{EM}+T_{EMRA}	7-2	2-1	2	19-1	22		
T19F5x2067	F5	III+14	T_{EM}+T_{EMRA}	5-6	2-5	2	5-1	15	1	

T19F5x2073	F5	III+14	T_{EM}+T_{EMRA}	5-6	1-2	1	26-1	37	2	
T19F5x2130	F5	III+14	T_{EM}+T_{EMRA}	28-1	2-1	2	12-1	49	42	Non-reactive
T19F5x2143	F5	III+14	T_{EM}+T_{EMRA}	7-6	1-5	1	26-1	49	60	Non-reactive
T19F5x2167	F5	III+14	Tet+	5-6	1-1	1	14DV4	53		
T19F5x2187	F5	III+14	Tet+	7-8	2-1	2	21-1	49	1	
T19F5x2201	F5	III+14	Tet+	9-1	2-1	2	13-1	15	63	Non-reactive
T19F5x2231	F5	III+14	Tet+	6-2	2-1	2	14DV4	9		
T19F5x2242	F5	III+14	Tet+	7-9	1-6	1	8-2	12	25	Non-reactive
T19F5x2245	F5	III+14	Tet+	28-1	2-3	2	12-1	49	63	Non-reactive
T19F5x2251	F5	III+14	Tet+	27	1-2	1	19	44	52	Non-reactive
T19F5x4251	F5	II+28	T_{EM}+T_{EMRA}	20-1	2-7	2	27	24	50	Non-reactive
T19F5x4261	F5	II+28	T_{EM}+T_{EMRA}	28	2-3	2	12-2	49	54	Non-reactive
T19F5x4313	F5	II+28	T_{EM}+T_{EMRA}	20-1	2-5	2	30	12	22	Non-reactive
T19F5x4336	F5	II+28	T_{EM}+T_{EMRA}	28	2-1	2	2	26	9	Non-reactive
T19F5x4355	F5	II+28	T_{EM}+T_{EMRA}	25-1	1-3	1	13-2	16	34	Non-reactive
T19F5x4356	F5	II+28	T_{EM}+T_{EMRA}	28	1-6	1	12-2	28	67	Non-reactive
T19F5x4361	F5	II+28	T_{EM}+T_{EMRA}	12-4	2-3	2	12-2	49	64	Non-reactive
T19F5x4395	F5	II+28	T_{EM}+T_{EMRA}	5-8	2-7	2	14/DV4	8	69	Non-reactive
T19F5x4452	F5	II+28	Tet+	5-6	2-3	2	13-2	48	48	Non-reactive
T19F5x4469	F5	II+28	Tet+	27	1-4	1	21	26	69	Non-reactive
T19F5x4473	F5	II+28	Tet+	7-6	1-4	1	26-2	43	66	Non-reactive
T19F5x4501	F5	II+28	Tet+	9	1-5	1	19	30	41	Non-reactive
T19F5x4523	F5	II+28	Tet+	12-4	2-7	2	1-1	41	25	Non-reactive
T19F5x4529	F5	II+28	Tet+	27	2-1	2	21	26	63	Non-reactive
T19F5x4714	F5	II+28	T_{EM}+T_{EMRA}	27-1	2-2	2	27-1	50		
T19F5x4720	F5	II+28	T_{EM}+T_{EMRA}	28-1	2-1	2	2-1	26		
T19F5x4785	F5	II+28	T_{EM}+T_{EMRA}	27-1	2-7	2	26-1	34		
T19F5x4805	F5	II+28	T_{EM}+T_{EMRA}	5-6	1-1	1	14DV4	53		
T19F5x4850	F5	II+28	Tet+	19-1	2-1	2	12-2	42	28	Non-reactive
T19F5x4856	F5	II+28	Tet+	5-6	1-1	1	30-1	37		
T19F5x4886	F5	II+28	Tet+	7-8	2-4	2	21-1	49	47	Non-reactive
T19F5x4889	F5	II+28	Tet+	18-1	2-7	2	20-1	18		
T19F5x4909	F5	II+28	Tet+	20-1	2-7	2	8-6	52	62	Non-reactive
T19F5x4912	F5	II+28	Tet+	19-1	1-5	1	12-2	30	57	Non-reactive

T19F5x4947	F5	II+28	Tet+	4-3	2-3	2	24-1	42		
T20F2x1197	F2	II+28	T_{EM}+T_{EMRA}	6-1	2-7	2	26-2	43	1	
T20F2x2439	F2	III+14	T_{EM}+T_{EMRA}	5-8	2-5	2	13-2	45	52	Non-reactive
T20F2x2549	F2	III+14	Tet+	10-2	2-7	2	12-1	37	53	Non-reactive
T20F2x3429	F2	II+28	CD8	11-2	2-5	2	13-1	45	1	
T20F3x2019	F3	II+28	T_{EM}+T_{EMRA}	29-1	2-1	2	38-1	48		
T20F3x2177	F3	II+28	T_{EM}+T_{EMRA}	2-1	2-7	2	13-1	24	62	Non-reactive
T20F4x2594	F4	III+14	T_{EM}+T_{EMRA}	6-2	2-4	2	17	10	2	
T20F4x2630	F4	III+14	T_{EM}+T_{EMRA}	3-1	1-5	1	12-2	23	51	Non-reactive
T20F4x2632	F4	III+14	T_{EM}+T_{EMRA}	12-3	2-7	2	26-2	20		
T20F4x2658	F4	III+14	T_{EM}+T_{EMRA}	12-3	2-7	2	26-1	20	68	Non-reactive
T20F4x3312	F4	II+28	CD8	7-9	1-2	1	29/DV5	52	30	EBV reactive
T20F5x2724	F5	II+28	Tet+	18-1	2-5	2	16-1	39	31	Non-reactive
T20F5x2754	F5	II+28	Tet+	2-1	1-1	1	41-1	45		
T20F5x2765	F5	II+28	Tet+	6-2	2-7	2	10-1	34	35	Non-reactive
T20F5x2819	F5	II+28	Tet+	19-1	2-1	2	29DV5	26	1	
T20F5x2820	F5	II+28	Tet+	7-2	2-2	2	1-2	31	60	Non-reactive
T20F5x2837	F5	II+28	Tet+	12-4	1-5	1	26-2	48		
T20F5x2874	F5	II+28	Tet+	5-6	2-3	2	4-1	5	67	Non-reactive
T20F5x2913	F5	II+28	T_{EM}+T_{EMRA}	11-3	2-1	2	12-2	20	36	Non-reactive
T20F5x2934	F5	II+28	T_{EM}+T_{EMRA}	27-1	2-2	2	8-3	35	3	
T20F5x2984	F5	II+28	T_{EM}+T_{EMRA}	3-1	2-1	2	8-4	54	55	Non-reactive
T20F5x2993	F5	II+28	T_{EM}+T_{EMRA}	19-1	2-1	2	2-1	27		
T20F5x3019	F5	II+28	T_{EM}+T_{EMRA}	20-1	2-5	2	8-2	12	56	Non-reactive
T20F5x3039	F5	II+28	T_{EM}+T_{EMRA}	15-1	2-1	2	22-1	35	57	Non-reactive
T20F5x3873	F5	III+14	T_{EM}+T_{EMRA}	10-3	2-5	2	8-3	13	53	Non-reactive
T20F5x3966	F5	III+14	T_{EM}+T_{EMRA}	13-1	2-1	2	21-1	42		
T20F5x3992	F5	III+14	T_{EM}+T_{EMRA}	6-5	1-2	1	4-1	41	73	Non-reactive
T20F5x4028	F5	III+14	T_{EM}+T_{EMRA}	7-9	2-1	2	19-1	12	20	
T20F5x4029	F5	III+14	T_{EM}+T_{EMRA}	19-1	2-7	2	19-1	42	36	Non-reactive
T20F5x4035	F5	III+14	T_{EM}+T_{EMRA}	6-2	1-5	1	41-1	45	54	Non-reactive
T20F5x4038	F5	III+14	T_{EM}+T_{EMRA}	3-1	1-1	1	38-2-DV8	40	63	Non-reactive
T20F5x4043	F5	III+14	T_{EM}+T_{EMRA}	19-1	1-2	1	41-1	45	1	
T20F5x4044	F5	III+14	T_{EM}+T_{EMRA}	27-1	1-1	1	1-1	24	1	
T20F5x4068	F5	III+14	T_{EM}+T_{EMRA}	28-1	2-7	2	8-6	30		

