THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

The adaptive immune response to Epstein-Barr virus

Melissa Jane Rist (nee Bell)

B. App. Sc. (Hons)

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2015

School of Medicine

Abstract

Major histocompatibility complex (MHC) class I molecules form medleys with peptide antigens which are expressed on the cell surface for recognition by $CD8^+$ T cells. Derived from antigens synthesized in the cytoplasm, these peptides are generally 8-10 amino acids in length. For most MHC alleles two of the pockets within the peptide binding groove display a marked preference for one or two amino acids at certain anchor positions within the peptide. This was the breakthrough discovery that enabled more efficient CTL epitope mapping. Dependent on this information, web-based algorithms used to predict $CD8^+$ T cell epitopes were designed to include peptides limited to between 8 and 10 residues. The apparent dominance of MHC class I peptides of 8 to 10 amino acids in length may be misleading and result from this bias of widely used algorithms. Several studies have clearly shown that some longer peptides are naturally processed and presented for recognition by $CD8^+$ T cells. A number of these noncanonical epitopes have been identified in Epstein-Barr virus (EBV). In order to ascertain the role of long peptides in $CD8^+$ T cell responses, studies were required which utilise overlapping peptide screening and not web-based algorithms to predict peptides. The latent herpes virus Epstein-Barr virus is a superb model to determine the relative prevalence of noncanonical T cell epitopes.

A number of highly immunogenic antigens of EBV were the focus of this study. Many novel class I EBV epitopes were identified with a significant proportion arising from the BZLF1 antigen. Overall, 28 latent epitopes from EBNA3A, EBNA3B, EBNA3C and EBNA1 and 13 lytic epitopes from BRLF1, BMLF1 and BMRF1 were identified in addition to novel epitopes from BZLF1, including five noncanonical $CD8^+$ T cell epitopes. The EBV BZLF1 protein demonstrated a propensity for epitope clustering. In total, I have defined 13 novel epitopes from this EBV protein which are restricted to a number of HLA alleles and appear to be clustered and in many instances overlapping. These included two overlapping epitopes of different length that nevertheless conform to the binding motif of the large and abundant HLA-B*44 supertype. While HLA-$B*18:01^+$ individuals responded strongly and exclusively to an octamer peptide ^{173}SELEIKRY^{180}, HLA-$B*44:03^+$ individuals responded to the atypically large dodecamer peptide ^{169}EECDSELEIKRY^{180} that encompasses the octamer peptide. Moreover, the octamer peptide bound more stably to HLA-B*18:01 than the dodecamer peptide while conversely, HLA-B*44:03 bound only the longer peptide. Furthermore, crystal structures of these viral peptide-HLA complexes showed that the antigen-binding cleft of HLA-B*18:01 was more ideally suited to bind shorter peptides, while HLA-B*44:03 exhibited characteristics that favoured the presentation of longer peptides. Mass spectrometric identification of over a thousand naturally-presented ligands revealed that HLA-B*18:01 was more biased towards presenting shorter peptides than HLA-B*44:03. Collectively,

these data highlight a mechanism through which polymorphism within an HLA class I supertype can diversify determinant selection and immune responses by varying peptide length preferences.

Interestingly, a human protein sequence (D<u>ELEIK</u>A<u>Y</u>) was identified with sequence homology to the octamer EBV epitope. The peptide was shown to bind stably to HLA-B*1801, and peptide elution/mass spectrometric studies showed it is presented by this HLA molecule on the surface of human cells. A significant proportion of T cells raised against the SELEIKRY EBV epitope cross-reacted with this HLA-B*1801-binding self-peptide. Of note, only a limited number of HLA-B*1801^{+} healthy individuals showing strong IFN-γ responses to SELEIKRY had a detectable response to DELEIKAY. These cross-reactive T cells were shown to express a diverse array of T cell receptors. The potential for self-reactivity by these CTLs is presumably kept under rigorous control by normal self-tolerance mechanisms. However, these EBV/self cross-reactive T cells could pose an autoimmune threat if HLA-B*1801-DELEIKAY levels increased or the T cell activation threshold is reduced as a result of cytokine release during inflammation and tissue damage. These cross-reactive T cell populations should be considered for their potential role in self-reactivity following viral infection.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

Original research articles (peer-reviewed)

Rist MJ, Hibbert KM, Croft NP, Smith C, Neller MA, Burrows JM, Miles JJ, Purcell AW, Rossjohn J, Gras S, Burrows SR. T cell cross-reactivity between a highly immunogenic Epstein-Barr virus epitope and a self-peptide naturally presented by HLA-B*18:01^{+} cells. *The Journal of Immunology*. 2015 May; 194: 4668-4675.

Rist MJ, Neller MA, Burrows JM, Burrows SR. T Cell Epitope Clustering in the Highly Immunogenic BZLF1 Antigen of Epstein-Barr virus. *Journal of Virology.* 2015 Jan; 89 (1):703-712.

Rist MJ, Theodossis A, Croft NP, Neller MA, Welland A, Chen Z, Sullivan LC, Burrows JM, Miles JJ, Brennan RM, Gras S, Khanna R, Brooks AG, McCluskey J, Purcell AW, Rossjohn J, Burrows SR. HLA Peptide Length Preferences Control CD8^{+} T Cell Responses. *The Journal of Immunology*. 2013 June; 191 561-571.

Neller MA, Burrows JM, **Rist MJ**, Miles JJ, Burrows SR. High frequency of herpesvirus-specific clonotypes in the human T cell repertoire can remain stable over decades with minimal turnover. *Journal of Virology*. 2013 87 (1) 697-700.

Horst D, Burrows SR, Gatherer D, van Wilgenburg B, **Bell MJ**, Boer IG, Ressing ME, and Wiertz EJ. Epstein-Barr virus isolates retain their capacity to evade T cell immunity through BNLF2a despite extensive sequence variation. *Journal of Virology*. 2012 86 (1):572-7.

Conference abstracts (peer-reviewed)

Rist MJ, Theodossis A, Croft NP, Neller MA, Welland A, Chen Z, Sullivan LC, Burrows JM, Miles JJ, Brennan RM, Gras S, Khanna R, Brooks AG, McCluskey J, Purcell AW, Rossjohn J, Burrows SR. HLA Peptide Length Preferences Control CD8^{+} T Cell Responses.
Australasian Society of Immunology Annual Meeting. December 2013. Wellington, New Zealand. (oral)

Rist MJ, Theodossis A, Croft NP, Neller MA, Welland A, Chen Z, Sullivan LC, Burrows JM, Miles JJ, Brennan RM, Gras S, Khanna R, Brooks AG, McCluskey J, Purcell AW, Rossjohn J, Burrows SR. HLA Peptide Length Preferences Control CD8^{+} T Cell Responses.
Brisbane Immunology Group Annual Meeting. August 2013. Gold Coast, Australia. (oral)

Rist MJ, Theodossis A, Croft NP, Neller MA, Welland A, Chen Z, Sullivan LC, Burrows JM, Miles JJ, Brennan RM, Gras S, Khanna R, Brooks AG, McCluskey J, Purcell AW, Rossjohn J, Burrows SR. HLA Peptide Length Preferences Control CD8^{+} T Cell Responses.
Frontiers in Immunology Research Network Conference. July 2013. Monte Carlo, Monaco. (oral)

Rist MJ, Theodossis A, Croft NP, Neller MA, Welland A, Chen Z, Sullivan LC, Burrows JM, Miles JJ, Brennan RM, Gras S, Khanna R, Brooks AG, McCluskey J, Purcell AW, Rossjohn J, Burrows SR. HLA Peptide Length Preferences Control CD8^{+} T Cell Responses.
Australasian Society of Immunology Annual Meeting. December 2012. Melbourne, Australia. (poster, Three minute mini-oral)

Publications included in this thesis

Chapter 4:

Rist MJ, Neller MA, Burrows JM, Burrows SR. T Cell Epitope Clustering in the Highly Immunogenic BZLF1 Antigen of Epstein-Barr virus. *Journal of Virology.* 2015 Jan; 89 (1):703-712

Contributor	Statement of contribution
Rist MJ (Candidate)	Experiment design (80%) Sample collection and processing (100%) Data acquisition/lab work (90%) Analysis and interpretation of data (80%) Manuscript preparation (80%)
Neller MA	Data acquisition/lab work (5%) Reviewed and edited manuscript
Burrows JM	Data acquisition/lab work (5%) Reviewed and edited manuscript
Burrows SR	Project conception (100%) Experiment design (20%) Analysis and interpretation of data (20%) Manuscript preparation (20%)

Chapter 5:

Rist MJ, Theodossis A, Croft NP, Neller MA, Welland A, Chen Z, Sullivan LC, Burrows JM, Miles JJ, Brennan RM, Gras S, Khanna R, Brooks AG, McCluskey J, Purcell AW, Rossjohn J, Burrows SR. HLA Peptide Length Preferences Control $CD8^+$ T Cell Responses. *The Journal of Immunology*. 2013 June; 191:561-571.

Contributor	Statement of contribution
Rist MJ (Candidate)	Experiment design (30%) Sample collection and processing (60%) Data acquisition/lab work (35%) Analysis and interpretation of data (30%) Manuscript preparation (50%)
Theodossis A	Experiment design (20%) Sample collection and processing (15%) Data acquisition/lab work (25%) Analysis and interpretation of data (15%) Manuscript preparation (15%)
Croft NP	Experiment design (20%)

	Sample collection and processing (15%) Data acquisition/lab work (25%) Analysis and interpretation of data (15%) Manuscript preparation (15%)
Neller MA	Data acquisition/lab work (5%) Reviewed and edited manuscript
Welland A	Data acquisition/lab work (5%) Reviewed and edited manuscript
Chen Z	Sample collection and processing (5%) Reviewed and edited manuscript
Sullivan LC	Made reagents (tetramers) Reviewed and edited manuscript
Burrows JM	Data acquisition/lab work (5%) Reviewed and edited manuscript
Miles JJ	Analysis and interpretation of data (5%) Reviewed and edited manuscript
Brennan RM	Sample collection and processing (5%) Reviewed and edited manuscript
Gras S	Analysis and interpretation of data (10%) Reviewed and edited manuscript
Khanna R	Analysis and interpretation of data (5%) Reviewed and edited manuscript
Brooks AG	Made reagents (tetramers) Reviewed and edited manuscript
McCluskey J	Experiment design (10%) Reviewed and edited manuscript
Purcell AW	Experiment design (10%) Reviewed and edited manuscript
Rossjohn J	Project conception (20%) Reviewed and edited manuscript
Burrows SR	Project conception (80%) Experiment design (10%) Analysis and interpretation of data (20%) Manuscript preparation (20%)

Chapter 6:

Rist MJ, Hibbert KM, Croft NP, Smith C, Neller MA, Burrows JM, Miles JJ, Purcell AW, Rossjohn J, Gras S, Burrows SR. T cell cross-reactivity between a highly immunogenic Epstein-Barr virus epitope and a self-peptide naturally presented by HLA-B*18:01^{+} cells. *The Journal of Immunology*. 2015 May; 194: 4668-4675.

Contributor	Statement of contribution
Rist MJ (Candidate)	Experiment design (40%) Sample collection and processing (95%) Data acquisition/lab work (55%) Analysis and interpretation of data (70%) Manuscript preparation (70%)
Hibbert KM	Experiment design (20%) Sample collection and processing (5%) Data acquisition/lab work (25%) Reviewed and edited manuscript
Croft NP	Experiment design (10%) Data acquisition/lab work (10%) Analysis and interpretation of data (10%) Reviewed and edited manuscript
Smith C	Experiment design (5%) Reviewed and edited manuscript
Neller MA	Data acquisition/lab work (5%) Reviewed and edited manuscript
Burrows JM	Experiment design (5%) Data acquisition/lab work (5%) Reviewed and edited manuscript
Miles JJ	Project conception (10%) Reviewed and edited manuscript
Purcell AW	Project conception (10%) Reviewed and edited manuscript
Rossjohn J	Project conception (20%) Reviewed and edited manuscript
Gras S	Experiment design (10%) Manuscript preparation (10%)
Burrows SR	Project conception (60%) Experiment design (10%)

	Analysis and interpretation of data (20%) Manuscript preparation (20%)

Contributions by others to the thesis

My supervisor, Professor Scott Burrows, contributed significantly to the conception and design of the projects described in this thesis. No other significant contributions were made.

Statement of parts of the thesis submitted to qualify for the award of another degree

None

Acknowledgements

First and foremost, I would like to thank my supervisor, Professor Scott Burrows. His constant support and guidance has been instrumental during my PhD and instilled in me the confidence to grow as a scientist. I genuinely appreciate the opportunities he has given to me and could not imagine a more patient supervisor. He has been an outstanding mentor and will continue to be. I would like to thank my associate supervisor Professor Rajiv Khanna for his support and guidance. Many thanks to Dr Corey Smith who has been a constant source of knowledge and patience and I greatly appreciate his invaluable friendship and advice, Mrs Jacqueline Burrows has been an unwavering support both in and out of the lab and a wealth of molecular knowledge and Dr Viviana Lutzky for her advice, friendship and support. Thank you to Professor Denis Moss for his guidance and advice. Thank you to Cellular Immunology lab member Dr Rebekah Brennan for her advice. I would also like to acknowledge collaborative efforts from Professor Jamie Rossjohn, Dr Stephanie Gras and Professor James McCluskey.

My family and friends have been an invaluable source of encouragement and support. Especially my parents whose love, support and unwavering belief that I am capable of achieving anything is instrumental in all that I do and my husband Michael, for your constant love, support and our two feline distractions.

Finally, I am extremely grateful for the funding sources that supported this project. My PhD scholarship was funded by the National Health and Medical Research Council and travel grant by the Australasian Society of Immunology.

Keywords

T Cell immunity, epitope, Epstein-Barr virus, T cell receptor repertoire, adaptive immunity, sequence polymorphism, major histocompatibilty complex

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 110704, Cellular Immunology, 80%

ANZSRC code: 110309, Infectious Diseases, 10%

ANZSRC code: 110804, Medical Virology, 10%

Fields of Research (FoR) Classification

FoR code: 1107, Immunology, 80%

FoR code: 1103, Clinical Sciences, 10%

FoR code: 1108, Medical Microbiology, 10%

Table of Contents

List of Abbreviations

α	alpha
β	beta
δ	delta
ε	epsilon
γ	gamma
ζ	zeta
APC	antigen presenting cell
APC	allophycocyanin
BCIP/NBT	5'bromo-4 chloro-3 indole and inorganic phosphate/ nitro blue tetrazolium chloride
BL	Burkitt's lymphoma
CCL2	chemokine (C-C motif) ligand 2
cDNA	complementary DNA
CDR	complementarity determining region
CLIP	class II-associated invariant chain peptide
cTEC	cortical thymic epithelial cell
CTL	cytotoxic T lymphocyte
CTLp	$CD8^+$ T cell precursor
DC	dendritic cell
DMSO	dimethyl sulfoxide
DN	double negative
DNA	deoxyribonucleic acid
DP	double positive
EBNA	Epstein-Barr virus nucleic acid
EBV	Epstein-Barr virus
ELISpot	enzyme linked immunospot assay
ER	endoplasmic reticulum
FACS	fluorescence activated cell sorter
FCS	fetal calf serum
FITC	fluorescein isothiocyanate
HC	heavy chain
HCMV	Human Cytomegalovirus
HHV	Human herpes virus
HIV	Human Immunodeficiency Virus
HL	Hodgkin's lymphoma

HLA	human leukocyte antigen
HSV	Herpes Simplex virus
ICS	intracellular cytokine staining
IDDM	Insulin-dependent diabetes mellitus
IFN	interferon
Ig	immunoglobulin
Ii	invariant chain
IL	interleukin
IM	infectious mononucleosis
ITAM	immunoreceptor tyrosine based activation motif
kb	kilobase
KSHV	Kaposi's Sarcoma-associated herpesvirus
LCL	lymhoblastoid cell line
LMP	latent membrane protein
LPS	lipopolysaccharide
mAb	monoclonal antibody
MFI	mean fluorescence intensity
MHC	major histocompatibility complex
mg	milligram
µg	microgram
mM	millimolar
µM	micromolar
MS	Multiple Sclerosis
NK	natural killer cells
NPC	Nasopharyngeal Carcinoma
PAMPs	pathogen-associated molecular patterns
PBMC	peripheral blood mononuclear cell
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PE	phycoerythrin
PerCP	peridinin-chlorophyll proteins
PerCP-Cy5.5	PerCP with a cyanine dye (Cy5.5)
PFA	paraformaldehyde
PHA	phytohemagglutinin
p-HLA	peptide-Human Leukocyte Antigen

pMHC	peptide-MHC complex
PRR	pattern recognition receptors
PTLD	post transplant lymphoproliferative disorder
RAG	recombinase-activating gene
RBC	red blood cell
RNA	ribonucleic acid
rpm	revolutions per minute
RPMI	Roswell Park Memorial Institute medium
RT	room temperature
SIV	simian immunodeficiency virus
SLE	systemic lupus erythematosus
SP	single positive
TAP	transporter associated with antigen processing
TCGF	T cell growth factor
TCM	T cell media
TCR	T cell receptor
TNFα	tumour necrosis factor-alpha
TR	tandom repeat
Vα	variable region of the T cell receptor alpha chain
Vβ	variable region of the T cell receptor beta chain
VZV	Varicella-Zoster virus
WT	wild type

Chapter 1

Introduction and Literature Review

1.1 The Immune System

1.1.1 Introduction

The immune system is an interesting and complex network of biological sentinels that defend the body from foreign invaders. Although, the human body has many levels of defence including skin, mucous membranes and stomach acids it is ultimately the immune system that detects invasion by bacterial, parasitic and viral agents distinguishing and eliminating these organisms without causing damage to self. Traditionally the immune system is divided into two streams, innate and adaptive.

1.1.2 Innate Immunity

It is believed that during evolution the innate immune system appeared prior to the adaptive immune system. Innate immunity is activated immediately following infection to control replication of the infecting pathogen. Mediated by germ-line-encoded receptors, innate immune recognition is genetically predetermined. These receptors have defined specificities for infectious microorganisms and are expressed on the surface of the professional antigen-presenting cells (APC) of the innate immune system including macrophages, dendritic cells (DC) and B cells (Medzhitov and Janeway Jr 2000).

As there is a limit to the number of genes that can be encoded by a host's genome, relatively few germ-line encoded molecules bear the task of detecting a vast array of molecular structures, or patterns associated with infectious microorganisms. This role is fulfilled by pattern- recognition receptors (PRRs) expressed on APCs focused on detecting a selection of highly conserved patterns, pathogen-associated molecular patterns (PAMPs), shared by a large group of pathogens. PAMPs include conserved features of microbial pathogens such as bacterial and fungal cell-wall components and viral nucleic acids (Janeway Jnr 1989). Importantly, none of these structures are produced by the host organism and are therefore distinct from self-antigens presenting the innate immune system with the ability to discriminate between self and non-self (Medzhitov and Janeway Jr 1997).

For example, synthesized only by bacteria, lipopolysaccharides (LPS) are recognized by PRRs on macrophages alerting the host to the presence of an infecting organism. Upon activation macrophages secrete cytokines (e.g. interleukin-1 (IL-1) and tumour necrosis factor-alpha (TNF-α)) and chemokines (e.g. chemokine (C-C motif) ligand 2 (CCL2)) directing monocytes and neutrophils to exit the bloodstream and enter the surrounding tissue at the site of infection, initiating the process of inflammation (Janeway Jnr, Travers et al. 2001). Ultimately, the sensing of microbes by cells of

the innate immune system initiate the adaptive immune response by instructing the direction of the adaptive immune system (Medzhitov and Janeway Jr 1997).

1.1.3 Adaptive Immunity

The adaptive immune system provides long-lasting antimicrobial protection and is mediated by T and B lymphocytes. Formed in the bone marrow, B cells secrete microbial-specific antibodies upon proliferation and activation, whilst following maturation in the thymus, T cells recognize foreign antigens derived from pathogens and protect the host from infection (Janeway Jnr, Travers et al. 2001). These lymphocytes bear antigen receptors generated somatically. Each lymphocyte displays a structurally unique receptor, as these receptors are not predestined and generated randomly this results is an extremely diverse repertoire of receptors.

Lymphocytes recognizing antigens specific for pathogens are subsequently selected for clonal expansion, however, the receptor repertoire contains binding sites that have the capacity to react not only with pathogens and environmental antigens but also with self. This diverse repertoire of antigen receptors of the adaptive immune system is almost infinite as there are approximately 10^{14} and 10^{18} different somatically generated immunoglobulin (Ig) receptors and T cell receptors, respectively (Medzhitov and Janeway Jr 2000).

1.1.4 Introduction to T-cell immunity

There are two distinct subsets of naïve T cells, those expressing the co-receptor CD8 or T cells expressing the co-receptor CD4. These T cells bear the αβ T cell receptor (TCR). Following activation of a naïve $CD4^+$ T cell, it differentiates into one of four further subsets: Th1, Th2, Th17 and regulator T cells (Treg). Once the naïve $CD8^+$ T cells are activated they develop into cytotoxic T cells (CTL). T cells committed to the CD8 lineage depend on the binding of the αβ TCR to major histocompatibility complex (MHC) class I molecules, where as the commitment of the T cells of the CD4 lineage is to MHC class II molecules (Borgulya, Kishi et al. 1991).

In 1974, Rolf Zinkernagel and Peter Doherty published in Nature the biological significance of MHC Class I molecules, showing that TCRs are designed to recognise the presented antigen in addition to the MHC, i.e. CTLs, induced by viral infection recognise cells in a virus specific, class I-restricted manner. (Zinkernagel and Doherty 1974). $CD4^+$ helper T cells contribute to the expansion of $CD8^+$ T cells and provide specific cytokines such as IL-2 as well as assisting B cells in proliferation and differentiation and switching the class of antibody they produce, whilst CTLs kill infected target cells.

1.1.5 Antigen Processing and Presentation

1.1.5a Structure of MHC Class I molecules

A mature MHC I complex consists of three noncovalently associated sections, a heavy chain (45kDa), a light chain (12kDa) also called β_2-microglobulin and a short peptide antigen. In humans, there are three class Ia loci, HLA-A, B and C. Both alleles of the loci are codominantly expressed, indicating that an individual may express up to six different class Ia molecules (Parham, Adams et al. 1995). The heavy chain (HC) consists of a cytoplasmic region, a transmembrane region, and an extracellular region composed of three α domains (α1, α2 and α3) with one to three N-linked glycosylation sites. It is the α3 region of the HC that contains the binding site for the CD8 receptor on CTL. The α1 and α2 domains interact with the TCR on CTL and fold together to form a groove that binds and displays the antigenic peptides. The base of the cleft is formed by a beta-pleated sheet and the walls are comprised of two alpha helices. It is primarily in these residues in and around the cleft of the HC where allelic polymorphisms occur and allow an alteration of peptide-binding specificity of the MHC I molecules (Parham, Adams et al. 1995) (Bjorkman, Saper et al. 1987).

1.1.5b MHC Class I pathway

MHC class I molecules display fragments of intracellular expressed antigens from pathogens or self for monitoring by the immune system. Generally these proteins are in the form of small peptides around 9 amino acids in length (Rock and Goldberg 1999). CTLs can identify peptide-MHC complexes (pMHC) and distinguish between infected and uninfected cells. This presentation pathway is the result of a number of biochemical processes from proteolysis to the transportation of fragments to the plasma membrane. Predominantly it is endogenous proteins that are presented to the MHC I molecule following ubiquitination in the cytosol and delivered to the proteasome to initiate fragmentation. An adenosine triphosphate (ATP)-dependent multisubunit protease, the proteasome is considered the main machinery behind the generation of antigenic peptides (usually 8-10 residues in length) (Rock, Gramm et al. 1994).

Class I peptides are then transported into the lumen of the endoplasmic reticulum (ER) by the transporter associated with antigen presentation (TAP) complex. Peptides are required to be 7 residues or longer for transportation by TAP, with the reduced efficiency with peptides longer than 12 residues (Shepherd, Schumacher et al. 1993). Following transportation to the ER, peptides are loaded onto the MHC class I molecule. This sequence of events involves a partially folded MHC class I α chain binding to calnexin until β2-microglobin (β2:M) binds and the calnexin dissociates. Additional molecules join the complex including the transmembrane glycoprotein, tapasin (which

links TAP to peptide free MHC class I molecules) and the chaperones calreticulin and ERp57 (Wright, Kozik et al. 2004). Within the ER, peptides bind to the MHC class I-β_2m heterodimer that is released from the peptide loading complex. The fully assembled peptide loaded MHC molecules are translocated from the ER through the Golgi apparatus onto the cell surface to be scrutinised by the TCR of $CD8^+$ T cells (Kloetzel 2001, Cresswell, Ackerman et al. 2005).

1.1.5c Cross Presentation

In some instances it is antigens from the extracellular environment that are presented on the cell surface in association with MHC class I molecules and stimulate $CD8^+$ T-cell immunity. This process is referred to as cross-presentation. Now known as a major mechanism of the immune system, cross-presentation is the only pathway the immune system can utilise to detect and advance on viral infection and mutations that occur predominantly in the parenchymal cells. These viral antigens are processed and presented by specialised APCs which acquire the viral antigen from the infected cells, in most cases by phagocytosis and occasionally using fluid-phase endocytosis (Watts and Amigorena 2001). The principal cell which cross presents antigens is the DC (Mellman and Steinman 2001). Internalised antigens may then be processed through one of two known pathways.

The phagosome-to-cytosol pathway and the vacuolar pathway are the two current pathways attributed with cross-presentation. The dominant of these mechanisms is the TAP dependent phagosome-to-cytosol pathway. In this pathway the antigen is internalised into phagosomes or micropinosomes and transferred into the cytosol. Once in the cytosol, proteosomes hydrolyse the antigen into oligopeptides which are transported by TAP to the MHC class I molecules in the ER or "ER-phagosome" vesicles (Kovacsovics-Bankowswki and Rock 1995). The second and less understood mechanism, the vacuolar pathway is TAP independent, however, requires endosomal protease, cathepsin S. Phagosomes internalise antigens to be degraded into oligopeptides, predominantly by cathepsin S. The resulting peptides are then presumably loaded onto the MHC class I molecules that have arrived into the vesicle from the plasma membrane or ER (Rock and Shen 2005).

1.1.5d MHC Class II pathway

MHC Class II processing pathway focuses MHC class II-antigen complex transporting to the cell surface for presentation to $CD4^+$ T cells. Although predominantly associated with peptides derived from exogenous proteins, studies have shown that endogenous proteins too come into play. MHC class II proteins are expressed on professional APCs including DCs, B cells and macrophages, also cortical and medullary thymic epithelial cells have been shown to be involved. There are two MHC class II pathways, the classical, MHC Class II-mediated exogenous antigen

processing pathway and the less understood, non-classical, cytoplasmic antigen presentation pathway.

In the exogenous pathway, exogenous proteins are degraded by acidic proteases located within endosomes and lysosomes, MHC class II molecules are assembled in the ER with the chaperone invariant chain (Ii). The Ii targets the Ii-MHC class II complex to the endosomal pathway, where once entered into the endosomal/lyosomal compartments, Ii is degraded into fragments, one of these small fragments, class II–associated invariant chain peptide (CLIP) is retained and involved with the MHC class II binding groove. An MHC-encoded molecule human leukocyte antigen HLA-DM is required for the removal of CLIP and capture of antigenic peptides by the MHC class II molecules. The resulting peptide-MHC class II complexes are subsequently transported to the cell surface for scrutiny by $CD4^+$ T cells (Bryant and Ploegh 2004, Li, Gregg et al. 2005). MHC class II molecules are also capable of binding peptides that are derived from endogenous proteins and cytoplasmic proteins, via the cytoplasmic antigen presentation pathway. It remains unknown how peptides from cytoplasmic antigens gain access to MHC class II molecules, but one possibility is that they are acquired exogenously by bystander APCs and then follow the classical MHC class II pathway.

1.1.5e Immunodominance

Immunodominance primarily refers to the phenomenon by which only a small fraction of the possible thousands of peptides encoded by complex foreign (especially viral) antigens are selected for presentation on the surface of MHC class I alleles and induce measurable responses. Limitations of the $CD8^+$ T cell repertoire play a major role in the resultant non-immunogenic status of peptides as does inefficient antigen processing.

Only ~ 1/2000 of foreign antigen peptides achieve immunodominant status within a given class I allele, with roughly the equivalent number of peptides acquiring subdominant status, inducing a weak to non-detectable $CD8^+$ T cell response (Yewdell and Bennink 1999). A central feature of the $CD8^+$ T cell response, immunodominance has been observed in relation to a number of viruses including influenza (Turner, Kedzierska et al. 2005) and Epstein-Barr virus (EBV) (Argaet, Schmidt et al. 1994). There are a number of determinants which may influence the immunogenicity of a peptide including the pMHC complexes' abundance on the surface of APCs, the peptide binding preferences of the MHC class I molecules and the ability to suppress subdominant epitope responses (Yewdell and Bennink 1999) (Tynan, Elhassen et al. 2005).

Although an abundance of particular CTL epitopes presented on the surface of APCs does account for immundominance in some instances, the correlation between abundance and immunodominance is not always the case (Crotzer, Christian et al. 2000). An example of this has

been demonstrated in the T cell response to EBV. Crotzer and colleagues characterized the expression of four epitopes from the latent proteins EBNA3B (RRARSLSAERY), EBNA3C (RRIYDLIEL and FRKAQIQGL) and LMP2 (RRRWRRLTV). These epitopes are recognized by HLA-B27 subtypes HLA-B*2702, B*2704 or B*2705. RRIYDLIEL the immunodominant epitope was recognized in the context of all three B27 subtypes, whilst the remaining three epitopes induced a response in only one subtype, in the following combinations RRARSLSAERY/B*2702, RRRWRRLTV/B*2704 and FRKAQIQGL/B*2705. The apparent immunodominance of RRIYDLIEL did not correlate with the abundance of this epitope and showed an inverse relationship. The immunodominant epitope was least abundant with less than or equal to the abundance of the nonimmunogenic epitopes and their respective HLA-B27 subtypes (Crotzer, Christian et al. 2000). This phenomenon may be the result of factors such as the T cell repertoire or T cell affinity influencing the degree of the T cell response.

Another influential element on immunodominance involves CTLs specific for dominant peptides suppressing responses to other competing peptides (immunodomination). A number of mechanisms may play a role in the suppression of subdominant peptides including a rapid response by $CD8^+$ T cells specific for immunodominant peptides cascading towards a reduction in antigenic load and suboptimal expression of subimmunodominant peptides for CTL activation. Suppression of a subdominant peptide by an immunodominant peptide has been repeatedly observed (Deng, Yewdell et al. 1997) (Mylin, Bonneau et al. 1995). Additionally, competition at the APC level for $CD8^+$ T cell activation or systemic suppression of subimmunodominant peptides by $CD8^+$ T cells specific for the immunodominant peptides (Zinkernagel, Althage et al. 1978, Yewdell and Bennink 1999). This form of suppression results in enhanced responses to subimmunodominant peptides under conditions when responses to immunodominant peptides are eliminated by changes or removal of the determinant, its MHC class I restriction element or by removal of the immunodominant peptide-specific $CD8^+$ T cell response (Doherty, Biddison et al. 1978, Zinkernagel, Althage et al. 1978). Tynan *et al*, demonstrated with their investigation into EBV BZLF1 epitope APQPAPENAY, that peptide conformation can have a dramatic impact on the immunogenicity of an MHC-peptide complex and play a role in controlling immunodominance in T cell responses (Tynan, Elhassen et al. 2005).

A further key predictor of immunodominance is supported by findings of La Gruta *et al.* This work suggests that $CD8^+$ T cell precursor (CTLp) frequency and epitope density are the prime determinants of immundominance hierarchy after influenza A virus infection in mice (La Gruta, Kedzierska et al. 2006). Other studies have also shown a relationship between naïve CTLp frequencies and the magnitude of the CTL immune response for a variety of both viral and nonviral epitopes (Moon, Chu et al. 2007, Kotturi, Scott et al. 2008, Obar, Khanna et al. 2008). However,

demonstrating the complexity of this field La Gruta and colleagues have reported a differential contribution of CTLp frequency to immunodominance. Recent work by this group suggests that for subdominant epitopes there is no correlation between the CTLp frequency and the immunodominance hierarchy following intranasal infection of mice with influenza A virus (La Gruta, Rothwell et al. 2010).

In another study that was focused on the role of viral coinfection on the inter and intra-individual immunodominance of CTL responses against known HIV and EBV-derived, HLA-class I restricted epitopes, the impact of HLA-allele usage, functional avidity and HLA binding affinity were compared. Observations show the two major aspects of immunodominance, the magnitude and frequency of recognition, are related to each other, whilst the functional avidity appears to be of more importance to the magnitude of responses than peptide binding affinity to HLA molecules (Bihl, Frahm et al. 2006).

Recent reports by Hansen and colleagues portray a shift in the current immunodominant hierarchy paradigm. Their findings show simian immunodeficiency virus (SIV)-specific $CD8^+$ T cells recognize highly promiscuous epitopes that interestingly include dominant responses to epitopes restricted by MHC class II molecules. These data reveal flexibility in $CD8^+$ T cell recognition determining that the established rules of epitope recognition and immunodominance are not absolute, as well as uncovering unconventional $CD8^+$ T cell responses (Hansen, Sacha et al. 2013). What remains unclear is how these various factors balance out and influence immunodominance in a normal immune response.

1.2 Antigen Presentation and Recognition

1.2.1 Introduction to the T cell Receptor

T cell antigen receptors (TCR) are expressed on the surface of T lymphocytes in order to recognize antigens. These exceptional proteins detect foreign antigens that have been captured and presented by MHC molecules. TCRs have unique binding specificity determined by unique antigen-binding site, a result of variation in amino acid sequence. In order to respond to an infinite number of pathogens, humans possess a predicted 10^{12} lymphocytes with the theoretical number of an estimated 10^{18} receptors. This extraordinary complex system plays a critical role in the immune system.

1.2.2 αβ T cells, γδ T cells and coreceptor molecules

T cells fall into one of two lineages depending on whether the TCR is comprised of an α-chain paired with a β-chain or a γ-chain combined with a δ-chain. The vast majority of the body's

circulating T cells are of the αβ lineage, whilst γδ TCRs are less abundant and represent around 5% of T cells in the circulating lymphocyte population (De Rosa, Andrus et al. 2004). The γδ T cells are mainly located in the epithelial sites of the body such as the gut and the skin. These cells appear to form the first line of defence, recognising structures such as lipids and heat shock proteins, and are therefore an integral part of the innate immune system (Taghon and Rothenberg 2008). The precise details of the γδ lineage are not well understood.

αβ TCRs are expressed on the surface of around 95% of all T lymphocytes. These TCRs interact with peptide antigens in association with MHC molecules (pMHC) and are divided into two classes determined by their effector functions and accessory molecules. T cells bearing the CD8 coreceptors recognize MHC class I molecules and are referred to as CTLs or killer T cells, whilst those bearing CD4 coreceptors bind MHC class II molecules on APCs including DCs and macrophages and are called helper T cells (Wilson and Garcia 1997). It has been postulated whether the role of the coreceptors is to increase affinity of the TCR for MHC molecules or to recruit sufficient intracellular protein tyrosine kinase, $p56^{lck}$ (Lck) (Zamoyska 1998). One of the primary events which occur immediately upon TCR engagement with pMHC is the recruitment of T-cell-specific Lck via CD4 or CD8 coreceptors, activating the T cell response (Wange and Samelson 1996).

The CD3 complex consists of CD3δ, CD3ε, and CD3γ. The presence of the CD3 complex is required for stable cell surface expression and normal development of αβ TCRs in addition to the signalling transduction cascade initiated by the TCR binding to the pMHC (Rudolph, Stanfield et al. 2006). The intracytoplasmic homodimer CD3ζζ signals to the interior of the cell upon antigen binding. Activated Lck phosphorylates the immunoreceptor tyrosine based activation motifs (ITAMs), present on the cytoplasmic tails of the CD3 chains, which play a critical role in the transmitting of signals onward. The TCR complex is issued with ten ITAMs, which may offer flexibility to signalling (Janeway Jnr, Travers et al. 2001).

1.2.3 T Cell Receptor V (D) J recombination

At the centre of the adaptive immune response is the recognition of antigenic peptides bound to MHC molecules by TCRs. Generated somatically through site specific DNA recombination reactions, TCR diversity is a result of the random assembly of variable (V), diversity (D) and joining (J) gene segments. The four TCR polypeptides (α, β, γ and δ) each have a V, J and C region, however, β and δ also have D regions. The thymus is the site where TCR genes are initially rearranged and expressed during the earliest stages of T cell differentiation. Undergoing rearrangement first is TCR β, γ and δ genes followed by TCR α (Davis and Bjorkman 1988). V gene segments are carried in the germline of humans (and mice) and encode around the first 90

residues of mature TCR α and TCR β chains. These are assembled in the double negative (DN) stage (CD4-CD8-) of T cell development of the TCR β chain gene. TCR β gene assembly involves a short D gene segment which is juxtaposed to a short J segment. V segment rearrangement follows and is assembled downstream to the DJ segment. In contrast, the TCR α chain is rearranged last at the double positive (DP) stage (CD4+CD8+) of T cell development and contains no D region (Goldrath and Bevan 1999). The human TCR β locus comprises 42 Vβ gene segments, 2Dβ gene segments and 12 Jβ gene segments and the human TCR α locus has 43 Vα gene segments and 58 Jα gene segments (Turner, Doherty et al. 2006).

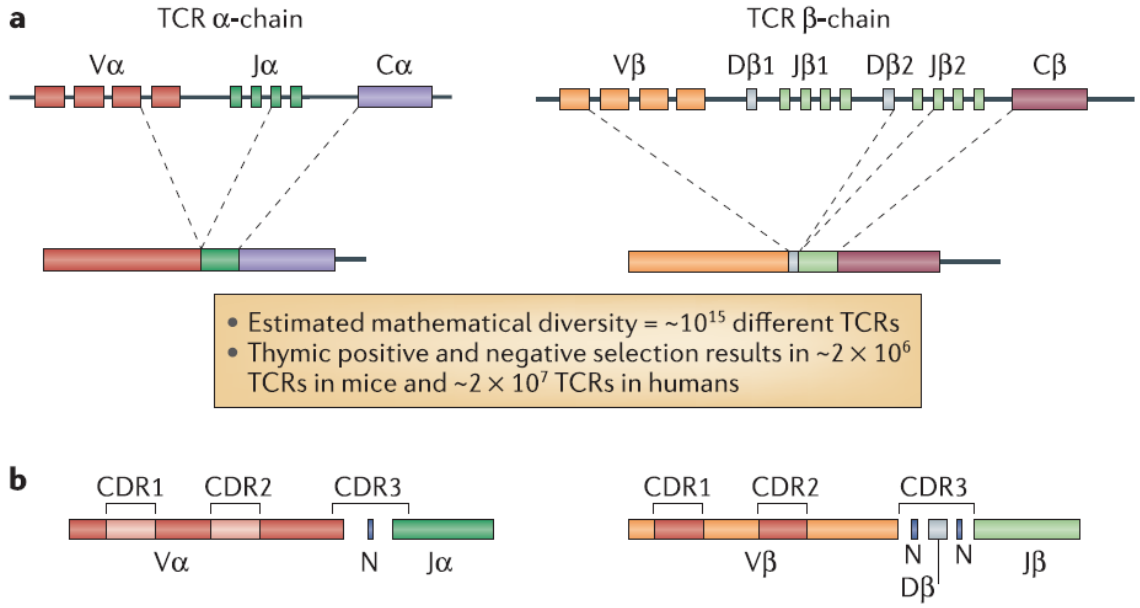


Figure 1. Schematic representation of TCR V(D)J Recombination. **a**) αβ TCR heterodimers consist of an α-chain and a β-chain generated by somatic gene recombination. **b**) Complimentarity-determining regions (CDRs) are regions of hypervariability encoded in the V gene segments. (Turner, Doherty et al. 2006).

V gene segments encode the regions of hypervariability known as complimentarity-determining regions (CDRs). The immunoglobulin-like fold results in three loops or CDRs. CDR loops 1 and 2 are encoded by the V gene segments and have a limited diversity resulting from the germline V gene segments (Goldrath and Bevan 1999). The CDR3 loop, however, is comprised of both germline residues from the V (D) J segments and nongermline components derived from N nucleotides that are added at the V (D) J joints during recombination (Davis and Bjorkman 1988). The CDR3 region is a critical component of the TCR as it can mediate contact with the antigenic peptide-MHC complex (Davis and Bjorkman 1988, Shortman, Egerton et al. 1990). Diversity within the CDR3 arises as a result of random nucleotide additions and DNA splicing by recombinase activating genes (RAG1 and RAG2). V (D) J recombination requires the coordinated

activity of both RAG1 and RAG2. These genes are only expressed in developing lymphocytes and determine irreversible T cell lineage commitment (Yui and Rothenberg 2014). The mathematical estimation of potential TCR diversity is around 10^{18} unique αβ TCRs, although after considering thymic positive and negative selection events the size of the naïve TCR αβ repertoire is estimated to be around 2×10^7 TCRs for each human (Arstila, Casrouge et al. 1999). Following translation, a functional TCR heterodimer comprised of a folded α-chain and β-chain, fused via a cysteine-cysteine linker is transported to the membrane surface.

1.2.4 Thymic Selection

The thymus is the site where positive and negative selection of T lymphocytes occurs. Their fate is determined by their surface receptors interaction with self-pMHC complexes displayed by thymic APCs (Klein, Kyewski et al. 2014). This process begins when DN (CD4-CD8-) progenitors are transmitted signals for positive selection from a pre-TCRβ complex generating immature DP (CD4+CD8+) thymocytes (Guidos 1996). Expressing a vast array of TCRαβ complexes for interactions with self-pMHC complexes on cortical thymic epithelial cells (cTECs), DP thymocytes will die by neglect if they fail to positively engage with self-pMHC complexes within the first 3-4 days. Negative selection (or clonal deletion) eliminates DP or single positive (SP) thymocytes that express TCR with high affinity for self antigens or due to excessive interactions (Guidos 1996, Starr, Jameson et al. 2003, Klein, Kyewski et al. 2014). Both these events of selection occur in discrete thymic microenvironments (cortex and medulla, respectively) composed of different types of APCs.

DP precursors bearing low densities of self MHC TCRαβ are positively selected and induced to differentiate into TCR^{hi} CD4 or CD8 SP thymocytes (Guidos 1996, Klein, Kyewski et al. 2014). Less than 5% of DP progenitors proceed to downregulate one of the coreceptors and become mature SP thymocytes (Goldrath and Bevan 1999). This process of "mainstream" αβ T cell positive selection occurs with a single APC type, the cTECs These specialized APCs are structurally arranged to support intimate interactions with DN and DP thymocytes (Klein, Kyewski et al. 2014). The intricate mechanisms involved in development and timing determining the commitment of TCR to either CD4 or CD8 lineage still remains to be completely understood. The process of commitment to either CD4 or CD8, however, has been affiliated with positive selection. Generally, DP precursors bearing class I MHC-specific TCRαβ will retain the expression of CD8, whilst those bearing the class II MHC-specific TCRαβ will retain CD4 expression (Fowlkes and Schweighoffer 1995, Guidos 1996). The affinity model proposed by Klein and colleagues envisages positive selection of T cells with a range of affinities for self peptides including T cells

with high affinity for self providing rapid, short lived immune responses balanced by T cells with a low affinity for self providing a sustained response (Klein, Kyewski et al. 2014).

1.2.5 TCR Bias

TCR diversity peaks in the naïve compartment, however, throughout an individuals' lifetime the antigen-experienced repertoire becomes skewed depending on antigen exposure and is often characterised by biases in the TCR repertoire. In addition individuals sharing the same HLA types display TCR biases with public clonotypes dominating the immune response. The extent of the bias may be one of three classifications depending on the degree of bias in the Vα/Vβ and CDR3 usage. TCR bias type 1 is characterised by the bias of at least one V region selection (generally, Vβ region), however, there is no display of CDR3 conservation. In contrast to type 1, type 2 TCR bias does involve sequence conservation within CDR3, often characterised by "motifs" within the CDR3 of the TCR Vα or Vβ chains. The least common form of TCR bias, type 3, is defined by conserved Vα and/or Vβ usage in combination with conservation within the CDR3 sequence and J region (Turner, Doherty et al. 2006, Gras, Kjer-Nielsen et al. 2008). As with TCR diversity, examples of TCR bias have been observed in persistent viral infections including EBV (Miles, Borg et al. 2006) and HIV (Stewart-Jones, Gillespie et al. 2005) as well as autoimmunity (Babbe, Roers et al. 2000) and lipid-based immunity (Borg, Wun et al. 2007).

The degree of antigen-driven TCR bias both within and between individuals appears to be determined by multiple factors including continuous antigenic stimulation, convergent recombination and structural constraints (Turner, Doherty et al. 2006, Gras, Kjer-Nielsen et al. 2008, Li, Ye et al. 2012). Such structural constraints playing a role in the emergence of TCR biases arise from the need to recognize atypical MHC landscapes including featureless and bulged epitopes (Gras, Kjer-Nielsen et al. 2008). In an effort to understand TCR bias structural studies have investigated features of immunodominant responses and unearthed fundamental elements of the basis of TCR bias and TCR-MHC interactions. Patterns of recognition between TCR and MHC have been identified indicating that a minimum of three positions on the MHC are contacted as a base requirement for MHC restriction (Tynan, Burrows et al. 2005). The adaptability of TCRs is highlighted by the capacity to recognize peptide-based or lipid-based antigens through various docking strategies ranging from parallel to orthogonal (Gras, Kjer-Nielsen et al. 2008).

Finite changes in the peptide ligand may result in viral escape as well as impact profoundly on the TCR repertoire. Thus, the TCR repertoire is a complex system determined by numerous factors including genetics and antigen exposure and further studies are required to unlock the mechanisms involved which may shed light on ways to optimize immunotherapeutic development.

1.2.6 The human TCR in Immune Responses

As mentioned previously epitope-specific TCR repertoires are selected from a pool of naïve TCRαβ heterodimers in the range of around 2 x 10^7 (Arstila, Casrouge et al. 1999). The adaptive immune system relies on T cells patrolling the blood stream to generate receptors with the ability to recognize and destroy virtually any pathogen. TCR diversity is provided by gene rearrangement within the variable regions (CDR loops) resulting from the V (D) J recombination process (Attaf, Huseby et al. 2015). Viral control, viral escape prevention and enhanced recognition of heterologous viruses are associated advantages of the utilization of highly diverse epitope-specific TCRs in response to viral infection (Charini, Kuroda et al. 2001, Messaoudi, Guevara Patino et al. 2002, Cornberg, Chen et al. 2006, Wang, Dash et al. 2012). In addition, TCR diversity has been shown to play a critical role towards homeostasis of T_{reg} cells and suppressor function (Fohse, Suffner et al. 2011). Of these 2 x 10^7 human TCRs, some clonotypes may be referred to as either "public" when found in different, unrelated individuals, or "private" when unique to the individual. This tendency of the clonotype appears to be encoded by the naïve repertoires (Thomas, Handel et al. 2013). Many studies have demonstrated distinct TCR repertoires in response to acute or persistent viral infections including influenza (Lehner, Wang et al. 1995), EBV (Argaet, Schmidt et al. 1994, Lim, Trautmann et al. 2000), human cytomegalovirus (HCMV) (Trautmann, Rimbert et al. 2005) and Human Immunodeficiency Virus (HIV) (Kalams, Johnson et al. 1994).

In one example, following acute influenza A virus infection, the $CD8^+$ T cell response in HLA-$A2^+$ individuals generally is directed predominantly towards a dominant influenza virus matrix protein epitope, $MP_{58-66.}$ This HLA-A2 specific public response is characterised by $CD8^+$ T cells specific for MP_{58-66} with a restricted TCR repertoire. These epitope specific $CD8^+$ T cells express a bias for Vβ17, and the highly conserved CDR3β sequences that contain the IRSSY motif (Moss, Moots et al. 1991, Lehner, Wang et al. 1995).

Lim *et al* performed extensive TCR analysis on the dominant HLA-A2 restricted EBV epitope GLCTLVAML derived from the BMLF1 lytic protein. Their findings showed several T cell clonotypes with recurrent Vβ subsets expressing Vβ2, Vβ4 or Vβ16 and highly conserved CDR3β motifs in epitope specific T cell lines derived from several unrelated individuals. Additionally, as the TCR α chains comprised the same TRAV region it is suggested that there is a hierarchical contribution of the TCR α-chain compared to the TCR β-chain CDR in the context of this particular pMHC complex (Lim, Trautmann et al. 2000).

Another example of an EBV public TCR clonotype is prominent in HLA-$B8^+$ individuals. This $CD8^+$ T cell response is directed at the epitope FLRGRAYGL, derived from the EBNA3A protein. Virtually all individuals select a particular T cell sequence denoted as the clonotype LC13 (Argaet, Schmidt et al. 1994). These CTL nearly all use the same TCRα and TCRβ chains derived

from identical Vα and Jα (TRAV 26-2*01, TRAJ52*01) and N region sequences and identical TCR Vβ, Dβ, and Jβ (TRBV7-8*03, TRBD1/D2, TRBJ2-7*01) sequences (Argaet, Schmidt et al. 1994). This highly conserved clonotype LC13 is the TCR used in the $CD8^+$ T cell response to the FLRGRAYGL epitope by most HLA-$B8^+$ individuals. $CD8^+$ T cells expressing this TCR are cross-reactive with the alloantigen HLA-B*4402. EBV^+ individuals expressing both HLA-B*8 and HLA-B*4402 demonstrate a reduced B8-restricted CTL response to the FLRGRAYGL epitope. Of most interest, is that these $CD8^+$ T cells recognise neither HLA-B*4402 nor express the LC13 TCR seen in the HLA-B8 response (Burrows, Silins et al. 1995).

An assessment of clonal diversity of CTL responses against HCMV by Trautmann and colleagues demonstrates a dramatic clonal focussing of $CD8^+$ T cells specific for the HLA-A2-restricted, pp65 epitope NLVPMVATV (NLV/A2). These NLV/A2 specific T cells derived from distinct individuals displayed public TCR features in immunodepressed or chronic inflammatory disease patients. In some instances, full conservation of the TCR Vα junctional region was observed. The NLV/A2 response in healthy individuals, however, was heterogenous with the $CD8^+$ T cell response comprising a number of Vβ regions. These findings suggest that the driving force behind this repertoire focusing in situations of repeated antigenic stimulation in the case of chronic inflammation and HCMV reactivation is TCR affinity/avidity (Trautmann, Rimbert et al. 2005). It remains to be determined if clonal focussing is linked to a high frequency in the preimmune repertoire, an improved immune response or a requirement to optimize interactions with the TCR.

1.2.7 TCR/pMHC Structures

In 1996 Garboczi and colleagues reported the original description of the crystal structure of a mouse TCR, providing some insight on the structure of the TCR/pMHC complex (Garboczi, Ghosh et al. 1996). Since then over thirty unique structures for TCR/pMHC class I complex structures and twenty unique TCR/pMHC class II complex structures have been solved broadening our knowledge of these significant interactions (Rossjohn, Gras et al. 2015). Generally, MHC molecules bind with peptides via a series of pockets within the peptide-binding groove. These pockets are referred to as either A-F or P1-P9 in the case of MHC class I and MHC class II, respectively (Adams and Luoma 2013). Predominantly, these are residues P5, P7 and P8 for nonamer peptides and P4, P6 and P7 for octamer peptides. Due to the extensive backbone interactions in the MHC class II binding groove, MHC class II peptides demonstrate a more uniform dispersal of key side chain contributions which are predominantly seen at P-1, P2, P3, P5 and P8 (Rudolph and Wilson 2002). Polymorphism within the MHC is centred around the antigen-binding cleft and is key in determining how and which peptides bind to any MHC molecule (Theodossis, Guillonneau et al. 2010).

Structural investigations showing docking geometry of TCR/pMHC complexes offers some insights as to how foreign peptides are recognised in the context of the same MHC and the role this may play in positive selection. Generally, it is a diagonal orientation that the TCR adopts above the pMHC complex. This orientation aligns the CDR1 and CDR2 loops over the α1 and α2 helices, or β1 helices of the MHC class I and MHC class II molecules, respectively (Rudolph and Wilson 2002). The basic affinity of the TCR for the generic MHC allele is provided by the interaction with the conserved germline CDR1 and CDR2 loops, whilst the CDR3 loop interacts with the peptide. In this context, the CDR1 and CDR2 loops are likely to play a role in positive selection, whereas the CDR3 loop's role lies within negative selection (Rudolph and Wilson 2002).

The structure of the TCR is comprised of an α- and a β- chain and resembles the overall scaffold of an Fab antibody fragment. The assembly of the germline-encoded regions within the complex indicates the CDR1 and CDR2 regions mostly contact the helices of the MHC, whereas the CDR3 region is in contact with the peptide (Wilson and Garcia 1997). Cradled within the MHC-Ag binding cleft, the peptide is positioned to enable the CDR3 loops of the TCR to sit above the MHC, permitting direct contact with the up-pointing peptide side chains. It is the regions of the peptide outwardly oriented that can directly contact the TCR, whilst buried regions of the peptide may have an indirect effect on TCR binding (Theodossis, Guillonneau et al. 2010). The interaction between the TCR and the pMHC occurs approximately in a diagonal orientation (Rudolph and Wilson 2002). This docking method permits the TCR binding surface to fit between the high areas that are situated near the amino termini of each of the two MHC α helices, maximising the few peptide residues that are exposed to the TCR (Mazza and Malissen 2007). The ability of the CDR3 loops to undergo conformational changes upon interaction with the pMHC underpins its role in TCR cross-reactivity (Rudolph, Stanfield et al. 2006).

The structural characteristics of the determined TCR/pMHC interactions have been compared and show that many closely related complexes vary in terms of the bound peptide, minor differences in MHC structure and subtle changes in TCR gene usage. Interestingly, what appears to be common to all TCR/pMHC I complexes solved to date is the docking topology, whereby, the V α-chain sits over the α2-helix and the V β-chain is positioned over the α1-helix of the MHC class I molecule (Rossjohn, Gras et al. 2015).

An elegant example is the archetypical LC13 TCR-HLA-B*0801-FLRGRAYGL structure (Figure 2). Following antigen recognition, this immunodominant TCR undergoes extensive conformational change in the CDRs, including disruption of the CDR1α and CDR2α loops resulting in an enhanced fit with the pMHC complex. With the focus on the long axis of the HLA-B8 peptide binding groove, the LC13 TCR binds in an approximate diagonal (60°) orientation. Of note, the usage of the variable domains Vα (58%) and Vβ (42%) is approximately equal which is consistent

with public TCRα and TCRβ selection (Kjer-Nielsen, Clements et al. 2003). The CDR3 loops were shown to be the driving force behind the interaction in a somewhat "peg and notch" manner with the P7-Tyr (peg) protruding into the central cavity (notch) of the LC13 TCR (Kjer-Nielsen, Clements et al. 2002, Kjer-Nielsen, Clements et al. 2003).

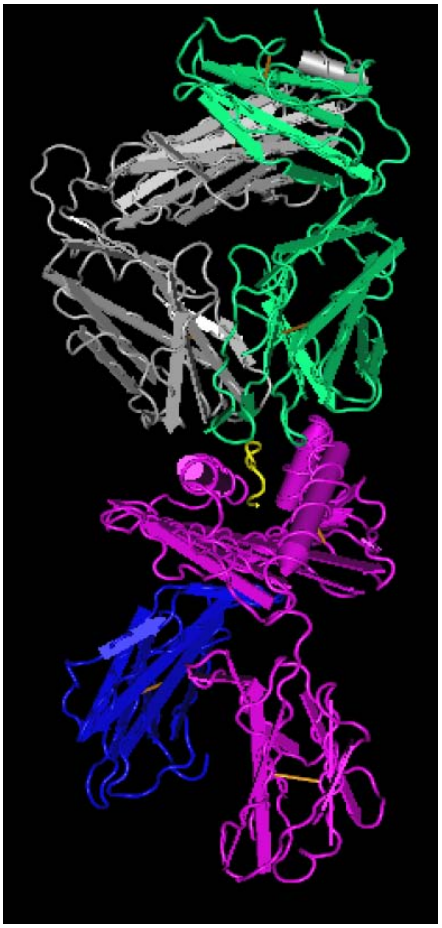


Figure 2. LC13 TCR Complexed to HLA-B8/FLRGRAYGL.

Ribbon representation of the LC13-HLA-B8-FLRGRAYGL complex. The bound EBV peptide is shown in yellow. The α-chain and β-chain of the LC13 TCR are shown in green and grey, respectively. The HLA-B*0801 molecule is shown in purple and the blue ribbons represent the β2-microglobulin domain (Kjer-Nielsen, Clements et al. 2003).

Crystal structures of TCR/pMHC have demonstrated how the peptide contacts the TCR/MHC interface. With the determining of more than 50 unique TCR-pMHC structures the field has seen a progression in knowledge in the area, however, there still remains fundamentals of T cell biology to be explored.

1.3 Immune response to a herpes virus

1.3.1 *Herpesviridae*

In 2008 the International Committee on Taxonomy of Viruses updated the former family *Herpesviridae* by splitting members into three families and incorporating them into a new order *Herpesvirales.* The family *Herpesviridae* retains mammal, bird and reptile viruses, while the new families *Alloherpesviridae* and *Malacoherpesviridae* incorporate fish and frog viruses and bivalve viruses, respectively (Davison, Eberle et al. 2009). Herpesviruses have distinct morphology from

all other viruses. A T = 16 icosahedral capsid contains a linear, double-stranded DNA genome of 125-290 kbp. This capsid is surrounded by a proteinaceous matrix dubbed the tegument and then by a lipid envelope containing membrane-associated proteins (Pellet and Roizman 2006).

The family *Herpesviridae* is divided into subfamilies *Alphaherpesviridae*, *Betaherpesviridae* and *Gammaherpesviridae.* A vast proportion of the human population is infected with one or more of the eight known human herpesviruses, including the alphaherpesviruses herpes simplex viruses 1 and 2 (HSV-1 and HSV-2) and varicella-zoster virus (VZV), the betaherpesviruses human cytomegalovirus (HCMV) and human herpesviruses 6 and 7 (HHV-6 and HHV-7) and the gammaherpesviruses Epstein-Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV) (Davison, Eberle et al. 2009). Despite the presence of a powerful immune response directed against herpesviruses they have the ability to establish lifelong infections which have been associated with the development of several malignancies (Horst, Verweij et al. 2011).

1.3.2 A Brief Overview of EBV

EBV, a lymphotropic gamma-1 herpesvirus, is widespread in all human populations, infecting around 95% of the adult population worldwide. The transmission of this lifelong persistent virus is via the oral route. After close contact, viral particles present in the saliva of infected individuals enter into the oral cavity of the naïve individual, the virus is amplified by replicative (lytic) infection in permissive cells (probably epithelial) in the oropharynx. This lytic infection results in vast amounts of virus shedding into the throat. Simultaneously, the infection of the mucosal B cells occurs and the virus initiates latent, persistent infection of the B cell pool (Kutok and Wang 2006, Hislop, Taylor et al. 2007, Long, Taylor et al. 2011). In developing countries, EBV is associated with an asymptomatic infection occurring usually in infancy or early childhood. In developed countries, however, acute infection is often delayed until adolescence or early adulthood and in around 25% of cases can result in the self-limiting lymphoproliferative disorder, infectious mononucleosis (IM). IM was etiologically linked to EBV in 1968 by Henle *et al*, symptoms associated with IM range from fever and sore throat to lymphadenopathy and splenomegaly (Henle, Henle et al. 1968, Kutok and Wang 2006, Rickinson and Kieff 2007).

EBV types 1 and 2 (also denoted as types A and B) circulate in most populations. These viral types differ largely in nuclear protein genes that encode EBNA-LP, EBNA2, EBNA3A, EBNA3B and EBNA3C, which are expressed in the latency III stage, with up to 47% amino acid sequence difference between types 1 and 2 for EBNA 2 alone. Type 1 is predominantly seen in developed countries, whereas type 2 is prevalent in equatorial Africa and New Guinea populations (Kieff and Rickinson 2007). Although EBV infection is generally benign and uneventful it is

currently associated with a number of malignancies. Diseases of the immunocompetent host associated with EBV include Burkitt's Lymphoma (BL), Hodgkin's Lymphoma (HL) and nasopharyngeal carcinoma (NPC) and those of the immunocompromised host include post-transplantation lymphoproliferative disorder (PTLD), AIDS associated B cell lymphomas, Wiskott-Aldrich syndrome-associated B cell lymphomas and Leiomyosarcoma (Kutok and Wang 2006).

