Towards a universally effective HLA-E-targeted HIV vaccine - a structural and biophysical exploration



Lucy Walters

Linacre College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2019

Abstract

Doctor of Philosophy in Clinical Medicine, Lucy Walters, Linacre College, Trinity 2019

Towards a universally effective HLA-E-targeted HIV vaccine - a structural and biophysical exploration

Through major histocompatibility complex class I leader sequence-derived peptide (VL9) binding and CD94/NKG2 receptor engagement, human leucocyte antigen E (HLA-E) reports cellular health to NK cells. Previous studies demonstrated a strong bias for VL9 binding by HLA-E, a preference subsequently supported by structural analyses. However, *Mycobacteria tuberculosis* (Mtb) infection and Rhesus *cytomegalovirus*-vectored SIV vaccination revealed contexts where HLA-E and its rhesus orthologue, Mamu-E, present sequence-diverse pathogen-derived peptides to $CD8^+$ T cells, respectively. In light of unprecedented vaccine-induced Mamu-E-restricted $CD8^+$ T cell-mediated protection against SIV, it is now important to re-explore the peptide binding specificity of the human orthologue, HLA-E, from the view of developing a universal HLA-E-targeted HIV vaccination strategy in humans. Through the development and optimisation of a high throughput *in vitro* peptide binding assay, we demonstrated an HLA-E binding capacity for a number of pathogen-derived epitopes containing non-canonical anchor residues. As previous crystal structures of HLA-E were exclusively solved in complex with canonical VL9 peptides, we also investigated the structural basis underlying pathogen-derived peptide binding. We obtained a panel of HLA-E crystal structures in complex with a number of Mtb-derived epitopes in addition to the HIV counterpart peptide to the vaccine-identified SIV Gag 'supertope$_{69}$'. Strikingly, despite the presence of canonical primary anchor residues, the HIV peptide adopted an alternative conformational motif within the HLA-E peptide binding groove, deviating from that of canonical VL9 and in turn generating a distinct solvent exposure profile, with potential implications for T cell receptor recognition and immunogenicity. Further, combined structural and mutagenesis analyses illustrated a greater tolerance for hydrophobic and polar residues in the primary pockets of the HLA-E binding groove than previously anticipated. Finally, biophysical analyses demonstrated the impact of suboptimal peptide binding on the overall conformational ensemble of HLA-E complexes in solution.

Acknowledgements

Thank you to Professor Geraldine Gillespie and Professor Sir Andrew McMichael of Oxford University for being the supervisory dream team! You have both repeatedly gone above and beyond as supervisors and it has been an absolute honour to be your DPhil student - thank you for making the last four years so much fun! I'm not sure it would be possible to find another person who loves HLA-E as much as Andrew McMichael - thank you for the opportunity to be a part of your legendary, world-renowned lab! And Gerry, I'm not sure how you manage to know everything there is to know in the lab and in the literature and still find the time to be the most personable and supportive supervisor on the planet.

I would also like to thank Professor E. Yvonne Jones, Dr Karl Harlos and Dr Daniel Rozbesky of Oxford University for their excellent structural supervision and Professor Robert Rambo of Diamond Light Source for his invaluable biophysical input. Karl, I will certainly not forget the times you kindly taught me how to operate a synchrotron beamline at 4am on a Saturday morning - what a highlight!

I would like to thank my parents for their unconditional support and for putting up with me moving back in with them to write up my thesis - although I'm sure they would agree I have been an easy-going stress-free absolute delight throughout the entire process. I would like to thank my sister for her moral support - largely in the form of relentless sarcasm - throughout our DPhils, and also my little brother for being reliably lovely and even bringing me pizza on the darkest of thesis days. Thank you to my husband Wes for your total dedication, for sharing my passion for science and for keeping me (relatively) sane during the final few months.

Importantly, I would like to thank the Bill & Melinda Gates Foundation for funding my DPhil and research. And finally, I would like to thank the science gods for making my crystals grow (and diffract).

Parts of the Methods and materials chapter and Abstract of this thesis are adapted from my first author publication (Walters et al., 2018) *or from my Master's thesis, however all such text and figures were originally my own work.*

List of publications

1. Walters, L.C., Harlos, K., Brackenridge, S., Rozbesky, D., Barrett, J.R., Jain, V., Walter, T.S., O'Callaghan, C.A., Borrow, P., Toebes, M., Hansen, S.G., Sacha, J., Abdulhaqq, S., Greene, J.M., Früh, K., Marshall, E., Picker, L.J., Jones, E.Y., McMichael, A.J., Gillespie, G.M., 2018. Pathogen-derived HLA-E bound epitopes reveal broad primary anchor pocket tolerability and conformationally malleable peptide binding. Nature Communications.

2. Hansen, S.G., Wu, H.L., Burwitz, B.J., Hughes, C.M., Hammond, K.B., Ventura, A.B., Reed, J.S., Gilbride, R.M., Ainslie, E., Morrow, D.W., Ford, J.C., Selseth, A.N., Pathak, R., Malouli, D., Legasse, A.W., Axthelm, M.K., Nelson, J.A., Gillespie, G.M., Walters, L.C., Brackenridge, S., Sharpe, H.R., López, C.A., Früh, K., Korber, B.T., McMichael, A.J., Gnanakaran, S., Sacha, J.B., Picker, L.J., 2016. Broadly targeted CD8+ T cell responses restricted by major histocompatibility complex E. Science

Contents

Abstract……i

Acknowledgements……ii

List of publications……iv

List of abbreviations……x

1. Introduction……1

Basic overview……1

The ongoing need for a vaccine against human immunodeficiency virus……2

Cytomegalovirus as a potential vector for next generation HIV vaccination……6

Genetic manipulation of Cytomegalovirus-based vaccine vectors……12

HLA-E-restricted $CD8^+$ T cell responses specific for *Mycobacterial* antigen……14

The original characterisation and the primary functional role of HLA-E……17

Structural insights into MHC class I leader sequence-derived peptide-bound HLA-E……24

Insights from ligand-receptor co-complex crystal structures involving HLA-E……32

Alternative contexts of HLA-E/Qa-1-restricted $CD8^+$ T cell responses and HLA-E/Qa-1 binding peptides……34

Aims and Objectives……42

2. Materials and methods……44

Expression Constructs……44

Protein Expression……46

Inclusion body isolation and protein purification……47

Peptide synthesis……48

Peptide binding assays……48

Sandwich ELISA……49

Protein refolding and purification……50

Crystallisation screening……51

Crystallographic analysis……52

Size exclusion chromatography-coupled small angle X-ray scattering……53

3. Results chapter 1 - HLA-E peptide binding assays……54

Introduction……54

Micro-refold-based HLA-E peptide binding assay……55

Comparison of micro-refolding versus UV-induced peptide exchange-based HLA-E peptide binding assays……65

Optimisation of the peptide exchange-based binding assay……71

Photo-assisted peptide exchange buffer conditions……71

L-Arginine monohydrochloride concentration……78

Optimisation of pH conditions in the UV-assisted peptide exchange reaction……80

Sandwich ELISA protein incubation temperature……82

Duration of UV irradiation in the peptide exchange reaction……83

Peptide exchange reaction in the absence of UV illumination……87

Peptide exchange-based binding assay screening of HLA-E peptide panels……91

Peptide panel 1 - *Previously described pathogen- and self-derived HLA-E binding peptides*……91

Peptide panel 2 - *11 optimal SIV-derived RhCMV 68-1 vaccine-identified Mamu-E restricted epitopes……94*

Peptide panel 3 - *RhCMV 68-1 vaccine-identified HIV Gag-derived optimal 15mers and overlapping 9mers……97*

Peptide panel 4 - *HIV Gag- and Pol-derived peptides computationally predicted to bind HLA-E by Bette Korber's Group……100*

Peptide panel 5 - *HIV-derived NetMHC-predicted HLA-E binding peptides……102*

Peptide panel 6 - *Mycobacterial peptides computationally predicted to bind HLA-E……104*

Peptide panel 7 - *HLA-E-restricted Mycobacterial peptides eluted from infected cells……106*

HLA-E sequence motifs based on screening of peptide panels in the peptide-exchange assay……108

Discussion……111

Peptide binding assay optimisation……111

Rhesus versus human MHC-E peptidomes……114

Peptide binding assay screening……117

4. Results chapter 2 - Structural analysis of pathogen-derived peptide binding to HLA-E……119

Introduction……119

Crystal structures of Mtb-derived peptide-bound HLA-E*0103……121

Structural determination of HLA-E*0103 in complex with the Mtb44 peptide……123

Overall structural architecture and alignment of Mtb44-bound HLA-E molecules……125

Root mean square deviation between molecules of the Mtb44-bound HLA-E*0103 structure……127

HLA-E heavy chain-derived alpha-3 domain and β2M interface analysis……129

Temperature factors……130

Alignment of Mtb44 versus VL9-bound HLA-E molecules……132

Positioning of the Mtb44 peptide in the four molecules of the asymmetric unit……133

Mtb44 peptide positioning within the HLA-E binding groove……135

Pocket-binding in the HLA-E*0103-Mtb44 structure……137

Hydrogen bonding……141

Structural determination of HLA-E*0103 in complex with Mtb14……143

Mtb14-bound HLA-E*0103 molecule alignment……143

Temperature factor analysis……147

Mtb14 peptide positioning between molecules of the asymmetric unit……152

Mtb14 peptide positioning within the HLA-E binding groove……154

Combined structural and mutagenesis analyses of HLA-E*0103 in complex with Mtb44 position 2 and 9 variants……160

Screening of Mtb44 and VL9 primary anchor variants in the peptide-exchange binding assay……161

Structural determination of HLA-E*0103 in complex with Mtb44 position 2 and 9 variants……166

HLA-E*0103-Mtb44*P2-Phe structural interpretation……169

HLA-E*0103-Mtb44*P2-Gln structural interpretation……177

HLA-E*0103-Mtb44*P9-Phe structural interpretation……182

Comparative analysis of peptide positioning between Mtb44 variant structures……185

Pockets……188

Inter-chain hydrogen bonding between Mtb44 variant peptides and the HLA-E heavy chain……190

Discussion……191

Implications for T cell recognition and vaccine-design……191

Conformational flexibility of the α3 domain……193

Conformational flexibility in the α2 helix short-arm……198

Pocket-forming residue rearrangement in response to anchor residue variation……200

5. Results chapter 3 - HLA-E-restricted conformationally distinct peptide binding motifs and a biophysical analysis of suboptimal peptide binding to HLA-E……204

Introduction……204

Structural analysis……206

Structural determination of HLA-E*0103 in complex with the RL9HIV peptide…206

HLA-E*0103-RL9HIV molecule alignment……210

Alignment of RL9HIV-HLA-E molecules to previous VL9-bound structures of HLA-E……213

RL9HIV peptide positioning within the HLA-E binding groove……214

Hydrogen bonding network between the RL9HIV peptide and HLA-E heavy chain……217

Pocket occupancy in the RL9HIV-HLA-E*0103 structure……219

Structural determination of HLA-E*0103 in complex with the IL9 peptide……223

Alignment of HLA-E*0103-IL9 molecules from the asymmetric unit……224

Alignment of IL9-HLA-E molecules to previous VL9-bound structures of HLA-E……227

IL9 peptide positioning within the HLA-E binding groove……229

Hydrogen bonding network between the IL9 peptide and HLA-E heavy chain……233

Pocket binding in the HLA-E*0103-IL9 structure……236

Size exclusion chromatography-coupled small angle X-ray scattering analysis of HLA-E in complex with peptides of differential binding affinity……243

Intra-sample conformational heterogeneity……252

A different buffer system trialled in the SEC-SAXS analysis of HLA-E……258

Excess peptide in the HPLC elution buffer……262

Structural modelling of SAXS intensity profiles……275

Discussion……276

Conformationally distinct HLA-E-restricted peptide binding motifs……276

Differential structural positioning of the IL9 peptide main chain……279

Insights from SEC-SAXS analyses of HLA-E refolds……280

Unsuccessful attempts to crystallise low affinity peptide-bound HLA-E……284

6. Overall discussion and concluding remarks……286

7. References……296

List of abbreviations

AIDS - Acquired immunodeficiency syndrome

ART - Anti-retroviral therapy

AS - Ammonium sulphate

AU - Asymmetric unit

β2M - β-2 microglobulin

BCG - Bacillus Calmette–Guérin vaccine

bnAbs - Broadly neutralising antibodies

BSA - Bovine serum albumin

CMV - *Cytomegalovirus*

CTL - Cytotoxic T lymphocyte

d_{max} - Maximum surface dimension

DTT - Dithiothreitol

EBV - Epstein-Barr virus

EDTA - Ethylenediamineteraacetic acid

ELISA - Enzyme-linked immunosorbent assay

ER - Endoplasmic reticulum

ERAAP - Endoplasmic reticulum aminopeptidase associated with antigen processing

FPLC - Fast protein liquid chromatography

HCMV - Human *Cytomegalovirus*

HCV - Hepatitis C virus

HIV - Human immunodeficiency virus

HLA-E - Human leukocyte antigen E

HPLC - High-performance liquid chromatography

Hsp60sp - Heat shock protein derived signal peptide

iIEL - Intestinal intraepithelial lymphocyte

IPTG - Isopropyl-β-D-thiogalactoside

K_D - Dissociation constant

LB - Luria-Bertani

LTNP - Long term non-progressor

Mtb - *Mycobacteria tuberculosis*

NaCl - Sodium Chloride

PBMCs - Peripheral blood mononuclear cells

PBS - Phosphate buffered saline

PPD - Purified protein derivative

Prdx5 - Peroxiredoxin 5

Qdm - Qa-1 determinant modifier

R_g - Radius of gyration

RhCMV - Rhesus *Cytomegalovirus*

Rmsd - Root mean square deviation

SEC-SAXS - Size exclusion chromatography-coupled small angle x-ray scattering

SIV - Simian immunodeficiency virus

STDEV - Standard deviation

TAP - Transporter associated with antigen presentation

T_{CM} - Central memory $CD8^+$ T cell

TCR - T cell receptor

T_{EM} - Effector memory $CD8^+$ T cell

TFH - $CD4^+$ follicular helper T-cells

T_m - Melting transition temperature

TMB - tetramethyl benzidine

T_{reg} - T regulatory cell

VL9 - MHC class I leader sequence-derived signal peptide

Introduction

Basic overview

An effective vaccine against HIV is not currently available. Several challenges have prevented the development of an effective HIV vaccine thus far, such as the ability of the virus to remain dormant inside infected cells preventing detection by vaccine-induced immune responses. These same challenges also prohibit the development of natural immunity to HIV meaning that following initial infection the virus spreads rapidly throughout the body and infects long-lived cells resulting in life-long infection. HIV specifically infects vital cells of the immune system and thus overtime, HIV-driven depletion of such cells results in a compromised immune status. Eventually, the immune system becomes so weak that HIV-infected individuals become highly susceptible to acquiring additional types of infection - often involving micro-organisms which would otherwise not pose a serious threat to an immune-competent HIV-negative individual. Additionally, a weakened immune system results in compromised immune surveillance of tumour cells leading to a particularly high incidence of cancer in such individuals. It is at this point, when the levels of certain immune cells have dropped below a given threshold resulting in increased susceptibility to infection and malignancy, that an HIV-positive individual is described as having progressed to AIDS (Acquired Immune Deficiency Syndrome). As naturally-elicited immune responses that protect against HIV infection have not been identified, this further confounds efforts to design an effective HIV vaccine, since different vaccines are specifically engineered to induce a certain type of immune response. However, there exists an urgent and significant global need for the development of an effective HIV vaccine. In 2017, 2 million previously uninfected people developed an HIV infection, bringing the total number of HIV infected individuals worldwide close to 40 million (World Health Organization (WHO/ CDS/HIV), 2019). Further, in the absence of an HIV vaccine,

it has been predicted that an additional 49 million people will become infected with HIV by 2035 (Medlock et al., 2017).

Recent advances in the HIV vaccine field have yielded promising results. Notably, one vaccine strategy elicited immune responses in vaccinated animals which conferred unprecedented, robust protection from infection with SIV, the simian equivalent of HIV (Hansen et al., 2016). Unexpectedly, upon further investigation it was found that these putatively protective immune responses were functionally distinct from conventional immune responses. Specifically, an unconventional, vaccine-induced T cell subset was able to recognise infected cells via a molecule called HLA-E in humans – a molecule that is not traditionally thought to participate in T cell immunity in this way. Although specific types of T cell have evolved to detect virally-infected cells, such detection is not usually mediated through HLA-E. Moreover, this unusual HLA-E-restricted T cell subset has also recently been identified to play a role in the immune response against tuberculosis. Similar to HIV, tuberculosis is considered a major global disease accounting for 1.6 million related-deaths in 2017 alone, according to the World Health Organisation. Although a vaccine against tuberculosis is currently available, it offers poor protection against the most common manifestation of the infection in adults - pulmonary disease. Such findings have consequently generated significant interest in the potential role of HLA-E in the immune response to HIV and tuberculosis. It is now important to explore this alternative functional role of HLA-E in T cell responses with the ultimate aim of developing an HLA-E-targeted vaccine against HIV, and also possibly a more effective vaccine against tuberculosis, in humans.

The ongoing need for a vaccine against human immunodeficiency virus

A prophylactic vaccine engaging both arms of the adaptive immune system is considered the most cost-effective approach in overcoming the ongoing human immunodeficiency virus 1 (HIV-1) pandemic, accountable for 37.9 million HIV-infected individuals worldwide (Fauci and

Marston, 2014) (Corey and Gray, 2017) (World Health Organization (WHO/ CDS/HIV), 2019). Controlling HIV transmission is complicated by factors such as non-specific symptoms that typify seroconversion followed by extended periods of viral latency (Tindall et al., 1988) (Weintrob et al., 2003) (Schacker et al., 1996). Consequently, only 56% of HIV-infected individuals in countries with high disease burdens, such as those in eastern and southern Africa, are aware of their seropositive status, of which only 54% were receiving anti-retroviral therapy (ART) in 2016 (Joint United Nations Programme on HIV/AIDS, 2016). Further, mathematical modelling projections based on the current rate of ART administration in 127 countries, estimated that 49 million new HIV infections will be acquired worldwide between 2015 and 2035 in the absence of a prophylactic vaccine (Medlock et al., 2017).

Despite extensive vaccination trials, following the identification of HIV as the etiological agent of AIDS in 1983 (Barre-Sinoussi et al., 1983), the highest protective efficacy elicited thus far has been limited to 31.2% in the phase 3 Rv144 Thai trial - a result reported 42 months post-vaccination and determined with borderline statistical significance ($p=0.04$) (Rerks-Ngarm et al., 2009). The high p-value corresponds to a 25% chance of zero vaccine efficacy according to standard Bayesian models (Gilbert et al., 2011). However, subsequent functional studies which identified immune correlates of infection risk, including putatively protective non-neutralising IgG, support a degree of efficacy (Haynes et al., 2012) (Corey et al., 2015) (Bonsignori et al., 2012). The Rv144 Thai trial comprised a prime-boost regimen of conventional vaccination techniques - a non-replicating canarypox vector recombinant for HIV immunogens Gag Pol and Env (ALVAC) in combination with subsequent recombinant glycoprotein-120 (AIDSVAX) boosting (Rerks-Ngarm et al., 2009). Notably, independent efficacy trials had previously been conducted for AIDSVAX in which no significant protective effects were detected (Pitisuttithum et al., 2006). Such HIV envelope-focused approaches were originally of particular interest following the reported protective effects of passive transfer of neutralising simian immunodeficiency virus (SIV)-specific antibody in non-human primates (Shibata et al., 1999).

Non-replicating viral vectors were also administered in the phase 2b STEP trial which deployed Adenovirus 5 vectors recombinant for HIV Gag, Pol or Nef, with the aim of inducing cytotoxic T cell responses to control viral load (Gray et al., 2010). Pre-clinical trials in primate models yielded promising results with vaccine-induced T cell-mediated reductions in viral set points following challenge with SHIV (SIV engineered to express an HIV envelope) (Shiver et al., 2002). T cell-targeted HIV vaccine approaches appeared particularly attractive after it was demonstrated that HIV Gag-specific cytotoxic T cell responses to functionally constrained epitopes were enriched in non-progressive infection (Pontesilli et al., 1998). Further, *in vivo* $CD8^+$ T cell depletion of SIV-infected non-human primates with viremia (Schmitz et al., 1999) or elite control (Friedrich et al., 2007) yielded elevated viral replication levels and increased plasma loads, illustrating the importance of T cell-mediated responses in controlling SIV infection. However regarding the STEP trial, an increased incidence of HIV infection among Adenovirus 5 pre-exposed individuals was observed in the vaccine group relative to the control group leading to premature trial termination (Iaccino et al., 2008). It was later speculated that pre-existing Adenovirus 5 immunity likely created a conducive environment for HIV replication with the recruitment and activation of Adenovirus 5-specific $CD4^+$ memory T cells - suitable host cells for HIV infection (Sekaly, 2008). Alternative, previously attempted vaccination approaches include pre-clinical trials in non-human primates with live attenuated simian immunodeficiency virus (SIV), such as Nef-deficient SIVmac239, which yielded robust protection against subsequent intravenous SIV challenge (Daniel et al., 1992) (Johnson and Desrosiers, 1998). However, the possibility of reversion to virulence poses inherent safety risks and hence such approaches have not since been translated into the human clinic. Moreover, the extent of live vaccine attenuation was frequently found to inversely correlate with its protective efficacy, further constraining the development of live attenuated vaccines against HIV (Johnson and Desrosiers, 1998).

The limited success of HIV vaccine trials to date supports the theory that next generation approaches will require unconventional strategies to overcome the plethora of immune evasion mechanisms employed by the virus which prohibits the establishment of natural immunity in the majority of individuals and limits efficacy in conventional vaccination platforms (Kirchhoff, 2010) (Agostini et al., 2002). Namely, antigenic hypervariability, generated by the low fidelity, recombinogenic reverse transcriptase (Kati et al., 1992) (Telesnitsky and Goff, 1997), in addition to extensive glycosylation of HIV Env and transient exposure of functionally conserved domains during viral entry (Fenouillet et al., 1994), promote immune escape and evasion. Additionally, accessory proteins encoded by the HIV genome target anti-viral host restriction factors or components of the antigen presentation pathway, for example HIV Nef which mediates the downregulation of HLA-A and HLA-B thwarting T cell recognition (Schwartz et al., 1993) (Collins et al., 1998). Viral latency following early integrative events into the host genome (Bushman et al., 1990) and direct cell-to-cell spread across virological synapses (Bracq et al., 2018), further facilitate immune evasion resulting in persistent viral reservoirs. Although certain immune factors identified in highly exposed seronegative individuals, including inhibitory killer Ig-like receptors (KIR) expression in the absence of cognate HLA genes, have been associated with resistance to HIV acquisition (Jennes et al., 2006), robust models of natural immunity and resistance are lacking, which complicates vaccine design. A range of next generation HIV vaccine approaches are currently under investigation. Notably, the discovery of potent, broadly neutralising antibodies (bnAbs) specific for conserved regions of HIV Env surface trimers in a minority of HIV positive patients with disseminated infection has prompted the design of HIV immunogens aimed at eliciting bnAbs in HIV naïve individuals (Burton and Hangartner, 2016). However, the main focus of further discussion shall constitute an unconventional T cell-targeted HIV vaccine strategy.

Cytomegalovirus as a potential vector for next generation HIV vaccination

Replicating viral vectors represent attractive HIV vaccine candidates due to their superior immunogenicity relative to protein-based or non-replicating viral vectors, in addition to improved safety profiles compared to live attenuation-based approaches (Robert-Guroff, 2007). However, the spatial and temporal signatures of anamnestic central memory $CD8^+$ T cell (T_{CM}) responses typically elicited by traditional, non-persistent replicating viral vectors, such as recombinant Orthopoxvirus or Adenovirus, render them poorly suited for controlling early HIV infection (Haase, 2010) (Johnson et al., 2013). By contrast, a next generation approach trialled by Hansen et al. in rhesus macaques deployed a replicating, SIV-recombinant *Cytomegalovirus* (CMV) vector; a persistent Beta-herpesvirus that provides long-term, sporadic, low-level antigenic exposure maintaining unusually high frequencies of effector memory $CD8^+$ T cells (T_{EM}) without causing immune exhaustion (Virgin et al., 2009) (Hansen et al., 2011). The rationale behind this approach was to exploit the persistent nature of the vector to drive $CD8^+$ T_{EM}-biased responses, thereby overcoming temporal and spatial limitations of anamnestic T_{CM} responses which typically dominate in conventional vaccination settings (Robert-Guroff, 2007). T_{CM} cells exhibit distinct homing patterns which are largely restricted to secondary lymphoid organs and must undergo differentiation into T_{EM} cells upon antigen re-exposure prior to trafficking to peripheral sites of infection where such cells subsequently perform cytotoxic effector functions (Sallusto et al., 2004). However, considering the genetic bottleneck of infecting founder virus clones and rapid HIV replication kinetics, it is thought HIV vaccines should ideally aim to target such homogenous populations of virus at the peripheral portal of entry in the short window following initial infection prior to dissemination and the establishment of hyper-variable quasispecies and latent reservoirs (Haase, 2010). Therefore, despite their limited expansion potential relative to T_{CM} cells, T_{EM} cells are considered to represent desirable HIV vaccine targets since such cells home to inflamed, non-lymphoid

tissues and remain in an effector-differentiated state capable of immediate effector function following initial antigen exposure (Sallusto et al., 2004) (Genescà et al., 2009).

The multivalent rhesus CMV-based vaccine vector recombinant for SIV genes Gag-Pol-Rev-Nef-Tat-Env (RhCMV 68-1-SIV) elicited unprecedented, potent early immune-control with apparent viral clearance in around 50% of vaccinated macaques, irrespective of the route of challenge with the highly pathogenic, neutralisation-resistant SIVmac239 (Scott G Hansen et al., 2013). The precise underlying mechanisms determining this dichotomy of protection have not been fully elucidated, although distinct inflammatory profiles have been identified post-vaccination, via transcriptomic analyses of peripheral blood mononuclear cells (PBMCs), which can distinguish the protected 50% of rhesus macaques relative to the 50% with progressive SIV infection (Louis Picker, Unpublished data). Early SIV infection was evident post-challenge in 100% of RhCMV 68-1-SIV-vaccinated rhesus macaques, with viral replication detected at systemic haematogenous and lymphoid sites including the draining lymph nodes, spleen and bone marrow for up to six weeks - although the protected group exhibited lower viral loads and smaller, transient viral reservoirs (Scott G Hansen et al., 2013). However, the subsequent course of infection exhibited striking differences in protected versus unprotected RhCMV 68-1-SIV-vaccinated macaques; unprotected macaques developed progressive SIV infection whereas early SIV infection was arrested in the protected 50% with viral plasma loads plummeting to undetectable levels with intermittent spikes of transient viremia which decreased in frequency overtime and eventually subsided altogether (Hansen et al., 2011).

As predicted, a T_{EM} bias dominated RhCMV 68-1-SIV vaccine-induced $CD8^+$ T cell memory responses, with the majority of cells displaying a fully differentiated $T_{EM}2$ phenotype (CD28-CCR7-) and cytotoxic effector function (Hansen et al., 2009). By contrast, a non-protective T_{CM} bias was present in a control group of rhesus macaques which developed progressive SIV infection following vaccination with a non-persistent, replication-defective Adenovirus 5

vector recombinant for SIV genes (Hansen et al., 2009). The exclusion of SIV Vif from the RhCMV 68-1 vector enabled discrimination of vaccine-induced SIV-specific T cell responses from those elicited by SIV challenge. Although Vif-specific $CD8^+$ T cell responses waned overtime, maintenance of RhCMV 68-1-SIV-elicited T_{EM} cell responses at high frequencies 700 days post-challenge was indicative of robust, long-term, vaccine-mediated protection. Further, SIV replication did not ensue following monoclonal antibody-mediated $CD8^+$ T cell depletion in protected RhCMV 68-1-SIV-vaccinated rhesus macaques, or following adoptive transfer of haematolymphoid cells from long-term protected, RhCMV 68-1-SIV-vaccinated rhesus macaques into SIV naïve recipients - such results are consistent with the established criteria of a functional HIV cure and strongly indicate viral clearance (Hansen et al., 2009). Further, the inability of multiple ultra-sensitive techniques to detect plasma or tissue-associated SIV or discriminate between uninfected and RhCMV 68-1-SIV protected rhesus macaques 3 years post challenge is unparalleled, especially considering HIV or SIV reservoirs persist in all other settings of control previously investigated including long term non-progressors (LTNP), elite controllers and anti-retroviral therapy responders (Poropatich and Sullivan, 2011) (Okulicz et al., 2009).