T20F5x4097	F5	III+14	Tet+	19-1	2-1	2	8-2	30	48	Non-reactive
T20F5x4118	F5	III+14	Tet+	7-8	1-1	1	4-1	9		
T20F5x4133	F5	III+14	Tet+	6-2	2-1	2	9-2	40	1	
T20F5x4140	F5	III+14	Tet+	18-1	2-7	2	12-1	44	37	Non-reactive
T20F5x4161	F5	III+14	Tet+	28-1	2-3	2	12-2	49		
T20F5x4162	F5	III+14	Tet+	20-1	1-5	1	12-2	24	79	Non-reactive
T20F5x4208	F5	III+14	Tet+	6-2	1-1	1	12-2	12	55	Non-reactive
T20F5x4215	F5	III+14	Tet+	19-1	1-2	1	38-2-DV8	27	54	Non-reactive
T19F5x743	F5	III+14	T_{EM}+T_{EMRA}	2-1	2-2	2	12-1	12	70.8	Non-reactive
T19F4x2853	F4	III+14	T_{EM}+T_{EMRA}	12-3	1-2	1	12-1	31	44.3	Non-reactive
T19F5x4436	F5	II+28	T_{EM}+T_{EMRA}	28-1	2-1	2	1-2	27	13.7	Non-reactive
T19F4x2920	F4	III+14	T_{EM}+T_{EMRA}	12-3	1-2	1	12-2	49		
T19F4x2912	F4	III+14	T_{EM}+T_{EMRA}	11-2	1-1	1	12-2	11		
T19F4x3016	F4	III+14	T_{EM}+T_{EMRA}	12-3	1-1	1	12-2	31	23.5	Non-reactive
T19F4x2888	F4	III+14	T_{EM}+T_{EMRA}	5-6	2-5	2	17-1	22		
T19F3x915	F3	III+14	Tet+	20-1	2-7	2	24-1	34	50.9	Non-reactive
T19F5x4265	F5	II+28	T_{EM}+T_{EMRA}	27-1	2-2	2	27-1	50	36.6	Non-reactive
T19F3x856	F3	III+14	T_{EM}+T_{EMRA}	12-4	1-2	1	35-1	50	38.5	Non-reactive

Supplementary Table 4: TCR sequence features and reactivity of TCRs cloned from *in vitro* expanded, activated CD8 T cells.

TCR ID	Donor	Time point	population	TRBV	TRBJ	TRBC	TRAV	TRAJ	TCR exp.	Reactivity
T21F1x582	F1	II+28	CD8	19-1	1-3	1	29DV5	43	4.32	
T21F1x615	F1	II+28	CD8	15-1	1-1	1	12-2	30	24.3	Non-reactive
T21F3x625	F3	II+28	CD8	7-9	1-5	1	21-1	30	0.85	
T21F3x649	F3	II+28	CD8	12-4	1-1	1	29DV5	30	14.1	Non-reactive
T21F3x661	F3	II+28	CD8	5-1	2-7	2	25-1	32	56.2	Non-reactive
T21F3x820	F3	II+28	CD8	19-1	1-6	1	10-1	42	1.22	
T21F3x857	F3	II+28	CD8	3-1	1-3	1	8-6	23	61.3	Non-reactive
T21F3x863	F3	II+28	CD8	20-1	2-7	2	24-1	34	78.6	Non-reactive
T21F3x883	F3	II+28	CD8	12-4	1-2	1	35-1	50	52.4	Non-reactive
T21F3x1001	F3	II+28	CD8	12-4	2-7	2	27-1	40	70.5	Non-reactive
T21F3x1010	F3	II+28	CD8	27-1	2-1	2	6-1	29	17.2	Non-reactive
T21F3x1011	F3	II+28	CD8	11-2	2-5	2	12-1	39	2.7	
T21F3x1025	F3	II+28	CD8	29-1	2-7	2	35-1	20	63	Non-reactive

T21F3x1031	F3	II+28	CD8	5-1	1-2	1	29DV5	40	29.5	Non-reactive
T21F3x1037	F3	II+28	CD8	12-4	2-4	2	22-1	29	55	Non-reactive
T21F3x1053	F3	II+28	CD8	5-5	1-3	1	22-1	48	77.6	Non-reactive
T21F3x1091	F3	II+28	CD8	27-1	1-1	2	26-2	47	0.5	
T21F3x1101	F3	II+28	CD8	5-1	1-2	1	17-1	53	61.4	Non-reactive
T21F3x1113	F3	II+28	CD8	5-1	1-1	1	19-1	57	59.2	Non-reactive
T21F7x1566	F7	II+28	CD8	7-9	1-2	1	12-2	26	7.06	Non-reactive
T21F7x1757	F7	II+28	CD8	30-1	2-2	2	12-2	15	57.3	Non-reactive
T21F7x1811	F7	II+28	CD8	19-1	2-1	2	12-1	8	30.9	Non-reactive
T21F9x1889	F9	II+28	CD8	7-2	1-2	1	36DV7	53	47	CS.T3
T21F7x1907	F7	II+28	CD8	30-1	2-2	2	12-2	15	62.9	Non-reactive
T21F9x2160	F9	II+28	CD8	7-2	1-4	1	1-2	20	66.6	Non-reactive
T21F7x2213	F7	II+28	CD8	7-9	2-1	2	12-2	24	34.2	Non-reactive
T21F7x2225	F7	II+28	CD8	5-8	2-3	2	29DV5	45	51.2	Non-reactive
T21F4x2344	F4	II+28	CD8	5-1	2-7	2	1-2	33	69.4	Non-reactive
T21F4x2345	F4	II+28	CD8	20-1	2-1	2	26-1	9	64.6	Non-reactive
T21F4x2349	F4	II+28	CD8	19-1	2-7	2	12-3	40	0.41	
T21F4x2356	F4	II+28	CD8	6-2	2-4	2	17-1	10	0.35	
T21F4x2375	F4	II+28	CD8	19-1	2-1	2	19-1	7	0	
T21F4x2376	F4	II+28	CD8	4-1	1-2	1	8-3	35	13.1	Non-reactive
T21F4x2396	F4	II+28	CD8	19-1	2-7	2	22-1	24	58.3	Non-reactive
T21F4x2402	F4	II+28	CD8	20-1	2-2	2	39-1	16	49.9	Non-reactive
T21F4x2454	F4	II+28	CD8	4-1	2-7	2	5-1	18		
T21F4x2483	F4	II+28	CD8	5-1	2-3	2	25-1	21	5.92	Non-reactive
T21F4x2507	F4	II+28	CD8	12-3	2-2	2	1-2	33	68.7	Non-reactive
T21F4x2512	F4	II+28	CD8	20-1	2-7	2	13-1	13	46.8	Non-reactive
T21F4x2528	F4	II+28	CD8	5-1	2-1	2	17-1	57	51.7	Non-reactive