There are a number of forms of EBV latency, apart from Latency 0 infection in which all antigen expression is suppressed, all forms express EBER1 and EBER2 RNAs and the Bam HIA rightward transcripts. During Latency I infection EBNA-1 is also expressed. In Latency II infection EBNA-1 is expressed as well as LMP from promoters in the Bam HIN region of the genome. Latency III infection sees the full spectrum of latent proteins expressed, including EBNA-1, 2, 3A, 3B, 3C and –LP (Rickinson and Kieff 2007). Some five decades on since the discovery of EBV by Epstein and colleagues (Epstein, Achong et al. 1964) and a vaccine continues to elude immunologists the world over.

1.3.3 Structure of EBV

EBV, as with other herpesviruses, has a toroid-shaped DNA core in a nucleocapsid with 162 capsomeres, a protein tegument lies between the nucleocapsid and an outer envelope with external glycoprotein spikes. The virus contains a linear, double stranded, 184kb DNA genome, features of which include short and long unique sequence domains, internal (IR1) and terminal (TR) tandem repeat and direct repeats, as well as perfect and imperfect DNA repeats, most of which are within the open reading frame (ORF). The linear termini of the genome join intracellularly to form circular, episomal DNA. The number of TR in any given linear genome is highly variable, a result of the random TR cleavage that creates full length genomes. A signature number of TRs is created when the linear genome circularises into the viral episome (Figure 3a). This signature remains during episomal replication in dividing cells and all subsequent daughter cells derived from the single infected cell share this signature which can be used to determine clonal populations of EBV-infected cells (Kutok and Wang 2006, Kieff and Rickinson 2007).

1.3.4 Lytic EBV Proteins of interest

Around 80 EBV proteins are expressed during lytic viral replication, which involves the sequential expression of immediate early, early and late proteins. Lytic cycle activation is initiated by the expression of two immediate early proteins BZLF1 (Zta) and BRLF1 (Rta). Cox *et al* noted that Rta and Zta appeared to regulate gene expression at the level of transcription and that these transactivators work in concert to facilitate reactivation (Cox, Leahy et al. 1990). BZLF1 is a member of the basic leucine zipper family which binds to the AP1-like Z response elements in EBV

early promoters (Zalani, Holley-Guthrie et al. 1996). BRLF1 has distant homology to c-myb which interacts with BZLF1 transactivating the BMRF1 promoter (Kenney, Holly-Guthrie et al. 1992). BMRF1, an early protein has proved to be essential for the production of infectious EBV progeny as well as being associated with enhanced expression of BHLF1 (Neuhierl and Delecluse 2006). The early transactivator protein, BMLF1 is activated by the expression of BZLF1 and BRLF1(Lieberman, O'Hare et al. 1986).

1.3.5 Latent EBV Proteins of interest

During latent infection, EBV genes hijack normal B cell mechanisms leading to the establishment of lymphoblastoid cell lines (LCLs). At least nine latency-associated viral antigens are expressed by LCLs and include the six nuclear antigens, EBV nuclear antigen 1 (EBNA1), EBNA2, EBNA3A, EBNA3B, EBNA3C and EBNA leader protein (LP) and latent membrane proteins LMP1, LMP2A and LMP2B (Figure 3b). As discussed earlier when all nine of these antigens are expressed it is referred to as latency III form of EBV infection (Kutok and Wang 2006, Rickinson and Kieff 2007). The focus will now shift to EBNA1 and the three EBNA3 genes, EBNA3A, -3B and -3C. EBNA1 is essential for the persistence and maintenance of the EBV episome and enhancing genome expression. A unique feature of EBNA1 is the glycine-alanine repeat (GAr) domain which inhibits MHC class I presentation of cis-linked epitopes. Also of great importance, EBNA1 is the only viral protein that is expressed in all forms of latency and in all EBV-associated malignancies (Rickinson and Kieff 2007, Long, Taylor et al. 2011).

Tandemly placed in the EBV genome, EBNA3A, -3B and -3C share a similar gene structure. These three stable proteins can interact with cellular proteins involved in transcription regulation and bind to the cellular repressor-binding protein (RBP)-Jκ (Le Roux, Kerdiles et al. 1994, Young and Rickinson 2004). EBNA3A and EBNA3C are both critical for B-lymphocyte growth transformation and survival. Interestingly, EBNA3B is not essential for latent infection of B lymphocytes and cell survival, although its presence in EBV isolates of healthy carriers indicates that it plays a part in natural infection (Kieff and Rickinson 2007). Recent observations from White and colleagues investigating viruses lacking this gene in both a humanized mouse model and patients with aggressive lymphomas suggest that EBNA3B behaves as a tumour suppressor. This appears to apply across all EBV strains, not just the B95-8 strain and may play a role in minimizing an oncogenic risk to the host (White, Ramer et al. 2012).

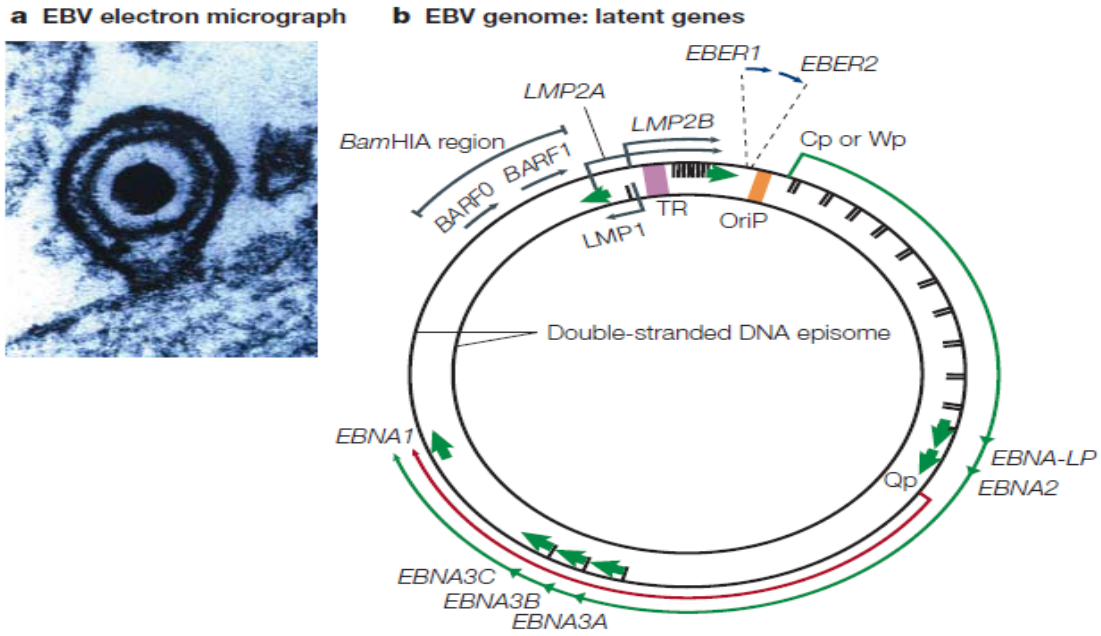


Figure 3. Epstein-Barr virus genome. a) Electron micrograph of the EBV virion. **b)** Diagramatic representation of the location and transcription of the EBV latent genes on the double-stranded viral DNA episome (Young and Rickinson 2004).

1.3.6 Cellular Immune Responses to EBV

$CD8^+$ T cells recognise peptides derived from viral proteins associated with MHC molecules. The adaptive immune response, in particular T cells, is vital in controlling EBV infected B-cell proliferation and is of immense importance during persistent infection (Hislop, Taylor et al. 2007). Studies focusing on the $CD8^+$ T cell response during IM portray a pattern of reactivity that to an extent mirrors the EBV infection cycle. Individual lytic epitopes have been shown to constitute up to 50% of the $CD8^+$ T cell population. EBV epitope-specific reactivities are typically seen at a higher frequency towards epitopes from immediate early (BZLF1 and BRLF1) antigens or a subset of early antigens (including BMRF1 and BMLF1). Delayed early and in particular late lytic antigens have a tendency for less frequent or smaller responses (Hislop, Taylor et al. 2007, Abbott, Quinn et al. 2013). Indeed this hierarchy correlates with the varying levels at which epitopes from these lytic proteins are presented on the surface of infected cells and portrays a system whereby the virus's array of immune evasion proteins seeks to impair the cell's antigen processing capacity (Rowe and Zuo 2010).

Responses to latent epitopes are present at a vastly reduced frequency in IM compared to lytic epitope responses, with individual epitope specificities contributing up to only 5% of $CD8^+$ T cells (Hislop, Kuo et al. 2005). The largest focus of responses is towards epitopes in the EBNA3 family of proteins (EBNA3A, EBNA3B and EBNA3C), which appears to likely reflect this group of proteins accounting for around 60% of the latent genes coding capacity. Responses towards

other latent antigens including EBNA1, EBNA2, EBNA-LP and LMP2 have also been seen at an equal intensity (Hislop, Taylor et al. 2007, Taylor, Long et al. 2015). Qualitative changes in the content of EBV-specific $CD8^+$ T cell responses have been noted over time following infection. A lytic epitope restricted by HLA-A2 with reactivties equal to other sustained epitopes has been shown to disappear post-IM, whilst responses to latent epitopes may not be detectable until 3-4 months post primary infection (Hislop, Annels et al. 2002). These changes result in a complex distribution of EBV-specific reactivities whereby $CD8^+$ T cell responses for one individual may present quite differently when at the post-IM state compared to at the height of acute infection (Hislop, Taylor et al. 2007).

A number of $CD8^+$ T cell responses to EBV latent and lytic proteins of interest have been previously described and are shown in Table 1 and Table 2, respectively. Lytic proteins BRLF1 and BZLF1 have a number of defined $CD8^+$ T cell epitopes presented by HLA-A, B and C. For example, Pepperl *et al* defined several BRLF1 epitopes including the HLA-A24-restricted decamer epitope DYCNVLNKEF, and ATIGTAMYK restricted to HLA-A11, as well as the HLA-Cw4-restricted ERPIFPHPSKPTFLP epitope and HLA-B61 restricted QKEEAAICGQMDLS epitope (Pepperl, Benninger-Doring et al. 1998). Examples of BZLF1 epitopes are HLA-B8-restricted RAKFKQLL (Bogedain, Wolf et al. 1995) and the HLA-B*3501-restricted EPLPQGQLTAY epitope (Saulquin, Ibisch et al. 2000).

There are less known epitopes from the early lytic proteins BMLF1 and BMRF1. Steven *et al* have however defined a number of epitopes from both these proteins. These include the HLA-A2-restricted GLCTLVAML and HLA-B18-restricted DEVEFLGHY epitopes which are both from the BMLF1 protein (Steven, Annels et al. 1997). BMRF1 epitopes include the HLA-Cw6 YRSGIIAVV peptide which has also been found to be restricted by HLA-B39 by Pudney *et al* and the FRNLAYGRTCVLGKE epitope with HLA-Cw3 restriction, for which the minimal active sequence has not been precisely mapped (Steven, Annels et al. 1997, Pudney, Leese et al. 2005). Results from Pudney and colleagues suggest a focusing of $CD8^+$ T cell responses toward epitopes from immediate early and early proteins and a hierarchy of immunodominance amongst the EBV lytic cycle antigens (Pudney, Leese et al. 2005).

The elusive search for an EBNA1 CTL epitope came to fruition in 1997 when Blake *et al* showed that $CD8^+$ T cell responses to EBNA1 epitopes were in fact detectable in humans with the discovery of two EBNA1 epitopes HPVGEADYFEY restricted to HLA-B*3501 and the HLA-A*0203-restricted VLKDAIKDL (Blake, Lee et al. 1997). In addition to this previous work, in 2000 Blake and colleagues mapped another three EBNA1 epitopes, two HLA-B7-restricted epitopes RPQKRPSCI and IPQCRLTPL and the HLA-B53-restricted epitope HPVGEADYF, which lies within the HLA-B*3501 HPVGEADYFEY sequence (Blake, Haigh et al. 2000).

There are a number of immunodominant epitopes attributed to the EBNA3 proteins. One particular immunodominant EBNA3A epitope is HLA-B8-restricted FLRGRAYGL (FLR) which dominates over another HLA-B8-restricted epitope, QAKWRLQTL (Burrows, Sculley et al. 1990) (Burrows, Gardner et al. 1994). FLR may be considered the most effective of CTL epitopes as this domination by a single TCR has been shown during acute primary infection (Callan, Annels et al. 1998) as well as the major clonotype for at least 18 years (Miles, Silins et al. 2005). This immunodominance phenomena is also seen in the response to the EBNA3B HLA-A11-restricted epitopes, where IVTDFSVIK is dominant over AVFDRKSDAK (Gavioli, Kurilla et al. 1993). Interestingly, most of the known EBNA3C epitopes are HLA-B restricted including the HLA-B*2705-restricted LRGKWQRRYR and FRKAQIQGL epitopes. Some epitopes are restricted by multiple class I HLA alleles such as the immunodominant RRIYDLIEL epitope which is presented by the HLA-B27 subtypes –B*2702, -B*2704 and –B*2705 (Brooks, Murray et al. 1993, Hill, Worth et al. 1995).

1.3.7 Unusually long EBV CTL Epitopes

Generally peptide epitopes presented by MHC class I molecules are 8 to 10 residues in length. For most MHC class I alleles two pockets within the peptide binding groove display a marked preference for two or three amino acids at certain anchor positions within the peptide (Rammensee, Falk et al. 1993). Although most peptides naturally presented for recognition by MHC class I molecules conform to this canonical length (8-10 residues), it has been demonstrated by several groups investigating a diverse range of viruses and several melanoma antigens that peptides of 11-14 residues can bind to particular MHC class I molecules (Burrows, Rossjohn et al. 2006). Indeed, earlier findings from our group demonstrate unusually long peptides binding to MHC class I molecules (Green, Miles et al. 2004, Burrows, Bell et al. 2008), (M. J. Rist, J. M. Burrows and S. R. Burrows, unpublished data).

Based on information including proteasome fragment length preference (Wenzel, Eckerskorn et al. 1994), class I antigen processing machinery (Pamer and Cresswell 1998) and TAP peptide length preferences (Momburg, Roelse et al. 1994), web-based algorithms have been designed to predict MHC class I peptides between 8-10 residues (Rammensee, Bachmann et al. 1999, Nussbaum, Kuttler et al. 2003). Despite these limitations and using alternate approaches to these web-based algorithms, there are numerous reports defining unusually long $CD8^+$ T cell epitopes over 10 residues in length presented by MHC class I molecules (Burrows, Rossjohn et al. 2006). Twenty-six CTL epitopes greater than 10 amino acids in length have been reported in viruses including HIV (Sipsas, Kalams et al. 1997), HCMV (Weekes, Wills et al. 1999), hepatitis B virus (Schirmbeck, Melber et al. 1994) and melanoma antigens (Aarnoudse, van den Doel et al.

1999, Kawakami, Wang et al. 2001). Of note, is the 14mer epitope derived from an alternate opening reading frame of the macrophage colony stimulating factor gene. This CTL epitope is naturally presented by HLA-B*3501 (Probst-Kepper, Stroobant et al. 2001, Probst-Kepper, Hecht et al. 2004).

Several MHC class I bound, immunogenic peptides that are longer than 10 residues have undergone structural characterisation that has shown in each case, the peptide adopts a bulging confirmation whereby the extra length "inserts" between the primary anchor residues. This adaptation maintains the hydrogen bonding with the peptide N- and C- termini (Probst-Kepper, Hecht et al. 2004, Tynan, Borg et al. 2005, Miles, Borg et al. 2006). Six of these defined unusually long CTL epitopes are derived from EBV antigens (Blake, Lee et al. 1997, Rickinson and Moss 1997, Green, Miles et al. 2004, Pudney, Leese et al. 2005, Burrows, Bell et al. 2008). For example, the highly immunogenic 13-mer epitope, ^{52}LPEPLPQGQLTAY^{64} (LPEP) from the BZLF1 antigen of EBV that is presented by HLA-B*3508 (Green, Miles et al. 2004). Also from BZLF1 and completely overlapped by the 13mer is the 11mer ^{54}EPLPQGQLTAY^{64} naturally presented by the HLA-B*3501 allele (Miles, Elhassen et al. 2005). ^{116}RPQGGSRPEFVKL^{128} is a 13mer epitope that binds to HLA-B*0702 from the lytic EBV antigen BMRF1 (Pudney, Leese et al. 2005). Another example of a noncanonical epitope is the 11-mer ^{407}HPVGEADYFEY^{417} (HPV), derived from the EBNA1 protein of EBV (Blake, Lee et al. 1997). HPV has been found to be highly immunogenic and binds promiscuously to two HLAs and may be recognised by individuals expressing either HLA-B*3501 or HLA-B*3508. Structural studies indicate HPV bulges out of the MHC molecule, in HLA-B*3501 the bulged section was disordered, whereas the bulged epitope adopts an ordered confirmation in relation to HLA-B*3508 (Miles, Borg et al. 2006).

A number of these unusually long CTL epitopes completely overlap with shorter sequences within the peptides that conform to the conventional binding length and score well using the binding algorithms, however, have not been found to be immunogenic. Both HPV and LPEP are such examples. HPV has been shown to be immundominant over shorter peptides HPVGEADY and HPVGEADYF that both conform to the peptide binding motif of HLA-B*3501 (Blake, Lee et al. 1997). Additionally, LPEP has been shown to bind with a higher affinity to HLA-B*3508 compared to the shorter sequences EPLPQGQLTAY and LPQGQLTAY which both conform to the binding motif of the HLA-B*3508 allele (Green, Miles et al. 2004).

A number of MHC class I molecules have demonstrated the ability to present unusually long CTL epitopes (Burrows, Rossjohn et al. 2006). Interestingly, five of the six EBV noncanonical epitopes are presented by members of the HLA-B7 supertype family. These alleles include HLA-B*0702, HLA-B*3501 and HLA-B*3508 which prefer peptides with a Proline at position 2. It has been postulated by Burrows and colleagues that members of this supertype present a high number

of long peptides due to a unique feature of the class I antigen processing pathway (Burrows, Rossjohn et al. 2006). As a result of the ER-associated aminopeptidases inability to cleave the bond at the N- side of Proline, an accumulation of peptides with x-Pro-x_n sequence is likely to occur (Serwold, Gaw et al. 2001), with many of these peptides likely to be long. As the HLA-B7 supertype has a preference for Prolines at position 2 these alleles are more likely to have unusually long CTL epitopes available for presentation within the ER compared to other HLA alleles that are not members of this supertype (Burrows, Bell et al. 2008).

Four out of the six unusually long EBV CTL epitopes are located in lytic antigens BZLF1 and BMRF1, with the remaining two epitopes from the latent antigens EBNA1 and EBNA3B. With such numbers it is difficult to ascertain if this is representative of a pattern or disproportionate lean towards the lytic antigens containing a majority of the noncanonical epitopes. It may reflect the individual epitope-specific populations that are typically 0.2-2% and 0.05-1% of the CTL population for lytic and latent epitopes, respectively (Bihl, Frahm et al. 2006). In addition, this observation mirrors to an extent the relative immunodominance of CTL responses of the lytic and latent cycle antigens (Hislop, Taylor et al. 2007). What remains to be determined is the frequency with which unusually long CTL epitopes are recognized by $CD8^+$ T cells and if there are other MHC class I alleles capable of naturally presenting such epitopes.

Table 1 Latent EBV CD8$^+$ T cell epitopes (adapted from Hislop et al, 2007)

EBV Antigen	Epitope Coordinates	Epitope Sequence	HLA Restriction	Reference
Latent				
EBNA1	72-80	RPQKRPSCI	B7	Blake et al, 2000
	407-415	HPVGEADYF	B53	Blake et al, 2000
	407-417	HPVGEADYFEY	B35.01	Blake et al, 1997
	528-536	IPQCRLTPL	B7	Blake et al, 2000
	574-582	VLKDAIKDL	A2.03	Blake et al, 1997
	566-574	LQTHIFAEV	A2	(Marescotti, Destro et al. 2009)
EBNA3A	158-166	QAKWRLQTL	B8	(Burrows, Gardner et al. 1994)
	176-184	AYSSWMYSY	A30.02	(Steven, Leese et al. 1996)
	246-253	RYSIFFDY	A24	Burrows et al, 1994
	325-333	FLRGRAYGL	B8	(Burrows, Sculley et al. 1990)
	378-387	KRPPIFIRRL	B27	Brooks J, unpub
	379-387	RPPIFIRRL	B7	Hill et al, 1995
	406-414	LEKARGSTY	B62	Rickinson and Moss, 1997
	450-458	HLAAQGMAY	?	Burrows et al, 1994
	458-466	YPLHEQHGM	B35.01	Burrows et al, 1994
	491-499	VFSDGRVAC	A29	Rickinson and Moss, 1997
	502-510	VPAPAGPIV	B7	Rickinson and Moss, 1997
	596-604	SVRDRLARL	A2	Burrows et al, 1994
	603-611	RLRAEAQVK	A3	(Hill, Lee et al. 1995)
	617-625	VQPPQLTLQV	B46	(Whitney, Chan et al. 2002)
EBNA3B	149-157	HRCQAIRKK	B27.05	Rickinson and Moss, 1997
	217-225	TYSAGIVQI	A24.02	Rickinson and Moss, 1997
	244-254	RRARSLSAERY	B27.02	Brooks et al, 1993
	279-287	VSFIEFVGW	B58	(Lee, Chan et al. 2000)
	399-408	AVFDRKSDAK	A11	(Gavioli, Kurilla et al. 1993)
	416-424	IVTDFSVIK	A11	Gavioli et al, 1993
	488-496	AVLLHEESM	B35.01	Rickinson and Moss, 1997
	657-666	VEITPYKPTW	B44	Rickinson and Moss, 1997
EBNA3C	163-171	EGGVGWRHW	B44.03	(Morgan, Wilkinson et al. 1996)
	213-222	QNGALAINTF	B62	(Kerr, Kienzle et al. 1996)
	249-258	LRGKWQRRYR	B27.05	Brooks et al, 1993
	258-266	RRIYDLIEL	B27.02/.04/.05	Brooks et al, 1993
	271-278	HHIWQNLL	B39	Rickinson and Moss, 1997
	281-290	EENLLDFVRF	B44.02	(Burrows, Misko et al. 1990)
	284-293	LLDFVRFMGV	A2.01	Kerr et al, 1996
	285-293	LDFVRFMGV	B37	(Shi, Smith et al. 1997)
	335-343	KEHVIQNAF	B44.02	(Khanna, Burrows et al. 1992)
	343-351	FRKAQIQGL	B27.05	Brooks et al, 1993

Table 2 Lytic EBV CD8$^+$ T cell epitopes (adapted from Hislop et al, 2007)

EBV Antigen	Epitope Coordinates	Epitope Sequence	HLA Restriction	Reference
Lytic				
BRLF1	25-39	LVSDYCNVLNKEFT	B18	Pepperl et al, 1998
	25-33	LVSDYCNVL	A2.05	Annels NE, unpub
	28-37	DYCNVLNKEF	A24	Pepperl et al, 1998
	91-99	AENAGNDAC	B45	Pudney et al, 2005
	101-115‡	IACPIVMRYVLDHLI	A24/Cw2	Pudney et al, 2005
	109-117	YVLDHLIVV	A2.01	Saulquin et al, 2000
	121-135	FFIQAPSNRVMIPAT	Unknown	Pepperl et al, 1998
	134-142	ATIGTAMYK	A11	Pepperl et al, 1998
	148-156	RVRAYTYSK	A3	Benninger-Doring et al, 1999
	225-239	RALIKTLPRASYSSH	A2	Pepperl et al, 1998
	393-407	ERPIFPHPSKPTFLP	Cw4	Pepperl et al, 1998
	441-455	EVCQPKRIRPFHPPG	Unknown	Pepperl et al, 1998
	529-543	QKEEAAICGQMDLS	B61	Pepperl et al, 1998
BZLF1	52-63	LPEPLPQGQLTAY	B35.08	Green et al, 2004
	54-63	EPLPQGQLTAY	B35.01	Saulquin et al, 2000
	77-86	APQPAPENAY	B35.08	Tynan et al, 2005
	79-89	QPAPENAYQAY	B35.01/08	(Burrows, Bell et al. 2008)
	172-183	DSELEIKRYKNR	B18	(Saulquin, Ibisch et al. 2000)
	190-197	RAKFKQLL	B8	Bogedain et al, 1995
	197-205	LQHYREVAA	Cw8	Pudney et al, 2005
	209-217	SENDRLRLL	B60	(Scotet, David-Ameline et al. 1996)
BMLF1	265-273	KDTWLDARM	Unknown	Steven et al, 1997
	259-267	GLCTLVAML	A2.01	Steven et al, 1997
	397-405	DEVEFKGHY	B18	Steven et al, 1997
	435-444	SRLVRAILSP	B14	Annels NE, unpub
BMRF1	20-28	CYDHAQTHL	A24	Pudney et al, 2005
	86-100‡	FRNLAYGRTCVLGKE	Cw3/10	Steven et al, 1997
	116-128	RPQGGSRPEFVKL	B7	Pudney et al, 2005
	208-216	TLDYKPLSV	A2.01	(Hislop, Annels et al. 2002)
	268-276	YRSGIIAVV	Cw6	Steven et al, 1997
	268-276	YRSGIIAVV	B39	Pudney et al, 2005
	286-295	LPLDLSVILF	B53	Annels NE, unpub

1.4 Scope of Thesis: Aims and Hypotheses

The primary aim of this study was to explore highly immunogenic antigens of Epstein-Barr virus to identify novel epitopes. This was conducted in an unbiased investigation of both epitope length and HLA involvement.

Aim One: Determine the relative prevalence of peptide epitopes of over 10 amino acids in length in $CD8^+$ T cell recognition of EBV.

> ***Hypothesis One:*** Peptide epitopes of over 10 amino acids in length are abundant in antiviral $CD8^+$ T cell responses.

Aim Two: Determine if the highly immunogenic BZLF1 antigen of EBV includes novel $CD8^+$ T cell epitopes.

> ***Hypothesis Two***: There are many unidentified $CD8^+$ T cell epitopes to be defined in the BZLF1 antigen.

Aim Three: Determine if the peptide length preferences of class I human leukocyte antigens influence epitope selection in the EBV-specific T cell response.

> ***Hypothesis Three:*** Different class I human leukocyte antigens have distinct peptide length preferences, and this influences epitope selection in the EBV-specific T cell response.

Aim Four: Characterise a cross-reactive T cell response to an EBV epitope and self-peptide presented by HLA-B*1801.

> ***Hypothesis Four:*** The T cell receptor repertoire that cross-reacts with an EBV epitope and a self-peptide are identical between individuals.

Chapter 2

Methods

2.1 Cell Culture

2.1.1 Ficoll-Paque Density Gradient – lymphocyte isolation

Blood (obtained with informed consent and approval from the QIMR Berghofer Medical Research Institute Human Research Ethics Committee, Brisbane, Queensland, Australia) was diluted 1:1 with warmed RPMI1640 medium (37°C) and mixed. 10mL of Ficoll-Paque (StemCell Technologies, USA) was aliquoted into a 30ml flat bottom tube and slowly overlayed with up to 20ml of blood/RPMI1640 mixture. These tubes were centrifuged at 500 x g for 20 minutes without the brake. Separation of the blood results in a number of defined layers. The top layer consists of plasma, the middle layer is lymphocyte rich, including monocytes, followed by the layer of Ficoll and the base layer of red blood cells (RBC) and granulocytes. The central, lymphocyte-rich layer made up of peripheral blood mononuclear cells (PBMCs) was harvested into 10mL tubes containing 6mL of RPMI1640. PBMCs were pelleted by centrifugation at 244 x g for 10 minutes with brake. The supernatant was aspirated using a vacuum pipette and the PBMCs were again pelleted by centrifugation at 244 x g for 10 minutes with brake. If RBC contamination was apparent, subsequent washes were performed. Following the final wash, PBMCs were resuspended into 5mL of RPMI1640 supplemented with 10% fetal calf serum (FCS). PBMC concentration was determined by counting the cells using a haemocytometer and microscope.

2.1.2 Cryopreservation and thawing of cells

Cryopreservation of cells in a 1ml cryovial routinely consists of a 50/50 ratio of 500µL cells resuspended in RPMI1640 with 10%FCS (R10) and 500µL of cryoprotectant solution, R10 containing 20% dimethyl sulfoxide (DMSO) (Sigma, USA). Tubes containing the required resuspension of cells and appropriate volume of cryoprotectant solution were placed in a beaker of chilled water maintained at 4°C for at least 30 minutes. 500µL of cells was mixed thoroughly with 500µL of the R10/DMSO solution (final DMSO concentration of 10%) and aliquoted into cryovials (Nunc, USA). These cryovials were placed in a Mr Frosty cooling container (Nalgene, USA) and immediately stored in the -80°C freezer. The Mr Frosty container consists of an isopropanol-containing wall that enables the cooling rate of -1°C/minute rate essential for successful cryopreservation. After 24 hours cyrovials of cells may be moved into liquid nitrogen for long term storage.

Following removal from liquid nitrogen tanks, cells were thawed by placing the cryovial in beakers of heated water (37°C). These thawed cells were transferred into 10mL tubes containing 9mL of warmed R10. Cells were centrifuged at 244 x g for 10 minutes and resuspended in fresh R10 to ensure residual DMSO removal.

2.1.3 Generating Phytohaemagglutinin blasts

A mitogen for T cells, phytohaemagglutinin (PHA), rapidly causes T cells to divide and remain viable for 6-8 weeks. Leucoagglutinin (PHA-L) was added at a concentration of 10μg/mL to $2x10^6$ PBMCs in 10ml of R10, aliquoted into a T75 flask (Nunc, Denmark). On day 4, T cell media (TCM) was added to the culture. TCM consists of RPMI1640 supplemented with 10% FCS, 30% T cell growth factor (TCGF) (supernatant from Gibbon lymphosarcoma T cell line, MLA-144) and 20 units/mL of interleukin-2 (IL-2) (Hoffman-LaRoche, USA). The PHA blasts were maintained with addition of fresh TCM bi-weekly.

2.1.4 Generating Lymphoblastoid Cell Lines

Lymphoblastoid cell lines (LCLs) were generated by the exogenous transformation of B cells with EBV derived from B95-8 or QIMR-Wil cell lines. $2x10^6$ PBMCs were pelleted in a 10mL tube and 1mL of filtered (0.65μM) supernatant from the B95-8 cell line was added and lightly vortexed. This was then incubated for 1 hour at 37°C. A final volume of 10mL R10 was added to the tube and the cells were centrifuged at 244 x g for 10 minutes. The cell pellet was resuspended in 5mL CSA media, consisting of RPMI1640 medium with 10% FCS and cyclosporin A (0.1 mg/ml) (Sigma-Aldrich, USA). Dilutions of these cells were added to 6 wells of a 24 well tissue culture plate. 2mls was added to the first well, 1ml to the second well, adding a lesser volume of cells across the remaining 4 wells with the last well containing 300μL of cells. All wells were made up to a final volume of 2mLs with CSA media and incubated in a 37°C, 5%CO_2/95% air humidified incubator. Bi-weekly addition of CSA media for 4 weeks enabled EBV-infected, transformed B cells to divide and establish into an immortal lymphoblastoid cell line. Mature LCLs can be subcultured indefinitely with R10 media.

Spontaneous LCLs were also established without the exogenous addition of EBV. These LCLs arise from the spontaneous outgrowth from PBMCs of EBV sero-positive donors. This involved culturing the PBMCs in 96 well flat-bottom plates at various cell densities for 2 months in CSA media (RPMI1640 medium with 10% FCS and cyclosporin A (0.1 mg/ml).

2.1.5 Generating LCL-stimulated T cell lines

EBV-specific CTL cultures were raised by culturing PBMCs ($2x10^6$/2mL well) with autologous lymphoblastoid cell lines (LCLs) ($2x10^5$/2mL well) (responder/stimulator = 10:1). LCLs were γ-irradiated (8000 rad) in 10mL tubes. R10 media was added to the tubes of irradiated LCLs to a final volume of 10mL and centrifuged at 244 x g for 10 minutes. The responders and stimulators were combined at a 10:1 ratio in 24 well plates and incubated for 3 days. Cultures were

supplemented with TCM on day 3, split on day 7, and subcultured for up to 18 days. There was no HLA preference in the production of CTL cultures, only EBV sero-positive status.

2.1.6 Generating peptide-stimulated T cell lines

CTL bulk cultures were raised by culturing PBMCs ($2x10^6$/2mL well) with autologous PBMCs that were precoated with the required peptide (0.1μM for 1 hour, responder/stimulator = 2:1). The peptide-coated stimulators were washed 3 times (final volume of 10mL R10, centrifuged at 244 x g for 10 minutes, supernatant aspirated) following the one hour incubation to remove unbound peptide. The responders and stimulators were combined in 24 well plates and incubated for 3 days. Cultures were supplemented with TCM on day 3, split on day 7 and subcultured for up to 18 days.

2.2 Cell surface staining and functional assays

2.2.1 Flow cytometric multimer staining

PBMCs or CTLs were stained with pHLA multimers by incubation for 30 minutes at 4°C with an allophycocyanin-labelled (APC) HLA-Cw*06-CRAKFKQLL (National Institute of Health, USA) or APC-labelled HLA-B*1801-SELEIKRY or a phycoerythrin-labelled (PE) HLA-B*4403-EECDSELEIKRY multimer (Manufactured by collaborator Lucy Sullivan, University of Melbourne, Australia) (Briefly, recombinant HLA B1801 or HLA B4403 and human b2 microglobulin, produced in Escherichia coli, were solubilised in urea and injected together with SELEIKRY or EECDSELEIKRY, respectively into a refolding buffer consisting of 100mM Tris (pH 8.0), 400mM arginine, 2mM EDTA, 5mM reduced glutathione, and 0.5mM oxidized glutathione. Refolded complexes were purified by anion exchange chromatography using DE52 resin (Whatman, Tewksbury, USA) followed by gel filtration through a Superdex 75 column (Amersham Pharmacia Biotech, Piscataway, USA). The refolded HLA-B1801-SELEIKRY and HLA-B4403-EECDSELEIKRY complexes were biotinylated by incubation for 16 hours at 30°C with the BirA enzyme (Avidity, USA). Multimeric HLA-peptide complexes were produced by the stepwise addition of extravidin-conjugated PE (Sigma, USA) or APC (Sigma, USA) to achieve a 1:4 molar ratio (extravidin-PE or APC:biotinylated class I)). Cells were then washed and incubated with peridinin-chlorophyll protein with a cyanine dye (PerCP-Cy5.5) –conjugated anti-human CD8 mAb (BioLegend), PE with a cyanine dye (PE-Cy7)-conjugated anti-human CD3 mAb (eBioscience) and Alexa Fluor 700-conjugated anti-human CD4 (BD Pharmingen) for 30 minutes at 4°C. Cells were washed twice in FACS buffer (2% FCS in phosphate buffered saline (PBS)), centrifuging at 1000 x g for 2 minutes and pouring off the supernatant between washes. These

samples were then analysed on a FACS LSR Fortessa flow cytometer using FACSDiva software (BD Biosciences). If the samples were not analysed immediately, they were fixed in 1% paraformaldehyde (PFA) in PBS.

2.2.2 Intracellular Cytokine Staining (ICS)

PBMCs or T cells were incubated for 4-5 hours at 37°C with various synthetic peptides (1μg/ml) in R10 supplemented with 5μg/ml Brefeldin A (BioLegend). These cells were then washed and incubated with PerCP-Cy5.5-conjugated anti-CD8 and Alexa flour 488-conjugated anti-human CD4 at 4°C for 30 minutes. Cells were washed, then fixed and permeabilised with Cytofix/Cytoperm (BD Pharmingen) at 4°C for 20 minutes. Cells were then washed in perm/wash (BD Pharmingen), incubated with PE-conjugated anti-Interferon (IFN)-γ (BD Pharmingen) at 4°C for 30 minutes, washed with perm/wash, resuspended in PBS and analysed on a FACS LSR Fortessa (BD Biosciences).

2.2.3 Flow cytometric analysis of TRBV usage

PBMCs or T cell lines (LCL-stimulated and/or peptide-stimulated) were incubated for 30 minutes at 4°C with an SELEIKRY-HLA-B*1801 allophycocyanin (APC-labelled) multimer (Melbourne). Cells were then washed and incubated with peridinin-chlorophyll protein-Cy5.5 (PerCP-Cy5.5)-labelled anti-human CD8 mAb (BioLegend), and one of the following phycoerythrin (PE) or fluorescein isothiocyanate (FITC)-labelled TCR β-chain-specific mAbs (Beckman Coulter): Vβ1 (TRBV9), Vβ2 (TRBV20-1), Vβ3 (TRBV28), Vβ4 (TRBV29), Vβ5.1 (TRBV5-1), Vβ5.2 (TRBV5-6), Vβ5.3 (TRBV5-5), Vβ6.7 (TRBV7-1), Vβ7 (TRBV4), Vβ7.2 (TRBV4-3), Vβ8 (TRBV12), Vβ9 (TRBV3), Vβ11 (TRBV25-1), Vβ12 (TRBV10), Vβ13.1 (TRBV6-5), Vβ13.2 (TRBV6-2), Vβ13.6 (TRBV6-6), Vβ14 (TRBV27), Vβ16 (TRBV14), Vβ17 (TRBV19), Vβ18 (TRBV18), Vβ20 (TRBV30), Vβ21.3 (TRBV11-1), Vβ22 (TRBV2) or Vβ23 (TRBV13). Cells were washed twice in FACS buffer (2% FCS in phosphate buffered saline (PBS)), centrifuging and tipping off the supernatant between washes. These samples were analysed on a FACS LSR Fortessa flow cytometer using FACSDiva software (BD Biosciences). If samples were not analysed immediately, they were fixed in 1% paraformaldehyde (PFA) in PBS.

2.2.4 Flow cytometric analysis of TRBV usage with ICS

In the absence of a multimer for TRBV usage assays, an ICS was required with the desired synthetic peptide followed by TRBV usage assay. PBMCs or T cells were incubated for 4-5 hours at 37°C with various synthetic peptides (1μg/ml) in R10 supplemented with 5μg/ml Brefeldin A (BioLegend). Cells were centrifuged at 1000g for 2 minutes and supernatant discarded. Cells were

resuspended in 25µl PBS and 1µl of one of the following phycoerythrin (PE) or fluorescein isothiocyanate (FITC)-labelled TCR β-chain-specific mAbs (Beckman Coulter): Vβ1 (TRBV9), Vβ2 (TRBV20-1), Vβ3 (TRBV28), Vβ4 (TRBV29), Vβ5.1 (TRBV5-1), Vβ5.2 (TRBV5-6), Vβ5.3 (TRBV5-5), Vβ6.7 (TRBV7-1), Vβ7 (TRBV4), Vβ7.2 (TRBV4-3), Vβ8 (TRBV12), Vβ9 (TRBV3), Vβ11 (TRBV25-1), Vβ12 (TRBV10), Vβ13.1 (TRBV6-5), Vβ13.2 (TRBV6-2), Vβ13.6 (TRBV6-6), Vβ14 (TRBV27), Vβ16 (TRBV14), Vβ17 (TRBV19), Vβ18 (TRBV18), Vβ20 (TRBV30), Vβ21.3 (TRBV11-1), Vβ22 (TRBV2) or Vβ23 (TRBV13) and incubated at room temperature for 20 minutes. Cells were then incubated with PerCP-Cy5.5-conjugated anti-CD8 mAb at 4°C for 20 minutes. Cells were washed, then fixed and permeabilised with Cytofix/Cytoperm (BD Pharmingen) at 4°C for 20 minutes. Cells were then washed in perm/wash (BD Pharmingen), incubated with APC-conjugated anti-Interferon (IFN)-γ (BD Pharmingen) at 4°C for 30 minutes, washed with perm/wash, resuspended in PBS and analysed on a FACS LSR Fortessa (BD Biosciences).

2.2.5 Interferon-γ ELISpot assay

Interferon (IFN)-γ ELISpot assays were performed using cytokine capture and detection reagents according to manufacturer's instructions (ELISpot ^{PRO} for Human IFN-γ, Mabtech). Briefly, 96-well nitrocellulose plates pre-coated with anti-IFN-γ Abs were seeded with approximately 50 000 autologous LCL T cells or PBMCs or peptide-stimulated CTL cultures and the appropriate peptide at various concentrations (10µM and 1µM or 10-fold dilutions from 10µM-0.000001µM). After incubation for 16 hours at 37°C in 5% CO_2, the cells were discarded and captured IFN-γ was detected with a biotinylated anti-IFNγ Ab, followed by development with an alkaline phosphatase substrate solution (BCIP/NBT-plus). Spots were counted using an automated plate counter (AID).

2.3 Flow cytometry reagents

2.3.1 Antibodies

Recognises	Conjugate	Isotype	Species	Raised in	Clone	Company
CD3	PE-Cy7	IgG1	Human	Mouse	SK7	BD Pharmingen
CD4	Alexa Flour 488	IgG1	Human	Mouse	RPA-T4	BioLegend
CD4	Alexa Fluor 700	IgG1	Human	Mouse	RPA-T4	BD Pharmingen
CD4	FITC	IgG1	Human	Mouse	RPA-T4	BD Pharmingen
CD4	PE	IgG1	Human	Mouse	L120	BD Pharmingen
CD8	PerCP-Cy5.5	IgG1	Human	Mouse	RPA-T8	BioLegend
CD14	Pacific Blue	IgG1	Human	Mouse	M5E2	BD Pharmingen
CD19	eFluor 450 (Pacific Blue)	IgG1	Human	Mouse	HIB19	eBioscience
IFN-γ	APC	IgG1	Human	Mouse	B27	BD Pharmingen
IFN-γ	PE	IgG1	Human	Mouse	B27	BD Pharmingen

2.3.2 Multimers

Sequence	HLA Restriction	Conjugate	Company
CRAKFKQLL	Cw*0602	APC	NIH
EECDSELEIKRY	B*4403	PE	Melbourne
SELEIKRY	B*1801	APC	Melbourne

Chapter 3:

Aim One:
Determine the relative prevalence of peptide epitopes of over 10 amino acids in length in $CD8^+$ T cell recognition of EBV.

Hypothesis One:
Peptide epitopes of over 10 amino acids in length are abundant in antiviral $CD8^+$ T cell responses.

3.1 Introduction

Class I major histocompatibility complex (MHC) molecules form medleys with peptide antigens which are expressed on the cell surface for recognition by $CD8^+$ T cells. Derived from antigens synthesized in the cytoplasm, these peptides are generally 8-10 amino acids in length. For most MHC alleles two of the pockets within the peptide binding groove display a marked preference for one or two amino acids at certain anchor positions within the peptide (Falk, Rotzschke et al. 1991). This was the breakthrough discovery that enabled more efficient CTL epitope mapping. Dependent on this information, web-based algorithms used to predict $CD8^+$ T cell epitopes were designed to include peptides limited to between 8 and 10 residues. The apparent dominance of MHC class I-presented epitopes of 8 to 10 amino acids in length may be misleading and result from this bias of widely used algorithms. Indeed, many predicted epitopes have been shown to be nonimmunogenic failing to be processed and presented on APCs or as a result of limitations in the TCR (Eisenlohr, Yewdell et al. 1992).

It has been demonstrated by several groups investigating viruses as well as melanoma antigens that peptides between 11-14 amino acids in length can bind to particular MHC class I molecules (Burrows, Rossjohn et al. 2006). Of the 27 $CD8^+$ T cell epitopes greater than 10 residues in length that have been reported in viruses including HIV (Sipsas, Kalams et al. 1997), HCMV (Weekes, Wills et al. 1999), and hepatitis B virus (Schirmbeck, Melber et al. 1994), seven unusually long $CD8^+$ T cell epitopes are derived from EBV antigens (Blake, Lee et al. 1997, Rickinson and Moss 1997, Green, Miles et al. 2004, Pudney, Leese et al. 2005, Burrows, Bell et al. 2008, Rist, Theodossis et al. 2013). The most dramatic example is the 14mer epitope, ^{4}LPAVVGLSPGEQEY^{17}, derived from an alternative open reading frame of the macrophage colony-stimulating factor gene. This extreme epitope is naturally presented by HLA-B*3501 and is recognised by tumour-infiltrating $CD8^+$ T cells (Probst-Kepper, Stroobant et al. 2001, Probst-Kepper, Hecht et al. 2004).

Several MHC class I bound, immunogenic peptides that are longer than ten residues have been characterized structurally and shown that, in each case, the peptide adopts a bulging confirmation whereby the extra length "inserts" between the primary anchor residues (Tynan, Borg et al. 2005, Miles, Borg et al. 2006). For example, the highly immunogenic 13-mer epitope ^{52}LPEPLPQGQLTAY^{64} (LPEP) from the BZLF1 antigen of EBV that is the target of a $CD8^+$ T cell response by HLA-B*3508$^+$ EBV$^+$ individuals (Green, Miles et al. 2004). Another example of a noncanonical epitope is the 11-mer ^{407}HPVGEADYFEY^{417}(HPV), derived from the EBNA1 protein of EBV (Miles, Borg et al. 2006). In addition, our recent findings of the BZLF1 dodecamer ^{169}EECDSELEIKRY^{180}, restricted to HLA-B*4403, also showed a bulged conformation (Rist,

Theodossis et al. 2013). Both LPEP and HPV share another characteristic in that these epitopes completely overlap with shorter sequence peptides that conform to the binding motif of their respective HLA allele and score well using predictive web-based algorithms but have proven to be nonimmunogenic. These unusually long CTL epitopes not only bind MHC class I molecules sufficiently to stimulate $CD8^+$ T cell responses they also show a superior immunodominance in comparison to the shorter peptides (Blake, Lee et al. 1997, Green, Miles et al. 2004, Burrows, Rossjohn et al. 2006).

Interestingly, many of the unusually long CTL epitopes are presented in the context of HLA-B*0702, HLA-B*3501 or HLA-B*3508. These HLA alleles are all members of the HLA-B7 supertype family which share primary anchor residues at position 2 and the C-terminus (Burrows, Rossjohn et al. 2006, Burrows, Bell et al. 2008). Burrows and colleagues have previously proposed that members of this supertype present a high number of long peptides due to a unique feature of the class I antigen processing pathway (Burrows, Rossjohn et al. 2006). As a result of the ER-associated aminopeptidases inability to cleave the bond at the N- side of Proline, an accumulation of peptides with x-Pro-x_n sequence is likely to occur (Serwold, Gaw et al. 2001), with many of these peptides likely to be long. As the HLA-B7 supertype has a preference for Prolines at position 2 these alleles are more likely to have unusually long CTL epitopes available for presentation within the ER compared to other HLA alleles that are not members of this supertype (Burrows, Bell et al. 2008). Indeed around half of the defined noncanonical CTL epitopes have a Proline at position 2 and this may continue as a characteristic with newly identified unusually long CTL epitopes.

In determining the presence of noncanonical $CD8^+$ T cell epitopes, the latent herpes virus EBV is a superb model due to the strong responses detected in healthy virus carriers. A number of highly immunogenic antigens of EBV were the focus of this study, including the lytic antigens BRLF1, BMLF1 and BMRF1 and the latent antigens EBNA3A, EBNA3B, EBNA3C and EBNA1. Immediate early (IE), early (E) and late (L) proteins are sequentially expressed during lytic viral replication. Incorporating around 80 EBV proteins the lytic cycle is activated by the expression of two immediate early proteins BZLF1 (*Zta*) and BRLF1 (*Rta*) (Zalani, Holley-Guthrie et al. 1996). Feederle and colleagues have demonstrated that both transactivators, BZLF1 and BRLF1 are required for early and late EBV gene expression as well as for viral DNA expression. The expression of late protein gp350 appears to require BRLF1 but does seem to be directly influenced by BZLF1 (Feederle, Kost et al. 2000). Early protein BMRF1 has proved to be essential for the production of infectious EBV progeny as well as being associated with enhanced expression of BHLF1 (Neuhierl and Delecluse 2006).

During latent infection, EBV genes hijack normal B cell mechanisms leading to the establishment of lymphoblastoid cell lines (LCLs). At least nine latency-associated viral antigens

are expressed by LCLs and include the six nuclear antigens, EBV nuclear antigen 1 (EBNA1), EBNA2, EBNA3A, EBNA3B, EBNA3C and EBNA leader protein (LP) and latent membrane proteins LMP1, LMP2A and LMP2B. When these nine antigens are expressed it is referred to as a latency III form of EBV infection (Kutok and Wang 2006, Rickinson and Kieff 2007). The essential antigen required for the persistence and maintenance of the EBV episome and enhancing genome expression is EBNA1. A unique feature of EBNA1 is the glycine-alanine repeat (GAr) domain which inhibits MHC class I presentation of cis-linked epitopes. Also of great importance, EBNA1 is the only viral protein that is expressed in all forms of latency and in all EBV-associated malignancies (Rickinson and Kieff 2007, Long, Taylor et al. 2011).

Tandemly placed in the EBV genome, EBNA3A, -3B and -3C share a similar gene structure. These three stable proteins can interact with cellular proteins involved in transcription regulation and bind to the cellular repressor-binding protein (RBP)-Jκ (Le Roux, Kerdiles et al. 1994, Young and Rickinson 2004). EBNA3A and EBNA3C are both critical for B-lymphocyte growth transformation and survival. Interestingly, EBNA3B is not essential for latent infection of B lymphocytes and cell survival, although its presence in EBV isolates of healthy carriers indicates that it plays a part in natural infection (Kieff and Rickinson 2007). Maruo and colleagues showed that EBNA3C plays an essential role in maintaining LCL growth and that this role cannot be filled by EBNA3A or EBNA3B (Maruo, Wu et al. 2006).

In order to determine the relative prevalence of peptide epitopes of over 10 amino acids in length in $CD8^+$ T cell recognition of EBV, this study utilised overlapping peptides, not web-based algorithms to predict peptides. This approach removes any bias towards peptide length and proposed HLA allele restriction. Many novel $CD8^+$ T cell epitopes were identified with examples from all EBV antigens investigated, including four unusually long $CD8^+$ T cell epitopes. As the vast majority of $CD8^+$ T cell epitopes that have been defined to date are 8 to 10 residues in length, it is unlikely that the wider immunology community accepts that longer peptides play a significant role. However, this apparent dominance may be misleading and due to the bias of widely used algorithms to predict class-I binding peptides no longer than 10 amino acids in length.

3.2 Results

IFN-γ ELISpot assays were used to test overlapping 20-amino acid (aa) peptides corresponding to the lytic antigens BRLF1, BMLF1 and BMRF1 and the latent antigens EBNA3A, EBNA3B, EBNA3C and EBNA1. These overlapping peptide sets were screened for recognition by T cell lines. These T cell lines were raised by *in vitro* stimulation of PBMCs from 41 healthy EBV-seropositive individuals with their autologous LCLs (Table 1). Results for a number of individuals screened with the overlapping peptide sets of the EBV antigens EBNA3A, BMLF1, BRLF1, EBNA3C, BMRF1, EBNA3B and EBNA1 are shown in Fig. 1 A to G, respectively.

TABLE 1 Class I HLA types of donors used in this study

Donor ID	HLA-A	HLA-B	HLA-C
B33	24, 29	1801, 4403	5, 16
B9	1, 30	8, 18	5, 7
JC33	3, 25	18, 4403	Unknown
DP55	2	15, 18	1, 5
CG59	2, 3	7, 18	7, 12
B16	2, 11	15, 44	3, 5
HB49	3, 32	7, 8	7
MR67	2	15, 57	3, 6
AO83	2, 24	18, 44	5, 7
B28	2, 23	4402, 53	5, 6
B20	3, 24	7	7
B5	11, 24	15, 35	3, 4
B22	3, 23	41, 47	6, 17
B2	1, 11	40, 52	3, 12
B3	1, 3	8, 40	2, 7
B39	26, 68	27, 4402	2, 5
DM42	24, 29	4403, 4405	2
KK68	1, 29	8, 4403	7, 16
B4	26, 32	40, 4402	1, 3
B17	1, 2	7, 8	7
B21	32, 68	7, 14	7, 8
B23	3, 11	35, 4402	3, 5
B36	3, 24	7, 4402	5, 7
B32	2, 68	7, 39	7
B34	1, 32	7	7
B8	2, 66	4402, 49	7
B15	1, 26	27, 62	1, 3
B25	1, 2	35, 51	4, 14
B26	2, 11	35, 58	4, 7
MW65	1, 3	8, 3501	4
LC50	1	8, 18	7
SB60	2	35, 57	4, 6
B7	2, 3	7, 37	6, 7
B13	1, 2	27, 57	2, 6
B38	2, 3	7, 57	6, 7
B10	11, 31	7, 51	7, 15
B11	11, 24	50, 62	4, 6
B14	2, 11	51	15, 16
B19	1, 3	8, 4403	4, 7
B21	32, 68	7, 14	7, 8
B24	2	40	3

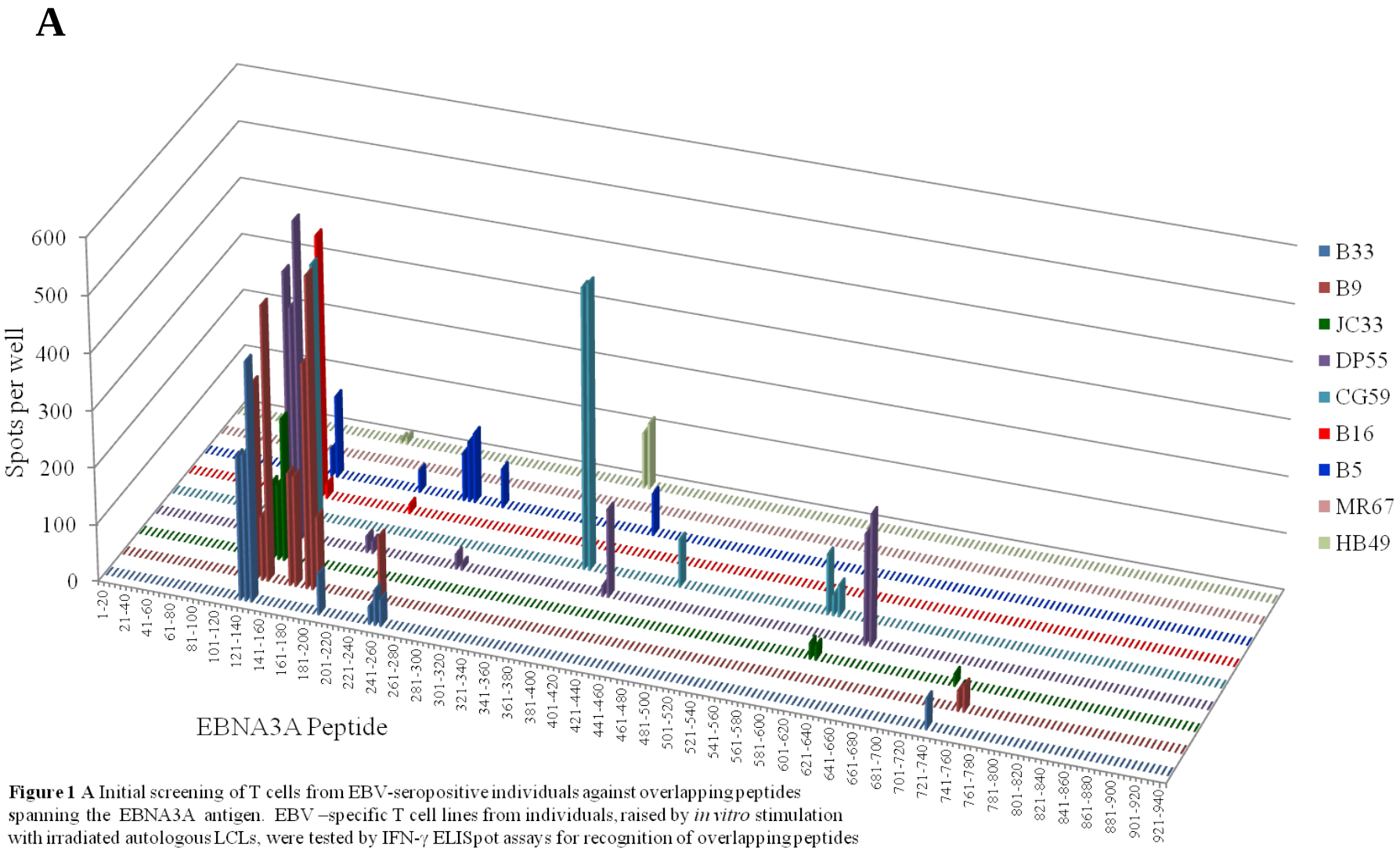


Figure 1 A Initial screening of T cells from EBV-seropositive individuals against overlapping peptides spanning the EBNA3A antigen. EBV –specific T cell lines from individuals, raised by *in vitro* stimulation with irradiated autologous LCLs, were tested by IFN-γ ELISpot assays for recognition of overlapping peptides (10μg/ml) corresponding to the EBNA3A antigen.

B

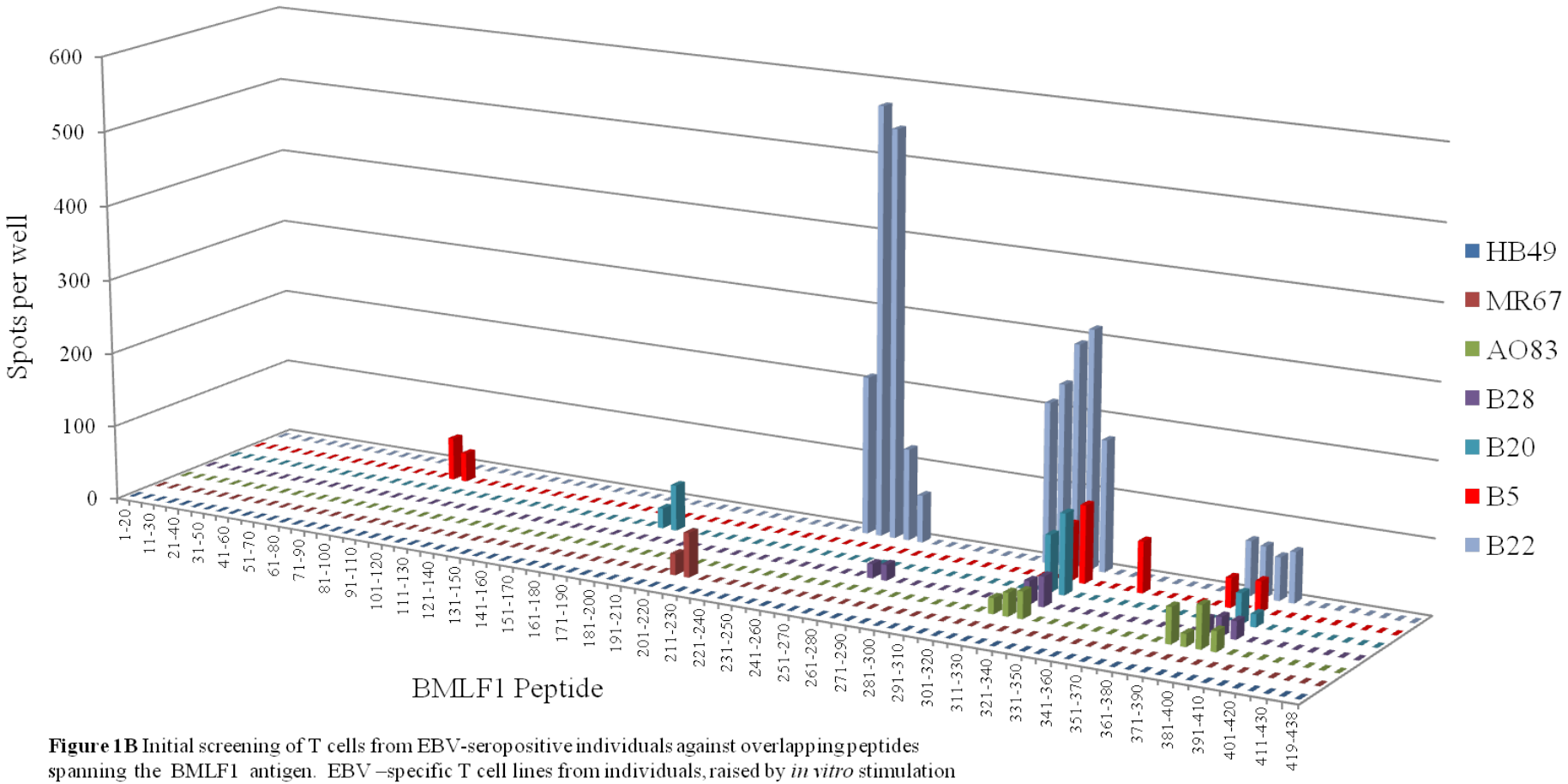


Figure 1B Initial screening of T cells from EBV-seropositive individuals against overlapping peptides spanning the BMLF1 antigen. EBV –specific T cell lines from individuals, raised by *in vitro* stimulation with irradiated autologous LCLs, were tested by IFN-γ ELISpot assays for recognition of overlapping peptides (10μg/ml) corresponding to the BMLF1 antigen.