Unexpectedly, RhCMV 68-1-elicited SIV-specific $CD8^+$ T_{EM} responses defied fundamental T cell restriction paradigms: 63% were restricted by MHC class II and 35% by the rhesus non-classical MHC class I homologue of HLA-E, Mamu-E (2% of responses were undeterminable) (Scott G. Hansen et al., 2013) (Hansen et al., 2016). Remarkably, a dramatically enhanced breadth of recognition was observed for these unconventionally-restricted RhCMV 68-1-SIV-elicited $CD8^+$ T cell responses - three-fold greater than that typically observed for conventionally-restricted SIV-specific $CD8^+$ T cell responses. These peptide epitopes, which exhibited extraordinary promiscuity comprising diverse amino acid sequences which lacked a common motif, covered 66% of the SIV Gag protein (Hansen et al., 2016). Further, the epitopes did not constitute an extension of previously identified, conventionally-restricted SIV Gag epitopes targeted in

conventional vaccine settings or natural SIV infection, illustrating the uniqueness of the distinct, non-overlapping RhCMV 68-1 vector-driven unconventionally-restricted repertoire (Scott G. Hansen et al., 2013). Superior epitope recognition breadth in RhCMV 68-1-SIV vaccinated macaques is considered a critical factor contributing to the protective efficacy of this vaccine approach given the inverse correlation between $CD8^+$ recognition breadth and the probability of the emergence of de novo escape mutants in HIV infection (van Deutekom et al., 2013). Further, two 'supertopes' were identified among Mamu-E-restricted epitopes which represent peptides recognised by 100% of vaccinated macaques, illustrating the universal potential of an MHC-E (Major histocompatibility complex E)-targeted HIV vaccine approach if reproducible in humans (Hansen et al., 2016). The 11 optimal Mamu-E-restricted SIV Gag epitopes identified in RhCMV 68-1-SIV vaccination studies, including the two supertopes, are listed in Table 1.

Peptide ID	Sequence	Aas	Organism	Protein	Ref.
AE9	APLVPTGSE	9	*Simian immunodeficiency virus*	Gag_{65-73}	Hansen, S,G. et al (2016)
SN9	SENLKSLYN			Gag_{72-80}	
EA9	EKVKHTEEA			Gag_{92-100}	
KS9	KTSRPTAPS			$Gag_{120-128}$	
GI9	GGNYPVQQI			$Gag_{132-140}$	
QE9	QIIRDIINE			$Gag_{201-209}$	
VI9	VGNIYRRWI			$Gag_{259-267}$	
$RL9_{supertope}$	RMYNPTNIL			$Gag_{276-284}$	
KG9	KPIKCWNCG			$Gag_{389-397}$	
LQ9	LGKQQREKQ			$Gag_{476-484}$	
$EK9_{supertope}$	EKQRESREK			$Gag_{482-490}$	

Table 1: Optimal Mamu-E-restricted SIV epitopes

A. Table detailing 11 optimal SIV epitopes restricted by Mamu-E identified in RhCMV 68-1-SIV vaccination studies. Peptide ID, sequence, length and origin are reported with the corresponding reference.

Subsequent, unpublished vaccination studies involving RhCMV 68-1-SIV vectors re-engineered to exclusively elicit Mamu-E-restricted $CD8^+$ T cell responses in the absence of MHC class II-restricted $CD8^+$ T cell responses and vice versa, confirmed that the Mamu-E-restricted element of the vaccine-elicited response is primarily responsible for conferring protection against SIV challenge (Louis Picker, unpublished data). However, it is important to note that vaccine-induced broadly-targeting Mamu-E-restricted $CD8^+$ T cell responses are indistinguishable between protected rhesus macaques and those that develop progressive SIV infection. Thus, it appears such responses are necessary, but not sufficient for protection.

Consistent with this, several lines of evidence suggest Mamu-E-restricted $CD8^+$ T cell responses could be uniquely poised to achieve superior SIV control. SIV-infected $CD4^+$ follicular helper T-cells (TFH), restricted to B-cell follicles constitute an inaccessible, residual site of productive infection in contexts of effective immune- or drug-mediated control such as elite controllers or anti-retroviral responders, respectively (Fukazawa et al., 2015). It is thought conventionally-restricted $CD8^+$ T cells do not typically express permissive chemokine receptor profiles for B-cell follicle entry and are thus largely concentrated in extrafollicular regions of the lymph node (Connick et al., 2007) (Vinuesa and Cyster, 2011). However, recent evidence suggests that a subset of regulatory $CD8^+$ T cells, which were shown to expand during chronic SIV infection in non-human primates, express $CXCR5^{hi}CD44^{hi}$ receptor profiles enabling entry into otherwise inaccessible follicular sanctuaries (Miles et al., 2016). This regulatory $CD8^+$ T cell subset were also shown to induce apoptosis of HIV-infected TFH via interactions with HLA-E *in vitro*. Consistent with these results, elevated Mamu-E surface expression has been reported to occur on SIV-infected TFH cells (Louis Picker, unpublished data). Given that low level SIV replication was detected in peripheral lymphoid tissues following SIV challenge of RhCMV 68-1-SIV vaccinated rhesus macaques that later achieved immune clearance, it is conceivable that protective vaccine-induced Mamu-E-restricted $CD8^+$ T cell responses are capable of

penetrating the B cell follicular barrier and targeting otherwise inaccessible reservoirs of SIV-infected cells.

As previously mentioned, the polyfunctional HIV accessory protein Nef has been shown to target the cytoplasmic tail of MHC class I molecules in an allele-specific fashion resulting in endocytosis, disrupted intracellular trafficking and reduced cell surface expression, facilitating $CD8^+$ T cell evasion (M. Williams et al., 2002). Selective interactions between Nef and HLA-A or HLA-B have previously been described although no such interaction was initially reported for HLA-C or HLA-E (Cohen et al., 1999). Further, enhanced HLA-E surface expression was detected on HIV-positive $CD4^+$ T cells isolated from infected individuals and HIV infection of PBMCs *in vitro* also resulted in upregulated surface levels of HLA-E (Nattermann et al., 2005b) (Martini et al., 2005). Together, such observations are not only consistent with increased Mamu-E surface expression on SIV-infected $CD4^+$ TFH cells and the unprecedented efficacy of SIV-specific RhCMV 68-1-elicited Mamu-E-restricted $CD8^+$ T cell responses, but also with evolutionary selective pressures to avoid NK cell-mediated lysis - a process discussed in greater detail in a subsequent section (Martini et al., 2005).

However, previous studies focusing on interactions between HIV Nef and MHC class I molecules were obtained using cell line-adapted HIV strains or viral isolates from HIV-infected individuals who had progressed to AIDS (Cohen et al., 1999). Conversely, Nef-mediated downregulation of HLA-E cell surface expression has more recently been demonstrated following *in vitro* $CD4^+$ T cell infection with a subset of primary HIV strains (Thans et al., 2019). Interestingly, the extent of HLA-E downregulation induced by certain primary HIV strains decreased in the absence of Vpu, highlighting a potential HLA-E-targeting role for this additional HIV accessory protein. Such observations perhaps indicate differential selection pressures are exerted at various stages of HIV infection - selective pressures which led to combined Nef- and Vpu-mediated HLA-E targeting early in HIV infection, possibly to evade

HLA-E-restricted $CD8^+$ T cell recognition, might have outweighed the selection pressure to maintain HLA-E surface expression for evasion of NK cell-mediated lysis (Martini et al., 2005).

Genetic manipulation of Cytomegalovirus-based vaccine vectors

The asymptomatic nature of CMV infection in the majority of immune-competent individuals potentially renders this β-herpesvirus a clinically suitable vaccine vector (Betts and Hanshaw, 1977). Further, it appears unlikely that high CMV seroprevalence (global estimates of 83%) would compromise vaccine efficacy since RhCMV 68-1 vectors not only demonstrated superinfection in rhesus macaques with pre-existing RhCMV immunity but also pre-exposure to RhCMV did not compromise vector immunogenicity (Scott G. Hansen et al., 2013) (Zuhair et al., 2019). Interestingly, it was found that genetic manipulation of the RhCMV68-1 vector altered the makeup of the vector-elicited immune response. For example, RhCMV Rh189, the orthologue of human CMV (HCMV) US11, was shown to be responsible for the suppression of conventionally-restricted $CD8^+$ T cell responses, with its deletion resulting in a reversal to canonical epitope recognition (Scott G. Hansen et al., 2013). The underlying mechanism for this suppression is currently unknown since the previously established functional role of US11 – a modest downregulation of classical MHC class I molecules – is considered unlikely to be fully accountable (Tirosh et al., 2005).

In addition, expression defects in Rh157.5, Rh157.4 and Rh157.6 (which correspond to HCMV UL128, UL130 and UL131, respectively) present in the RhCMV 68-1 vector strain are crucial for the generation of unconventionally-restricted $CD8^+$ T cells (Scott G. Hansen et al., 2013). Wild-type Rh157.4-6 assemble to form part of an epithelial and endothelial cell tropism-promoting pentameric complex on the surface of the CMV virion (Lilja and Shenk, 2008) (Wang and Shenk, 2005), highlighting the possible importance of the specific host cell range of the vectoring RhCMV 68-1 strain. Consistent with this, it has subsequently been shown that the myeloid cell-tropic nature of the RhCMV 68-1 strain is necessary for the generation of

protective, unconventionally-restricted responses (Louis Picker, Unpublished data). Intriguingly, the observation that defects in Rh157.4-6 unleash unconventional priming of broadly targeting, Mamu-E-restricted $CD8^+$ T cell responses to non-canonical epitopes perhaps suggests these proteins may have originally evolved specifically to suppress such responses - for example by directing viral tropism to cell types that do not typically prime MHC-E-restricted $CD8^+$ T cells.

Further, Rh67 - the functionally analogous RhCMV homologue of the early HCMV protein, UL40 - in the RhCMV 68-1 vector is essential for the generation of SIV-specific Mamu-E-restricted $CD8^+$ T cell responses (Louis Picker, unpublished data). Both Rh67 and UL40 encode peptides within their leader sequences which mimic the natural MHC class I leader sequence-derived peptide ligands (VL9) of Mamu-E and HLA-E, respectively (Tomasec et al., 2000). It was previously demonstrated that UL40 expression regulated HLA-E surface expression in transporter associated with antigen presentation (TAP)-defective cell lines (Tomasec et al., 2000) (Ulbrecht et al., 2000). Thus, in the face of RhCMV-mediated MHC class I downmodulation by the gene products of the Rh182-189 locus (homologues of HCMV US2-4 and 11), Rh67 (or HCMV UL40) provides a compensatory supply of VL9 peptides which restore the refolding, trafficking and cell surface expression of Mamu-E (or human HLA-E), in turn preventing NK cell-mediated lysis (Tomasec et al., 2000) (Wang et al., 2002). Since US6 directly targets TAP-dependent peptide transport into the endoplasmic reticulum (ER), the N-terminal residues of the UL40 protein facilitate TAP-independent localisation into the ER thereby promoting HLA-E surface expression despite virally-mediated TAP inhibition (Lehner et al., 1997). In summary, genetic manipulation of the RhCMV vector background can re-programme the corresponding immune response. It will be important to elucidate the precise underlying mechanisms and pathways that prime broadly-targeted Mamu-E-restricted $CD8^+$ T cell responses which confer protection against SIV infection. Although CMV vaccine vectors may have acceptable safety profiles in immune-competent individuals, CMV infection is known to

cause morbidity in immune compromised individuals and unborn foetuses. Ultimately, following the mechanistic delineation of how RhCMV 68-1 vectors induce such responses, their induction could be recapitulated in a synthetic, more controlled manner mitigating the need to vaccinate with live, replicating persistent virus.

HLA-E-restricted CD8$^+$ T cell responses specific for *Mycobacterial* antigen

Non-classical, HLA-E-restricted CD8$^+$ T cell responses were initially implicated in *Mycobacteria tuberculosis* (Mtb) infection following their identification as the dominant CD8$^+$ T cell response in individuals with latent infection (Heinzel et al., 2002). Although the specific sequences of Mtb-derived HLA-E-restricted epitopes were not identified in these original studies, it was established that such unconventional antigen presentation was mediated independently of TAP, similar to SIV-derived Mamu-E-presented epitopes identified following RhCMV 68-1-SIV vaccination (Lewinsohn et al., 1998) (Hansen et al., 2016). More recently, 69 Mtb encoded peptides were computationally predicted to bind HLA-E *in silico*, via combined bioinformatics approaches (Joosten et al., 2010). A subset of these peptides was shown to induce proliferation of CD8$^+$ T cells isolated from purified protein derivative (PPD)-responsive or Bacillus Calmette–Guérin (BCG)-vaccinated individuals when presented by HLA-E transfected cell lines lacking expression of classical MHC class I molecules. Subsequently, phenotypic characterisation of HLA-E-restricted CD8$^+$ T cells with specificities for several of the 69 Mtb peptides isolated from either actively or latently Mtb-infected individuals, revealed an unusual cytotoxic and immunoregulatory dual functionality (van Meijgaarden et al., 2015). These multifunctional CD8$^+$ T cells not only expressed Th1 cytokines such as IFN-γ and cytolytic effector molecules such as perforin, granulysin and granzyme A/B, but also expressed the Th2-associated lineage-determining transcriptional regulator, GATA3, which promoted the production of Th2 cytokines including IL-4, 5, 10 and 13 upon peptide stimulation (Prezzemolo et al., 2018) (Yagi et al., 2011). Further, anti-mycobacterial intracellular growth activity was

demonstrated for this unconventionally-restricted T cell subset and was independent of CD94/NKG2 class receptor ligation indicating effector functions were presumably mediated via the T cell receptor (TCR) (van Meijgaarden et al., 2015). However, as previously mentioned Mtb-encoded HLA-E-presented epitopes recognised in these studies were originally identified using algorithms based on formerly established binding motifs for HLA-E, thus these peptide sequences are less diverse and more canonical relative to Mamu-E-restricted SIV-derived epitopes characterised in RhCMV 68-1-SIV vaccine studies (Joosten et al., 2010) (Hansen et al., 2016).

Naturally processed and presented Mtb-encoded HLA-E-restricted epitopes were subsequently identified in a physiological setting. An HLA-E-restricted $CD8^+$ T cell clone isolated from an individual with latent Mtb infection was found to recognise a post-translationally modified peptide derived from the N-terminal of the Mtb glycoprotein protein, MPT32 (Rv1860) (Harriff et al., 2017). N-terminal O-linked mannosylation of MPT32, catalysed by an Mtb expressed mannosyltransferase encoded by the Rv1022c gene, was a prerequisite for HLA-E-restricted $CD8^+$ T cell recognition illustrating, for the first time, HLA-E presentation of a glycosylated peptide - this phenomenon has previously been reported for classical MHC class I and II molecules primarily in the context of aberrantly glycosylated tumour-associated antigen (Galli-Stampino et al., 1997) (Haurum et al., 1999). It was previously shown that O-β-linked GlcNAc (N-acetylglucosamine) monosaccharide glycosylation of cytosolic peptides did not disrupt TAP-dependent transport into the endoplasmic reticulum permitting MHC class I presentation of such glycosylated peptides (Haurum et al., 1999). However, the TAP-independent nature of Mtb- and SIV-derived antigen presentation by HLA-E and Mamu-E in the settings of *Mycobacterial* exposure or infection and RhCMV 68-1-SIV vaccination, respectively, might permit the presentation of peptide that has been subject to additional types of post-translational modification which would otherwise be incompatible with TAP-mediated transport (Lewinsohn et al., 1998).

Additionally, 28 novel HLA-E ligands derived from 13 Mtb encoded proteins - largely comprising secreted antigen or membrane-bound proteins - were identified via mass spectrometry following elution from purified HLA-E*0103 secreted from Mtb-infected cells (McMurtrey et al., 2017). Nine such peptides were recognised by TCRαβ$^+$ CD8$^+$ T cells, as evidenced by IFN-γ production. A lack of recognition of the remaining 19 peptides might be due to the limited number of screened donors. Alternatively it is possible that such peptides were recognised by Mtb-specific HLA-E-restricted CD8$^+$ T cells displaying the alternative, previously described, Th2-like phenotype for which IFN-γ readouts were inappropriate (van Meijgaarden et al., 2015). One of the 28 peptides, comprising residues 19-29 of the conserved hypothetical protein Rv0634A (EIEVDDDLIQK), demonstrated an immunodominant profile with responses detected in 81% of tested donors with active or latent Mtb infection. HLA-E-restriction of the immunodominant response was subsequently confirmed via peptide-specific HLA-E tetramer staining. Given that such peptides were identified in the physiological context of Mtb infection, this study provided insight into the naturally presented and processed HLA-E ligandome. Interestingly, the 28 HLA-E-restricted Mtb-encoded ligands deviated in both sequence and length from canonical MHC class I leader sequence peptides which dominate the HLA-E repertoire in the steady state (Braud et al., 1997). Such peptides comprised 8 to 20 amino acids in length and a small subset contained Met oxidation or N-terminal acetylation modifications.

Despite detectable anti-mycobacterial cytotoxic effector functions (van Meijgaarden et al., 2015), the relative contribution of HLA-E-restricted CD8$^+$ T cells in the complex, multifaceted immune response to Mtb remains to be elucidated. In support of a significant protective role, β-2 microglobulin (β2M) knockout mice exhibit greater susceptibility to Mtb infection than mice with defective classical MHC class I alleles (H2-K^b/H2-D^b double knockouts) (Rolph et al., 2001). Further, as previously mentioned, HLA-E-restricted CD8$^+$ T cells constitute the dominant Mtb-specific CD8$^+$ T cell response in latently infected individuals (Heinzel et al., 2002).

Additionally, although Mtb-mediated downmodulation of classical MHC class I molecules may compromise classically-restricted $CD8^+$ T cell recognition, HLA-E has been found to be enriched in Mtb-infected phagosomes (Grotzke et al., 2009). Upregulated HLA-E expression has also been reported to occur upon monocyte to macrophage differentiation, with the majority of newly synthesised β2M-associated HLA-E complexes detected in LC3- or LAMP1-positive autophagosomal-lysosomal vesicles where they are exposed to exogenous antigenic peptides (Camilli et al., 2016). Similar to HLA-E enrichment in Mtb-infected phagosomes (Grotzke et al., 2009), HLA-E upregulation in newly differentiated macrophage is uncoupled from classical MHC class I protein expression and does not correlate with superior evasion from NK cell-mediated lysis, indicating a potentially distinct function for autophagosomally-sequestered HLA-E (Camilli et al., 2016). Together, these results indicate that under certain physiological conditions, HLA-E exhibits an alternative form of trafficking which appears to intersect with the MHC class II pathway (Neefjes et al., 2011). Specifically, under the unusual conditions of RhCMV 68-1 vaccination or Mtb infection, when the classical MHC class I trafficking pathway is disrupted by virally-encoded immunomodulatory proteins and the supply of MHC class I leader peptide is concomitantly reduced, possible re-routing of HLA-E to peripheral intracellular compartments may permit access to exogenous antigen in a TAP-independent manner (Grotzke et al., 2009). It is clear that both Rh157.4-6-defective strains of RhCMV and Mtb possess intrinsic properties which trigger unconventional MHC-E-restricted $CD8^+$ T cell responses (Scott G. Hansen et al., 2013) (Joosten et al., 2010) (McMurtrey et al., 2017). However, elucidating the underlying mechanisms of this unconventional T cell priming pathway will prove essential in enabling its manipulation in therapeutic or vaccine settings.

The original characterisation and the primary functional role of HLA-E

HLA-E is a member of the non-classical MHC class Ib family in addition to HLA-G and HLA-F. HLA-G has a well-established role in mediating maternal tolerance to the foetus with its

placental extravillous cytotrophoblast cell-restricted expression and ligation with the universally NK cell-expressed inhibitory receptor, KIR2DL4 (killer cell immunoglobulin-like receptor) (Rajagopalan and Long, 1999) (Ferreira et al., 2017). However, the function of HLA-F has remained more elusive, although recent evidence points toward a putative role in regulating NK cell function via interactions with open MHC class I conformers and certain KIRs, including KIR3DS1 (Burian et al., 2016) (Lin and Yan, 2019). Other, non-classical MHC class I-like proteins exist in humans such as the non-MHC-encoded CD1 molecules which fulfil specialised roles in presenting lipid antigen (Beckman et al., 1994) (Sieling et al., 1995). In contrast to classical MHC class Ia molecules which mediate adaptive immune responses involving $CD8^+$ cytotoxic T lymphocytes (CTL) (Townsend et al., 1985), non-classical MHC class I molecules are typically characterised by innate-like effector functions (Rodgers and Cook, 2005).

Similar to classical MHC class I molecules, the HLA-E heavy chain non-covalently associates with the β2-microglobulin invariant light chain and a peptide ligand - such heterotrimeric complex formation is dependent on a suitable peptide providing sufficient binding energy to the refolding reaction and is a prerequisite for intracellular trafficking and peptide presentation at the cell surface (Ulbrecht et al., 1992). Classical MHC class I molecules exhibit unprecedented polymorphism, with over 1000 alleles identified at the HLA-A, HLA-B and HLA-C loci which have evolved by overdominant selection (Hughes and Nei, 1988). The 9 overarching 'supertypes', into which the majority of classical MHC class I molecules can be assigned, are characterised by similar residue profiles in the peptide binding domain which in turn yield overlapping peptide binding motifs (Sette and Sidney, 1999) (Sidney et al., 2008). Polymorphic residues are predominantly concentrated within the peptide binding groove (α1 and α2 domains of the MHC class I heavy chain) with the degree of inter-allelic variation typically ranging from a single substitution, denominated as 'micro-polymorphism', to around 30 residues (Parham et al., 1988). Importantly, polymorphic residues of the peptide binding

domain dictate the peptide binding preferences of the MHC class I allele, with differential usage of typically two to three of six binding pockets (A-F) within the groove which selectively accommodate complementary peptide anchor side chains (Bjorkman, 1990) (Macdonald et al., 2003). Binding pockets vary in stereochemistry between allelic variants according to the nature of the pocket-lining residues and compatible peptide anchor side chains can be accommodated depending on shape complementarity, electrostatic, hydrophobic or van der Waals interactions and hydrogen bonding (Johansen et al., 1997) (Zhang et al., 1998). Thus, extensive polymorphism at MHC class I loci ensures a large range of distinct inter-allelic peptide binding repertoires are presented, providing broad coverage at the population level of the diverse range of antigenic peptide encoded by human pathogens.

By contrast, limited allelic polymorphism distinguishes non-classical MHC class Ib molecules from their classical counterpart molecules with the HLA-E locus displaying the lowest degree of polymorphism of any non-classical MHC class I molecule in humans. The HLA-E locus is dimorphic with only two non-synonymous functional alleles identified in human populations - HLA-E*0101 and HLA-E*0103 - which differ by a single Arg or Gly amino acid residue, respectively, at position 107 on a surface loop connecting β-strands of the peptide binding groove (Geraghty et al., 1992) (Grimsley et al., 2002) (O'Callaghan et al., 1998). The two HLA-E allelic variants are maintained in ethnically diverse populations at frequencies which are considered too similar to reflect a neutrally evolving gene (Geraghty et al., 1992) (Grimsley and Ober, 1997). Thus, balancing selection is thought to operate at the HLA-E locus, positioned on the short arm of chromosome 6, which implies that the two allelic variants perform functionally distinct roles. Although similar intracellular proteins levels and peptide binding preferences have previously been described for the two HLA-E alleles, HLA-E*0103 exhibits higher cell surface expression - a property likely afforded by superior peptide binding affinity and greater resultant thermal stability relative to HLA-E*0101 for all VL9 variant peptides tested (Ulbrecht et al., 1999) (Strong et al., 2003). Despite superior thermal stability compared

to HLA-E*0101, peptide-refolded HLA-E*0103 exhibits melting temperatures (T_ms) (≤52°C) that correspond to the low end of the range previously reported for classical MHC class I molecules (T_m range of 52-61°C) (Fahnestock et al., 1992) (Morgan et al., 1997) (Strong et al., 2003).

Importantly, minimal polymorphism at the HLA-E locus (Geraghty et al., 1992) in addition to similar peptide binding motifs of the two functional alleles HLA-E*0101 and HLA-E*0103 (Strong et al., 2003) makes HLA-E a particularly attractive HIV vaccine target since protective responses would be universally effective across the population. Given that superior HIV control in LTNPs or elite controllers is associated with the presence of specific classical MHC class I alleles such as HLA-B*5701 (Migueles et al., 2000) and HLA-B*27 (Carrington and O'Brien, 2003), it is predicted that potential protective benefits of conventional vaccination strategies that elicit classically-restricted $CD8^+$ T cell responses, would be confined to individuals with a particular MHC class I allotype.

Given the protective SIV-specific Mamu-E-restricted $CD8^+$ T cell responses elicited by RhCMV 68-1-SIV vaccination of rhesus macaques, deciphering the precise degree of functional similarity between the human and rhesus MHC-E homologues will be critical in determining the potential efficacy of an HLA-E-targeted universal HIV vaccine in humans. Although the two human alleles HLA-E*0101 and HLA-E*0103 differ by a single residue at position 107, a greater degree of sequence diversity is observed between the more numerous (22) Mamu-E alleles (Wu et al., 2018). However, heavy chain residues forming the pockets of the peptide binding domain are highly conserved between HLA-E and Mamu-E (Walters et al., 2018), indicating that these homologues are likely capable of accommodating similar peptide repertoires. Accordingly, both HLA-E and Mamu-E were equally capable of presenting identical SIV epitopes and stimulating Mamu-E restricted $CD8^+$ T cells isolated from RhCMV 68-1 vaccinated macaques, further illustrating the functional equivalence of the human and rhesus homologues (Wu et al., 2018).

Similar to classical MHC class I molecules and in contrast to its counterpart non-classical MHC class I molecules, HLA-E exhibits a broad tissue distribution and is present on the surface of nearly all human cell types including placental and trophoblastic cells (Ulbrecht et al., 1992) (Ni Lee et al., 1998), albeit with lower relative surface expression levels than classical MHC class Ia molecules (V. M. Braud et al., 1998). In addition, it was found that HLA-E follows a similar trafficking pathway to that adopted by classical MHC class Ia molecules and interacts in a similar manner with the multiple ER-resident chaperones of antigen presentation; both TAP- and tapasin-dependent peptide loading in the endoplasmic reticulum correlated with HLA-E surface expression (V. M. Braud et al., 1998). TAP and tapasin constitute fundamental components of the MHC class I trafficking pathway. TAP transports proteasomally degraded cytosolic peptide fragments into the ER lumen and associates with the MHC class I-specific chaperone, tapasin (Neefjes et al., 2011). Tapasin, for which a spectrum of dependency has been established for different MHC alleles, imparts its chaperoning effect by stabilising high energy peptide-free MHC class I transition states (Grandea and Van Kaer, 2001). Subsequent tapasin-mediated peptide exchange yields optimally loaded peptide-MHC complexes which are then released from the ER-based multi-molecular peptide loading complex for cell surface presentation (A. P. Williams et al., 2002). Calreticulin - a lectin-like ER-resident chaperone which stabilises the interaction between MHC class I molecules and tapasin in a glycan-dependent manner and forms part of the peptide loading complex in addition to TAP, tapasin and its covalently associated thiol oxidoreductase ERp57 (Wearscha et al., 2011) - has also been demonstrated to interact with HLA-E (V. M. Braud et al., 1998).

Regarding the peptide binding repertoire, it was shown that HLA-E binds signal peptides comprising residues 3-11 of MHC class I leader sequences, with the exception of HLA-F and HLA-E which possess N-terminally truncated leader sequences which do not generate HLA-E-binding ligands (Braud et al., 1997) (N Lee et al., 1998). Crucially, loading of MHC class I signal peptides correlated with the dissociation of HLA-E from TAP, while the availability of MHC

class I signal peptides correlated with HLA-E surface expression (V. M. Braud et al., 1998). Although MHC class I leader sequence derived peptides, extended by one or two residues from the naturally flanking N- or C-terminal regions also bound to HLA-E, they exhibited lower binding affinities relative to the optimal nonameric counterpart peptides, highlighting the preference of the HLA-E peptide binding groove for nonapeptides (Braud et al., 1997). Together, these results demonstrate the dependence of HLA-E assembly and trafficking on the supply of MHC class I signal peptides and imply that the HLA-E peptidome is likely largely restricted to a limited number of highly conserved nonameric leader peptide variants.

The HLA-E peptide binding sequence motif was initially explored via Ala and Gly substitution experiments along the HLA-B*0801 leader peptide (VMAPRTVLL), for which strong sequence selectivity was uncovered at primary anchor positions 2 (Met) and 9 (Leu), with weaker preferences observed for position 7 (Val) and 3 (Ala) (Braud et al., 1997). The primary anchor residue at position 2 appeared particularly critical to the HLA-E peptide binding motif with both Ala and Gly substitutions abrogating binding of the HLA-B*0801 leader peptide to HLA-E (Braud et al., 1997). Consistent with these findings, a Met to Thr substitution present at position 2 in a subset of HLA-B27 (VTAPRTLLL) and HLA-B15 (VTAPRTVLL) leader peptides was shown to diminish HLA-E cell surface expression due to reduced peptide affinity and complex stability (N Lee et al., 1998) (Brooks et al., 1999). Such results are similar to those obtained for the murine homologue of HLA-E, Qa-1, which exhibited a strong, TAP-dependent binding preference for the Qa-1 determinant modifier (Qdm) peptide nonamer (AMAPRTLLL) derived from the signal sequence of murine classical MHC class I molecules, H-2D and H-2L (Aldrich et al., 1994) (DeCloux et al., 1997). The murine Qdm peptide exhibits striking sequence conservation with human MHC class I leader sequence peptides and Qdm substitution experiments revealed that, analogous to HLA-E, Qa-1 has a preference for Leu at position 9 of the peptide, highlighting the similar peptide binding motifs evolved by the murine and human orthologues (Cotterill et al., 1997) (Braud et al., 1997).