14.3. Primer sequences

Supplementary Table 5: PCR primers for generation of TCR alpha amplicons. the common linker sequence is highlighted in red.

First PCR forward primers

hTRAV1-1	CCAGGGTTTTCCCAGTCACGACCTGCAGGACAAAGCCTTGAGC
hTRAV1-2	CCAGGGTTTTCCCAGTCACGACGATGGGAGGCACTACAGGAC
hTRAV2	CCAGGGTTTTCCCAGTCACGACGCAAGGACCAAGTGTTTCAGCC
hTRAV3	CCAGGGTTTTCCCAGTCACGACGAGAGCTCAGTCAGTGGCTCAGC
hTRAV4	CCAGGGTTTTCCCAGTCACGACCCTTGCTAAGACCACCCAGCC
hTRAV5	CCAGGGTTTTCCCAGTCACGACGAGTAGAGGAGAGGATGTGGAGC
hTRAV6	CCAGGGTTTTCCCAGTCACGACGTGGACTGGGTGAAGAGCC
hTRAV7	CCAGGGTTTTCCCAGTCACGACGTCTTGGCTGGGCAAATGGAG
hTRAV8-1	CCAGGGTTTTCCCAGTCACGACGATGCCAGAGCCCAGTCTGTG
hTRAV8-2,4,6	CCAGGGTTTTCCCAGTCACGACGGAGGAACCAGAGCCCAGTC
hTRAV8-3	CCAGGGTTTTCCCAGTCACGACCCAGAGCCCAGTCAGTGACC
hTRAV9-1	CCAGGGTTTTCCCAGTCACGACGGAGATTCAGTGGTCCAGACAG
hTRAV9-2	CCAGGGTTTTCCCAGTCACGACCGTGGAGATTCAGTGACCCAG
hTRAV10	CCAGGGTTTTCCCAGTCACGACGGGAATGGCAAAAACCAAGTGG
hTRAV12-1	CCAGGGTTTTCCCAGTCACGACGAGCCAACGGAAGGAGGTG
hTRAV12-2,3	CCAGGGTTTTCCCAGTCACGACGAGCCAACAGAAGGAGGTGG
hTRAV13-1	CCAGGGTTTTCCCAGTCACGACGGAGAGAATGTGGAGCAGCATCC
hTRAV13-2	CCAGGGTTTTCCCAGTCACGACGGTGAGCAGAGGAGAGAGTGTGG
hTRAV14DV4	CCAGGGTTTTCCCAGTCACGACGGCATTGCCCAGAAGATAACTC
hTRAV16	CCAGGGTTTTCCCAGTCACGACGGAACAAGAGCCCAGAGAGTGAC
hTRAV17	CCAGGGTTTTCCCAGTCACGACGAACAGTCAACAGGGAGAAGAGG

hTRAV18	CCAGGGTTTTCCCAGTCACGACCCAGTGGAGACTCGGTTACCCAG
hTRAV19	CCAGGGTTTTCCCAGTCACGACGGCTCAGAAGGTAACTCAAGCG
hTRAV20	CCAGGGTTTTCCCAGTCACGACGAGTGGAGAAGACCAGGTGACGC
hTRAV21	CCAGGGTTTTCCCAGTCACGACGAGCAGCAAACAGGAGGTGACG
hTRAV22	CCAGGGTTTTCCCAGTCACGACGGAATACAAGTGGAGCAGAGTCC
hTRAV23DV6	CCAGGGTTTTCCCAGTCACGACGTGAGTGGCCAACAGAAGGAG
hTRAV24	CCAGGGTTTTCCCAGTCACGACGTGAGCAGCATACTGAACGTGG
hTRAV25	CCAGGGTTTTCCCAGTCACGACGGTGAATGGACAACAGGTAATGC
hTRAV26-1	CCAGGGTTTTCCCAGTCACGACGATGCTAAGACCACCCAGCC
hTRAV26-2	CCAGGGTTTTCCCAGTCACGACGTGATGCTAAGACCACACAGCC
hTRAV27	CCAGGGTTTTCCCAGTCACGACGGTGAGCACCCAGCTGCTG
hTRAV29DV5	CCAGGGTTTTCCCAGTCACGACGTCAACAGAAGAATGATGACCAGC
hTRAV30	CCAGGGTTTTCCCAGTCACGACGAGAAGCCAACAACCAGTGCAG
hTRAV34	CCAGGGTTTTCCCAGTCACGACCAGTAGCCAAGAACTGGAGCAG
hTRAV35	CCAGGGTTTTCCCAGTCACGACGACATGGGTCAGTGGTCAACAG
hTRAV36DV7	CCAGGGTTTTCCCAGTCACGACGTGAGCAGTGAAGACAAGGTGG
hTRAV38-1,2	CCAGGGTTTTCCCAGTCACGACGGCYCAGACAGTCACTCAGTCTC
hTRAV39	CCAGGGTTTTCCCAGTCACGACGGTTAAGTGGAGAGCTGAAAGTGG
hTRAV40	CCAGGGTTTTCCCAGTCACGACCCAGCAGCAATTCAGTCAAGC
hTRAV41	CCAGGGTTTTCCCAGTCACGACCAGCTAAGCTGTGTAAGTGCCG

First PCR reverse primer

$TRAC_{217-240}$	GCTGTTGTTGAAGGCGTTTGCAC

Second PCR forward primer

Linker	CCAGGGTTTTCCCAGTCACGAC

Second PCR reverse primer

$TRAC_{13-33}$	CTGGTACACGGCAGGGTCAG

Supplementary Table 6: PCR primers for generation of TCR beta amplicons. The common linker sequence is highlighted in red.