C

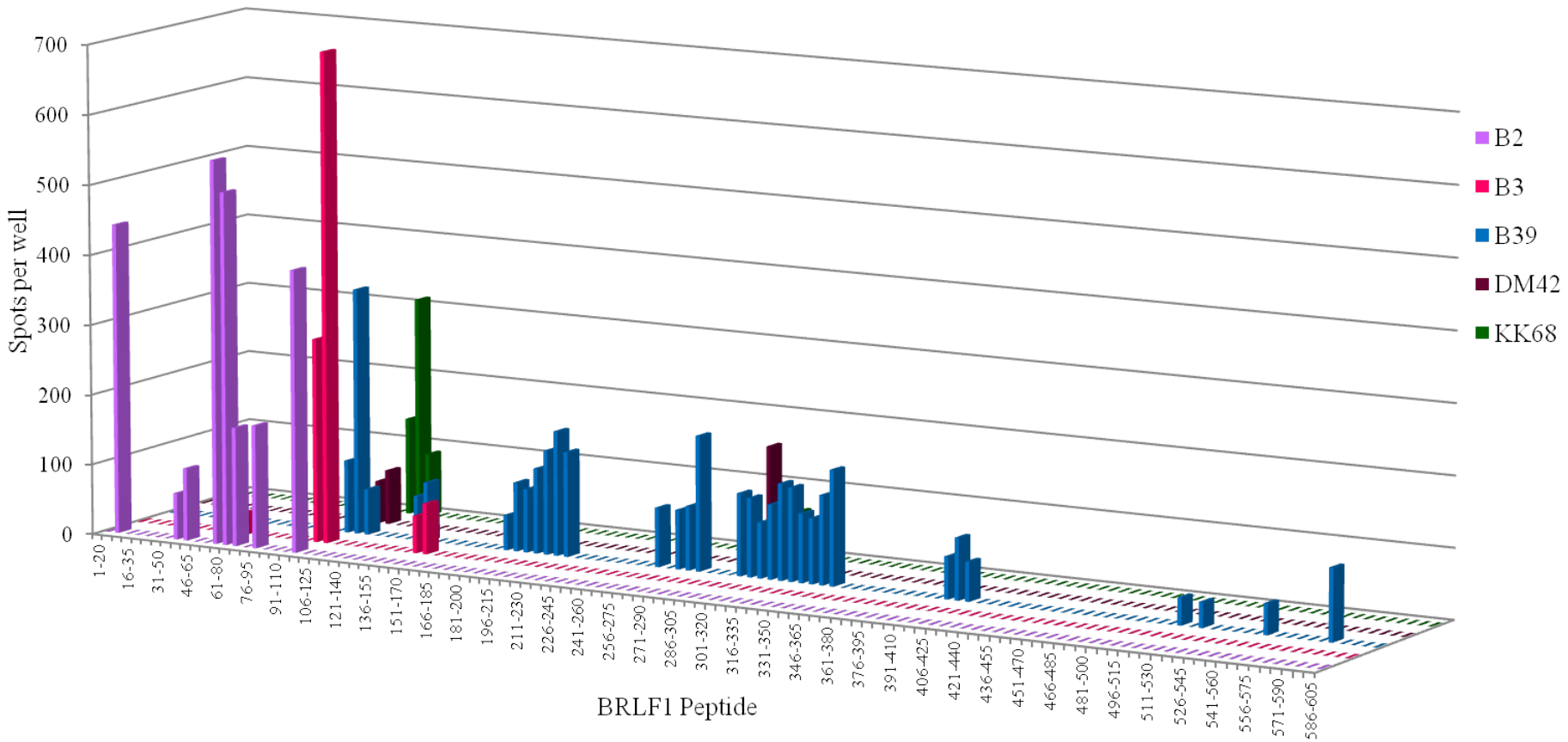


Figure 1C Initial screening of T cells from EBV-seropositive individuals against overlapping peptides spanning the BRLF1 antigen. EBV –specific T cell lines from individuals, raised by *in vitro* stimulation with irradiated autologous LCLs, were tested by IFN-γ ELISpot assays for recognition of overlapping peptides (10μg/ml) corresponding to the BRLF1 antigen.

D

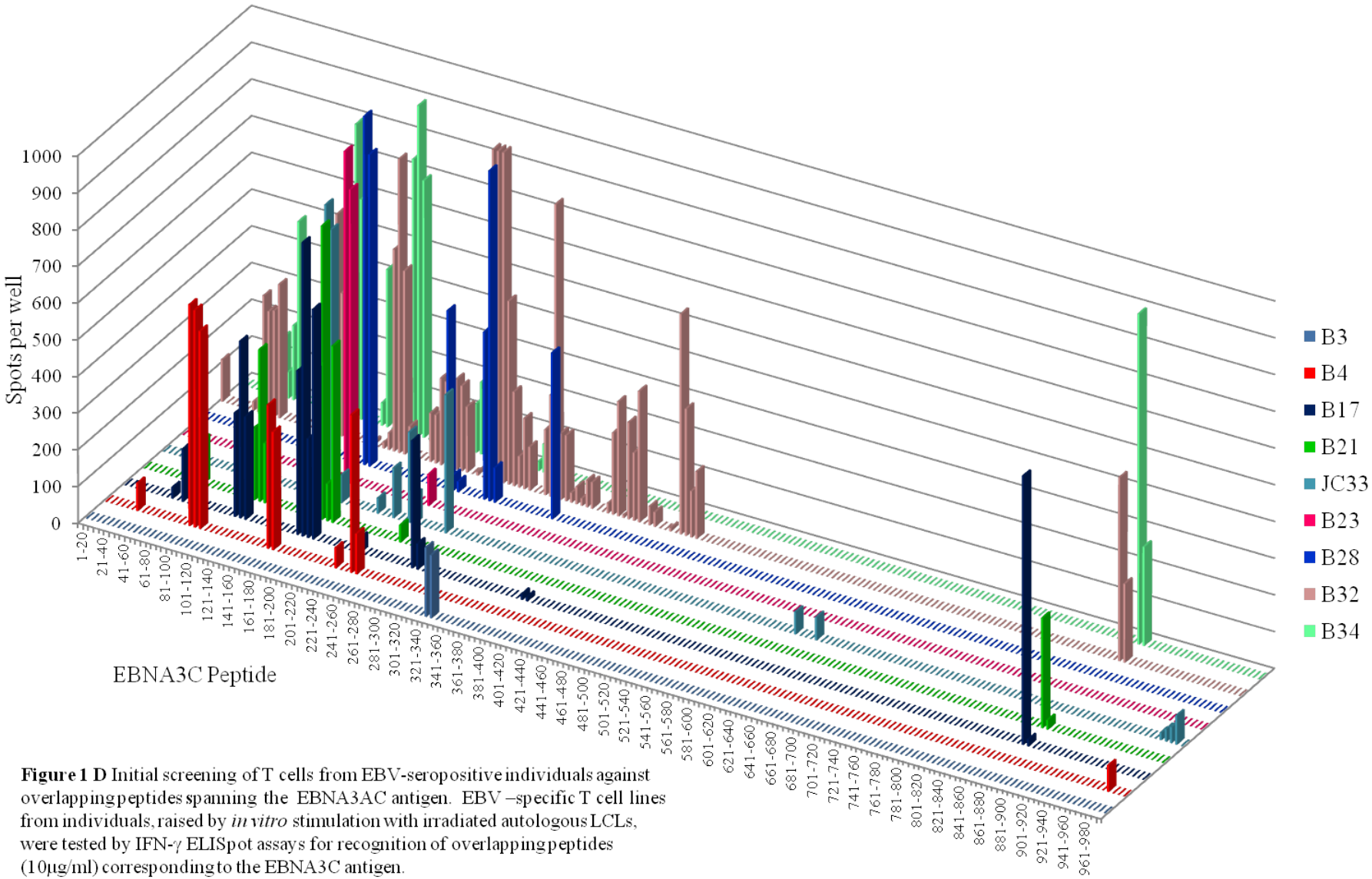


Figure 1 D Initial screening of T cells from EBV-seropositive individuals against overlapping peptides spanning the EBNA3AC antigen. EBV –specific T cell lines from individuals, raised by *in vitro* stimulation with irradiated autologous LCLs, were tested by IFN-γ ELISpot assays for recognition of overlapping peptides (10μg/ml) corresponding to the EBNA3C antigen.

E

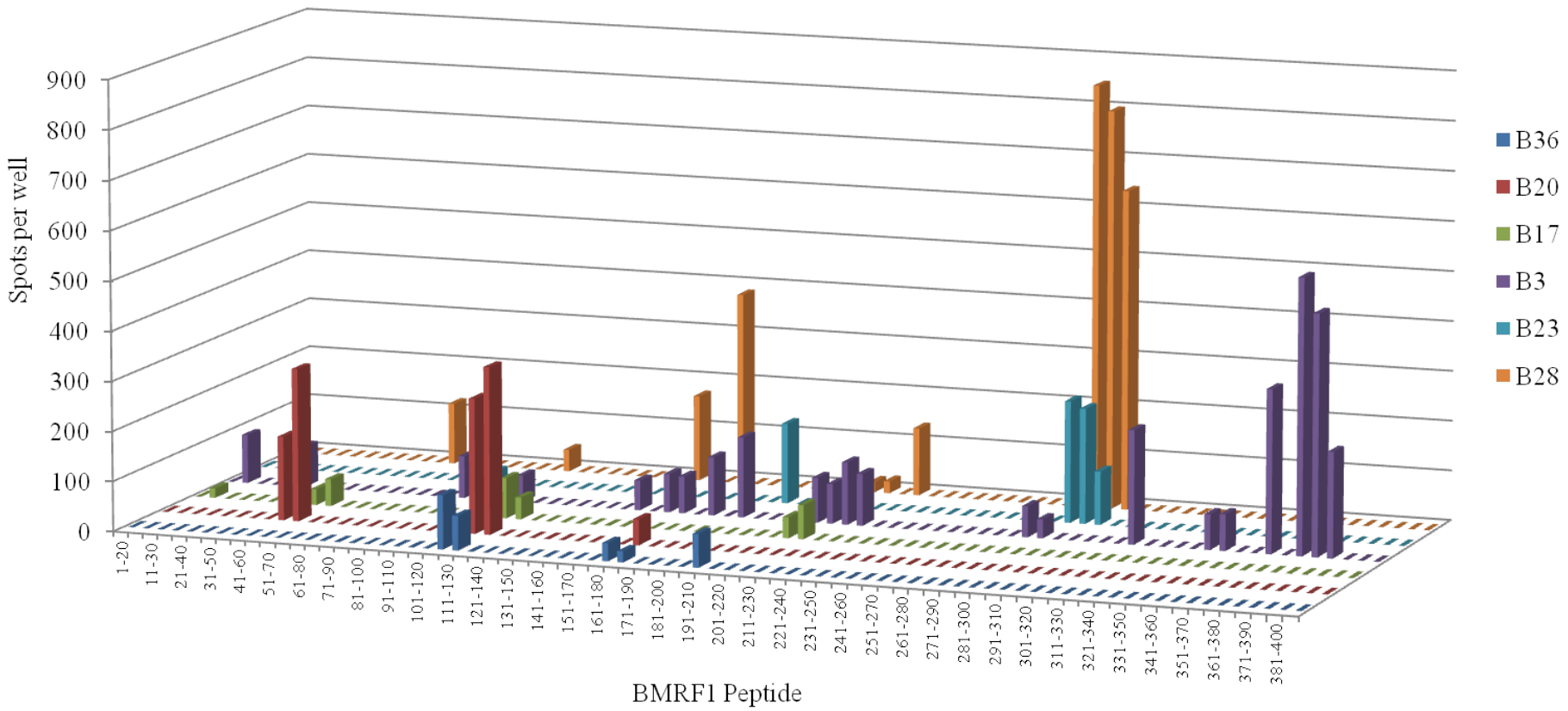


Figure 1 E Initial screening of T cells from EBV-seropositive individuals against overlapping peptides spanning the BMRF1 antigen. EBV –specific T cell lines from individuals, raised by *in vitro* stimulation with irradiated autologous LCLs, were tested by IFN-γ ELISpot assays for recognition of overlapping peptides (10μg/ml) corresponding to the BMRF1 antigen.

F

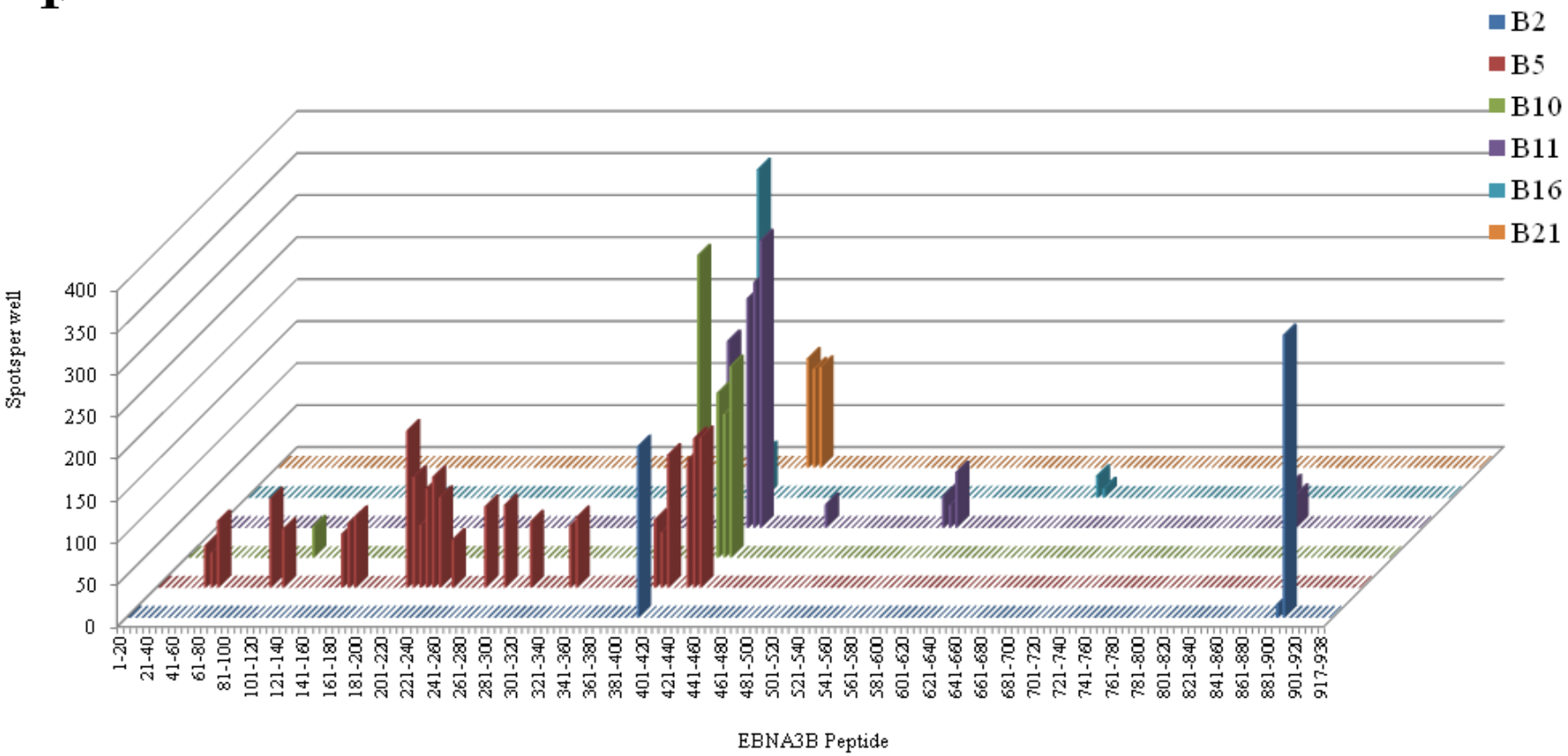


Figure 1 F Initial screening of T cells from EBV-seropositive individuals against overlapping peptides spanning the EBNA3B antigen. EBV –specific T cell lines from individuals, raised by *in vitro* stimulation with irradiated autologous LCLs, were tested by IFN-γ ELISpot assays for recognition of overlapping peptides (10 μg/ml) corresponding to the EBNA3B antigen.

G

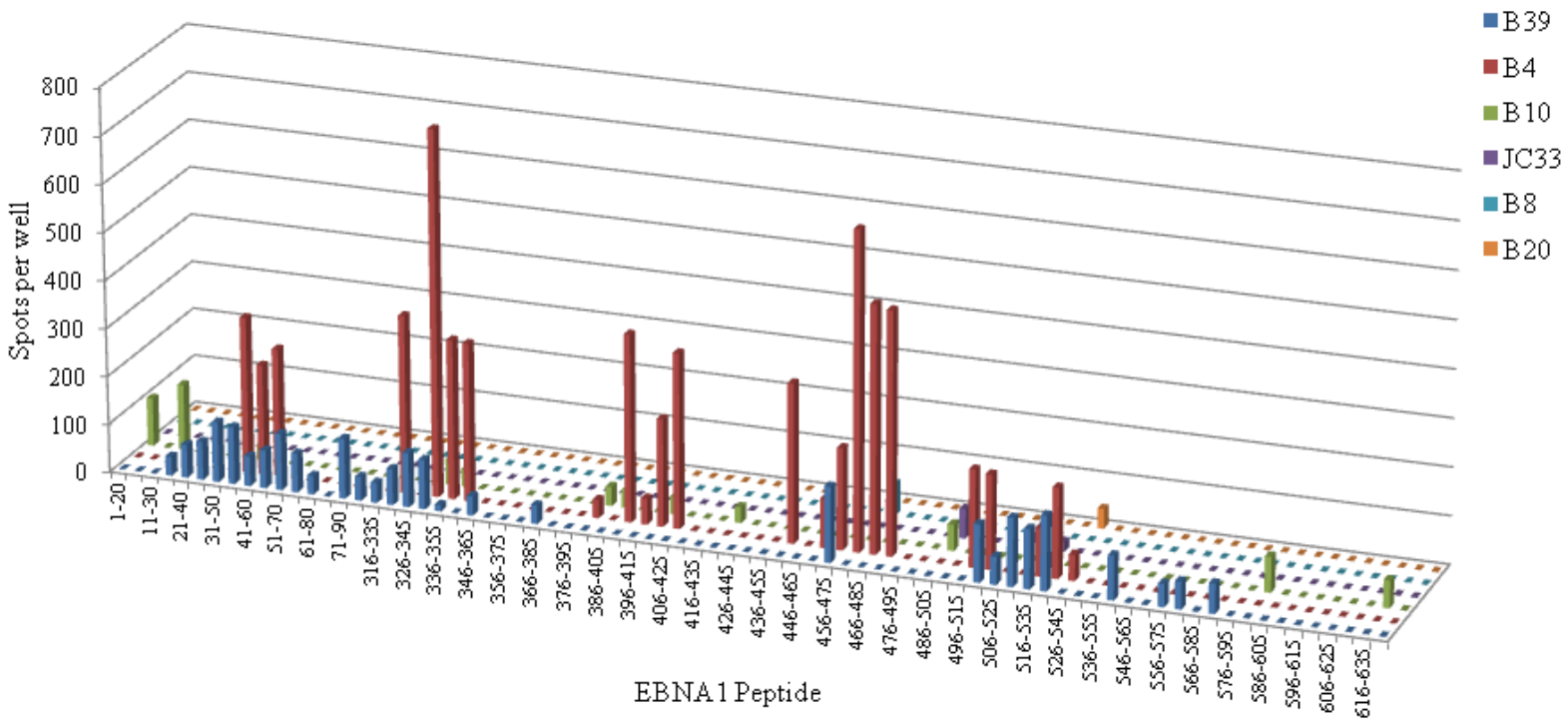


Figure 1 G Initial screening of T cells from EBV-seropositive individuals against overlapping peptides spanning the EBNA1 antigen. EBV –specific T cell lines from individuals, raised by *in vitro* stimulation with irradiated autologous LCLs, were tested by IFN-γ ELISpot assays for recognition of overlapping peptides (10 μg/ml) corresponding to the EBNA1 antigen. Note: GAr residues 91-336 are omitted.

With nearly 1000 peptides screened for each of the 41 healthy EBV carriers it is not possible to show the results of each individual to all peptides. Figure 1 A to G shows examples of results with these overlapping peptides for 25 of the 41 individuals screened (Table 1). The results are indicative of those attained for all of the EBV antigens tested, with a portion of the peptides eliciting an IFN-γ response whilst the majority of screened peptides gave a negative result. Collated from the results of all individuals tested, is the list of novel epitopes defined in this study (Table 2).

The greatest number of epitopes defined in the antigens investigated was 12 novel epitopes within EBNA3A. There was a degree of epitope clustering and overlapping in this antigen (Table 2 & Fig.1A). Depicted in Fig. 1A are the initial results for the three highly immunogenic overlapping peptides ^{116}DQFFGSQISNTEMYIMYAMA^{135}, ^{121}SQISNTEMYIMYAMAIRQAI^{140} and ^{126}TEMYIMYAMAIRQAIRDRRR^{145}. All of the responders were HLA-B*18 positive and all had the strongest IFN-γ response to peptide 126-145. HLA-B*18 negative donors (MR67 and HB49) displayed no IFN-γ response to these peptides (Fig. 1A). The peptides TEMYIMYAMA, TEMYIMYAM and TEMYIMYA, which have sequences included within these 20mer peptides, were screened to map the EBNA3A epitope. The peptides were titrated from 10μg/ml to 0.000001μg/ml and tested with HLA-B18^{+} donors B9 and B33 (Fig. 2A and 2B, respectively). The nonamer ^{131}TEMYIMYAM^{139} was determined to be the minimal epitope titrating out to a final peptide concentration of 0.0001μg/ml with B9 and 0.000001μg/ml with B33. It is notable that this nonamer peptide conforms to the HLA-B*1801 binding motif (Hillen, Mester et al. 2008).

Another EBNA3A epitope which overlaps TEMYIMYAM is ^{126}SQISNTEMY^{134}. Figure 2C displays results from donor B16 and peptides FGSQISNTEMY, SQISNTEMY and QISNTEMY tested with IFN-γ ELISpot titrations from 10μg/ml to 0.000001μg/ml, showing SQISNTEMY and QISNTEMY as equally active epitopes. The nonamer conforms to the HLA-B*1501 binding motif with a Glutamine at position 2 and Tyrosine at the C terminus (Falk, Rotzschke et al. 1995). However, it cannot be ruled out that both the nonamer SQISNTEMY and octamer QISNTEMY are epitopes in this donor as the dose response curves for these two peptides are very similar.

Figure 2D demonstrates results indicating that another novel epitope from EBNA3A is ^{34}WENVLIEL^{41}. Donors B3 and B4 show significant IFN-γ responses to a number of peptides in this truncated peptide set. All peptides that included the nonamer sequence ^{34}WENVLIEL^{41} were recognized by the T cells. As HLA-B*40 is shared by both donors B3 and B4 and the nonamer conforms to the published binding motif of HLA-B*40 with anchor residues Glutamic acid at position 2 and Leucine at the C-terminus (Falk, Rotzschke et al. 1995) it is likely that this epitope is restricted by HLA-B*40.

TABLE 2 Summary of Novel EBV $CD8^+$ T cell epitopes defined in this study

Protein	Epitope Coordinates	Epitope Sequence	Epitope Length (AA)	Likely HLA Restriction
Lytic				
IE BRLF1	64-73	REWGGLMATL	10	B40
(605AA)	101-109	IACPIVMRY	9	A29 and Cw2/3
	198-206	TYPVLEEMF	9	A24
	260-268	KAVETPILV	9	B51
	293-301	SESGQFHAF	9	B44
E BMLF1	209-218	VAAHPEIGAW	10	B57/58
(438AA)	244-252	KDTWLDARM	9	B47
	320-328	DYNFVKQLF	9	A24/23
	376-384	DEVEFLGHY	9	Cw6
E BMRF1	116-125	RPQGGSRPEF	10	B7
(404AA)	146-156	MPYMPPASDRL	11	B53
	286-293	LPLDLSVILF	10	B35
	369-377	LERPLAVQL	9	B40
Latent				
EBNA3A	34-41	WENVLIEL	8	B40 (B60/61)
(944AA)	126-134	SQISNTEMY	9	B1501(62)
	131-139	TEMYIMYAM	9	B18
	133-141	MYIMYAMAI	9	A23
	137-145	YAMAIRQAI	9	B51
	170-180	WPMGYQAYSSW	11	B53
	187-196	HQTTPTFVHL	10	B52 or Cw3
	283-290	LQRTDLSY	8	B1501(62)
	509-517	IVRPWEPSL	9	Unknown
	631-640	MEGPLVPEQQ	10	B18
	639-647	QQMFPGAPF	9	B18
	642-650	FPGAPFSQV	9	B51
EBNA3B	168-176	KLLSSCRSW	9	A32
(938AA)	178-186	MGYRTHNLK	9	A3
	198-206	HPVLVTATL	9	B3501
	279-287	VSFIEFVGW	9	B57
	416-424	IVTDFSVIK	9	A68
	888-898	SHSPVVILENV	11	B52
EBNA3C	90-98	TEDNVPPWL	9	B40
(992AA)	160-168	LVAEGGVGW	9	A25
	162-170	AEGGVGWRHW	10	B4402
	180-189	WPMGYRTATL	10	B7
	244-253	AEVRFLRGKW	10	B4402/03
	335-343	KEHVIQNAF	9	B40
	343-351	FRKAQIQGL	9	Cw6
	650-658	WEMRAGREI	9	B4901
	740-750	QPAPQAPYQGY	11	B3501
EBNA1				
(641AA)	499-508	DEGTWVAGVF	10	B18

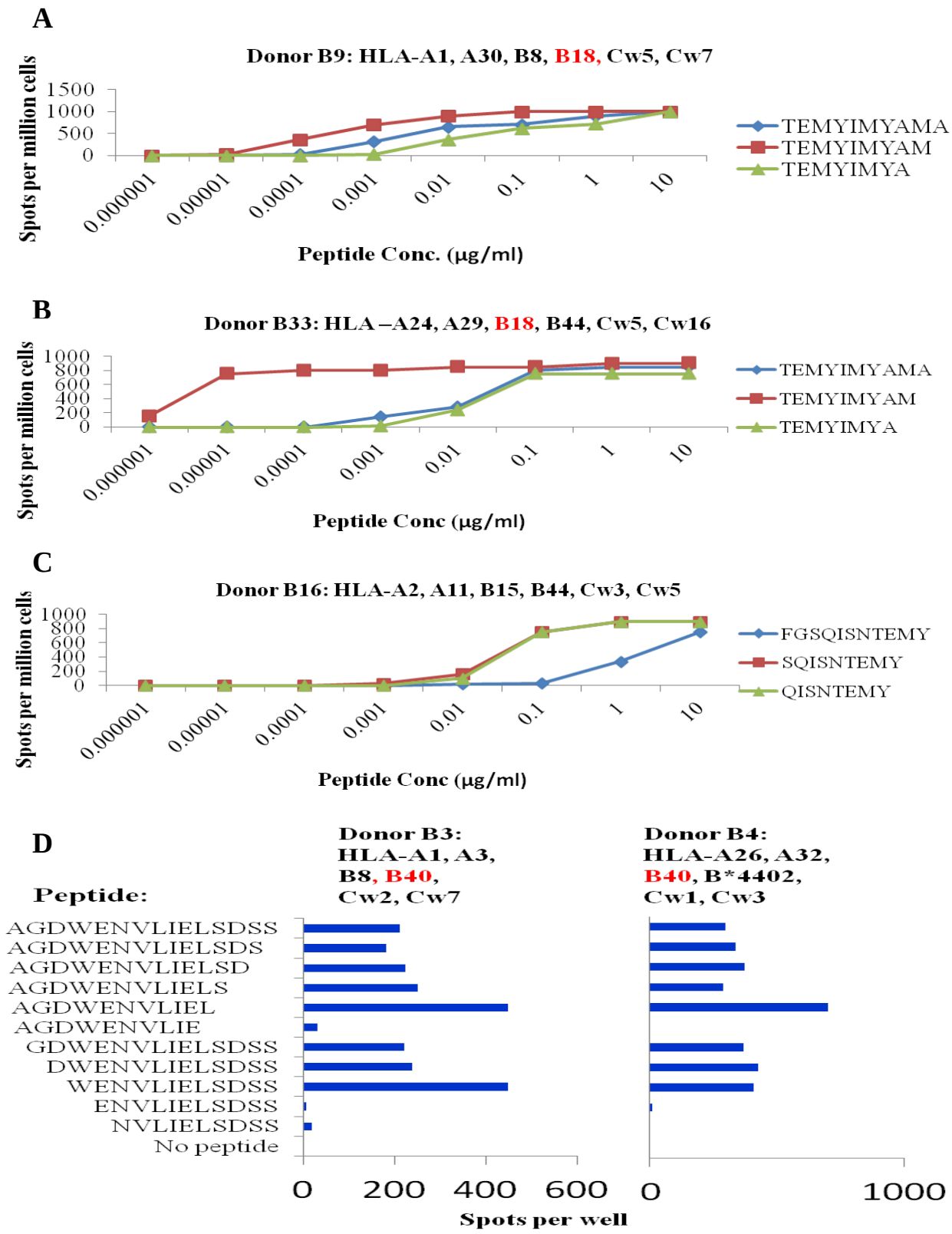


FIGURE 2. Novel epitopes from the latent EBV antigen EBNA3A. EBV-specific T cell lines from donors **(A)** B9, **(B)** B33 and **(C)** B16 were tested with IFN-γ ELISpot assays to precisely map the minimal epitopes. Cells were tested with various concentrations of the indicated peptides. **(D)** EBV-specific T cell lines, raised by *in vitro* stimulation with irradiated autologous LCLs from donors B3 and B4 were tested in IFN-γ ELISpot assays for recognition of overlapping peptides (10µg/ml) corresponding to a region of the EBNA3A protein sequence.

Four novel epitopes were defined from the lytic antigen BMLF1 (Table 2). Abbott *et al* recently noted ^{205}SMLAVAAHPEIGAWQ^{219} as a novel epitope (Abbott, Quinn et al. 2013) and, from results in this study (Fig. 3A), the decamer ^{209}VAAHPEIGAW^{218} appears to be the minimal epitope with HLA-B*57/58 the likely restricting allele, based on the published motif of these two alleles (Falk, Rotzschke et al. 1995). Concurring with the findings of Abbott *et al*, the responding donor in those studies was also HLA-B*58 positive.

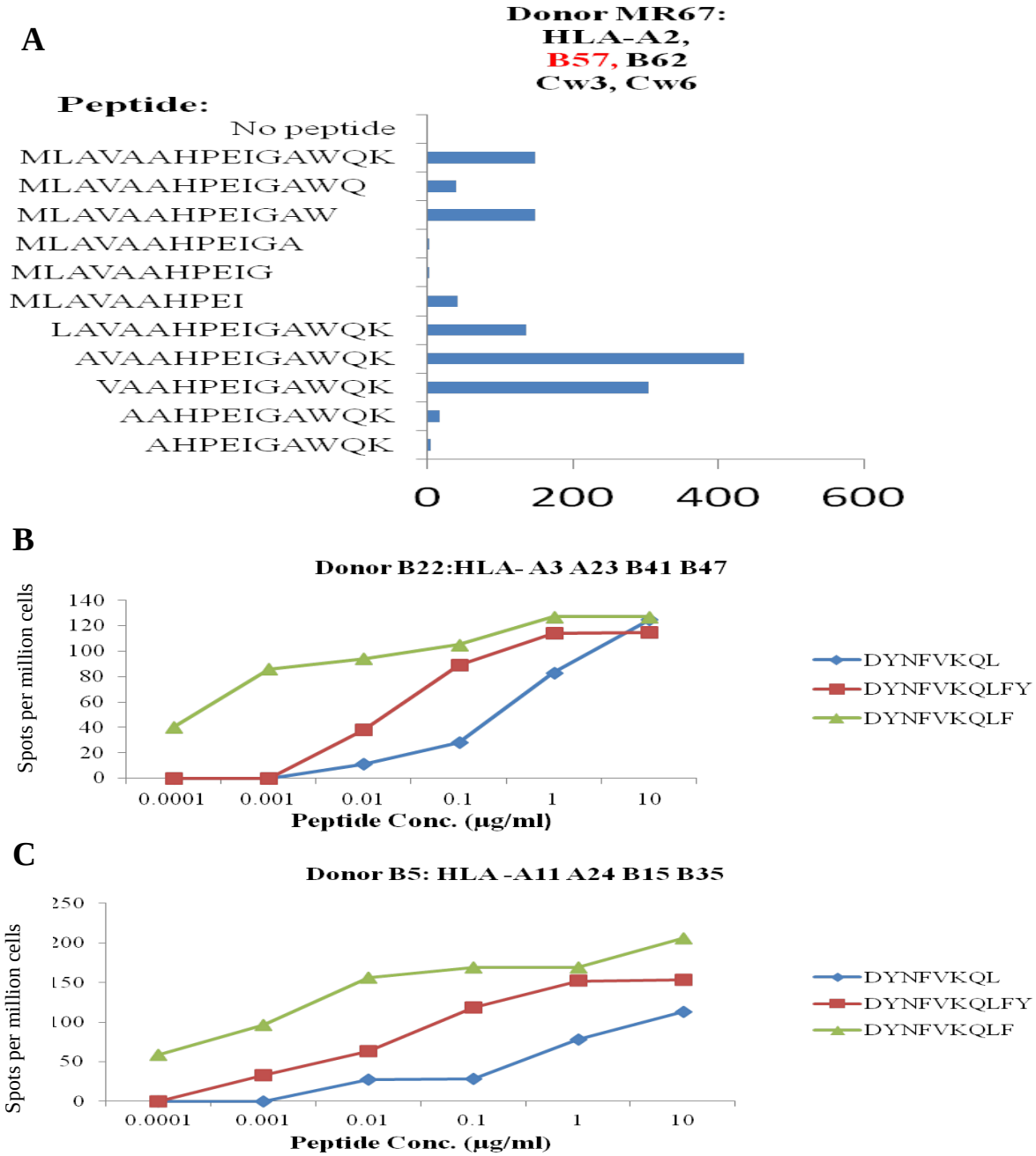


FIGURE 3. Mapping epitopes from the BMLF1 protein. **(A)** EBV-specific T cell lines from two HLA-B57^{+} individuals MR67 and B38 were tested with the overlapping peptides as shown by IFN-γ ELISpot assays at a final concentration of 10µg/ml. Various concentrations of the indicated peptides were tested with EBV-specific T cell lines from **(B)** an HLA-A23^{+} donor B22 and **(C)** an HLA-A24^{+} donor B5 identifying the minimal epitope.

In 1997, Steven *et al* described ^{244}KDTWLDARM^{252} (included within peptides ^{236}EFFTKSTNKDTWLDARMQAI^{255} and ^{241}STNKDTWLDARMQAIQNAGL^{260}, Fig. 1B) as a new epitope, however no HLA-restriction was noted (Steven, Annels et al. 1997). Based on the previously described binding motif for HLA-B47 (Hillen, Mester et al. 2008) and the HLA typing of the responding donor in our study (Donor B22), it is likely that this epitope is restricted by HLA-B47.

Another newly defined BMLF1 epitope mapped from several donors is ^{320}DYNFVKQLF^{328} (included within peptides ^{311}APCFLPNTRDYNFVKQLFYI^{330} and ^{316}PNTRDYNFVKQLFYITCATA^{335}, Fig. 1B). Three peptides were designed to map this new BMLF1 epitope: DYNFVKQL, DYNFVKQLFY and DYNFVKQLF. These peptides were titrated from 10μg/ml to 0.0001μg/ml and tested with donors B22 (HLA-A*23-Fig. 3B) and B5 (HLA-A*24-Fig. 3C). The minimal epitope was found to be the nonamer ^{320}DYNFVKQLF^{328} which titrated out to a final peptide concentration of 0.0001μg/ml. Of most interest was that DYNFVKQLF had a significant response in both the HLA-A*23 and HLA-A*24 donor. Degeneracy in HLA restriction has been shown previously for other EBV epitopes (Burrows, Elkington et al. 2003, Frahm, Yusim et al. 2007). Such epitopes could be particularly valuable for EBV vaccine development allowing one epitope to deliver protection through more than one HLA allele.

Five novel epitopes were defined from the lytic antigen BRLF1 (Table 2). Figure 1C shows a selection of donors with significant IFN-γ responses to a number of the BRLF1 overlapping peptides. Peptides ^{91}AENAGNDACSIACPIVMRYV^{110}, ^{96}NDACSIACPIVMRYVLDHLI^{115} and/or ^{101}IACPIVMRYVLDHLIVVTDR^{120} (Fig. 1C) elicited IFN-γ responses in donors B2, B3 and B39, DM42 and KK68. The results from donors B2, B3 and B39 tested with a truncated peptide set are shown in Fig. 4 A to C. The minimal epitope with maximum activity in all three donors is the nonamer, ^{101}IACPIVMRY^{109}. Interestingly, this epitope appears to be presented by a number of HLA alleles including HLA-Cw2 and HLA-Cw3 (shared by donors B2, B3 and B39). Previous studies have noted a longer version of this epitope ^{101}IACPIVMRYVLDHLI^{115} (Pudney, Leese et al. 2005) (Abbott, Quinn et al. 2013), concluding that the epitope was HLA-A*2402 and Cw*0202 restricted in the case of Pudney *et al* and HLA-B*58 restricted in the Abbott *et al* study although the responding donor was also HLA-Cw3.

BRLF1 overlapping peptides ^{56}INEAKAHGREWGGLMATLNI^{75}, ^{61}AHGREWGGLMATLNICNFWA^{80} and ^{66}WGGLMATLNICNFWAILRNN^{85} (Fig. 1C) induced significant IFN-γ responses from donor B2. Figure 4D depicts the results of a truncated peptide set to determine the minimal epitope in this donor. The decamer ^{64}REWGGLMATL^{73} is the minimal epitope with maximum activity, with a likely HLA-restriction of HLA-B*40 based on the published

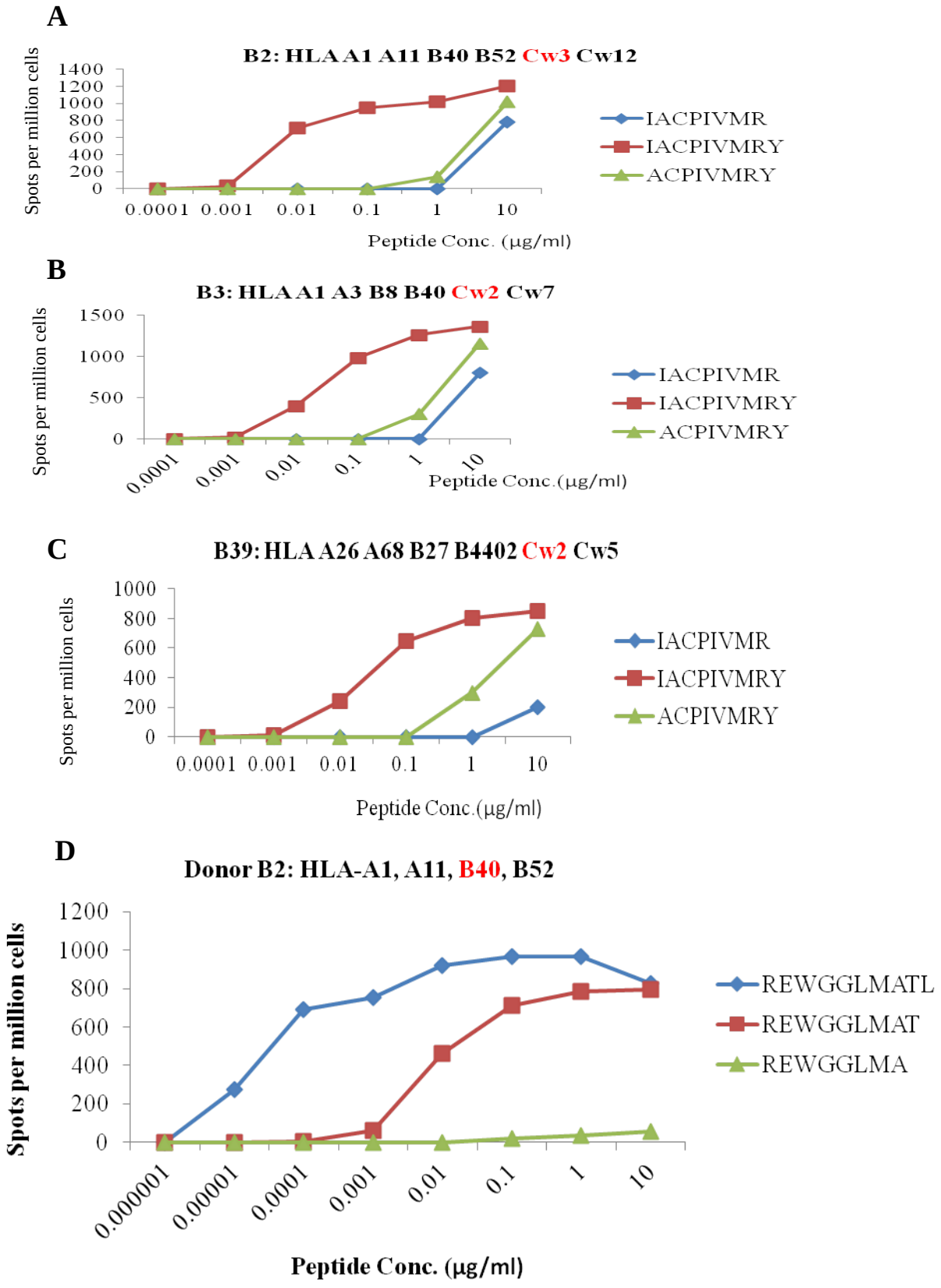


FIGURE 4. Identification of epitopes mapped from the BRLF1 antigen of EBV. EBV-specific T cell lines from **(A)** an HLA-Cw3^{+} individual, **(B)** and **(C)** two HLA-Cw2^{+} individuals and **(D)** an HLA-B40^{+} individual, raised by in vitro stimulation with irradiated autologous LCLs, tested in IFN-γ ELISpot assays for recognition of indicated peptides corresponding to regions of the BRLF1 protein sequence. Cells were tested with various concentrations of the indicated peptides.

binding motif with anchor residues Glutamic acid at position 2 and Leucine at the C-terminus (Falk, Rotzschke et al. 1995). Abbott *et al* also noted that a longer version of this peptide ^{61}AHGREWGGLMATLNI^{75} stimulated T cell responses in their study, with an unknown HLA restriction (Abbott, Quinn et al. 2013). The responding donor in the Abbott *et al* study is noted as HLA-B*60 which is a split antigen that recognizes certain B40 serotypes.

Latent antigen EBNA3C was the source of the second largest number of novel epitopes, with nine defined in this study (Table 2). Figure 1D displays a number of donors that responded to peptides ^{236}EQERYAREAEVRFLRGKWQR^{255} and ^{241}AREAEVRFLRGKWQRRYRRI^{260}. Data from a truncated peptide set corresponding to the overlapping region of these two peptides, tested with donors B23 and B28 are shown in Fig. 5A and 5B demonstrating the novel decamer ^{244}AEVRFLRGKW ^{253} as the minimal epitope with maximum activity. This sequence conforms to the HLA-B*4402/03 binding motif (Hillen, Mester et al. 2008), an allele shared by donors B23 and B28. This epitope overlaps with the previously mapped HLA-B*2705-restricted epitope ^{249}LRGKWQRRYR^{258} (Brooks, Murray et al. 1993). It is also possible that the 12mer peptide ^{242}REAEVRFLRGKW^{253} is an epitope presented by HLA-B44 since it also conforms to the binding motif of this HLA molecule (Hillen, Mester et al. 2008).

A second novel decamer from EBNA3C was mapped from donor B32. Significant IFN-γ responses were observed from this donor to peptides ^{171}WLLTSPSQSWPMGYRTATLR^{190} and ^{176}PSQSWPMGYRTATLRTLTPV^{195} (Fig. 1D). Results from a truncated peptide set with donor B32 demonstrate that the minimal epitope with maximum activity was ^{180}WPMGYRTATL^{189} (Fig. 5C) which conforms to the HLA-B7 binding motif (Maier, Falk et al. 1994). Also shown in Fig. 1D is the response to peptides ^{326}NPYHARRGIKEHVIQNAFRK^{345} and ^{331}RRGIKEHVIQNAFRKAQIQG^{350} from donor B3. These overlapping peptides incorporate the nonamer ^{335}KEHVIQNAF^{343} that was previously mapped as a HLA-B*4402 restricted epitope (Khanna, Burrows et al. 1992). Although donor B3 is not HLA-B*4402^{+}, another member of the HLA-B44 supertype, HLA-B40 is expressed, suggesting that the ^{335}KEHVIQNAF^{343} epitope is presented by this allele as well as HLA-B*4402.

As with the other lytic antigens included in this study, BMRF1 was found to include several novel epitopes. Pudney *et al* mapped the unusually long HLA-B*0702-restricted epitope ^{116}RPQGGSRPEFVKL^{128} in a 2005 study (Pudney, Leese et al. 2005). Figure 1E shows IFN-γ responses to the overlapping peptides ^{106}VEQASLQFYKRPQGGSRPEF^{125} and ^{111}LQFYKRPQGGSRPEFVKLTM^{130} from donors B17, B20 and B36. Figure 6A depicts results from donor B20 screened for recognition of a truncated peptide set corresponding to the overlapping region. The minimal epitope with maximum activity was found to be the decamer, ^{116}RPQGGSRPEF^{125}. Interestingly, another HLA-B7^{+} donor, B21, only responded to the longer

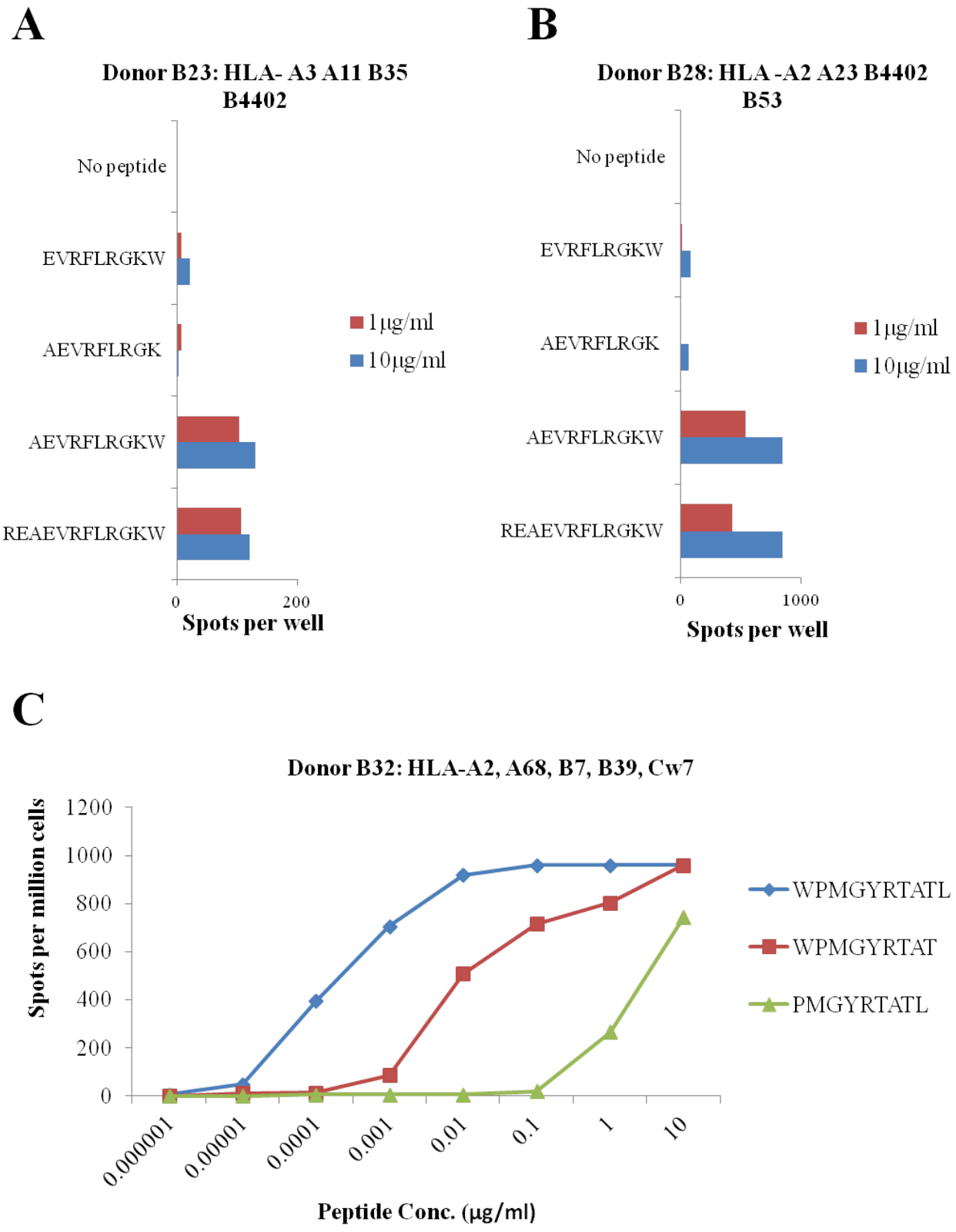


FIGURE 5. Defining two decamer EBNA3C epitopes. IFN-γ ELISpot assays were used to define minimal epitopes, by screening EBV-specific T cell lines from two HLA-B*4402^{+} individuals **(A)** donor B23 and **(B)** donor B28 with truncated peptides at10µg/ml and 1µg/ml and **(C)** a T cell line from the HLA-B7^{+} individual, donor B32, against various concentrations of the indicated peptides.

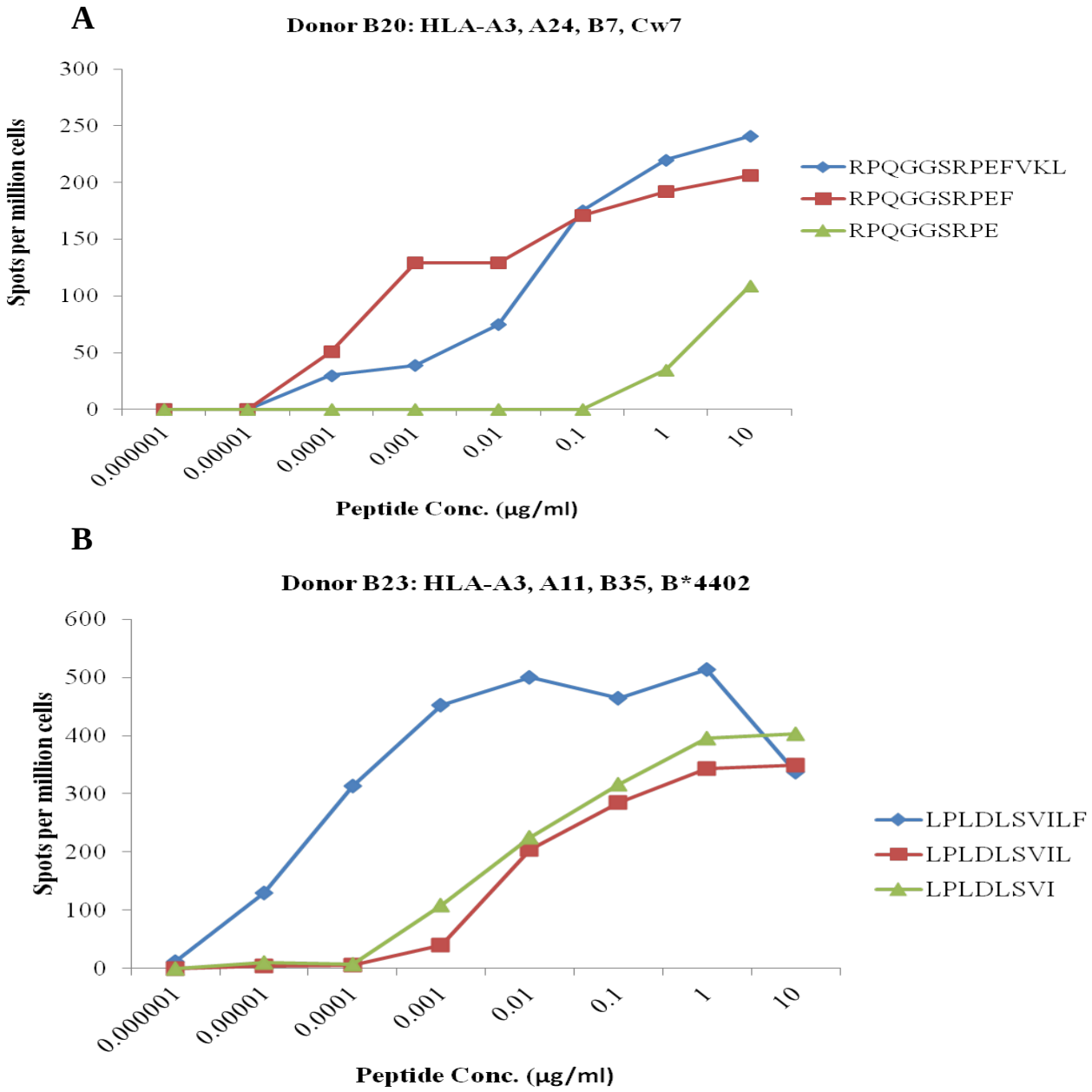


FIGURE 6. Identification of minimal epitopes from the EBNA3C antigen of EBV. EBV-specific T cell lines from HLA-B7^{+} individual **(A)** donor B20 and HLA-B35^{+} individual **(B)** donor B23 were screened with various concentrations of the indicated peptides by IFN-γ ELISpot assays.

13mer peptide as described by Pudney *et al* (data not shown). A novel epitope was also mapped from donor B23. As shown in Fig. 6B, the decamer ^{286}LPLDLSVILF^{295} was recognised at relatively low peptide concentrations, and this peptide conforms to the HLA-B*35 binding motif (Falk, Rotzschke et al. 1993). Interestingly, this epitope was previously noted as HLA-B*53 restricted (Chapter 1, Table 2, Annels, NE unpublished) and both HLA-B*35 and HLA-B*53 are members of the HLA-B7 supertype (Sidney, Peters et al. 2008). Together these findings provide a significant list of novel epitopes from a number of EBV lytic and latent antigens restricted to various HLA -A, -B and -C alleles. In order to definitively determine the HLA restricting elements

of these novel epitopes future studies could focus on culturing $CD8^+$ T cell clones specific for each novel peptide. These cloned populations can be tested for recognition of peptide-loaded target cells, either autologous donors or allogeneic donors partially matched specific for either HLA-A, -B or -C alleles. Results from such studies will identify the epitope's restricting allele (Feederle, Kost et al. 2000).

3.3 Discussion

The initial aim of this study was to address the relative prevalence of noncanonical $CD8^+$ T cell epitopes of over 10 amino acids in length in selected lytic and latent antigens of EBV. During the investigation it became apparent that the system is very complex, with various factors influencing epitope selection. The "big picture" view needs to incorporate, the virus's interaction with the host and the role that each antigen plays in that, and the HLA allele with respect to anchor residue preferences in addition to what length of epitope each allele prefers. Noncanonical epitopes were identified albeit in very limited number. However, an expansive list of new EBV CTL epitopes have been defined, allowing an assessment of the hierarchy of immunodominance between the EBV antigens and highlighting the highly immunogenic regions within several of the antigens where epitope clustering has been observed.

Considering the immunodominance between the EBV antigens included in this study of healthy individuals, as assessed by the total number of $CD8^+$ T cell epitopes, the latent antigen EBNA3A contains the greatest number of epitopes. Supporting our observations Ning *et al* showed immunodominance of EBNA3 proteins over other latent proteins including EBNA1 in $CD8^+$ T cell responses (Ning, Xu et al. 2011). Hislop *et al*, having compared the relative immunodominance of lytic and latent EBV antigens in healthy donors, showed a hierarchy which largely concurs with our findings (Hislop, Taylor et al. 2007). The relative immunodominance hierarchy in this study (with the exclusion of BZLF1), based on the identification of 28 latent and 13 lytic epitopes, is EBNA3A, EBNA3C, EBNA3B, BRLF1, BMLF1 and BMRF1 and lastly EBNA1.

Latent antigens EBNA3A, EBNA3B and EBNA3C all play an important role in the interaction with RBPJκ binding activity (Robertson, Lin et al. 1996). This RBPJκ binding activity region in all three of the EBNA 3 proteins is the location of around half of the novel epitopes identified in this study. It is interesting that, with a vast level of homology between these three antigens, that EBNA3A appears to have the most epitopes that are recognized by EBV-specific cytotoxic T lymphocytes. Indeed it is the conserved region in these three EBNA3 proteins from amino acids 90-320 that account for the location of around half of the epitopes for these antigens. Only a handful of EBNA 3C epitopes are located outside of this conserved region. EBNA3A also has regions of overlapping and clustered epitopes, albeit not to the extent of BZLF1, although this is

probably influenced by antigen size as EBNA3A is 944aa compared to BZLF1's 245aa (refer to Table 2 and Chapter 1,Table 1).

The EBNA 3 proteins contain many immunodominant $CD8^+$ T cell epitopes. Highlighted by Bihl and colleagues study with healthy virus carriers, six of the defined EBNA3A epitopes are immunodominant, as are seven epitopes for both EBNA 3B and EBNA3C (Bihl, Frahm et al. 2006). However, the frequency with which individuals respond to these immunodominant epitopes appears to be variable. A study by Benninger-Doring *et al* with eight HLA-B8-positive EBV carriers, examined for reactivities to HLA-B8-restricted EBV epitopes, showed 50% of donors responded to the immunodominant EBNA3A-^{325}FLRGRAYGL^{333} epitope while only 25% of donors reacted against ^{158}QAKWRLQTL^{166}, also from EBNA3A. This is compared to all eight donors recognising the BZLF1 ^{190}RAKFKQLL^{197} epitope (Benninger-Doring, Pepperl et al. 1999).

It has been suggested that differences in $CD8^+$ T cell responses may be greatly influenced by the source of the EBV epitope (e.g. lytic or latent antigen) (Hislop, Annels et al. 2002). The rationale for this being that high viral loads in acute infection may propel responses of low avidity against highly expressed lytic antigens, whilst during chronic infection the preference may lean to high avidity responses against latent antigens expressed during later infection stages (Woodberry, Suscovich et al. 2005). Indeed, previous work has demonstrated a noticeable shift from lytic to latent antigen-specific responses between the acute and chronic stages of EBV infection (Hislop, Annels et al. 2002, Woodberry, Suscovich et al. 2005). Furthermore, the reactivities towards CTL epitopes will reflect the donor cohort and may be quite variable between studies depending on the individual's EBV status i.e. acute IM, memory IM or healthy carrier.