In physiological conditions VL9 and Qdm peptides appear to dominate the HLA-E and Qa-1 repertoires respectively (Cotterill et al., 1997) (McMurtrey et al., 2017), a finding supported by peptide scanning substitutions which delineated a shared optimal peptide binding motif for the human and murine homologues (Braud et al., 1997) (Kraft et al., 2000). However, various lines of evidence support a capacity for the human and murine orthologues to bind a range of suboptimal, sequence diverse peptides. Similar results were obtained via *in vitro* refolding of either HLA-E (Miller et al., 2003) or Qa-1 (Kraft et al., 2000) with respective pools of VL9 or Qdm primary and secondary anchor residue mutant peptides followed by acid elution and mass spectrometric analysis of the bound peptide. Both studies identified similar patterns of promiscuity for suboptimal VL9- and Qdm-based peptide binding to HLA-E and Qa-1, respectively. Notably, HLA-E and Qa-1 tolerated a diverse range of amino acids at the secondary anchor position 3, including a positively charged Lys which minimally impacted binding affinity to HLA-E. Additionally, a position 2 Met to Leu or Ile substitution was well tolerated by HLA-E and Qa-1, with Leu at position 2 of Qdm/VL9 permitting Qa-1/HLA-E binding with similar affinity relative to the wild-type peptide.

It was subsequently demonstrated that peptide-HLA-E complexes constitute major ligands for heterodimeric inhibitory CD94-NKG2A/B and activating CD94-NKG2C/E receptors, predominantly expressed on NK cells (V. M. Braud et al., 1998) (Ni Lee et al., 1998) (Borrego et al., 1998) (Kaiser et al., 2005). Engagement of leader peptide-HLA-E complexes with the inhibitory receptor CD94/NKG2A was shown to confer protection against NK cell-mediated lysis and the nature of this interaction was influenced by the stability of the HLA-E complex in addition to the sequence of the bound peptide (Brooks et al., 1999) (Llano et al., 1998). Thus, considering HLA-E surface expression in the steady state is both dependent on TAP-mediated transport and a supply of MHC class I leader peptides, HLA-E not only functions to report the downregulation of MHC class I molecules to NK cells via CD94/NKG2 receptor interaction, but it also relays information regarding the functional integrity of TAP. Such signalling enables the

detection and elimination of virally infected or malignant cells since MHC class I molecules, in addition to components of the antigen presentation pathway such as TAP, are commonly targeted by virally-encoded immunomodulatory proteins or by the selective pressures for immune evasion in tumour evolution (Barnes and Grundy, 1992) (Früh et al., 1995) (Petersen et al., 2003) (Vinay et al., 2015).

The optimal Qdm nonamer was found to rapidly dissociate from Qa-1 (half-life of 1.5 hours at 37°C) compared to the slower dissociation kinetics observed for classical MHC class I molecules bound to natural peptide ligands (half-lives of up to 31 hours at 37°C) (Kambayashi et al., 2004). Such results are corroborated by low thermal stabilities of HLA-E refolded with VL9 peptide variants relative to the melting temperatures of peptide-bound classical MHC class I molecules (Fahnestock et al., 1992) (Morgan et al., 1997) (Strong et al., 2003). Interestingly, it has been postulated that such features might be integral to the optimal function of the human and murine homologues; since HLA-E and Qa-1 report the intracellular integrity of the antigen presentation pathway, prolonged complex stability at the cell surface would serve to delay the detection of classical MHC class I or TAP downmodulation. Thus, low affinity peptide binding and rapid dissociation kinetics may have evolved for HLA-E and Qa-1 to ensure peptide presentation at the cell surface is an accurate and current reflection of intracellular conditions.

Structural insights into MHC class I leader sequence-derived peptide-bound HLA-E

The first crystal structure of HLA-E (1MHE) was solved in complex with the HLA-B7/8/14/39/42/48-derived leader sequence signal peptide, VMAPRTVLL, to 2.85 Å (Figure 1) (O'Callaghan et al., 1998). Subsequent structures of HLA-E were determined in complex with highly conserved VL9 peptide variants derived from the leader sequences of HLA-B27 (VTAPRTLLL) (Strong et al., 2003), HLA-G*01 (VMAPRTLFL) (Hoare et al., 2008) and HLA-Cw*07 (VMAPRALLL) (Hoare et al., 2008). Pairwise alignment of HLA-E molecules from within the

asymmetric unit or between structures revealed minimal repositioning (global rmsd <1Å) with the greatest adjustment, which is thought to be attributable to crystal packing contacts, concentrated in the surface loop region between residues 216-230 of the heavy chain α3 domain (220's loop) – a region which also exhibits flexibility in structures of classical MHC class I molecules and corresponds to the CD8 co-receptor main interacting region (Gao et al., 2000) (Strong et al., 2003).

The structural architecture of the HLA-E complex was found to resemble that of classical MHC class I molecules (Bjorkman et al., 1987a) (Bjorkman et al., 1987b). For example, perhaps unsurprisingly given that 76% of the extracellular domains of HLA-E (1MHE) and HLA-A2 exhibit conserved sequence identity, minimal pairwise root mean square deviation (<1Å) was observed following α-carbon superposition of the two MHC class I molecules (O'Callaghan et al., 1998). Analogous to classical MHC class I molecules, the α1 and α2 domains of the HLA-E heavy chain formed a peptide binding groove comprising two helices positioned either side of the bound peptide with an underlying floor consisting of eight anti-parallel β-strands (Bjorkman et al., 1987a) (O'Callaghan et al., 1998). The α3 domain and non-covalently associated β2M invariant chain form β-sandwich immunoglobulin-like folds on the membrane-proximal side of the binding groove - the intracellular and transmembrane regions were removed prior to crystallisation.

Structural determination of both HLA-E allelic variants, HLA-E*0103 (Gly-107) and HLA-E*0101 (Arg-107), in complex with the same VL9 peptide (VMAPRTVLL) did not provide insight into their postulated differential functional roles implied by balancing selection (Strong et al., 2003). Despite the loss of a single hydrogen bond in HLA-E*0103, formed between Arg-107 and His-3 in the HLA-E*0101 variant, minimal inter-allelic structural repositioning (>0.8Å) was observed for α-carbons in the flexible surface loop region (residues 104-109) containing the only mismatched residue. Notably, α-carbons of the position 104-109 loop region were

separated by greater distances following alignment of molecules from the same asymmetric unit of the same allelic variant - such movements likely arose from crystal packing contacts between this loop region and the α3 domain of a neighbouring symmetry-related molecule. The loss of a hydrogen bond at position 107 of the HLA-E*0103 structure might suggest inferior stability, however contrary to such predictions and as previously discussed, HLA-E*0103 variant complexes consistently demonstrate superior thermal stability relative to HLA-E*0101 complexes for every VL9 peptide ligand tested (Strong et al., 2003).

The binding configuration of the VL9 peptide in the HLA-E binding groove resembles the prototypical kinked conformation adopted by the majority of 9 amino acid peptides bound to classical MHC class I molecules (Madden, 1995) (O'Callaghan et al., 1998) (Strong et al., 2003) (Hoare et al., 2008). A superposable conformational motif was also adopted by Qdm peptides bound to the murine HLA-E orthologues, Qa-1a and Qa-1b (Zeng et al., 2012) (Ying et al., 2017). Analogous to peptides bound to classical MHC class I molecules, the canonically orientated VL9 peptide runs parallel to the HLA-E α1 helix and exhibits anchored amino and carboxy termini, with a more loosely constrained central region. Given that the length of the MHC class I binding groove is insufficient to accommodate a fully extended linear nonameric peptide ligand and that the ends of the MHC class I binding groove are occluded by the conserved bulky side chains of Tyr-84 and Trp-167, a kink in the peptide main chain at a central position is typically formed (Madden, 1995). The centrally positioned peptide kink region with increased conformational malleability is in part permitted by a preserved network of hydrogen bonds mediated by highly conserved clusters of heavy chain residues present at the termini of the binding groove, which primarily form bonds with main chain atoms at the peptide termini rendering the central peptide region subject to fewer constraints (Madden, 1995) (O'Callaghan et al., 1998). The kink following the VL9 position 3 α-carbon projects the peptide main chain away from the floor of the binding groove yielding greater conformational flexibility and solvent exposure of the central peptide residues at positions 4 and 5

(O'Callaghan et al., 1998). Side chains of positions 1 and 8 of VL9 peptides also consistently project toward the solvent and are thus potential candidates for receptor interaction in addition to the peptide side chains within the central solvent-accessible kinked region.

Peptide binding motifs for classical MHC class I molecules typically involve the accommodation of two to three anchor residue side chains within respective groove pockets, however, HLA-E-bound VL9 displayed an unusually high density of pocket usage (Madden, 1995) (O'Callaghan et al., 1998) (Strong et al., 2003) (Hoare et al., 2008). Not only did primary anchor side chains at positions 2 and 9 project into their respective B and F pockets, but positions 3, 6 and 7 also constitute secondary anchor residues which project side chains into the shallow D, C and deeper E pockets, respectively (Figure 1). As previously mentioned, pocket-anchor side chain shape complementarity in addition to specific pocket-based contacts such as hydrogen bonding, electrostatics and van der Waals give rise to preferential pocket binding to specific amino acids (Madden, 1995). Thus, sequence restrictions are imposed on the peptide ligand at such anchor positions, reducing the overall magnitude of the potential peptide binding repertoire. Accordingly, the more promiscuous yet high affinity peptide binding characteristic of classical MHC class I molecules is thought to be attributable to the limited number of anchor residues, thus relieving sequence restrictions at remaining positions along the peptide, with certain MHC class Ia alleles having been reported to present thousands of distinct peptide ligands (Madden, 1995). It is therefore conceivable that the requirement to occupy a greater number of pockets, as observed in VL9-bound structures of HLA-E, would correlate with a less diverse peptidome (O'Callaghan et al., 1998).

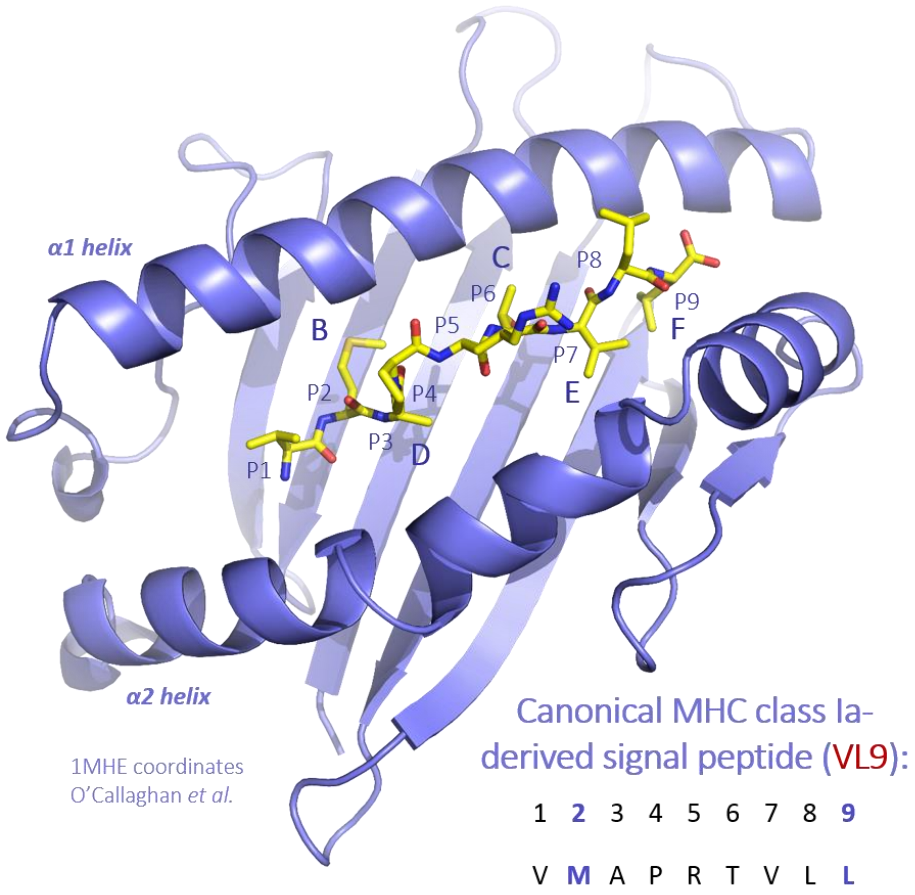


Figure 1: Canonical VL9 peptide accommodation by the HLA-E peptide binding groove

A PyMol-generated visualisation of the HLA-E peptide binding groove bound to the HLA-B7-derived leader peptide (VMAPRTVLL) from 1MHE - the first published structure of HLA-E (O'Callaghan et al., 1998). The groove comprises two α-helices derived from the α1 and α2 domains of the HLA-E heavy chain with a β-sheet floor to which both domains contribute (depicted in purple cartoon). The canonical MHC class Ia-derived leader sequence signal peptide, VL9, is displayed in yellow stick-form with residues at positions 1-9 labelled as PX. The positioning of primary B and F pockets in addition to secondary D, C and E pockets within the groove is indicated next to the corresponding anchor side chain. The α3 domain of the HLA-E heavy chain and the invariant light chain, β2M, were omitted for clarity.

A deep, hydrophobic B pocket that accommodates the highly conserved peptide position 2 Met primary anchor side chain is present in the HLA-E peptide binding groove – the B pocket-lining residue, Met-45, confers hydrophobicity whereas the small Ala-67 bestows depth

(O'Callaghan et al., 1998). Interestingly, minimal structural repositioning is observed in the B pocket when the preferred Met anchor residue is substituted to Thr in the HLA-B27 leader peptide (Strong et al., 2003). The reduced quality of fit of the smaller Thr side chain combined with a lack of compensatory structural readjustment result in the expansion of a B pocket-based cavity from 31 to 53 Å - a phenomenon that has previously been linked to complex instability, consistent with reduced thermal stability and surface expression of HLA-E refolded with the HLA-B27-derived signal peptide (Xu et al., 1998). Such structural features are in line with recent molecular dynamics simulations which predict the HLA-E binding groove retains a greater degree of structural integrity than HLA-A2 in the absence of a bound peptide ligand (Hansen et al., 2016).

The structural rigidity of the HLA-E peptide binding groove is unusual among its classical MHC class I counterpart molecules, in which groove lining residues frequently exhibit structural malleability to optimally accommodate peptide residue variation (Madden, 1995) (Kjer-Nielsen et al., 2002) (Badrinath et al., 2012). For example a Gly to larger Phe substitution at position 1 of an HLA-B8-bound peptide results in dramatic conformational adjustments of pocket lining residues; not only is the HLA-B8 Arg-62 side chain displaced from the A pocket but also Trp-167 is reoriented 1.9 Å higher in the A pocket to accommodate for the larger Phe side chain (Kjer-Nielsen et al., 2002). Similarly, a Gly to Leu substitution at position 2 of the HLA-B8-bound peptide displaces the B pocket lining residue Phe-67, which instead adopts an alternative rotamer projecting away from the pocket cavity permitting the accommodation of the larger position 2 Leu side chain. Further, Phe-67 repositioning has knock-on structural ramifications in the HLA-B8 binding groove, triggering the readjustment of Phe-22, Phe-36 and Glu-45 which in turn forces the reorientation of Arg-75, to avoid steric clashes between repositioned side chains.

The shallow nature of the peptide position 3-accommodating D pocket in the HLA-E binding groove is conferred by the large pocket-forming residue, Tyr-159 (O'Callaghan et al., 1998). The non-polar Trp-97 of the HLA-E heavy chain further reduces the size of the D pocket and gives rise to a preference for small, hydrophobic secondary anchor residues at position 3 of the peptide, consistent with a small, highly conserved Ala at position 3 of VL9 peptides. Such structural features are largely unique to HLA-E since position 97 in classical MHC class I molecules is usually polar, projects away from the D pocket and is thus not typically involved in imposing restraints on D pocket binding (Madden, 1995). Further, residues Trp-97 and Phe-116 of the HLA-E heavy chain form a stack in the peptide binding groove conferring hydrophobicity to both the shallow, solvent accessible peptide position 6-accommodating C pocket and the deeper peptide position 7-accommodating E pocket (O'Callaghan et al., 1998). Similar to the structural rigidity of the HLA-E B pocket in response to position 2 primary anchor variation, substitution at position 6 of the HLA-Cw*07 leader peptide, from a highly conserved Thr to Ala, does not trigger rearrangement of C pocket-lining residues to compensate for the smaller side chain (Hoare et al., 2008). Further, minimal rearrangement of E pocket-lining residues ensued following Leu to Val substitution at position 7 of the HLA-E-bound VL9 peptide (Strong et al., 2003). Multiple F pocket-forming residues such as Leu-81, Leu-95 and Leu-124 of the HLA-E heavy chain confer depth and hydrophobicity generating a preference for the highly conserved hydrophobic Leu at position 9 of VL9 peptides (O'Callaghan et al., 1998). In summary, the unique composition of pocket-forming residues in the HLA-E binding groove confers preferential binding to the hydrophobic primary and secondary anchor residues conserved among VL9 peptide variants.

Both HLA-E and the murine homologue Qa-1 lack residues that are highly conserved among classical MHC class I heavy chains at positions 143 and 147 - namely Thr-143 and Trp-147 are both substituted to Ser in HLA-E and Qa-1 (Braud et al., 1997). Consequently, the reduced volume occupied by the smaller substituted Ser side chains enlarges the E pocket in the HLA-E

and Qa-1 binding grooves (O'Callaghan et al., 1998) (Ying et al., 2017). This in turn permits secondary anchor peptide position 7 side chain occupancy of the discrete E pocket recess which is typically less accessible and obstructed by the larger Thr-143 and Trp-147 side chains present in classical MHC class I molecules (O'Callaghan et al., 1998). Consequently, HLA-E- and Qa-1-bound peptides are positioned deeper in the groove at position 7 relative to those bound to classical MHC class I molecules (Zeng et al., 2012). Additionally, Trp-97, which contributes to E pocket hydrophobicity in HLA-E and replaces a polar residue present in the majority classical MHC class I molecules, is also present in Qa-1, highlighting the unique structural properties of this region of the binding groove for the human and murine homologues (Zeng et al., 2012) (Ying et al., 2017). The enlarged, well-defined E pocket of the HLA-E binding groove is consistent with previous Ala and Gly substitution experiments along the VL9 peptide which identified strong sequence selectivity for peptide position 7 (Braud et al., 1997).

The aforementioned Trp-147 to Ser substitution in HLA-E and Qa-1 also results in the loss of a conserved hydrogen bond formed between the heavy chain and main chain of peptide position 8 in structures of classical MHC class I molecules (Cotterill et al., 1997) (O'Callaghan et al., 1998) (Madden et al., 1993). Despite this loss, an extensive hydrogen bonding network secures the VL9 peptide main chain within the HLA-E peptide binding groove; compensatory hydrogen bonds are formed between heavy chain Asn-77 and position 7 and 9 of the peptide main chain and heavy chain Gln-156 and the main chain of peptide position 5 (O'Callaghan et al., 1998). In summary, structural features that appear to impose sequence restraints on HLA-E-binding peptides, including heavy pocket usage and extensive peptide-heavy chain hydrogen bonding, are consistent with the targeted functional role of HLA-E in presenting MHC class I-derived VL9 peptides to CD94/NKG2 class receptors to evade NK cell-mediated lysis (V. Braud et al., 1998) – a role which would presumably require a restricted and specific peptidome. Thus, it is conceivable that HLA-E evolved to accommodate a limited number of peptides and

that the steady state HLA-E peptidome predominantly comprises MHC class I leader sequence VL9 peptides (O'Callaghan et al., 1998).

Insights from ligand-receptor co-complex crystal structures involving HLA-E

Previous studies illustrated that substitutions in HLA-E-binding VL9 peptides at positions 5, 6 and 8 influenced interactions with the C-type lectin-like CD94/NKG2 class receptors (Llano et al., 1998) (Miller et al., 2003). Strikingly, VL9 peptide position 5 Arg to Lys substitution abrogated the interaction between HLA-E and CD94/NKG2A (Miller et al., 2003). Additionally, substitution of position 6 from a highly conserved Thr to Ala in the HLA-Cw*07 leader peptide (VMAPRALLL) significantly reduced the affinity of HLA-E-CD94/NKG2A binding resulting in a failure of this HLA-E-peptide complex to prevent NK cell-mediated lysis when expressed at physiological levels (Kaiser et al., 2005) (Valés-Gómez et al., 1999). By contrast, the naturally occurring substitution of the largely conserved Leu to Phe at position 8 of the HLA-E-binding HLA-G*01 leader peptide (VMAPRTLFL) yielded the highest affinity interactions with CD94/NKG2A of any VL9 peptide tested (K_D=1μM) (Kaiser et al., 2005) (Valés-Gómez et al., 1999). Structural determination of HLA-E in complex with the HLA-Cw*07 and HLA-G*01 leader peptides revealed that the α-carbon at position 6 of the HLA-Cw*07 peptide resides 1 Å lower in the C pocket of the HLA-E binding groove relative to that of the HLA-G*01 signal peptide, rendering it more solvent inaccessible and potentially accounting for reduced affinity interactions with CD94/NKG2A (Hoare et al., 2008).

A co-complex crystal structure was subsequently solved comprising HLA-G*01 leader peptide-loaded HLA-E bound to the CD94/NKG2A receptor (Petrie et al., 2008) (Kaiser et al., 2008). The interaction involved minimal structural repositioning of HLA-E or the CD94/NKG2A receptor relative to the non-ligated structures, consistent with the previously documented nature of innate-like receptor-ligand engagement (Hoare et al., 2008) (Sullivan et al., 2007). The binding interface was characterised by strong charge complementarity, a large buried area (~2100 $Å^2$),

extensive hydrogen bonding but poor shape complementarity - especially poor for the NKG2A subunit (Hoare et al., 2008). Interfacing contacts were predominantly formed with the invariant CD94 subunit which exhibited a larger footprint not only on HLA-E but also on the bound HLA-G*01 leader peptide. Consistent with previous studies which implicated peptide positions 5, 6 and 8 to form key contacts in the CD94/NKG2A interface (Kaiser et al., 2005) (Valés-Gómez et al., 1999), a multitude of interactions were formed between the CD94 subunit and positions 5, 6 and 8 of the HLA-E-bound peptide, with the position 5 Arg side chain projecting into a cavity between the CD94 and NKG2A subunits (Petrie et al., 2008).

As previously mentioned, the HCMV UL40 protein encodes a peptide mimic, VMAPRTLIL, which is identical to the HLA-Cw03 leader peptide (although modest sequence variation is observed among HCMV strains), and acts as a TAP-independent surrogate for VL9 peptides facilitating evasion from NK cell-mediated lysis in the face of virally-mediated MHC class I downregulation and TAP inhibition (Tomasec et al., 2000) (Ulbrecht et al., 2000). In CMV seropositive individuals where the MHC class I genotype encodes VL9 peptide sequences that are mismatched from the UL40-derived mimic peptide, UL40-specific HLA-E-restricted TCR$\alpha\beta^+$ $CD8^+$ T cell responses have been detected; such responses are present at similar frequencies to classically-restricted $\alpha\beta$ $CD8^+$ T cell responses (Hoare et al., 2006) and have been shown to kill CMV-infected cells (Mazzarino et al., 2005).

Unsurprisingly, co-complex crystal structures of $\alpha\beta^+$ TCRs bound to HLA-E loaded with a UL40 mimic peptide revealed a highly specific mode of recognition with TCR footprints centred around peptide residue 8 – the residue of UL40 mimic peptides that differs from host MHC class I genotype encoded VL9 peptides in responding individuals (Hoare et al., 2006) (Sullivan et al., 2017). The TCR-HLA-E interaction was low affinity in nature (30μM) as it is possible high affinity TCRs recognising the UL40 mimic peptide would have undergone clonal deletion in the thymus due to cross-reactivity with highly similar self VL9 peptide. Thus, the low affinity

nature of this interaction in addition to the C-terminally shifted docking of the TCR onto HLA-E centring on peptide position 8 might be specific to the recognition of UL40-derived mimic peptides due to their 89% shared sequence identity with self-encoded VL9 peptides. It is possible alternative binding affinities and modes of docking exist for Mamu-E-restricted TCRs specific for SIV epitopes identified in RhCMV 68-1-SIV vaccination (Hansen et al., 2016). Although such structural studies might not reflect TCR engagement with non-canonical peptide-bound HLA-E, they did establish a unique basis of TCR restriction to HLA-E; HLA-E heavy chain residues Asp-69 and His-155 form contacts at the TCR interface, and such residues are not conserved with classical MHC class I molecules (Hoare et al., 2006) (Sullivan et al., 2017). Notably, the substitution of HLA-E His-155 to Ala abrogated recognition by HLA-E-restricted T cells emphasising the importance of this specific residue in HLA-E TCR restriction (Hoare et al., 2006).

Alternative contexts of HLA-E/Qa-1-restricted $CD8^+$ T cell responses and HLA-E/Qa-1 binding peptides

In addition to broadly-targeted SIV-specific Mamu-E-restricted $CD8^+$ T cell responses conferring protection against SIV challenge following RhCMV 68-1 vaccination, and Mtb-specific HLA-E-restricted $CD8^+$ T cells demonstrating anti-mycobacterial activity, non-VL9 peptides with binding affinity for and T cell responses restricted by HLA-E or Qa-1 have been identified in a range of alternative contexts.

Although Qdm peptides dominate the Qa-1 eluted ligandome (Cotterill et al., 1997), distinct populations of Qa-1-restricted TCR$\alpha\beta^+$ and TCR$\gamma\delta^+$ $CD8^+$ T cells have been identified which recognise alternative Qa-1-presented peptides. Early studies identified a subset of Qa-1-restricted regulatory $CD8^+$ T cells (T_{regs}) that were capable of mediating feedback inhibition by suppressing activated $CD4^+$ T follicular helper cells (Noble et al., 1998). Subsequently, Qa-1-restricted $CD8^+$ T_{regs} recognising a peptide derived from the TCR variable β8 chain (Vβ8) were

shown to ameliorate immunopathology in experimental murine models of autoimmune disease including type 1 diabetes via depletion of activated, autoreactive Vβ8$^+$ CD4$^+$ T cells (Panoutsakopoulou et al., 2004). Similarly, a regulatory subset of Qa-1-restricted CD8αα$^+$ TCRαβ$^+$ T cells specific for a conserved peptide derived from the TCR Vβ8.2 chain (position 42-50, GLRLIHYSY) were found to promote tolerance and prevent the onset of autoimmune disease by exclusively killing activated self-reactive Vβ8.2$^+$ CD4$^+$ T cells in the context of autoimmune encephalomyelitis (Tang et al., 2006). It is thought that upregulated Qa-1 surface expression on activated Vβ8.2$^+$ CD4$^+$ T cells renders them susceptible to cytolysis by the novel Qa-1-restricted CD8$^+$ T_{reg} subset whilst sparing naïve, resting Vβ8.2$^+$ CD4$^+$ T cells (Stanton and Carbon, 1982).

In addition, a Qa-1-restricted immunosuppressive CD8$^+$ T cell subset were reported to recognise the Heat shock protein derived signal peptide (Hsp60sp) presented by Qa-1 on autoreactive CD4$^+$ T cells, thereby mediating protection from autoimmune encephalopathy in a murine model of encephalomyelitis (Chen et al., 2007). Similar findings were recapitulated in a murine model of rheumatoid arthritis in which immunisation with dendritic cells pulsed with a Qa-1-binding peptide from a distinct position within the Hsp60 (residues 216-224) but not a control self-peptide, triggered an expansion of Qa-1-restricted CD122+, Ly49+ T_{regs} (Leavenworth et al., 2013). This immunosuppressive T cell subset mediated perforin-dependent cytolysis of immunopathological CD4$^+$ T follicular helper and Th17 cells, preventing collagen-induced arthritis independently of traditional Foxp3+ CD4+ T_{reg} induction. It is possible that the emergence of the peptide derived from the stress-related Hsp60 within the Qa-1 binding repertoire combined with upregulation of Qa-1 surface expression on activated CD4$^+$ T cells in contexts of cellular stress such as autoimmunity, infection or transformation, provides an opportunity for these cells to be eliminated by Qa-1-restricted T_{regs}. In humans the Hsp60 signal peptide has also been shown to stabilise HLA-E surface expression in stressed

cells and abrogate inhibitory interactions with CD94/NKG2A leading to diminished protection from NK cell-mediated lysis (Michaëlsson et al., 2002).