First PCR forward primers

hTRBV2	CCAGGGTTTTCCCAGTCACGACCACAGAACCTGAAGTCACCCAG
hTRBV3	CCAGGGTTTTCCCAGTCACGACGGACACAGCTGTTTCCCAGACTC
hTRBV4-1,2,3	CCAGGGTTTTCCCAGTCACGACACACCTGGTCATGGGAATGAC
hTRBV5-1,4,5,6,8	CCAGGGTTTTCCCAGTCACGACCTGATCAAAACGAGAGGACAGC
hTRBV6-1,2,3,5,6,8,9	CCAGGGTTTTCCCAGTCACGACGAATGCTGGTGTCACTCAGACCC
hTRBV6-4	CCAGGGTTTTCCCAGTCACGACGACTGCTGGGATCACCCAGG
hTRBV7-2,3	CCAGGGTTTTCCCAGTCACGACCCAGTAACAAGGTCACAGAGAAGG
hTRBV7-4,6,7,8	CCAGGGTTTTCCCAGTCACGACGGTGCTGGAGTCTCCCAGTC
hTRBV7-9	CCAGGGTTTTCCCAGTCACGACGATACTGGAGTCTCCCAGGACCC
hTRBV9	CCAGGGTTTTCCCAGTCACGACGGATTCTGGAGTCACACAAACCC
hTRBV10	CCAGGGTTTTCCCAGTCACGACGGATGCTGRAATCACCCAGAGC
hTRBV11-1	CCAGGGTTTTCCCAGTCACGACGAAGCTGAAGTTGCCCAGTCCC
hTRBV11-2	CCAGGGTTTTCCCAGTCACGACGAAGCTGGAGTTGCCCAGTCTCC
hTRBV11-3	CCAGGGTTTTCCCAGTCACGACGAAGCTGGAGTGGTTCAGTCTCC
hTRBV12-3,4,5	CCAGGGTTTTCCCAGTCACGACGGTGACAGAGATGGGACAAGAAG
hTRBV13	CCAGGGTTTTCCCAGTCACGACGGCTGCTGGAGTCATCCAGTC
hTRBV14	CCAGGGTTTTCCCAGTCACGACGAAGCTGGAGTTACTCAGTTCCC

hTRBV15	CCAGGGTTTTCCCAGTCACGACGGATGCCATGGTCATCCAG
hTRBV16	CCAGGGTTTTCCCAGTCACGACGGTGAAGAAGTCGCCCAGACTC
hTRBV18	CCAGGGTTTTCCCAGTCACGACGGCGTCATGCAGAACCCAAG
hTRBV19	CCAGGGTTTTCCCAGTCACGACGGATGGTGGAATCACTCAGTCC
hTRBV20	CCAGGGTTTTCCCAGTCACGACGCTTGGTGCTGTCGTCTCTCAAC
hTRBV24	CCAGGGTTTTCCCAGTCACGACGGATGCTGATGTTACCCAGACCC
hTRBV25	CCAGGGTTTTCCCAGTCACGACGGAAGCTGACATCTACCAGACCC
hTRBV27	CCAGGGTTTTCCCAGTCACGACGGAAGCCCAAGTGACCCAGAAC
hTRBV28	CCAGGGTTTTCCCAGTCACGACGCCTCGTAGATGTGAAAGTAACCC
hTRBV29	CCAGGGTTTTCCCAGTCACGACGGACTAGGCTCTGTGTTCAGTGC
hThRBV30	CCAGGGTTTTCCCAGTCACGACCCTGGGCACTTTCTTTGGG

First PCR reverse primer

$TRBC_{115-135}$	CAGCTCAGCTCCACGTGGTC

Second PCR forward primer

Linker	CCAGGGTTTTCCCAGTCACGAC

Second PCR reverse primer

$TRBC_{30-51}$	GATGGCTCAAACACAGCGACC

Supplementary Table 7: Specific PCR primers for amplification and subsequent cloning of *TRA* genes. The sequences encoding parts of the TRAV13-1 signal peptide, used as overlap for Gibson assembly cloning, are highlighted in blue.

Specific forwards primers

TRAV1-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAAAGCCTTGAGCAGCCC
TRAV1-2_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAAAACATTGACCAGCCCA

TRAV2-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAAAGGACCAAGTGTTTCAGCC
TRAV3-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGTCAGTGGCTCAGCCG
TRAV4-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAAAGACCACCCAGCCCATC
TRAV5-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAGAGGATGTGGAGCAGAGTCTT
TRAV6-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAAAAGATAGAACAGAATTCCGAGG
TRAV7-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAGAAAACCAGGTGGAGCACA
TRAV8-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGTCTGTGAGCCAGCATAACC
TRAV8-2_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGTCGGTGACCCAGCTT
TRAV8-3_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGTCAGTGACCCAGCCTG
TRAV8-4_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGTCGGTGACCCAGCTT
TRAV8-6_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGTCTGTGACCCAGCTTGA
TRAV8-7_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGTCGGTGACCCAGCTT
TRAV9-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAGATTCAGTGGTCCAGACAGAAG
TRAV9-2_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAGATTCAGTGACCCAGATGGA
TRAV10-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAAAAAACCAAGTGGAGCAGAGTC
TRAV12-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAACGGAAGGAGGTGGAG
TRAV12-2_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAACAGAAGGAGGTGGAGC
TRAV12-3_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAACAGAAGGAGGTGGAGC
TRAV13-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAGAGAATGTGGAGCAGCATCC
TRAV13-2_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAGAGAGTGTGGGGCTGCAT
TRAV14DV4_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGAAGATAACTCAAACCCAACC
TRAV16-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGAGAGTGACTCAGCCCGA
TRAV17-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAACAGGGAGAAGAGGATCC
TRAV18-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAGACTCGGTTACCCAGACAGAA

TRAV19-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGAAGGTAACTCAAGCGCA
TRAV20-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAGAAGACCAGGTGACGCAG
TRAV21-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAAAACAGGAGGTGACGCAGA
TRAV22-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAATACAAGTGGAGCAGAGTCCTCC
TRAV23DV6_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAACAGAAGGAGAAAAGTGACCA
TRAV24-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAATACTGAACGTGGAACAAAGTCC
TRAV25-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAACAGGTAATGCAAATTCCTCA
TRAV26-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAAAGACCACCCAGCCCAC
TRAV26-2_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAAAGACCACACAGCCAAATTCA
TRAV27-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGCTGCTGGAGCAGAGC
TRAV29DV5_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAACAGAAGAATGATGACCAGC
TRAV30-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAACAACCAGTGCAGAGTCC
TRAV34-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAAGAACTGGAGCAGAGTCCT
TRAV35-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAACAGCTGAATCAGAGTCCTC
TRAV36DV7_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAGAAGACAAGGTGGTACAAAGCC
TRAV38-1,2_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGACAGACAGTCACTCAGTCTCAACC
TRAV39-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAGAGCTGAAAGTGGAACAAAACC
TRAV40-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAAATTCAGTCAAGCAGACGGG
TRAV41-1_sPCR	TCCTGTGGCTGCAGCTGGACTTGGTGAATGGAGCCAAAAATGAAGTGGAGC

Specific reverse primer

$TRAC_{13\text{-}33}$	CTGGTACACGGCAGGGTCAG

Supplementary Table 8: Specific PCR primers for amplification and subsequent cloning of *TRB* genes. The sequences encoding parts of the TRBV5-1 signal peptide, used as overlap for Gibson assembly cloning, are

highlighted in blue. Orange marks sequences of the respective V gene which are missing in the template sequence and were attached during specific PCR.