The remaining latent antigen examined, EBNA1, is essential for the persistence and maintenance of the EBV episome and enhancing genome expression. A unique feature of EBNA1 is the GAr domain which inhibits MHC class I presentation of cis-linked epitopes. Also of great importance, EBNA1 is the only viral protein that is expressed in all forms of latency and in all EBV-associated malignancies (Rickinson and Kieff 2007, Long, Taylor et al. 2011). Supporting the notion that the GAr domain of EBNA1 reduces its immunogenicity, only one epitope was defined in this study (Table 2) compared to the large number of epitopes that were identified in the other EBV antigens investigated. A number of donors did respond to known EBNA 1 epitopes, although not all defined epitopes elicited CTL reactivities (Fig.1G). This is not the first study to describe little or no responses to EBNA1. A study by Murray *et al* also noted no detection of CTL components to EBNA1 when testing 16 EBV-immune donors (Murray, Kurilla et al. 1992). It is notable that the C-terminal half of this protein is a rich source of $CD4^+$ T cell epitopes. These epitopes covering a range of HLA class II alleles, have been recognised in Caucasian donors

(Steigerwald-Mullen, Kurilla et al. 2000) (Voo, Fu et al. 2002) (Leen, Meij et al. 2001) and Chinese donors (Tsang, Lin et al. 2006) .

The results of this study confirm and extend the findings of the recent work of Abbott *et al* on immunodominance among EBV lytic cycle antigens. These authors found a number of new epitopes in EBV antigens BZLF1, BRLF1, BMLF1 and BMRF1 (among other lytic antigens not examined in this study). Not all epitopes identified in the Abbott *et al* work elicited CTL responses in the donors in this study. We have mapped several new epitopes from these antigens while additionally identifying the likely HLA-restriction. An example is the novel BRLF1 epitope ^{101}IACPIVMRY^{109} which overlaps with the previously defined HLA-A*0201-restricted epitope ^{109}YVLDHLIVV^{117}. Initially defined by Pudney *et al* as ^{101}IACPIVMRYVLDHLI^{115} with HLA-A24 or HLA-Cw2 restriction, Abbott *et al* also noted this 15-mer to be HLA-B58 restricted. We have shown that the minimal epitope ^{101}IACPIVMRY^{109} appears to be presented by a number of HLA alleles including HLA-A*29 and possibly HLA-Cw2 and HLA-Cw3 (shared by donors B2, B3 and B39). It cannot be ruled out that this nonamer epitope is unusually promiscuous with regards to HLA restriction and can also elicit responses from other HLA alleles including HLA-B58 and HLA-A24.

The pattern of EBV lytic antigen immunodominance described by Hislop *et al* in healthy virus carriers (Hislop, Taylor et al. 2007) with immediate early (IE) proteins being immunodominant over early (E) proteins, is maintained in our study, with the IE BRLF1 antigen demonstrating five novel epitopes and the E antigens BMLF1 and BMRF1 each containing four novel epitopes (Table 2). When the IE antigen BZLF1 is considered (refer to Chapter 4 and Chapter 5) this pattern becomes more evident, with a total of 13 novel epitopes arising from this antigen. The role of transcription factor BRLF1 in concert with BZLF1 is the initiation of the lytic cycle from the latent state (Ragoczy, Heston et al. 1998), and full expression of both these antigens is required during the lytic phase (Feederle, Kost et al. 2000). Previously defined BRLF1 epitopes are found in all domains of this antigen ranging from the DNA binding domain (aa 1-320), the dimerization domain (aa 1-232) to the transcriptional domain (aa 520-605) (Manet, Rigolet et al. 1991), while the proline rich domains (aa 352-410 and aa 450-500) (Manet, Rigolet et al. 1991) contains only one. However, novel BRLF1 epitopes defined in this study remained confined to the DNA binding domain. Interestingly, sequence analysis of BRLF1 in various lymphoma patients shows that amino acid mutations do not occur in the sequence of any novel or already known epitopes (Yang, Jia et al. 2014).

Early lytic antigen BMRF1 has recently been shown to play a pivotal role in nuclear translocalization when coexpressed with molecular chaperone Hsp90 (Kawashima, Kanda et al. 2013). This process is dependent on the nuclear localisation signal domain located at the C-

terminal region of BMRF1 (aa 378-404) (Zhang, Holley-Guthrie et al. 1999) . It is of note that no epitopes are contained within this region of BMRF1 and all previously defined and novel epitopes with the exception of ^{369}LERPLAVQL^{377}, are located in the DNA binding domains (aa 44-194 and aa 238-302). An interesting observation from this antigen is that there appears to be a 13-mer epitope ^{116}RPQGGSRPEFVKL^{128} with HLA-B*0702 restriction noted by Pudney *et al* in a 2005 study (Pudney, Leese et al. 2005) which overlaps with the novel decamer epitope, ^{116}RPQGGSRPEF^{125} that also appears to be restricted by HLA-B*0702. Interestingly, donor B21 only responds to the longer 13-mer peptide as described by Pudney *et al* (data not shown). This preference to respond to the longer epitope may be due to a subtle HLA subtype difference in this donor and may warrant further investigation.

In contrast to other EBV antigens all of the BMLF1 epitopes are located in the C-terminal half of the antigen. The results from this study confirmed a number of known BMLF1 epitopes, albeit with additional likely HLA allele restriction, as well as mapping a minimal epitope previously described by Abbott *et al* (Abbott, Quinn et al. 2013). Furthermore, we have shown the novel epitope ^{320}DYNFVKQLF^{328} to have HLA-restriction to both the HLA-A23 and HLA-A24 alleles.

With the exception of BMLF1, there appears to be a trend that a great majority of the epitopes defined in this study were located in the N-terminal region of the EBV antigens examined indicating that this region is an important target for the T cell response. If this observation indeed reflects the distribution of epitopes this information can be harnessed for many applications. Initially, the location of epitopes within an antigen can provide insight into the functional roles and importance of certain domains within an antigen. For example, EBNA3C resiudes aa130-159 bind to IRF4 or IRF8 (Banerjee, Lu et al. 2013) in addition to coactivating the EBV LMP-1 promoter with EBNA-2 via an SPI1 site in the absence of RBPJ (Zhao and Sample 2000, Lin, Johannsen et al. 2002). Interestingly, there are no EBNA3C epitopes located in this region. These areas bereft of epitopes may play a role in immune evasion and host mechanisms for such. This has been observed in relation to GAr region of EBNA1, whereby EBNA1 peptides are poorly presented in the context of MHC Class I molecules and somewhat protected from recognition by CTLs (Levitskaya, Shapiro et al. 1997, Fogg, Kaur et al. 2005). This information can be further used to enhance epitope prediction algorithms by providing further parameters to increase peptide prediction success. Future studies could investigate these “epitope poor” regions of antigens in search of immune evasion mechanisms.

A point of difference between this study and many previous investigations that have mapped EBV T cell epitopes is the samples which were utilised. Previous work by many has focused on the $CD8^+$ T cell responses from IM samples (Callan 2003) (Steven, Annels et al. 1997, Hislop, Annels et al. 2002, Pudney, Leese et al. 2005), whereas this study has used T cell lines raised by *in vitro*

stimulation of PBMCs from 41 healthy EBV-seropositive individuals with their autologous LCLs. IM samples, however, have been shown to have somewhat skewed reactivities towards a lytic response and possibly epitopes that are only apparent during the initial stages of infection that later become extinguished (Callan 2003). Sample sizes and donor cohort differences, in addition to the overlapping peptide screening method may lead to variations in epitope detection.

It is difficult to compare the findings of this study with those focused on IM samples as in some cases the HLA of the indivduals and scope of the study was limited to that of already defined epitopes (Hislop, Annels et al. 2002) or limited to lytic antigens (Steven, Annels et al. 1997, Pudney, Leese et al. 2005). In this study a significant proportion of the epitopes defined are from latent antigens compared to the lytic antigens with 28 and 13 epitopes defined, respectively. This observation in itself somewhat defies the notion of skewed reactivities towards lytic responses, however comparisons have not been made with respect to the number of $CD8^+$ T cells specific for each of these newly defined epitopes. Of note are some similarities between the work of Pudney and colleagues (Pudney, Leese et al. 2005) examining $CD8^+$ T cell responses in IM donors and this study examining healthy carriers. Both studies identified near identical epitopes in BRLF1 and BMRF1 (the minimal epitope was not determined in the Pudney study, refer to Table 2 and Chapter 1, Table 2) demonstrating that near identical CTL responses can be seen in both IM and healthy donors.

As this study utilised T cell lines that were generated by exposure to autologous LCLs, there remains a possibility of skewing epitope-specific responses towards the latent antigens of EBV. However,this does not appear to be the case since similar reactivities were observed when peptides were tested with PBMCs or T cell lines generated by peptide stimulation (data not shown, (Rist, Hibbert et al. 2015)). Although a potential bias may exist for reactivities towards latent responses, LCLs typically have 1-5% of cells in the lytic cycle, and this appears to be sufficient to stimulate many large lytic antigen-specific responses.

In order to ascertain the extent of the abundance of noncanonical epitopes the overall numbers of novel epitopes of this study were combineed as well as results from previous studies. Of the 41 novel epitopes that were defined in this study, 4 (~10%) were noncanonical 11mers (Table 2). The distribution of these 11mer epitopes was in antigens BMRF1, EBNA3A, EBNA3B and EBNA3C. Interestingly, three of these epitopes have a Proline at the second residue and are recognised by members of the HLA–B7 supertype family concurring with previous findings (Burrows, Rossjohn et al. 2006, Burrows, Bell et al. 2008). To determine the percentage of noncanonical epitopes for each studied antigen the number of new epitopes defined in this study were added to the number of known epitopes to establish an overall number of epitopes. For the latent antigens EBNA1, EBNA3A, EBNA3B and EBNA3C the percentage of noncanonical

epitopes is 11%, 4%, 14% and 5%, respectively. As minimal epitopes have been determined for all of the epitopes within these latent antigens it is clear that epitopes of over 10 residues in length play a significant role in the $CD8^+$ T cell response to these antigens.

However, this is less clear for the lytic antigens as many of the known epitopes are not defined as minimal epitopes, although and there are a number for which the minimal epitope has been determined in this study. For example, there are eight known BMRF1 epitopes including a defined 13mer and two 15mers which may or may not be the minimal epitope. In addition four novel epitopes have been defined in this study, one of which is an 11mer. The percentage of noncanonical epitopes in BMRF1 is therefore 16% if the 15mers are excluded. There were no noncanonical epitopes found in BMLF1 or BRLF1 in this study.

To determine if the 41 novel epitopes identified in this study would have been predicted using standard epitope prediction algorithms, each antigen was assessed with the web-based prediction tool, SYFPEITHI. Interestingly, 15 of the 41 novel epitopes, including the four 11 mer epitopes were not predicted by this tool. This failure to predict epitopes was due to a number of factors, including limitations with HLA coverage and epitope length. There are no prediction parameters for HLA-A23, -A29, -A32, -B15, -B40, -B52, -B53 and the HLA-C alleles and although there appears to be an 11mer prediction function when this was chosen for all antigens the program came back with no search results. These findings indicate that many parameters of standard epitope prediction algorithms require reassessment and updating in order to reflect accurate epitope length and wider HLA coverage.

Although it does appear that long $CD8^+$ T cell epitopes are not as plentiful as we initially hypothesised, four 11-mer epitopes have been defined including one in each of the EBNA3 proteins and another in BMRF1 (Table 2). Our recent studies have shown that epitope length preference is heavily influenced by structural features of different HLA molecules. The noncanonical dodecamer that was defined in this study, ^{169}EECDSELEIKRY^{180} from BZLF1 helped highlight that HLA-B*4403 favours the presentation of longer peptides (Rist, Theodossis et al. 2013) (Chapter Five). Future work in this area is required to determine the peptide length preference of other HLA alleles which may be used to improve web-based algorithms for epitope prediction in addition to assisting strategies for peptide-based vaccine design. However, the present investigation has revealed many new epitopes including some that are clustered together and overlapping with previously defined epitopes and some that are restricted by several HLA alleles.

Chapter 4:

Aim Two:
Determine if the highly immunogenic BZLF1 antigen of EBV includes novel $CD8^+$ T cell epitopes.

Hypothesis Two:

There are many unidentified $CD8^+$ T cell epitopes to be defined in the BZLF1 antigen.

Publication: Original research article

Rist MJ, Neller MA, Burrows JM, Burrows SR. T Cell Epitope Clustering in the Highly Immunogenic BZLF1 Antigen of Epstein-Barr virus. *Journal of Virology.* 2015 Jan; 89 (1):703-712.

JVI
Journals.ASM.org

T Cell Epitope Clustering in the Highly Immunogenic BZLF1 Antigen of Epstein-Barr Virus

Melissa J. Rist,[a,b] Michelle A. Neller,[a] Jacqueline M. Burrows,[a] Scott R. Burrows[a,b]

Cellular Immunology Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, Australia[a]; School of Medicine, The University of Queensland, Brisbane, Australia[b]

ABSTRACT

Polymorphism in the human leukocyte antigen (HLA) loci ensures that the CD8$^+$ T cell response to viruses is directed against a diverse range of antigenic epitopes, thereby minimizing the impact of virus escape mutation across the population. The BZLF1 antigen of Epstein-Barr virus is an immunodominant target for CD8$^+$ T cells, but the response has been characterized only in the context of a limited number of HLA molecules due to incomplete epitope mapping. We have now greatly expanded the number of defined CD8$^+$ T cell epitopes from BZLF1, allowing the response to be evaluated in a much larger proportion of the population. Some regions of the antigen fail to be recognized by CD8$^+$ T cells, while others include clusters of overlapping epitopes presented by different HLA molecules. These highly immunogenic regions of BZLF1 include polymorphic sequences, such that up to four overlapping epitopes are impacted by a single amino acid variation common in different regions of the world. This focusing of the immune response to limited regions of the viral protein could be due to sequence similarity to human proteins creating "immune blind spots" through self-tolerance. This study significantly enhances the understanding of the immune response to BZLF1, and the precisely mapped T cell epitopes may be directly exploited in vaccine development and adoptive immunotherapy.

IMPORTANCE

Epstein-Barr virus (EBV) is an important human pathogen, associated with several malignancies, including nasopharyngeal carcinoma and Hodgkin lymphoma. T lymphocytes are critical for virus control, and clinical trials aimed at manipulating this arm of the immune system have demonstrated efficacy in treating these EBV-associated diseases. These trials have utilized information on the precise location of viral epitopes for T cell recognition, for either measuring or enhancing responses. In this study, we have characterized the T cell response to the highly immunogenic BZLF1 antigen of EBV by greatly expanding the number of defined T cell epitopes. An unusual clustering of epitopes was identified, highlighting a small region of BZLF1 that is targeted by the immune response of a high proportion of the world's population. This focusing of the immune response could be utilized in developing vaccines/therapies with wide coverage, or it could potentially be exploited by the virus to escape the immune response.

Epstein-Barr virus (EBV), a lymphotropic gamma-1 herpesvirus, is widespread in all human populations, infecting around 95% of the adult population worldwide. Transmission of this lifelong persistent virus is via the oral route. After close contact, viral particles present in the saliva of infected individuals enter the oral cavity of the naive individual, and the virus is amplified by replicative (lytic) infection in permissive cells in the oropharynx. This lytic infection results in vast amounts of virus shedding into the throat. Simultaneously, infection of mucosal B cells occurs, and the virus initiates latent, persistent infection of the B cell pool (1–3). In developing countries, EBV is associated with an asymptomatic infection, usually occurring in infancy or early childhood. In developed countries, however, acute infection is often delayed until adolescence or early adulthood and, in around 25% of cases, can result in the self-limiting lymphoproliferative disorder infectious mononucleosis (IM). IM was etiologically linked to EBV in 1968 by Henle et al. (4). Symptoms associated with IM range from fever and sore throat to lymphadenopathy and splenomegaly (3–5).

Around 80 EBV proteins are expressed during lytic viral replication, which involves the sequential expression of immediate early, early, and late proteins. Lytic cycle activation is initiated by the expression of two immediate early proteins, BZLF1 (Zta) and BRLF1 (Rta). Cox et al. noted that *Rta* and *Zta* appeared to regulate gene expression at the level of transcription and that these transactivators work in concert to facilitate reactivation (6). Our protein of interest, BZLF1, or ZEBRA (Z EBV replication activator), is a member of the basic leucine zipper family which binds to the AP1-like Z response elements in EBV early promoters (7) and exhibits sequence similarity to c-Fos (8). In addition to the transactivation domain (amino acids [aa] 1 to 166), BZLF1 contains a basic DNA recognition domain (aa 178 to 194) and a coiled-coil dimerization domain (aa 198 to 225) (9, 10).

The adaptive immune response is vital in controlling EBV-infected B-cell proliferation and is of immense importance during persistent infection (1). CD8$^+$ T cells recognize peptides derived

Received 11 September 2014 Accepted 20 October 2014
Accepted manuscript posted online 29 October 2014
Citation Rist MJ, Neller MA, Burrows JM, Burrows SR. 2015. T cell epitope clustering in the highly immunogenic BZLF1 antigen of Epstein-Barr virus. J Virol 89:703–712. doi:10.1128/JVI.02642-14
Editor: R. M. Longnecker
Address correspondence to Scott R. Burrows, scott.burrows@qimrberghofer.edu.au.

doi:10.1128/JVI.02642-14

from viral proteins associated with major histocompatibility complex (MHC) molecules. Results from Pudney et al. suggest a focusing of $CD8^+$ T cell responses toward epitopes from immediate early and early proteins and a hierarchy of immunodominance among the EBV lytic cycle antigens (11). The lytic protein BZLF1 includes a number of defined $CD8^+$ T cell epitopes presented by HLA-B and -C alleles. The most widely studied is the HLA-B8-restricted RAKFKQLL epitope. Tan et al. demonstrated that up to 5.5% of $CD8^+$ T cells in the peripheral circulation of healthy virus carriers are specific for this epitope (12). Other $CD8^+$ BZLF1 epitopes include the highly immunogenic 13-mer LPEPLPQGQLTAY, presented by HLA-B*3508, which overlaps with the HLA-B*3501-restricted EPLPQGQLTAY epitope (13). Recently, we defined two novel overlapping epitopes from BZLF1: an HLA-B*1801-restricted octamer, SELEIKRY, which is encompassed by the HLA-B*4403-restricted EECDSELEIKRY dodecamer (14).

In the present study, we further investigated the $CD8^+$ T cell response to the BZLF1 protein, identifying 11 novel epitopes, many presented by common HLA alleles. Interestingly, these epitopes appear to be clustered within certain domains of the protein, with many overlapping sequences.

MATERIALS AND METHODS

Generation of T cell lines. PBMCs were isolated by Ficoll-Hypaque centrifugation into RPMI 1640 medium supplemented with 10% fetal calf serum (FCS) (R10). Blood donors were healthy, EBV-seropositive individuals who had given written informed consent. Approval for this research was obtained from the QIMR Berghofer Medical Research Institute Human Ethics Committee (Brisbane, Australia). EBV-specific T cell cultures were raised by culturing peripheral blood mononuclear cells (PBMCs) (2×10^6/2-ml well) with irradiated autologous lymphoblastoid cell lines (LCLs) transformed with the B95-8 strain of EBV (2×10^5/2 ml well). Cultures were supplemented with recombinant interleukin 2 (rIL-2) (120 IU/ml) from day 3 and analyzed on day 18.

Peptides. Forty-six overlapping peptides (20-mers overlapping by 15 aa) covering the entire length of BZLF1 were designed based on the sequence of the B95.8 strain of EBV. Synthetic peptides were purchased from either New England Peptides (Gardner, MA) or GL Biochem (Shanghai, China) and dissolved in dimethyl sulfoxide (DMSO).

ELISpot assays. Gamma interferon (IFN-γ) enzyme-linked immunosorbent spot (ELISpot) assays were performed using cytokine capture and detection reagents according to the manufacturer's instructions ($ELISpot^{PRO}$ for human IFN-γ; Mabtech, Stockholm, Sweden). Briefly, 96-well nitrocellulose plates precoated with anti-IFN-γ monoclonal antibody (MAb) were seeded with approximately 5×10^4 EBV-specific T cells and peptide at various concentrations. After incubation for 16 h at 37°C in 5% CO_2, the cells were discarded and captured IFN-γ was detected with a biotinylated anti-IFN-γ antibody (Ab), followed by development with an alkaline phosphatase substrate solution (5-bromo-4-chloro-3-indolyl-phosphate [BCIP]–nitroblue tetrazolium [NBT]-plus). All samples were tested in duplicate, and spots were counted using an automated plate counter.

Tetramer labeling. PBMCs or T cell lines were labeled with an allophycocyanin (APC)-conjugated HLA-Cw*06-CRAKFKQLL tetramer (National Institutes of Health Tetramer Core Facility, Emory University, Atlanta, GA) by incubation for 30 min at 4°C. Cells were then washed and incubated with Cy5.5-peridinin chlorophyll protein (PerCP)-conjugated anti-human CD8 MAb (BioLegend, San Diego, CA), Cy7-phycoerythrin (PE)-conjugated anti-human CD3 MAb (eBioscience, San Diego, CA), and Alexa Fluor 700-conjugated anti-human CD4 (BD PharMingen, San Diego, CA) for 30 min at 4°C. Cells were washed and analyzed on a BD LSR Fortessa flow cytometer using the FACSDiva software program (BD Biosciences).

TABLE 1 Class I HLA types of donors used in this study

Donor ID[a]	Allele(s)		
	HLA-A	HLA-B	HLA-C
B2	1, 11	40, 52	3, 12
B4	26, 32	40, 4402	1, 3
B5	11, 24	62, 35	3, 4
B8	2, 66	4402, 49	7, 7
B9	1, 30	8, 18	5, 7
B15	1, 26	62, 27	1, 3
B21	32, 68	7, 14	7, 8
B25	1, 2	35, 51	4, 14
B26	2, 11	35, 58	4, 7
JC33	3, 25	18, 4403	Unknown
MW65	1, 3	8, 3501	4
DP55	2	62, 18	1, 5
CG59	2, 3	7, 18	7, 12
B33	24, 29	1801, 4403	5, 16
LC50	1	8, 18	7
MR67	2	62, 57	3, 6
SB60	2	35, 57	4, 6
HB49	3, 32	7, 8	7
B7	2, 3	7, 37	6, 7
B13	1, 2	27, 57	2, 6
B22	3, 23	41, 47	6, 17
AO83	2, 24	18, 44	5, 7
KK68	1, 29	8, 4403	7, 16
B28	2, 23	4402, 53	5, 6
B38	2, 3	7, 57	6, 7
B3	1, 3	8, 40	2, 7
B10	11, 31	7, 51	7, 15
B11	11, 24	62, 50	4, 6
B14	2, 11	51	15, 16
B16	2, 11	62, 44	3, 5
B17	1, 2	7, 8	7
B19	1, 3	8, 4403	4, 7
B20	3, 24	7	7
B21	32, 68	7, 14	7, 8
B23	3, 11	35, 4402	3, 5
B24	2	40	3

[a] Donor identification.

Intracellular cytokine staining. EBV-specific T cell lines raised by *in vitro* stimulation with an irradiated autologous LCL were incubated in R10 containing various stimuli for 4 h at 37°C in the presence of 5 μg/ml brefeldin A (BioLegend). T cell lines were stimulated with either peptide-pulsed (incubation with 1 μg/ml peptide for 1 h) HLA-deficient T2 cell lines stably transfected to express HLA-A3, A24, or B7 or serially diluted peptide (0.0001 to 10 μg/ml). The cells were then washed and incubated at 4°C for 30 min with Live/Dead fixable aqua dead cell stain (Life Technologies, Grand Island, NY) and fluorescently labeled MAb specific for cell surface markers (CD4-fluorescein isothiocyanate [FITC] [BD PharMingen]; CD8-Cy5.5-PerCP [BioLegend]). Cells were washed and then fixed and permeabilized with Cytofix/Cytoperm fixation/permeabilization solution (BD PharMingen) at 4°C for 20 min. Next, the cells were washed with Perm/Wash buffer (BD PharMingen), incubated with IFN-γ–APC MAb (BD PharMingen) at 4°C for 30 min, and then washed again with Perm/Wash. Cells were resuspended in phosphate-buffered saline (PBS) for analysis on a BD FACSCanto II flow cytometer (BD Biosciences).

RESULTS

BZLF1 is a rich source of $CD8^+$ T cell epitopes. IFN-γ ELISpot assays were used to screen overlapping 20-aa peptides corre-

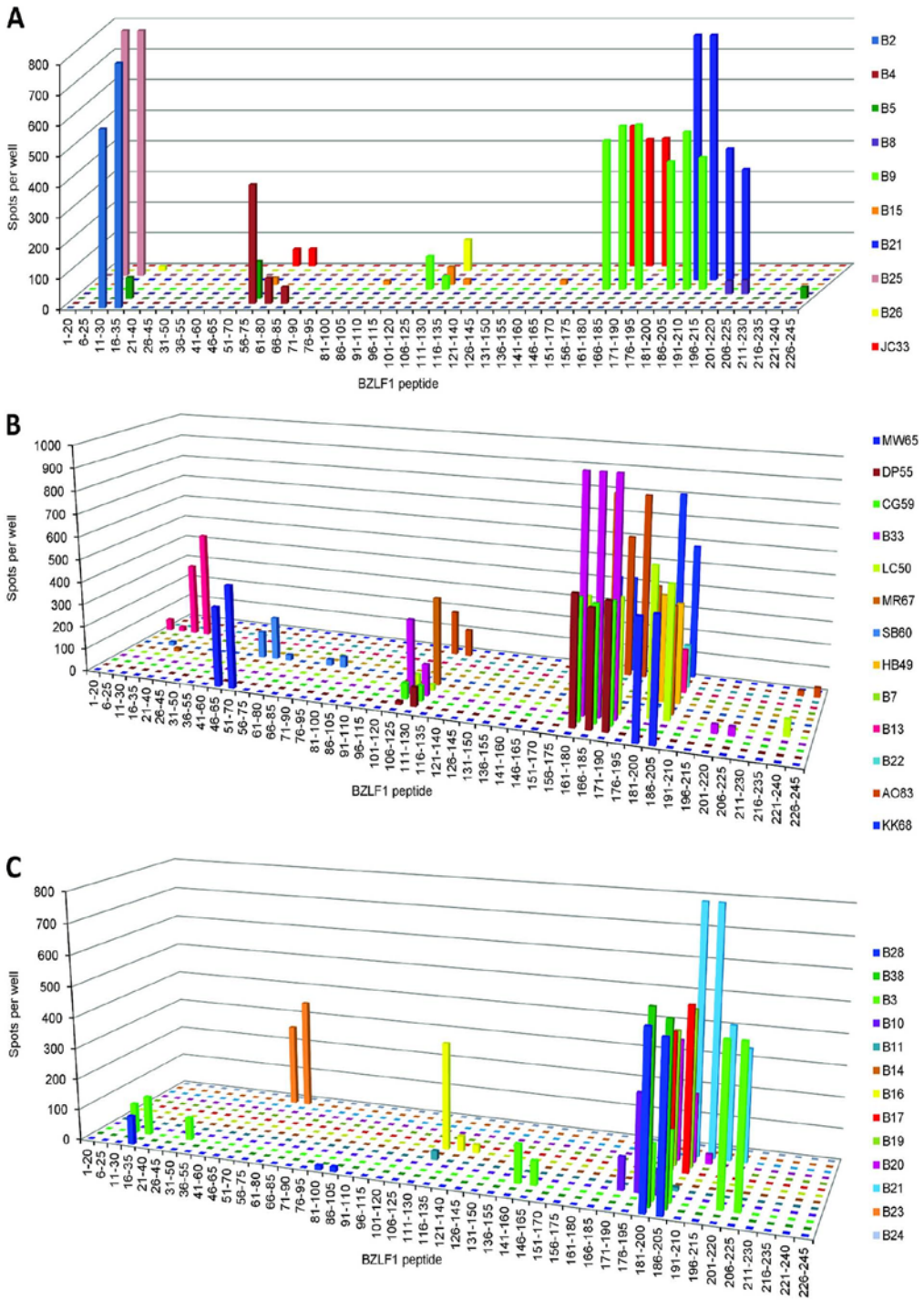


FIG 1 Initial screening of T cells from EBV-seropositive individuals against overlapping peptides spanning the BZLF1 antigen. EBV-specific T cell lines from 36 individuals, raised by *in vitro* stimulation with irradiated autologous LCLs, were tested by IFN-γ ELISpot assays for recognition of overlapping peptides (10 μg/ml) corresponding to the lytic EBV BZLF1 antigen.

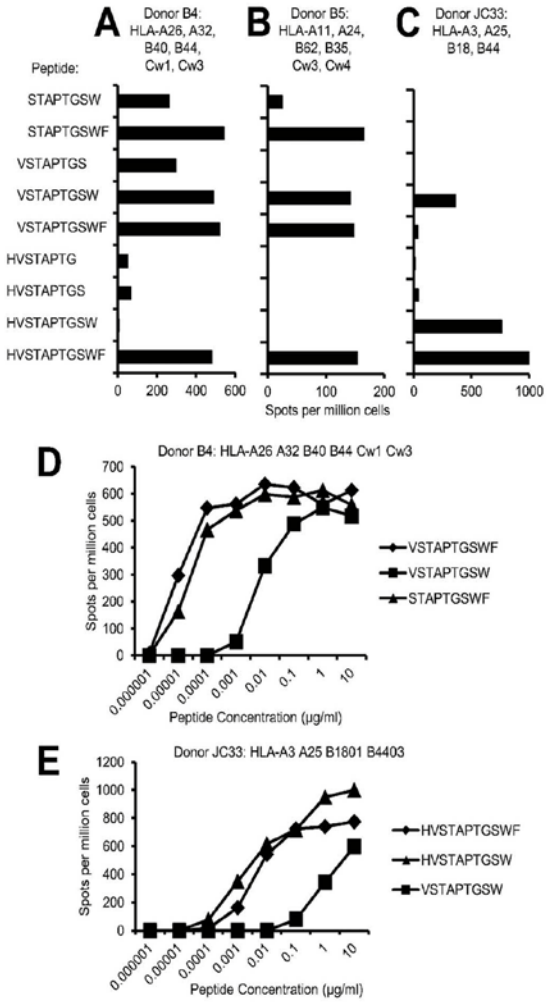


FIG 2 Overlapping length-variant epitopes mapped from the BZLF1 antigen of EBV. (A to C) EBV-specific T cell lines from three individuals, raised by *in vitro* stimulation with irradiated autologous LCLs, were tested in IFN-γ ELISpot assays for recognition of overlapping peptides (1 μg/ml) corresponding to a region of the BZLF1 protein sequence. (D and E) EBV-specific T cell lines from donor B4 (D) or donor JC33 (E), tested with IFN-γ ELISpot assays to precisely map the minimal epitopes. Cells were tested with various concentrations of the indicated peptides. The experiments described for Fig. 2C, D, and E were each performed with one other donor expressing the relevant HLA allele, with similar results.

sponding to the BZLF1 protein sequence for recognition by T cell lines. These T cells were raised by *in vitro* stimulation of PBMCs from 36 healthy EBV-seropositive individuals with their autologous LCLs (Table 1). Results for all individuals screened with this overlapping peptide set are shown in Fig. 1A to C. Confirming the importance of this small antigen as a target for T cell recognition, only 8% (3/36) of screened individuals showed no response to any of the BZLF1 overlapping peptides. It is notable that these three individuals also showed no responses to overlapping peptides from several other EBV antigens that were screened alongside the BZLF1 peptides (data not shown), so this may reflect a general weakness in the EBV-specific T cell response in these three donors.

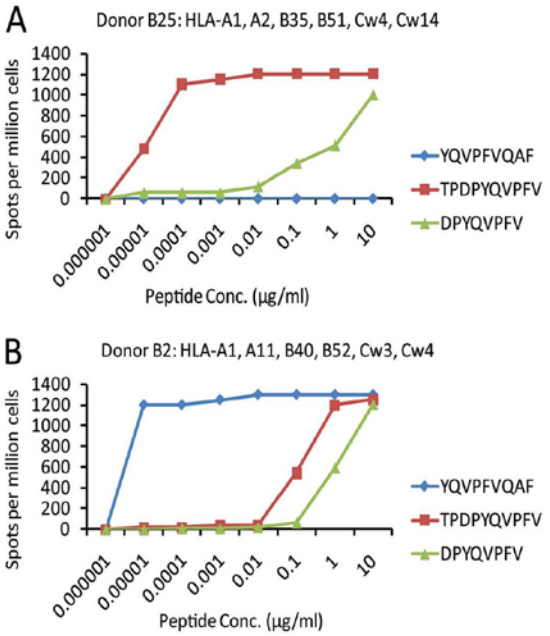


FIG 3 Identification of overlapping minimal epitopes from the BZLF1 antigen of EBV. EBV-specific T cell lines were screened with various concentrations of the indicated peptides by IFN-γ ELISpot assays. Results shown are from donor B25 (A) or donor B2 (B). These experiments were each performed with one other donor expressing the relevant HLA allele, with similar results.

It is also clear that the T cell response was biased toward restricted regions of this antigen, with a strong focus on residues 160 to 220.

We next aimed to map the epitopes more precisely within the BZLF1 20-mer sequences. Several individuals responded to the overlapping peptides ^{56}LPQGQLTAYHVSTAPTGSWF^{75} and ^{61}LTAYHVSTAPTGSWFSAPQP^{80}, and so smaller peptides from the overlapping region were tested for recognition by T cell lines from three of these donors. Figure 2A to C shows results from donors B4 (HLA-A26, A32, B40, B4402, Cw1, Cw3), B5 (HLA-A11, A24, B62, B35, Cw3, Cw4) and JC33 (HLA-A3, A25, B18, B44). Since the pattern of recognition was quite distinct between B4/B5 and JC33, it appeared that more than one epitope was localized in this region. Subsequent peptide dose-response IFN-γ ELISpot assays (Fig. 2D and E) confirmed the presence of two overlapping epitopes. The data indicate that the minimal sequence recognized by donor B4 was the ^{67}STAPTGSWF^{75} nonamer peptide. This novel epitope conforms to the binding motif of HLA-Cw3, an HLA allele shared by donors B4 and B5 (Fig. 2D) (15). However, we cannot rule out the possibility that the ^{66}VSTAPTGSWF^{75} sequence is also an epitope for this donor, because the dose-response curves for these two sequences are very similar. In contrast, T cells from donor JC33 most efficiently recognized the overlapping ^{65}HVSTAPTGSW^{74} decamer, which conforms to the HLA-A25 binding motif (Fig. 2E) (16). Interestingly, Abbott et al. recently published data mapping an epitope to the ^{66}VSTAPTGSWF^{75} sequence, with an ascribed HLA restriction of HLA-B*58:01 (17); however, it is notable that the responding donor in this study was also HLA-Cw3^{+}.

Subtle differences in the recognition of the overlapping 20-mers corresponding to the extreme N terminus of BZLF1 were

TABLE 2 BZLF1 $CD8^+$ T cell epitopes

Epitope coordinates (aa)	Epitope sequence	Likely HLA restriction	Reference
14–23	TPDPYQVPFV	B51	New data
18–26	YQVPFVQAF	Cw3	New data
52–64	LPEPLPQGQLTAY	B35.08	13
54–64	EPLPQGQLTAY	B35.01	19
65–74	HVSTAPTGSW	A25	New data
67–75	STAPTGSWF	Cw3	New data
77–86	APQPAPENAY	B35.08	20
79–89	QPAPENAYQAY	B35.01/08	21
122–130	VQTAAAVVF	B62(15.01)	New data
123–130	QTAAAVVF	B58	New data
169–180	EECDSELEIKRY	B44.03	14
173–180	SELEIKRY	B18.01	14
179–187	RYKNRVASR	A31	New data
188–196	KCRAKFKQL	B7	New data
189–197	CRAKFKQLL	Cw6	New data
190–197	RAKFKQLL	B8	22
192–200	KFKQLLQHY	A30	New data
196–205	LQHYREVAA	Cw8	11
209–217	SENDRLRLL	B60	23
209–217	SENDRLRLL	B49	New data

also noted between donors B2 and B25. While donor B2 showed strong responses toward peptides 11 to 30 and 16 to 35, donor B25 responded to peptides 6 to 25 and 11 to 30 (Fig. 1A). To investigate the basis for these differences, truncated versions of these peptides were assessed for T cell recognition, and distinct epitopes were mapped for each donor. As shown in Fig. 3A, ^{14}TPDPYQVPFV^{23} is a strongly recognized target epitope for T cells from donor B25, with a likely HLA restriction of HLA-B51 based on the published peptide binding motif for this HLA molecule (18). In contrast, the overlapping ^{18}YQVPFVQAF^{26} nonamer is the more potent target epitope for donor B2, with HLA-Cw4 the likely restricting molecule due to its preference for binding peptides with the primary anchor residue Phe at the C terminus (Fig. 3B) (15). With a total of six epitopes mapped to the first 75 amino acids of the BZLF1 protein (Table 2), this N-terminal region is clearly an important target for the T cell response.

T cell recognition for some donors was also focused on a central region of the antigen, between residues 106 and 130 (Fig. 1). T cells from donors MR67 and DP55 were found to recognize the nonamer ^{122}VQTAAAVVF^{130} (Fig. 4A and B), which conforms to the binding motif of HLA-B62 (24), an allele shared by both donors. This epitope encompasses another T cell determinant, recognized by donor B26: an octamer (^{123}QTAAAVVF^{130}) which conforms to the HLA-B58 binding motif (25) (Fig. 4C). Abbott and colleagues noted that the decamer ^{121}TVQTAAAVVF^{130} stimulated T cell responses in an HLA-$B62^+$ individual, supporting our conclusion that this region of BZLF1 includes an HLA-B62-restricted epitope (17).

"Hot spot" of T cell epitopes between residues 161 to 225 of BZLF1. In 1995, Bogedain et al. first described the highly immunogenic HLA-B8-restricted ^{190}RAKFKQLL^{197} epitope and, in the same study, an HLA-Cw6-restricted response was localized within the sequence ^{187}RKCRAKFKQLLQHYR^{201} (22). Our results confirm the presence of an HLA-Cw6-restricted response in this region of BZLF1 and demonstrate the minimal sequence to be ^{189}CRAKFKQLL^{197}, which encompasses the HLA-B8-binding octamer epitope (Fig. 5A and C). T cell lines from three HLA-$Cw6^+$ and HLA-$B8^-$ donors demonstrated IFN-γ responses to almost all overlapping peptides that included the nonamer ^{189}CRAKFKQLL^{197} (Fig. 5A). In addition, flow cytometric analysis using the APC-conjugated HLA-Cw*06-CRAKFKQLL tetramer confirmed that HLA-$Cw6^+$ PBMCs and EBV-specific T cell lines included significant populations of cells that recognized the novel nonamer peptide (Fig. 5C).

An additional epitope from BZLF1 that overlaps with ^{190}RAKFKQLL^{197} and ^{189}CRAKFKQLL^{197} was discovered from a donor who was negative for both HLA-B8 and Cw6. The results from donor B20 show that all overlapping peptides that included the nonamer ^{188}KCRAKFKQL^{196} stimulated T cell IFN-γ production (Fig. 5B). Since this sequence did not conform to the peptide binding motif of any of the HLA alleles expressed by this donor,

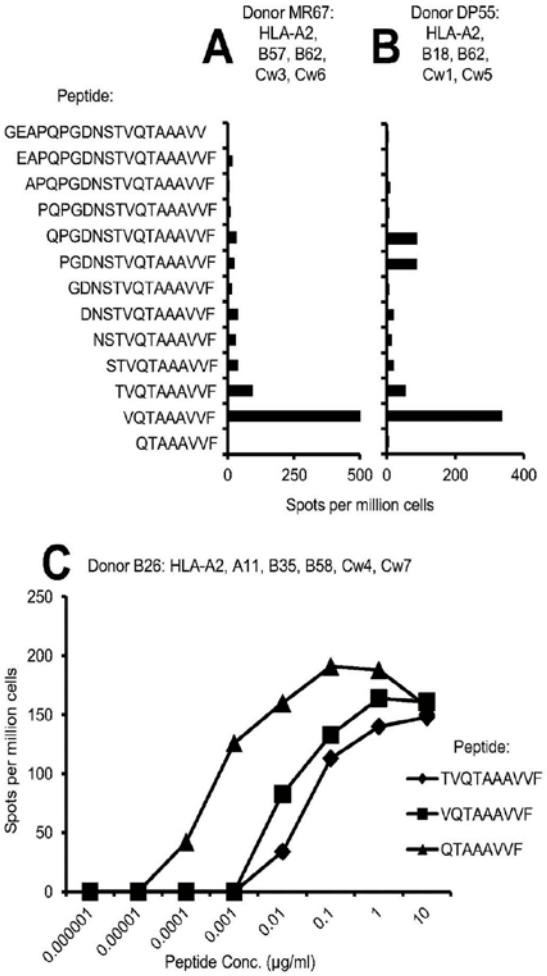


FIG 4 Mapping of minimal epitopes from the central region of the BZLF1 protein. IFN-γ ELISpot assays were used to define minimal epitopes by screening EBV-specific T cell lines from HLA-$B62^+$ individuals, donor MR67 (A) and donor DP55 (B), against truncated peptides at 1 μg/ml and a T cell line from the HLA-$B58^+$ individual, donor B26 (C), against various concentrations of the indicated peptides. The experiment described for panel C was performed with one other donor expressing the relevant HLA allele, with similar results.

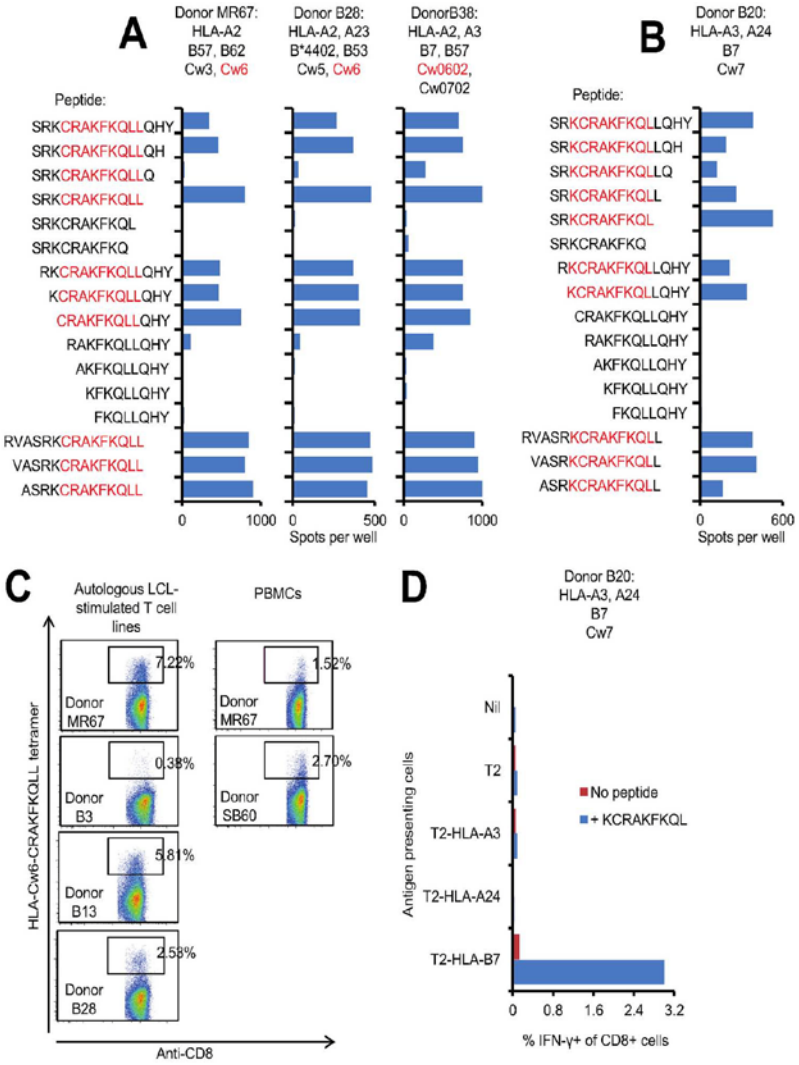


FIG 5 Overlapping epitopes in the coiled-coil dimerization domain of the BZLF1 EBV antigen. (A and B) EBV-specific T cells from three HLA-Cw6$^+$ individuals (A) and an HLA-B7$^+$ individual (B) were tested in IFN-γ ELISpot assays against overlapping peptides (1 μg/ml) corresponding to a region of the BZLF1 protein sequence. (C) PBMCs and EBV-specific T cell lines from HLA-Cw6$^+$ individuals were analyzed by flow cytometry with the HLA-Cw6-CRAKFKQLL tetramer. (D) HLA-deficient T2 cell lines transfected to express either HLA-A3, A24, or B7 were presensitized with the ^{188}KCRAKFKQL^{196} peptide (0.1 μg/ml) or left uncoated and then washed and exposed to EBV-specific T cells from donor B20. After incubation, these T cells were analyzed for IFN-γ production by intracellular cytokine staining.

we utilized the HLA-deficient T2 cell line, which was transfected to express either HLA-A3, A24, or B7 as peptide-presenting cells in order to determine the HLA restriction of this response. Intracellular cytokine staining showed that a T cell line from donor B20, which was raised by *in vitro* stimulation with an autologous LCL, recognized the ^{188}KCRAKFKQL^{196} peptide when presented by T2-HLA-B7 but not other antigen-presenting cells (Fig. 5D).

Since these three epitopes differed by just one or two amino acids, experiments were conducted to assess if T cells have the capacity to cross-react between them or if the responses are distinct and peptide specific. As shown in Fig. 6A to C, EBV-specific T cell populations from individuals expressing either HLA-B8, Cw6, or B7 were specific for ^{190}RAKFKQLL^{197}, ^{189}CRAKFKQLL^{197}, or ^{188}KCRAKFKQL^{196}, respectively, although the HLA-B8$^+$ T cells also recognized ^{189}CRAKFKQLL^{197} at relatively high concentrations (Fig. 6A). These data indicate that distinct T cell receptor repertoires are utilized against each peptide-HLA complex.

The Gln residue at position 195 of BZLF1 is mutated to a charged His residue in type 2 EBV strains and the dominant type 1 EBV strains in China. To investigate the impact of this sequence polymorphism on the HLA-Cw6-restricted response, T cells from donor MR67 were tested for recognition of various concentrations of the ^{189}CRAKFKQLL^{197} epitope versus the ^{189}CRAKFKHLL^{197} variant (the variant residue is underlined). As shown in Fig. 6D, this single amino acid polymorphism had a significant impact on T cell recog-

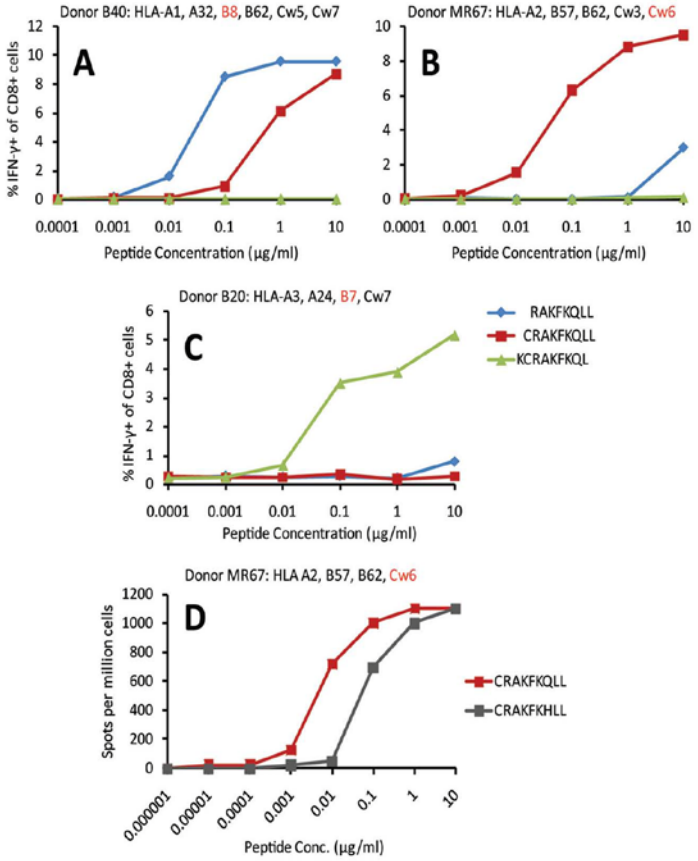


FIG 6 T cell recognition of the clustered overlapping epitopes between residues 188 and 197. EBV-specific T cell populations from individuals expressing either HLA-B8 (A), Cw6 (B), or B7 (C) were tested for recognition of the peptides ^{190}RAKFKQLL^{197}, ^{189}CRAKFKQLL^{197}, and ^{188}KCRAKFKQL^{196} at various concentrations by intracellular cytokine staining. (D) T cells from the HLA-Cw6$^+$ donor MR67 were also tested in an ELISpot assay for recognition of the variant peptide ^{189}CRAKFKHLL^{197} (the variant residue is underlined), in comparison to the results with ^{189}CRAKFKQLL^{197} epitope.

nition, with approximately 10-fold-higher concentrations of the variant peptide required for equivalent T cell recognition.

The intense immunogenicity of this small region of BZLF1 was further highlighted by the mapping of a fourth epitope overlapping with these three T cell determinants. Cells from the HLA-B8$^+$ donor B9 responded not only to the ^{190}RAKFKQLL^{197} epitope (data not shown) but also to a distinct overlapping determinant further toward the protein C terminus. Figure 7 displays data from donor B9, screened for T cell recognition of 8-, 9-, 10-, and 11-aa peptides from this region. The best-recognized peptides in the initial screen (FKQLLQHYR, QLLQHYREV, and KFKQLLQHY) were further examined by peptide titration ELISpot assays, confirming the nonamer ^{192}KFKQLLQHY^{200} as the minimal epitope, with a likely restriction of HLA-A30 based on the published peptide binding motif for this HLA molecule (26).

The final epitope to be mapped was from the only HLA-A31$^+$ donor included in the study. Donor B10 responded to this region of BZLF1 (peptides 171 to 190 and 176 to 195) (Fig. 1), and the target epitope was mapped to the nonamer ^{179}RYKNRVASR^{187}. This epitope is likely to be presented by HLA-A31 based on the published motif, which has an Arg primary anchor at the C terminus and Tyr and Val as secondary anchors at the second and sixth residues, respectively (Fig. 7C) (27). This epitope lies directly adjacent to the ^{188}KCRAKFKQL^{196} epitope, confirming the intense focusing of CD8$^+$ T cell epitopes in this C-terminal region of BZLF1.

DISCUSSION

The BZLF1 antigen of EBV is an essential component of the viral life cycle, since it plays an indispensable part in the switch from latent to lytic gene expression. The present study explored the T cell response to this protein by testing a large cohort of individuals against an overlapping peptide library covering the entire antigen and using T cell lines raised against autologous LCLs to optimize the sensitivity of the ELISpot assays. Our data provide an extensive list of novel epitopes within this antigen, with evidence of epitope clustering and responses restricted through a large number of HLA alleles (Fig. 8).

BZLF1 is clearly an important target for the T cell response to EBV and should therefore be considered in epitope-based vaccine strategies. Of the nine previously defined BZLF1 CD8$^+$ T cell epitopes (Table 2), restriction was limited to seven HLA alleles, including both HLA-B and -C alleles but not HLA-A alleles. In this study, we have defined a further 11 BZLF1 epitopes (Table 2),

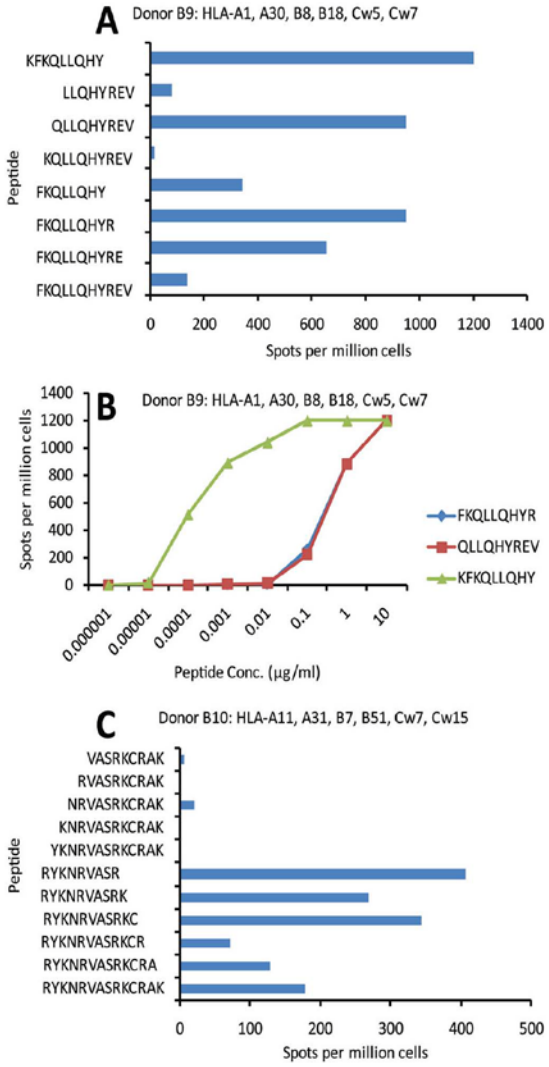


FIG 7 Identification of two HLA-A-restricted epitopes in the BZLF1 EBV antigen. (A) EBV-specific T cells from the HLA-A30$^+$ individual, donor B9, were tested by IFN-γ ELISpot assays against overlapping peptides (1 μg/ml) corresponding to a region of the BZLF1 antigen. (B) Subsequent IFN-γ ELISpot assays mapped the minimal epitope by screening the indicated peptides at various concentrations. (C) EBV-specific T cells from the HLA-A31$^+$ individual, donor B10, were tested by IFN-γ ELISpot assays with overlapping peptides (1 μg/ml) as shown. The experiment described for panel A was performed with one other donor expressing the relevant HLA allele, with similar results.

including epitopes restricted through HLA-A as well as HLA-B and -C alleles, thereby significantly increasing the population coverage. Indeed, 8 out of the 36 individuals tested in this study showed responses directed toward these novel epitopes but not previously defined epitopes.

Woodberry and colleagues showed that BZLF1 responses are significant in acute infectious mononucleosis, as well as during persistent EBV infection (28). Considering that the BZLF1 protein is only 245 aa in length, it is notable that almost half of these amino acids (105, or 42.9%) are included in epitopes. The transactivation domain (1 to 166) includes 10 epitopes, six of which were defined in this study. Furthermore, in a region of just 37 aa between residues 169 and 205, up to eight epitopes are present. This region includes the DNA recognition domain (178 to 194) and part of the coiled-coil dimerization domain. Located within the coiled-coil dimerization domain is the previously mapped HLA-B60-restricted nonamer epitope ^{209}SENDRLRLL^{217} (23). Interestingly, we have also found an HLA-B60-negative individual whose cells responded to this epitope (data not shown). Cells from this individual (donor B8; HLA-A2, A66, B4402, B49, Cw7) expressed HLA-B49, which is likely to be the restricting allele based on the binding motif of HLA-B49, which includes anchor residues Glu at position 2 and Leu at the C terminus. It is notable that McDonald and colleagues found the motif ^{209}SENDRLR^{215}, located midway along the coiled-coil dimerization region, to play a key role in the BZLF1 structure and to be required for EBV DNA replication (29). In addition, the leucine at residue 217 is suggested to be important for dimerization, since it is positioned adjacent to the interacting faces of the helices (10). This presumably explains why this sequence and the nonamer epitope are conserved in both type 1 and 2 EBV strains.

Why are some regions of the BZLF1 sequence so rich in CD8$^+$ T cell epitopes while other areas are not targeted at all? One possibility is that sequence similarity with human protein sequences results in immune tolerance to parts of the BZLF1 antigen. Several previous studies have shown the phenomenon of epitope clustering. Examples range from bacteria and tumor antigens to HIV. Kim and DeMars found that the major outer membrane protein of *Chlamydia trachomatis* displays a coclustering of class I and class II epitopes within a 20-mer region of the protein (30). Furthermore, Valmori and colleagues identified clusters of closely overlapping epitopes in the NY-ESO-1 protein (expressed in several types of tumors) (31). HIV studies have shown epitope clustering in the Nef and Gag proteins, corresponding to areas of protein hydrophobicity. Interestingly, CD8$^+$ T cell epitopes are predicted to be quite evenly distributed throughout the BZLF1 sequence according to the Immune Epitope Analysis Resource (http://tools.immuneepitope.org/main/) for 10 common HLA class I alleles (data not shown), which uses predictors of proteasomal processing, TAP transport, and MHC binding to produce an overall score for each peptide's intrinsic potential for being a T cell epitope.

HLA polymorphism has evolved to ensure that a wide range of antigenic epitopes are presented across the population in order to reduce the impact of amino acid mutation within T cell epitopes. The clustering of epitopes and the sharing of amino acids between multiple epitopes has the potentially dangerous consequence that viral escape mutants may impact on a large proportion of the population. An example of this is the Gln residue at position 195 of BZLF1, which is included within four epitopes restricted by either HLA-A30, -B7, -B8, or -Cw6. Importantly, the Gln residue is mutated to a charged His residue in type 2 EBV strains and the dominant type 1 EBV strains in China (32). Since HLA-Cw6 is expressed at a high frequency in parts of China, we examined the impact of this sequence polymorphism on recognition by HLA-Cw6-restricted T cells raised against the ^{189}CRAKFKQLL^{197} epitope. T cell recognition was significantly reduced by the Q^{195}-to-H^{195} variation, raising the interesting possibility that this variant may have arisen in the Chinese population as an escape mutant.

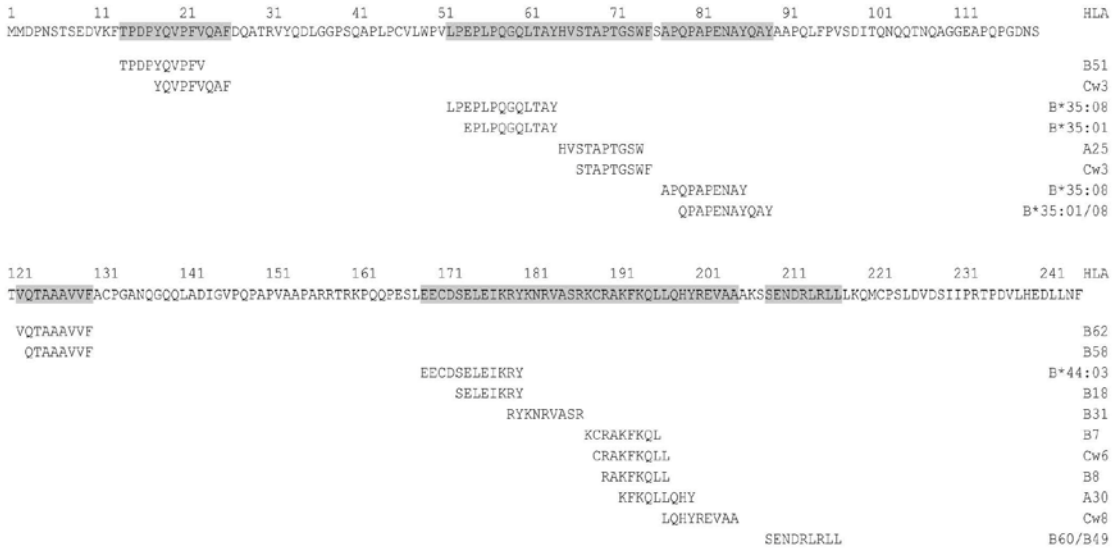


FIG 8 Clustered distribution of CD8$^+$ T cell epitopes within BZLF1. The 245-amino-acid sequence of BZLF1 is derived from the B95-8 strain of EBV. Regions that are recognized by CD8$^+$ T cells are shaded in gray and represent 105 amino acids (42.9% of the protein). Epitopes are listed below the full sequence, and the likely HLA restriction of each epitope is shown on the right of the figure.

Our work adds significantly to the understanding of the immune response to BZLF1, the most highly immunogenic T cell target antigen from EBV. Precisely mapped T cell epitopes have become a critical tool for analyzing the immune response to pathogens through, for example, HLA-peptide multimers. Furthermore, they are also being used directly in vaccine development and adoptive immunotherapy, particularly for tumor-associated viruses, such as EBV, from which individual proteins are potentially oncogenic (33). We have greatly expanded the number of defined CD8$^+$ T cell epitopes from BZLF1 and highlighted their clustered distribution along the length of the sequence. This information is likely to be widely utilized in future studies aimed at manipulating the immune system to treat or prevent EBV-associated diseases.