The immunosuppressive role of Qa-1-restricted $CD8^+$ T_{regs} in promoting immune homeostasis was demonstrated by a Qa-1 mutant knock-in mouse model containing a position 227 Asp to Lys substitution which abrogated Qa-1-TCR interactions via blockade of CD8 co-receptor engagement (Lu et al., 2008). Such Qa-1 mutant mice exhibited dysregulated $CD4^+$ T cell responses and increased susceptibility to experimental autoimmune encephalomyelitis. Equally, autoimmune-prone mice containing the same point mutation (Asp227Lys) at the Qa-1 locus, presented with 5-6 fold increases in $CD4^+$ T follicular helper cell levels and enlarged germinal centres relative to wild-type mice due to defective Qa-1-restricted $CD8^+$ T_{reg}-mediated cytolysis of activated $CD4^+$ follicular helper T cells (Kim et al., 2010). Strikingly, such mice spontaneously developed a systemic lupus-like disease characterised by tissue-specific autoantibodies, dysregulated expansions of $CD4^+$ follicular helper T cells and glomerulonephritis, further illustrating the importance of this tolerogenic Qa-1-restricted $CD8^+$ T_{reg} subset. Consistent with results obtained in mice, a population of regulatory $CD8^+$ T cells that recognised Hsp60 derived peptide presented by HLA-E were identified in humans and shown to manifest functional defects in patients with type 1 diabetes (Jiang et al., 2010). Further, a population of HLA-E-restricted $TCR\alpha\beta^+$ $CD8^+$ T_{regs} were capable of recognising and mediating granule-dependent cytolysis of autoreactive myelin-specific $CD4^+$ T cells isolated from patients with multiple sclerosis (Correale and Villa, 2008). Notably, the frequencies and cytolytic potential of such HLA-E-restricted $CD8^+$ T_{regs} were lower in the peripheral blood of patients experiencing disease exacerbation relative to those in remission or when compared to healthy controls.

Interestingly, a non-Qdm Fam49b-encoded Qa-1-binding nonapeptide, FL9 (FYAEATPML), was found to be exclusively presented when the ER aminopeptidase associated with antigen

processing, (ERAAP), but not TAP, was defective (Nagarajan et al., 2012). Given that high frequencies of cytotoxic Qa-1-restricted $CD8^+$ T cell responses specific for the immunodominant FL9 epitope were capable of eliminating ERAAP-deficient cells *in vitro* and *in vivo*, it was proposed this alternative Qa-1-presented peptide serves to report the presence of specific antigen processing defects to T cells. Further, FL9-bound Qa-1 complexes were not recognised by CD94/NKG2A – a somewhat unsurprising finding considering CD94-contacting peptide residues are not conserved between VL9 and FL9 peptides (Kaiser et al., 2008) (Nagarajan et al., 2012).

In terms of pathogen-derived HLA-E or Qa-1 binding peptides, a highly conserved HLA-E-binding peptide encoded by the Hepatitis C Virus (HCV) core protein (YLLPRRGPRL) was identified that contained a central position 5 Arg residue conserved with VL9 peptides – a residue that was previously shown to be critical for CD94/NKG2A interaction (Nattermann et al., 2005a) (Kaiser et al., 2008) (Miller et al., 2003). Accordingly, it was shown that HLA-E bound to the HCV core peptide and that this complex mediated inhibitory interactions with CD94-NKG2A receptors (Nattermann et al., 2005a). Thus, it was concluded that HCV has evolved a strategy to evade NK cell lysis via the upregulation of HLA-E in infected hepatocytes coupled with the expression of a VL9-like HLA-E-restricted peptide to preserve inhibitory CD94/NKG2A interactions. Subsequently, HLA-E-restricted $CD8^+$ T cell responses that recognised the HLA-E-restricted HCV peptide were identified and found to occur at higher frequencies in individuals homozygous for the HLA-E*0101 (position 107-Arg) allele (Schulte et al., 2009). Further, a homozygous HLA-E*0101 genotype correlated with reduced susceptibility to chronic HCV infection.

Interestingly the HLA-E-restricted HCV epitope exhibits dual binding specificity whereby it also constitutes a well-characterised HLA-A2 epitope recognised by HLA-A2-restricted $CD8^+$ T cells in HCV infection (Pietra et al., 2010). This is perhaps unsurprising given that HLA-E and HLA-A2

share overlapping peptide binding repertoires and similar binding motifs with preferences for hydrophobic residues at primary anchor positions 2 and 9 (Parker et al., 1992) (Braud et al., 1997). Accordingly, a number of additional pathogen-derived epitopes with dual HLA-A2/HLA-E binding specificity have been identified including peptides from the Epstein-Barr virus (EBV) BZLF1 protein (SQAPLPCVL) and the Influenza Matrix protein (ILGFVFTLT), which were capable of stabilising HLA-E surface expression (Ulbrecht et al., 1998). Both such peptides exhibit sequence differences to canonical VL9 peptides with the EBV peptide containing a polar Gln in place of the hydrophobic Met at primary anchor position 2 and a polar Cys in place of a hydrophobic Val or Leu at secondary anchor position 7. Similarly the Influenza peptide contains a non-canonical Thr in place of the highly conserved Leu at primary anchor position 9. Interestingly, a C-terminally truncated 7mer version of the EBV peptide was able to stabilise HLA-E surface expression, albeit at lower levels than the nonameric version, supporting the hypothesis that the HLA-E binding groove can tolerate peptides of different lengths in addition to peptide sequence variation (Ulbrecht et al., 1998). Additionally, two self-peptides (AMAPIKTHL and AMAPIKVRL) encoded by alternatively spliced isoforms of the stress-inducible peroxiredoxin 5 (Prdx5), both of which contained the canonical HLA-E binding motif, also exhibited dual HLA-E/HLA-A2 binding specificity (Sensi et al., 2009). Both Prdx5 self-peptides stabilised HLA-E and HLA-A2 surface expression and elicited HLA-A2- and HLA-E-restricted $CD8^+$ cytotoxic T lymphocytes, respectively, whilst abrogating interactions between HLA-E and CD94/NKG2A. Given the inducible nature of expression of the anti-oxidant enzyme, Prdx5, in contexts of oxidative stress, it was hypothesised that such T cell responses may serve to eliminate stressed cells that have, for example, become virally infected or undergone transformation.

A *Salmonella typhimurium*-derived immunodominant epitope, GroEL (KMLRGVNVL), was initially shown to elicit Qa-1-restricted $CD8^+$ T cell responses in infected mice (Lo et al., 2000), and subsequently GroEL derived peptides were identified as targets of HLA-E-restricted $CD8^+$ T

cell-mediated granule-dependent cytotoxicity in *Salmonella enterica* serovar Typhi-vaccinated humans (Salerno-Gonçalves et al., 2004). Interestingly, GroEL-specific Qa-1-restricted $CD8^+$ T cells were shown to cross react with a Hsp60-derived peptide, with implications for autoimmune pathology (Lo et al., 2000). In addition to Qa-1-restricted $TCR\alpha\beta^+$ $CD8^+$ T cells, a distinct Qa-1-restricted $TCR\gamma\delta^+$ $CD8^+$ T cell subset devoid of CD94/NKG2 class receptor expression that comprised 15% of the overall γδ intestinal intraepithelial lymphocyte (iIEL) population, was also identified following oral *Salmonella* infection; although the functional relevance of and precise ligand(s) recognised by this novel subset are yet to be determined (Davies et al., 2004). In summary, a range of self- and pathogen-derived HLA-E- or Qa-1-binding peptides, listed in Table 2, have been identified in various contexts of infection, autoimmune disease and immune homeostasis - a number of which lack canonical primary anchor residues indicating that HLA-E is capable of presenting a more diverse peptidome.

Peptide ID	Sequence	Aas	Organism	Protein	Ref.
GroEL Salmonella	KMLRGVNVL	9	*Salmonella enterica* serovar Typhi	$GroEL^{15-23}$	Salerno-Goncalves, R. et al. (2004)
EBV BZLF1	SQAPLPCVL	9	*Epstein Barr Virus*	$BZLF1^{39-47}$	Jørgensen, P,B. et al. (2012)
InflM	ILGFVFTLT	9	*Influenza*	$Matrix^{59-67}$	Ulbrecht, M. et al. (1998)
HCV Core 10mer	YLLPRRGPRL	10	*Hepatitis C Virus*	$HCV\ core^{35-44}$	Jiang, H. et al. (2010)
AA9 HIV p24	AISPRTLNA	9	*Human Immunodeficiency virus*	$P24^{14-22}$	Davis, Z,B. et al (2016)
Prdx5	AMAPIKTHL	9	Human	$Prdx5^{52-60}$	Sensi, M. et al. (2009)
Hsp4 Self	QMRPVSRVL	9	Human	$Hsp60sp^{10-18}$	Michaëlsson, J. et al. (2002)
Hsp5 self	GMKFDRGYI	9	Human	$Hsp60.4^{216-224}$	Michaëlsson, J. et al. (2002)
Fam49B protein	FYAEATPML	9	Human	$Fam49b^{190-198}$	Nagarajan, N,A. et al. (2012)

Table 2: HLA-E and Qa-1 binding peptides identified in various contexts of infection or autoimmune disease

A. Table detailing pathogen- or self-derived peptides reported to bind HLA-E or Qa-1 in various settings of infection, autoimmune disease and immune homeostasis. Peptide ID, sequence, length and origin are reported with the corresponding reference.

Interestingly, a diversification of the Qa-1 peptide binding repertoire away from the Qdm peptide is observed in tumours with full or partial defects in the antigen processing pathway, primarily involving TAP or the classical murine MHC class I molecules H-2K^{b} or H-2D^{b} (Oliveira et al., 2010). Qa-1-bound eluted peptides from TAP-deficient tumours were derived from a range of endogenous house-keeping self-proteins, ranged between 8-18 amino acids in length and exhibited striking sequence diversity with substantial deviation from the canonical, hydrophobic Qdm motif of position 2 Met and position 9 Leu. Similar to the sequence diverse, motif-less TAP-independent SIV-derived epitopes identified in RhCMV 68-1-SIV vaccination studies (Hansen et al., 2016), Qa-1 eluted peptides in Tap-negative tumour cells contained a range of non-canonical residues at both primary anchor position 2 (including Ala, Gly, Ser, Gln, Glu, Asp and Asn) and 9 (including Phe, Lys, Gln, Ala, Ile and Tyr) (Oliveira et al., 2010). The emergence of non-Qdm peptides presented by Qa-1 resulted in a loss of recognition by CD94/NKG2A and subsequent NK cell-mediated lysis of TAP-deficient tumour cells. Moreover, a population of Qa-1-restricted TCRαβ^{+} CD8^{+} cytotoxic lymphocytes (CTLs) were capable of recognising the immunogenic non-Qdm self-derived peptides presented by Qa-1, thus contributing to anti-tumour immunity (Van Hall et al., 2006). Such findings are consistent with the identification of TAP-independent Mtb and SIV epitope presentation by HLA-E- and Mamu-E in contexts where the classical antigen presentation pathway is disrupted by Mtb infection or RhCMV 68-1-SIV vaccination, respectively (van Meijgaarden et al., 2015) (Hansen et al., 2016). Such results also support the hypothesis that VL9 or Qdm peptides dominate the HLA-E or Qa-1 peptide binding repertoires in the steady state, respectively, however in specific contexts where the antigen presentation pathway becomes compromised by infection or cellular transformation, a diversification of these binding repertoires can ensue facilitating unconventional CD8^{+} T cell recognition.

Upregulated HLA-E surface expression is common in human tumours, leukaemia-derived cells and conditions that reflect the tumour environment such as combined hypoxia and glucose

deprivation (Marín et al., 2003) (Gooden et al., 2011) (Sasaki et al., 2016) (Seliger et al., 2016) (Abd Hamid et al., 2019). Aberrant tumour cell HLA-E surface expression was correlated with partial downregulation of classical MHC class I alleles and a resultant increase in the availability of free β2M (Marín et al., 2003). Although co-expression of MHC class I molecules containing HLA-E-binding leader peptides was not essential for HLA-E surface expression in malignant tissues, a supply of MHC class I leader peptide was required for inhibitory interactions with CD94/NKG2A-positive T cells and NK cells (Monaco et al., 2011). Accordingly, elevated levels of surface expressed HLA-E promoted tumour escape and was correlated with poorer prognoses in patients with breast carcinomas co-expressing HLA-G - the non-classical MHC class Ib molecule which encodes a particularly high affinity HLA-E-binding leader peptide (de Kruijf et al., 2010). Further, multiple studies have associated enhanced tumour expression of HLA-E with the functional impairment and attenuation of CD94/NKG2A-positive tumour infiltrating $CD8^+$ T cells (Gooden et al., 2011) (Abd Hamid et al., 2019). The administration of Monalizumab, an anti-NKG2A monoclonal antibody, has yielded promising therapeutic effects in a phase 2 clinical trial for head and neck cancer by restoring anti-tumour effector functions of infiltrating $CD8^+$ T cells (André et al., 2018).

Although upregulated HLA-E surface expression is likely selected for during tumour evolution to maintain CD94/NKG2A inhibitory signals and thus evade NK cell-mediated lysis (de Kruijf et al., 2010) (Monaco et al., 2011), this feature could be exploited in an HLA-E-targeted vaccine setting. For example, it may be possible to elicit potent HLA-E-restricted anti-tumour $CD8^+$ T cell responses via vaccination with appropriate CMV vector strains recombinant for HLA-E-binding neoantigens that frequently arise in certain types of cancer, similar to SIV-specific RhCMV 68-1-SIV-elicited Mamu-E-restricted $CD8^+$ T cell responses which confer protection against SIV challenge (Hansen et al., 2016). Higher levels of HLA-E surface expression reported for a number of neoplastic cells relative to their untransformed counterpart cells would likely render them particularly vulnerable to such unconventional T cell targeting.

Aims and Objectives

In conclusion, although in the steady state the primary functional role of HLA-E is to regulate NK-cell activity via MHC class I signal peptide presentation, specific contexts permit diversified HLA-E peptide binding repertoires and HLA-E-restricted $CD8^+$ T-cell priming. In light of the unprecedented Mamu-E-restricted $CD8^+$ T cell-mediated protection against SIV challenge in RhCMV 68-1-SIV-vaccinated rhesus macaques, it is now important to explore the alternative antigen presentation function of HLA-E with the aim of developing a universal, HLA-E-targeted $CD8^+$ T cell vaccine against HIV in humans.

It is currently unclear whether HLA-E shares the binding breadth of its rhesus counterpart molecule, Mamu-E, and thus the capacity of HLA-E to accommodate sequence diverse, non-canonical peptides identified in RhCMV 68-1-SIV vaccine studies needs to be explored. Further, although a range of Qa-1 or HLA-E binding peptides have been identified in various contexts of infection or autoimmune disease, a comparative and comprehensive analysis of relative HLA-E peptide binding affinity is lacking. The underlying structural basis of suboptimal, non-canonical peptide accommodation by the hydrophobic HLA-E binding groove also remains to be determined.

We therefore aimed to address some of the unanswered questions regarding suboptimal epitope binding to HLA-E from a molecular, structural and biophysical perspective. Via the development and optimisation of a high-throughput *in vitro* HLA-E peptide binding assay and subsequent screening of large numbers of predicted HLA-E binding peptides, we aimed to re-evaluate the breadth of the HLA-E binding repertoire and compare relative binding strengths of suboptimal peptides to HLA-E. A structural exploration of HIV and Mtb epitope binding to HLA-E aimed to determine how non-VL9 peptides are accommodated by the HLA-E binding groove and subsequent biophysical analyses involving small angle x-ray scattering aimed to

investigate the impact of suboptimal peptides on the overall conformational ensemble of HLA-E complexes in solution.

Materials and methods

Expression Constructs

A pre-existing construct (generated by Professor G. Gillespie), comprising the extracellular domains (residues 1-274) of the HLA-E*0103 heavy chain in linkage with a C-terminally extended biotin substrate (AviTag) motif was used as a template to generate a non-tagged HLA-E*0103 construct. The specific DNA region encompassing residues 1-274 was amplified using the REDTaq ReadyMix PCR kit (Sigma Aldrich) - specific reaction conditions and temperatures are specified in Table 1. The forward primer comprised 5'-tgaaga**CATATG**GGTTCTCAT TCTTTAAAATAT-3', where lower case letters represent non-specific overhangs to facilitate restriction digests, CATATG is the NdeI restriction site with an ATG start codon and the remainder are nucleotides which correspond to the HLA-E 'GSHSLKY' amino acid codons. The reverse-complemented reverse primer comprised 5'-gaaga**GGATCCTTATTA**TGGCTTCCATCTCAG -3', whereby lower case letters represent non-specific overhangs to facilitate restriction digests, GGATCC is the BamHI restriction site, TTATTA is a double stop codon and the remainder are the nucleotides corresponding to amino acid codons for HLA-E 'LRWKP', where P is residue 276. Successful PCR amplification was confirmed by electrophoresis where 10µL of the PCR product was resolved for 30mins at 120V on a 1% TAE gel, made by dissolving 0.5g Agarose into 50mL TAE buffer and subsequently adding 2.5µL SYBR Safe DNA Gel Stain (ThermoFisher Scientific). The remainder of the PCR product (40µL) was prepped to remove free nucleotide and primers using the supplied buffers and specified protocol of the Qiagen Gel Extraction Kit, and subsequently digested using the restriction enzymes NdeI & BamHI (New England BioLabs). A total digestion reaction volume of 36µL comprising 30µL of the purified PCR product, 1.2µL of each restriction enzyme (24 Units of each in the final reaction volume) and 3.6µL 10X CutSmart buffer (New England BioLabs) was incubated at 37°C for 3hrs and subsequently at 65°C for 10mins to heat

inactivate the restriction enzymes. Following digestion, the sample was resolved on a 1% TAE gel for 30mins at 120V, following which the gel band was excised using an x-tracta gel extraction tool (Sigma) and purified according to the Qiagen Gel Extraction Kit protocol. The PET22B+ expression vector was also digested using the restriction enzymes NdeI & BamHI (New England BioLabs). Specifically, 1µg plasmid DNA, 1.5µL of each restriction enzyme (30 Units each in a final volume of 35µL) and 3.5µL CutSmart 10X buffer (New England BioLabs) were made up to a final volume of 35µL in MiliQ water and incubated for 2.5hrs at 37°C. 1µL Antarctic Phosphatase (5 Units in the final reaction volume) (New England BioLabs) was then added to the reaction which was subsequently incubated for a further 1hr at 37°C to dephosphorylate digested 5' and 3' plasmid ends to prevent plasmid re-circularisation. Finally, the digestion reaction was heat inactivated at 65°C for 10mins. The plasmid was subsequently purified by gel electrophoresis on a 1% TAE gel for 1hr at 100V and the gel band excised using an x-tracta gel extraction tool (Sigma) prior to purification according to the Qiagen Gel Extraction Kit protocol. The concentrations of both the digested PCR product and plasmid were obtained by measurement of the absorbance at 260nm using a NanoDrop ND-1000 Spectrophotometer.

RedTaq ReadyMix PCR:	
Reaction conditions:	**Reactions Temperature:**
25uL readymix	X1 Cycle: 95° for 5 mins
0.75uL forward primer 0.75uL reverse primer	X35 Cycles: 95° for 1 min 62° for 1 min 72° for 3 mins
22.5uL MiliQ Water	X1 Cycle: 95° for 1 min
1uL template	62° for 1 min 72° for 6 mins

Figure 1: PCR amplification conditions

Table outlining reaction conditions and temperatures for PCR amplification of HLA-E*0103 using the REDTaq ReadyMix kit.

Amplified HLA-E*0103 inserts were subsequently ligated into the commercial PET22b+ prokaryotic expression vector (Novagen) at 2 test concentrations of the PCR insert (5.9ng of PCR insert for the 1:1 insert to plasmid ratio and 17.8ng for the 3:1 insert to plasmid ratio) and 1 concentration of the plasmid (35ng). In addition to insert and plasmid DNA, ligation reactions contained 1µL T4 DNA Ligase (New England BioLabs) and 2µL 10X T4 DNA Ligase Reaction Buffer (New England BioLabs) – such reactions were made up to a total volume of 20µL in MiliQ water and incubated for 12 hours at 16°C. 10µL of each test ligation was incubated with chemically competent DH5α bacterial cells (ThermoFisher Scientific) for 20mins on ice, heat shocked at 42°C for 20 seconds and recovered in 300µL SOC medium (Sigma Aldrich) prior to plating 125µL on pre-cast Luria-Bertani (LB)-Ampicillin plates (100 ug/mL Ampicillin concentration) which were incubated overnight at 37°C. Colonies were subsequently picked and used to inoculate tubes containing 5mL LB broth (containing 100 ug/mL freshly added Ampicillin) which were incubated for 12hrs at 37°C and subsequently purified for sequence analysis according to the instructions of the QIAprep Spin Miniprep Kit (Qiagen).

Protein Expression

Approximately 100ng of plasmid containing the HLA-E*0103 insert was added to 50µL of chemically competent BL21DE3pLysS *E.coli* (ThermoFisher) and incubated on ice for 30 minutes. Following heat-shock at 42°C for 30 seconds, bacterial cells were spread on LB-Ampicillin bacterial agar plates – such plates were poured from LB which was prepared with 5g tryptone, 2.5g yeast extract, 2.5g NaCl and 8g Agar in 500mL MiliQ water and autoclaved prior to the addition of Ampicillin to achieve a concentration of 100 ug/mL (all such reagents were purchased from Sigma Aldrich). A previously generated human β2M construct cloned into the T7 expression plasmid (PGMT7) was also expressed in BL21 competent bacteria. An overnight starter culture was prepared by inoculating 1L of LB broth (composition detailed above), to which Ampicillin had been freshly added to achieve a concentration of 100 ug/mL, with a

single bacterial colony. Following an overnight incubation at 37°C at a low rotation speed, 150mL of starter culture was equally divided between three flasks each containing 1L of LB broth with freshly added Ampicillin (100 ug/mL). After further incubation at 37°C and once the cultures reached an OD at 600nm of 0.5, protein expression was induced by the addition of 0.5M Isopropyl-β-D-thiogalactoside (IPTG) per litre of LB broth. Following a further 4 hours incubation at 37°C, bacterial cells were harvested via centrifugation (3000g for 30 minutes at 4°C in a Thermo Scientific Heraeus Megafuge 40R bench top centrifuge with a swing-bucket TX-1000 rotor). Following disposal of supernatants, bacterial cell pellets were re-suspended (by repeated pipetting in 10mL ice-cold PBS), transferred into 50mL Falcon tubes and stored at -80°C until required.

Inclusion body isolation and protein purification

Bacterial pellets of inclusion-body expressed HLA-E*0103 and β2M were purified by a process of sonication and homogenisation in a Triton-based buffer comprising 0.5% Triton X-100 (Sigma Aldrich), 50mM Tris pH 8 (ThermoFisher Scientific), 100mM NaCl (pre-prepared lab stock), 1mM Ethylenediamineteraacetic acid (EDTA) (ThermoFisher Scientific), 2mM Dithiothreitol (DTT) (ThermoFisher Scientific) and 0.1% sodium azide (Sigma Aldrich), prepared in MiliQ water. Sonication was conducted for 6 minutes using a 13mm tip diameter probe head in a Sonics Vibra cell instrument whereas centrifugation was conducted in an Avanti JXN-26 centrifuge with a JA17 rotor. Sonicated samples were spun in Nalgene x10 Centrifuge oak ridge tubes (50mL volume) at 27,000 g for 14 minutes at 4°C and each pellet was subsequently re-solubilised in 10mL of the Triton-based wash buffer (detailed above) in a potter type homogeniser (Kartell), before being topped up to 50mL with the Triton-based buffer. Centrifugation and pellet re-solubilisation steps were repeated until three Triton-based buffer washes had been completed. Following the final spin, the pellet was re-solubilised into a re-suspension buffer (50mM Tris pH 8, 100mM NaCl, 1mM EDTA and 2mM DTT, prepared in

MiliQ water with all reagents being purchased from ThermoFisher Scientific except NaCl which was a pre-prepared lab stock) and spun again at 27,000 g for 14 minutes at 4°C. Final re-solubilisation of the pellet was achieved in 20mL of buffer containing 8M urea (Sigma Aldrich), 50 mM MES pH 6.5 (pre-prepared lab stock), 0.1 mM EDTA (ThermoFisher Scientific) and 0.1 mM DTT (ThermoFisher Scientific). The re-solubilised pellet was incubated at 4°C for 12 hours after which the protein concentration was obtained by measurement of the absorbance at 280nm using a NanoDrop ND-1000 Spectrophotometer. Aliquots were subsequently snap-frozen on dry ice and stored at -80°C. Prior to addition in micro-scale refolds in peptide binding assays, β2M was subsequently refolded in a buffer containing 100mM Tris pH 8.0, 400mM L-Arginine monohydrochloride, 2mM EDTA, 5mM reduced Glutathione and 0.5mM oxidised Glutathione, prepared in MiliQ water (all such reagents were purchased from ThermoFisher Scientific except L-Arginine monohydrochloride which was from Sigma Aldrich).

Peptide synthesis

Synthetic peptides were generated by Fmoc (9- fluorenylmethoxy carbonyl) chemistry to a purity of 85% by Genscript USA. All peptides were provided as lyophilised powder, reconstituted in DMSO to a concentration of 200mM, and stored at −80 °C. A UV photo-labile version of the HLA-B leader sequence peptide, VMAPRTLVL, incorporating a UV-sensitive 3-amino-3-(2-nitrophenyl)-propionic acid residue (J residue) substitution at position 5, was synthesised by Dris Elatmioui at LUMC, The Netherlands. The UV-sensitive peptide (termed 7MT2) was stored as lyophilised powder, and dispensed/reconstituted as required.

Peptide binding assays

A peptide binding assay was adapted from published micro-scale refold enzyme-linked immunosorbent assay (ELISA)-based methods (Sylvester-Hvid et al., 2002) (Strong et al., 2003). 1 µM heavy chain and 1.5 µM pre-refolded β2M were refolded in 0.33 mM Tris-Maleate (Sigma Aldrich) and 0.5% Lutrol-F68 (Sigma Aldrich) in the presence of 70 µM peptide, pre-

diluted to 2 mM working stocks in 100mM Tris-HCL pH 8. Such micro-scale refolds, which had a final total volume of 400µL, were assembled in Eppendorf tubes and incubated at room temperature for 48hrs, with a rotation of 7rpm on a Stuart Equipment Tube Rotator, before the relative capacity of each peptide to support stable HLA-E-β2M-peptide complex formation was quantified by ELISA.

An alternative UV-induced peptide exchange-based binding assay was developed using HLA-E-β2M complexes that had been pre-refolded in the presence of UV-labile 7MT2 peptide and purified by size exclusion chromatography, as outlined previously (Rodenko et al., 2006) and as outlined below in the section titled 'Protein refolding and purification'. For UV photo-cleavage and peptide exchange, 3µg of pre-refolded HLA-E-7MT2 monomer was incubated in the presence of 100 µM 'exchange' peptide in polypropylene V-shaped 96-well plates (Greiner Bio-One). The final volume of each well was adjusted to 125µL by adding exchange buffer (20mM Tris pH 7.4 (ThermoFisher Scientific), 150mM NaCl (pre-prepared lab stock)) and the reaction was subsequently exposed to long-wave UV radiation at 366nm for a period of 1hr in a Camag UV cabinet on ice. The optimisation of certain variables of this technique is detailed in the text of Results Chapter 1, specifically in the section discussing peptide binding assay optimisation.

Sandwich ELISA

96 well ELISA plates (Life Technologies) were coated with the anti-human HLA-E monoclonal capture antibody, 3D12 (BioLegend), diluted to 10µg/mL in freshly prepared ELISA Coating Buffer (BioLegend) and incubated for 12 hours at 4°C. Wells were washed three times in 200µL phosphate buffered saline (PBS), and subsequently incubated in 300µL of 2% IgG-free bovine serum albumin (BSA) (Sigma Aldrich) for 2 hours at room temperature to block unoccupied surfaces on the ELISA plate wells to prevent non-specific protein binding. Following plate blocking, plates were washed five times with 200µL freshly prepared 0.05% Tween-based ELISA Wash Buffer (BioLegend) and a final time in 200µL phosphate buffered saline (PBS) (all

subsequent plate washing was carried out in an identical manner). 50µL of the HLA-E reaction mixture from either micro-scale refolding or UV-induced peptide exchange, diluted 1:100 in 2% BSA, was added to each well of the ELISA plate and left to incubate for 1 hour at room temperature. Following protein incubation and subsequent plate washing, 50µL of polyclonal anti-human β2M HRP-conjugated IgG detection antibody raised in rabbits (ThermoFisher Scientific), diluted 1:2500 in 2% BSA (0.2µg/mL), was added to each well to ensure detection of β2M-associated forms of HLA-E only. After 30mins of dark incubation at 4°C, ELISA plates were washed and 50µL of a goat anti-rabbit HRP-conjugated IgG enhancement antibody (EnVision+ System-HRP from Agilent), diluted 1:15 in 2% BSA containing 1% mouse serum to prevent cross-reactivity, was added to each well to amplify the signal. ELISA plates were incubated for 15mins in the dark at room temperature, and following a final round of plate washing, developed in 100µL of 3,3',5,5'-tetramethyl benzidine (TMB) substrate (BioLegend) and incubated in the dark at room temperature for a further 10 minutes. Finally, the reactions were terminated in 100µL of H_2SO_4 STOP solution (BioLegend) and absorbance readings were obtained at 450nm on a FLUOstar OMEGA plate reader.