Specific forwards primers

TRBV2-1_sPCR	AGCAGGCCCAGTAAAGGCTGAACCTGAAGTCACCCAGACTC
TRBV3-1_sPCR	AGCAGGCCCAGTAAAGGCTGCTGTTTCCCAGACTCCAAA
TRBV4-1_sPCR	AGCAGGCCCAGTAAAGGCTGAAGTTACCCAGACACCAAAACACCTGGTCATGG GAATGAC
TRBV4-2,3_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTTACGCAGACACCAAGACACCTGGTCATGG GAATGAC
TRBV5-1_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCACTCAAACTCCAAGATATCTGATCAAAAC GAGAGGACAGC
TRBV5-(3-7)_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCACCCAAAGTCCCACACACCTGATCAAAAC GAGAGGACAGC
TRBV5-8_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCACACAAAGTCCCACACACCTGATCAAAAC GAGAGGACAGC
TRBV6-(1-3),(5-9)_sPCR	AGCAGGCCCAGTAAAGGCTGGTGTCACTCAGACCCCAA
TRBV6-4_sPCR	AGCAGGCCCAGTAAAGGCTGGGATCACCCAGGCACCA
TRBV7-1_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCTCCCAGTCCCTGA
TRBV7-2_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCTCCCAGTCCCCCAGTAACAAGGTCACAG AGAAGG
TRBV7-3_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCTCCCAGACCCCCAGTAACAAGGTCACAG AGAAGG
TRBV7-4_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCTCCCAGTCCCCAA
TRBV7-6,7_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCTCCCAGTCTCCCA
TRBV7-8_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCTCCCAGTCCCCTAG
TRBV7-9_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCTCCCAGGACCCCA
TRBV9-1_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCACACAAACCCCAA

TRBV10-1_sPCR	AGCAGGCCCAGTAAAGGCTGAAATCACCCAGAGCCCAA
TRBV10-2,3_sPCR	AGCAGGCCCAGTAAAGGCTGGAATCACCCAGAGCCCA
TRBV11-1_sPCR	AGCAGGCCCAGTAAAGGCTGAAGTTGCCCAGTCCCCCA
TRBV11-2_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTTGCCCAGTCTCCCA
TRBV11-3_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTGGTTCAGTCTCCCAG
TRBV12-3_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTTATCCAGTCACCCCGCCATGAGGTGACAG AGATGGGACAAGA
TRBV12-4_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTTATCCAGTCACCCCGGCACGAGGTGACAG AGATGGGACAAGA
TRBV12-5_sPCR	AGCAGGCCCAGTAAAGGCTAGAGTCACCCAGACACCAAGGGACAAGGTGACAG AGATGGGACAAGA
TRBV13-1_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTCATCCAGTCCCCAA
TRBV14-1_sPCR	AGCAGGCCCAGTAAAGGCTGGAGTTACTCAGTTCCCCAGC
TRBV15-1_sPCR	AGCAGGCCCAGTAAAGGCTATGGTCATCCAGAACCCAAG
TRBV16-1_sPCR	AGCAGGCCCAGTAAAGGCTGAAGAAGTCGCCCAGACTCC
TRBV18-1_sPCR	AGCAGGCCCAGTAAAGGCTGGCGTCATGCAGAACCCA
TRBV19-1_sPCR	AGCAGGCCCAGTAAAGGCTGGAATCACTCAGTCCCCAAA
TRBV20-1_sPCR	AGCAGGCCCAGTAAAGGCTGCTGTCGTCTCTCAACATCCG
TRBV24-1_sPCR	AGCAGGCCCAGTAAAGGCTGATGTTACCCAGACCCCAAGG
TRBV25-1_sPCR	AGCAGGCCCAGTAAAGGCTGACATCTACCAGACCCCAAGATACC
TRBV27-1_sPCR	AGCAGGCCCAGTAAAGGCTCAAGTGACCCAGAACCCAAG
TRBV28-1_sPCR_S	AGCAGGCCCAGTAAAGGCTAAAGTAACCCAGAGCTCGAGATATC
TRBV29-1_sPCR	AGCAGGCCCAGTAAAGGCTGCTGTCATCTCTCAAAAGCCA
TRBV30-1_sPCR	AGCAGGCCCAGTAAAGGCTCAGACTATTCATCAATGGCCAGC

Specific reverse primer	
$TRBC_{30-51}$	GATGGCTCAAACACAGCGACC

Supplementary Table 9: Primers used for colony PCR of cloned *TR* genes.

Forward primer	
colP_Psi_in_F	CAAGCCCTTTGTACACCCT
Reverse primer	
$TRBC_{115-135}$	CAGCTCAGCTCCACGTGGTC

14.4. Supplementary Material

14.4.1. Antibodies

Mouse anti-human HLA-DR (unconjugated, clone L243)	BioLegend GmbH, Fell, Germany
Mouse anti-human HLA-DQ (unconjugated, clone SPVL3)	Abeomics, San Diego, USA
Mouse anti-human HLA-DP (unconjugated, clone B7/21)	Leico Technologies, St. Louis, USA
Mouse anti-human HLA-A (unconjugated, clone 108-2C5)	Abeomics, San Diego, USA
Mouse anti-human HLA-B (unconjugated, clone JOAN-1)	Abeomics, San Diego, USA
Mouse anti-human HLA-C (unconjugated, clone DT-9)	BioLegend GmbH, Fell, Germany
Mouse anti-human HLA-E (unconjugated, 3D12HLA-E)	Invitrogen GmbH, Darmstadt, Germany

14.4.2. FACS analysis antibodies

CellTrace™ Violet Cell Proliferation Kit	Invitrogen GmbH, Darmstadt, Germany
7-AAD	Invitrogen GmbH, Darmstadt, Germany
LIVE/DEAD™ Fixable Near-IR Dead Cell Stain Kit	Invitrogen GmbH, Darmstadt, Germany
Mouse anti-human CD3 (FITC-conjugated, clone OKT3)	BioLegend GmbH, Fell, Germany
Rat anti-human CD4 (APC-Cy7-conjugated , clone A161A1)	BioLegend GmbH, Fell, Germany
Mouse anti-human CD8a (AF700-conjugated, clone SK1)	BioLegend GmbH, Fell, Germany
Mouse anti-human CD19 (PE-Cy7-conjugated, clone SJ25C1)	BD Biosciences GmbH, Heidelberg, Germany
Mouse anti-human CD45RA (BV510-conjugated, clone HI100)	BioLegend GmbH, Fell, Germany
Mouse anti-human CD45RA (PE-conjugated, clone HI100)	BioLegend GmbH, Fell, Germany
Rat anti-human CD185/CXCR5 (AF647-conjugated, clone RF8B2)	BD Biosciences GmbH, Heidelberg, Germany
Mouse anti-human CD196/CCR6 (PE-conjugated, clone G034E3)	BioLegend GmbH, Fell, Germany
Mouse anti-human CD197/CCR7 (BV711-conjugated, clone G043H7)	BioLegend GmbH, Fell, Germany
Hamster anti-human CD278/ICOS (PE-Cy7-conjugated, clone C398.4A)	BioLegend GmbH, Fell, Germany
Mouse anti-human CD279/PD-1 (BV605-conjugated, clone EH12.2H7)	BioLegend GmbH, Fell, Germany
Mouse anti-human TCR alpha/beta (PE-Cy7-conjugated, clone IL26)	BioLegend GmbH, Fell, Germany