ACKNOWLEDGMENTS

S.R.B. is an NHMRC principal research fellow (APP1021452). This work was supported by an NHMRC project grant (APP1021620) and an NHMRC postgraduate scholarship to M.J.R. (APP1017834).

The following tetramer was obtained through the NIH Tetramer Facility: HLA-Cw*06-CRAKFKQLL.

REFERENCES

1. Hislop AD, Taylor GS, Sauce D, Rickinson AB. 2007. Cellular responses to viral infection in human: lessons from Epstein-Barr virus. Annu Rev Immunol 25:587–617. http://dx.doi.org/10.1146/annurev.immunol.25.022106.141553.
2. Long HM, Taylor GS, Rickinson AB. 2011. Immune defence against EBV and EBV-associated disease. Curr Opin Immunol 23:258–264. http://dx.doi.org/10.1016/j.coi.2010.12.014.
3. Kutok JL, Wang F. 2006. Spectrum of Epstein-Barr virus-associated diseases. Annu Rev Pathol 1:375–404. http://dx.doi.org/10.1146/annurev.pathol.1.110304.100209.
4. Henle G, Henle W, Diehl V. 1968. Relation of Burkitt's tumor-associated herpes-type virus to infectious mononucleosis. Proc Natl Acad Sci U S A 59:94–101. http://dx.doi.org/10.1073/pnas.59.1.94.
5. Rickinson A, Kieff E. 2007. Epstein-Barr virus, p 2655–2700. *In* Knipe DM, Howley PM (ed), Fields virology, 5th ed, vol 2. Lippincott, Williams and Wilkins, Philadelphia, PA.
6. Cox MA, Leahy J, Hardwick JM. 1990. An enhancer within the divergent promoter of Epstein-Barr virus responds synergistically to the R and Z transactivators. J Virol 64:313–321.
7. Zalani S, Holley-Guthrie E, Kenney S. 1996. Epstein-Barr viral latency is disrupted by the immediate-early BRLF1 protein through a cell-specific mechanism. Proc Natl Acad Sci U S A 93:9194–9199. http://dx.doi.org/10.1073/pnas.93.17.9194.
8. Farrell PJ, Rowe DT, Rooney CM, Kouzarides T. 1989. Epstein-Barr virus BZLF1 trans-activator specifically binds to a consensus AP-1 site and is related to c-fos. EMBO J 8:127–132.
9. Flemington EK, Borras A, Lytle MJP, Speck SH. 1992. Characterization of the Epstein-Barr virus BZLF1 protein transactivation domain. J Virol 66:922–929.
10. Flemington E, Speck SH. 1990. Evidence for coiled-coil dimer formation by an Epstein-Barr virus transactivator that lacks a heptad repeat of leucine residues. Proc Natl Acad Sci U S A 87:9459–9463. http://dx.doi.org/10.1073/pnas.87.23.9459.
11. Pudney VA, Leese AM, Rickinson AB, Hislop AD. 2005. CD8+ immunodominance among Epstein-Barr virus lytic cycle antigens directly reflects the efficiency of antigen presentation in lytically infected cells. J Exp Med 201:349–360. http://dx.doi.org/10.1084/jem.20041542.
12. Tan LC, Gudgeon N, Annels NE, Hansasuta P, O'Callaghan CA, Rowland-Jones S, McMichael AJ, Rickinson AB, Callan MF. 1999. A re-evaluation of the frequency of CD8+ T cells specific for EBV in healthy virus carriers. J Immunol 162:1827–1835.
13. Green KJ, Miles JJ, Tellam J, van Zuylen WJ, Connolly G, Burrows SR. 2004. Potent T cell response to a class I-binding 13-mer viral epitope and the influence of HLA micropolymorphism in controlling epitope length. Eur J Immunol 34:2510–2519. http://dx.doi.org/10.1002/eji.200425193.
14. Rist MJ, Theodossis A, Croft NP, Neller MA, Welland A, Chen Z, Sullivan LC, Burrows JM, Miles JJ, Brennan RM, Gras S, Khanna R, Brooks AG, McCluskey J, Purcell AW, Rossjohn J, Burrows SR. 2013. HLA peptide length preferences control CD8$^+$ T cell responses. J Immunol 191:561–571. http://dx.doi.org/10.4049/jimmunol.1300292.
15. Falk K, Rotzschke O, Grahovac B, Schendel D, Stevanovic S, Gnau V, Jung G, Strominger JL, Rammensee HG. 1993. Allele-specific peptide ligand motifs of HLA-C molecules. Proc Natl Acad Sci U S A 90:12005–12009. http://dx.doi.org/10.1073/pnas.90.24.12005.
16. Kruger T, Schoor O, Lemmel C, Kraemer B, Reichle C, Dengjel J, Weinschenk T, Muller M, Hennenlotter J, Stenzl A, Rammensee HG, Stevanovic

S. 2005. Lessons to be learned from primary renal cell carcinomas: novel tumor antigens and HLA ligands for immunotherapy. Cancer Immunol Immunother 54:826–836. http://dx.doi.org/10.1007/s00262-004-0650-5.
17. Abbott RJ, Quinn LL, Leese AM, Scholes HM, Pachnio A, Rickinson AB. 2013. CD8+ T cell responses to lytic EBV infection: late antigen specificities as subdominant components of the total response. J Immunol 191:5398–5409. http://dx.doi.org/10.4049/jimmunol.1301629.
18. Falk K, Rotzschke O, Takiguchi M, Gnau V, Stevanovic S, Jung G, Rammensee HG. 1995. Peptide motifs of HLA-B51, -B52 and -B78 molecules, and implications for Behcet's disease. Int Immunol 7:223–228. http://dx.doi.org/10.1093/intimm/7.2.223.
19. Saulquin X, Ibisch C, Peyrat MA, Scotet E, Hourmant M, Vie H, Bonneville M, Houssaint E. 2000. A global appraisal of immunodominant CD8 T cell responses to Epstein-Barr virus and cytomegalovirus by bulk screening. Eur J Immunol 30:2531–2539. http://dx.doi.org/10.1002/1521-4141(200009)30:9<2531::AID-IMMU2531>3.0.CO;2-O.
20. Tynan FE, Elhassen D, Purcell AW, Burrows JM, Borg NA, Miles JJ, Williamson NA, Green KJ, Tellam J, Kjer-Nielsen L, McCluskey J, Rossjohn J, Burrows SR. 2005. The immunogenicity of a viral cytotoxic T cell epitope is controlled by its MHC-bound conformation. J Exp Med 202:1249–1260. http://dx.doi.org/10.1084/jem.20050864.
21. Burrows JM, Bell MJ, Brennan R, Miles JJ, Khanna R, Burrows SR. 2008. Preferential binding of unusually long peptides to MHC class I and its influence on the selection of target peptides for T cell recognition. Mol Immunol 45:1818–1824. http://dx.doi.org/10.1016/j.molimm.2007.09.026.
22. Bogedain C, Wolf H, Modrow S, Stuber G, Jilg W. 1995. Specific cytotoxic T lymphocytes recognize the immediate-early transactivator Zta of Epstein-Barr virus. J Virol 69:4872–4879.
23. Scotet E, David-Ameline J, Peyrat MA, Moreau-Aubry A, Pinczon D, Lim A, Even J, Semana G, Berthelot JM, Breathnach R, Bonneville M, Houssaint E. 1996. T cell response to Epstein-Barr virus transactivators in chronic rheumatoid arthritis. J Exp Med 184:1791–1800. http://dx.doi.org/10.1084/jem.184.5.1791.
24. Prilliman K, Lindsey M, Zuo Y, Jackson KW, Zhang Y, Hildebrand W. 1997. Large-scale production of class I bound peptides: assigning a signature to HLA-B*1501. Immunogenetics 45:379–385. http://dx.doi.org/10.1007/s002510050219.
25. Falk K, Rotzschke O, Takiguchi M, Gnau V, Stevanovic S, Jung G, Rammensee HG. 1995. Peptide motifs of HLA-B58, B60, B61, and B62 molecules. Immunogenetics 41:165–168. http://dx.doi.org/10.1007/BF00182333.
26. Krausa P, Munz C, Keilholz W, Stevanovic S, Jones EY, Browning M, Bunce M, Rammensee HG, McMichael A. 2000. Definition of peptide binding motifs amongst the HLA-A*30 allelic group. Tissue Antigens 56:10–18. http://dx.doi.org/10.1034/j.1399-0039.2000.560102.x.
27. Falk K, Rotzschke O, Takiguchi M, Grahovac B, Gnau V, Stevanovic S, Jung G, Rammensee HG. 1994. Peptide motifs of HLA-A1, -A11, -A31, and -A33 molecules. Immunogenetics 40:238–241. http://dx.doi.org/10.1007/BF00167086.
28. Woodberry T, Suscovich TJ, Henry LM, Davis JK, Frahm N, Walker BD, Scadden DT, Wang F, Brander C. 2005. Differential targeting and shifts in the immunodominance of Epstein-Barr virus-specific CD8 and CD4 T cell responses during acute and persistent infection. J Infect Dis 192:1513–1524. http://dx.doi.org/10.1086/491741.
29. McDonald CM, Petosa C, Farrell PJ. 2009. Interaction of Epstein-Barr virus BZLF1 C-terminal tail structure and core zipper is required for DNA replication but not for promoter transactivation. J Virol 83:3397–3401. http://dx.doi.org/10.1128/JVI.02500-08.
30. Kim S-K, DeMars R. 2001. Epitope clusters in the major outer membrane protein of *Chlamydia trachomatis*. Curr Opin Immunol 13:429–436. http://dx.doi.org/10.1016/S0952-7915(00)00237-5.
31. Valmori D, Levy F, Godefroy E, Scotto L, Souleimanian NE, Karbach J, Tosello V, Hesdorffer CS, Old LJ, Jager E, Ayyoub M. 2007. Epitope clustering in regions undergoing efficient proteasomal processing defines immunodominant CTL regions of a tumor antigen. Clin Immunol 122:163–172. http://dx.doi.org/10.1016/j.clim.2006.09.005.
32. Yang Y, Jia Y, Wang Y, Wang X, Sun Z, Luo B. 2014. Sequence analysis of EBV immediate-early gene BZLF1 and BRLF1 in lymphomas. J Med Virol 86:1788–1795. http://dx.doi.org/10.1002/jmv.23911.
33. Smith C, Khanna R. 2012. A new approach for cellular immunotherapy of nasopharyngeal carcinoma. Oncoimmunology 1:1440–1442. http://dx.doi.org/10.4161/onci.21286.

Chapter 5:

Aim Three:

Determine if the peptide length preferences of class I human leukocyte antigens influence epitope selection in the EBV-specific T cell response.

Hypothesis Three:

Different class I human leukocyte antigens have distinct peptide length preferences, and this influences epitope selection in the EBV-specific T cell response.

Publication: Original research article

Rist MJ, Theodossis A, Croft NP, Neller MA, Welland A, Chen Z, Sullivan LC, Burrows JM, Miles JJ, Brennan RM, Gras S, Khanna R, Brooks AG, McCluskey J, Purcell AW, Rossjohn J, Burrows SR. HLA Peptide Length Preferences Control $CD8^+$ T Cell Responses. *The Journal of Immunology*. 2013 June; 191:561-571

HLA Peptide Length Preferences Control CD8$^+$ T Cell Responses

Melissa J. Rist,[*,†,1] Alex Theodossis,[‡,1] Nathan P. Croft,[‡,1] Michelle A. Neller,[*] Andrew Welland,[‡] Zhenjun Chen,[§] Lucy C. Sullivan,[§] Jacqueline M. Burrows,[*] John J. Miles,[*,†,¶] Rebekah M. Brennan,[*] Stephanie Gras,[‡] Rajiv Khanna,[*,†] Andrew G. Brooks,[§] James McCluskey,[§] Anthony W. Purcell,[‡] Jamie Rossjohn,[‡,¶] and Scott R. Burrows[*,†]

Class I HLAs generally present peptides of 8–10 aa in length, although it is unclear whether peptide length preferences are affected by HLA polymorphism. In this study, we investigated the CD8$^+$ T cell response to the BZLF1 Ag of EBV, which includes overlapping sequences of different size that nevertheless conform to the binding motif of the large and abundant HLA-B*44 supertype. Whereas HLA-B*18:01$^+$ individuals responded strongly and exclusively to the octamer peptide ^{173}SELEIKRY^{180}, HLA-B*44:03$^+$ individuals responded to the atypically large dodecamer peptide ^{169}EECDSELEIKRY^{180}, which encompasses the octamer peptide. Moreover, the octamer peptide bound more stably to HLA-B*18:01 than did the dodecamer peptide, whereas, conversely, HLA-B*44:03 bound only the longer peptide. Furthermore, crystal structures of these viral peptide–HLA complexes showed that the Ag-binding cleft of HLA-B*18:01 was more ideally suited to bind shorter peptides, whereas HLA-B*44:03 exhibited characteristics that favored the presentation of longer peptides. Mass spectrometric identification of > 1000 naturally presented ligands revealed that HLA-B*18:01 was more biased toward presenting shorter peptides than was HLA-B*44:03. Collectively, these data highlight a mechanism through which polymorphism within an HLA class I supertype can diversify determinant selection and immune responses by varying peptide length preferences. *The Journal of Immunology*, **2013, 191: 561–571.**

In humans, MHC molecules, or HLAs, play a central role in Ag presentation and are characterized by a high level of polymorphism concentrated within the Ag-binding cleft. The six pockets (A–F) of the HLA class I Ag-binding cleft vary in their depth, electrostatic potential, and hydrophobicity, thereby determining the individual specificity of the peptide–HLA I interaction (1). For most HLA alleles, two of these pockets display a marked preference for one or two amino acids at certain positions within the peptide. In addition, residues at secondary anchor positions can also enhance or inhibit allele-specific binding (2, 3). A degree of degeneracy in HLA–peptide binding has been demonstrated, whereby multiple class I alleles can share common sequence motifs (referred to as supermotifs) owing to homology of amino acids within the major pockets of the peptide-binding cleft, and these groups of alleles are referred to as HLA supertypes. On the basis of these HLA structural similarities and overlapping peptide-binding motifs, nine major HLA supertypes have been proposed (4). The HLA-B*44 supertype, for example, includes the common alleles HLA-B*18:01 and HLA-B*44:03, which share a preference for peptides with Glu at position 2 and Phe or Tyr at the C terminus (5–8).

MHC class I–peptide binding also involves a conserved hydrogen bonding network at the peptide N and C termini and MHC H chain residues within pockets A and F, respectively, that is highly conserved between different class I allomorphs. These interactions contribute more binding energy to MHC class I–peptide complexes than do peptide side-chain interactions, and are thought to be the major factor limiting the size of peptides presented by MHC class I molecules to predominantly 8–10 aa in length (9). It is clear, however, that some longer peptides can also bind (10–15). Indeed, several MHC class I–bound peptides that are between 11 and 16 residues in length have been characterized structurally, and in each case the peptide was shown to adopt a bulging conformation whereby the "extra length" between the primary anchor residues protrudes outward from the cleft, thereby maintaining the conserved hydrogen bonding of the H chain with the peptide N and C termini (16–25). Thus, central peptide residues often make no, or limited, contact with the MHC within these complexes and, accordingly, can show a high degree of flexibility (16, 21). These bulged epitopes are considered to represent challenging targets for TCR ligation (26, 27)

*Centre for Immunotherapy and Vaccine Development, Queensland Institute of Medical Research, Brisbane, Queensland 4029, Australia; †School of Medicine, University of Queensland, Brisbane, Queensland 4006, Australia; ‡Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia; §Department of Microbiology and Immunology, University of Melbourne, Parkville, Victoria 3010, Australia; and ¶Institute of Infection and Immunity, Cardiff University School of Medicine, Heath Park, Cardiff CF14 4XN, United Kingdom

[1]M.J.R., A.T., and N.P.C. contributed equally to this work.

Received for publication January 30, 2013. Accepted for publication May 13, 2013.

This work was supported by the National Health and Medical Research Council (NHMRC) of Australia and the Australian Research Council (ARC). J.J.M. is supported by an NHMRC Career Development Fellowship, S.R.B. by an NHMRC Principal Research Fellowship, A.W.P. by an NHMRC Senior Research Fellowship, S.G. by an ARC Future Fellowship, and J.R. by an NHMRC Australia Fellowship.

The atomic coordinates and structure factors presented in this article have been submitted to the Protein Data Bank (http://www.pdb.org/pdb/home/home.do) under accession numbers 4JQV and 4JQX.

Address correspondence and reprint requests to Prof. Scott R. Burrows, Prof. Jamie Rossjohn, or Prof. Anthony W. Purcell, Queensland Institute of Medical Research, Brisbane, QLD 4029, Australia (S.R.B.), or Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, VIC 3800, Australia (J.R. and A.W.P.). E-mail addresses: scott.burrows@qimr.edu.au (S.R.B), jamie.rossjohn@monash.edu (J.R.), or anthony.purcell@monash.edu (A.W.P.)

The online version of this article contains supplemental material.

Abbreviations used in this article: LCL, lymphoblastoid cell line; Tm, melting temperature.

www.jimmunol.org/cgi/doi/10.4049/jimmunol.1300292

and are generally recognized by specific TCRs with a capacity to accommodate the atypical bulged peptide conformation (28).

The BZLF1 or ZEBRA (Z EBV replication activator) protein of EBV is responsible for the switch from latent to lytic gene expression and is an important target for T cell recognition. In this article, we show that the widely expressed HLA-B*44 supertype molecules, HLA-B*18:01 and HLA-B*44:03, present overlapping epitopes from BZLF1 that share a common C-terminal primary anchor residue but markedly differ in size. An explanation for this observation was provided by analysis of the self-peptide repertoire of these two closely related HLA molecules, revealing distinct peptide length preferences. Furthermore, structural analysis of the viral peptide–HLA complexes provided a molecular basis for this inherent bias for presenting peptides of different length. Collectively, we demonstrate that class I HLA polymorphism can skew peptide length preferences without altering anchor residue preferences, thereby having impact on epitope selection in an antiviral response and highlighting an additional mechanism through which the $CD8^+$ T cell responses to foreign Ags are diversified.

Materials and Methods

T cell cultures

PBMCs were isolated from healthy, EBV-seropositive blood donors after giving informed consent and approval was obtained from the Queensland Institute of Medical Research Human Research Ethics Committee (Brisbane, QLD, Australia). EBV-specific CTL cultures were raised from volunteers who were positive for either HLA-B*18:01 or HLA-B*44:03, as determined by HLA sequence analysis, by culturing PBMCs (2×10^6/2-ml well) with autologous lymphoblastoid cell lines (LCLs) (2×10^5/2-ml well). Cultures were supplemented with rIL-2 on day 7 and analyzed on day 10. CTLs were also stained with pHLA multimers by incubation for 30 min at 4°C with an allophycocyanin-labeled HLA-B*18:01–SELEIKRY or a PE-labeled HLA-B*44:03–EECDSELEIKRY tetramer. Cells were washed and analyzed on a FACSCanto flow cytometer with FACSDiva software (Becton Dickinson).

Cytotoxicity assays

CTLs from two donors were tested in duplicate for cytotoxicity in standard 5-h chromium release assays. In brief, CTLs were assayed against ^{51}Cr-labeled autologous PHA blast targets (E:T ratio = 2:1) that were pretreated with various concentrations of synthetic peptide or left untreated. Peptides were synthesized by Mimotopes. Toxicity testing of all peptides was performed before use by adding peptide to ^{51}Cr-labeled PHA blasts in the absence of CTL effectors. A β scintillation counter (TopCount Microplate; PerkinElmer) was used to measure ^{51}Cr levels in assay supernatant samples. The mean spontaneous lysis for target cells in the culture medium was always < 20%, and the variation from the mean specific lysis was < 10%.

ELISPOT assays

IFN-γ ELISPOT assays were performed using cytokine capture and detection reagents, according to the manufacturer's instructions (Mabtech). In brief, anti–IFN-γ Abs were coated on the wells of a 96-well nitrocellulose plate, and duplicate wells were seeded with T cells and peptide at 1 μM. After incubation for 16 h, captured IFN-γ was detected with a biotinylated anti–IFN-γ Ab, followed by development with streptavidin–HRP complex and chromogenic substrate, and spots were counted using an automated plate counter.

Thermal stability assay

Soluble constructs of the HLA-B*18:01–octamer and HLA-B*18:01–dodecamer binary complexes were expressed, refolded, and purified as described previously (29). Thermal denaturation of both complexes was monitored by circular dichroism spectroscopy over a temperature range of 20–90°C, using an experimentally determined absorption minimum of 218.5 nm. Data were acquired at two different protein concentrations for each of the complexes and then analyzed in PRISM (v5.0c) to determine their respective melting temperature (Tm) values. The raw data were normalized and transformed before being fitted to a dose–response curve by nonlinear regression.

Purification of MHC-bound peptides

C1R-B*18:01 and C1R-B*44:03 cell pellets ($3–5 \times 10^8$) were ground in a Retsch Mixer Mill MM 400 under cryogenic conditions; resuspended in 0.5% IGEPAL, 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, and protease inhibitors (Complete Protease Inhibitor Cocktail Tablet; Roche Molecular Biochemicals) at a density of 5×10^7 cells ml^{-1}; and incubated with rotation for 1 h at 4°C (15). Lysates were cleared by ultracentrifugation (180,000 × *g*) and HLA–peptide complexes sequentially immunoaffinity purified using solid-phase bound anti-Bw4 (Rm7.9.63) and anti-Bw6 (HB152-SRF) Abs. Bound complexes were washed and peptides eluted by acidification with 10% acetic acid, as described (14, 30, 31). The mixture of peptides and MHC protein chains was fractionated on a 4.6-mm internal diameter × 50-mm-long reversed-phase C18 HPLC column (Chromolith SpeedROD; Merck), using an ÄKTAmicro HPLC system (GE Healthcare) running on a mobile phase buffer A of 0.1% trifluoroacetic acid and buffer B of 80% acetonitrile/0.1% trifluoroacetic acid and at a flow rate of 1 ml/min.

Identification of MHC-bound peptides using LC–tandem mass spectrometry

Peptide-containing reversed-phase HPLC fractions were concentrated using a Labconco CentriVap concentrator. Samples were concentrated down to a volume of 15 μl 0.1% formic acid in water, sonicated in a water bath for 10 min, centrifuged for 10 min at 18,000 × *g*, and finally transferred to mass spectrometry vials for analysis by an AB SCIEX TripleTOF 5600 mass spectrometer equipped with a NanoSpray III ion source and coupled on-line to an Eksigent nanoLC-Ultra cHiPLC system (15, 30). Samples were loaded onto the microfluidic trap column packed with ChromXP C_{18}-CL 3-μm particles (300 Å nominal pore size; equilibrated in 0.1% formic acid/2% acetonitrile) at 5 μl min^{-1}. An analytical (15 cm × 75 μm ChromXP C_{18}-CL 3-μm particles) microfluidic column was then switched in line, and peptides were separated by linear gradient elution of 0–80% acetonitrile over 90 min (300 nl min^{-1}). Mass spectrometry conditions were set to accumulate up to 30 tandem spectrometry spectra per second. Data were analyzed with ProteinPilot software (AB SCIEX), and peptide identities were determined subject to strict bioinformatic criteria that included the use of a decoy database to calculate the false discovery rate. A false discovery rate cutoff of 5% was applied, and the filtered dataset was further analyzed manually to exclude redundant peptides, known contaminants, and nonspecific peptides (i.e., for C1R-B*18:01 those identified with anti-Bw4, and for C1R-B*44:03 those identified with anti-Bw6). MHC motif analysis was based on the frequency of amino acids occurring at P2 and PΩ.

Structural characterization

Soluble constructs of the HLA-B*18:01–octamer and HLA-B*44:03–dodecamer binary complexes were expressed, refolded, and purified according to standard protocols, described previously (29). Both complexes were crystallized, using the hanging drop vapor diffusion technique, in 0.1 M citrate (pH 5.6), 0.2 M NH4OAc, and 16–26% polyethylene glycol 4000, by seeding from existing crystals of HLA-B*44. Before flash freezing in liquid nitrogen, crystals were cryoprotected by equilibration in mother liquor supplemented with 10–20% glycerol. Complete datasets for HLA-B*18:01–octamer and HLA-B*44:03–dodecamer complexes were collected from single crystals at 100K, at the Advanced Photon Source in Chicago, IL, and the Australian Synchrotron in Melbourne, Australia, respectively. Both datasets were processed in MOSFILM (32) and SCALA (33). Molecular replacement was carried out in PHASER (34), using protein coordinates of HLA-B*44:03 [PDBid 1SYS (35)] as a search model. Refinement and model building were carried out in PHENIX (36) and COOT (37), respectively. Structure validation was performed using MOLPROBITY (38) and tools in the PHENIX package. A summary of data collection and refinement statistics is provided in Supplemental Tables I and II, respectively.

The molecular volume of each Ag-binding cleft (corresponding to the Connolly surface) was calculated using the CASTp server with a 1.4-Å probe radius (39). Electrostatic calculations were carried out using only the H chain and β_2-microglobulin coordinates of each structure. Briefly, coordinate preparation (placing and optimization of hydrogen atoms, assuming standard protonation at pH 7.0, followed by assignment of atomic charge and radii parameters using the PARSE forcefield) was accomplished using the PDB2PQR server (v1.8) (40). Electrostatic calculations were subsequently carried out using the APBS plugin in PyMOL (v1.1.0 and v1.2.x, respectively) with a 0.15 M concentration for the +1 and −1 ion species (41).

Accession numbers

The atomic coordinates and structure factors are available from the Protein Data Bank (http://www.pdb.org/pdb/home/home.do) under the following accession numbers: 4JQV and 4JQX.

Results

*EBV-specific T cell responses restricted through HLA-B*18:01 or HLA-B*44:03 recognize overlapping peptides of different length*

The BZLF1 protein from EBV stimulates large and well-characterized CTL responses restricted through HLA-B*08:01 and HLA-B*35:01 (42-46). Strong BZLF1-specific T cell responses restricted through HLA-B*18:01 and HLA-B*44:03 have also been described, but the precise epitope sequences have not been mapped (44). To address this issue, overlapping 20-aa peptides corresponding to the BZLF1 protein sequence were screened for recognition by T cell lines raised by in vitro stimulation of PBMCs from healthy EBV-exposed individuals with their autologous LCLs. These IFN-γ ELISPOT assays showed that donors expressing either HLA-B*44:03 or HLA-B*18:01 recognized two or three peptides between residues 161 and 190 of BZLF1 (Supplemental Fig. 1). The peptide ^{168}LEECDSELEIKRY^{180} from this region was later found to be recognized by both HLA-B*44:03$^+$ and HLA-B*18:01$^+$ EBV-specific T cells, and to more precisely map the target BZLF1 epitope within this 13-aa sequence, truncated versions of this peptide were screened for T cell recognition. All possible 8-, 9-, 10-, 11-, and 12-aa peptides from within this sequence were tested with T cells from five HLA-B*18:01$^+$ and three HLA-B*44:03$^+$ individuals, all of whom were EBV seropositive. Of interest, HLA-B*18:01$^+$ individuals recognized all peptides that included the ^{173}SELEIKRY^{180} octamer sequence, whereas HLA-B*44:03$^+$ individuals recognized the dodecamer peptide ^{169}EECDSELEIKRY^{180} strongly, but not shorter peptides (Fig. 1A). Note that both sequences have Glu at position 2 and the shared Tyr^{180} residue at the C terminus, which are the preferred primary anchor residues for both HLA-B*18:01 and HLA-B*44:03 (6, 7).

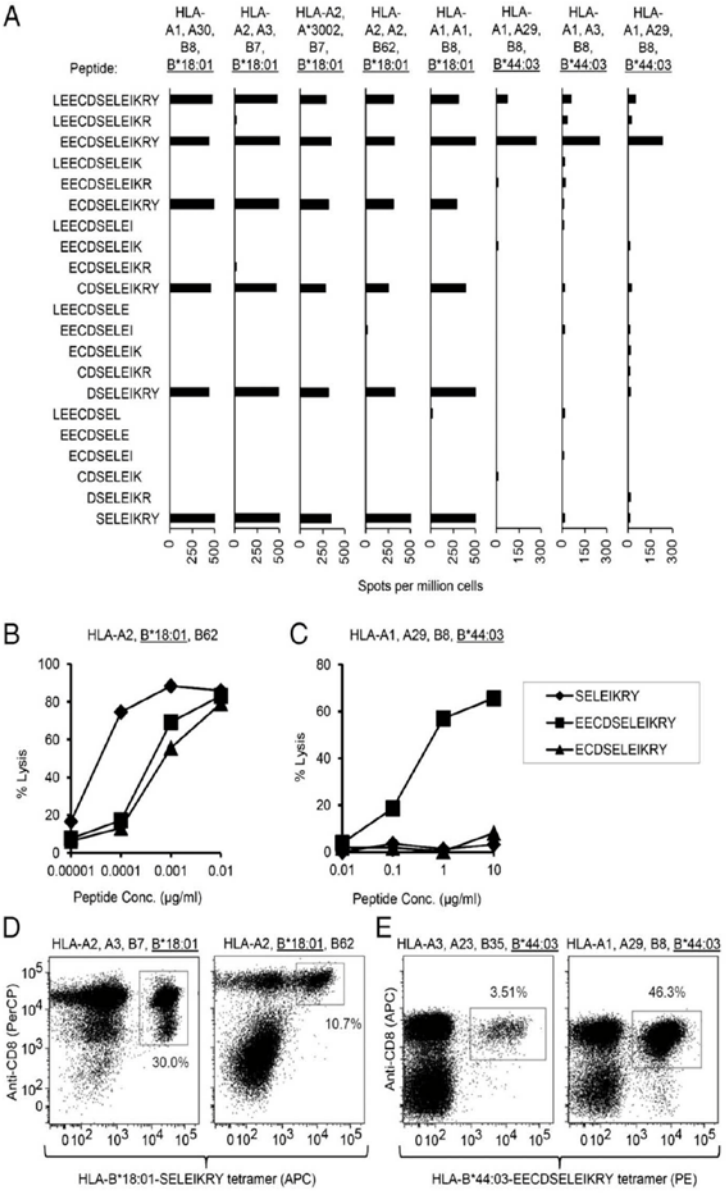


FIGURE 1. Overlapping length-variant epitopes mapped from the BZLF1 Ag of EBV for the HLA-B*44 supertype-restricted T cell response. (**A**) EBV-specific T cell lines from five HLA-B*18:01$^+$ and three HLA-B*44:03$^+$ individuals, raised by in vitro stimulation with irradiated autologous LCLs, tested in IFN-γ ELISPOT assays for recognition of overlapping peptides corresponding to a region of the BZLF1 protein sequence. EBV-specific T cell lines from (**B**) an HLA-B*18:01$^+$ and (**C**) an HLA-B*44:03$^+$ individual, tested in chromium-release assays for recognition of HLA-B*18:01$^+$ and HLA-B*44:03$^+$ target cells, respectively, treated with various concentrations of the indicated peptides. EBV-specific T cell lines from (**D**) two HLA-B*18:01$^+$ and (**E**) two HLA-B*44:03$^+$ individuals, analyzed by flow cytometry with tetramers of the HLA-B*18:01–SELEIKRY and HLA-B*44:03–EECDSELEIKRY complexes, respectively. These experiments were conducted at least twice with similar results.

One possible explanation for these patterns of T cell recognition is that HLA-B*18:01 presents SELEIKRY, whereas HLA-B*44:03 presents the unusually large EECDSELEIKRY peptide from the BZLF1 protein. To investigate this issue, we tested a T cell line from an HLA-B*18:01$^+$ individual (HLA-A2, B*18:01, B62), using chromium-release assays for recognition of HLA-B*18:01$^+$ target cells that had been treated with various concentrations of these two peptides. T cells from an HLA-B*44:03$^+$ individual (HLA-A1, A29, B8, B*44:03) were also tested, using HLA-B*44:03$^+$ target cells. The 11-aa peptide ECDSELEIKRY was included to verify that T cell recognition in the context of HLA-B*44:03 requires the full 12-aa sequence. These experiments confirmed that the octamer was the minimal peptide with maximal activity and therefore the likely target epitope presented by HLA-B*18:01 (Fig. 1B). In contrast, HLA-B*44:03$^+$ T cells lysed only target cells presenting the HLA-B*44:03–EECDSELEIKRY complex (Fig. 1C). Furthermore, flow cytometric analysis with tetramers of the HLA-B*18:01–SELEIKRY or HLA-B*44:03–EECDSELEIKRY complexes confirmed recognition of the different-length peptides by EBV-specific T cells restricted by each of these HLA-B*44 supertype alleles (Fig. 1D, 1E). The specificity of these tetramers for each pMHC complex was confirmed by including T cell lines from HLA-B*18:01$^-$ and HLA-B*44:03$^-$ individuals that showed no staining (data not shown).

To study mechanisms controlling the epitope selection differences of HLA-B*18:01 and HLA-B*44:03, we assessed the stability of the HLA–peptide complexes by measuring their thermal stability. HLA-B*18:01 was successfully refolded and purified in complex with both SELEIKRY and EECDSELEIKRY. The thermal unfolding properties of the two HLA-B*18:01 complexes were analyzed by circular dichroism spectroscopy (Fig. 2). Data for each complex were acquired at two different protein concentrations, resulting in an average Tm of 60.9°C ± 0.1°C for HLA-B*18:01–SELEIKRY and 52.2°C ± 1.9°C for HLA-B*18:01–EECDSELEIKRY. The resulting Tm difference of 8.7°C ± 2.8°C is consistent with the observed preference of this HLA allele for the shorter of the two peptide variants. Notably, attempts at refolding the HLA-B*44:03–SELEIKRY complex were unsuccessful, and so it was not possible to compare the binding stability of the length-variant peptides to this HLA molecule. However, this observation strongly indicates weak binding, which is consistent with the preference of HLA-B*44:03 for the 12-aa BZLF1 peptide.

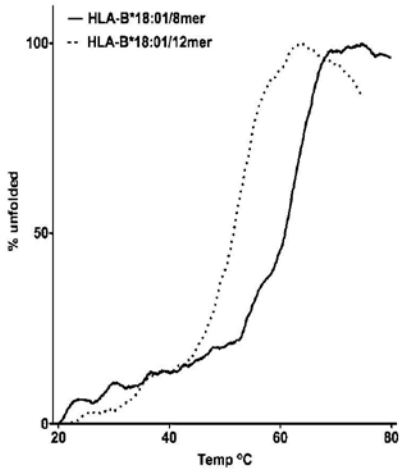


FIGURE 2. Thermal stability of the HLA-B*18:01–octamer and HLA-B*18:01–dodecamer complexes. Thermal unfolding of HLA-B*18:01 complexes as measured by circular dichroism spectroscopy at 218.5 nm. Representative curves for the 8-aa SELEIKRY peptide and the EECDSELEIKRY 12-aa peptide complexes are shown.

To indirectly assess presentation of the 8-aa and 12-aa epitopes, and determine which residues are important for T cell recognition, peptides into which Ala substitutions were introduced were tested for their capacity to activate T cells from HLA-B*18:01$^+$ or HLA-B*4403$^+$ individuals. Ala substitution at position 2 or the C terminus (pE2 and pY-Ω) reduced T cell activation by both peptides, and this is consistent with the likely role of these residues as primary anchor residues (Fig. 3). Other amino acid substitutions that negatively affected T cell recognition were at positions 3, 4, 5, 6, and 7 within the SELEIKRY peptide, and positions 6, 7, 8, and 11 within EECDSELEIKRY. These are therefore possible exposed residues contacted by the Ag receptors of the T cells used in these assays, or residues important in stabilizing the immunogenic conformation of the pHLA complex. Collectively, these data, together with the aforementioned data, provide proof-of-concept that polymorphism within an HLA class I supertype can diversify T cell epitope selection by influencing the preferred length of the presented peptide from within an Ag.

*HLA-B*18:01–SELEIKRY and HLA-B*44:03–EECDSELEIKRY structures*

To investigate the molecular basis for the preferential binding of the 8-aa and 12-aa peptides to HLA-B*18:01 and HLA-B*44:03, respectively, the structures of each pHLA binary complex were determined to high resolution (see Supplemental Tables I and II for data collection and refinement statistics). For the HLA-B*18:01–octamer complex, all residues corresponding to the SELEIKRY peptide were well ordered and clearly visible (Fig. 4A, 4B). The structure of the HLA-B*44:03–dodecamer complex revealed clear, unbiased density for the N and C termini of the

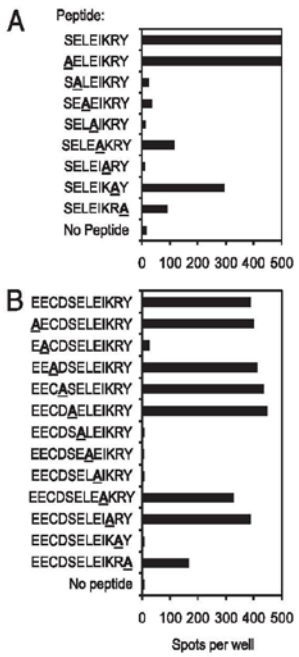


FIGURE 3. Impact on T cell recognition of Ala substitutions within the length-variant peptide epitopes. EBV-specific T cell lines from (**A**) an HLA-B*18:01$^+$ and (**B**) an HLA-B*44:03$^+$ individual, tested in IFN-γ ELISPOT assays for recognition of HLA-B*18:01$^+$ and HLA-B*44:03$^+$ target cells, respectively, treated with 1 μM of the peptides listed on the vertical axis. The Ala replacement residues are shown in bold underlined text. These experiments were conducted at least twice with similar results.

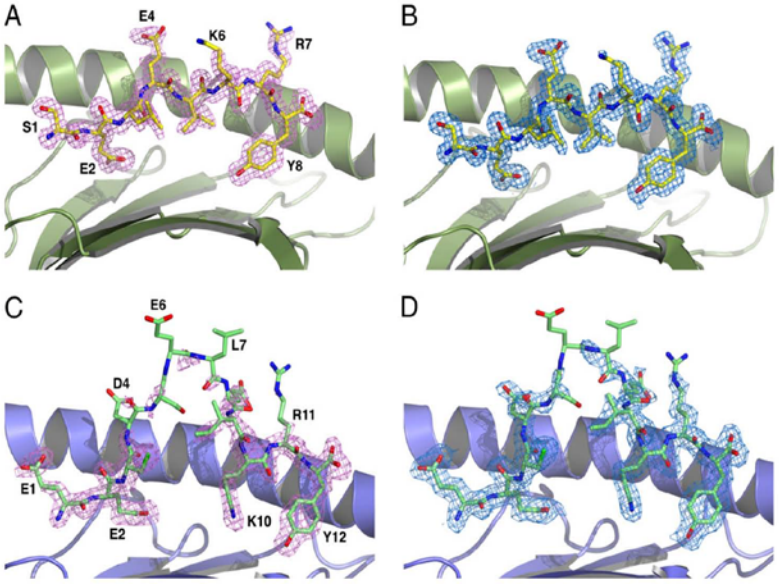


FIGURE 4. The crystal structures of HLA-B*18:01–SELEIKRY and HLA-B*44:03–EECDSELEIKRY. (**A**) Refined coordinates for the binary complex between HLA-B*18:01 (green sections) and the 8-aa peptide SELEIKRY (yellow sticks) overlaid with the unbiased Fo-Fc density for the peptide contoured at 2.5 σ (magenta mesh). (**B**) Refined coordinates for the HLA-B*18:01–8-aa complex overlaid with the final 2Fo-Fc density for the peptide contoured at 1 σ (blue mesh). In both images the α2-helix has been removed for clarity. (**C** and **D**) Equivalent images for the binary complex between HLA-B*44:03 (blue sections) and the 12-aa peptide EECDSELEIKRY (lime green sticks).

peptide (pE1-pD4 and pI9-pY12; Fig. 4C), whereas the bulged region was less well resolved, consistent with a greater degree of mobility of this 12-aa peptide (Fig. 4D).

In complex with HLA-B*18:01, the SELEIKRY peptide adopted an extended conformation in which the peptide was characterized by minimal intramolecular contacts, consistent with an unconstrained conformation (47). The side chains of pE2, pI5, and pY8 pointed down toward the floor of the cleft, whereas pL3 was directed toward the α2-helix. The side chains of pS1, pE4, pK6, and pR7 pointed out of the cleft, with pE4, pK6, and pR7 highly solvent accessible and thus likely to play a role in directly contacting the TCR, consistent with the observed impact on T cell recognition of their substitution by an Ala (Fig. 3A). The SELEIKRY peptide was anchored to the cleft of HLA-B*18:01 primarily by pE2 and pY8 (Fig. 5A, 5B; Table I). Namely, the side chain of pE2 formed hydrogen bonds to S24 and Y99, and a potential salt bridge to H9, as well as forming a network of water-mediated and van der Waals contacts with HLA-B*18:01. In addition to sharing extensive Van

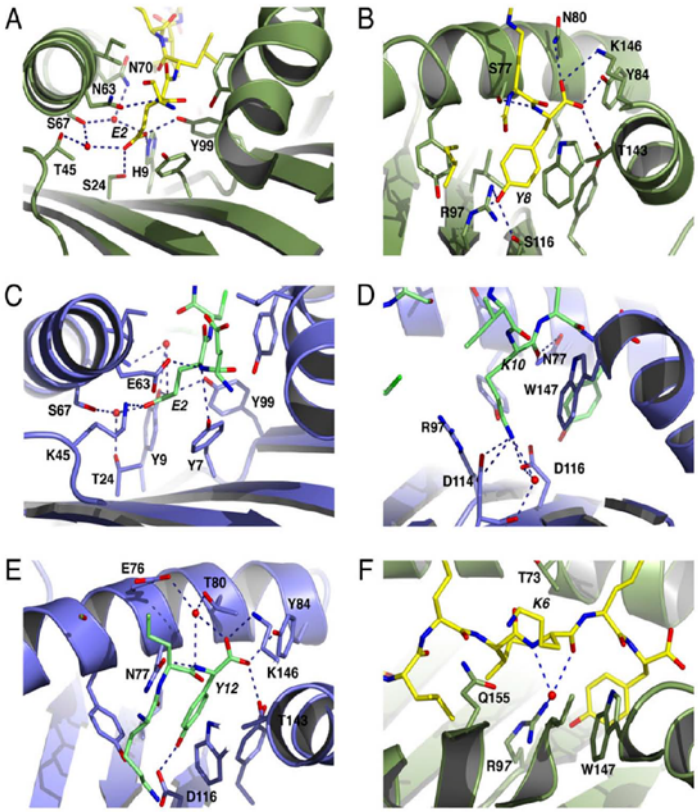


FIGURE 5. Differences in primary and secondary peptide anchoring between HLA-B*18:01–SELEIKRY and HLA-B*44:03–EECDSELEIKRY. Refined coordinates of the HLA-B*18:01–octamer (**A**, **B**, and **F**) and HLA-B*44:03–dodecamer (**C**–**E**) structures. In each case, the H chain is rendered schematically (green and blue, respectively), with selected side chains shown as sticks. The 8-aa and 12-aa peptides are presented as sticks (yellow and lime green, respectively). Selected water molecules are represented by red spheres. Hydrogen bonding and potential salt bridge interactions are shown as dashes. (A and C) Interactions between the primary peptide anchor pGlu-2 and either HLA-B*18:01 or HLA-B*44:03, respectively. All H chain residues involved in interactions with pGlu-2 (whether van der Waals, hydrogen bonding, or ionic) are shown in stick format. (B and E) Interactions between the primary pΩ anchor (Tyr^8 or Tyr^{12}) and either HLA-B*18:01 or HLA-B*44:03, respectively. (D and F) In the HLA-B*44:03–dodecamer complex, the side chain of pLys-10 is buried deep within the cleft. In the HLA-B*18:01–octamer structure, the equivalent peptide residue (pLys-6) adopts a solvent-accessible conformation.

Table I. Peptide contacts in the HLA-B*18:01–SELEIKRY and HLA-B*44:03–EECDSELEIKRY binary structures

HLA-B*18:01–SELEIKRY			HLA-B*44:03–EECDSELEIKRY		
Peptide Residue	Interaction Type	Contact [a,b]	Peptide Residue	Interaction Type	Contact
Ser^{1}	VDW	Y7, Y59, N63, Y159, W167	Glu^{1}	VDW	M5, Y7, Y59, R62, E63, Y159, L163, S167, R170, Y171
$Ser^{1\alpha}$	HB	$Y159^{O\eta}$	$Glu^{1N,\ O}$	HB	$Y7^{O\eta}$, $Y159^{O\eta}$, $S167^{O\gamma}$, $Y171^{O\eta}$
Ser^{1N}	Wm	$Y7^{O\eta}$, $Y59^{O\eta}$, $H171^{N\varepsilon 2}$			
$Ser^{1O\gamma}$	HB	$N63^{N\delta 2}$	$Glu^{1O\varepsilon 1,\ O\varepsilon 2}$	HB, SB	$R62^{N\varepsilon,\ N\eta 2}$, $R170^{N\eta 1,\ N\eta 2}$
			$Glu^{1O\varepsilon 1,\ O\varepsilon 2}$	Wm	$R62^{N\varepsilon}$, $E63^{O\varepsilon 2}$, $R170^{N\eta 1}$
Glu^{2}	VDW	Y7, H9, S24, N63, I66, S67, Y99, Y159	Glu^{2}	VDW	Y7, Y9, T24, K45, E63, I66, S67, N70, Y99, Y159, L163
Glu^{2N}	HB	$N63^{O\delta 1}$	Glu^{2N}	HB	$E63^{O\varepsilon 2}$
$Glu^{2O\varepsilon 1,\ O\varepsilon 2}$	HB, SB	$H9^{N\varepsilon 2}$, $S24^{O\gamma}$, $Y99^{O\eta}$	$Glu^{2O\varepsilon 1,\ O\varepsilon 2}$	HB, SB	$Y9^{O\eta}$, $K45^{N\zeta}$, $Y99^{O\eta}$
$Glu^{2O\varepsilon 1,\ O\varepsilon 2}$	Wm	$H9^{N\varepsilon 2}$, $T45^{O\gamma 1}$, $S67^{O\gamma}$, $N70^{N\delta 2}$	$Glu^{2O\varepsilon 1,\ O\varepsilon 2}$	Wm	$Y9^{O\eta}$, $T24^{O\gamma 1}$, $S67^{O\gamma}$, $N70^{N\delta 2}$
			Cys^{3}	VDW	I66, Y99, Y159
			Cys^{3N}	HB	$Y99^{O\eta}$
			$Cys^{3\alpha}$	Wm	$Y9^{O\eta}$, $N70^{N\delta 2}$
			Asp^{4}	VDW	I66
			Ser^{5}	VDW	*I9*
Leu^{3}	VDW	I66, N70, Y99, Q155, L156, Y159			
$Leu^{3N,\ O}$	HB	$N70^{N\delta 2}$, $Y99^{O\eta}$			
Glu^{4}	VDW	I66, T69, N70	Glu^{8}	VDW	T73, *R11*
$Glu^{4\alpha}$	HB	$Q155^{N\varepsilon 2}$	$Glu^{8\alpha}$		*R11*$^{N,\ N\varepsilon}$
$Glu^{4O\varepsilon 1,\ O\varepsilon 1}$	Wm	$T69^{O\gamma 1}$, $Q155^{N\varepsilon 2}$, *K6*$^{N\zeta}$	$Glu^{8O\varepsilon 1}$	HB	$T73^{O\gamma 1}$
Ile^{5}	VDW	N70, T73, Y74, R97	Ile^{9}	VDW	W147, V152, Q155, *S5*
Ile^{5N}	HB	$N70^{O\delta 1}$	$Ile^{9\alpha}$	HB	$W147^{N\varepsilon 1}$
Ile^{5N}	Wm	$N70^{O\delta 1}$, $T73^{O\gamma 1}$			
Lys^{6}	VDW	T73, R97, W147, V152, Q155	Lys^{10}	VDW	T73, Y74, N77, R97, D114, D116, W147
			$Lys^{10\alpha}$	HB	$N77^{O\delta 1\ N\delta 2}$
$Lys^{6N,\ O}$	Wm	$R97^{N\eta 2}$			
			$Lys^{10N\zeta}$	HB, SB	$D114^{O\delta 1\ O\delta 2}$, $D116^{O\delta 2}$
			$Lys^{10N\zeta}$	Wm	$D114^{\circ}$, $D116^{O\delta 2}$
Arg^{7}	VDW	T73, E76, S77, N80, W147	Arg^{11}	VDW	T73, E76, N77, W147, *E8*
$Arg^{7\alpha}$	HB	$W147^{N\varepsilon 1}$	$Arg^{11N,\ O}$	HB	$W147^{N\varepsilon 1}$, *E8*$^{\alpha}$
$Arg^{7N\varepsilon,\ N\eta 2}$	HB, SB	$E76^{O\varepsilon 1,\ O\varepsilon 2}$	$Arg^{11N\varepsilon}$	HB	*E8*$^{\alpha}$
$Arg^{7N\eta 1}$	Wm	$N80^{N\delta 2}$			
Tyr^{8}	VDW	Y74, S77, N80, L81, Y84, L95, R97, S116, Y123, T143, K146, W147	Tyr^{12}	VDW	Y74, N77, T80, Y84, I95, D116, Y123, T143, K146, W147
$Tyr^{8N,\ O,\ OXT}$	HB, SB	$S77^{O\gamma}$, $N80^{N\delta 2}$, $Y84^{O\eta}$, $T143^{O\gamma 1}$, $K146^{N\zeta}$	$Tyr^{12N,\ O,\ OXT}$	HB	$N77^{O\delta 1}$, $Y84^{O\eta}$, $T143^{O\gamma 1}$, $K146^{N\zeta}$
			$Tyr^{12N,\ OXT}$	Wm	$E76^{O\varepsilon 2}$, $T80^{O\gamma 1}$
$Tyr^{8O\eta}$	HB	$R97^{N\varepsilon,\ N\eta 2}$, $S116^{O\gamma}$	$Tyr^{12O\eta}$	HB	$D116^{O\delta 1}$

HB, hydrogen bond interactions, defined as contact distances of ≤ 3.5 Å between suitable atoms at appropriate angles, as determined in COOT; SB, ionic interactions, defined as contact distances of ≤ 4.0 Å between suitable residue atoms; Wm, hydrogen bond mediated by one water molecule (see HB); VDW, Van der Waals interactions, defined as non–hydrogen bond contact distances of ≤ 4.0 Å.

[a]Atomic contacts determined using the CCP4i implementation of CONTACT and a cutoff of 4.0 Å.

[b]Selected contacts between peptide residues are set in italics.

der Waals contacts with the MHC, the side chain of pY8 formed hydrogen bonds with R97 and S116 (Fig. 5B).

In complex with HLA-B*44:03, the EECDSELEIKRY peptide adopted a bulged conformation, with pR11 and the residues of the central bulged region (pD4–pI9) highly solvent exposed and therefore representing potential TCR contact points (Fig. 4C). This finding was consistent with the Ala substitution functional data (Fig. 3B), showing that pR11 and residues at the apex of the peptide bulge were critical for T cell recognition. Few interactions were observed between residues in the central region of EECDSELEIKRY and the HLA (Table I), which is consistent with the observed disorder in this region. Anchoring of EECDSELEIKRY to the cleft of HLA-B*44:03 was mediated predominantly by residues pE1, pE2, pK10, and pY12 (Table I), all of which were involved in multiple direct polar and ionic interactions with the MHC, in addition to forming a number of van der Waals and water-mediated contacts. Namely, the side chain of pE1 formed a salt bridge to R62 and R170 (not shown), whereas the side chain of pE2 interacted with Y9, K45, and Y99 (Fig. 5C) and formed water-mediated interactions with T24, S67, and N70. The side chain of pK10 salt bridged with D114 and D116 (Fig. 5D), whereas the side chain of pY12 hydrogen bonded to D116 (Fig. 5E).

The polymorphic differences between the HLA-B*18:01 and HLA-B*44:03 alleles markedly alter the shape and electrostatic properties of their Ag-binding clefts (Fig. 6A–C) and remodel the primary peptide anchor pockets (Fig. 5). Consequently, interactions between these two alleles and positions pE-2 and pY-Ω of the 8- and 12-aa peptides differed. The pE-2 of each peptide interacted directly with up to five polymorphic positions (positions 9, 24, 45, 63, and 163), four of which are nonconservative substitutions (H9Y, T45K, N63E, and T163L) (Fig. 6C). Of note, these substitutions did not alter the number of hydrogen bonding contacts between each allele and pE2, and a salt bridge interaction was also maintained in both (Table I). However, for these interactions to be maintained and steric clashes to be avoided, pE-2 was "pushed" toward the α2-helix in HLA-B*44:03. The observed shift was large (≤1.5-Å movement in the side chain) and had an impact on the backbone conformation of adjacent peptide positions (Fig. 7).

HLA polymorphism had a similar effect at pΩ, where the Tyr interacted directly with five residues that are different between

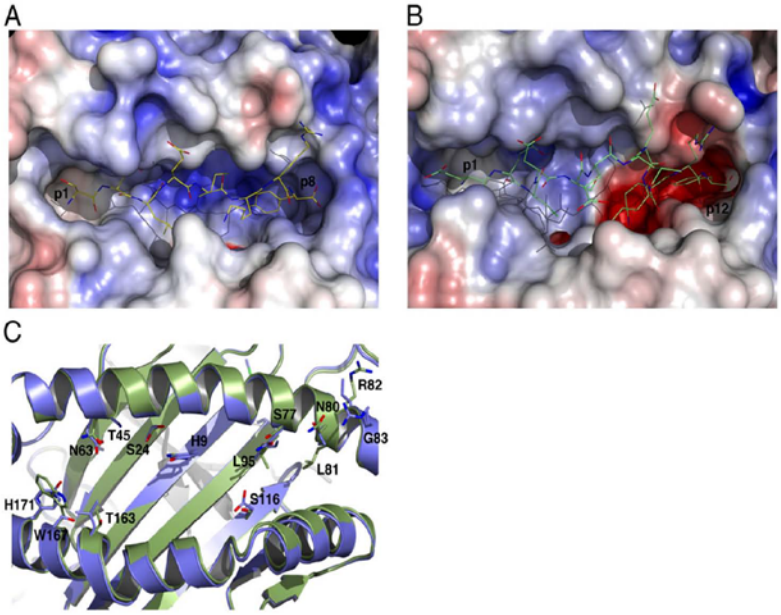


FIGURE 6. Polymorphism alters the electrostatic properties of the Ag-binding cleft in HLA-B*18:01 and HLA-B*44:03. Surface representation of the Ag-binding clefts of HLA-B*18:01 (**A**) and HLA-B*44:03 (**B**) as observed in the 8-aa and 12-aa complex structures, respectively. In each case the ± 5 kT/e electrostatic potential from the solvent-accessible surface of the MHC has been rendered on the molecular surface. The 8-aa and 12-aa peptides are shown in stick format and are colored yellow and lime, respectively. (**C**) Superposition of HLA-B*18:01 (green) and HLA-B*44:03 (blue), showing the Ag-binding cleft from above. The heavy chains and β_2-microglobulin (gray) are rendered schematically, with polymorphic residues shown as sticks. Selected polymorphic residues are labeled according to the HLA-B*18:01 sequence.

HLA-B*18:01 and HLA-B*44:03 (positions 77, 80, 81, 95, and 116), although only one of these is a nonconservative substitution (S116D) (Fig. 6C). Nevertheless, these substitutions resulted in a significant shift in the position of pY-Ω (≤2.8 Å in the side chain), which in turn altered the backbone conformation of the C-terminal region of the two peptides (Fig. 7). Compared with the HLA-B*18:01–octamer complex, pY-Ω was raised out of the cleft in B*44:03 and lost two direct hydrogen bond contacts with the H chain.

The HLA polymorphisms between HLA-B*18:01 and HLA-B*44:03 did not affect the primary anchor positions alone. In the B*44:03 complex, for example, pK10 served as a secondary anchor, with its side chain buried deep within the cleft, where it interacted with D114 and the polymorphic residue D116, as well as forming a number of van der Waals contacts with several other MHC residues (Fig. 5D). In the HLA-B*18:01–octamer complex, the equivalent peptide residue of SELEIKRY (pK6) adopted a very different conformation, with its side chain facing out of the cleft (Fig. 5F). In the HLA-B*18:01 complex, the pocket that would otherwise be occupied by the pK6 was occupied by R97 and pY-Ω. pY-Ω sat in the observed position because of the polymorphic substitutions A81L and I95L, whereas the Y9H substitution may have made the observed R97 conformation more energetically favorable.

Binding of the SELEIKRY peptide to HLA-B*44:03 is predicted to be less favorable owing to steric clashes between pL3, pI5, and R97. Because of its shorter length and anchoring requirements, the SELEIKRY peptide is likely to be forced to sit flat along the center of the cleft of HLA-B*44:03. However, the depth of the central region of the HLA-B*18:01 cleft observed in the structure was deeper than that of HLA-B*44:03, as illustrated in Fig. 8 (Fig. 8A, 8B, respectively). This difference is the result of R97 adopting a conformation in HLA-B*18:01, in which it was folded toward the C-terminal end of the cleft. This conformation of R97 was dependent on Y74, which was in turn influenced by the polymorphic positions H9 and S24 (Fig. 8C). In the HLA-B*18:01 structure, these residues allowed the side chain of Y74 to rest in a forward position, where it hydrogen bonded to R97. The H9Y substitution in HLA-B*44:03, as well as the S24T substitution, forced the aromatic ring of Y74 to swing toward the back of the cleft. In this position, Y74 would have clashed with the folded conformation of R97, which instead was observed in an upright conformation, thereby reducing the depth of the cleft. Collectively, the structural data show that the binding of shorter peptides

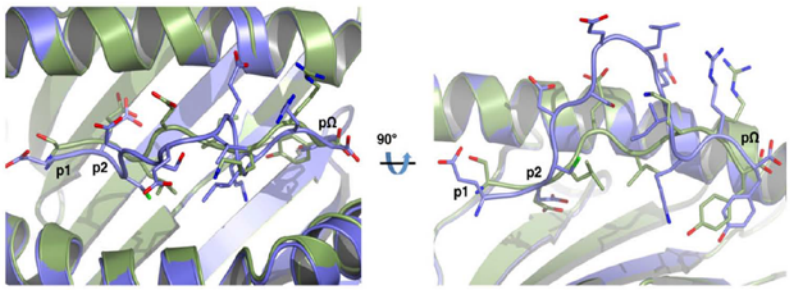


FIGURE 7. Superposition of the crystal structures of HLA-B*18:01–SELEIKRY and HLA-B*44:03–EECDSELEIKRY. Superposition of the refined coordinates of the HLA-B*18:01–octamer and HLA-B*44:03–dodecamer complexes, using main chain atoms from residues 1–181 of the H chain (root mean square deviation = 0.30 Å), viewed either from above or through the α2-helix (*left* and *right panels*, respectively). The heavy chains of HLA-B*18:01 (green) and HLA-B*44:03 (blue) are rendered schematically. Peptide atoms are shown as sticks, using the same color scheme. In this superposition, main chain atoms of the peptide anchors Glu-2 and Tyr-Ω show displacements >1 Å and higher than for any H chain residue in the range to 1 from 181.