Protein refolding and purification

250mL HLA-E protein refolds were assembled as follows. β2M in Urea-Mes, at a final concentration of 2µM, was refolded at 4 °C for 30mins in a macro-refolding buffer prepared in MiliQ water containing 100 mM Tris pH8.0, 400mM L-arginine monohydrochloride, 2mM EDTA, 5mM reduced glutathione and 0.5mM oxidised Glutathione (all such reagents were purchased from ThermoFisher Scientific except L-Arginine monohydrochloride which was from Sigma Aldrich). Test peptide that was previously reconstituted in DMSO, was subsequently added to the 250mL β2M refold to achieve a peptide concentration of 20-50µM. HLA-E*0103 heavy chain was subsequently pulsed into the refolding buffer until a final concentration of 1µM was reached. To avoid aggregation, the HLA-E heavy chain was added via a method of

dilution whereby refolding buffer was first slowly added to the appropriate volume of the HLA-E heavy chain protein preparation in Urea-Mes, before the buffer-diluted protein sample was added to the refolding mix. Following incubation for 72hrs at 4 °C, HLA-E refolds were filtered through 1.0µm cellular nitrate membranes to remove aggregates prior to concentration by the VivaFlow 50R system with a 10kDa molecular weight cut-off (Sartorius), until the 250mL refold had reached a volume of 30mL. Subsequent concentration was carried out in 10kDa cut-off VivaSpin Turbo Ultrafiltration centrifugal devices (Sartorius) until the sample reached a volume of 2mL. Samples were separated according to size into 20 mM Tris pH8 (ThermoFisher Scientific), 100 mM NaCl (pre-prepared lab stock) by fast protein liquid chromatography (FPLC) on an AKTA Start System using a Superdex S75 16/60 column. Elution profiles were visualised by UV absorbance at 280mAU, enabling differentiation of correctly refolded HLA-E-β2M-pepide complexes from smaller non-associated β2M and larger misfolded aggregates. FPLC-eluted protein peak fractions were combined and concentrated to 10 mg/mL using 10kDa cut-off VivaSpin Turbo Ultrafiltration centrifugal devices (Sartorius) prior to the assembly of crystallisation plates. The final protein concentration was obtained by measurement of the absorbance at 280nm using a NanoDrop ND-1000 Spectrophotometer. 1µL aliquots were analysed by non-reducing SDS-PAGE electrophoresis on NuPAGE™ 12% Bis-Tris protein gels (ThermoFisher Scientific) to confirm the presence of non-aggregated HLA-E heavy chain and β2m.

Crystallisation screening

A total of 100nL protein, at 10mg/mL, and 100nL reservoir buffer were mixed in crystallisation wells by the mosquito protein crystallisation robot (TTP LabTech) and equilibrated by sitting drop vapour-diffusion at 20 °C (Walter et al., 2005). A multitude of commercial sparse matrix screens from Hampton Research including PEGRx1 and SaltRx1 in addition to ammonium sulphate versus pH grid screens (Hampton Research), were used to identify optimal

crystallisation conditions, around which ammonium sulphate fine gradient and additive screens were subsequently setup. Automated plate imaging in combination with manual monitoring via the web-based protein crystal information management system, xtalPiMS, enabled identification of protein crystals. Crystals were cryopreserved in 25% glycerol by Karl Harlos, Oxford University and diffraction data were collected at Diamond Light Source Beamlines I04 and I24. Precise crystallisation conditions in addition to data collection and refinement statistics are detailed in Results Chapters 2 and 3.

Crystallographic analysis

Diffraction images from multiple isomorphous crystals were merged in Xia2 to increase the completeness of datasets (Kabsch, 2010) (Evans and Murshudov, 2013). Diffraction data were auto-indexed by Xia2 DIALS using the default parameters since 2015: I/sig(I) > 0.25, merged I/sig(I) > 1 and CC ½ > 0.558–60. Initial phasing was carried out using the coordinates of the VL9-bound HLA-E*01:01 structure (PDB code 1MHE), stripped of peptide, hydrogens and waters as the search model in MolRep of the CCP4i suite (Winn et al., 2011) (Vagin and Teplyakov, 2010) (Murshudov et al., 2011). Molecular replacement for datasets corresponding to RL9HIV, IL9, Mtb14 and Mtb44 variant peptide refolded HLA-E, was subsequently carried out in Phaser in Phenix (Adams et al., 2010) using the refined Mtb44-HLA-E structure coordinates as the phasing model. Rigid body, restrained and TLS refinement were computed by the CCP4i REFMAC5 (Murshudov et al., 2011) or Phenix.refine (Afonine et al., 2012) applying non-crystallographic symmetry restraints between iterative cycles of manual model building in Coot (Emsley et al., 2010). Models were validated using MolProbity to ensure acceptable geometry (Chen et al., 2010), visualised using the PyMOL Molecular Graphics System, version 2.0 (Schrödinger, LLC) and further investigated using PDBePISA (Krissinel and Henrick, 2007) and PDBeFOLD (Krissinel and Henrick, 2004).

Size exclusion chromatography-coupled small angle X-ray scattering

125mL HLA-E-β2M-peptide refolds were assembled in the L-Arginine Tris pH 8 macro-refolding buffer according to the protein refolding section detailed above and incubated for 72hrs at 4°C prior to concentration with the VivaFlow 50R system with a 10kDa molecular weight cut-off (Sartorius). 45µL of pre-refolded HLA-E-β2M-peptide sample concentrated to 10-20mgs/mL, was subsequently injected onto a high-performance liquid chromatography (HPLC) KW402.5 column at Diamond Light Source Beamline B21, several hours post-concentration. HPLC elution buffers corresponded exactly to the L-Arginine Tris pH 8 macro-refolding injection buffer minus the protein components – although 60µM or 120µM excess peptide was added to the elution buffer for certain SEC-SAXS experiments. SEC-SAXS data were collected at Diamond Light Source Beamline B21 with images being taken every 3 seconds of X-ray-exposed HPLC-eluted material over the course of a 32 minute elution period. Scattering data were circularly integrated prior to buffer subtraction followed by manual Guinier fitting and pairwise distribution function calculations, which were performed in the SAXS-dedicated software, ScÅtter, developed by Robert Rambo (Franke et al., 2017). Average bead models were generated by DAMMINF (Svergun, 1999) and the superposition of bead models to structural coordinates was carried out using the Supcomb package of ATSAS (Kozin and Svergun, 2001).

Results chapter 1

HLA-E peptide binding assays

Introduction

A number of HLA-E- and Qa-1-restricted peptides from a range of self- and pathogen-derived proteins have previously been reported, including peptides from Salmonella (Lo et al., 2000),(Salerno-Gonçalves et al., 2004), Hepatitis C virus (Nattermann et al., 2005a), Epstein-Barr virus (EBV) (Ulbrecht et al., 1998), Cytomegalovirus (Tomasec et al., 2000) and Influenza (Ulbrecht et al., 1998). Self-derived peptides from Fam49B (Nagarajan et al., 2012), Hsp60 (Michaëlsson et al., 2002) and Prdx5 (Sensi et al., 2009) have also been reported to bind HLA-E or Qa-1 in the context of cellular stress and autoimmune disease.

More recently, two predominant contexts have been found where large numbers of MHC-E-restricted epitopes are presented – RhCMV 68-1-SIV/HIV vaccination (Hansen et al., 2016) and *Mycobacteria tuberculosis* (Mtb) infection or BCG vaccination (Joosten et al., 2010),(Joosten et al., 2016). Specifically, eleven optimal sequence diverse SIV Gag-derived epitopes were originally shown to elicit SIV-protective Mamu-E-restricted $CD8^+$ T cell responses in RhCMV 68-1-SIV vaccinated rhesus macaques (Hansen et al., 2016). Additionally, a diverse range of Mamu-E-restricted epitopes from Gag, Pol and Nef of both SIV and HIV have recently been shown to stimulate Mamu-E-restricted $CD8^+$ T cells responses following vaccination with RhCMV 68-1 vectors recombinant for SIV or HIV (Hansen et al., 2016) (Louis Picker, unpublished data). Regarding Mtb infection or BCG vaccination, over 70 *Mycobacterial* peptides were predicted to bind HLA-E *in silico* using peptide motif-based algorithms, of which a number were subsequently shown to stimulate HLA-E-restricted $CD8^+$ T cell responses in Mtb-immune individuals (Joosten et al., 2010). Additionally, a smaller set of HLA-E-restricted

Mtb-derived peptides were eluted from infected cells and validated as $CD8^+$ T cell epitopes (McMurtrey et al., 2017).

Together, over 100 HLA-E, Mamu-E or Qa-1 binding peptides have been identified in diverse contexts, primarily involving RhCMV 68-1-vectored SIV/HIV vaccination or Mtb infection. Although many of these peptides have been shown to elicit Mamu-E-, HLA-E- or Qa-1-restricted $CD8^+$ T cell responses in rhesus macaques, humans or mice, respectively, relative peptide binding strengths have not been established. We recognised the need for a high throughput *in vitro* peptide binding assay to validate and relatively quantitate binding to HLA-E. Comparative screening analyses of the numerous, recently identified MHC-E binding peptides could ultimately enable the identification of novel peptide sequence motifs and help re-characterise the HLA-E peptide binding repertoire.

Further, it is critical to determine whether Mamu-E-restricted SIV/HIV-derived epitopes identified in RhCMV 68-1 vaccination are also capable of binding the human orthologue, HLA-E. In order for an MHC-E-targeted HIV vaccine approach to translate into the human clinic, HLA-E must share the broad peptide binding properties of its rhesus macaque counterpart, Mamu-E. An approach that allowed large panels of Mamu-E-restricted SIV- and HIV-derived epitopes to be screened for their capacity to bind HLA-E, would not only enable the breadth of the HLA-E peptidome to be re-evaluated but would also shed light on the extent of overlap in the peptide binding repertoires of the human and rhesus MHC-E orthologues.

Micro-refold-based HLA-E peptide binding assay

We adapted a sandwich ELISA-based assay which had originally been developed to determine peptide binding affinity to classical MHC class I molecules (Sylvester-Hvid et al., 2002). Specifically, this assay was initially optimised to quantify high affinity binding, within the low nM range, of natural peptide ligands to HLA-A*0201 and HLA-A*1101. The method was subsequently used to assess VL9 peptide binding affinity to the HLA-E*0101 and HLA-E*0103

allelic variants, which occur at equal frequencies across ethnically diverse populations (Strong et al., 2003) (Grimsley and Ober, 1997). We adapted this technique to investigate non-canonical epitope binding to HLA-E – in particular, RhCMV 68-1 vaccine-identified SIV/HIV-derived Mamu-E-restricted epitopes (Hansen et al., 2016). We carried out a number of optimisation strategies with the aim of enhancing assay sensitivity and concomitantly enabling detection of the lower affinity, non-canonical MHC-E restricted peptides, many of which contain suboptimal primary anchor residues at positions 2 and 9. Further, we focused primarily on HLA-E*0103 in the peptide binding assays since this allelic variant was previously shown to exhibit superior affinity for all VL9 variant peptides tested (Strong et al., 2003). In addition, HLA-E*0103 yielded higher melting transition temperatures (T_m), which reflect thermal stability, and greater levels of surface expression, relative to HLA-E*0101.

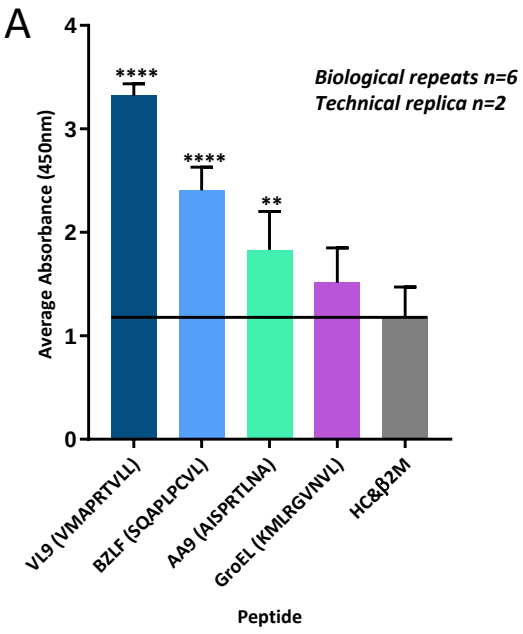
Stable heterodimeric assembly of classical MHC class I heavy chain and β2M light chain is energetically unfavourable in the absence of a binding peptide (Elliott et al., 1991). Thus, the optical density readout of this peptide binding assay reflects the concentration of stably refolded HLA complexes in the refolding buffer, which primarily depends on the affinity of the test peptide. Hence, the ELISA signal is an indirect measure of relative peptide binding strength for HLA-E. Further, anti-HLA-E heavy chain-specific antibody capture with consecutive anti-β2M-specific detection in the sandwich ELISA ensures β2M-bound HLA-E complexes exclusively contribute to the absorbance reading. Computational structure and sequence-based analyses originally ascribed a conformational and discontinuous nature to the epitope of the anti-HLA-E capture antibody, 3D12, with recognition requiring folded secondary structure elements in the HLA-E heavy chain of β2M-associated complexes (Ravindranath et al., 2011). However, subsequent studies demonstrated 3D12 recognition of monomeric, denatured HLA-E heavy chain (Tremante et al., 2015). Therefore, the sandwich nature of the ELISA with anti-β2M detection eliminates unfolded or aggregated β2M-free HLA-E heavy chain

generating signal. In addition, soluble, un-associated β2M would fail to become immobilised by 3D12 and subsequently be lost during the multiple wash steps.

Another advantage afforded by the sandwich ELISA-based binding assay is its high throughput nature. It is possible to simultaneously screen entire panels of test peptides and relatively quantitate binding to HLA-E. This negates the need for more laborious and cost-ineffective approaches, such as those requiring macro-scale protein refolding and size exclusion chromatography.

Experiments using this assay successfully established hierarchies of peptide binding, however, high standard deviation was observed between biological repeats, as represented by error bars in Figure 1, A. Despite the presence of six biological repeats, standard deviation for three test peptides ranged between 0.23 and 0.37 – this is particularly problematic given the low optical density readings (Figure 1, B). In addition, the micro-refold sandwich ELISA method yielded unexpectedly high signals for the peptide-free negative control comprising HLA-E heavy chain and β2M-only (Figure 1, A). A number of optimisation strategies were therefore carried out to address the high background and poor assay reliability. A selection of HLA-E binding peptides were used in optimisation assays according to peptide stock availability including the EBV-derived BZLF (SQAPLPCVL) (Ulbrecht et al., 1998), the HIV p24-derived AA9 (AISPRTLNA) (Nattermann et al., 2005b), the Salmonella-derived GroEL (KMLRGVNVL) (Lo et al., 2000) and the HIV Gag-derived RL9HIV (RMYSPTSIL) (Walters et al., 2018). All such peptides are listed in Figure 1, B and have previously been reported to be HLA-E-restricted in the literature.

A



B

Optimisation peptides

Peptide ID	Sequence	Aas	Organism	Protein	ELISA STDEV	Ref.
GroEL Salmonella	KMLRGVNVL	9	*Salmonella enterica* serovar Typhi	GroEL$^{15-23}$	0.33	Salerno-Goncalves, R. et al. (2004)
EBV BZLF1	SQAPLPCVL	9	*Epstein Barr Virus*	BZLF1$^{39-47}$	0.23	Jørgensen, P,B. et al. (2012)
AA9 HIV p24	AISPRTLNA	9	*Human Immunodeficiency virus*	P24$^{14-22}$	0.37	Davis, Z,B. et al. (2016)
RL9HIV	RMYSPTSIL	9	*Human Immunodeficiency virus*	Gag$_{275-283}$	N/A	Walters, L,C. et al. (2018)

Figure 1: Micro-refold sandwich ELISA-based peptide binding assay

A. Average absorbance readings for 3 HLA-E binding pathogen-derived peptides sampled in the micro-refolding peptide binding assay. The Y-axis denotes average absorbance readings at 450nm and the X-axis specifies the test peptide. **B.** Table detailing previously reported HLA-E binding peptides used in peptide binding assay optimisation. Peptide ID, sequence and origin are reported along with the corresponding reference. Standard deviation (STDEV) from the peptide binding assay displayed in A, is also reported.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background (HC&β2M) in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=6, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

A multitude of parameters from both the micro-refold and sandwich ELISA stages of the binding assay were optimised in an attempt to improve assay reliability. A selection of assay optimisations are depicted in Figure 2, A, B and C including different micro-refolding buffer systems in addition to differential ELISA incubation and development time-points. The Lutrol-Tris-maleate-based refolding buffer was tested in the HLA-E peptide binding assay due to its reported efficacy in classical MHC class I-based peptide binding affinity assays, in which its low pH of 6.6 was thought to optimally support peptide loading (Sylvester-Hvid et al., 2002).

However, we typically prepare peptide working stocks in Tris-HCL pH 8.0 in the lab. Hence, when these peptide working stocks were added to the Lutrol-Tris-maleate micro-refolding buffer, an increase in pH to 8.0 was measured, presumably due to the superior buffering capacity of Tris-HCL. We therefore trialled alternative peptide working stock buffers which did not alter the pH of the Lutrol-Tris-maleate micro-refolding buffer including MiliQ H_2O, PBS and Lutrol-Tris-maleate. When Tris-HCL pH 8.0 was replaced by Lutrol-Tris-maleate, a decrease in standard deviation was observed for AA9 (AISPRTLNA), as represented by the error bars in Figure 2, A. However, AA9 binding to HLA-E was not significant in the Lutrol-Tris-maleate peptide working stock buffer and hence this condition was not pursued. A similar effect was observed when determining the optimal duration for TMB development (Figure 2, C). Although restricting TMB development to one minute effectively reduced the background signal relative to the standard 5 minute development period, the average absorbances of both test peptides and the positive control VL9 peptide also decreased resulting in a loss of significance for AA9 (AISPRTLNA) binding.

In summary, despite the various assay optimisation strategies trialled, the micro-refold sandwich ELISA continued to generate variable peptide binding signals between biological repeats. An additional caveat of this approach was that variability was not only detected between biological repeats of the same assay, but also considerable inter-assay variation was

observed. For example, the HIV p24 derived AA9 peptide produced signals above the heavy chain and β2M background in some assays (Figure 2, A) but was below background in others (Figure 2, B).

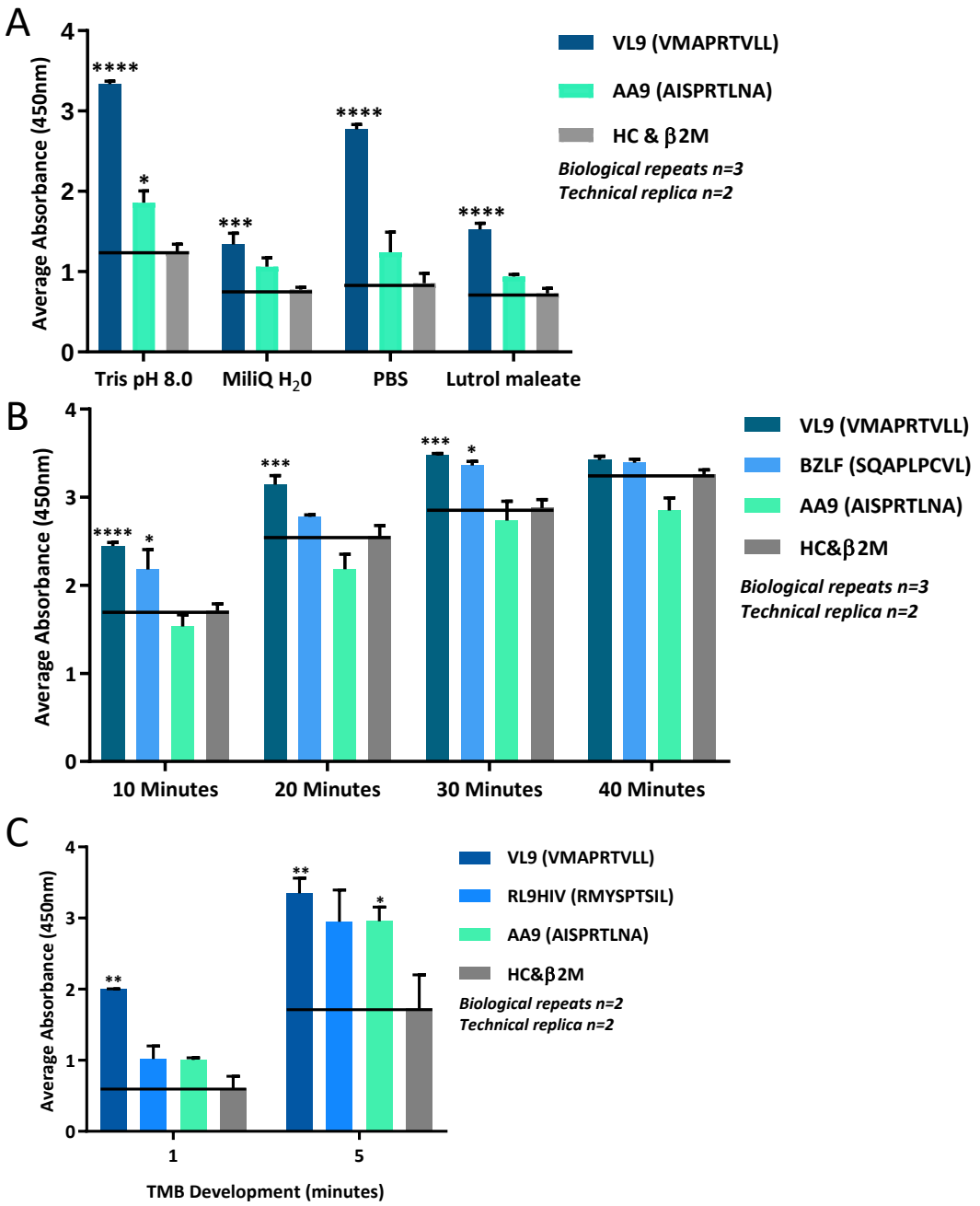


Figure 2: Micro-refold sandwich ELISA-based peptide binding assay optimisation

A. Average absorbance readings for 4 sandwich ELISA assays where 4 different peptide working stock buffer systems were trialled. The standard method of diluting peptide master stocks into Tris pH 8.0 was

compared to MiliQ H_20, PBS and Lutrol maleate. **B.** Average absorbance readings from ELISA optimisation assays where the EnVision™ + System HRP amplification incubation period was varied between 10 and 40 minutes. **C.** Average absorbance readings from TMB development time-point optimisation in the peptide binding assay.

*The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes test peptides. Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background (HC&β2M) in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Optimisation experiments were also carried out to investigate the unexpectedly high background signals generated by HLA-E heavy chain and β2M in the absence of a peptide ligand. The molar ratio of MHC class I heavy chain to β2M was previously optimised to ensure minimal background signals from peptide-free micro-refolds (Sylvester-Hvid et al., 2002). Although low background signals were originally observed for peptide-free HLA-A*0201 and HLA-A*1101 micro-refolds at such molar ratios (Sylvester-Hvid et al., 2002), background signals for HLA-E were considerably higher. It is possible, this could arise from a superior intrinsic propensity of the HLA-E heavy chain to associate with β2M in the absence of peptide. However, it is also possible such signals could derive from contaminating peptide fragments in the protein preparations which interact with the HLA-E binding groove and stabilise complex formation.

To address this, potential peptide fragment contaminants originating from bacterial lysates were removed from protein preparations via pre-purification of urea-solubilised HLA-E heavy chain and β2M material by size exclusion chromatography prior to addition to the micro-refold. Complete elimination is not, however, possible since protein degradation upon thawing and during the 48 hour micro-refold incubation will potentially produce peptide fragments.

Despite this, it is reasonable to assume the rate of contaminating peptide fragment generation would be comparable between HLA-E, HLA-A*0201 and HLA-A*1101, given the similar micro-refold conditions. Unexpectedly, following pre-purification of protein preparations the peptide-free background signals in the sandwich ELISA increased (Figure 3, A), resulting in inferior peptide to background ratios for all test peptides including the positive control VL9 peptide (Figure 3, B). This indicates that elevated background signals might instead be attributable to a superior capacity for peptide-free β2M association and complex formation by the HLA-E heavy chain relative to HLA-A*0201 or HLA*1101.

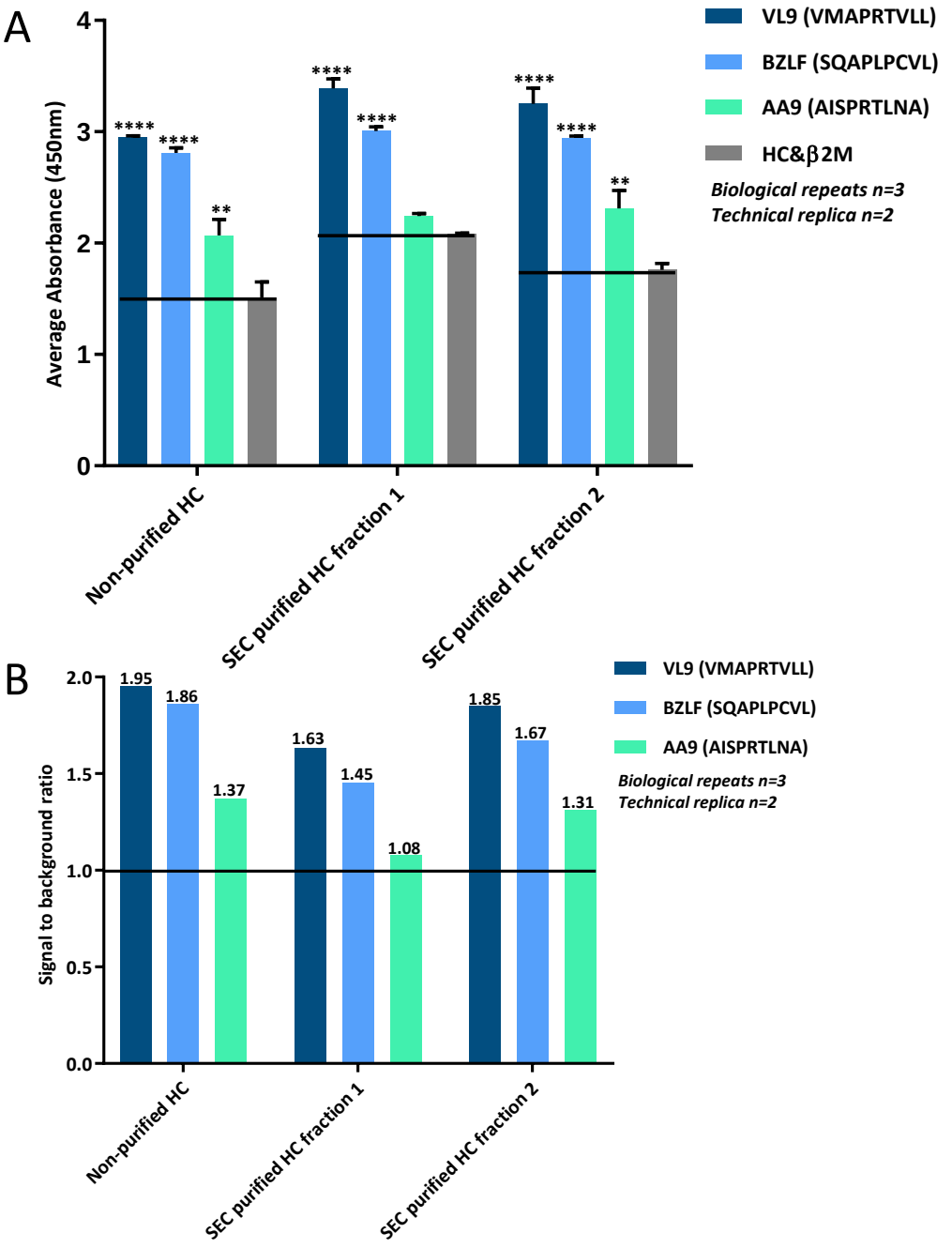


Figure 3: Pre-purification of HLA-E heavy chain in the peptide binding assay

A. Average absorbance readings from peptide binding assays where the HLA-E heavy chain was pre-purified via size exclusion chromatography prior to micro-refolding and sampling on the sandwich ELISA plate. The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes the condition of the HLA-E heavy chain added to the micro-refold. 'Standard' indicates non-purified HLA-E heavy chain, whereas 'SEC purified HC fraction 1' and 'SEC purified HC fraction 2' indicate size exclusion purified HLA-E heavy chain from the leading and tailing fractions of the FPLC elution peak, respectively.