Mouse anti-human CD137/4-1BB (BV786 conjugated, clone C65-485)	BD Biosciences GmbH, Heidelberg, Germany
Mouse anti-human CD69 (Alexa647 conjugated, clone FN50)	BioLegend GmbH, Fell, Germany
Mouse anti-human CD134/OX40 (BV421 conjugated, clone ACT-35)	BioLegend GmbH, Fell, Germany
Mouse anti-human CD25 (PE conjugated, clone BC96)	BioLegend GmbH, Fell, Germany
Mouse anti-human ß-2-microglobulin (PE conjugated, clone BBM.1)	Santa Cruz Biotechnology, Dallas, USA

14.4.3. Peptide pool

CSP1-10 (MMRKLAILSV)	PSL GmbH, Heidelberg, Germany
CSP4-72 (SLKKNSRSL)	PSL GmbH, Heidelberg, Germany
CSP7-16 (ILSVSSFLFV)	PSL GmbH, Heidelberg, Germany
CSP86-94 (LRKPKHKKL)	PSL GmbH, Heidelberg, Germany
CSP285-293 (MPNDPNRNV)	PSL GmbH, Heidelberg, Germany
CSP310-319 (EPSDKHIKEY	PSL GmbH, Heidelberg, Germany
CSP319-327 (YLNKIQNSL)	PSL GmbH, Heidelberg, Germany
CSP336-345 (VTCGNGIQVR)	PSL GmbH, Heidelberg, Germany
CSP353-360 (KPKDELDY)	PSL GmbH, Heidelberg, Germany
CSP386-394 (GLIMVLSFL)	PSL GmbH, Heidelberg, Germany

CSP region	Sequence	Position	Expansion pools	Re-stimulation pools
N-term	MMRKLAILSVSSFLF	1-15		Pool 1
	LAILSVSSFLFVEAL	5-19		Pool 1
	SVSSFLFVEALFQEY	9-23		Pool 1
	FLFVEALFQEYQCYG	13-27		Pool 1
	EALFQEYQCYGSSSN	17-31		Pool 1
	QEYQCYGSSSNTRVL	21-35	Mix 1	Pool 1
	CYGSSSNTRVLNELN	25-39	Mix 1	Pool 1
	SSNTRVLNELNYDNA	29-43	Mix 1	Pool 1
	RVLNELNYDNAGTNL	33-47	Mix 1	Pool 1
	ELNYDNAGTNLYNEL	37-51	Mix 1	Pool 1
	DNAGTNLYNELEMNY	41-55	Mix 1	Pool 2
	TNLYNELEMNYYGKQ	45-59	Mix 1	Pool 2
	NELEMNYYGKQENWY	49-63	Mix 1	Pool 2
	MNYYGKQENWYSLKK	53-67	Mix 1	Pool 2
	GKQENWYSLKKNSRS	57-71	Mix 1	Pool 2
	ENWYSLKKNSRSLGE	61-75	Mix 1	Pool 2
	SLKKNSRSLGENDDG	65-79	Mix 1	Pool 2
	NSRSLGENDDGNNED	69-83	Mix 1	Pool 2
	LGENDDGNNEDNEKL	73-87	Mix 1	Pool 2
	DDGNNEDNEKLRKPK	77-91	Mix 1	Pool 2
	NEDNEKLRKPKHKKL	81-95	Mix 1	Pool 3
	EKLRKPKHKKLKQPA	85-99	Mix 1	Pool 3
	KPKHKKLKQPADGNP	89-103	Mix 1	Pool 3

	KKLKQPADGNPDPNA	93-107	Mix 1	Pool 3
	QPADGNPDPNANPNV	97-111	Mix 1	Pool 3
	GNPDPNANPNVDPNA	101-115	Mix 1	Pool 3
Junction +	PNANPNVDPNANPNV	105-119	Mix 1	Pool 3
Repeat	PNVDPNANPNVDPNA	109-123	Mix 1	Pool 3
	NANPNANPNANPNAN	257-271	Mix 1	Pool 3
C-term	NPNANPNANPNKNNQ	263-277	Mix 1	Pool 4
	NPNANPNKNNQGNGQ	267-281	Mix 1	Pool 4
	NPNKNNQGNGQGHNM	271-285	Mix 2	Pool 4
	NNQGNGQGHNMPNDP	275-289	Mix 2	Pool 4
	NGQGHNMPNDPNRNV	279-293	Mix 2	Pool 4
	HNMPNDPNRNVDENA	283-297	Mix 2	Pool 4
	NDPNRNVDENANANS	287-301	Mix 2	Pool 4
	RNVDENANANSAVKN	291-305	Mix 2	Pool 4
	ENANANSAVKNNNNE	295-309	Mix 2	Pool 4
	ANSAVKNNNNEEPSD	299-313	Mix 2	Pool 4
	VKNNNNEEPSDKHIK	303-317	Mix 2	Pool 5
	NNEEPSDKHIKEYLN	307-321	Mix 2	Pool 5
	PSDKHIKEYLNKIQN	311-325	Mix 2	Pool 5
	HIKEYLNKIQNSLST	315-329	Mix 2	Pool 5
	YLNKIQNSLSTEWSP	319-333	Mix 2	Pool 5
	IQNSLSTEWSPCSVT	323-337	Mix 2	Pool 5
	LSTEWSPCSVTCGNG	327-341	Mix 2	Pool 5
	WSPCSVTCGNGIQVR	331-345	Mix 2	Pool 5

SVTCGNGIQVRIKPG	335-349	Mix 2	Pool 5
GNGIQVRIKPGSANK	339-353	Mix 2	Pool 5
QVRIKPGSANKPKDE	343-357	Mix 2	Pool 6
KPGSANKPKDELDYA	347-361	Mix 2	Pool 6
ANKPKDELDYANDIE	351-365	Mix 2	Pool 6
KDELDYANDIEKKIC	355-369	Mix 2	Pool 6
DYANDIEKKICKMEK	359-373	Mix 2	Pool 6
DIEKKICKMEKCSSV	363-377	Mix 2	Pool 6
KICKMEKCSSVFNVV	367-381	Mix 2	Pool 6
MEKCSSVFNVVNSSI	371-385	Mix 2	Pool 6
SSVFNVVNSSIGLIM	375-389		Pool 6
NVVNSSIGLIMVLSF	379-393		Pool 6
SSIGLIMVLSFLFLN	383-397		Pool 6