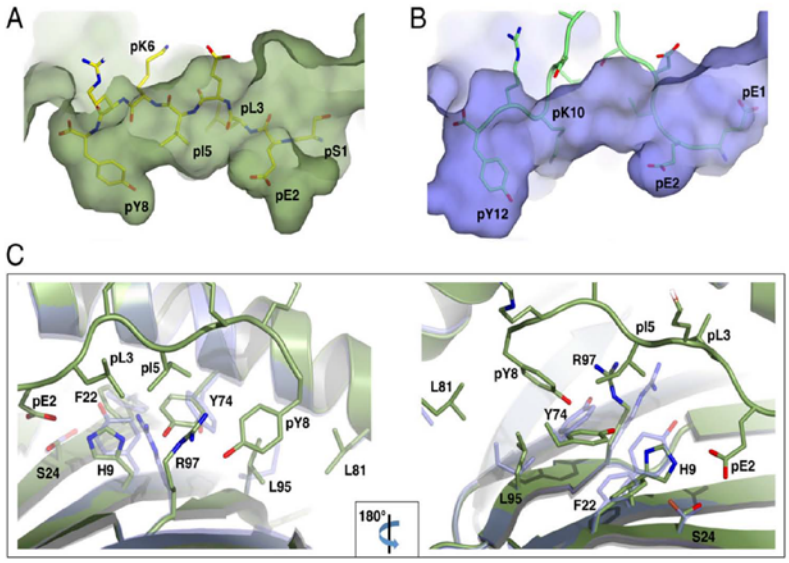


FIGURE 8. Polymorphic differences between HLA-B*18:01 and HLA-B*44:03 alter the shape of their Ag-binding clefts. (**A** and **B**) Surface representation of the Ag-binding clefts of HLA-B*18:01 (A) and HLA-B*44:03 (B) viewed through the α1-helix. The peptide pockets are rendered in semitransparent mode, with the 8-aa and 12-aa peptides shown in stick format. Polymorphic differences between these two alleles lead to remodeling of their B and F pockets, the latter of which can accommodate a Trp residue in the case of HLA-B*44:03 alone. Moreover, the central region of the HLA-B*44:03 cleft is rendered shallower than that of HLA-B*18:01 (the region corresponding to association of the 8-aa peptide positions pL3–pK6). (**C**) Detail of the Ag-binding cleft in the HLA-B*18:01–8-aa complex viewed through the α2- and α1-helices (*left* and *right panels*, respectively). The H chain and peptide are presented schematically (colored green), with key side chains shown as sticks. Superimposed is the structure of the HLA-B*44:03–12-aa complex (colored blue), drawn in semitransparent mode. Only H chain residues of HLA-B*44:03 are shown. Polymorphism at key positions within the cleft has an impact on the conformation of surrounding H chain residues, including the conserved Arg^{97}. In HLA-B*18:01, the side chain of Arg^{97} is able to adopt a folded conformation that points toward the C terminus of the cleft, whereas in the HLA-B*44:03 complex, Arg^{97} is forced into an upright position. The folded conformation of Arg^{97} deepens the central region of the HLA-B*18:01 cleft sufficiently to accommodate the side chains of the pL3 and pI5 residues of the 8-aa epitope.

by HLA-B*44:03 is more likely to be disfavored owing to steric clashes, as such peptides need to sit flatter within the cleft.

*The binding specificities of HLA-B*18:01 and HLA-B*44:03 are intrinsically biased toward peptides of different length*

To investigate the possibility that epitope selection in the T cell response to BZLF1 is influenced by inherent variability in the peptide length specificity of different HLA molecules, the global self-peptide repertoires of HLA-B*18:01 and HLA-B*44:03 were compared. We employed MHC peptide elution following immunoprecipitation of class I complexes from C1R cells engineered to express either of the two alleles (C1R-B*18:01 and C1R-B*44:03) (14, 15, 30). Peptides were detected by LC–tandem mass spectrometry, and endogenous human sequences analyzed by the ProteinPilot algorithm. We first analyzed the binding motif of the two alleles. Although both motifs have been described previously (5–7), we confirm in this study the nearly exclusive preference for a p2 Glu, and dominant Tyr or Phe at the C terminus, with an additional codominance of Trp in HLA-B*44:03 (but not B*18:01) (Fig. 9A, 9B). The preference for Trp in HLA-B*44:03 alone can be accounted for fully by the observed structural influence of polymorphic substitutions on the F pockets of these two alleles. More importantly, analysis of > 1200 individual sequences across the two alleles revealed a striking difference in peptide length distribution (Fig. 9C, 9D). Of the total peptides 8–12 aa in length, 8-aa (17%) and 9-aa (76%) peptides dominated the B*18:01 repertoire, with 10-aa (4.5%), 11-aa (1.7%), and 12-aa (0.6%) peptides making a small contribution. In contrast, peptides eluted from HLA-B*44:03$^+$ cells were dominated by the longer lengths (9 aa, 48%; 10 aa, 33%; 11 aa, 14%; and 12 aa, 4.5%), with 8-aa peptides present at a frequency of just 0.8%. These data thus provide a basis for understanding functional and biochemical findings with the BZLF1 epitopes SELEIKRY and EECDSELEIKRY, in that these alleles are intrinsically biased toward differing peptide lengths.

Discussion

Polymorphism at the MHC locus enhances immune defense across the population by ensuring wide variation in the T cell response to infecting pathogens, and this is primarily achieved through presentation of a diverse array of antigenic peptide sequences. The present study points out that class I MHC polymorphism influences not only amino acid preferences at different positions within peptide ligands but also the length of presented peptides. Furthermore, this report, to our knowledge, is the first to show that variability in the peptide length preferences of different MHC molecules can influence epitope selection in an Ag-specific T cell response, thus highlighting a novel mechanism through which MHC class I polymorphism can further diversify immune responses.

The CD8$^+$ T cell response to the BZLF1 Ag of EBV was specific for either the SELEIKRY octamer or the EECDSELEIKRY dodecamer peptide, depending on whether it was restricted by HLA-B*18:01 or HLA-B*44:03, respectively. The likely mechanism for this intriguing observation was revealed by pMHC binding assays showing that the octamer peptide bound more stably to HLA-B*18:01 than did the dodecamer peptide, whereas HLA-B*44:03 bound only the longer peptide. Subsequent structural studies of these viral pMHC complexes indicated that the binding clefts of these two members of the HLA-B*44 supertype may favor the presentation of peptides of different lengths, with the more "shallow" groove of HLA-B*44:03 leading to steric clashes with

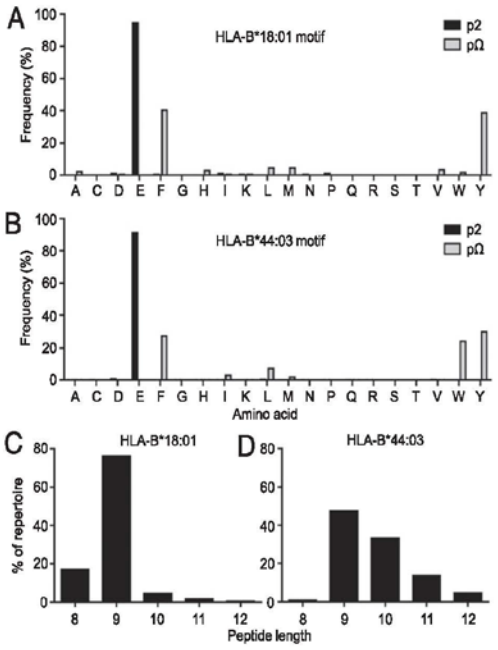


FIGURE 9. Binding specificity of HLA-B*18:01 and HLA-B*44:03 assessed by analysis of endogenous human peptide ligands. The global self-peptide repertoires of HLA-B*18:01 and HLA-B*44:03 were compared by eluting peptides from the C1R-B*18:01 and C1R-B*44:03 cell lines, followed by LC–tandem mass spectrometry. (**A**) The frequency of amino acids occurring at P2 and PΩ in peptides eluted from HLA-B*18:01 and (**B**) HLA-B*44:03 is shown. (**C**) The distribution of peptide lengths from the datasets for HLA-B*18:01 and (**D**) HLA-B*44:03 is also shown.

octamer peptides. Indeed, the intrinsic bias of these two HLA alleles for peptides of different length was confirmed by mass spectrometric identification of natural ligands for HLA-B*18:01 and HLA-B*44:03, showing that 18% of peptides presented by HLA-B*18:01 were octamers versus 0.8% for HLA-B*44:03, whereas peptides of $>$ 10 aa were more frequently presented by HLA-B*44:03 (18.5%) than by HLA-B*18:01 (2.3%).

It should be noted that peptide secondary anchor residues could also play a role in influencing the relative binding stability of the two EBV length-variant peptides to HLA-B*18:01 or HLA-B*44:03, although no direct evidence was found for this in our global analysis of the self-peptide repertoires of these two HLA-B*44 supertype allomorphs. For example, there was no preference for Ser at P1 within self-peptides eluted from HLA-B*18:01, nor for Glu at P1 within self-peptides eluted from HLA-B*44:03 (data not shown). Nonetheless, pE1 of the 12-aa peptide and pS1 of the 8-aa peptide do interact differently with HLA-B*44:03/B*18:01, with a more extensive network of van der Waals and hydrogen bonds observed with pE1 of the 12-aa peptide than with pS1 of the 8-aa peptide. Major differences were also observed with the Lys residue of each peptide. The pK6 of SELEIKRY points out of the HLA cleft, whereas the pK10 of EECDSELEIKRY plays an important secondary anchor residue role.

In all structures of HLA-B*44:03 determined to date, in complex with different peptide ligands, Arg^{97} has been shown to adopt the same "upright" conformation seen in the HLA-B*44:03–12-aa complex presented in this article (8, 29, 35, 48). This residue plays a major role in reducing the size of the peptide-binding cleft and is shared by HLA-B*44:02, which differs from HLA-B*44:03 by only a single $Asp^{156}Leu$ substitution, and which also displays the same aversion for presenting 8-aa peptides (6). In fact, comparison of all published structures of HLA-B*44:02 and HLA-B*44:03 in complex with a variety of peptides reveals that their Ag-binding grooves share a highly conserved conformation. This observation would suggest that the clefts of these two alleles lack plasticity, and this may be another factor influencing their reduced tolerance for binding shorter peptides.

Other HLA class I allomorphs have shown a preference for peptides of a certain length. For example, different members of the HLA-B*41 allelic group show marked differences in peptide length preferences (24). Although the HLA-B*41:03 and HLA-B*41:04 alleles were able to present peptides of 9 aa in length or longer (up to 16 aa in the case of HLA-B*41:03 and 15 aa in the case of HLA-B*41:04), octamers were eluted only from HLA-B*41:04 and accounted for $>$ 20% of peptides. Of interest, HLA-B*41:04 is characterized by an $Arg^{97}Ser$ substitution in comparison with HLA-B*41:03, which increases the size of the HLA-B*41:04 Ag-binding cleft by 140 $Å^3$. In addition, HLA-A*02:01 has been shown to disfavor the presentation of octamer peptides (14). Like HLA-B*41:03 and HLA-B*44:02/03, HLA-A*02:01 has an Arg residue at position 97. When compared with HLA-A*02:01, the Ag-binding cleft of HLA-B*08:01, an allele with an apparent propensity for binding 8-aa peptides (49, 50), is considerably deeper owing to a number of polymorphic substitutions, including $Phe^{9}Asp$, $His^{74}Asp$, $Arg^{97}Ser$, and $His^{114}Asn$. Clearly, the prediction of which HLA alleles favor presentation of octamer peptides cannot simply be reduced to the identity of a particular polymorphic residue, such as position 97. The alleles HLA-B*57:01 and HLA-B*57:03 have a relatively small residue (Val) at position 97, yet both have been shown to disfavor the presentation of octamer peptides (15). Instead, these two alleles are characterized by different bulky substitutions within the Ag-binding cleft when compared with HLA-B*18:01. These include $His^{9}Tyr$ and $Ser^{67}Met$, which are common to both HLA-B*57:01 and HLA-B*57:03, and $Ser^{116}Tyr$, which is specific to HLA-B*57:03. Furthermore, volumetric measure of the Ag-binding cleft may not always reflect the peptide length preferences of HLA alleles. In the case of HLA-B*18:01 and HLA-B*44:03, the overall volume of their respective clefts differs by only 50$Å^3$, although polymorphic differences between the two do render the central region of the HLA-B*44:03 cleft shallower than that of HLA-B*18:01.

The importance of relative depth and contour features of the Ag-binding cleft in peptide length preferences has also been highlighted for different murine class I alleles. In particular, the ability of H2-K^b, but not H2-D^b, to bind peptides shorter than 9 residues has been attributed to a hydrophobic ridge present within the Ag-binding cleft of the latter, but not the former, allele (51, 52). By contrast, a structurally distinct hydrophobic ridge present in the cleft of H2-L^d (53, 54) does not impose the same minimum length requirements on bound peptides, presumably owing to its ability to be more readily remodeled in response to binding of octamer peptides. Importantly, the distinct structural features between these alleles have also been shown to have an impact on allore-activity (55).

Structural features of the MHC class I peptide-binding cleft also appear to limit the capacity to present peptide ligands of $>$ 10 amino acids in length. Unusually long peptide ligands have not been described for allotypes such as H2-K^b, HLA-B*08:01, and HLA-B*14, presumably because these molecules generally require a primary anchor residue in the central part of the peptide, as well as primary or secondary anchors at the extremities, which is unlikely to be compatible with the looping conformation of long peptides. Thus, through various mechanisms, MHC class I polymorphism promotes the presentation of epitopes of variable size,

thereby contributing to the diversity of Ag-specific immune responses across the population and ensuring effective immune surveillance by $CD8^+$ T cells. Evolutionary pressure on MHC class I molecules to maintain polymorphisms that enhance the presentation of a diverse array of peptide sizes could be provided by pathogens, such as human CMV, that interfere with trimming of pathogen-derived peptides, leading to reduced susceptibility of infected cells to T cell recognition (56, 57).

Acknowledgments

We thank the staff at the Advanced Photon Source (Chicago, IL) and the Australian Synchrotron (Melbourne, Australia) for assistance with data collection.

Disclosures

The authors have no financial conflicts of interest.

References

1. Clements, C. S., M. A. Dunstone, W. A. Macdonald, J. McCluskey, and J. Rossjohn. 2006. Specificity on a knife-edge: the alphabeta T cell receptor. *Curr. Opin. Struct. Biol.* 16: 787–795.
2. Ruppert, J., J. Sidney, E. Celis, R. T. Kubo, H. M. Grey, and A. Sette. 1993. Prominent role of secondary anchor residues in peptide binding to HLA-A2.1 molecules. *Cell* 74: 929–937.
3. Burrows, J. M., K. K. Wynn, F. E. Tynan, J. Archbold, J. J. Miles, M. J. Bell, R. M. Brennan, S. Walker, J. McCluskey, J. Rossjohn, et al. 2007. The impact of HLA-B micropolymorphism outside primary peptide anchor pockets on the CTL response to CMV. *Eur. J. Immunol.* 37: 946–953.
4. Sette, A., and J. Sidney. 1999. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. *Immunogenetics* 50: 201–212.
5. Sidney, J., S. Southwood, V. Pasquetto, and A. Sette. 2003. Simultaneous prediction of binding capacity for multiple molecules of the HLA B44 supertype. *J. Immunol.* 171: 5964–5974.
6. Hillen, N., G. Mester, C. Lemmel, A. O. Weinzierl, M. Müller, D. Wernet, J. Hennenlotter, A. Stenzl, H. G. Rammensee, and S. Stevanović. 2008. Essential differences in ligand presentation and T cell epitope recognition among HLA molecules of the HLA-B44 supertype. *Eur. J. Immunol.* 38: 2993–3003.
7. Hickman, H. D., A. D. Luis, R. Buchli, S. R. Few, M. Sathiamurthy, R. S. VanGundy, C. F. Giberson, and W. H. Hildebrand. 2004. Toward a definition of self: proteomic evaluation of the class I peptide repertoire. *J. Immunol.* 172: 2944–2952.
8. Macdonald, W. A., A. W. Purcell, N. A. Mifsud, L. K. Ely, D. S. Williams, L. Chang, J. J. Gorman, C. S. Clements, L. Kjer-Nielsen, D. M. Koelle, et al. 2003. A naturally selected dimorphism within the HLA-B44 supertype alters class I structure, peptide repertoire, and T cell recognition. *J. Exp. Med.* 198: 679–691.
9. Bouvier, M., and D. C. Wiley. 1994. Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. *Science* 265: 398–402.
10. Probst-Kepper, M., V. Stroobant, R. Kridel, B. Gaugler, C. Landry, F. Brasseur, J. P. Cosyns, B. Weynand, T. Boon, and B. J. Van Den Eynde. 2001. An alternative open reading frame of the human macrophage colony-stimulating factor gene is independently translated and codes for an antigenic peptide of 14 amino acids recognized by tumor-infiltrating CD8 T lymphocytes. *J. Exp. Med.* 193: 1189–1198.
11. Samino, Y., D. López, S. Guil, L. Saveanu, P. M. van Endert, and M. Del Val. 2006. A long N-terminal-extended nested set of abundant and antigenic major histocompatibility complex class I natural ligands from HIV envelope protein. *J. Biol. Chem.* 281: 6358–6365.
12. Burrows, J. M., M. J. Bell, R. Brennan, J. J. Miles, R. Khanna, and S. R. Burrows. 2008. Preferential binding of unusually long peptides to MHC class I and its influence on the selection of target peptides for T cell recognition. *Mol. Immunol.* 45: 1818–1824.
13. Burrows, S. R., J. Rossjohn, and J. McCluskey. 2006. Have we cut ourselves too short in mapping CTL epitopes? *Trends Immunol.* 27: 11–16.
14. Scull, K. E., N. L. Dudek, A. J. Corbett, S. H. Ramarathinam, D. G. Gorasia, N. A. Williamson, and A. W. Purcell. 2012. Secreted HLA recapitulates the immunopeptidome and allows in-depth coverage of HLA A*02:01 ligands. *Mol. Immunol.* 51: 136–142.
15. Illing, P. T., J. P. Vivian, N. L. Dudek, L. Kostenko, Z. Chen, M. Bharadwaj, J. J. Miles, L. Kjer-Nielsen, S. Gras, N. A. Williamson, et al. 2012. Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. *Nature* 486: 554–558.
16. Speir, J. A., J. Stevens, E. Joly, G. W. Butcher, and I. A. Wilson. 2001. Two different, highly exposed, bulged structures for an unusually long peptide bound to rat MHC class I RT1-Aa. *Immunity* 14: 81–92.
17. Guo, H. C., T. S. Jardetzky, T. P. Garrett, W. S. Lane, J. L. Strominger, and D. C. Wiley. 1992. Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. *Nature* 360: 364–366.
18. Probst-Kepper, M., H. J. Hecht, H. Herrmann, V. Janke, F. Ocklenburg, J. Klempnauer, B. J. van den Eynde, and S. Weiss. 2004. Conformational restraints and flexibility of 14-meric peptides in complex with HLA-B*3501. *J. Immunol.* 173: 5610–5616.
19. Tynan, F. E., N. A. Borg, J. J. Miles, T. Beddoe, D. El-Hassen, S. L. Silins, W. J. van Zuylen, A. W. Purcell, L. Kjer-Nielsen, J. McCluskey, et al. 2005. High resolution structures of highly bulged viral epitopes bound to major histocompatibility complex class I. Implications for T-cell receptor engagement and T-cell immunodominance. *J. Biol. Chem.* 280: 23900–23909.
20. Miles, J. J., D. Elhassen, N. A. Borg, S. L. Silins, F. E. Tynan, J. M. Burrows, A. W. Purcell, L. Kjer-Nielsen, J. Rossjohn, S. R. Burrows, and J. McCluskey. 2005. CTL recognition of a bulged viral peptide involves biased TCR selection. *J. Immunol.* 175: 3826–3834.
21. Miles, J. J., N. A. Borg, R. M. Brennan, F. E. Tynan, L. Kjer-Nielsen, S. L. Silins, M. J. Bell, J. M. Burrows, J. McCluskey, J. Rossjohn, and S. R. Burrows. 2006. TCR alpha genes direct MHC restriction in the potent human T cell response to a class I-bound viral epitope. *J. Immunol.* 177: 6804–6814.
22. Stewart-Jones, G. B., G. Gillespie, I. M. Overton, R. Kaul, P. Roche, A. J. McMichael, S. Rowland-Jones, and E. Y. Jones. 2005. Structures of three HIV-1 HLA B*5703-peptide complexes and identification of related HLAs potentially associated with long-term nonprogression. *J. Immunol.* 175: 2459–2468.
23. Wynn, K. K., Z. Fulton, L. Cooper, S. L. Silins, S. Gras, J. K. Archbold, F. E. Tynan, J. J. Miles, J. McCluskey, S. R. Burrows, et al. 2008. Impact of clonal competition for peptide-MHC complexes on the CD8+ T-cell repertoire selection in a persistent viral infection. *Blood* 111: 4283–4292.
24. Bade-Döding, C., A. Theodossis, S. Gras, L. Kjer-Nielsen, B. Eiz-Vesper, A. Seltsam, T. Huyton, J. Rossjohn, J. McCluskey, and R. Blasczyk. 2011. The impact of human leukocyte antigen (HLA) micropolymorphism on ligand specificity within the HLA-B*41 allotypic family. *Haematologica* 96: 110–118.
25. Ebert, L. M., Y. C. Liu, C. S. Clements, N. C. Robson, H. M. Jackson, J. L. Markby, N. Dimopoulos, B. S. Tan, I. F. Luescher, I. D. Davis, et al. 2009. A long, naturally presented immunodominant epitope from NY-ESO-1 tumor antigen: implications for cancer vaccine design. *Cancer Res.* 69: 1046–1054.
26. Tynan, F. E., S. R. Burrows, A. M. Buckle, C. S. Clements, N. A. Borg, J. J. Miles, T. Beddoe, J. C. Whisstock, M. C. Wilce, S. L. Silins, et al. 2005. T cell receptor recognition of a 'super-bulged' major histocompatibility complex class I-bound peptide. *Nat. Immunol.* 6: 1114–1122.
27. Gras, S., S. R. Burrows, S. J. Turner, A. K. Sewell, J. McCluskey, and J. Rossjohn. 2012. A structural voyage toward an understanding of the MHC-I-restricted immune response: lessons learned and much to be learned. *Immunol. Rev.* 250: 61–81.
28. Ekeruche-Makinde, J., J. J. Miles, H. A. van den Berg, A. Skowera, D. K. Cole, G. Dolton, A. J. Schauenburg, M. P. Tan, J. M. Pentier, S. Llewellyn-Lacey, et al. 2013. Peptide length determines the outcome of TCR/peptide-MHCI engagement. *Blood* 121: 1112–1123.
29. Archbold, J. K., W. A. Macdonald, S. Gras, L. K. Ely, J. J. Miles, M. J. Bell, R. M. Brennan, T. Beddoe, M. C. Wilce, C. S. Clements, et al. 2009. Natural micropolymorphism in human leukocyte antigens provides a basis for genetic control of antigen recognition. *J. Exp. Med.* 206: 209–219.
30. Dudek, N. L., C. T. Tan, D. G. Gorasia, N. P. Croft, P. T. Illing, and A. W. Purcell. 2012. Constitutive and inflammatory immunopeptidome of pancreatic β cells. *Diabetes* 61: 3018–3025.
31. Tan, C. T., N. P. Croft, N. L. Dudek, N. A. Williamson, and A. W. Purcell. 2011. Direct quantitation of MHC-bound peptide epitopes by selected reaction monitoring. *Proteomics* 11: 2336–2340.
32. Leslie, A. G. W., and H. R. Powell. 2007. Processing Diffraction data with mosflm. In *Evolving Methods for Macromolecular Crystallography.* Read and Sussman, eds. Springer, Dordrecht, the Netherlands, p. 41–51.
33. Evans, P. R. 2011. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.* 67: 282–292.
34. McCoy, A. J., R. W. Grosse-Kunstleve, L. C. Storoni, and R. J. Read. 2005. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D Biol. Crystallogr.* 61: 458–464.
35. Zernich, D., A. W. Purcell, W. A. Macdonald, L. Kjer-Nielsen, L. K. Ely, N. Laham, T. Crockford, N. A. Mifsud, M. Bharadwaj, L. Chang, et al. 2004. Natural HLA class I polymorphism controls the pathway of antigen presentation and susceptibility to viral evasion. *J. Exp. Med.* 200: 13–24.
36. Adams, P. D., P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, et al. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66: 213–221.
37. Emsley, P., B. Lohkamp, W. G. Scott, and K. Cowtan. 2010. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66: 486–501.
38. Chen, V. B., W. B. Arendall, III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, and D. C. Richardson. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66: 12–21.
39. Dundas, J., Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, and J. Liang. 2006. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* 34 (Web Server issue): W116–118.
40. Dolinsky, T. J., J. E. Nielsen, J. A. McCammon, and N. A. Baker. 2004. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 32(Web Server issue): W665–667.

41. Baker, N. A., D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* 98: 10037–10041.
42. Tan, L. C., N. Gudgeon, N. E. Annels, P. Hansasuta, C. A. O'Callaghan, S. Rowland-Jones, A. J. McMichael, A. B. Rickinson, and M. F. Callan. 1999. A re-evaluation of the frequency of CD8+ T cells specific for EBV in healthy virus carriers. *J. Immunol.* 162: 1827–1835.
43. Green, K. J., J. J. Miles, J. Tellam, W. J. van Zuylen, G. Connolly, and S. R. Burrows. 2004. Potent T cell response to a class I-binding 13-mer viral epitope and the influence of HLA micropolymorphism in controlling epitope length. *Eur. J. Immunol.* 34: 2510–2519.
44. Saulquin, X., C. Ibisch, M. A. Peyrat, E. Scotet, M. Hourmant, H. Vie, M. Bonneville, and E. Houssaint. 2000. A global appraisal of immunodominant CD8 T cell responses to Epstein-Barr virus and cytomegalovirus by bulk screening. *Eur. J. Immunol.* 30: 2531–2539.
45. Tynan, F. E., H. H. Reid, L. Kjer-Nielsen, J. J. Miles, M. C. Wilce, L. Kostenko, N. A. Borg, N. A. Williamson, T. Beddoe, A. W. Purcell, et al. 2007. A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I molecule. *Nat. Immunol.* 8: 268–276.
46. Bogedain, C., H. Wolf, S. Modrow, G. Stuber, and W. Jilg. 1995. Specific cytotoxic T lymphocytes recognize the immediate-early transactivator Zta of Epstein-Barr virus. *J. Virol.* 69: 4872–4879.
47. Theodossis, A., C. Guillonneau, A. Welland, L. K. Ely, C. S. Clements, N. A. Williamson, A. I. Webb, J. A. Wilce, R. J. Mulder, M. A. Dunstone, et al. 2010. Constraints within major histocompatibility complex class I restricted peptides: presentation and consequences for T-cell recognition. *Proc. Natl. Acad. Sci. USA* 107: 5534–5539.
48. Macdonald, W. A., Z. Chen, S. Gras, J. K. Archbold, F. E. Tynan, C. S. Clements, M. Bharadwaj, L. Kjer-Nielsen, P. M. Saunders, M. C. Wilce, et al. 2009. T cell allorecognition via molecular mimicry. *Immunity* 31: 897–908.
49. Bell, M. J., J. M. Burrows, R. Brennan, J. J. Miles, J. Tellam, J. McCluskey, J. Rossjohn, R. Khanna, and S. R. Burrows. 2009. The peptide length specificity of some HLA class I alleles is very broad and includes peptides of up to 25 amino acids in length. *Mol. Immunol.* 46: 1911–1917.
50. DiBrino, M., K. C. Parker, J. Shiloach, R. V. Turner, T. Tsuchida, M. Garfield, W. E. Biddison, and J. E. Coligan. 1994. Endogenous peptides with distinct amino acid anchor residue motifs bind to HLA-A1 and HLA-B8. *J. Immunol.* 152: 620–631.
51. Fremont, D. H., M. Matsumura, E. A. Stura, P. A. Peterson, and I. A. Wilson. 1992. Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb. *Science* 257: 919–927.
52. Young, A. C., W. Zhang, J. C. Sacchettini, and S. G. Nathenson. 1994. The three-dimensional structure of H-2Db at 2.4 A resolution: implications for antigen-determinant selection. *Cell* 76: 39–50.
53. Balendiran, G. K., J. C. Solheim, A. C. Young, T. H. Hansen, S. G. Nathenson, and J. C. Sacchettini. 1997. The three-dimensional structure of an H-2Ld-peptide complex explains the unique interaction of Ld with beta-2 microglobulin and peptide. *Proc. Natl. Acad. Sci. USA* 94: 6880–6885.
54. Reddehase, M. J., J. B. Rothbard, and U. H. Koszinowski. 1989. A pentapeptide as minimal antigenic determinant for MHC class I-restricted T lymphocytes. *Nature* 337: 651–653.
55. Speir, J. A., K. C. Garcia, A. Brunmark, M. Degano, P. A. Peterson, L. Teyton, and I. A. Wilson. 1998. Structural basis of 2C TCR allorecognition of H-2Ld peptide complexes. *Immunity* 8: 553–562.
56. Kim, S., S. Lee, J. Shin, Y. Kim, I. Evnouchidou, D. Kim, Y. K. Kim, Y. E. Kim, J. H. Ahn, S. R. Riddell, et al. 2011. Human cytomegalovirus microRNA miR-US4-1 inhibits CD8(+) T cell responses by targeting the aminopeptidase ERAP1. *Nat. Immunol.* 12: 984–991.
57. Blanchard, N., T. Kanaseki, H. Escobar, F. Delebecque, N. A. Nagarajan, E. Reyes-Vargas, D. K. Crockett, D. H. Raulet, J. C. Delgado, and N. Shastri. 2010. Endoplasmic reticulum aminopeptidase associated with antigen processing defines the composition and structure of MHC class I peptide repertoire in normal and virus-infected cells. *J. Immunol.* 184: 3033–3042.

Chapter 6:

Aim Four:

Characterise a cross-reactive T cell response to an EBV epitope and self–peptide presented by HLA-B*1801.

Hypothesis Four:

The T cell receptor repertoire that cross-reacts with an EBV epitope and a self-peptide are identical between individuals.

Published research article:

Rist MJ, Hibbert KM, Croft NP, Smith C, Neller MA, Burrows JM, Miles JJ, Purcell AW, Rossjohn J, Gras S, Burrows SR. T cell cross-reactivity between a highly immunogenic Epstein-Barr virus epitope and a self-peptide naturally presented by HLA-B*18:01^{+} cells. *The Journal of Immunology.* 2015, May; 194: 4668-4675.

T cell cross-reactivity between a highly immunogenic Epstein-Barr virus epitope and a self-peptide naturally presented by HLA-B*18:01^{+} cells

Melissa J. Rist,[1*,†], Kelly M. Hibbert,[1‡,§], Nathan P. Croft,[‡] Corey Smith,* Michelle A. Neller,* Jacqueline M. Burrows,* John J. Miles,*[,†,‖] Anthony W. Purcell,[‡] Jamie Rossjohn,[‡,§,‖,#] Stephanie Gras[‡,§,#] and Scott R. Burrows*[,†,#]

*QIMR Berghofer Medical Research Institute, Brisbane, Australia; [†]School of Medicine, University of Queensland, Brisbane, Australia; [‡]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Australia; [§]ARC Centre of Excellence in Advanced Molecular Imaging, Monash University, Clayton, Australia; [‖]Institute of Infection and Immunity, Cardiff University School of Medicine, Heath Park, Cardiff, United Kingdom

[1] M.J.R. and K.M.H. contributed equally to this work.
[#] S.R.B., S.G. and J.R. contributed equally to this work.

Address correspondence and reprint requests to Prof. Scott R. Burrows, QIMR Berghofer Medical Research Institute, 300 Herston Road, Herston, Brisbane, 4029, Australia, Ph: +61738453793, or Dr. Stephanie Gras or Prof. Jamie Rossjohn, Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia, Ph: +61399050254; E-mail addresses: Scott.Burrows@qimrberghofer.edu.au (S.R.B.) or Stephanie.Gras@monash.edu (S.G.) or Jamie.Rossjohn@monash.edu (J.R.)

Running Title: Antiviral T cell cross-reactivity with a self-peptide

This work was supported by the Australian Research Council (Future Fellowship FT120100416 to S.G.) and the National Health and Medical Research Council (NHMRC) of Australia (Australia Fellowship to J.R., Research Fellowships to S.R.B. & A.W.P; Career Development Fellowship to J.J.M.; and Dora Lush Scholarship to M.J.R.), and by grants from the NHMRC (1023141).

Abstract

T cell cross-reactivity underpins the molecular mimicry hypothesis in which microbial peptides sharing structural features with host peptides stimulate T cells that cross-react with self-peptides, thereby initiating and/or perpetuating autoimmune disease. EBV represents a potentially important factor in the pathogenesis of several autoimmune disorders, with molecular mimicry a likely mechanism. Here we describe a human self-peptide (DELEIKAY) that is a homologue of a highly immunogenic EBV T cell epitope (SELEIKRY) presented by HLA-B*18:01. This self-peptide was shown to bind stably to HLA-B*18:01, and peptide elution/mass spectrometric studies showed it is naturally presented by this HLA molecule on the surface of human cells. A significant proportion of $CD8^+$ T cells raised from some healthy individuals against this EBV epitope cross-reacted with the self-peptide. A diverse array of T cell receptors was expressed by the cross-reactive T cells, with variable functional avidity for the self-peptide, including some T cells that appeared to avoid autoreactivity by a narrow margin, with only 10-fold more of the self-peptide required for equivalent activation as compared to the EBV peptide. Structural studies revealed that the self-peptide-HLA-B*18:01 complex is a structural mimic of the EBV peptide-HLA-B*18:01 complex, and that the strong antiviral T cell response is primarily dependent on the alanine/arginine mismatch at position 7. To our knowledge, this is the first report confirming the natural presentation of a self-peptide cross-recognized in the context of self-HLA by EBV-reactive $CD8^+$ T cells. These results illustrate how aberrant immune responses and immunopathological diseases could be generated by EBV infection.

Introduction

The vast majority of the body's circulating $CD8^+$ T cells express αβ T cell receptors (TCRs) on the cell surface and play an integral role in the adaptive immune response, recognizing peptides bound to major histocompatibility complex (MHC) class I molecules [1]. T cells undergo rigorous selection in the thymus, eliminating those that recognize self-peptides associated with MHC molecules, whilst positively selecting those with the potential to recognize antigenic peptide epitopes [2]. Importantly, self or autoreactive T cells with a very low affinity/avidity can escape negative selection in the thymus and, under certain circumstances, can facilitate a variety of autoimmune conditions in humans [3] [4].

Generated somatically through site specific DNA recombination reactions, TCR diversity is a result of the random assembly of variable (V), diversity (D) and junctional (J) gene segments. The TCR α chain is comprised of V and J gene segments whilst V, D and J genes encode the TCR β chain [5]. Each chain has three CDRs that make the Ag-binding site of the TCR. The CDR3 loop is a critical component of the TCR, as it generally mediates contact with the antigenic peptide-MHC complex [6]. The mathematical estimation of potential TCR diversity is around 10^{15} different TCRs, although after considering thymic positive and negative selection events, the size of the naïve TCRαβ repertoire is estimated to be around 2×10^7 TCRs for each human [5].

As the binding sites of TCRs arise from random genetic mechanisms, these sites have the capacity to react not only with pathogens and environmental antigens but also with self antigens. Indeed, autoimmune disease and allergies can occur when activation of the adaptive immune response is directed at antigens that are self or environmental. Forming the basis of the molecular mimicry hypothesis is T cell cross-reactivity, whereby it is postulated that T cells stimulated by an epitope derived from an infectious agent can cross-react with a self antigen that may share sequence or structural homology with that epitope. This phenomenon may initiate and/or perpetuate autoimmune disease [7]. Consistent with this theory a number of animal studies have demonstrated that antigen mimicry can induce a T cell response in addition to autoimmune disease in the form of murine autoimmune oophoritis [8], and limited sequence homology can evoke a multiple sclerosis (MS)-like disease [9].

Epstein-Barr virus (EBV), a lymphotropic gamma-1 herpesvirus, is widespread in all human populations, infecting around 95% of the adult population worldwide [10] [11]. EBV infection is associated with several autoimmune conditions including MS, for which a molecular mimicry mechanism has been proposed [12]. Recently, we investigated the $CD8^+$ T cell response to the immediate early, BZLF1 antigen of EBV and described an octamer peptide ^{173}SELEIKRY^{180} (termed SEL) that is recognized by HLA-B*18:01^+ individuals [13]. In the present study, we have

identified a human protein sequence (DELEIKAY; termed DEL) with sequence homology to the octamer EBV epitope. A peptide corresponding to this sequence from the cleavage and polyadenylation specific factor 3-like protein (CPSF3L) was shown by peptide elution/mass spectrometry be presented by the HLA-B*18:01 molecule on the surface of human cells. A significant proportion of T cells raised against the EBV epitope cross-reacted with this HLA-B*18:01-binding synthetic self-peptide, and structural studies revealed the molecular basis of this cross-reactivity. These results underscore the exquisite specificity of the immune system but highlight the potential danger of self-reactivity by T cells expanded in response to common viral infection.

Materials and Methods

Purification of MHC-bound peptides

C1R-HLA-B*18:01 cell pellet (5 x 10^8 cells) were ground in a Retsch Mixer Mill MM 400 under cryogenic conditions, resuspended in 0.5 % IGEPAL, 50 mM Tris-HCl pH 8.0, 150 mM NaCl and protease inhibitors (Complete Protease Inhibitor Cocktail Tablet; Roche Molecular Biochemicals) at a density of 5 x 10^7 cells/mL and incubated with rotation for 1 h at 4 °C [14]. Lysates were cleared by ultracentrifugation (180,000 x g) and HLA-peptide complexes sequentially immunoaffinity purified using solid-phase bound anti-Bw6 (HB152-SRF) antibodies. Bound complexes were washed and peptides eluted by acidification with 10 % acetic acid, as described [15-17]. The mixture of peptides and HLA proteins was fractionated on a 4.6 mm internal diameter x 50 mm long monolithic reversed-phase C18 HPLC column (Chromolith Speed Rod, Merck) using an ÄKTAmicro™ HPLC system (GE Healthcare). Peptide-containing fractions were resolved from β2-microglobulin and solubilized HLA heavy chain using a linear gradient of aqueous 0.1% trifluoroacetic acid (TFA) (buffer A) to 80% acetonitrile/0.1% TFA (buffer B) at a flow rate of 1 mL/min.

Identification of MHC-bound peptides using LC-MS/MS

Peptide-containing RP-HPLC fractions were concentrated using a Labconco CentriVap concentrator and reconstituted in 15 µL of 0.1% formic acid in water. After sonication in a water bath for 10 min, the reconstituted fractions were centrifuged for 10 min at 18,000 × g and transferred to mass spectrometry vials for analysis by an AB SCIEX 5600 TripleTOF mass spectrometer equipped with a Nanospray III ion source and coupled on-line to an Eksigent NanoUltra cHiPLC system. [14,16]. Samples were loaded onto the microfluidic trap column packed with ChromXP C_{18}-CL 3-µm particles (300 Å nominal pore size; equilibrated in 0.1% formic acid/5% acetonitrile) at $5\,\mu l\,min^{-1}$. An analytical ($15\,cm \times 75\,\mu m$ ChromXP C_{18}-CL 3) microfluidic column was then switched in line and peptides separated by linear gradient elution of 0–80% acetonitrile over 90 min ($300\,nl\,min^{-1}$). Mass spectrometry conditions were set to accumulate up to 30 MS/MS spectra per second. Data were analysed with Protein Pilot software (AB SCIEX) and peptide identities were determined subject to strict bioinformatic criteria that included the use of a decoy database to calculate the false discovery rate (FDR). A FDR cut-off of 5 % was applied consistent with the accepted cut off for non-tryptic peptides [18, 19], and the filtered data set was further analyzed manually to exclude redundant peptides, known contaminants and non-specific peptides (i.e. those identified in an anti-HLA-Bw4 control peptide elution experiment).

T cell cultures

Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll-Hypaque centrifugation into RPMI 1640 medium supplemented with 10% FCS (R10 medium). Blood donors were healthy, EBV-seropositive individuals who had given written informed consent. Approval was obtained from the QIMR Berghofer Medical Research Institute Human Ethics Committee (Brisbane, QLD, Australia). T cell cultures were raised by culturing PBMCs ($2x10^6$ per 2ml well) with either irradiated autologous lymphoblastoid cell lines (LCLs) ($2x10^5$ per 2ml well) for EBV-specific cultures or PBMCs that had been pre-coated with the EBV-SEL peptide (1μM for 1h, responder:stimulator ratio 2:1) for SEL-specific cultures. Cultures were supplemented with rIL-2 (20U/ml) from day 3 and analyzed on day 18. Synthetic EBV peptide SEL and self-peptide DEL were synthesized by use of the Merrifield solid-phase method and purchased from New England Peptides (SEL) (Gardner, MA) and Mimotopes (DEL) (Clayton, Australia).

ELISPOT assays

IFN-γ ELISPOT assays were performed using cytokine capture and detection reagents according to manufacturer's instructions (ELISpot^{PRO} for Human IFN-γ, Mabtech, Stockholm, Sweden). Briefly, 96-well nitrocellulose plates pre-coated with anti-IFN-γ mAb were seeded with approximately 5 x 10^4 EBV-specific T cells and peptides SEL or DEL at various concentrations. After incubation for 16 hours at 37°C in 5% CO_2, the cells were discarded, and captured IFN-γ was detected with a biotinylated anti-IFNγ Ab, followed by development with an alkaline phosphatase substrate solution (BCIP/NBT-plus). Spots were counted using an automated plate counter.

pMHC multimer and TRBV staining

T cell cultures (LCL- or peptide-stimulated) from healthy virus carriers were incubated for 30 minutes at 4°C with an SEL-HLA-B*18:01 allophycocyanin (APC)-labelled tetramer [13]. Cells were then washed and incubated with PerCP-Cy5.5 conjugated anti-human CD8 mAb (BioLegend, San Diego, CA), and one of the following phycoerythrin or fluorescein isothiocyanate-labeled TCR β-chain-specific mAbs (Beckman Coulter): Vβ1 (*TRBV9*), Vβ2 (*TRBV20-1*), Vβ3 (*TRBV28*), Vβ4 (*TRBV29*), Vβ5.1 (*TRBV5-1*), Vβ5.2 (*TRBV5-6*), Vβ5.3 (*TRBV5-5*), Vβ6.7 (*TRBV7-1*), Vβ7 (*TRBV4*), Vβ7.2 (*TRBV4-3*), Vβ8 (*TRBV12*), Vβ9 (*TRBV3*), Vβ11 (*TRBV25-1*), Vβ12 (*TRBV10*), Vβ13.1 (*TRBV6-5*), Vβ13.2 (*TRBV6-2*), Vβ13.6 (*TRBV6-6*), Vβ14 (*TRBV27*), Vβ16 (*TRBV14*), Vβ17 (*TRBV19*), Vβ18 (*TRBV18*), Vβ20 (*TRBV30*), Vβ21.3 (*TRBV11-1*), Vβ22 (*TRBV2*) or Vβ23 (*TRBV13*). Cells were washed twice and fixed in 1% paraformaldehyde (PFA) in PBS. These samples were analysed on a FACS LSR-Fortessa flow cytometer using FACSDiva software (BD Biosciences).

Intracellular cytokine staining assay

T cell lines from EBV-seropositive, HLA-B*18:01^{+} donors, raised by stimulation of PBMCs with the SEL peptide, were incubated for 4 hours at 37°C with either SEL or DEL peptide (1μg/ml) in R10 medium supplemented with 5μg/ml Brefeldin A (BioLegend). These cells were then washed and incubated at 4°C for 30 min with fluorescently-labeled mAb specific for cell surface markers (CD8-PerCP-Cy5.5 (Biolegend) and CD4-Alexa flour 488 (Biolegend)). Cells were washed, then fixed and permeabilized with Cytofix/Cytoperm fixation/permeabilization solution (BD PharMingen) at 4°C for 20 min. Next, the cells were washed in Perm/Wash (BD PharMingen), incubated with IFN-γ-PE (BD PharMingen) at 4°C for 30 min, washed with Perm/Wash, resuspended in PBS and analysed on a FACS LSR-Fortessa (BD Biosciences). In experiments evaluating the TRBV staining of DEL-reactive T cells, TRBV staining was performed prior to the fixation and permeabilization steps.

Protein expression, purification and crystallisation

HLA-B*18:01$_{1-276}$ and β2m were transformed in BL21-RIL *E.coli* competent cells and isolated from inclusion bodies similarly as previously described [20]. Thirty milligrams of HLA-B*18:01 heavy chain (hc), 20 mg of β2m and 10 mg of each peptide (DEL or SEL) were dissolved in refolding buffer containing 100 mM Tris-HCl pH 8.0, 400 mM L-Arginine-HCl, 2 mM EDTA, 0.2 mM phenylmethylsulfonyl fluoride, 0.5 mM oxidised glutathione and 5 mM reduced glutathione. Two further 30 mg aliquots of the hc were added to the refolding buffer, leaving at least 3 h between each injection. After 48 h, the protein was dialysed three times against 10 L of 10 mM Tris-HCl pH 8.0 over a period of 24 h. The refolded protein was then purified by elution over anion exchange chromatography.

Crystals of the DEL-HLA-B*18:01 complex were grown by the hanging-drop, vapour-diffusion method at 20°C with a protein/reservoir drop ratio of 1:1, at a concentration of 10 mg/mL in TBS (10 mM Tris-HCl pH 8.0, 150 mM NaCl) using 14- 20% PEG 4K, 0.2 M Amonium Acetate and 0.1 M Na-Citrate pH 5.6.

Data collection and structure determination

The DEL-HLA-B*18:01 crystals were soaked in a cryoprotectant solution containing mother liquor solution with a final PEG concentration of 30% (w/v) PEG 4K and then flash frozen in liquid nitrogen. The data were collected on the MX1 beamline at the Australian Synchrotron (Clayton, Vic) using the ADSC-Quantum 210 CCD detector (at 100K). Data were processed using iMOSFLM [21] and scaled using SCALA on CCP4 suite [22]. The structure was determined by molecular replacement using the Phaser molecular replacement program in PHENIX [23] with the

SEL-HLA-B*18:01 (Protein Data Bank accession number, 4JQV) [13] as a search model minus the peptide. The model was then refined using PHENIX [23] and manually built using COOT [24]. The final model has been validated using the Protein Data Base validation web site and the final refinement statistics are summarized in Table 1. Coordinates have been submitted to PDB database, code: 4XXC. All molecular graphics representations were created using PyMol [25].

Thermal stability assay

To assess the difference in stability between HLA-B*18:01 bound to the DEL self-peptide and SEL EBV peptide, a thermal shift assay was performed. The fluorescent dye Sypro orange was used to monitor the protein unfolding. The thermal stability assay was performed in the Real Time Detection system (Corbett RotorGene 3000), originally designed for PCR. Data was acquired for the two pHLA complexes at two concentrations (5 and 10 mM) in TBS buffer (10 mM Tris-HCl pH8, 150 mM NaCl) in duplicate and temperature was increased at a rate of 1°C/min over a range of 29 to 90°C. The fluorescence intensity was measured with excitation at 530 nm and Emission at 555 nm. The thermal melt (Tm) point represents the temperature for which 50% of the protein is unfolded. The raw data were processed using GraphPad prism software.

Results

*Self-peptide homology to an EBV epitope presented by HLA-B*18:01*

The BZLF1 antigen of EBV includes a highly immunogenic octamer epitope ^{173}SELEIKRY^{180} that is recognized by $CD8^+$ T cells from HLA-B*18:01^+ individuals [13]. Database searches identified a human protein sequence (DELEIKAY) with sequence homology to this EBV epitope. This sequence is from CPSF3L which is a component of the multiprotein integrator complex associated with the C-terminal domain of RNA polymerase II [26], as well as a member of a group of the zinc-dependent hydrolases called the metallo-β-lactamase family [27]. To establish if a peptide corresponding to this sequence can bind to HLA-B*18:01, the stability of this HLA molecule bound to the EBV or self-peptide was compared using a thermal stability assay. The two amino acid difference between the two peptides had little effect on the overall stability of the p-HLA-B*18:01 complexes, with the two peptides showing a similar Tm point of ~70°C (Fig. 1).

In order to determine whether this self-peptide is presented by cells bearing the HLA-B*18:01 allele, we screened mass spectrometry spectra from peptides eluted from C1R cells engineered to express HLA-B*18:01 [13]. ProteinPilot™-based searches detected the octamer DELEIKAY, although the confidence level assigned (~74%) was below the stringent false-discovery rate cut-off that would normally be applied to refine the data from such searches. We also identified the C-terminally extended nonamer peptide DELEIKAYY at a confidence of 99%. Reasoning that both peptides may be presented rather than just the nonamer, with the differential abundance potentially accounting for a higher confidence scoring of the nonamer, we manually interrogated the spectra for the presence of precursor MS1 ions from both peptides (Fig. 2). Doubly charged precursor ions were detected for both the octamer (theoretical m/z of 490.7514; Fig. 2A) and nonamer (theoretical m/z of 572.2821; Fig. 2C) sequences. MS2 fragmentation spectra (Fig. 2B and 2D) triggered from both of these precursors confirmed their correct sequence assignment and thus their presence on B*18:01+ve cells. It is worth noting that the signal intensity for the MS1 and MS2 spectra for the octamer was ~10-fold lower than that observed for the nonamer, and the relatively low level of the former was likely to account for the lower confidence scoring by the ProteinPilot™ algorithm. These data also suggested that the octamer is therefore presented at lower abundance than the nonamer, although without accounting for relative differences in ionisation efficiency between the two peptides it is not possible to state this conclusively without further investigation. We failed to detect either peptide from cells not expressing HLA-B*18:01 (data not shown). These experiments therefore confirm the natural presentation of the DELEIKAY self-peptide homologue of the HLA-B*18:01-restricted EBV T cell epitope on the surface of human cells.

*Cross-recognition by EBV-reactive CD8$^+$ T cells of a self-peptide presented by HLA-B*18:01*

In order to determine whether the DEL self-peptide is cross-recognized by EBV-reactive T cells, polyclonal T cell lines were generated from the PBMCs of healthy EBV-exposed, HLA-B*18:01$^+$ individuals by stimulation with either SEL peptide or autologous EBV-infected cells (LCLs), followed by screening for recognition of the SEL and DEL peptides. IFN-γ ELISPOT assays using various concentrations of each peptide were conducted. All donors showed strong IFN-γ responses to the SEL EBV peptide, while cross-recognition of the human peptide was variable (Fig. 3). Only donors 1 and 2 recognized the DEL self-peptide although, in both cases, higher concentrations of DEL were required for equivalent levels of recognition observed with the EBV peptide. Nonetheless, the EBV-reactive T cell repertoire of Donor 1 appears to avoid autoreactivity by a surprisingly narrow margin, with only 10-fold more of the self-peptide required for equivalent activation compared with the EBV peptide.

To determine the proportion of CD8$^+$ T cells within the SEL-reactive repertoire that cross-recognizes the DEL self-peptide, T cell lines were raised from healthy EBV-exposed, HLA-B*18:01$^+$ individuals by stimulation with SEL peptide and screened by intracellular cytokine staining for recognition of SEL and DEL (Fig. 4). Results showed variability in the percentage of SEL-reactive T cells cross-recognizing DEL, ranging from 60% in Donor 1, 8% in Donor 2, and 20% in Donor 5. This experiment was repeated using a T cell culture from Donor 2 that was generated by using stimulation with the autologous LCL and, although the frequency of SEL (17.0%) and DEL (1.6%)-reactive T cells was lower, the percentage of SEL-reactive T cells cross-recognizing DEL (9%) was very similar to that measured in the T cell line from Donor 2 generated by using SEL peptide stimulation (data not shown). These data therefore demonstrate that T cells with significant avidity for a naturally-presented self-peptide are expanded and activated by EBV infection of some HLA-B*18:01$^+$ individuals.

TCR repertoire of cells cross-recognising the self-peptide and EBV epitope

To determine if SEL/DEL cross-reactivity is mediated by a monoclonal, oligoclonal or polyclonal population of T cells, the next set of experiments analyzed TCR ß-chain variable (TRBV) gene usage of the cross-reactive T cells from Donors 1, 2 and 5. Before examining the cross-reactive subset, we firstly analyzed the total SEL-reactive population using flow cytometry with monoclonal TRBV antibodies in combination with an SEL-HLA-B*18:01 tetramer (Fig. 5A). Although the constraints imposed by tolerance to a self-peptide homologue could be expected to restrict the diversity of a T cell response, relatively diverse TRBV gene usage was observed. Bias towards particular TRBV genes was noted, however, with strong selection of *TRBV10* in all three donors, *TRBV4* in Donors 2 and 5, *TRBV6.5* in Donor 2, and *TRBV25* in Donor 5. We next analyzed the

DEL-cross-reactive subset; however it was not possible to utilize a DEL-HLA-B*18:01 tetramer because preliminary experiments revealed that this tetramer failed to label T cells from all donors (data not shown), presumably due to the low avidity of some of these T cells for this self-peptide-HLA complex, as suggested by the peptide dose-response data shown in Fig. 3. Instead, the DEL-reactive T cells were assessed for intracellular IFN-γ following stimulation with this peptide. As shown in Fig. 5B, this DEL-cross-reactive subset was also surprisingly polyclonal, but with dominant usage of *TRBV10* in Donors 1 and 2, co-dominance of *TRBV25* in Donor 2, and strong usage of both *TRBV11* and *TRBV20* in Donor 5, indicating that a diverse array of TCRs have the capacity to cross-recognize these two peptides.

The self-peptide structure is a mimic of the EBV epitope

To have a better understanding of the cross-reactivity observed for the SEL-specific T cells we next crystallized the DEL self-peptide in complex with the HLA-B*18:01 molecule to compare it with the previously solved HLA-B*18:01-SEL complex structure [13]. The structure of the HLA-B*18:01-DEL complex was solved at a resolution of 1.43 Å (Table 1) revealing a clear electron density for the peptide (Fig. 6A & 6B) in the cleft of HLA-B*18:01. We superimposed the structures of the self and EBV peptides in complex with the HLA-B*18:01 (Fig. 7). The HLA structures were very similar, with a root mean square deviation on the Cα atoms of the α1-α2 domains of 0.143 Å between the two pHLA complexes. The peptides were also adopted a very similar confirmation, with a root mean square deviation on the peptide Cα atoms of 0.17 Å. The major differences between the two structures occurring at the P1 and P7 where the substitutions between the two peptides are located (Fig. 7B). Overall the structure of HLA-B*18:01 in complex with the DEL self-peptide shows that the substitutions with the EBV epitope did not affect the structure of the pHLA complex, and that the viral epitope was a structural mimic of the self-peptide.

We previously showed that the substitution of Ser with Ala at P1 of the EBV peptide did not affect recognition by a polyclonal T cell line, but that Ala substitution at P7 decreased T cell recognition of the EBV epitope [13]. Since the self-peptide contains a P7-Ala, it is likely that this substitution is the basis for the lower TCR avidity implied by the dose-response curves. The P7-Arg is presumably the focus of the EBV-specific T cell response in healthy virus carriers in order to avoid self-reactivity.

Discussion

Although T cells have the capacity to discriminate between an enormous number of peptide-MHC complexes, there is a limit to their specificity due to the small size of the target epitope. A small, strongly immunogenic $CD8^+$ T cell epitope from EBV was found to share six of its eight residues with a sequence from the human CPSF3L protein, including the primary anchor residues, raising questions on the potential of this anti-viral response to be "dangerous" or auto-reactive under certain circumstances. Expression of the CPSF3L protein has been shown to be upregulated by the HIV-1 Tat protein and disrupted by RNA interference demonstrating that CPSF3L is an important regulatory protein for both viral and cellular gene expression [28]. This protein also plays a critical role in snRNA 3' end formation [29] and has been shown to be expressed in both LCLs and human osteoblasts [30].

Our investigations first showed that this self-peptide does bind well to HLA-B*18:01, and is presented on human cells at levels detectable by LC-MS/MS. Furthermore, some T cells raised from healthy donors against this EBV peptide could recognize the human peptide. Although self-tolerance had ensured there was a window of differential sensitivity between T cell recognition of the viral versus self-peptide, in some cases only 10-fold more of the self-peptide was required for equivalent T cell activation to that observed using the EBV peptide – a surprisingly small buffer zone of functional avidity to prevent autoimmunity. The TCR repertoire for the EBV peptide was not highly restricted by the self-peptide homology, and many different TRBV genes were also expressed by the subset of T cells that cross-reacted with the self-peptide. This was unexpected, given that structural analysis showed that the viral-peptide-HLA and self-peptide-HLA complexes were virtually indistinguishable, with the exception of the two mismatched peptide side-chains.

This is the first report, to our knowledge, confirming the natural presentation on human cells of a self-peptide cross-recognized by EBV-reactive $CD8^+$ T cells, for which the presenting HLA allele for the EBV and self-peptide is the same. Earlier work on alloantigen cross-reactivity by EBV-reactive $CD8^+$ T cells demonstrated that molecular mimicry also plays a role in alloreactivity. In this case, EBV-reactive T cells were shown to cross-recognize an HLA-B*08:01-bound viral epitope and an HLA-B*44:02/5-bound peptide from a human ATP binding cassette protein [31]. Previous studies have also described T cell cross-reactivity between EBV and peptides from the MS autoantigen, myelin basic protein (MBP). In one case, $CD4^+$ T cells raised against EBV nuclear antigen 1 cross-reacted with MBP and other myelin proteins [32], and in another example, $CD4^+$ T cells were shown to cross-react with an epitope from the DNA-polymerase protein of EBV (BALF5), and MBP (residues 83–99) [33]. In the latter example, the EBV peptide was presented

by HLA-DRB5*0101 and the MBP peptide was presented by DRB1*1501, both MS-associated restriction elements.

These studies support the notion that EBV-reactive T cells could be involved in autoimmunity via molecular mimicry. There is now a large body of evidence that infection with EBV is an environmental trigger for many chronic autoimmune diseases [12, 34, 35]. Many observations implicate EBV in the pathogenesis of MS; for example, MS patients are almost universally seropositive for EBV, but not for other viruses [34, 36]. Moreover, a clinical history of acute infectious mononucleosis (IM) increases the risk of MS, with a relative risk of 2.3 [37]. Patients with MS also have elevated levels of serum IgG antibodies against many EBV antigens [38]. EBV may also be involved in the salivary gland tissue damage leading to Sjögren's syndrome [39]. Data suggesting this possibility include a higher incidence of EBV reactivation in Sjögren's syndrome patients, and increased levels of EBV antigens and DNA in salivary infiltrating lymphocytes [40]. It has also been shown that the saliva of Sjögren's syndrome patients can stimulate transactivation of BZLF1 (the viral antigen from which the SEL peptide is derived), and that this may be responsible for the autoimmune response in the salivary glands of Sjögren's syndrome patients [41]. An increased seroprevalence of EBV has been noted in North Americans of different ethnicities with systemic lupus erythematosus (SLE) [42, 43]. SLE patients also have an increased EBV genome load in blood compared to healthy individuals [44], with viral loads peaking after initiation of disease flares [45]. Importantly, in the context of the present report, SLE patients also have aberrant expression of the BZLF1 antigen in the blood [46]. It is also relevant to highlight that EBV-reactive $CD8^+$ T cells are enriched in or near the diseased organs of patients with rheumatoid arthritis [47, 48] and MS [49], including a high proportion of T cells specific for BZLF1 [50]. This could reflect a local immune response against EBV in the diseased organs, or non-specific homing of virus-specific T cells to inflamed sites [51].

It is possible that autoreactivity by EBV-reactive T cells is partly responsible for the immunopathology associated with acute IM following primary EBV infection. BZLF1-specific T cells are known to be expanded to huge numbers during acute IM, with one report of T cells specific for a single BZLF1 epitope accounting for 44% of the $CD8^+$ T cells in IM blood [52]. Indeed, an earlier study by Misko *et al.* demonstrated another example of T cell cross-reactivity between a BZLF1 octamer peptide and a self-peptide, both presented by HLA-B8, although in this case, the self peptide was not shown to be naturally presented [53]. Mouse studies certainly support the notion that autoimmunity following infection with pathogens is mediated by effector T cells with low avidity for self and which are not normally primed by endogenous amounts of self antigen [54, 55].

As well as highlighting the potential danger of self-reactivity by T cells expanded in response to common viral infection, the present report also underscores the exquisite specificity of the immune system by showing that a strong T cell response, utilizing a variety of TCRs, can be mounted against a foreign peptide-HLA complex that is essentially identical to a self-peptide-HLA complex with the exception of one TCR-accessible peptide side-chain. Although immune tolerance ensures that the healthy subjects examined in this study are ignorant towards their own cells, increasing self-peptide levels above those naturally presented (by addition of exogenous self-peptide) leads to readily detectible T cell activation. The functional avidity of T cells from one healthy subject was only 10-fold higher for the viral epitope compared to the self-peptide. Such minor differences in T cell activation threshold are presumably sufficient to avoid autoreactivity in most individuals, as supported by previous studies showing that small avidity differences can have a major functional impact on T cell recognition [56, 57].