B. Bar chart depicting peptide signal to background ratios for the ELISA data shown in A.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background (HC&β2M) in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Despite some of the inconsistencies experienced, we used the micro-refold sandwich ELISA-based peptide binding assay to explore binding of eleven optimal SIV-derived epitopes identified in RhCMV 68-1-SIV vaccination (Hansen et al., 2016). A statistical analysis of independent assays using Wilcoxon signed-rank tests on matched samples established significant binding for six of the eleven epitopes, including both 'supertopes' against which Mamu-E-restricted $CD8^+$ T cell responses were elicited in 100% of vaccinated rhesus macaques (Figure 4). This statistical approach was chosen to tackle inter-assay variations typical of this technique. These results not only suggest extensive overlap in the peptide binding repertoires of the human and rhesus MHC-E orthologues, but also indicate that HLA-E is capable of binding non-canonical and diverse peptides, albeit with lower relative binding strength than canonical VL9 peptides. These results were published in Hansen, S. et al. (2016).

Optimal 9mer peptide	Peptide ID	Sequence	Recognition frequency	Binding assay significance	Ref
Gag_{65-73}	AE9	APLVPTGSE	44	**	Hansen, S. et al. (2016)
Gag_{72-80}	SN9	SENLKSLYN	34	n/s	
Gag_{92-100}	EA9	EKVKHTEEA	49	n/s	
$Gag_{120-128}$	KS9	KTSRPTAPS	41	**	
$Gag_{132-140}$	GI9	GGNYPVQQI	30	n/s	
$Gag_{201-209}$	QE9	QIIRDIINE	56	*	
$Gag_{259-267}$	VI9	VGNIYRRWI	51	n/s	
$Gag_{276-284}$	RL9	RMYNPTNIL	100	***	
$Gag_{389-397}$	KG9	KPIKCWNCG	61	n/s	
$Gag_{476-484}$	LQ9	LGKQQREKQ	71	**	
$Gag_{482-490}$	EK9	EKQRESREK	100	*	

Figure 4: Table of optimal SIV-derived epitope binding to HLA-E in the micro-refold peptide binding assay

Peptide ID, position within SIV Gag and peptide sequence of the eleven optimal SIV-derived epitopes from RhCMV 68-1 SIV vaccination with previously reported recognition frequencies (the percentage of vaccinated rhesus macaques in which the test peptide elicits Mamu-E-restricted $CD8^+$ T cell responses, *n=42*). Significant binding to HLA-E, as determined by the micro-refold sandwich ELISA-based peptide binding assay is shown. Stars in 'Binding assay significance' column reflect degree of significance in peptide binding by nonparametric t-tests: n/s=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001. Significance was calculated from eleven biological runs with 2 technical replica per test peptide via Wilcoxon signed-rank tests on matched pairs in GraphPad Prism.

However, due to persistently high background signals and considerable standard deviation between biological repeats, alternative peptide binding assay approaches were subsequently explored.

Comparison of micro-refolding versus UV-induced peptide exchange-based HLA-E peptide binding assays

Our lab hypothesised that micro-refolding native HLA-E heavy chain and β2M with test peptides *in vitro* may not reflect the RhCMV 68-1 vector-induced MHC-E protein refolding and trafficking pathway following vaccination *in vivo*. It has been demonstrated that the Rh67-derived VL9 peptide mimic encoded within the RhCMV 68-1 vector is essential for the generation of Mamu-E-restricted $CD8^+$ T cell responses, and that Rh67 can be substituted for UL40, the HCMV equivalent protein which functions analogously (Personal communication, Louis Picker and Klaus Früh). It was previously shown that the UL40 protein and its intrinsically encoded VL9 peptide mimic facilitated cell surface expression of HLA-E in a TAP-independent manner (Tomasec et al., 2000). Thus, it is thought the Rh67 VL9 peptide mimic serves to stabilise Mamu-E refolding permitting its exit from the endoplasmic reticulum and subsequent trafficking to the cell surface. Recycling of cell surface Mamu-E into endosomal/phagosomal

compartments represents a possible route enabling peptide exchange with highly abundant, low affinity SIV-derived epitopes in peripheral intracellular compartments. Although this hypothesis is presently speculative, multiple independent lines of evidence point toward a novel, context-dependent role for MHC-E in presenting phagosomally-derived antigen to αβ $CD8^+$ T cells.

Phagosomal compartments in Mtb-infected dendritic cells are specifically enriched for HLA-E, relative to classical MHC class Ia counterpart molecules (Grotzke et al., 2009). Similarly, it was demonstrated that following upregulation of HLA-E expression upon monocyte to macrophage differentiation, the majority of newly synthesised, β2M-associated HLA-E complexes are retained in LC3- or LAMP1-positive autophagosomal-lysosomal vesicles where they are exposed to exogenous antigenic peptides (Camilli et al., 2016). Further, upregulation of HLA-E in newly differentiated macrophage is uncoupled from classical MHC class I protein expression and does not correlate with superior evasion from NK cell-mediated lysis, perhaps indicating a distinct function for autophagosomally-sequestered HLA-E. Such evidence is suggestive of alternative HLA-E trafficking which intersects with the MHC class II pathway. HLA-E could become re-routed under the unusual conditions of RhCMV 68-1-SIV vaccination or Mtb infection when classical MHC class I trafficking pathways are disrupted.

In terms of the peptide binding assay, it could therefore be predicted that low affinity SIV-derived epitopes provide insufficient binding energy to instigate initial assembly and refolding of native HLA-E heavy chain and β2M *in vitro*. However, such peptides could potentially be exchanged into previously refolded complexes with pre-formed tertiary structures and disulphide bridges. To reflect this, a peptide exchange-based sandwich ELISA binding assay was explored as a potential alternative to the micro-refold approach.

The micro-refolding step of the binding assay was substituted for a UV-induced peptide exchange reaction prior to sampling by sandwich ELISA. HLA-E heavy chain and β2M were

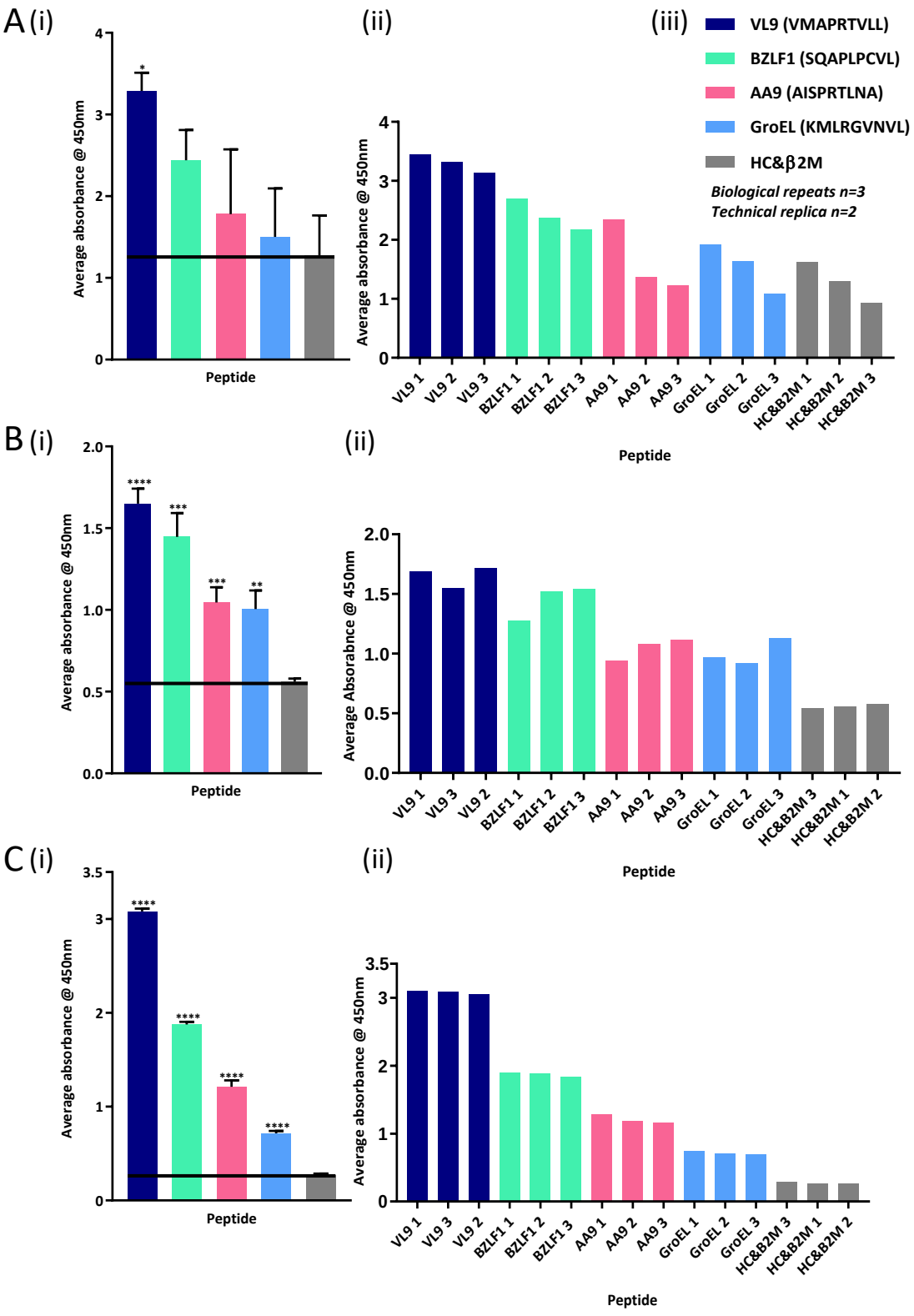
previously refolded with a UV-labile HLA-E binding peptide based on the VL9 sequence in which the solvent exposed Arg at position 5 of VL9 was replaced with a photo-cleavable, unnatural 2-nitrophenylamino acid residue termed J (Toebes et al., 2006). This photo-cleavable VL9-based peptide was synthesised by Paul Hekking at LUMC University. Such conditional peptide ligand technology was originally developed for HLA-A2.1 (Toebes et al., 2006) and subsequently adapted to facilitate peptide exchange in HLA-A1, -A3, -A11, -B7 and -B57 molecules (Bakker et al., 2008) (Brackenridge et al., 2011). Specifically, the nitrophenyl moiety within the substituted beta amino acid is sensitive to UV illumination at >350nm. This results in peptide cleavage and successive dissociation of the degraded peptide fragments from the MHC class I peptide binding groove, as previously validated by mass spectrometric analysis (Toebes et al., 2006). Consequently, in the absence of bound peptide, MHC class I complexes become destabilised and disintegrate. However, complex disintegration may be circumvented by peptide exchange with a rescue peptide that binds the MHC allele in question.

We conducted a direct comparison of the micro-refolding versus UV-assisted peptide exchange binding assays to determine which approach yielded the highest sensitivity and reliability in discriminating low affinity HLA-E peptide binders from background signals (Figure 5). Three biological repeats for a small set of HLA-E binding peptides were included in the comparison. Test peptides comprised AA9, BZLF and GroEL (Figure 1, B), for which a range of relative HLA-E binding affinity had previously been established in the micro-refold binding assay.

Notably, the hierarchy of relative peptide affinity remained constant across both approaches with BZLF exhibiting the strongest and AA9 the weakest binding to HLA-E (Figure 5, A and B). However, the average absorbance reading for each test peptide was statistically significant in the initial UV-induced peptide exchange assay, whereas only the VL9 reading achieved

significance in the micro-refold approach (Figure 5, A). Further, the standard deviation between biological repeats was lower for every peptide tested including the positive and negative controls in the initial UV-induced peptide exchange assay compared to the micro-refold approach (Figure 5, D). Standard deviation ranged between 0.37-0.79 for test peptides in the micro-refold approach but such values were substantially reduced to 0.09-0.15 in the initial UV peptide exchange assay, with further reductions to 0.03-0.07 following optimisation. The standard deviation for three biological repeats of the no peptide negative control decreased by over an order of magnitude in the initial UV exchange approach compared to the micro-refold assay (from 0.49 to 0.02). Considering the promising initial results of the UV exchange approach, which not only exhibited superior reliability with reduced standard deviation but also demonstrated increased sensitivity with significant binding of low affinity peptides to HLA-E, this approach was further optimised.

A multitude of optimisation strategies were investigated, a number of which are discussed in greater detail in subsequent sections. Following optimisation, a greater degree of statistical significance was achieved for all three test peptides and the standard deviation was further reduced for all peptides except the negative control which remained low at 0.02. Assay sensitivity, defined by ability to discriminate between low affinity HLA-E peptide binding and the no-peptide background, was considerably improved post-optimisation – test peptide to background signal ratios for BZLF and AA9 more than doubled relative to the initial UV exchange assay (Figure 5, E).


A (i)
(ii)
(iii)
VL9 (VMAPRTVLL)
BZLF1 (SQAPLPCVL)
AA9 (AISPRTLNA)
GroEL (KMLRGVNVL)
HC&β2M
Biological repeats n=3
Technical replica n=2
Average absorbance @ 450nm
Peptide
VL9 1
VL9 2
VL9 3
BZLF1 1
BZLF1 2
BZLF1 3
AA9 1
AA9 2
AA9 3
GroEL 1
GroEL 2
GroEL 3
HC&B2M 1
HC&B2M 2
HC&B2M 3
B (i)
(ii)
Average absorbance @ 450nm
Average Absorabnce @ 450nm
Peptide
VL9 1
VL9 3
VL9 2
BZLF1 1
BZLF1 2
BZLF1 3
AA9 1
AA9 2
AA9 3
GroEL 1
GroEL 2
GroEL 3
HC&B2M 3
HC&B2M 1
HC&B2M 2
C (i)
(ii)
Average absorbance @ 450nm
Peptide
VL9 1
VL9 3
VL9 2
BZLF1 1
BZLF1 2
BZLF1 3
AA9 1
AA9 2
AA9 3
GroEL 1
GroEL 2
GroEL 3
HC&B2M 3
HC&B2M 1
HC&B2M 2

D

Standard deviation

Assay Method:	VL9 (VMAPRTVLL)	BZLF1 (SQAPLPCVL)	AA9 (AISPRTLNA)	GroEL (KMLRGVNVL)	HC&β2M/No Rescue
Micro-Refold	0.22	0.37	0.79	0.59	0.49
Initial UV Exchange	0.09	0.15	0.09	0.11	0.02
Optimised UV Exchange	0.03	0.03	0.07	0.03	0.02

E

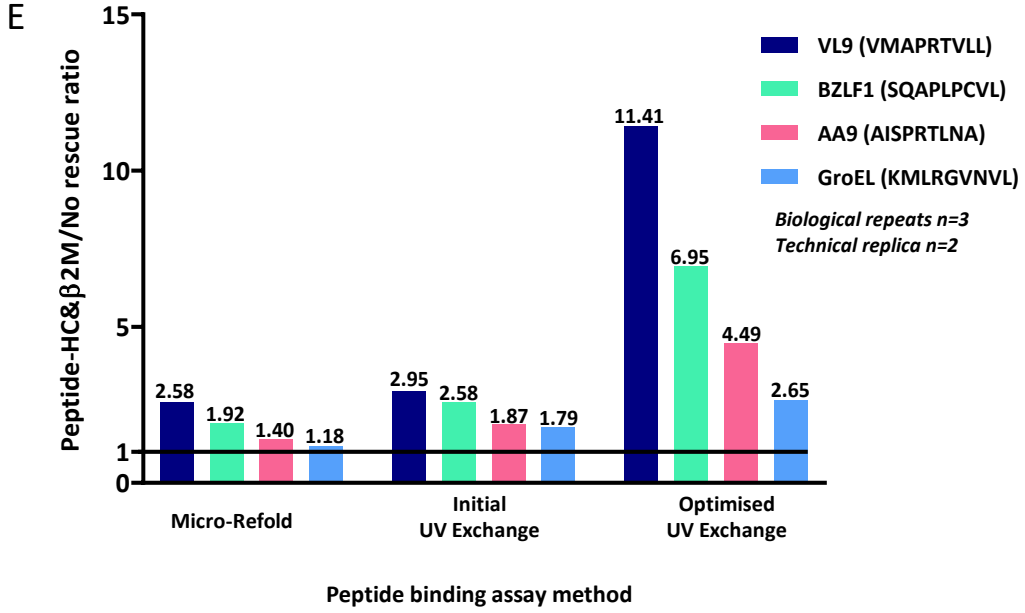


Figure 5: Comparison of micro-refold versus UV exchange peptide binding assays

A. (i) Average absorbance readings for 3 HLA-E binding pathogen-derived peptides sampled in the micro-refolding peptide binding affinity assay. The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes test peptides, which are colour-coded according to the figure legend in A (iii). **(ii)** Bar chart depicting variance between biological repeats in the micro-refold peptide binding assay, with each bar representing a single repeat. **(iii)** Figure legend for Figure 5 A-E with peptide ID and sequences denoted. **B. (i)** Average absorbance readings from initial assays using the UV-exchange peptide binding affinity approach. **(ii)** Bar chart depicting variance between biological repeats in the non-optimised UV exchange binding assay. **C. (i)** Average absorbance readings from the optimised UV exchange peptide binding assay approach. **(ii)** Depiction of variance between biological repeats in the optimised UV exchange peptide binding assay technique. **D.** Table of standard deviation for the micro-refold and the initial and optimised UV exchange peptide binding assay approaches. **E. (i)** Bar chart depicting peptide signal to background ratios for the micro-refold and the initial and optimised UV exchange peptide binding assay approaches.

Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue

*with peptide free background in grey. Negative controls for micro-refold approach consisted of HLA-E heavy chain and β2M only whereas the UV peptide exchange approach negative control comprised previously refolded HLA-E complexes with no added rescue peptide in the exchange reaction. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Optimisation of the peptide exchange-based binding assay

Following the initial identification of UV-induced peptide exchange as a superior approach to micro-refolding, a multitude of assay conditions were further investigated to optimise performance.

Photo-assisted peptide exchange buffer conditions

First, buffer conditions were optimised in the UV-induced peptide exchange reaction prior to sampling by sandwich ELISA. Three buffer systems were tested including the Lutrol-Tris-maleate-based buffer previously used in the micro-refolding peptide binding assay, a Tris-based buffer previously used for UV-induced peptide exchange with classical class I MHC molecules (Toebes et al., 2006), and finally an L-Arginine-Tris-based buffer traditionally used in macro-refolding of class I MHC proteins (Garboczi et al., 1992). The L-Arginine-Tris-based buffer yielded lower peptide-free background signals, concomitantly improving peptide to background ratios for all 4 medium-low affinity peptides tested (Figure 6).

It was previously established via micro-refolding-based peptide binding assays and size exclusion chromatography following macro-protein refolding, that SIV-derived epitopes identified in RhCMV 68-1 vaccination bind HLA-E weakly compared to VL9 peptides. Furthermore, the majority bind with such low affinity that average absorbance signals generated in the peptide binding assay lie marginally above the peptide-free background. Therefore, assay sensitivity, as indicated by statistically significant discrimination of low affinity

peptide binding from background, was a key feature selected for during optimisation of the UV-induced peptide exchange-based method.

The peptide to background ratio for the weakest binding peptide tested, QE9 (QIIRDIINE), increased from 1.12 and 0.71 respectively in the Tris and Lutrol-Tris-maleate-based buffers to 2.44 in the L-Arginine/Tris-based buffer. Furthermore, statistically significant binding was achieved only in the L-Arginine/Tris-based buffer for QE9. The assay was therefore repeated in the L-Arginine-Tris-based buffer system to validate increased sensitivity. QE9 binding remained statistically significant and the peptide to background ratios were again higher for all 4 peptides tested compared to the Tris-based or Lutrol-maleate buffers. Hence, the L-Arginine/Tris-based buffer was adopted in all subsequent optimisation and peptide screening assays.

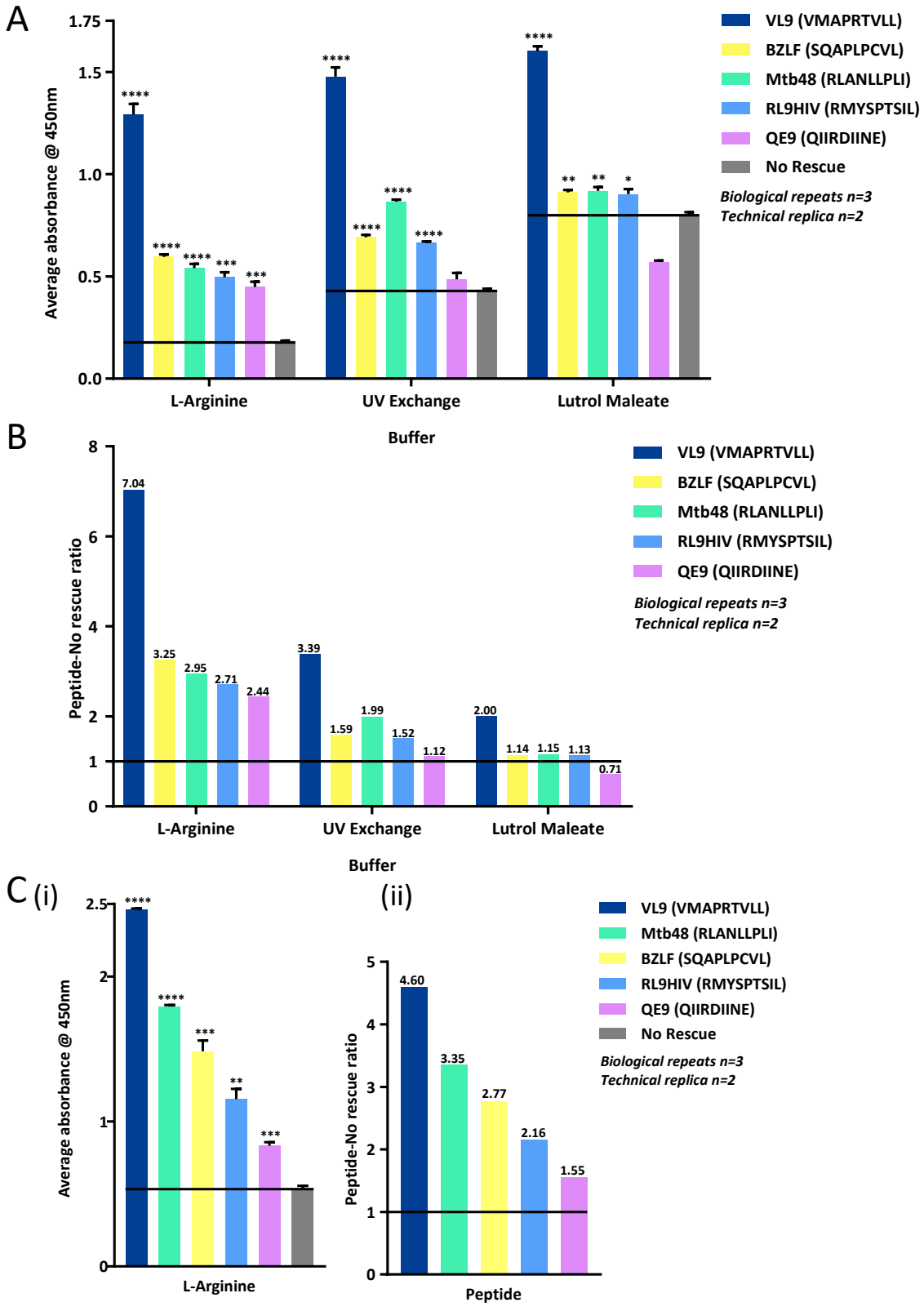


Figure 6: UV-induced peptide exchange buffer optimisation

A. Average absorbance readings from 3 different buffer systems trialled in the UV-induced peptide exchange sandwich ELISA peptide binding assay. The Y-axis denotes average absorbance readings at

450nm and the X-axis denotes the buffer system. Test peptide IDs and sequences are specified in the figure legend. **B.** Peptide to background signal ratios for data shown in A. **C. (i)** Average absorbance readings from an independent repeat assay using the L-Arginine UV-induced peptide exchange buffer. **(ii)** Peptide to background signal ratios of data shown in C (i).
*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

The buffer into which the UV-induced peptide exchange reaction is diluted for sandwich ELISA sampling was also optimised. According to the micro-refolding ELISA protocol, the micro-refold or peptide exchange reaction is diluted 1:100 into BSA before addition to the ELISA plate as a step to ensure the constant presence of blocking solution. However, to investigate the possibility of BSA obstructing HLA-E binding to the 3D12 coating antibody, the UV exchange reaction was diluted 1:100 back into the UV reaction buffer (Figure 7). Diluting the Tris-based UV peptide-exchange buffered reaction back into the Tris-buffer improved the signal to background ratios for the test peptides, Mtb48 and RL9HIV, compared to dilution into BSA. However, these improved ratios were still inferior to those generated for the same two peptides when the UV-induced peptide exchange reaction was carried out in the L-Arginine-Tris-based buffer and subsequently diluted into BSA.

Dilution of the L-Arginine-Tris-buffered peptide exchange reaction into the same buffer resulted in reduced signals for all four medium to low affinity peptides tested with two peptides losing statistical significance. Furthermore, a reduction in the average absorbance reading from 1.29 to 0.59 was observed for the positive control peptide, VL9. Therefore, UV-induced peptide exchange reactions were conducted in L-Arginine-Tris buffers and continually diluted into BSA in subsequent assays.

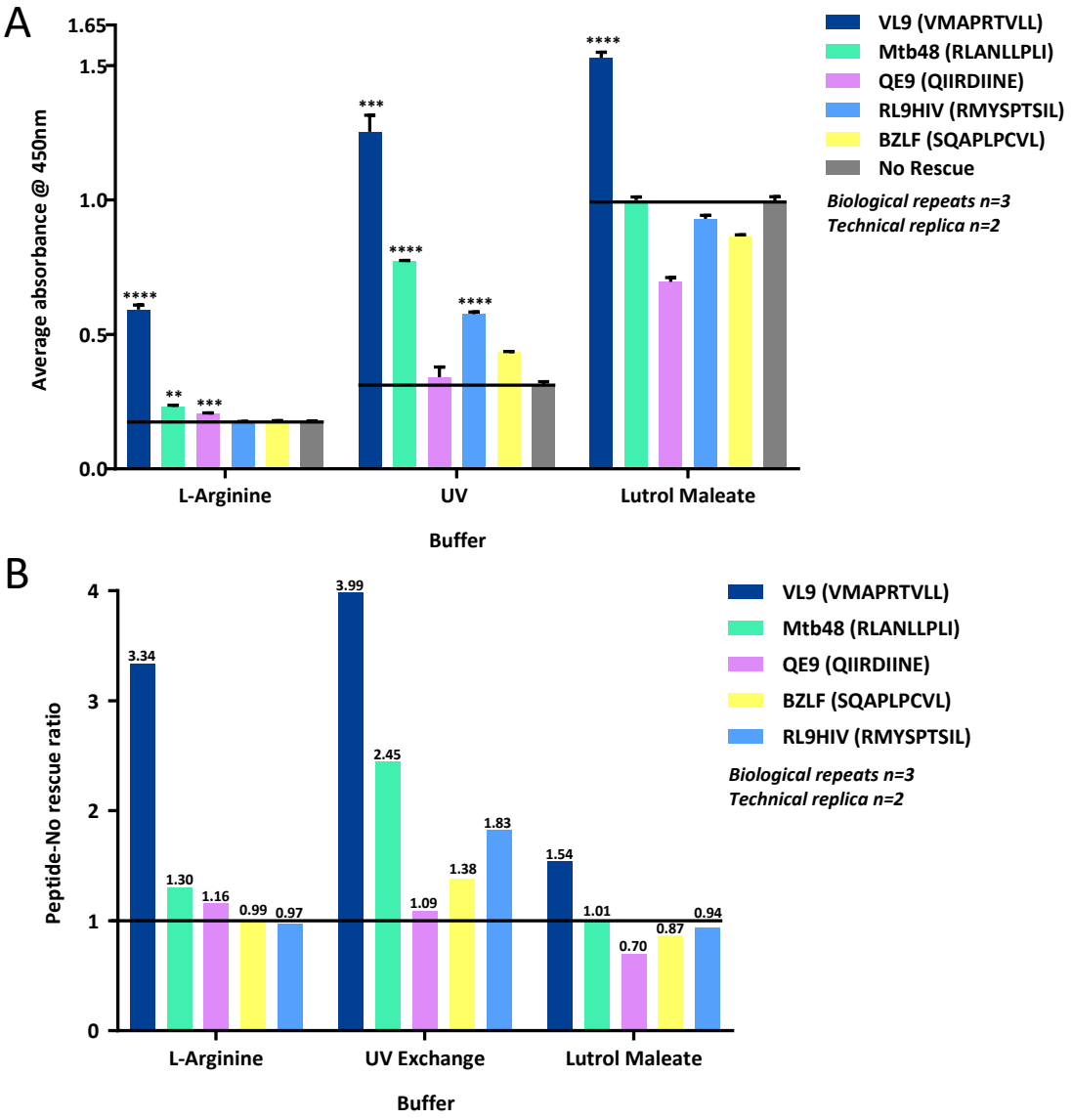


Figure 7: Optimisation of peptide-exchange reaction dilution buffer conditions prior to ELISA sampling

A. Average absorbance readings from 3 different buffer systems used not only for the UV-induced peptide exchange reaction but also instead of BSA for protein sample dilution prior to addition to the ELISA plate. The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes the buffer system. **B.** Peptide to background signal ratios of data shown in A.

Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error

*bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

To determine which component of the L-Arginine-Tris-based buffer system was responsible for reduced peptide-free background signals, further assays were carried out in which L-Arginine monohydrochloride or the glutathione redox system were individually excluded (Figure 8). In the absence of L-Arginine monohydrochloride, the average absorbance signal for the peptide-free background increased from 0.18 to 0.65. Whereas, in the absence of the redox system the background signal remained unaffected at 0.18.

Such results indicate that L-Arginine monohydrochloride is the component responsible for peptide-free background reduction. The mechanism for this is unknown, however it is possible that the presence of 400mM L-Arginine monohydrochloride competes the UV-degraded peptide fragments out the peptide binding groove, accelerating the destabilisation and dissociation of the protein complex in the absence of rescue peptide. Although the glutathione redox system didn't appear to play a role in decreasing background signals, the peptide to background ratio and degree of statistically significant binding both decreased for QE9 in its absence. Hence, the glutathione redox system was continually included in subsequent assays.

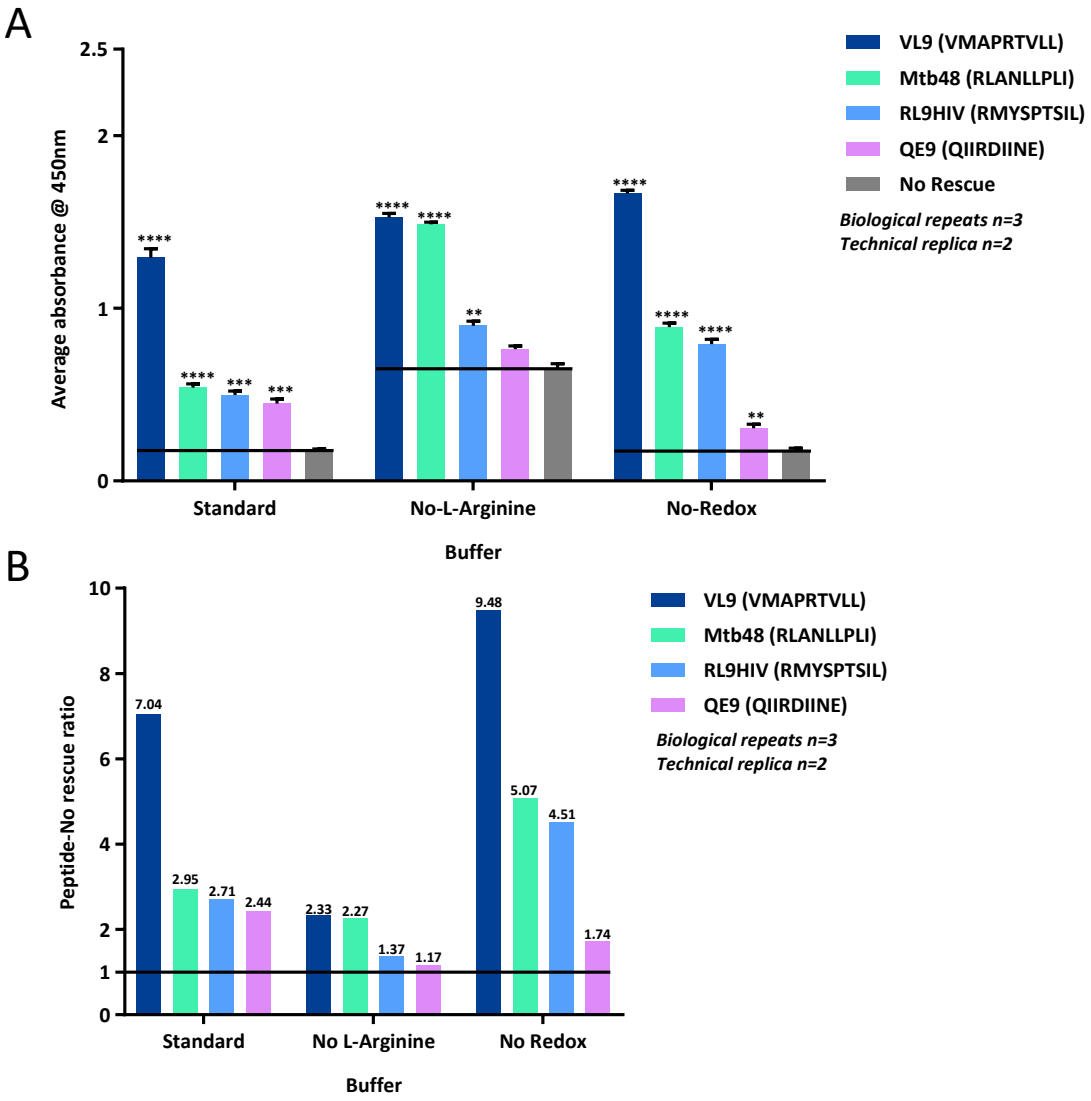


Figure 8: Evaluation of L-Arginine-based refolding buffer composition

A. Average absorbance readings from 3 peptide binding ELISA optimisation assays in which components of the L-Arginine-based refolding buffer were individually excluded. The 'standard' buffer contains all components of the L-Arginine-based macro-refolding buffer. L-Arginine monohydrochloride was excluded from the 'No L-Arginine' assay and the glutathione redox system excluded from the 'No-Redox' assay. The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes buffer composition. **B.** Bar chart depicting peptide to background signal ratios for data shown in A. *Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error*

*bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

L-Arginine monohydrochloride concentration

Given its capacity to reduce peptide-free background signals, the concentration of L-Arginine monohydrochloride was further explored. Although background signals further decreased at higher concentrations and plateaued at 800mM, peptide signals were also affected. The lowest peptide to background ratios for all peptides were observed at 1600mM, which constituted the highest L-Arginine monohydrochloride concentration tested (Figure 9). Similar peptide to background ratios were observed at 400mM and 800mM for both RL9HIV and RL9SIV, however greater standard deviation and decreased significance at 800mM resulted in the continued use of 400mM L-Arginine monohydrochloride in subsequent assays.

Although the degree of statistical significance remained constant for VL9 binding across all L-Arginine monohydrochloride concentrations tested, the average absorbance reading decreased from 1.93 at 200mM to 0.87 at 1600mM. It is therefore possible that at higher concentrations the mechanism employed by L-Arginine monohychloride - speculatively, to compete UV-degraded peptide fragments out the peptide binding groove - also extends to destabilising peptide exchange despite an excess of optimal HLA-E binding peptide.

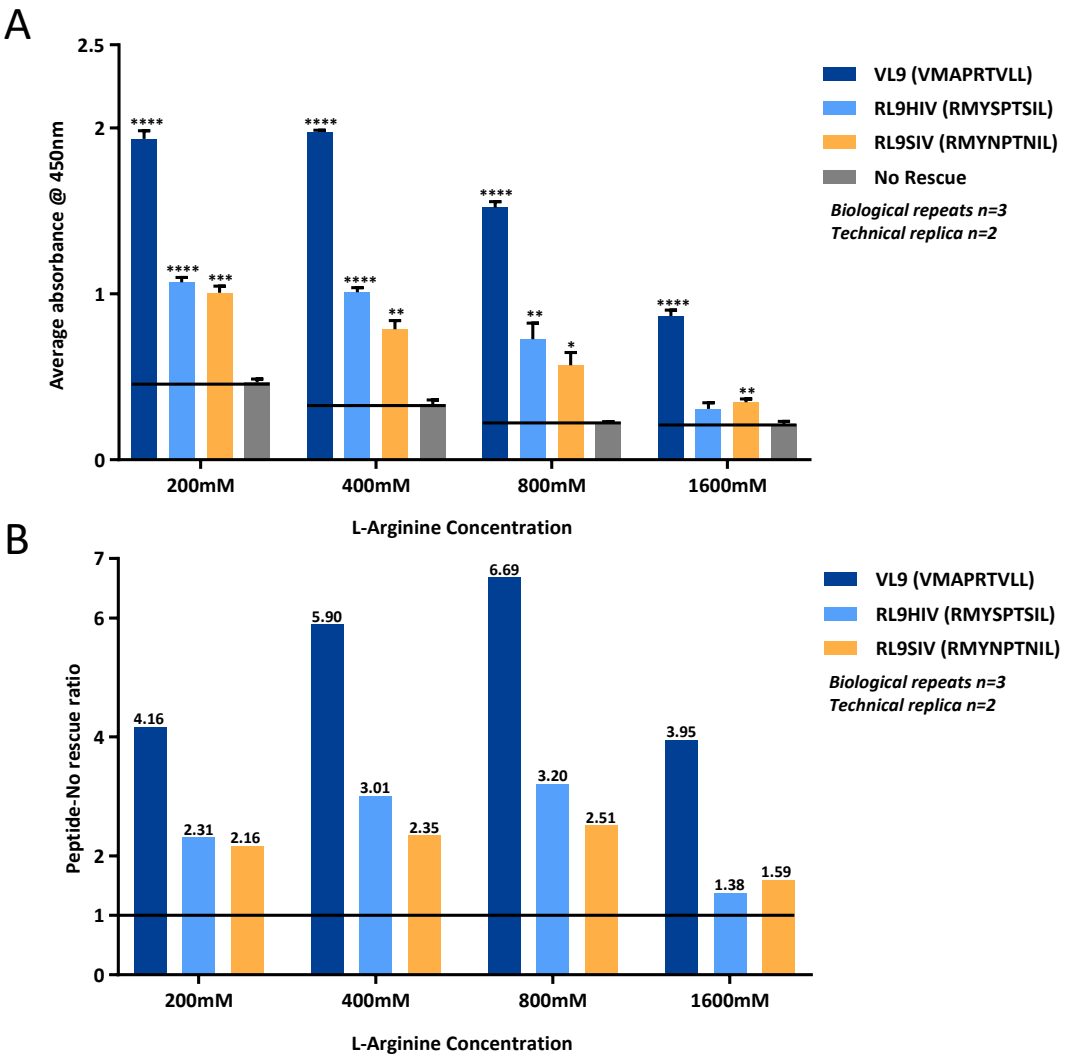


Figure 9: Optimisation of L-Arginine monohydrochloride concentration

A. Average absorbance readings from peptide binding assays in which L-Arginine monohydrochloride concentration was varied in the UV-induced peptide exchange reaction. The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes L-Arginine concentration. **B.** Bar chart depicting peptide to background signal ratios of data shown in A.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Optimisation of pH conditions in the UV-assisted peptide exchange reaction

UV-induced peptide exchange reactions were carried out over a pH range of 5.5 to 9 to investigate the potential impact of pH on peptide binding to HLA-E. Although a basic pH 8 typically favours disulphide bond formation facilitating overall complex assembly (Trivedi et al., 2009), a more acidic pH between 6-7 was found to optimally support peptide loading in six of seven classical MHC class I molecules tested previously (Pedersen et al., 2001). The micro-refolding peptide binding assay in which HLA-E complexes were assembled from individual native protein components required *de novo* disulphide bridge formation. However, it could be hypothesised that the UV-induced peptide exchange approach would function optimally at a lower pH since pre-refolded HLA-E material should eliminate the need for conditions that support disulphide bridge formation.

A neutral pH 7 was originally selected for reactions involving conditional photo-assisted peptide exchange technology to minimise destabilisation of β2M-bound MHC class I complexes (Toebes et al., 2006). However, as previously mentioned, the UV-induced peptide exchange platform was originally developed for classical MHC class I proteins whose tolerance to pH deviation away from neutral conditions might differ from that of their non-classical counterpart molecule, HLA-E. The presence of stable, β2M-associated HLA-E molecules in acidic Mtb-infected phagosomes or autophagosomal-lysosomal compartments in newly differentiated macrophage (Camilli et al., 2016) might reflect that β2M-bound HLA-E retains stability at lower pH conditions - similar to the non-classical lipid antigen presentation molecule, CD1 (Odyniec et al., 2010).

The pH optimisation peptide-exchange assays revealed that pH 8 yielded the greatest test peptide to background ratios for both RL9HIV and QE9 and thus this pH was continually adopted in subsequent assays (Figure 10).

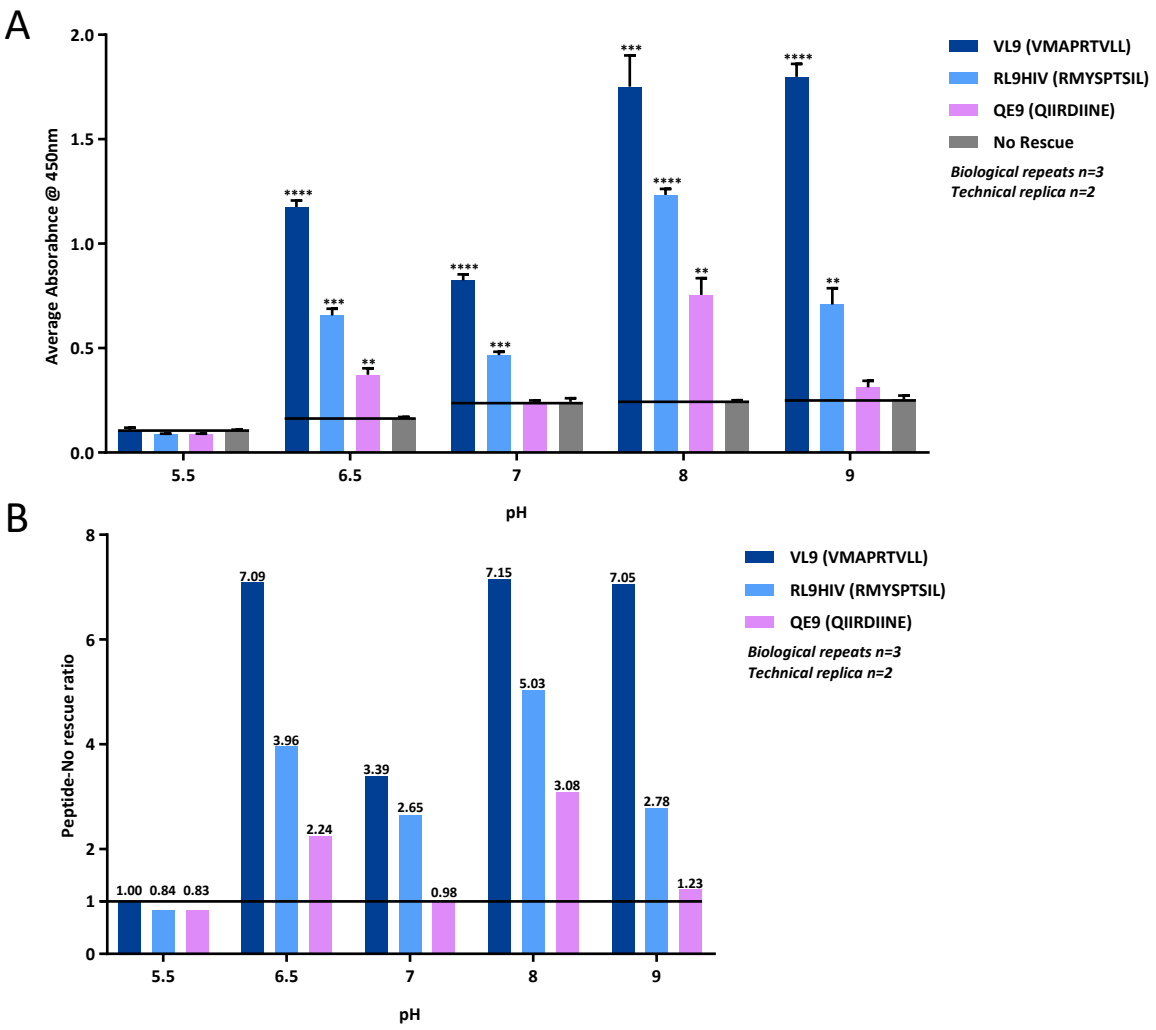


Figure 10: pH optimisation in the UV-induced peptide exchange reaction

A. Average absorbance readings from peptide binding assays where the pH of the L-Arginine-based buffer in the peptide-exchange reaction was varied between 5.5 and 9. The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes buffer pH. **B.** Bar chart depicting peptide to background signal ratios of data shown in A.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Sandwich ELISA protein incubation temperature

The temperature of protein sample immobilisation on the ELISA plate following the addition of diluted UV-induced peptide exchange reaction samples, was explored. A range of temperatures from 37-4°C was tested over a 1.5 hour incubation period in addition to 4°C for 24 hours, which has previously been reported to increase ELISA sensitivity due to reduced non-specific hydrophobic binding (Figure 11). However, the 1.5 hour incubation period at room temperature produced the greatest peptide to background ratios in addition to the greatest degrees of statistically significant binding for RL9HIV and RL9SIV, and was hence selected for subsequent assays.

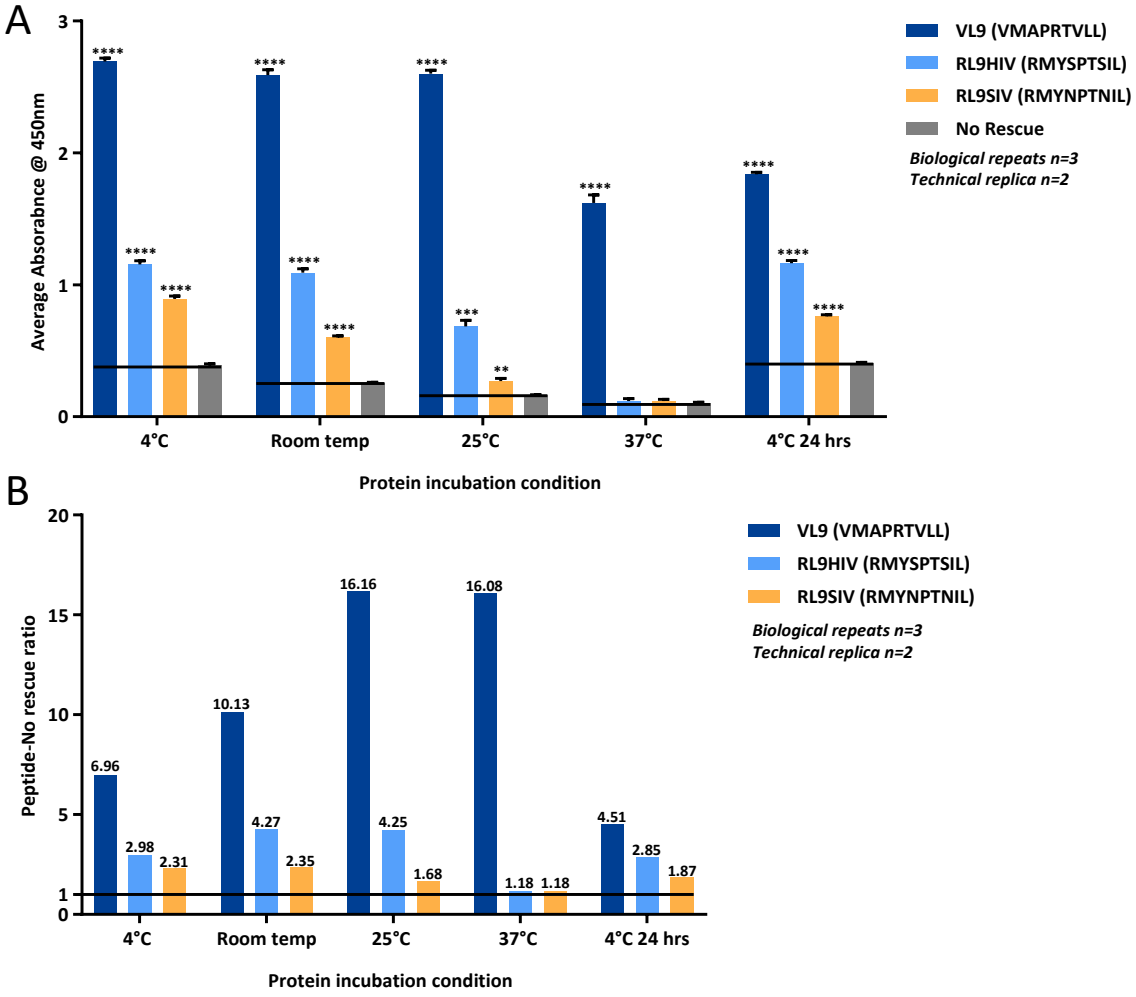


Figure 11: Optimisation of sandwich ELISA protein incubation temperature

A. Average absorbance readings from peptide binding assays where the temperature of protein sample immobilisation on the sandwich ELISA plate was varied between 37-4°C. The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes protein incubation temperature. **B.** Bar chart depicting peptide to background signal ratios of data shown in A.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Duration of UV irradiation in the peptide exchange reaction

The optimal duration of UV exposure for the peptide exchange reaction was investigated. The recommended duration of UV exposure in previously published protocols using UV-labile ligand technology is 1hr (Toebes et al., 2006). However, we hypothesised that extended exposure could enable further cleavage of UV-sensitive peptides resulting in enhanced destabilisation of previously refolded HLA-E complexes in the absence of rescue peptide, and that this would in turn further reduce background signals and concomitantly improve peptide to background ratios. Therefore we exposed the peptide exchange reaction to UV radiation for a period of zero, one or two hours before sampling by sandwich ELISA. As predicted, the peptide to no rescue background ratios of all three test peptides, RL9HIV, RL9SIV and EK9, and the positive control VL9 peptide were higher after two hours of UV exposure (Figure 12).

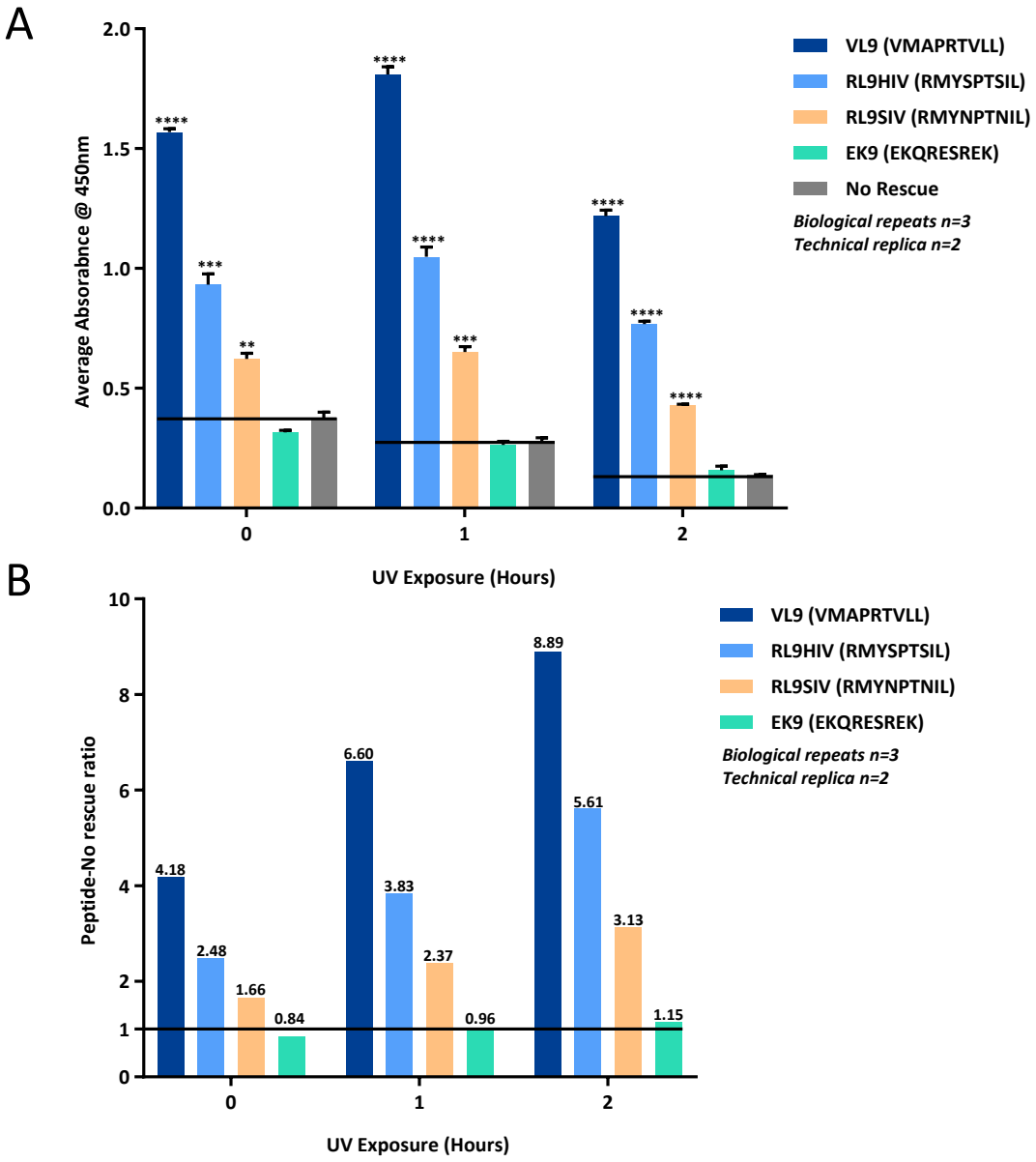


Figure 12: Optimisation of UV exposure duration in the peptide exchange reaction

A. Average absorbance readings from three sandwich ELISAs where the peptide exchange reaction was sampled at 0, 1, and 2 hours of UV exposure. The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes duration of UV exposure. **B.** Peptide to background signal ratios of data shown in A.

Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error

*bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Further optimisation was conducted to explore whether extended UV exposure beyond the 2 hour period continued to improve peptide to no rescue ratios of HIV and SIV-derived epitopes. However, further investigation revealed that peptide to background ratios of all three test peptides including the positive control VL9 were similar after two or three hours of UV exposure (Figure 13). Furthermore, RL9HIV and RL9SIV peptide to background ratios began to decrease at the four hour UV exposure time-point, perhaps due to radiation damage reducing overall protein yield in the reaction. Two hours was therefore selected as an optimal length of UV exposure for subsequent peptide exchange reactions.