14.4.4. Bacteria

E.coli DH10B	Clonetech Inc. Palo Alto, CA, USA

14.4.5. Buffers, solutions and chemicals

2x HBS (pH 7.05)	50 mM HEPES 280 mM sodium chloride 0.78 mM disodium phosphate
Dasatinib >98% HPLC	
Bovine serum albumin fraction V (BSA)	Carl Roth GmbH & Co. KG, Karlsruhe, Germany

Calcium chloride dihydrate (CaCl2 x 2 H2O)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Chloroquine diphosphate salt	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Dimethyl sulfoxide (DMSO)	Genaxxon bioscience GmbH, Ulm, Germany
Disodium phosphate (Na2HPO4)	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Ethanol	VWR International, Leicestershire, UK
Gel Loading Dye, Purple (6X)	New England Biolabs GmbH, Frankfurt am Main, Germany
GIBCO™ 10x PBS (pH 7.4)	Life Technologies GmbH, Darmstadt, Germany
GIBCO™ 1x PBS (pH 7.4)	Life Technologies GmbH, Darmstadt, Germany
GIBCO™ Trypan Blue Stain 0.4%	Life Technologies GmbH, Darmstadt, Germany
HEPES	Sigma Aldrich Chemie GmbH, Steinheim, Germany
PegGREEN	PEQLAB Biotechnologie GmbH, Erlangen, Germany
Protamine sulfate	MP Biomedicals, Santa Ana, California, USA
Puromycin dihydrochloride	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Resiquimod (R848)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
SeaKem® LE Agarose	Lonza, Rockland, ME, USA
Sodium chloride (NaCl)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Streptavidin-coated beads	Spherotech, Lake Forest, USA
SybrSafe DNA gel stain	Life Technologies GmbH, Darmstadt, Germany
TAE Buffer (50x)	PanReac AppliChem, Darmstadt, Germany

Tween® 20 (C58H114O26)	Carl Roth GmbH & Co. KG, Karlsruhe, Germany

14.4.6. Cell lines

Jurkat76	Kind gift from Dr. Mirjam Heemskerk, LUMC, Leiden, Netherlands
Phoenix Ampho cells	Kind gift from Florian Salopiata, DKFZ, Heidelberg, Germany

14.4.7. Cell culture media

AIM-V Medium	Life Technologies GmbH, Darmstadt, Germany
GIBCO™ DMEM GlutaMAXTM	Life Technologies GmbH, Darmstadt, Germany
GIBCO™ FBS	Life Technologies GmbH, Darmstadt, Germany
GIBCO™ L-glutamine	Life Technologies GmbH, Darmstadt, Germany
GIBCO™ MEM Non-Essential Amino Acids Solution	Life Technologies GmbH, Darmstadt, Germany
GIBCO™ PenStrep	Life Technologies GmbH, Darmstadt, Germany
GIBCO™ RPMI	Life Technologies GmbH, Darmstadt, Germany
GIBCO™ Sodium Pyruvate	Life Technologies GmbH, Darmstadt, Germany
GIBCO™ Trypsin EDTA 1x	Life Technologies GmbH, Darmstadt, Germany
GIBCO™ Fetal Bovine serum, certified, USA	Life Technologies GmbH, Darmstadt, Germany
Human recombinant IL-2	STEMCELL Technologies, Köln, Germany
Mouse anti human CD28 (clone CD28.2)	BD Biosciences GmbH, Heidelberg, Germany

14.4.8. Commercial kits

DNeasy Blood & Tissue Kit	Qiagen AG, Hilden, Germany
ELISA MAX™ Deluxe Set Human IL-2	BioLegend GmbH, Fell, Germany
High Sensitivity DNA Kit	Agilent, Santa Clara, United States
KAPA Library Quant Kit	Roche Diagnostics GmbH, Mannheim, Germany
NEBuilder HIFI DNA Assembly	New England Biolabs GmbH, Frankfurt am Main, Germany
NucleoBond® Xtra Midi / Maxi	Macherey-Nagel GmbH & Co. KG, Düren, Germany
NucleoSpin® Gel and PCR Clean-up	Macherey-Nagel GmbH & Co. KG, Düren, Germany
NucleoSpin® Plasmid Kit	Macherey-Nagel GmbH & Co. KG, Düren, Germany
Quant-iT PicoGreen ds DNA Kit	Life Technologies GmbH, Darmstadt, Germany
TruSeq Nano DNA LT	Illumina, San Diego, CA, USA

14.4.9. Enzymes and additives

BsrGI-HF, MfeI-HF	New England Biolabs GmbH, Frankfurt am Main, Germany
Calf Intestinal Alkaline Phosphatase (CIP)	New England Biolabs GmbH, Frankfurt am Main, Germany
Cutsmart buffer 10x	New England Biolabs GmbH, Frankfurt am Main, Germany
NEBuffer 2.1	New England Biolabs GmbH, Frankfurt am Main, Germany

T4 DNA Ligase	New England Biolabs GmbH, Frankfurt am Main, Germany
T4 DNA Ligase buffer 10x	New England Biolabs GmbH, Frankfurt am Main, Germany

14.4.10. RT and PCRs

10x PCR buffer	Qiagen AG, Hilden, Germany
5x First strand buffer (RT)	Life Technologies GmbH, Darmstadt, Germany
5X Phusion HF Buffer	New England Biolabs GmbH, Frankfurt am Main, Germany
DTT	Life Technologies GmbH, Darmstadt, Germany
Hotstart *Taq* DNA polymerase	Qiagen GmbH, Hilden, Germany
Magnesium chloride ($MgCl_2$)	Qiagen GmbH, Hilden, Germany
NP-40	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Nuclease free water	Eppendorf AG, Hamburg, Germany
Phusion® High-Fidelity DNA Polymerase	New England Biolabs GmbH, Frankfurt am Main, Germany
RNAsin®	Promega Inc., Madison, WI, USA
SuperScript™ IV Reverse Transcriptase	Life Technologies GmbH, Darmstadt, Germany

14.4.11. Nucleotides and nucleic acids

100 bp DNA Ladder	New England Biolabs GmbH, Frankfurt am Main, Germany

1kb Plus DNA Ladder	New England Biolabs GmbH, Frankfurt am Main, Germany
Desoxynucleotide Triphosphates (dNTPs)	Life Technologies GmbH, Darmstadt, Germany
Oligonucleotides	Eurofins Genomics Germany GmbH, Ebersberg, Germany
Random Hexamer Primers	Roche Diagnostics GmbH, Mannheim, Germany

14.4.12. Instruments and consumables

1.5 ml reaction tubes	Sarstedt AG, Nümbrecht, Germany
2 ml reaction tubes	Sarstedt AG, Nümbrecht, Germany
5 ml tubes	Sarstedt AG, Nümbrecht, Germany
96-well semi-skirted PCR plate	4titude, Surrey, UK
Adhesive aluminum foil seal	4titude, Surrey, UK
BD FACSAria™ III	BD Biosciences GmbH, Heidelberg, Germany
2100 Bioanalyzer Instrument	Agilent, Santa Clara, USA
Cell culture 6-well plate	Greiner Bio-One GmbH, Frickenhausen, Germany
Cell culture 24-well plate	Corning Inc., Corning, NY, USA
Cell culture 96-well plate, U-bottom	TPP, Trasadingen, Switzerland
CellStar sterile serological pipettes 2 ml, 5 ml, 10 ml, 25 ml, 50 ml	Greiner Bio-One GmbH, Frickenhausen, Germany
Centrifuge 5427 R (rotor FA-45-30-11)	Eppendorf
Centrifuge mikro 200R (rotor 2424-B)	Andreas Hettich GmbH & Co.KG, Tuttlingen, Germany