It is difficult to establish the mechanisms behind the EBV/autoimmunity association because EBV does not infect animals commonly used in models of autoimmune diseases. However, it is notable that recent studies with a humanized mouse model of EBV infection reported the development of erosive arthritis resembling rheumatoid arthritis in the majority of such mice [58]. The present study has provided support for a molecular mimicry mechanism by demonstrating that EBV-reactive $CD8^+$ T cell expansions can be identified in healthy people that can cross-react with a naturally presented self-peptide. The potential for self-reactivity by these T cells is presumably kept under rigorous control by normal self-tolerance mechanisms. However, these EBV/self cross-reactive T cells could pose an autoimmune threat if co-stimulatory adhesion molecules or HLA-B*18:01-DEL levels increased or the T cell activation threshold was reduced as a result of cytokine release during inflammation or acute infections [54]. The cross-reactive T cell populations identified in the present report should be considered for their potential role in autoimmune diseases associated with HLA-B*18:01 such as type 1 diabetes [59].

Acknowledgments

We thank the staff at the Advanced Photon Source and the Australian Synchrotron for assistance with data collection.

Disclosures

The authors have no conflicting financial interests.

Table 1. Data Collection and Refinement Statistics

Data Collection Statistics	
Space group	$\boldsymbol{P}2_12_12_1$
Cell Dimensions (a,b,c) (Å)	50.66, 80.80, 108.98
Resolution (Å)	24.67 - 1.43 (1.50 - 1.43)
Total number of observations	493100 (65298)
Number of unique observations	80489 (11097)
Multiplicity	6.1 (5.9)
Data completeness (%)	95.8 (91.7)
I/σ_I	7.7 (2.2)
R_{pim} [a] (%)	4.7 (28.3)
Refinement Statistics	
Non-hydrogen atoms	
Protein	3265
Water	516
R_{factor} [b] (%)	20.5
R_{free} [b] (%)	23.5
Rms deviations from ideality	
Bond lengths (Å)	0.006
Bond angles (°)	1.064
Ramachandran plot (%)	
Favored region	97.7
Disallowed region	0

[a] $R_{p.i.m} = \Sigma_{hkl}\,[1/(N-1)]^{1/2}\,\Sigma_i\,|\,I_{hkl,\,i} - \langle I_{hkl}\rangle\,|\,/\,\Sigma_{hkl}\,\langle I_{hkl}\rangle$. [b] $R_{factor} = \Sigma_{hkl}\;\;|\,|\,F_o\,|\,-\,|\,F_c\,|\,|\,/\,\Sigma_{hkl}\,|\,F_o\,|$ for all data except ≈ 5% which were used for R_{free} calculation.

References

1. Starr, T. K., S. C. Jameson, and K. A. Hogquist. 2003. Positive and Negative Selection of T Cells. *Annual Reviews Immunology* 21: 139-176.
2. Nikolic-Zugic, J., and M. J. Bevan. 1990. Role of self-peptides in positively selecting the T-cell repertoire. *Nature* 344: 65-67.
3. Hogquist, K. A., T. A. Baldwin, and S. C. Jameson. 2005. Central tolerance: learning self-control in the thymus. *Nat. Rev. Immunol.* 5: 772-782.
4. Zehn, D., and M. J. Bevan. 2006. T cells with low avidity for a tissue-restricted antigen routinely evade central and peripheral tolerance and cause autoimmunity. . *Immunity* 25: 261-270.
5. Turner, S. J., P. C. Doherty, J. McCluskey, and J. Rossjohn. 2006. Structural determinants of T-cell receptor bias in immunity. *Nature Reviews Immunology* 6: 883-894.
6. Gras, S., S. R. Burrows, S. J. Turner, A. K. Sewell, J. McCluskey, and J. Rossjohn. 2012. A structural voyage toward an understanding of the MHC-I-restricted immune response: lessons learned and much to be learned. *Immunol Rev* 250: 61-81.
7. Oldstone, M. B. 1987. Molecular mimicry and autoimmune disease. *Cell* 50: 819-820.
8. Luo, A.-M., K. M. Garza, D. Hunt, and K. S. K. Tung. 1993. Antigen Mimicry in Autoimmune Disease Sharing of Amino Acid Residues Critical for Pathogenic T Cell Activation. *J. Clin. Invest.* 92: 2117-2123.
9. Harkiolaki, M., S. L. Holmes, P. Svendsen, J. W. Gregersen, L. T. Jensen, R. McMahon, M. A. Friese, G. van Boxel, R. Etzensperger, J. S. Tzartos, K. Kranc, S. Sainsbury, K. Harlos, E. D. Mellins, J. Palace, M. Esiri, M., P. A. van der Merwe, E. Y. Jones, and L. Fugger. 2009. T Cell-Mediated Autoimmune Disease Due to Low-Affinity Crossreactivity to Common Microbial Peptides. *Immunity* 30: 348-357.
10. Hislop, A. D., G. S. Taylor, D. Sauce, and A. B. Rickinson. 2007. Cellular Responses to Viral Infection in Human: Lessons from Epstein-Barr Virus. *Annual Reviews Immunology* 25: 587-617.
11. Long, H. M., G. S. Taylor, and A. B. Rickinson. 2011. Immune defence against EBV and EBV-associated disease. *Current Opinion in Immunology* 23: 258-264.
12. Libbey, J. E., M. F. Cusick, and R. S. Fujinami. 2014. Role of pathogens in multiple sclerosis. *Int Rev Immunol* 33: 266-283.
13. Rist, M. J., A. Theodossis, N. P. Croft, M. A. Neller, A. Welland, Z. Chen, L. C. Sullivan, J. M. Burrows, J. J. Miles, R. M. Brennan, S. Gras, R. Khanna, A. G. Brooks, J. McCluskey, A. W. Purcell, J. Rossjohn, and S. R. Burrows. 2013. HLA peptide length preferences control CD8+ T cell responses. *J Immunol* 191: 561-571.
14. Illing, P. T., J. P. Vivian, N. L. Dudek, L. Kostenko, Z. Chen, M. Bharadwaj, J. J. Miles, L. Kjer-Nielsen, S. Gras, N. A. Williamson, S. R. Burrows, A. W. Purcell, J. Rossjohn, and J. McCluskey. 2012. Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. *Nature* 486: 554-558.
15. Scull, K. E., N. L. Dudek, A. J. Corbett, S. H. Ramarathinam, D. G. Gorasia, N. A. Williamson, and A. W. Purcell. 2012. Secreted HLA recapitulates the immunopeptidome and allows in-depth coverage of HLA A*02:01 ligands. *Mol Immunol* 51: 136-142.
16. Dudek, N. L., C. T. Tan, D. G. Gorasia, N. P. Croft, P. T. Illing, and A. W. Purcell. 2012. Constitutive and Inflammatory Immunopeptidome of Pancreatic beta-Cells. *Diabetes* 61: 3018-3025.
17. Tan, C. T., N. P. Croft, N. L. Dudek, N. A. Williamson, and A. W. Purcell. 2011. Direct quantitation of MHC-bound peptide epitopes by selected reaction monitoring. *Proteomics* 11: 2336-2340.

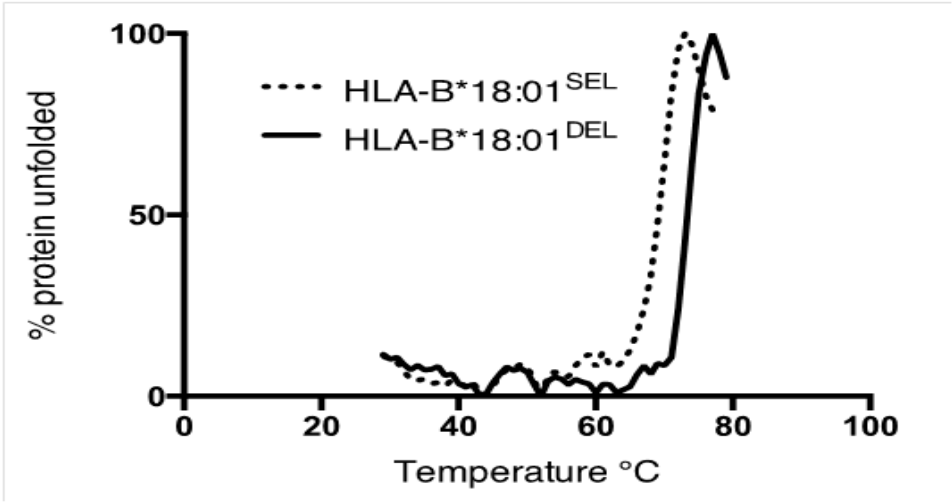
18. Illing, P. T., J. P. Vivian, N. L. Dudek, L. Kostenko, Z. Chen, M. Bharadwaj, J. J. Miles, L. Kjer-Nielsen, S. Gras, and N. A. Williamson. 2012. Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. *Nature* 486: 554-558.
19. Schittenhelm, R. B., T. C. Lim Kam Sian, P. G. Wilmann, N. L. Dudek, and A. W. Purcell. 2014. Revisiting the arthritogenic peptide theory: Quantitative not qualitative changes in the peptide repertoire of HLA-B27 allotypes. *Arthritis Rheumatol*.
20. Gras, S., S. R. Burrows, L. Kjer-Nielsen, C. S. Clements, Y. C. Liu, L. C. Sullivan, M. J. Bell, A. G. Brooks, A. W. Purcell, J. McCluskey, and J. Rossjohn. 2009. The shaping of T cell receptor recognition by self-tolerance. *Immunity* 30: 193-203.
21. Battye, T. G. G., L. Kontogiannis, O. Johnson, H. R. Powell, and A. G. W. Leslie. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* 67: 271-281.
22. Winn, M. D., C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannus, E. A. Potterton, H. R. Powell, R. Read, A. Vagin, and K. S. Wilson. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67: 235-242.
23. Adams, P., P. Afonine, G. Bunkoczi, V. Chen, N. Echols, J. Headd, L. Hung, S. Jain, G. Kapral, R. Grosse Kunstleve, A. McCoy, N. Moriarty, R. Oeffner, R. Read, D. Richardson, J. Richardson, T. Terwilliger, and P. Zwart. 2011. The Phenix software for automated determination of macromolecular structures. *Methods* 55: 94-106.
24. Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60: 2126-2132.
25. DeLano, W. L. 2002. The PyMOL Molecular Graphics System. *DeLano Scientific* Palo Alto, CA.
26. Baillat, D., M.-A. Hakimi, A. M. Naar, A. Shilatifard, N. Cooch, and R. Shiekhattar. 2005. Integrator, a Multiprotein Mediator of Small Nuclear RNA Processing, Associates with the C-Terminal Repeat of RNA Polymerase II. *Cell* 123: 265-276.
27. Callebaut, I., D. Moshous, J. P. Mornon, and J. P. de Villartay. 2002. Metallo-beta-lactamase fold within the nucleic acids processing enzymes: the beta CASP family. *Nucleic Acid Res* 30: 3592-3601.
28. Calzado, M. A., R. Sancho, and E. Munoz. 2004. Human Immunodeficiency Virus Type 1 Tat Increase Expression of Cleavage and Polyadenylation Specificity Factor 73-Kilodalton Subunit Modulating Cellular and Viral Expression. *J Virol* 78: 6846-6854.
29. Albrecht, T. R., and E. J. Wagner. 2012 snRNA 3' End Formation Requires Heterodimeric Association of Integrator Subunits. *Mol Cell Biol* 32: 1112-1123.
30. Kwan, T., E. Grundberg, V. Koka, B. Ge, K. C. L. Lam, C. Dias, A. Kindmark, H. Mallmin, O. Ljunggren, F. Rivadeneira, K. Estrada, J. B. van Meurs, A. Uitterlinden, M. Karlsson, C. Ohlsson, D. Mellstrom, O. Nilsson, T. Pastinen, and J. Majewski. 2009. Tissue Effect on Genetic Control of Transcript Isoform Variation. *PLoS Gen* 5: e1000608.
31. Macdonald, W. A., Z. Chen, S. Gras, J. K. Archbold, F. E. Tynan, C. S. Clements, M. Bharadwaj, L. Kjer-Nielsen, P. M. Saunders, and M. C. Wilce. 2009. T cell allorecognition via molecular mimicry. *Immunity* 31: 897-908.
32. Lunemann, J. D., I. Jelcic, S. Roberts, A. Lutterotti, B. Tackenberg, R. Martin, and C. Munz. 2008. EBNA1-specific T cells from patients with multiple sclerosis cross react with myelin antigens and co-produce IFN-gamma and IL-2. *J Exp Med* 205: 1763-1773.
33. Lang, H. L., H. Jacobsen, S. Ikemizu, C. Andersson, K. Harlos, L. Madsen, P. Hjorth, L. Sondergaard, A. Svejgaard, K. Wucherpfennig, D. I. Stuart, J. I. Bell, E. Y. Jones, and L. Fugger. 2002. A functional and structural basis for TCR cross-reactivity in multiple sclerosis. *Nat Immunol* 3: 940-943.
34. Ascherio, A., and K. L. Munger. 2007. Environmental risk factors for multiple sclerosis. Part I: the role of infection. *Ann Neurol* 61: 288-299.

35. Vento, S., L. Guella, F. Mirandola, F. Cainelli, G. Di Perri, M. Solbiati, T. Ferraro, and E. Concia. 1995. Epstein-Barr virus as a trigger for autoimmune hepatitis in susceptible individuals. *Lancet* 346: 608-609.
36. Pakpoor, J., G. Disanto, J. E. Gerber, R. Dobson, U. C. Meier, G. Giovannoni, and S. V. Ramagopalan. 2013. The risk of developing multiple sclerosis in individuals seronegative for Epstein-Barr virus: a meta-analysis. *Mult Scler* 19: 162-166.
37. Thacker, E. L., F. Mirzaei, and A. Ascherio. 2006. Infectious mononucleosis and risk for multiple sclerosis: a meta-analysis. *Ann Neurol* 59: 499-503.
38. Pender, M. P., and S. R. Burrows. 2014. Epstein-Barr virus and multiple sclerosis: potential opportunities for immunotherapy. *Clin Transl Immunology* 3: e27.
39. Igoe, A., and R. H. Scofield. 2013. Autoimmunity and infection in Sjogren's syndrome. *Curr Opin Rheumatol* 25: 480-487.
40. Mariette, X., J. Gozlan, D. Clerc, M. Bisson, and F. Morinet. 1991. Detection of Epstein-Barr virus DNA by in situ hybridization and polymerase chain reaction in salivary gland biopsy specimens from patients with Sjogren's syndrome. *Am J Med* 90: 286-294.
41. Inoue, H., K. Mishima, S. Yamamoto-Yoshida, R. Ushikoshi-Nakayama, Y. Nakagawa, K. Yamamoto, K. Ryo, F. Ide, and I. Saito. 2012. Aryl hydrocarbon receptor-mediated induction of EBV reactivation as a risk factor for Sjogren's syndrome. *J Immunol* 188: 4654-4662.
42. James, J. A., B. R. Neas, K. L. Moser, T. Hall, G. R. Bruner, A. L. Sestak, and J. B. Harley. 2001. Systemic lupus erythematosus in adults is associated with previous Epstein-Barr virus exposure. *Arthritis Rheum* 44: 1122-1126.
43. McClain, M. T., B. D. Poole, B. F. Bruner, K. M. Kaufman, J. B. Harley, and J. A. James. 2006. An altered immune response to Epstein-Barr nuclear antigen 1 in pediatric systemic lupus erythematosus. *Arthritis Rheum* 54: 360-368.
44. Moon, U. Y., S. J. Park, S. T. Oh, W. U. Kim, S. H. Park, S. H. Lee, C. S. Cho, H. Y. Kim, W. K. Lee, and S. K. Lee. 2004. Patients with systemic lupus erythematosus have abnormally elevated Epstein-Barr virus load in blood. *Arthritis Res Ther* 6: R295-302.
45. Larsen, M., D. Sauce, C. Deback, L. Arnaud, A. Mathian, M. Miyara, D. Boutolleau, C. Parizot, K. Dorgham, L. Papagno, V. Appay, Z. Amoura, and G. Gorochov. 2011. Exhausted cytotoxic control of Epstein-Barr virus in human lupus. *PLoS Pathog* 7: e1002328.
46. Gross, A. J., D. Hochberg, W. M. Rand, and D. A. Thorley-Lawson. 2005. EBV and systemic lupus erythematosus: a new perspective. *J Immunol* 174: 6599-6607.
47. David-Ameline, J., A. Lim, F. Davodeau, M. A. Peyrat, J. M. Berthelot, G. Semana, C. Pannetier, J. Gaschet, H. Vie, J. Even, and M. Bonneville. 1996. Selection of T cells reactive against autologous B lymphoblastoid cells during chronic rheumatoid arthritis. *J Immunol* 157: 4697-4706.
48. Scotet, E., J. David-Ameline, M. A. Peyrat, A. Moreau-Aubry, D. Pinczon, A. Lim, J. Even, G. Semana, J. M. Berthelot, R. Breathnach, M. Bonneville, and E. Houssaint. 1996. T cell response to Epstein-Barr virus transactivators in chronic rheumatoid arthritis. *J Exp Med* 184: 1791-1800.
49. Jaquiery, E., S. Jilek, M. Schluep, P. Meylan, A. Lysandropoulos, G. Pantaleo, and R. A. Du Pasquier. 2010. Intrathecal immune responses to EBV in early MS. *Eur J Immunol* 40: 878-887.
50. Tan, L. C., A. G. Mowat, C. Fazou, T. Rostron, H. Roskell, P. R. Dunbar, C. Tournay, F. Romagne, M. A. Peyrat, E. Houssaint, M. Bonneville, A. B. Rickinson, A. J. McMichael, and M. F. Callan. 2000. Specificity of T cells in synovial fluid: high frequencies of CD8(+) T cells that are specific for certain viral epitopes. *Arthritis Res* 2: 154-164.
51. Lossius, A., J. N. Johansen, O. Torkildsen, F. Vartdal, and T. Holmoy. 2012. Epstein-Barr virus in systemic lupus erythematosus, rheumatoid arthritis and multiple sclerosis-association and causation. *Viruses* 4: 3701-3730.

52. Callan, M. F., L. Tan, N. Annels, G. S. Ogg, J. D. Wilson, C. A. O'Callaghan, N. Steven, A. J. McMichael, and A. B. Rickinson. 1998. Direct visualization of antigen-specific CD8+ T cells during the primary immune response to Epstein-Barr virus In vivo. *J Exp Med* 187: 1395-1402.
53. Misko, I.S., S. M. Cross, R. Khanna, S. L. Elliott, C. Schmidt, S. J. Pye, and S. L. Silins. 1999. Crossreactive recognition of viral, self, and bacterial peptide ligands by human class I-restricted cytotoxic T lymphocyte clonotypes: implications for molecular mimicry in autoimmune disease. *Proc. Natl. Acad. Sci. U. S. A.* 96: 2279-2284.
54. Zehn, D., and M. J. Bevan. 2006. T cells with low avidity for a tissue-restricted antigen routinely evade central and peripheral tolerance and cause autoimmunity. *Immunity* 25: 261-270.
55. Oldstone, M. B., M. Nerenberg, P. Southern, J. Price, and H. Lewicki. 1991. Virus infection triggers insulin-dependent diabetes mellitus in a transgenic model: role of anti-self (virus) immune response. *Cell* 65: 319-331.
56. Hislop, A. D., U. Palendira, A. M. Leese, P. D. Arkwright, P. S. Rohrlich, S. G. Tangye, H. B. Gaspar, A. C. Lankester, A. Moretta, and A. B. Rickinson. 2010. Impaired Epstein-Barr virus-specific $CD8^+$ T-cell function in X-linked lymphoproliferative disease is restricted to SLAM family-positive B-cell targets. *Blood* 116: 3249-3257.
57. Burrows, S. R., S. J. Rodda, A. Suhrbrier, H. M. Geysen, and D. J. Moss. 1992. The specificity of recognition of a cytotoxic T lymphocyte epitope. *Eur. J. Immunol.*22: 191-195.
58. Kuwana, Y., M. Takei, M. Yajima, K. Imadome, H. Inomata, M. Shiozaki, N. Ikumi, T. Nozaki, H. Shiraiwa, N. Kitamura, J. Takeuchi, S. Sawada, N. Yamamoto, N. Shimizu, M. Ito, and S. Fujiwara. 2011. Epstein-Barr virus induces erosive arthritis in humanized mice. *PLoS One* 6: e26630.
59. Nejentsev, S., J. M. M. Howson, N. M. Walker, J. Szeszko, S. F. Field, H. E. Stevens, P. Reynolds, M. Hardy, E. King, J. Masters, J. Hulme, L. M. Maier, D. Smyth, R. Bailey, J. D. Cooper, G. Ribas, R. D. Campbell, T. W. T. C. C. Consortium, D. G. Clayton, and J. A. Todd. 2007. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. *Nature* 450: 887-892.

Figures and Figure Legends

Fig. 1



HLA-B*18:01 complex	T_m (°C)
DELEIKAY (self-peptide)	72.3 ± 1.3
SELEIKRY (EBV peptide)	68.8 ± 0.3

FIGURE 1. Thermal stability of SEL-HLA-B*18:01 and DELHLA-B*18:01 complexes. Thermal unfolding of peptide-HLA-B*18:01 complexes was measured by real time detection. Graph shows curves for the SELEIKRY peptide (dotted line) and DELEIKAY peptide (solid line) bound to HLA-B*18:01. T_m is the melting point temperature needed to unfold 50% of the protein.

Figure 2

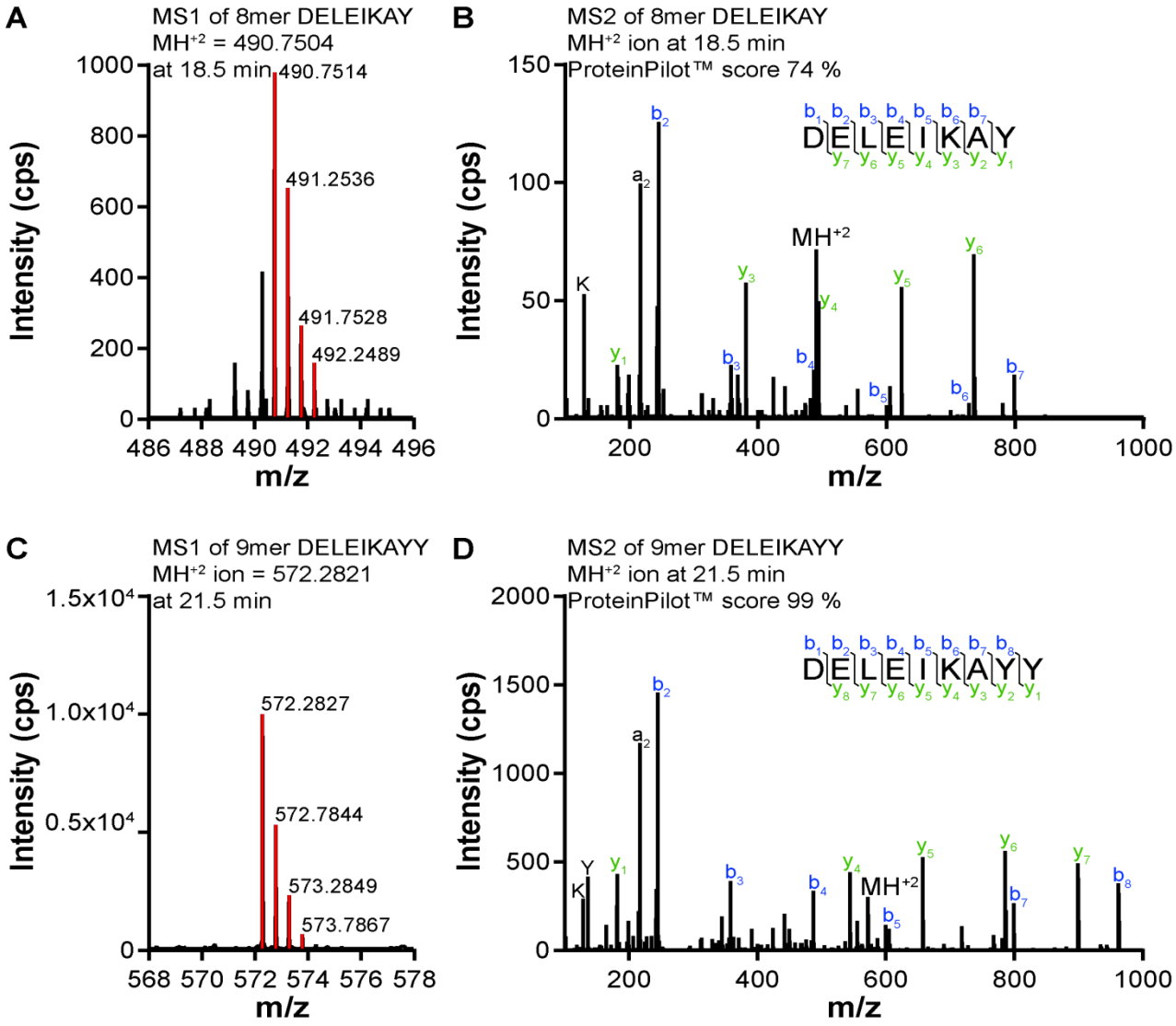


FIGURE 2. Mass spectrometry-based detection of the HLA-B*18:01-bound peptides DELEIKAY and DELEIKAYY. Peptides were eluted from C1R-B*18:01 cells and identified by liquid chromatography tandem mass spectrometry on an AB SCIEX TripleTOF® 5600+ mass spectrometer and data manually interrogated for the presence of the DEL 8mer and 9mer sequences. **(A)** MS1 isotope series (red) for the doubly charged ion of the 8mer DELEIKAY. **(B)** MS2 fragmentation spectra of DELEIKAY, highlighting dominant fragment ions. **(C)** MS1 isotope series (red) for the doubly charged ion of the 9mer DELEIKAYY. **(D)** MS2 fragmentation spectra of DELEIKAYY, highlighting dominant fragment ions.

Figure 3

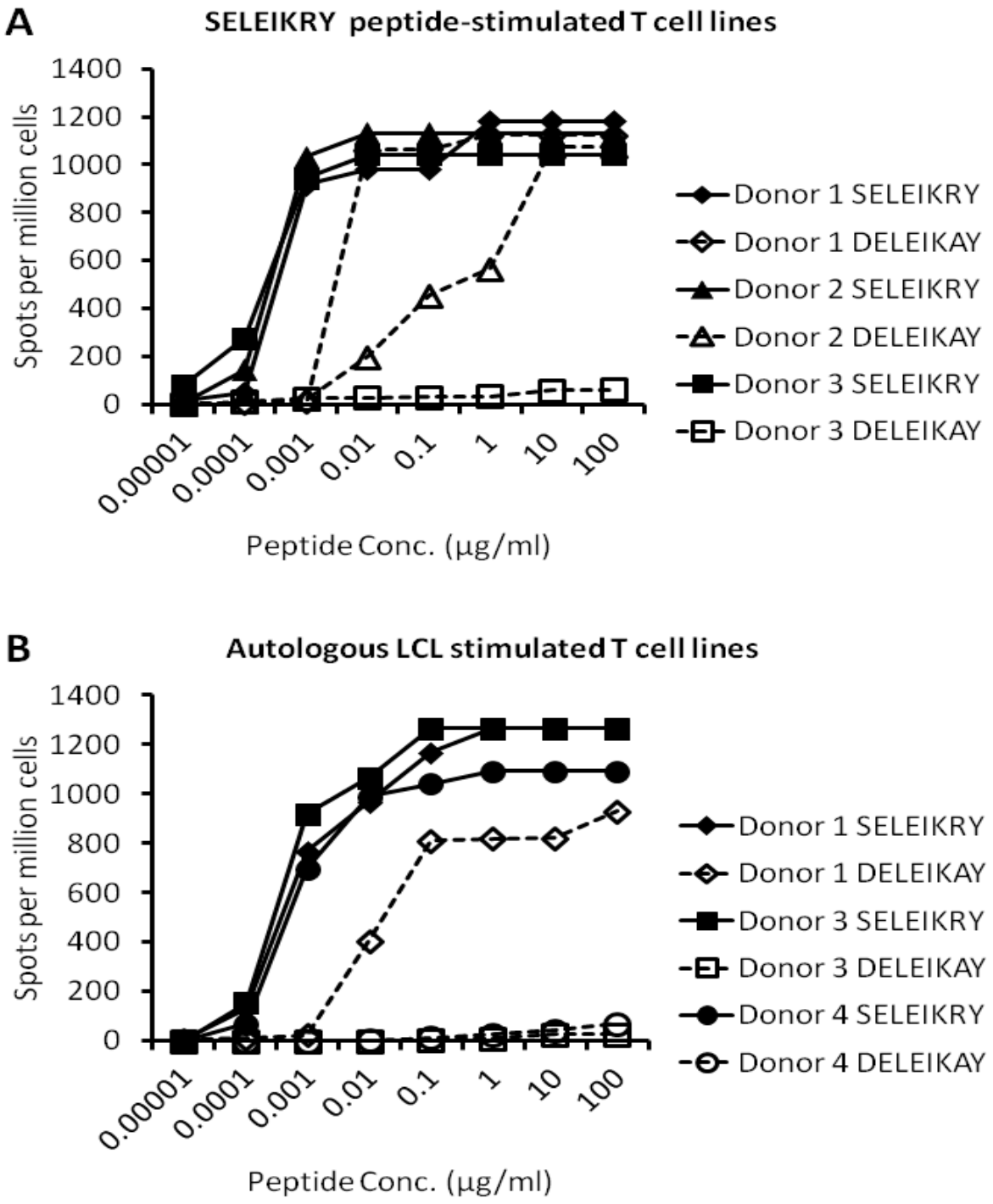


FIGURE 3. HLA-B*18:01^{+} individuals tested for recognition of SELEIKRY and DELEIKAY. IFN-γ ELISPOT assays were used to test (**A**) SEL-specific T cell lines from Donor 1 (HLA-A2, A3, B7, B18, Cw7, Cw12), Donor 2 (HLA-A2, B18, B62, Cw1, Cw5) and Donor 3 (HLA-A1, A30, B8, B18, Cw7, Cw7) and (**B**) EBV-specific T cell lines from Donor 1, Donor 3 and Donor 4 (HLA-A23, A25, B18, B49, Cw7, Cw12) for recognition of indicated peptides at various concentrations.

Figure 4

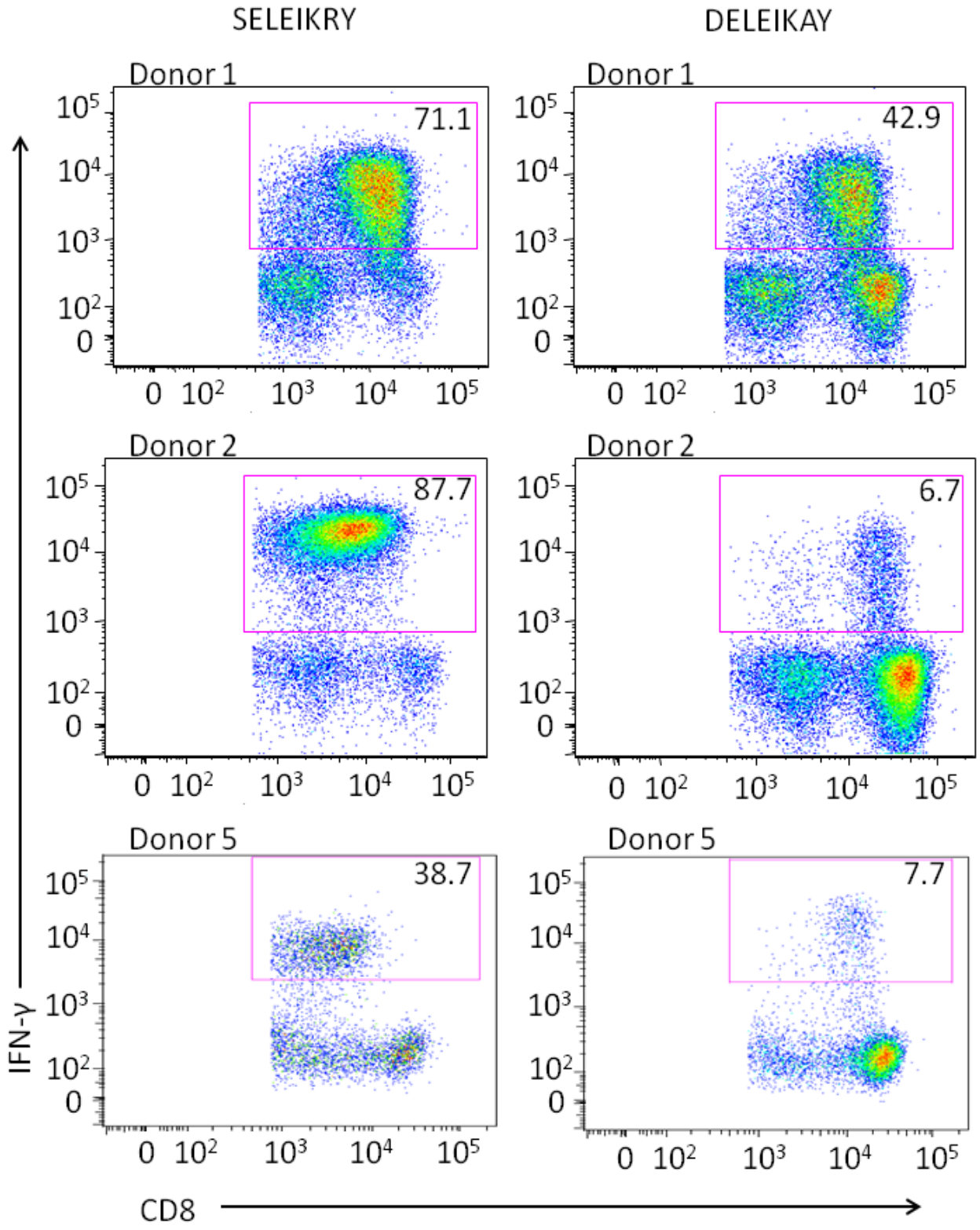


FIGURE 4. SEL-specific T cell lines from Donor 1, Donor 2 and Donor 5 (HLA-A3, A25, B18, B4403) were exposed to SEL or DEL peptide (1μg/ml) and assessed for IFN-γ production by intracellular cytokine staining.

Figure 5

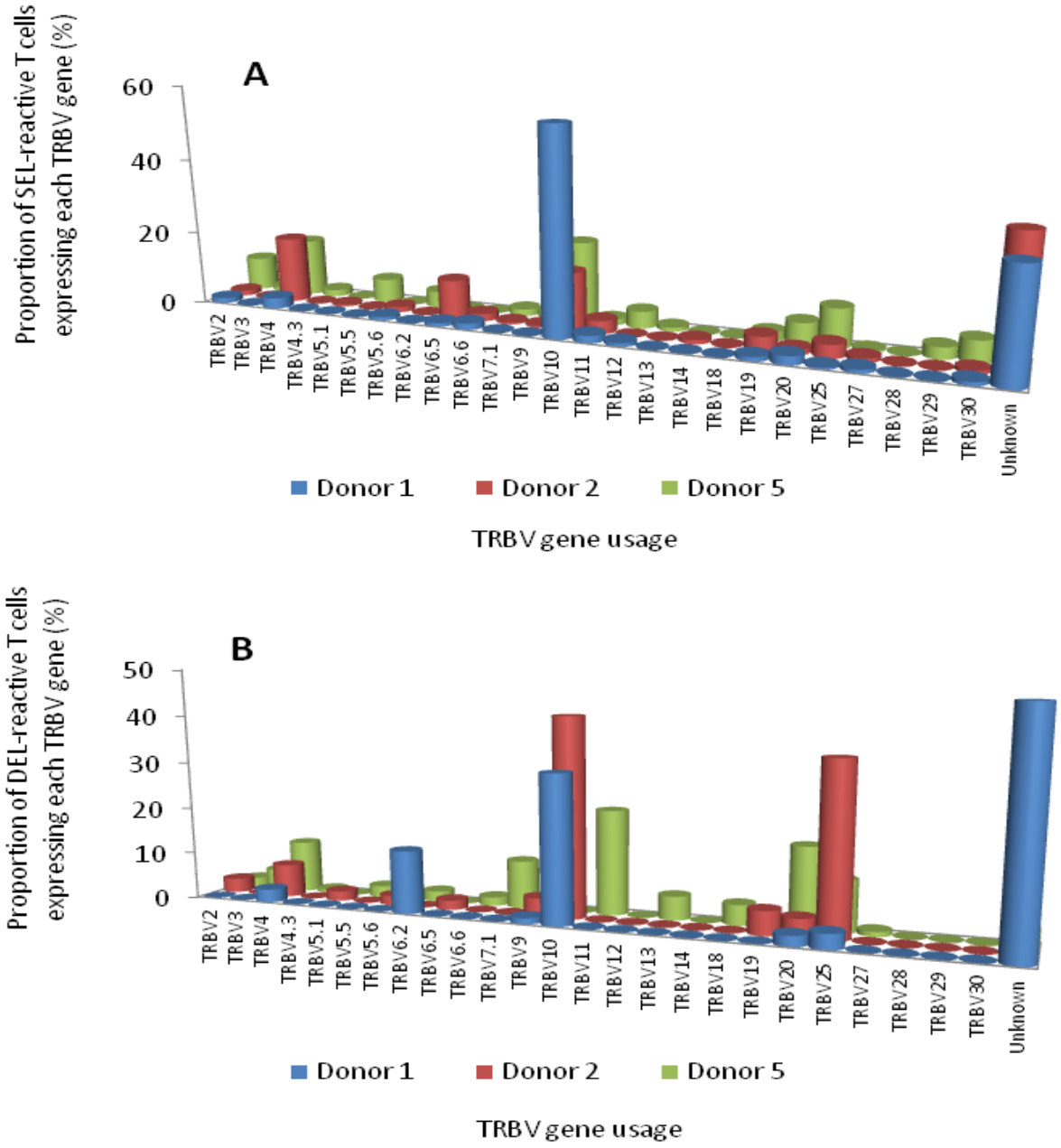


FIGURE 5. TRBV gene usage in the recognition of the SEL and DEL peptides by EBV-exposed, HLA-B*1801^{+} individuals. TRBV gene usage was assessed in T cell cultures using flow cytometry with a panel of TCR-β-chain-specific mAbs. (**A**) An SEL-HLA-B*1801 tetramer was used to examine the SEL-reactive T cell subset, and (**B**) intracellular IFN-γ production following stimulation with the DEL peptide was used to examine the DEL-reactive T cell subset.

FIGURE 6

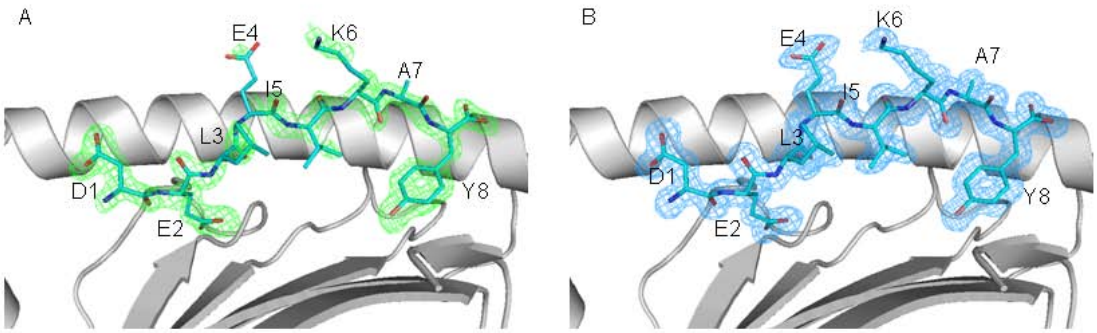


FIGURE 6. The crystal structure of HLA-B*18:01-DEL. **(A)** Structure of HLA-B*18:01 (white cartoon) in complex with the DEL peptide (cyan sticks) overlaid with the unbiased Fo-Fc density map for the peptide contoured at 3σ (green mesh). **(B)** Structure of HLA-B*18:01 (white cartoon) and the DEL peptide (cyan sticks) overlaid with the refined 2Fo-Fc density map contoured at 1σ (blue mesh). The α2 helix has been removed for clarity in this figure.

FIGURE 7

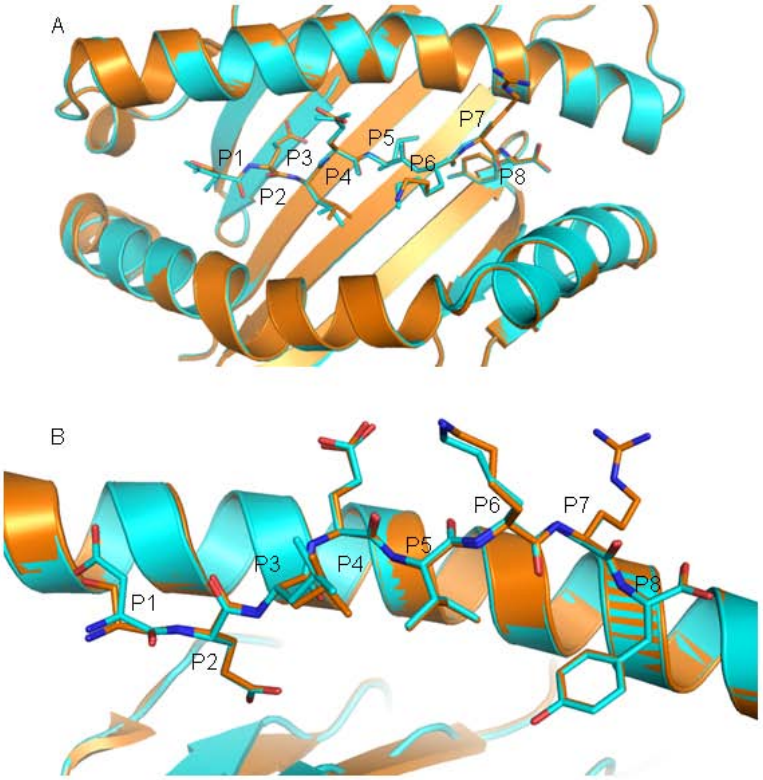


FIGURE 7. Superimposition of the HLA-B*18:01-DEL (cyan) with the HLA-B*18:01-SEL (orange) structures (PDB: 4JQV). Superimposition of the crystal structures using main chain atoms from residues 1 – 181 of the heavy chain (root mean square deviation = 0.144 Å), viewed from above **(A)** or the side **(B)**. The heavy chain atoms are shown as cartoons and the peptide atoms are shown as sticks.

Chapter 7

Discussion

7.1 Overview

In determining the presence of novel $CD8^+$ T cell epitopes, including noncanonical epitopes, the latent herpes virus EBV is a superb model due to the strong responses detected in healthy virus carriers. In this thesis, eight highly immunogenic EBV antigens including four latent antigens and four lytic antigens were chosen to address the first of three main aims. The initial aim was to determine the relative prevalence of peptide epitopes of over 10 amino acids in length in $CD8^+$ T cell recognition of EBV. This aim was achieved by investigating the $CD8^+$ T cell responses in healthy virus carriers toward seven of these EBV antigens (Chapter 3). The close examination of immediate early EBV antigen BZLF1 (Chapter 4) fulfilled the second aim by showing that this highly immunogenic antigen did indeed include novel $CD8^+$ T cell epitopes. The focus of the third aim was to determine if the peptide length preference of class I HLA influences epitope selection in the EBV-specific T cell response. Chapter 5 demonstrates that HLA peptide length preferences do indeed control EBV–specific $CD8^+$ T cell responses. The fourth aim was to characterise a cross-reactive T cell response to an EBV epitope and a self-peptide presented by HLA-B*1801. In Chapter 6 it is shown that individuals with T cells cross-reactive to both the EBV epitope and the self-peptide express a diverse array of T cell receptors. This thesis reports a comprehensive study of the $CD8^+$ T cell response to eight EBV antigens determining novel $CD8^+$ T cell epitopes, HLA peptide length preferences and cross-reactivity with a self antigen.

7.2 Determine the relative prevalence of peptide epitopes of over 10 amino acids in length in $CD8^+$ T cell recognition of EBV

Class I MHC molecules present peptides derived from viral antigens on the cell surface for recognition by $CD8^+$ T cells. These peptides are generally between 8 and 10 amino acids in length. The breakthrough discovery enabling more efficient CTL epitope mapping was that, for most MHC alleles, two of the pockets within the peptide binding groove have a preference for one or two amino acids at certain anchor positions within the peptide. In most cases, previous studies have utilised web-based algorithms designed to predict $CD8^+$ T cell epitopes to lengths of 8, 9 or 10 amino acids. These widely used algorithms may be behind this bias in the literature of MHC class I-presented epitopes of 8 to 10 amino acids in length. Several studies have shown noncanonical epitopes naturally processed and presented for $CD8^+$ T cell recognition including a number in EBV (Green, Miles et al. 2004) (Tynan, Borg et al. 2005, Miles, Borg et al. 2006, Rist, Theodossis et al. 2013).

Thus, investigating the relative prevalence of noncanonical epitopes in EBV by utilising overlapping peptides from immunogenic EBV antigens as opposed to web-based algorithm prediction may determine natural presentation of longer epitopes. In Chapter 3, 20-mer peptides

(overlapping by 15 aa) corresponding to the EBV lytic antigens BRLF1, BMLF1 and BMRF1 and latent antigens EBNA3A, -3B, -3C and EBNA1 were assayed for $CD8^+$ T cell recognition. IFN-γ ELISpot assays shown to be the best high-throughput screening method (Tischer, Dieks et al. 2014) were used to test forty-one healthy virus carriers with around 1000 peptides to address this aim. This study provides an extensive list of novel class I EBV epitopes identified from these antigens, including four noncanonical epitopes.

Results from Pudney *et al* suggest a focusing of $CD8^+$ T cell responses toward epitopes from immediate early (IE) and early (E) proteins and a hierarchy of immunodominance amongst the EBV lytic cycle antigens (Pudney, Leese et al. 2005). The findings of this study confirm the immunodominance pattern of IE proteins (BRLF1, in addition to BZLF1 (Chapter 4 and 5)) being immunodominant over E (BMLF1 and BMRF1) proteins. Immediate early proteins BRLF1 and BZLF1 activate the expression of early proteins including transactivator BMLF1 (Lieberman, O'Hare et al. 1986) and BMRF1, involved in viral DNA replication (Pearson, Luka et al. 1987) among others. Pepperl *et al* defined several BRLF1 epitopes including an HLA-A24-restricted decamer epitope ^{28}DYCNVLNKEF^{37} and ^{134}ATIGTAMYK^{142} restricted to HLA-A11, as well as the HLA-Cw4 restricted ^{393}ERPIFPHPSKPTFLP^{407} epitope and HLA-B61-restricted ^{529}QKEEAAICGQMDLS^{543} peptide (Pepperl, Benninger-Doring et al. 1998). Our results demonstrate five novel epitopes for the IE BRLF1 antigen and four novel epitopes from each of the E proteins BMLF1 and BMRF1. Included in the novel BMRF1 $CD8^+$ T cell epitopes is one of the four noncanonical epitopes, 11-mer epitope ^{146}MPYMPPASDRL^{156} (Table 2, Chapter 3).

Comparing the number of defined IE epitopes to the E epitopes there are fewer epitopes from the E lytic proteins BMLF1 and BMRF1. Steven *et al* have defined epitopes from both of these proteins. These include the HLA-A2-restricted ^{259}GLCTLVAML^{267} and HLA-B18-restricted ^{397}DEVEFLGHY^{405} epitopes which are both from the BMLF1 protein (Steven, Annels et al. 1997). BMRF1 epitopes include the HLA-Cw6 ^{268}YRSGIIAVV^{276} peptide which has also been found to be restricted by HLA-B39 by Pudney *et al* and the ^{86}FRNLAYGRTCVLGKE^{100} epitope with HLA-Cw3 restriction, for which the minimal active sequence has not been precisely mapped (Steven, Annels et al. 1997) (Pudney, Leese et al. 2005).

The hierarchy of immunodominance of both lytic and latent EBV antigens, however, depicts the latent EBNA-3A, -3B and -3C antigens as immunodominant over the lytic IE and E proteins, based on the total number of defined epitopes. The relative immunodominance hierarchy in this study (with the exclusion of BZLF1), based on the identification of 28 latent and 13 lytic CTL epitopes, is EBNA3A, EBNA3C, EBNA3B, BRLF1, BMLF1 and BMRF1 and lastly EBNA1, in accord with the findings of Hislop *et al* (Hislop, Taylor et al. 2007). This focus of CTL epitope responses towards the EBNA 3 group of proteins may be a reflection of these proteins accounting

for >60% of the latent antigen's coding capacity (Taylor, Long et al. 2015). Each antigen's position in the hierarchy of immunodominance likely correlates with the kinetics of the virus. The development of memory responses to EBV antigens post-IM involves the maintenance of slow latent responses while certain lytic responses disappear (Hislop, Annels et al. 2002, Woodberry, Suscovich et al. 2005). The strongest individual lytic epitope-specific responses appear to be directed towards epitopes that are efficiently presented on the surface of lytically infected cells (Pudney, Leese et al. 2005), and the CTL response becomes focused on targets expressed early in the lytic cycle (Hislop, Taylor et al. 2007).

Interestingly, an 11-mer epitope was defined in each of the EBNA3 proteins with the overall distribution of these noncanonical epitopes in antigens BMRF1, EBNA3A, EBNA3B and EBNA3C. The extent of the abundance or prevalence of noncanonical epitopes from antigens of interest was determined by combining the overall numbers of novel epitopes of this study with defined epitopes from previous studies. Of the 41 novel epitopes that were defined in this study, 4 (~10%) were noncanonical 11mers (Chapter 3, Table 2). Of note, three of these epitopes have a proline at the second residue and are recognised by members of the HLA–B7 supertype family concurring with previous findings (Burrows, Rossjohn et al. 2006, Burrows, Bell et al. 2008) and implicating an association between particular HLA alleles and noncanonical epitopes. Although these noncanonical epitopes may not be overtly abundant they do play a significant role in the antiviral $CD8^+$ T cell response.

The greatest number of epitopes defined in these antigens of interest were from EBNA3A which also displayed a degree of epitope clustering and overlapping. The RBPJκ binding activity region and the conserved region of all three EBNA3 proteins from amino acids 90-320 are the domains of epitope clustering. Epitope overlapping was observed between residues 126 and 145 with four epitopes contained in these 20 amino acids, in addition to residues 170 to 196 and 502 to 517 of EBNA 3A with 3 epitopes and 2 epitopes overlapping, respectively. Thus, these regions are clearly an important target of the T cell response.

In this study, the transcription and replication factor, EBNA1 was the antigen with the least $CD8^+$ T cell immune responses. It is likely that this low level of immune recognition by $CD8^+$ T cells is due to the GAr region. Almost the entire N-terminal region of EBNA1 is composed of the GAr that protects against degradation by the proteasome (Levitskaya, Shapiro et al. 1997). Of note, the EBNA1 $CD8^+$ T cell epitope ^{72}RPQKRPSCI^{80} is the only epitope located in the N-terminal half of the antigen and prior to the GAr region (Blake, Haigh et al. 2000). The remaining epitopes, including the novel EBNA1 epitope ^{499}DEGTWVAGVF^{508} from this study, are located in the C-terminal half of the antigen (Chapter 1, Table 1 and Chapter 3 Table 2). This concurs with the postulation of Capone and colleagues that EBNA1 contains two spatially distinct domains with the

N-terminal region hosting an immunotolerance function whilst the C-terminal region plays an antigenic role (Capone, Calabro et al. 2013).

Another possibility is that the weak immune recognition of EBNA1 is due to this antigen sharing areas of sequence with a number of self proteins including myelin basic protein (MBP), with cross-reactive antibodies described in multiple sclerosis (MS) patients (Bray, Luka et al. 1992), and the Ro 60 kDa protein with cross-reactive antibodies described in systemic lupus erythematosus (SLE) patients (McClain, Heinlen et al. 2005). It has recently been reported that EBNA1 shows significant peptide sharing with other human proteins. Almost ninety percent of this peptide sharing occurs in the GAr region along the N-terminus of this antigen, supporting the suggestion of molecular mimicry by EBNA1 playing a role in immunoevasion (Capone, Calabro et al. 2013). Thus, these high levels of sequence sharing between the host and EBNA1 may create a camouflage affect for the virus leading to tolerance and ultimately escaping immune recognition. Although current peptide elution/mass spectrometric technology is limited in its ability to detect EBV epitopes on LCLs, future studies could endeavour to determine if peptides from these homologous regions are presented by LCLs or if indeed these potentially reactive T cells have been removed from the repertoire. This low $CD8^+$ T cell response to EBNA1 may contribute to the immune escape of EBV-associated malignancies.

As many studies have used IM patient samples to find new epitopes, previously defined epitopes will differ significantly from our findings as it has been shown that large clonal expansions of $CD8^+$ T cells in acute IM do not persist following recovery (Callan, Steven et al. 1996). Indeed, previous work has demonstrated a noticeable shift from lytic to latent antigen-specific responses between the acute and chronic stages of EBV infection (Hislop, Annels et al. 2002, Woodberry, Suscovich et al. 2005) with some T cell specificities emerging whilst others are lost following resolution of IM (Hislop, Annels et al. 2002). Many of the novel epitopes defined in this study may arise from the difference in donor cohorts and account for some of the T cell specificities that emerge with long-term carriage of the virus. Furthermore, the reactivities towards CTL epitopes are likely to reflect the donor cohort and may be quite variable between studies depending on the individual's HLA type. Future studies could determine if novel epitopes found in this study elicit an immune response with IM samples.

Thus, this study has shown that epitopes of over 10 amino acids in length account for ~ 10% $CD8^+$ T cell epitopes as well as providing a significant list of novel epitopes from a number of immunogenic EBV antigens, identified from healthy virus carriers, restricted to various HLA-A, -B and -C alleles. In addition, the limitations of web-based prediction algorithms including HLA coverage and epitope length have been highlighted. Taken together these findings provide further understanding of the complex antiviral CTL immunity towards several highly immunogenic EBV

antigens and many new epitopes for further study and possible incorporation into peptide-based vaccines.

7.3 Determine if the highly immunogenic BZLF1 antigen of EBV includes novel $CD8^+$ T cell epitopes

The expression of BZLF1 in association with BRLF1 plays a critical role in EBV, initiating lytic cycle activation (Cox, Leahy et al. 1990). In addition to the seven highly immunogenic EBV antigens that were investigated in Chapter 3, the IE EBV antigen BZLF1 was closely examined in Chapter 4. T cells raised from *in vitro* stimulation of PBMCs from thirty-six healthy virus carriers with their autologous LCLs were screened with the 46 overlapping peptides (20-mers overlapping by 15aa) covering the entire length of BZLF1 based on the B95-8 strain of EBV. This study identified 11 novel $CD8^+$ T cell epitopes restricted through a large number of HLA alleles that were clustered within certain domains of the protein, with many overlapping sequences.

Importantly, this plethora of novel $CD8^+$ T cell epitopes are located in this antigen of only 245 residues, with almost half of these amino acids included in epitopes. The most significant clustering of overlapping epitopes was evident between residues 160 and 220, exquisitely focused from residues 188 to 200 with three epitopes that overlap the HLA-B8-restricted ^{190}RAKFKQLL^{197} epitope, first identified 20 years ago (Bogedain, Wolf et al. 1995). As our findings have shown ^{190}RAKFKQLL^{197} is located within an epitope cluster region of BZLF1 and is overlapped by the novel HLA-B7-restricted ^{188}KCRAKFKQL^{196}, HLA-Cw6-restricted ^{189}CRAKFKQLL^{197} and HLA-A30-restricted ^{192}KFKQLLQHY^{200} epitopes.

Previous work in healthy virus carriers by Benninger-Doring and colleagues analysed the frequency of $CD8^+$ T cells specific for peptides from BZLF1, BRLF1 and EBNA3A, restricted to either HLA-A3 or HLA-B8. Their findings showed that the BZLF1 HLA-B8-restricted ^{190}RAKFKQLL^{197} epitope was likely the dominant target for specific $CD8^+$ T cells in all $HLA\text{-}B8^+$ individuals tested. Responses to EBNA3A HLA-B8-restricted ^{325}FLRGRAYGL^{333} epitope were significantly lower and only seen in three individuals (Benninger-Doring, Pepperl et al. 1999). Future work incorporating tetramer labeling assays could determine if donors expressing HLA-B8 in conjunction with the HLA alleles restricted by the other cluster epitopes highlighted above (i.e. HLA-B8 and HLA-B7 and/or HLA-Cw6) maintain a dominant ^{190}RAKFKQLL^{197} immune response.

Sequence analysis of ^{190}RAKFKQLL^{197} -reactive clones from five $HLA\text{-}B8^+$ donors showed individuals responded with a unique or “private” TRBV repertoire which remained unfocused after long-term virus exposure. Importantly, specific $CD8^+$ T cells for this epitope have been shown to be present in individuals that undergo primary infection with either EBV type A or type B (Silins,

Cross et al. 1997). Given the strong immune response to this epitope and its conserved sequence in both type A and type B virus, ^{190}RAKFKQLL^{197} should be considered for inclusion in vaccine formulations aimed at reducing the viral load following primary EBV infection.

Luo and colleagues investigated sequence variation in BZLF1 in EBV-associated gastric carcinomas and NPC in Northern China and detected 17 mutations and a deletion. Interestingly, seven of these mutations and a deletion are located in epitopes. The most common of these mutations (73/89 samples) was the missense mutation at residue 195 resulting in Gln → His (Luo, Tang et al. 2011). This mutation is within the ^{189}CRAKFKQLL^{197} epitope and our assessment of T cell recognition of the variant ^{189}CRAKFK<u>H</u>LL^{197} indicated that this single sequence polymorphism had a significant impact with approximately 10 fold reduction in T cell recognition. It is of note, that this missense mutation is included within the four cluster epitopes restricted by either HLA-A30, -B7, -B8, or -Cw6.

This highly immunogenic antigen also appears to play a role in autoimmune disease. Gross *et al* have shown the aberrant expression of IE EBV antigen, BZLF1 in addition to LMP1 and LMP2A in the peripheral blood of SLE patients. Only a small fraction of cells expressed BZLF1, however, expression of BZLF1 has never been detected in the blood of healthy controls (Gross, Hochberg et al. 2005). Future work could delve into the immune responses of autoimmune disease patients and identify if there is a pronounced immune response to BZLF1 epitopes or if indeed the opposite is the case and sufferers of autoimmune conditions show a significantly decreased $CD8^+$ T cell response to this antigen. The present study has significantly expanded knowledge of the immune response to BZLF1 adding to the current pool of $CD8^+$ T cell epitopes that can be utilised in vaccine development and adoptive immunotherapy.

7.4 Determine if the peptide length preferences of class I human leukocyte antigens influence epitope selection in the EBV-specific T cell response

Fundamental to adaptive immunity, the HLA genes represent the greatest polymorphic region in the human genome culminating within the peptide-binding groove (Sette and Sidney 1998). This polymorphism gives rise to HLA molecules from different haplotypes with the ability to present an array of both self and microbial derived peptides (Doherty and Zinkernagel 1975). The amino acids accommodated within the six peptide-binding pockets, A to F, are distinct and determine the specificity of the peptide-HLA class I interaction (Clements, Dunstone et al. 2006). Generally HLA alleles have a preference for one or two amino acids at certain positions within the peptide. Based on this, peptide-binding motifs have been identified, designating primary anchors, with the greatest effect on binding and to a lesser extent on auxiliary or secondary anchors. In

many instances several HLA alleles share common motifs in addition to structure and are classified into groups or HLA supertypes accordingly (Doytchinova, Guan et al. 2004).