Centrifuge rotina 420R (rotor 4784-A)	Andreas Hettich GmbH & Co.KG, Tuttlingen, Germany
Centrifuge universal 320R (rotor 1689-A)	Andreas Hettich GmbH & Co.KG, Tuttlingen, Germany
Centrifuge 5180R (rotor 5-4-104, rotor A-2-DWPAT)	Eppendorf AG, Hamburg, Germany
CO_2 Incubator CB210	Binder GmbH, Tüttlingen, Germany
CoolCell LX / FTS30	Corning Inc., Corning, NY, USA
CountessTM II FL Automated Cell Counter	Life Technologies GmbH, Darmstadt, Germany
CountessTM Cell Counting Chamber Slides	Life Technologies GmbH, Darmstadt, Germany
CryoTube™ Vials, 1.8 ml	Thermo Fisher Scientific Inc., Darmstadt, Germany
EC 300 XL power supply	Thermo Scientific Inc., Rochester, NY, USA
ELISA plates (96-well, flat bottom)	Costar Inc., Corning, Action, MA, USA
FrameStar® 384	4titude, Surrey, UK Freedom
Evo200 automation platform	Tecan, Crailsheim, Germany
Gel DocTM XR+	Bio-Rad Laboratories GmbH, München, Germany
Luna Automated Cell Counter	Logos Biosystems Inc., Anyang, Korea Luna
Vell Counting Slides	Logos Biosystems Inc., Anyang, Korea
M1000Pro plate reader	Tecan, Crailsheim, Germany
Mastercycler® ep Gradient S	Eppendorf AG, Hamburg, Germany
Mastercycler® Pro 384	Eppendorf AG, Hamburg, Germany
Microscope Primovert	ZEISS Microscopy, Oberkochen, Germany
Model 680 Microplate Reader	Bio-Rad Laboratories GmbH, München, Germany

Multichannel pipets Eppendorf Research® plus	Eppendorf AG, Hamburg, Germany
0.5-10 µl, 10-100 µl, 30-300 µl, 120-1200 µl Multipette plus	Eppendorf AG, Hamburg, Germany
Multiply - µStrip Pro 8-strip (PCR tube strip)	Sarstedt AG, Nümbrecht, Germany
Multitron Pro (Bacteria shaker)	Infors HT, Bottmingen, CH
NanoQuant plateTM	Tecan, Crailsheim, Germany
Neubauer Counting Chamber by Marienfeld	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
OWL A3-1, OWL B3, OWL Easycast B1A	Thermo Scientific Inc., Rochester, NY, USA
Thinkpad Laptop	Lenovo
PCR foil seal	4titude, Surrey, UK Petri dishes (100 mm) Greiner Bio-One GmbH, Frickenhausen, Germany
Pipetboy acu 2	Integra Biosciences GmbH, Fernwald, Germany
Gilson™ PIPETMAN Classic™ 0.1-2 µl, 2-20 µl, 20- 200 µl, 100-1000 µl	Gilson S.A.S., Villiers-le-Bel, France
Polyolefin sealing foil for microtiter plate	HJ-Bioanalytik GmbH, Erkelenz, Germany
Polypropylene tubes (15 ml, 50 ml)	Sarstedt AG, Nümbrecht, Germany
Polystyrene round bottom tube (5 ml) with cell strainer cap	BD Biosciences GmbH, Heidelberg, Germany
T25 / T75 / T150 cm2 flask	TPP, Trasadingen, Switzerland
Thermomixer comfort	Eppendorf AG, Hamburg, Germany
Vortex genie 2	Scientific Industries Inc., Bohemia, NY, USA
Water bath with thermostat	Memmert GmbH + Co. KG, Schwabach, Germany

14.4.13. Software

Clone Manager Professional 9	Scientific & Educational Software, Westminster, USA
FlowJo v10.7.1	Treestar Systems Inc., Ashland, USA
GraphPad Prism Version 9.3.1	GraphPad Software Inc., La Jolla, USA
Affinity Designer 2	Serif (Europe) Ltd.
Microsoft® Office 2019	Microsoft GmbH, Stuttgart, Germany
EndNote	Version 21
R	Version 3.4.2
Rstudio version 1.3.1093	RStudio, Inc., Boston, MA US
Cytoscape	Version 3.9.1
DBeaver	Version 7.3.0

14.4.14. Web Resources

Clustal Omega	http://www.ebi.ac.uk/Tools/msa/clustalo/
Ensemble Genome Browser	http://www.ensembl.org/index.html
Lab collector	https://labcollector.com/
IMGT®	http://www.imgt.org/
NCBI Ig Blast	http://www.ncbi.nlm.nih.gov/igblast/
PlasmoDB	http://plasmodb.org/plasmo/
SignalP-4.1 Server	http://www.cbs.dtu.dk/services/SignalP/
VDJdb	https://vdjdb.cdr3.net/search
NetMHC pan 4.1	https://services.healthtech.dtu.dk/service.php?NetMHCpan-4.1

15. Acknowledgement

First of all, I would like to thank Prof. Dr. Hedda Wardemann for giving me the opportunity to perform my PhD work in her lab. Her scientific guidance and numerous productive discussions on scientific data supported me throughout the years to grow as a scientist. I gratefully appreciate my first examiner, Prof. Dr. Nina Papavasiliou, and her scientific input during annual thesis advisory committee meetings. Further I would like to thank Prof. Dr. Friedrich Frischknecht and Prof. Dr. Alexander Dalpke for reviewing my thesis and being part of my thesis examination committee.

The work presented in this study would not have been possible without Dr. Sheetij Dutta and the clinical team at the Walter Reed Army Institute as well as the clinical trial volunteers. Thank you for providing the samples and taking part in this collaborative study!

I would like to express my appreciation to Dr. Ilka Wahl, who supported me throughout the years, both personally and scientifically. Special thanks are extended to all members of the B cell immunology group. I appreciate your critical and motivating input during lab meetings, lunch and coffee breaks. Also, I want to thank Sara Kraker and Christine Niesik for helping me with numerous experiments. Importantly, I would like to thank Hendrik Feuerstein and Dr. Ilka Wahl for their helpful input on the thesis draft.

I thank Dr. Frank Momburg for helpful discussions on HLA-E peptide binding predictions and biochemical mechanisms in HLA-E peptide presentation.

I am sincerely thankful to my parents, Iris and Peter, for always supporting me throughout the years and my brother, Niklas, for his input and assistance during the thesis writing phase. My profound thanks go to Constantin for his support and encouragement during the last years, especially during stressful times.