One such group is the HLA-B*44 supertype consisting of some of the most common HLA alleles (Cao, Hollenbach et al. 2001), including HLA-B*1801 and HLA-B*4403 which share a preference for peptides with Glutamic Acid (E) at position 2 and Phenylalanine (F) or Tyrosine (Y) at the C terminus (Sidney, Southwood et al. 2003, Hillen, Mester et al. 2008). In addition to the 11 $CD8^+$ T cell epitopes identified in the EBV IE BZLF1 antigen and described in Chapter 4, Chapter 5 focuses on a further two novel overlapping peptides from this antigen of different size that both conform to the HLA-B*44 supertype binding motif. These peptides are, the HLA-B*1801-restricted octamer ^{173}SELEIKRY^{180} and the HLA-B*4403-restricted noncanonical dodecamer ^{169}EECDSELEIKRY^{180}, which encompasses the octamer peptide.

Previous work of Hillen and colleagues has shown different length variants in the HLA-B*44 supertype with the nonamer AEFKEAFQL presented by HLA-B*4001 and the decamer AEFKEAFQLF presented by HLA-B*4402 (Hillen, Mester et al. 2008). Also described in their study was the peptide SEIEAKVRY from the TLN1 gene that encodes a cytoskeletal protein concentrated in areas of cell-substratum and cell to cell contacts. This peptide is presented by HLA-B*4402 (Hillen, Mester et al. 2008) and shows striking sequence homology to the HLA-B*1801-restricted SELEIKRY EBV epitope. It would be interesting to determine whether the SEIEAKVRY peptide is also recognised by HLA-B*1801-restricted T cells reactive with the octamer EBV peptide. As it has been proposed that a peptide that binds to one member of a HLA supertype could be capable of binding to other members of that supertype (Sidney, Southwood et al. 2003) it would be intriguing to assess whether healthy virus carriers with the remaining HLA allotypes of the HLA-B*44 supertype HLA-B37, -B40, -B41, -B45, -B47, -B49 or -B50 recognise either EECDSELEIKRY or SELEIKRY.

To establish an explanation for the variable peptide length preferences of these HLA-B*44 supertype molecules, crystal structures were solved for the viral peptide-HLA complexes. This revealed that the antigen-binding cleft of HLA-B*1801 was clearly suited to bind shorter length peptides, whilst HLA-B*4403 was more suited to presenting longer peptides with a bulged conformation. It is predicted that binding of HLA-B*4403 to SELEIKRY is likely to lead to steric clashes as the peptide would be forced to sit flat in the binding cleft and the depth of the HLA-B*4403 is shallow compared to the depth of the central region of HLA-B*1801. Differences have been observed in the antigen presentation pathway and structure between members of the HLA-B*44 supertype. HLA-B*4402 and HLA-B*4405, which differ by a single residue alteration at position 116 (Asp→Tyr) were shown to present unique peptides as a result of the change in the F pocket of the binding cleft (Zernich, Purcell et al. 2004). Although this work gives no indication as

to peptide length preference of these HLA alleles, another study into HLA-B*4402 has shown this allele to have an intolerance for presenting 8 aa peptides (Hillen, Mester et al. 2008) in concordance with HLA-B*4403.

Thermal stability analysis provided further proof of the epitope selection differences between HLA-B*1801 and HLA-B*4403. Data for each of the HLA-peptide complexes indicated an observable preference for HLA-B*1801 and the shorter peptide SELEIKRY and HLA-B*4403 and the longer peptide EECDSELEIKRY. Most notable was the instability of the HLA-B*4403-SELEIKRY complex. Furthermore, mass spectrometry data confirmed these findings showing the natural ligands of cells expressing HLA-B*1801 or HLA-B*4403. Results showed 18% of peptides presented by HLA-B*1801 were octamers compared to 0.8% eluted off HLA-B*4403. In contrast, longer length peptides > 10aa were dominantly presented by HLA-B*4403 with 18.5% compared to 2.3% by HLA-B*1801.

Our work has recently been cited by Motozono and colleagues whose findings also portray overlapping peptides eliciting CTL responses. Their study investigating the highly immunogenic viral protein, HIV-1 Nef demonstrates overlapping peptides of different lengths (8mer, 9mer and 11mer) presented by the same HLA-molecule, HLA-B*3501 (Motozono, Yokoyama et al. 2014). It was shown that the TCR reactivity toward these peptide length variants was either mutually exclusive toward the 11mer or the 8mer or cross-reactive toward the 9mer and 11mer. Of note, the 8mer VY8 (VPLRPMTY) is a dominant epitope in the early phase of HIV-1 infection, while the overlapping 11mer RY11 (RPQVPLRPMTY) becomes the dominant epitope in the chronic phase (Ueno, Idegami et al. 2007). Interestingly, the presenting allele, HLA-B*3501, a member of the HLA-B7 supertype, has previously been associated with recognition of noncanonical 11mer epitopes (Burrows, Rossjohn et al. 2006, Bell, Brennan et al. 2008) making this allele an ideal candidate for future studies into peptide length preferences.

To our knowledge the present report is the first to show peptide length preference among MHC molecules and provide a structural mechanism for the phenomenon. Future work needs to expand on this knowledge and study other class I MHC molecules to determine peptide length preferences and enhance the quality of predictive algorithms ensuring more accurate peptide prediction tools.

7.5 T cell cross-reactivity between an EBV epitope and an abundant self-peptide presented by HLA-B*1801^{+} cells

The fourth aim of characterising a cross-reactive T cell response to an HLA-B*1801-restricted epitope, derived from the BZLF1 antigen of EBV (SELEIKRY) and an HLA-B*1801-bound self-peptide, from the cleavage and polyadenylation specific factor 3-like protein

(DELEIKAY) was addressed in Chapter 6. This study has provided support for a molecular mimicry mechanism in autoimmunity by showing that EBV-reactive $CD8^+$ T cell expansions can be identified in healthy HLA-B*1801^+ individuals that can cross-react with a naturally presented self-peptide.

Thymic clonal deletion is the most efficient mechanism for the removal of clones that have a high affinity for self antigens, although this process cannot eliminate all self-reactive lymphocytes (Hogquist, Baldwin et al. 2005). Recently it has been suggested that a crippling of the $CD8^+$ T cell repertoire and domination by T cells with a high affinity to self antigens can result from interference in antigen processing in the cortical thymic epithelial cell-specific pathway that diversifies the T cell repertoire (Klein, Kyewski et al. 2014). Another recent study has shown that mechanisms exist that allow T cells with low affinity for self ligands to avoid clonal deletion and apoptosis (Hogquist and Jameson 2014). It has also been suggested that nonstimulatory self peptide-MHC complexes may assist the reactivity of T cells to foreign peptide-MHC antigen (Krogsgaard, Juang et al. 2007).

Klein and colleagues postulate the beneficial aspects of having T cells with a range of both high and low affinities for self antigens and the role these may play in the response to infection with pathogens. T cells with high affinity for self may provide initial, rapid response whilst T cells with low self affinity maintain a sustained response (Klein, Kyewski et al. 2014). These studies were undertaken using transgenic mice and it remains to be determined how these observations correlate with the adaptive immune response in humans. Interestingly, the three HLA-B*1801^+ donors with cross-reactive $CD8^+$ T cells to both the EBV SEL epitope and the self-peptide DEL are also the donors that displayed the strongest CTL responses to the SEL epitope. TCR repertoire analysis of the cross-reactive T cells depicted a polyclonal population of T cells from all three donors as many different Vβ genes were expressed by these subsets of T cells.

Molecular mimicry is the mechanism by which it is hypothesized that an immune response by the host mounted against a specific determinant of an infectious microbial agent may cross-react with a mimicked self sequence and lead to autoimmunity and/or disease (Oldstone 1987). T cell cross-reactivity to viral antigens of EBV and self-antigens has been shown for autoimmune diseases including multiple sclerosis (MS), systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA) (Niller, Wolf et al. 2008). Cross-reactivity has been shown previously between peptides from EBV antigen BALF5 (627-641) and MBP (85-99). Although these peptides have different sequences, there are structural similarities and these peptides are recognised by the same TCR. Interestingly, recognition is through two different MHC class II alleles, HLA-DRB1*1501 (the strongest genetic risk to MS) presenting the MBP peptide and HLA-DRB5*0101 the BALF5 peptide (Lang, Jacobsen et al. 2002). Structural similarities of charge and overall shape between SEL and DEL indicate that molecular mimicry is the underlying mechanism for the observed cross-

reactivity. Crystal structures revealed that HLA-B*1801 in complex with the EBV epitope was a structural mimic of the self-peptide.

Malissen and colleagues first identified in clones originating from a B10.A mouse, dual reactivity resulting from the cross-reactive recognition of an allogeneic MHC product by a self MHC-restricted αβ TCR (Malissen, Trucy et al. 1988). Transgenic mice studies have demonstrated that low-avidity $CD8^+$ T cells are largely spared from elimination by central and peripheral tolerance mechanisms. These low-avidity self-reactive $CD8^+$ T cells were shown to have the ability to make IFN-γ following brief *in vitro* restimulation with high concentrations of peptide and be mediators of autoimmunity in the system (Zehn and Bevan 2006). This study correlates with our findings that the self-peptide DEL was recognised by cross-reactive $CD8^+$ T cells albeit requiring a higher concentration than the EBV SEL epitope. Further studies have shown that challenge with lymphocytic choriomeningitis virus (LCMV) following insertion of the LCMV NP (nucleoprotein) or GP (glycoprotein) genes in the germline of transgenic mice leads to the generation of an antiviral response that is anti-self, leading to Insulin-dependent diabetes mellitus (IDDM) (Oldstone, Nerenberg et al. 1991).

It has been suggested that the autoimmune disease IDDM might be the culmination of a series of inflammatory "hits", that affect particular areas of the pancreas in a relapsing-remitting cycle, and viruses, including *Herpesviridae*, may play a role in this (Schneider and von Herrath 2013). IDDM has been associated with host genes, autoimmune responses and environmental triggers including viral infection with coxsackie virus B (Bach 2002) and cytomegalovirus (Richer and Horwitz 2008). In addition, viruses that induce cross-reactivity and/or molecular mimicry are likely to be involved in the induction of autoimmunity which may ultimately end in autoimmune disease e.g. IDDM (Schneider and von Herrath 2014). It is reasonable to propose, therefore, that these EBV/self cross-reactive T cell populations identified in this study may potentially play a role in autoimmune disease associated with HLA-B*1801 such as type 1 diabetes (Nejentsev, Howson et al. 2007).

7.6 Conclusions

This thesis reports the comprehensive study of eight highly immunogenic EBV antigens to assess the adaptive immune response to this virus. Aim one (Chapter 3) focused on the relative prevalence of peptide epitopes of over 10 amino acids in length in $CD8^+$ T cell recognition of EBV. The hypothesis was that indeed peptide epitopes of over 10 amino acids in length are abundant in antiviral $CD8^+$ T cell responses. What has been determined is that there are various factors which influence epitope selection including the virus's interaction with the host and the role that each

antigen plays in that. Five noncanonical epitopes were identified and shown to account for ~10% of $CD8^+$ T cell epitopes. This study has also provided an expansive list of novel EBV CTL epitopes, demonstrating a hierarchy of immunodominance between the EBV antigens, as well as highly immunogenic regions within several of the antigens including epitope clustering. The relative immunodominance hierarchy (with the exclusion of BZLF1), based on the identification of 28 latent and 13 lytic epitopes, including four noncanonical epitopes, was EBNA3A, EBNA3C, EBNA3B, BRLF1, BMLF1, BMRF1 and lastly EBNA1. The limited number of $CD8^+$ T cell epitopes identified in EBNA1 is likely a result of the GAr region and the immunoevasion quality this presents.

The focus of the second aim was the detailed investigation of the IE EBV antigen BZLF1 identified 11 novel $CD8^+$ T cell epitopes, many presented by common HLA alleles (Chapter 4). Interestingly, these epitopes have a clustered distribution along the length of this highly immunogenic antigen, with many overlapping sequences. As BZLF1 has shown itself to elicit a strong $CD8^+$ T cell response in many individuals, this antigen would make an ideal candidate in the development of an EBV vaccine. Although it is unlikely that any EBV vaccine would provide prevention from primary infection, a more realistic approach may focus on a reduction in viral load, which could in turn reduce the incidence of IM, Hodgkin's lymphoma, PTLD and possibly other EBV-associated malignancies. For immunotherapeutic purposes the EBV-BZLF1 peptide pool may be a potential candidate as it is available commercially in GMP-grade and we have shown that this diminutive antigen is highly immunogenic and induces immune responses through a wide range of HLA alleles. Thus, these results can be widely used in future studies focused on manipulating the immune system and targeting EBV and its associated diseases.

The third objective (Chapter 5) was to determine if the peptide length preference of class I HLA influences epitope selection in the EBV-specific T cell response. In addition to the 11 novel BZLF1 $CD8^+$ T cell epitopes identified in Chapter 4, a further two novel epitopes were identified, the HLA-B*1801-restricted octamer ^{173}SELEIKRY^{180} and the HLA-B*4403-restricted noncanonical dodecamer ^{169}EECDSELEIKRY^{180}, which encompasses the octamer epitope. The common alleles HLA-B*1801 and HLA-B*4403 are members of the HLA-B*44 supertype which share a preference for peptides with Glu at position 2 and Phe or Tyr at the C terminus (Sidney, Southwood et al. 2003, Hillen, Mester et al. 2008). Findings of this study indicate that class I MHC polymorphism influences not only amino acid preferences at certain positions within the peptide, but in addition, a preference for particular lengths of presented peptides. Structural studies of these viral pMHC complexes depict the binding clefts of these two HLA-B*44 supertype members favouring peptides of different lengths. This was evident with the more "shallow" groove of HLA-B*4403 leading to steric clashes with octamer peptides, whilst the deep groove of HLA-B*1801 is

perfectly suited to these shorter epitopes. The peptide length preference was further confirmed with the results of mass spectrometry analysis showing a bias of HLA-B*1801 presenting peptides <10 aa, whereas HLA-B*4403 more frequently presented peptides >10 aa. In summary, this study highlights a novel mechanism through which MHC class I polymorphism can diversify the immune response, influencing epitope selection in an antigen specific T cell response. Future studies in this area are required to determine the peptide length preference of other HLA alleles which may be used to improve web-based algorithms for epitope prediction, in addition to assisting strategies for peptide-based vaccine design.

The final objective was to characterise a cross-reactive T cell response to an HLA-B*1801-bound EBV epitope and a self-peptide (Chapter 6). The hypothesis was that some healthy individuals carry T cells that cross-react with these two closely related peptides and that the T cell receptor repertoire that cross-reacts with an EBV epitope and a self-peptide is similar between individuals. The data confirmed that a significant minority of individuals carry T cells that can cross-recognise the viral and self-peptide but unexpectedly, a diverse array of T cell receptors was expressed by these cross-reactive T cells. 11 HLA-B*1801^{+} healthy virus carriers were screened for cross-reactive responses to the HLA-B*1801-restricted epitope, SEL and HLA-B*1801-restricted self-peptide, DEL, and interestingly, only 3 of these individuals were shown to elicit a CTL immune response to both SEL and DEL (all donors recognised SEL). The self-peptide DEL has been shown to share substantial sequence and structural homology with the HLA-B*1801-bound EBV peptide. The self-peptide was only recognised *in vitro* at relatively high concentrations, which explains how these auto-reactive T cells have evaded tolerance. To our knowledge, this report is the first showing the natural presentation on human cells of a self-peptide cross-recognized by EBV-reactive CD8^{+} T cells and presented by a self-HLA antigen. In future studies it will be important to establish if this cross-reactive T cell population plays a role in patients with IDDM, given the association of HLA-B*1801 with this autoimmune disease (Nejentsev, Howson et al. 2007).

It has been proposed that a peptide that binds to one member of a HLA supertype should, reasonably be capable of binding to other members of that supertype (Sidney, Southwood et al. 2003), and therefore epitopes that fit into this category would make ideal vaccine and/ or immunotherapy candidates. A number of epitopes that are presented by more than one HLA antigen have been identified in this study including BRLF1 ^{101}IACPIVMRY^{109} presented by HLA-A*24 and HLA-A*29, members of the HLA-A*24 supertype, and BMRF1 ^{286}LPLDLSVI/LF^{295} presented by HLA-B*35 and HLA-B*53, members of the HLA-B*07 supertype, and BZLF1 ^{209}SENDRLRLL^{217} presented by the HLA-B*44 supertype members HLA-B*60 and HLA-B*49. This study has contributed a significant list of novel CD8^{+} T cell EBV epitopes as well as

highlighting a novel mechanism through which MHC class I polymorphism can further diversify immune responses. Together the findings of this study will impact EBV vaccine and immunotherapy development, improve web-based algorithms for epitope prediction, and highlights a potential mechanism by which EBV infection could trigger autoimmune disease.

7.7 References List

Aarnoudse, C. A., et al. (1999). "Interleukin-2-induced, melanoma-specific T cells recognize CAMEL, an unexpected translation product of LAGE-1." Int. J. Cancer. **82**: 442-448.

Abbott, R. J. M., et al. (2013). "$CD8^+$ T Cell Responses to Lytic EBV Infection: Late Antigen Specificities as Subdominant Components of the Total Response." The Journal of Immunology **191**: 5398-5409.

Adams, E. J. and A. M. Luoma (2013). "The adaptable major histocompatability complex (MHC) fold: structure and function of nonclassical and MHC class I-like molecules." Annu. Rev. Immunol. **31**: 529-561.

Argaet, V. P., et al. (1994). "Dominant selection of an invariant T cell antigen receptor in response to persistent infection by Epstein-Barr virus." The Journal of Experimental Medicine **180**: 2335-2340.

Arstila, T. P., et al. (1999). "A Direct Estimate of the Human αβ T Cell Receptor Diversity." Science **286**(5441): 958-961.

Attaf, M., et al. (2015). "αβ T cell receptors as predictors of health and disease." Cellular and Molecular Immunology.

Babbe, H., et al. (2000). "Clonal Expansions of $CD8^+$T Cells Dominate the T Cell Infiltrate in Active Multiple Sclerosis Lesions as Shown by Micromanipulation and Single Cell Polymerase Chain Reaction." The Journal of Experimental Medicine **192**(3): 393-404.

Bach, J. F. (2002). "Immunotherapy of Type 1 diabetes: lessons for other autoimmune diseases." Arthritis Res **4**(Suppl 3): S3-S15.

Banerjee, S., et al. (2013). "The EBV latent antigen 3C inhibits apoptosis through targeted regulation og interferon regulatory factors 4 and 8. ." PLOS Pathogens **9**: e1003314.

Bell, M. J., et al. (2008). "Widespread Sequence Variation in Epstein-Barr Virus Nuclear Antigen 1 Influences the Antiviral T Cell Response." The Journal of Infectious Diseases **197**: 1594-1597.

Benninger-Doring, G., et al. (1999). "Frequency of $CD8^+$ T Lymphocytes Specific for Lytic and Latent Antigens of Epstein-Barr Virus in Healthy Virus Carriers." Virology **264**: 289-297.

Bihl, F., et al. (2006). "Impact of HLA-B Alleles, Epitope Binding Affinity, Functional Avidity, and Viral Coinfection on the Immunodominance of Virus-Specific CTL Responses." The Journal of Immunology **176**: 4094-4101.

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

Blake, N., et al. (2000). "The Importance of Exogenous Antigen in Priming the Human $CD8^+$ T Cell Response: Lessons from the EBV Nuclear Antigen EBNA1." The Journal of Immunology **165**: 7078-7087.

Blake, N., et al. (1997). "Human $CD8^+$ T Cell Responses to EBV EBNA1: HLA Class I Presentation of the (Gly-Ala)-Containing Protein Requires Exogenous Processing." Immunity **7**: 791-802.

Bogedain, C., et al. (1995). "Specific cytotoxic T-lymphocytes recognize the immediate-early transactivator ZTA of Epstein-Barr virus." Journal of Virology **69**: 4872-4879.

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

Borgulya, P., et al. (1991). "Development of the CD4 and CD8 lineage of T cells: instruction versus selection." The EMBO Journal **10**(4): 913-918.

Bray, P. F., et al. (1992). "Antibodies against Epstein-Barr nuclear antigen (EBNA) in multiple sclerosis CSF, and two pentapeptide sequence identities between EBNA and myelin basic protein." Neurology **42**: 1798-1804.

Brooks, J., et al. (1993). "Different HLA-B27 subtypes present the same immunodominant Epstein-Barr virus peptide." The Journal of Experimental Medicine **178**: 879-887.

Bryant, P. and H. Ploegh (2004). "Class II MHC peptide loading by the professionals." Current Opinion in Immunology **16**: 96-102.

Burrows, J. M., et al. (2008). "Preferential binding of unusually long peptides to MHC class I and its influence on the selection of target peptides for T cell recognition." Molecular Immunology **45**: 1818-1824.

Burrows, S. R., et al. (2003). "Promiscous CTL recognition of viral epitopes on multiple human leukocyte antigens: biological validation of the proposed HLA A24 supertype. ." The Journal of Immunology **171**(3): 1407-1412.

Burrows, S. R., et al. (1994). "Five new cytotoxic T cell epitopes identified within Epstein-Barr virus nuclear antigen 3." Journal of General Virology **75**: 2489-2493.

Burrows, S. R., et al. (1990). "An Epstein-Barr virus-specific cytotoxic T cell epitope present on A- and B-type transformants." Journal of Virology **64**: 3974-3976.

Burrows, S. R., et al. (2006). "Have we cut ourselves too short in mapping CTL epitopes?" TRENDS in Immunology **27**(1): 11-16.

Burrows, S. R., et al. (1990). "An Epstein-Barr Virus-Specific Cytotoxic T Cell Epitope in EBV Nuclear Antigen 3 (EBNA3)." The Journal of Experimental Medicine **171**: 345-349.

Burrows, S. R., et al. (1995). "T Cell Receptor Repertoire for a Viral Epitope in Humans Is Diversified by Tolerance to a Background Major Histocompatibility Complex Antigen." The Journal of Experimental Medicine **182**(6): 1703-1715.

Callan, M. F. C. (2003). "The Evolution of Antigen-Specific $CD8^+$ T Cell Responses after Natural Primary Infection of Humans with Epstein-Barr Virus." Viral Immunology **16**(1): 3-16.

Callan, M. F. C., et al. (1998). "T cell selection during the evolution of $CD8^+$ T cell memory in vivo." Eur. J. Immunol **28**: 4382-4390.

Callan, M. F. C., et al. (1996). "Large clonal expansions of $CD8^+$T cells in acute infectious mononucleosis." Nature Medicine **2**(8): 906-911.

Cao, K., et al. (2001). "Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution pattern in these populations." Hum. Immunol. **62**: 1009-1030.

Capone, G., et al. (2013). "Peptide matching between Epstein-Barr virus and human proteins." Pathogens and Disease **69**(3): 205-212.

Charini, W. A., et al. (2001). "Clonally diverse CTL response to a dominant viral eptope recognizes potential epitope variants." The Journal of Immunology **167**: 4996-5003.

Clements, C. S., et al. (2006). "Specificity on a knife edge: the alphabeta T cell receptor." Current Opinion in Structural Biology **16**: 787-795.

Cornberg, M., et al. (2006). "Narrowed TCR repertoire and viral escape as a consequence of heterologous immunity." The Journal of Clinical Investigation **116**(5): 1443-1456.

Cox, M. A., et al. (1990). "An enhancer within the divergent promoter of Epstein-Barr virus responds synergistically to the R and Z transactivators." Journal of Virology **64**(1): 313-321.

Cresswell, P., et al. (2005). "Mechanisms of MHC class I-restricted antigen processing and cross-presentation." Immunological Reviews **207**: 145-157.

Crotzer, V. L., et al. (2000). "Immunodominance among EBV-derived epitopes restricted by HLA-B27 does not correlate with epitope abundance in EBV-transformed B-lymphoblastoid cell lines." The Journal of Immunology **164**(12): 6120-6129.

Davis, M. M. and P. J. Bjorkman (1988). "T-cell antigen receptor genes and T-cell recognition. ." Nature **334**: 395-402.

Davison, A. J., et al. (2009). "The order Herpesvirales." Arch Virol **154**: 171-177.

De Rosa, S. C., et al. (2004). "Ontogeny of γδ T Cells in Humans. ." The Journal of Immunology **172**: 1637-1645.

Deng, Y., et al. (1997). "MHC Affinity, Peptide Liberation, T Cell Repertoire, and Immunodominance All Contribute to the Paucity of MHC Class I-Restricted Peptides Recognized by Antiviral CTL." The Journal of Immunology **158**: 1507-1515.

Doherty, P. C., et al. (1978). "Cytotoxic T-cell responses in mice infected with influenza and vaccinia viruses vary in magnitude with H-2 genotype." The Journal of Experimental Medicine **148**: 534-543.

Doherty, P. C. and R. M. Zinkernagel (1975). "A biological role for the major histocompatibility antigens." Lancet **1**: 1406-1409.

Doytchinova, I. A., et al. (2004). "Identifying human MHC supertypes using bioinformatic methods." J Immunol **172**: 4314-4323.

Eisenlohr, L. C., et al. (1992). "Flanking Sequences Influence the Presentation of an Endogenously Synthesized Peptide to Cytotoxic T Lymphocytes." The Journal of Experimental Medicine **175**: 481-487.

Epstein, M. A., et al. (1964). "Virus particles in cultured lymphoblasts from Burkitt's Lymphoma." Lancet **283**(7335): 702-703.

Falk, K., et al. (1993). "Peptide motifs of HLA-B35 and -B37 molecules." Immunogenetics **38**: 161-162.

Falk, K., et al. (1991). "Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules." Nature **351**: 290-296.

Falk, K., et al. (1995). "Peptide motifs of HLA-B58, B60, B61 and B62 molecules." Immunogenetics **41**: 165-168.

Feederle, R., et al. (2000). "The Epstein-Barr virus lytic program is controlled by the co-operative functions of two transactivators." The EMBO Journal **19**(12): 3080-3089.

Fogg, M. H., et al. (2005). "The CD8+ T-cell response to an Epstein-Barr virus-related gammaherpesvirus infecting rhesus macaques provides evidence for immune evasion by the EBNA-1 homologue." Journal of Virology **79**: 12681-12691.

Fohse, L., et al. (2011). "High TCR diversity ensures optimal function and homeostasis of Foxp3^{+} regulatory T cells." Eur. J. Immunol **41**: 3101-3113.

Fowlkes, B. J. and E. Schweighoffer (1995). "Positive selection of T cells." Curr. Opin. Immunol. **7**: 188-195.

Frahm, N., et al. (2007). "Extensive HLA class I allele promiscuity among viral CTL epitopes." European Journal of Immunology **37**: 2419-2433.

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

Gavioli, R., et al. (1993). "Multiple HLA A11-restricted cytotoxic T-lymphocyte epitopes of different immunogencities in the Epstein-Barr virus-encoded nuclear antigen 4." Journal of Virology **67**(3): 1572-1578.

Goldrath, A. W. and M. J. Bevan (1999). "Selecting and maintaining a diverse T-cell repertoire." Nature **402**: 255-262.

Gras, S., et al. (2008). "T-cell receptor bias and immunity." Current Opinion in Immunology **20**: 119-125.

Green, K. J., et al. (2004). "Potent T Cell response to a class I-binding 13-mer viral epitope and the influence of HLA micropolymorphism in controlling epitope length." European Journal of Immunology **34**: 2510-2519.

Gross, A. J., et al. (2005). "EBV and Systemic Lupus Erythematosus: A New Perspective." The Journal of Immunology **174**: 6599-6607.

Guidos, C. J. (1996). "Positive selection of $CD4^+$ and $CD8^+$ T cells." Curr. Opin. Immunol. **8**(2): 225-232.

Hansen, S. G., et al. (2013). "Cytomegalovirus Vectors Violate $CD8^+$T Cell Epitope Recognition Paradigms." Science **340**: 1237874.

Henle, G., et al. (1968). "Relation of Burkitt's Tumor-Associated Herpes-Type Virus to Infectious Mononucleosis." Proceedings of National Academy Science **59**: 94-101.

Hill, A., et al. (1995). "Characterization of two Epstein-Barr virus epitopes restricted by HLA-B7*." European Journal of Immunology **25**: 18-24.

Hill, A. B., et al. (1995). "Class I Major Histocompatibility Complex-restricted Cyotoxic T Lymphocytes Specific for Epstein-Barr Virus (EBV) Nuclear Antigens Fail to Lyse the EBV-transformed B lymphoblastoid Cell Lines against Which They Were Raised." The Journal of Experimental Medicine **181**: 2221-2228.

Hillen, N., et al. (2008). "Essential diffferences in ligand presentation and T cell epitope recognition among HLA molecules of the HLA-B44 supertype." European Journal of Immunology **38**(11): 2993-3003.

Hislop, A. D., et al. (2002). "Epitope-specific Evolution of Human $CD8^+$ T cell Responses from Primary to Persistent Phases of Epstein-Barr Virus Infection " The Journal of Experimental Medicine **195**(7): 893-905.

Hislop, A. D., et al. (2005). "Tonsillar homing of Epstein-Barr virus-specific CD8+ T cells and the virus-host balance." J. Clin. Investig. **115**: 546-555.

Hislop, A. D., et al. (2007). "Cellular Responses to Viral Infection in Human: Lessons from Epstein-Barr Virus." Annual Reviews Immunology **25**: 587-617.

Hogquist, K. A., et al. (2005). "Central tolerance: learning self-control in the thymus." Nat. Rev. Immunol. **5**: 772-782.

Hogquist, K. A. and S. C. Jameson (2014). "The self-obsession of T cells: how TCR signaling thresholds affect fate "decisions" and effector function." Nature Immunology **15**(9): 815-823.

Horst, D., et al. (2011). "Viral evasion of T cell immunity: ancient mechanisms offering new applications." Current Opinion in Immunology **23**: 96-103.

Janeway Jnr, C. A. (1989). "Approaching the asymptote? Evolution and revolution in immunology." Cold Spring Harbor Symposia on Quantitative Biology **54**: 1-13.

Janeway Jnr, C. A., et al. (2001). Immunobiology: The immune system in health and disease. New York, Garland.

Kalams, S. A., et al. (1994). "Longitudinal analysis of T cell receptor (TCR) gene usage by human immundeficiency virus 1 envelope-specific cytotoxic T lymphocyte clones reveals a limited TCR repertoire." The Journal of Experimental Medicine **179**(4): 1261-1271.

Kawakami, Y., et al. (2001). "Isolation of a new melanoma antigen, MART-2, containing a mutated epitope recognized by autologous tumor-infiltrating T lymphocytes." The Journal of Immunology **166**: 2871-2877.

Kawashima, D., et al. (2013). "Nuclear Transport of Epstein-Barr Virus DNA Polymerase Is Dependent on the BMRF1 Polymerase Processivity Factor and Molecular Chaperone Hsp90." Journal of Virology **87**(11): 6482-6491.

Kenney, S. C., et al. (1992). "The cellular oncogene c-myb can interact synergistically with the Epstein-Barr virus BZLF1 transactivator in lymphoid cells." Mol. Cell. Biol **12**(1): 136-146.

Kerr, B. M., et al. (1996). "Identification of type B-specific and cross-reactive cytotoxic T-lymphocyte responses to Epstein-Barr virus." Journal of Virology **70**(12): 8858-8864.

Khanna, R., et al. (1992). "Localization of Epstein-Barr Virus Cytotoxic T Cell Epitopes Using Recombinant Vaccinia: Implications for Vaccine Development." The Journal of Experimental Medicine **176**: 169-176.

Kieff, E. D. and A. B. Rickinson (2007). Epstein-Barr Virus and Its Replication. Fields Virology (5th ed). D. M. Knipe and P. M. Howley. Philadelphia, Lippincott, Williams & Wilkins. **2:** 2603-2654.

Kjer-Nielsen, L., et al. (2002). "The 1.5 A crystal structure of a highly selected antiviral T cell receptor provides evidence for a structural basis of immunodominance." Structure **10**: 1521-1532.

Kjer-Nielsen, L., et al. (2003). "A Structural Basis for the Selection of Dominant αβ T Cell Receptors in Antiviral Immunity." Immunity **18**(1): 53-64.

Klein, L., et al. (2014). "Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see)." Nature Reviews Immunology **14**: 377-391.

Kloetzel, P. M. (2001). "Antigen Processing by the Proteasome." Nature Reviews Molecular Cell Biology **2**: 179-187.

Kotturi, M., et al. (2008). "Naive precursor frequencies and MHC binding rather than the degree of epitope diversity shape $CD8^+$ T cell immunodominance." The Journal of Immunology **181**(3): 2124-2133.

Kovacsovics-Bankowswki, M. and K. L. Rock (1995). "A phagosome-to-Cytosol Pathway for Exogenous Antigens Presented on MHC Class I Moleciles." Science **267**: 243-245.

Krogsgaard, M., et al. (2007). "A role for "self" in T-cell activation " Semin. Immunol **19**: 236-244.

Kutok, J. L. and F. Wang (2006). "Spectrum of Epstein-Barr Virus-Associated Diseases." The Annual Review of Pathology: Mechanisms of Disease. **1**: 375-404.

La Gruta, N. L., et al. (2006). "A virus-specific $CD8^+$ T cell immunodominance hierarchy determined by antigen dose and precursor frequencies." Proceedings of the National Academy of Sciences **103**(4): 994-999.

La Gruta, N. L., et al. (2010). "Primary CTL respose magnitude in mice is determined by the extent of naive T cell recruitment and subsequent clonal expansion." The Journal of Clinical Investigation **120**(6): 1885-1894.

Lang, H. L. E., et al. (2002). "A functional and structural basis for TCR cross-reactivity in multiple sclerosis." Nature Immunology **3**(10): 940-943.

Le Roux, A., et al. (1994). "The Epstein-Barr virus determined nuclear antigens EBNA-3A, -3B, and -3C repress EBNA-2 mediated transactivation of the viral terminal protein 1gene promoter." Virology **205**: 596-602.

Lee, S. P., et al. (2000). "CTL Control of EBV in Nasopharyngeal Carcinoma (NPC): EBV-Specific CTL Responses in the Blood and Tumours of NPC Patients and the Antigen-Processing Function of the Tumour Cells." The Journal of Immunology **165**: 573-582.

Leen, A., et al. (2001). "Differential immunogenicity of Epstein-Barr virus latent-cycle proteins for human $CD4^+$ T-helper 1 responses." Journal of Virology **75**(18): 8649-8659.

Lehner, P., et al. (1995). "Human HLA-A0201-restricted cytotoxic T lymphocyte recognition of influenza A is dominated by T cells bearing the V beta 17 gene segment." The Journal of Experimental Medicine **181**(79-81).

Levitskaya, J., et al. (1997). "Inhibition of ubiquitin/proteasome-dependent protein degradation by the Gly-Ala repeat domain of the Epstein-Barr virus nuclear antigen 1." Proceedings of National Academy Science **94**: 12616-12621.

Li, H., et al. (2012). "Determinants of public T cell responses." Cell. Res. **22**: 33-42.

Li, P., et al. (2005). "Compartmentalization of class II antigen presentation: contribution of cytoplasmic and endosomal processing " Immunological Reviews **207**: 206-217.

Lieberman, P. M., et al. (1986). "Promiscuous *trans* activation of gene expression by an Epstein-Barr virus-encoded early nuclear protein." Journal of Virology **60**: 140-148.

Lim, A., et al. (2000). "Frequent Contribution of T Cell Clonotypes with Public TCR Features to the Chronic Response Against a Dominant EBV-Derived Epitope: Application to Direct Detection of Their Molecular Imprint on the Human Peripheral T Cell Repertoire." The Journal of Immunology **165**(4): 2001-2011.

Lin, J., et al. (2002). "Epstein-Barr virus nuclear antigen 3C putative repression domain mediates coactivation of the LMP1 promoter with EBNA-2." Journal of Virology **76**: 232-242.

Long, H. M., et al. (2011). "Immune defence against EBV and EBV-associated disease." Current Opinion in Immunology **23**: 258-264.

Luo, B., et al. (2011). "Sequence variation of Epstein-Barr virus (EBV) BZLF1 gene in EBV-associated gastric carcinomas and nasophayngeal carcinomas in Northern China." Microbes and Infection **13**: 776-782.

Maier, R., et al. (1994). "Peptide motifs of HLA-A3, -A24, and -B7 molecules as determined by pool sequencing." Immunogenetics **40**: 306-308.

Malissen, M., et al. (1988). "A T Cell Clone Expresses Two T Cell Receptor α Genes but Uses One αβ Heterodimer for Allorecognition and Self MHC-Restricted Antigen Recognition." Cell **55**: 49-59.

Manet, E., et al. (1991). "Domains of the Epstein-Barr virus (EBV) transcription factor R required for dimerization, DNA binding and activation." Nucleic Acid Res **19**(10): 2661-2667.

Marescotti, D., et al. (2009). "Characterization of an human leucocyte antigen A2-restricted Epstein-Barr virus nuclear antigen-1-derived cyotoxic T-lymphocyte epitope." Immunology **129**: 386-395.

Maruo, S., et al. (2006). "Epstein-Barr virus nuclear protein EBNA3C is required for cell cycle progression and growth maintenance of lymphoblastoid cells." Proceedings of National Academy Science **103**(51): 19500-19505.

Mazza, C. and B. Malissen (2007). "What guides MHC-restricted TCR recognition?" Seminars in Immunology **19**: 225-235.

McClain, M. T., et al. (2005). "Early events in lupus humoral autoimmunity suggest initiation through molecular mimicry." Nature Medicine **11**: 85-89.

Medzhitov, R. and C. A. Janeway Jr (1997). "Innate immunity: impact on the adaptive immune response." Current Opinion in Immunology **9**: 4-9.

Medzhitov, R. and C. A. Janeway Jr (2000). "Innate Immunity." The New England Journal of Medicine **343**(5): 338-344.

Mellman, I. and R. M. Steinman (2001). "Dendritic cells: specialized and regulated antigen processing machines." Cell **106**: 255-258.

Messaoudi, I., et al. (2002). "Direct link between mhc polymorphism, T cell avidity, and diversity in immune difense." Science **298**: 1797-1800.

Miles, J. J., et al. (2006). "TCRα Genes Direct MHC Restriction in the Potent Human T Cell Response to a Class I-Bound Viral Epitope." The Journal of Immunology **177**: 6804-6814.

Miles, J. J., et al. (2005). "CTL Recognition of a Bulged Viral Peptide Involves Biased TCR Selection." The Journal of Immunology **175**(3826-3834).

Miles, J. J., et al. (2005). "T-cell grit: large clonal expansions of virus-specific $CD8^+$ Tcells can dominate in the peripheral circulation for at least 18 years." Blood **106**: 4412-4413.

Momburg, F., et al. (1994). "Peptide Size Selection by the Major Histocompatibility Complex-encoded Peptide Transporter." The Journal of Experimental Medicine **179**: 1613-1623.

Moon, J., et al. (2007). "Naive $CD4(^+)$ T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude." Immunity **27**(2): 203-213.

Morgan, S., et al. (1996). "A recombinant adenovirus expressing an Epstein-Barr virus (EBV) target antigen can selectively reactivate rare components of EBV cytotoxic T-lymphocyte memory in vitro." Journal of Virology **70**: 2394-2402.

Moss, P., et al. (1991). "Extensive conservation of alpha and beta chains of the human T-cell antigen receptor recognising HLA-A2 and influenza A matrix peptide." Proceedings of National Academy Science **88**: 8987-8990.

Motozono, C., et al. (2014). "Cross-reactivity analysis of T cell receptors specific for overlapping HIV-1 Nef epitopes of different lengths." Microbes and Infection **16**: 320-327.

Murray, R., et al. (1992). "Identification of the target antigens for the human cytotoxic T cell response to Epstein-Barr virus (EBV): Implications for the immune control of EBV-positive malignancies." The Journal of Experimental Medicine **176**: 157-168.

Mylin, L. M., et al. (1995). "Hierarchy among multiple H-2^{b}-restricted cytotoxic T-lymphocyte epitopes within simian virus 40 T antigen." Journal of Virology **69**(11): 6665-6677.

Nejentsev, S., et al. (2007). "Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A." Nature **450**: 887-892.

Neuhierl, B. and H.-J. Delecluse (2006). "The Epstein-Barr Virus BMRF1 Gene Is Essential for Lytic Virus Replication " Journal of Virology **80**(10): 5078-5081.

Niller, H. H., et al. (2008). "Regulation and dysregulation of Epstein-Barr virus latency: Implications for the development of autoimmune diseases." Autoimmunity **41**(4): 298-328.

Ning, R. J., et al. (2011). "Long-term carriers generate Epstein-Barr virus (EBV)-specific CD4+ and CD8+ polyfunctional T-cell responses which show immunodominance hierarchies of EBV proteins." Immunology **134**: 161-171.

Nussbaum, A. K., et al. (2003). "Using the World Wide Web for predicting CTL epitopes. ." Current Opinion in Immunology **15**: 69-74.

Obar, J., et al. (2008). "Endogenous naive CD8^{+} T cell precursor frequency regulates primary and memory responses to infection. ." Immunity **28**(6): 859-869.

Oldstone, M. B. (1987). "Molecular mimicry and autoimmune disease." Cell **50**: 819-820.

Oldstone, M. B., et al. (1991). "Virus Infection Triggers Insulin-Dependent Diabetes Mellitus in a Transgenic Model: Role of Anti-Self (Virus) Immune Response." Cell **65**: 319-331.

Pamer, E. and P. Cresswell (1998). "Mechanisms of MHC Class I-Restricted Antigen Processing." Annual Reviews Immunology **16** 323-358.

Parham, P., et al. (1995). "The origins of HLA-A, B, C polymorphism." Immunol. Rev. **143**: 141-180.

Pearson, G. R., et al. (1987). "Identification of an Epstein-Barr virus early gene encoding a second component of the restricted early antigen complex." Virology **160**: 151-161.

Pellet, P. and B. Roizman (2006). The *Herpesviridae*: a brief introduction. Fields Virology, 5th edn. D. M. Knipe, P. M. Howley, D. E. Griffin et al. Philadelphia, Lippincott, Williams & Wilkins**:** 2479-2499.

Pepperl, S., et al. (1998). "Immediate-Early Transactivator Rta of Epstein-Barr Virus (EBV) Shows Multiple Epitopes Recognized by EBV-Specific Cytotoxic T Lymphocytes." Journal of Virology **72**(11): 8644-8649.

Probst-Kepper, M., et al. (2004). "Conformational Restraints and Flexibility of 14-Meric Peptides in Complex with HLA-B*3501." The Journal of Immunology **173**: 5610-5616.

Probst-Kepper, M., et al. (2001). "An alternative open reading frame of the human macrophage colony-stimulating factor gene is independently translated and codes for an antigenic peptide 14 amino acids recognized by tumor-infiltrating CD8 T lymphocytes. ." The Journal of Experimental Medicine **193**: 1189-1198.

Pudney, V. A., et al. (2005). "CD8 $^{+}$ immunodominance among Epstein-Barr virus lytic cycle antigens directly reflects the efficiency of antigen presentation in lytically infected cells " The Journal of Experimental Medicine **201**(3): 349-360.

Ragoczy, T., et al. (1998). "The Epstein-Barr virus Rta protein activates lytic cycle genes and can disrupt latency in B lymphocytes." Journal of Virology **72**: 7978-7984.

Rammensee, H.-G., et al. (1999). "SYFPEITHI: database for MHC ligands and peptide motifs." Immunogenetics **50**: 213-219.

Rammensee, H. G., et al. (1993). "Peptides naturally presented by MHC class I molecules." Annu. Rev. Immunol. **11**: 213-244.

Richer, M. J. and M. S. Horwitz (2008). "Viral infections in the pathogenesis of autoimmune diseases: focus on type 1 diabetes." Front Biosci **13**: 4241-4257.

Rickinson, A. and E. Kieff (2007). Epstein-Barr Virus. Fields Virology. D. M. Knipe and P. M. Howley. Philadelphia, Lippincott, Williams and Wilkins. **2:** 2655-2700.

Rickinson, A. B. and D. J. Moss (1997). "Human Cytotoxic T Lymphocyte Responses to Epstein-Barr Virus Infection." Annual Reviews Immunology **15**: 405-431.

Rist, M. J., et al. (2015). "T Cell Cross-Reactivity between a Highly Immunogenic EBV Epitope and a Self-Peptide Naturally Presented by HLA-B*18:01^{+} Cells." The Journal of Immunology **194**: 4668-4675.

Rist, M. J., et al. (2013). "HLA Peptide Length Preferences Control CD8^{+}T Cell Responses." The Journal of Immunology **191**: 561-571.

Robertson, E. E., et al. (1996). "The Amino-Terminal Domains of Epstein-Barr Virus Nuclear Proteins 3A,3B and 3C Interact with RBPJκ." Journal of Virology **70**(5): 3068-3074.

Rock, K. L. and A. L. Goldberg (1999). "Degradation of cell proteins and the generation of MHC class I-presented peptides." Annu. Rev. Immunol. **17**: 739-779.

Rock, K. L., et al. (1994). "Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules." Cell **78**(5): 761-771.

Rock, K. L. and L. Shen (2005). "Cross-presentation: underlying mechanisms and role in immune surveillance " Immunological Reviews **207**: 166-183.

Rossjohn, J., et al. (2015). "T Cell Antigen Receptor Recognition of Antigen-Presenting Molecules." Annu. Rev. Immunol. **33**: 169-200.

Rowe, M. and J. Zuo (2010). "Immune responses to Epstein-Barr virus: molecular interactions in the virus evasion of $CD8^+$T cell immunity." Microbes and Infection **12**: 173-181.

Rudolph, M. G., et al. (2006). "How TCRs Bind MHCs, Peptides, and Coreceptors." The Annual Review of Immunology **24**: 419-466.

Rudolph, M. G. and I. A. Wilson (2002). "The specificity of TCR/pMHC interaction." Current Opinion in Immunology **14**: 52-65.

Saulquin, X., et al. (2000). "A global appraisal of immunodominant CD8 T cell responses to Epstein-Barr virus and cytomegalovirus by bulk screening." European Journal of Immunology **30**(9): 2531-2539.

Schirmbeck, R., et al. (1994). "Immunization with soluble hepatitis B virus surface protein elicits murine H-2 class I-restricted CD8+ cytotoxic T lymphocyte responses *in vivo*." The Journal of Immunology **152**: 1110-1119.

Schneider, D. A. and M. G. von Herrath (2013). "Viruses and Type 1 diabetes: a dynamic labile equilibrium." Diabetes Manag (Lond) **3**(3): 217-223.

Schneider, D. A. and M. G. von Herrath (2014). "Potential viral pathogenic mechanism in human type 1 diabetes." Diabetologia **57**: 2009-2018.

Scotet, E., et al. (1996). "T Cell Response to Epstein-Barr Virus Transactivators in Chronic Rheumatoid Arthritis." The Journal of Experimental Medicine **184**: 1791-1800.

Serwold, T., et al. (2001). "ER aminopeptidases generate a unique pool of peptides for MHC class I molecules." Nature Immunology **2**: 644-651.

Sette, A. and J. Sidney (1998). "HLA supertypes and supermotifs: a functional perspective on HLA polymorphism." Current Opinion in Immunology **10**: 478-482.

Shepherd, J. C., et al. (1993). "TAP-I dependent peptide translocation in vitro is ATP dependent and peptide selective." Cell **74**: 577-584.

Shi, Y., et al. (1997). "Cytotoxic $CD8^+$ T cells recognize EBV antigen but poorly kill autologous EBV-infected B lymphoblasts: immunodominance is elicited by a peptide epitope that is presented at low levels in vitro. ." Journal of Immunology **159**: 1844-1852.

Shortman, K., et al. (1990). "The generation and fate of thymocytes." Semin. Immunol. **2**(1): 3-12.

Sidney, J., et al. (2008). "HLA class I supertypes: a revised and updated classification." BMC Immunology **9**(1).

Sidney, J., et al. (2003). "Simultaneous prediction of binding capacity for multiple molecules of the HLA B44 supertype. ." J Immunol **171**: 5964-5974.

Silins, S. L., et al. (1997). "Selection of a diverse TCR repertoire in response to an Epstein-Barr virus-encoded transactivator protein BZLF1 by $CD8^+$ cytotoxic T lymphocytes during primary and persistent infection." International Immunology **9**(11): 1745-1755.

Sipsas, N. V., et al. (1997). "Identification of type-specific cytotoxic T lymphocyte responses to homologous viral proteins in laboratory workers accidentially infected with HIV-1." J. Clin. Investig. **99**: 752-762.

Starr, T. K., et al. (2003). "Positive and Negative Selection of T Cells." Annual Reviews Immunology **21**: 139-176.

Steigerwald-Mullen, P., et al. (2000). "Type 2 cytokines predominate in the human $CD4^+$T-lymphocyte response to Epstein-Barr virus nuclear antigen 1." Journal of Virology **74**(6748-59).

Steven, N., et al. (1997). "Immediate Early and Early Lytic Cycle Proteins Are Frequent Targets of the Epstein-Barr Virus-induced Cytotoxic T Cell Response." The Journal of Experimental Medicine **185**(9): 1605-1617.

Steven, N. M., et al. (1996). "Epitope focusing in the primary cytotoxic T-cell response to Epstein-Barr virus and its relationship to T-cell memory." The Journal of Experimental Medicine **184**: 1801-1813.

Stewart-Jones, G., et al. (2005). "Structures of three HIV-1 HLA-B*5703-peptide complexes and identification of related HLAs potentially associated with long-term nonprogression." Journal of Immunology **175**: 2459-2468.

Taghon, T. and E. V. Rothenberg (2008). "Molecular mechanisms that control mouse and human TCR-αβ and TCR-γδ T cell development." Semin Immunopathol **30**: 383-398.

Taylor, G. S., et al. (2015). "The Immunology of Epstein-Barr Virus-Induced Disease." Annu. Rev. Immunol. **33**: 787-821.

Theodossis, A., et al. (2010). "Constraints within major histocompatibility complex class I restricted peptides: Presentation and consequences for T-cell recognition." Proceedings of National Academy Science **107**(12): 5534-5539.

Thomas, P. G., et al. (2013). "Ecological analysis of antigen-specific CTL repertoires defines the relationship between naive and immune T-cell populations." Proceedings of the National Academy of Sciences of the United States of America **110**: 1839-1844.

Tischer, S., et al. (2014). "Evaluation of suitable target antigens and immunoassays for high-accuracy immune monitoring of cytomegalovirus and Epstein-Barr virus-specific T cells as targets of interest in immunotherapeutic approaches." Journal of Immunological Methods **408**: 101-113.

Trautmann, L., et al. (2005). "Selection of T Cell Clones Expressing High-Affinity Publice TCRs within Human Cytomegalovirus-Specific CD8 T Cell Responses." The Journal of Immunology **175**(9): 6123-6132.

Tsang, C., et al. (2006). "$CD4^+$ T-cell responses to Epstein-Barr virus nuclearantigen EBNA1 in Chinese populations are highly focused on novel C-terminal domain-derived epitopes. ." Journal of Virology **80**(6): 8263-8266.

Turner, S. J., et al. (2006). "Structural determinants of T-cell receptor bias in immunity." Nature Reviews Immunology **6**: 883-894.

Turner, S. J., et al. (2005). "Lack of prominent peptide-major histocompatibility complex features limits repertoire diversity in virus-specific $CD8^+$ T cell populations." Nature Immunology **6**: 382-389.

Tynan, F. E., et al. (2005). "High Resolution Structures of Highly Bulged Viral Epitopes Bound to Major Histocompatibility Complex Class I." The Journal of Biological Chemistry **280**(25): 23900-23909.

Tynan, F. E., et al. (2005). "T cell receptor recognition of a "super-bulged" major histocompatibility complex class I-bound peptide." Nature Immunology **6**(11): 1114-1122.

Tynan, F. E., et al. (2005). "The immunogenicity of a viral cytotoxic T cell epitope is controlled by its MHC-bound conformation." The Journal of Experimental Medicine **202**(9): 1249-1260.

Ueno, T., et al. (2007). "Altering effects of antigenic variations in HIV-1 on antiviral effectiveness of HIV-specific CTLs." J Immunol **178**: 5513-5523.

Voo, K., et al. (2002). "Identification of HLA-DP3-restricted peptides from EBNA1 recognized by $CD4^+$ T cells." Cancer Research **62**: 7195-7199.

Wang, G. C., et al. (2012). "T cell receptor αβ diversity inversely correlates with pathogen-specific antibody levels in human cytomegalovirus infection." Sci Transl Med **4**: 128ra142.

Wange, R. and L. Samelson (1996). "Complex complexes: signaling at the TCR." Immunity **5**: 197-205.

Watts, C. and S. Amigorena (2001). "Phagocytosis and antigen presentation." Semin. Immunol. **13**: 373-379.

Weekes, M. P., et al. (1999). "The memory cytotoxic T-lymphocyte (CTL) response to human cytomegalovirus infection contains individual peptide-specific CTL clones that have undergone extensive expansion *in vivo*." Journal of Virology **73**: 2099-2108.

Wenzel, T., et al. (1994). "Existence of a molecular ruler in proteasomes by analysis of degradation products." FEBS Letters **349**: 205-209.

White, R. E., et al. (2012). "EBNA3B-deficient EBV promotes B cell lymphomagenesis in humanized mice and is found in human tumors." The Journal of Clinical Investigation **122**(4): 1487-1502.

Whitney, B. M., et al. (2002). "Frequency of Epstein-Barr Virus-Specific Cytotoxic T Lymphocytes in the Blood of Southern Chinese Blood Donors and Nasopharyngeal Carcinoma Patients." Journal of Medical Virology **67**: 359-363.

Wilson, I. A. and K. C. Garcia (1997). "T-cell receptor structure and TCR complexes." Current Opinion in Structural Biology **7**: 839-848.

Woodberry, T., et al. (2005). "Differential Targeting and Shifts in the Immunodominance of Epstein-Barr Virus-Specific CD8 and CD4 T Cell Responses during Acute and Persistent Infection." The Journal of Infectious Diseases **192**: 1513-1524.

Wright, C. A., et al. (2004). "Tapasin and other chaperones: models of the MHC class I loading complex." Biol. Chem **385**: 763-768.

Yang, Y., et al. (2014). "Sequence Analysis of EBV Immediate-Ealry Gene BZLF1 and BRLF1 in Lymphomas." Journal of Medical Virology **86**: 1788-1795.

Yewdell, J. W. and J. R. Bennink (1999). "Immunodominance in Major Histocompatibility Complex Class I-Restricted T Lymphocyte Responses." The Annual Review of Immunology **17**: 51-88.

Young, L. S. and A. B. Rickinson (2004). "Epstein-Barr virus: 40 years on." Nat. Rev. Cancer. **4**: 757-768.

Yui, M. and E. Rothenberg (2014). "Developmental gene networks: a triathlon on the course to T cell identity." Nature Reviews Immunology **14**: 529-545.

Zalani, S., et al. (1996). "Epstein-Barr viral latency is disrupted by the immediate-early BRLF1 protein through a cell-specific mechanism." Proceedings of National Academy Science **93**: 9194-9199.

Zamoyska, R. (1998). "CD4 and CD8: modulators of T-cell receptor recognition of antigen and of immune responses?" Current Opinion in Immunology **10**: 82-87.

Zehn, D. and M. J. Bevan (2006). "T cells with low avidity for a tissue-restricted antigen routinely evade central and peripheral tolerance and cause autoimmunity. ." Immunity **25**: 261-270.

Zernich, D., et al. (2004). "Natural HLA Class I Polymorphism Controls the Pathway of Antigen Presentation and Susceptibility to Viral Evasion." The Journal of Experimental Medicine **200**(1): 13-24.

Zhang, Q., et al. (1999). "Identification of transactivator and nuclear localization domains in the Epstein-Barr virus DNA polymerase accessory protein, BMRF1." Journal of General Virology **80**: 69-74.

Zhao, B. and C. A. Sample (2000). "Epstein-Barr virus nuclear antigen 3C activates the latent membrane protein 1 promoter in the presence of Epstein-Barr virus nuclear antigen 2 through sequences encompassing an spi-1/Spi-B binding site. ." Journal of Virology **74**: 5151-5160.

Zinkernagel, R. M., et al. (1978). "Ir-genes in H-2 regulate generation of anti-viral cytotoxic T cells. Mapping to K or D and dominance of unresponsiveness." The Journal of Experimental Medicine **148**: 592-606.

Zinkernagel, R. M. and P. C. Doherty (1974). "Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system." Nature **248**: 701-702.

Supplementary Data

Chapter 5:

Aim Three:

Determine if the peptide length preferences of class I human leukocyte antigens influence epitope selection in the EBV-specific T cell response.

Hypothesis Three:

Different class I human leukocyte antigens have distinct peptide length preferences, and this influences epitope selection in the EBV-specific T cell response.

Publication: Original research article

Rist MJ, Theodossis A, Croft NP, Neller MA, Welland A, Chen Z, Sullivan LC, Burrows JM, Miles JJ, Brennan RM, Gras S, Khanna R, Brooks AG, McCluskey J, Purcell AW, Rossjohn J, Burrows SR. HLA Peptide Length Preferences Control $CD8^+$ T Cell Responses. *The Journal of Immunology*. 2013 June; 191:561-571

Supplementary Material

Supplemental Figure 1: T cell recognition of overlapping peptides from BZLF1. T cell lines from nine EBV-seropositive donors, raised by *in vitro* stimulation with irradiated autologous LCLs, tested in interferon-γ ELISpot assays for recognition of five overlapping peptides corresponding to a region of the BZLF1 protein sequence.

Supplemental Table I: Data collection statistics

	HLA-B*18:01-SELEIKRY	HLA-B*44:03-EECDSELEIKRY
Wavelength (Å)	0.9794	0.9567
Resolution limits (Å) [a]	46.22 – 1.50 (1.58 - 1.50)	40.18 – 1.90 (2.00 – 1.90)
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 50.85, b = 81.25, c = 110.89	a = 50.82, b = 81.62, c = 110.34,
No. observations [a]	267,076 (38,021)	205,568 (28,897)
No. unique reflections [a]	73,568 (10,522)	35,075 (5,142)
Mosaicity	0.4	0.8
Completeness (%) [a]	99.1 (98.5)	95.5 (97.1)
R_{merge} (%) [a, b]	6.1 (52.5)	14.7 (56.9)
$R_{p.i.m.}$ (%) [a, c]	3.7 (31.6)	6.5 (24.7)
$< I/\sigma(I) >$ [a]	12.2 (2.4)	8.3 (2.6)
Multiplicity [a]	3.6 (3.6)	5.9 (5.6)

[a] Values in parentheses refer to the highest resolution bin.

[b] $R_{merge} = \Sigma_{(hkl)} \Sigma_i | I_{i(hkl)} - \langle I_{(hkl)} \rangle | / \Sigma_{(hkl)} \Sigma_i \langle I_{(hkl)} \rangle$ where I is the observed intensity and $< I >$ is the average intensity of multiple observations from symmetry-related reflections.

[c] $R_{p.i.m.} = \Sigma_{(hkl)} [1/(N-1)]^{1/2} \Sigma_i | I_{i(hkl)} - \langle I_{(hkl)} \rangle | / \Sigma_{(hkl)} \Sigma_i \langle I_{(hkl)} \rangle$ where N is the redundancy of the *hkl* reflection.

Supplemental Table II: Refinement statistics

	HLA-B*18:01-SELEIKRY	HLA-B*44:03-EECDSELEIKRY
Resolution (Å)	40.6 – 1.5	38.3 – 1.9
Working set reflections	73,505	35,037
Test set reflections	3,713	1,751
Data completeness (%)	98.9	94.90
R-factor (%) [a]	18.9	19.5
R-free (%) [b]	21.4	23.1
• Number of atoms:		
- protein	3112	3095
- peptide	73	105
- water	422	201
- other	4	14
• Average atomic B-factors (Å^2):		
- protein	18.7	19.4
- peptide	18.1	28.8
- water	28.4	24.5
- other	31.0	26.2
rmsd bonds (Å)	0.006	0.007
rmsd angles (°)	1.067	1.052
•*MOLPROBITY* results (%):		
- Ramachandran favoured	97.7	97.4
- Ramachandran outliers	0.0	0.0

[a] $R_{factor} = (\Sigma \, | \, |F_o| - |F_c| \, |) / (\Sigma \, |F_o|)$ - for all data except as indicated in footnote d.

[b] 5.0% of data were used for the R_{free} calculation (see footnote c).

Supplemental Figure S1