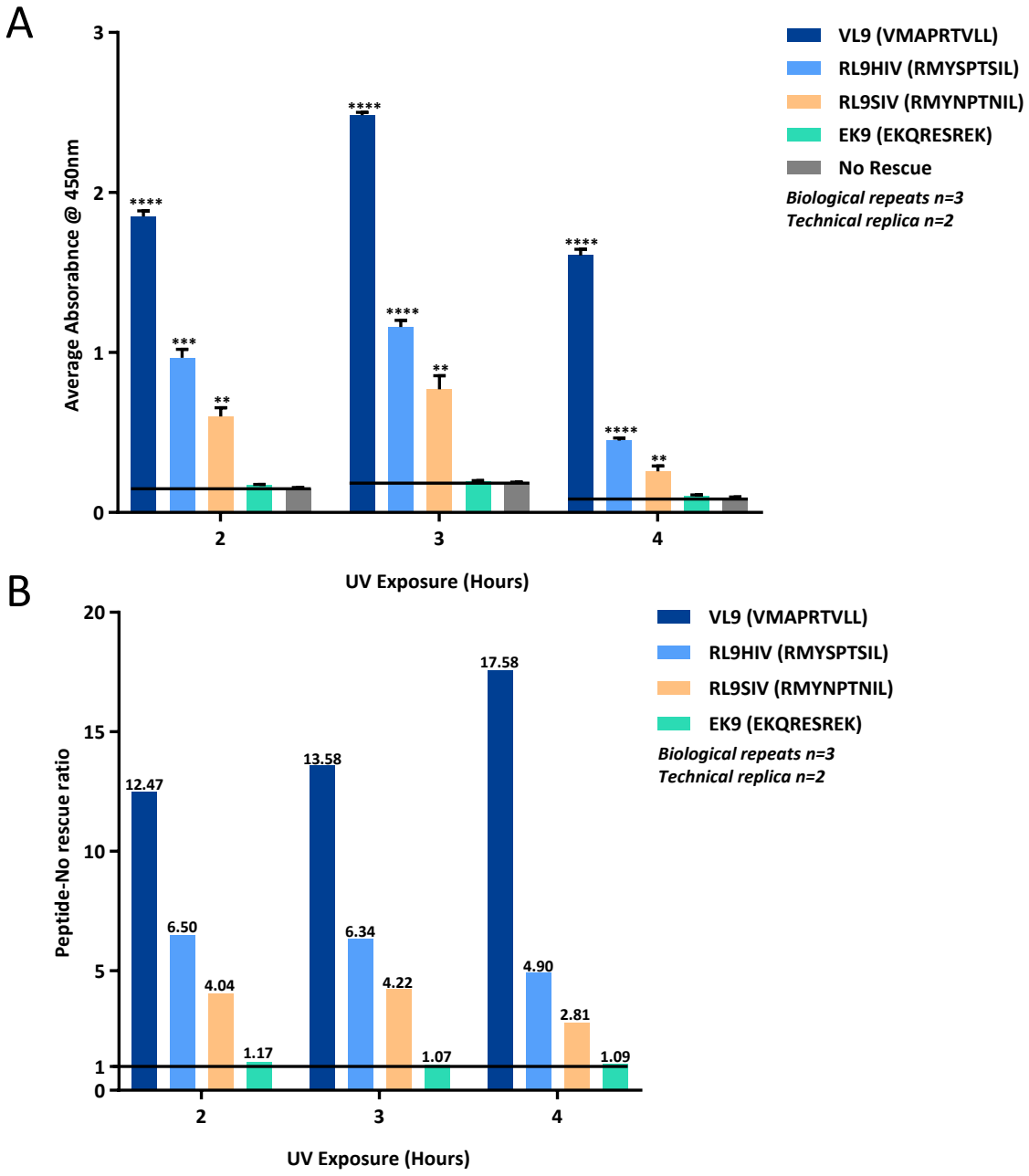


Figure 13: Optimisation of UV exposure duration in the peptide exchange reaction

A. Average absorbance readings from 3 sandwich ELISAs where the peptide exchange reaction was sampled at 2, 3, and 4 hours of UV exposure. The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes duration of UV exposure. **B.** Bar chart depicting peptide to background signal ratios of data shown in A.

Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error

*bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Peptide exchange reaction in the absence of UV illumination

As previously discussed, extended long-wavelength UV irradiation from one to two hours resulted in increased signal to background ratios for all peptides tested in the peptide binding assay (Figure 12 and 13). However, it remained uncertain whether increased assay sensitivity was attributable to prolonged photo-stimulation or simply due to the extended time period over which peptide exchange could occur. Furthermore, although exposure to light within the UV spectrum was originally deemed necessary for photo-cleavage and resultant exchange of HLA-A2.1-restricted UV-labile peptides (Toebes et al., 2006), this had not been fully elucidated for the UV-labile HLA-E binding peptide (VMAPJTLVL). Experiments to rule out the possibility of UV-independent peptide exchange were therefore conducted – particularly since unassisted peptide exchange was previously observed by Professor Geraldine Gillespie of Oxford University in the blue native gel system whereby pre-refolded VL9-bound HLA-E was capable of converting to Mtb44 peptide-loaded material in the presence of excess Mtb44, as evidenced by distinct gel band positioning for VL9 versus Mtb44 peptide-loaded HLA-E (Walters et al., 2018). Peptide exchange was thus subsequently carried out over a similar time course in the absence of UV illumination.

When the peptide exchange reaction was conducted at 4°C the peptide to no-rescue signal ratios exhibited similar trends over the zero to three hour time-points in the presence or absence of UV irradiation. This indicates that peptide exchange does not require UV incubation (Figure 14). This optimisation was also conducted at room temperature in the absence of UV illumination, however at this temperature the peptide to background ratios did not improve after the zero hour time point (Figure 15).

Interestingly, the trend of peptide to background ratios beyond the three hour time-point for the 4°C reaction in the absence of UV differed to that observed beyond three hours with UV exposure. In the presence of UV radiation the ratios began to decrease at four hours, perhaps from loss of protein due to radiation damage. However, the peptide to no-rescue ratios continue to increase beyond the three hour time-point up to the five hour time-point in the absence of UV exposure and potential radiation damage. Furthermore, at five hours, the ratios are similar to those observed at the 24 hour time point in the absence of UV radiation.

Of specific interest are peptide binding assay signals generated by the SIV-derived EK9 'supertope'. This epitope is of particular importance since it is recognised by 100% of vaccinated rhesus macaques despite possessing a highly charged sequence with Lys at both primary anchor positions 2 and 9 – such primary anchors are not predicted to bind HLA-E. Further, macro-refolding and size exclusion chromatography of HLA-E refolded in the presence of EK9 result in FPLC elution peaks which are similar in scale to those observed for HLA-E refolded with β2M in the absence of peptide. Such results indicate EK9 binds HLA-E very weakly, hence conditions that discriminate between EK9 binding and no rescue backgrounds in the peptide binding assay must be highly sensitive and enable the detection of very low affinity peptide binders. Therefore, peptide binding assay conditions that best support EK9 binding to HLA-E were selected in optimisation experiments for subsequent assays. Considering optimal peptide to background ratios for EK9 were generated in the five hour peptide exchange time-point, this duration was adopted in subsequent peptide exchange reactions which were conducted in the absence of UV irradiation.

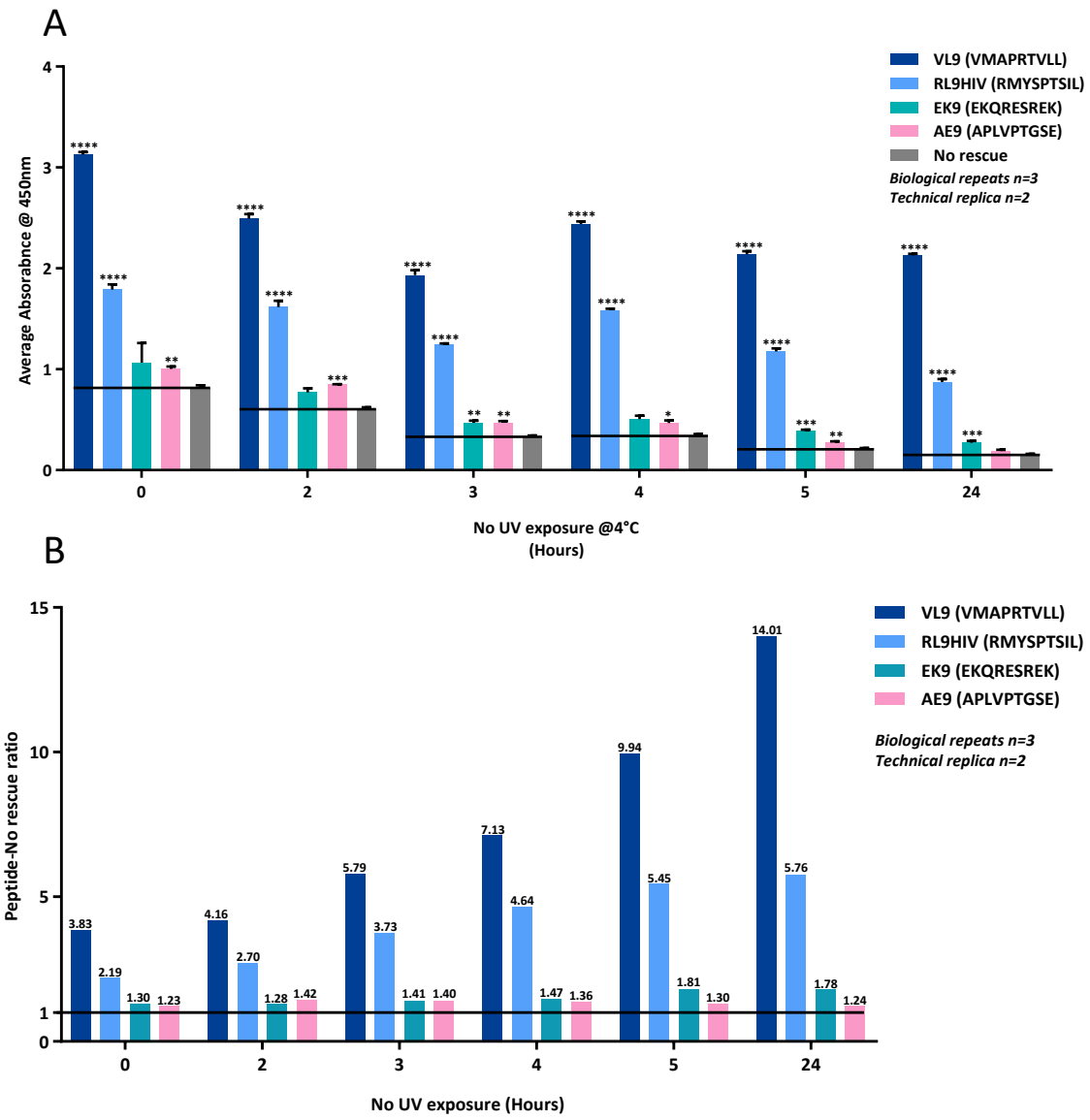


Figure 14: Optimisation of peptide exchange duration in the absence of UV exposure at 4°C

A. Average absorbance readings from 6 sandwich ELISAs where the peptide exchange reaction was sampled at 0, 2, 3, 4, 5, and 24 hours in the absence of UV exposure. The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes duration of the peptide exchange reaction at 4°C.

B. Bar chart depicting peptide to background signal ratios of data shown in A.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background in grey. Error bars depict standard deviation of biological repeats. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

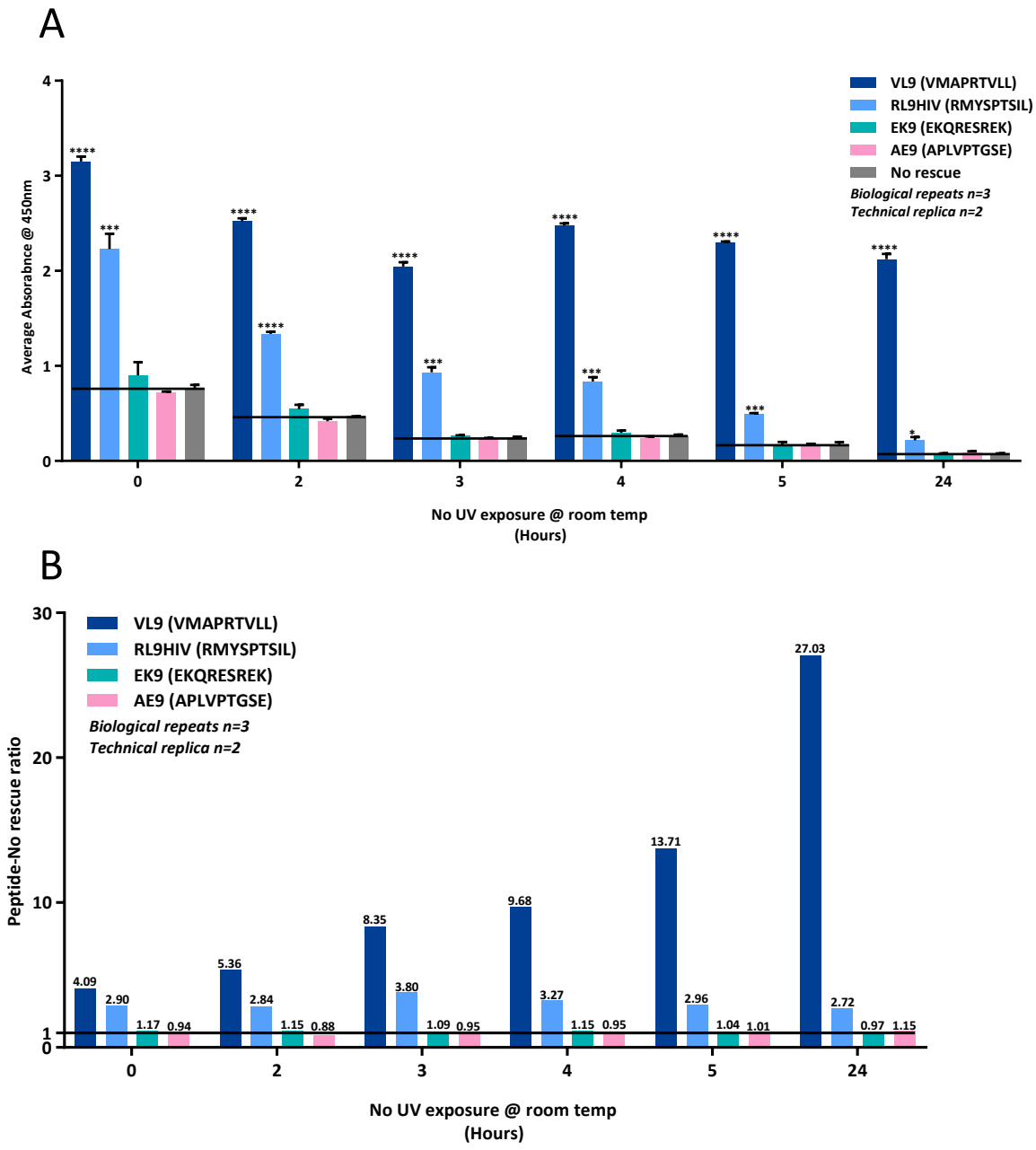


Figure 15: Optimisation of peptide exchange duration in the absence of UV exposure at room temperature

A. Average absorbance readings from 6 sandwich ELISAs where the peptide exchange reaction was sampled at 0, 2, 3, 4, 5, and 24 hours in the absence of UV exposure. The Y-axis denotes average absorbance readings at 450nm and the X-axis denotes duration of the peptide exchange reaction at room temperature. **B.** Bar chart depicting peptide to background signal ratios of data shown in A. *Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background in grey. Error bars depict standard deviation of biological repeats.*

*Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Peptide exchange-based binding assay screening of HLA-E peptide panels

Having optimised the peptide exchange-based binding assay to a level that achieved strong and reliable binding signals for weak binding epitopes, peptide panels comprising HLA-E-restricted epitopes, or peptides algorithmically-predicted to bind HLA-E, derived from a range of pathogen and self-proteins were assembled for screening using this improved system.

Peptide panel 1

Previously described pathogen- and self-derived HLA-E binding peptides

The first panel comprised a selection of previously described HLA-E-restricted peptides derived from a range of self-proteins or pathogens, including from Salmonella, EBV, Hepatitis C, Influenza and HIV (Figure 16, A) (Lo et al., 2000) (Nattermann et al., 2005a) (Ulbrecht et al., 1998) (Tomasec et al., 2000) (Nagarajan et al., 2012) (Michaëlsson et al., 2002) (Nattermann et al., 2005b). All except one of the nine peptides exhibited statistically significant binding to HLA-E, and a range of relative binding strengths were observed (Figure 16, B). The peptide (FYAEATPML) derived from the self-protein, Fam49B, demonstrated the strongest binding to HLA-E. Although the Influenza matrix protein-derived peptide, InflM (ILGFVFTLT), showed significant binding, the average absorbance reading was only marginally above that of the peptide-free negative control signal. Interestingly, the four peptides in this panel which contain canonical position 2 Met, three of which also contain canonical position 9 Leu, demonstrated weaker binding than the Fam49B- and the EBV BZLF-derived peptides which contain a non-canonical Tyr and Gln at position 2, respectively. Further, the three peptides containing both canonical primary anchor residues in addition to the positive control VL9

peptide, generate a wide range of relative binding affinities. This indicates that other key determinants impact HLA-E peptide binding affinity.

Each peptide in panel 1 is nonameric except the Hepatitis C virus (HCV) Core protein-derived peptide (YLLPRRGPRL) which comprises ten amino acids. HLA-E binding properties for the HCV decamer are conceivable considering MHC class I molecules typically present peptides of eight-ten amino acids in length. However the HCV 10mer contains an internal nonameric sequence (LLPRRGPRL) which aligns more optimally to the canonical binding motif of HLA-E, with hydrophobic residues at position 2 and 9. We thus compared binding of the LLPRRGPRL nonamer to the original decamer peptide. The average absorbance reading for the nonamer was more than double that of the original 10mer and also bound HLA-E with a greater degree of significance, indicating that this shorter peptide is perhaps the optimal epitope (Figure 16, C).

A

Peptide panel 1

Peptide ID	Sequence	Aas	Organism	Protein	Ref.
GroEL Salmonella	KMLRGVNVL	9	*Salmonella enterica* serovar Typhi	GroEL$^{15-23}$	Salerno-Goncalves, R. et al. (2004)
EBV BZLF1	SQAPLPCVL	9	*Epstein Barr Virus*	BZLF1$^{39-47}$	Jørgensen, P,B. et al. (2012)
InflM	ILGFVFTLT	9	*Influenza*	Matrix$^{59-67}$	Ulbrecht, M. et al. (1998)
HCV Core 10mer	YLLPRRGPRL	10	*Hepatitis C Virus*	HCV core$^{35-44}$	Jiang, H. et al. (2010)
AA9 HIV p24	AISPRTLNA	9	*Human Immunodeficiency virus*	P24$^{14-22}$	Davis, Z,B. et al (2016)
Prdx5	AMAPIKTHL	9	Human	Prdx5$^{52-60}$	Sensi, M. et al. (2009)
Hsp4 Self	QMRPVSRVL	9	Human	Hsp60sp$^{10-18}$	Michaëlsson, J. et al. (2002)
Hsp5 self	GMKFDRGYI	9	Human	Hsp60.4$^{216-224}$	Michaëlsson, J. et al. (2002)
Fam49B protein	FYAEATPML	9	Human	Fam49b$^{190-198}$	Nagarajan, N,A. et al. (2012)

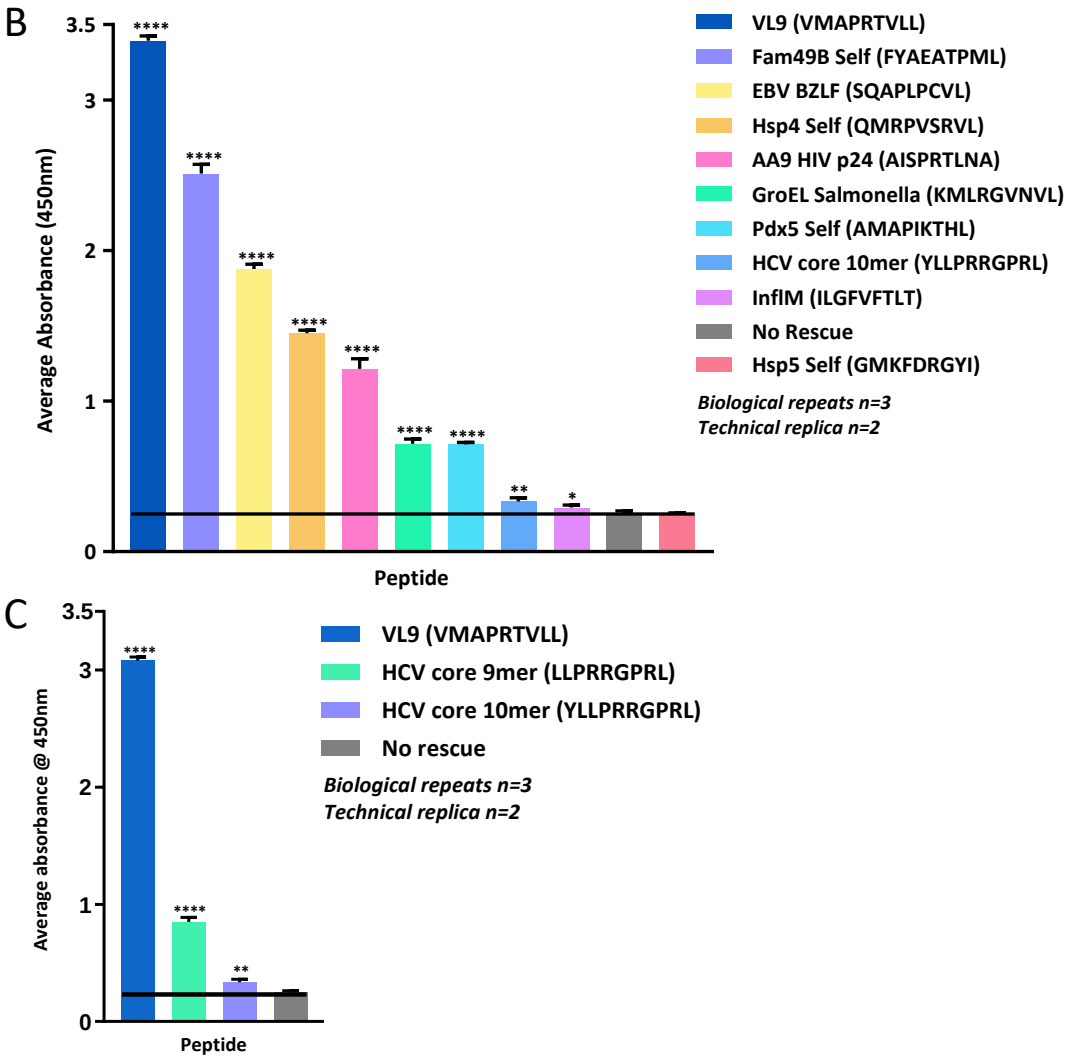


Figure 16: Screening of peptide panel 1 in the optimised peptide exchange-based binding assay

A. Table detailing peptides in panel 1. Peptide ID, sequence, length and origin are reported with the corresponding reference. **B.** Average absorbance readings for pathogen- and self-derived HLA-E binding peptides in panel 1 using the optimised peptide binding assay method. **C.** Peptide binding assay comparison of the previously described Hepatitis C virus-derived decamer peptide and the internal

nonameric variant lacking the N-terminal Tyr.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background (No rescue) in grey. Error bars depict standard error of the mean. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Peptide panel 2

11 optimal SIV-derived RhCMV 68-1 vaccine-identified Mamu-E restricted epitopes

We next screened the eleven optimal SIV-derived epitopes identified in RhCMV 68-1-vectored vaccination which were previously tested in the micro-refolding ELISA approach with significant binders listed in Figure 4 (Hansen et al., 2016). Seven of the eleven Mamu-E-restricted epitopes demonstrated statistically significant binding to the human orthologue HLA-E in the optimised peptide exchange-based assay (Figure 17a) – five of which also demonstrated significant binding in the micro-refolding approach, including RL9, EK9, AE9, QE9 and LQ9. Unsurprisingly, the strongest HLA-E binding peptide was the SIV-derived supertope, RL9SIV, which contained the canonical primary anchor residues for HLA-E – position 2 Met and position 9 Leu. However, six epitopes containing non-canonical primary anchor residues also demonstrated significant binding, including the second supertope, EK9, which contains a non-canonical Lys at both positions 2 and 9. Further, a range of non-canonical amino acids with biochemically dissimilar side chains to those present in canonical VL9 peptides were tolerated at position 2 including Pro, Ile, Lys and Glu. Similarly a collection of non-canonical side chains including those with both positive and negative charges were tolerated at position 9 including Glu, Lys, Asn, Gly and Ala.

A

Peptide Panel 2

Peptide ID	Sequence	Aas	Organism	Protein	Ref.
AE9	APLVPTGSE	9	*Simian immunodeficiency virus*	Gag_{65-73}	Hansen, S,G. et al (2016)
SN9	SENLKSLYN			Gag_{72-80}	
EA9	EKVKHTEEA			Gag_{92-100}	
KS9	KTSRPTAPS			$Gag_{120-128}$	
GI9	GGNYPVQQI			$Gag_{132-140}$	
QE9	QIIRDIINE			$Gag_{201-209}$	
VI9	VGNIYRRWI			$Gag_{259-267}$	
$RL9_{supertope}$	RMYNPTNIL			$Gag_{276-284}$	
KG9	KPIKCWNCG			$Gag_{389-397}$	
LQ9	LGKQQREKQ			$Gag_{476-484}$	
$EK9_{supertope}$	EKQRESREK			$Gag_{482-490}$	

B

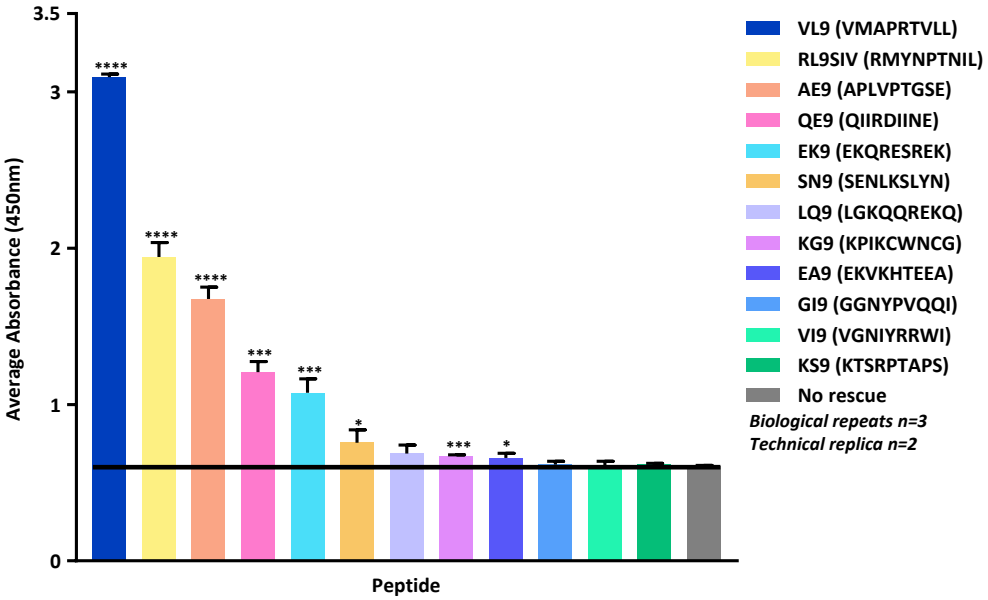


Figure 17a: Screening of peptide panel 2 in the optimised peptide exchange-based binding assay

A. Table detailing peptides in panel 2. Peptide ID, sequence, length and origin are reported with corresponding reference. **B.** Average absorbance readings are denoted on the Y-axis for the RhCMV 68-1 vaccine-identified SIV-derived 11 optimal epitopes tested in the optimised peptide exchange-based binding assay.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background (No rescue) in grey. Error bars depict standard error of the mean. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

The SIV-derived supertope RL9SIV (RMYNPTNIL) was aligned to its HIV Clade B-derived counterpart sequence, RMYSPTSIL (peptide ID: RL9HIV). The physiological relevance of the aligned HIV peptide was validated by the identification of RL9HIV-specific Mamu-E-restricted $CD8^+$ T cell responses in rhesus macaques vaccinated with RhCMV 68-1 vectors recombinant for HIV Gag (Walters et al., 2018). Relative HLA-E binding affinity of the SIV- and HIV-derived counterpart peptides was assessed in the peptide exchange assay. Although both RL9HIV and RL9SIV exhibited the greatest degree of statistically significant binding to HLA-E, the average absorbance signal for RL9HIV was consistently superior to that generated by RL9SIV (Figure 17b).

A

Peptide ID	Sequence	Aas	Organism	Protein	Ref.
$RL9SIV_{supertope}$	RMYNPTNIL	9	*Simian immunodeficiency virus*	$Gag_{276-284}$	Hansen, S,G. et al (2016)
RL9HIV	RMYSPTSIL		*Human immunodeficiency virus*	$Gag_{275-283}$	Walters, L,C. et al (2018)

B

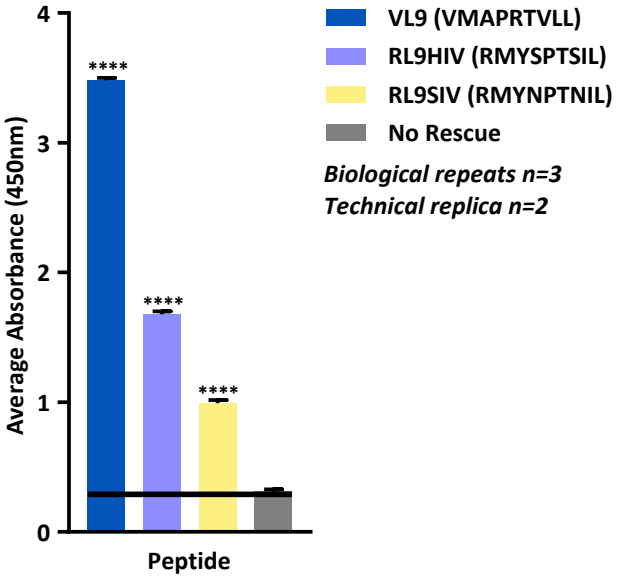


Figure 17b: RL9HIV versus RL9SIV peptide binding to HLA-E

A. Table detailing RL9SIV (RMYNPTNIL) and RL9HIV (RMYSPTSIL) peptides. Peptide ID, sequence, length and origin are reported with corresponding reference. **B.** Average absorbance readings are denoted on the Y-axis for RhCMV 68-1 vaccine-identified RL9SIV and RL9HIV peptides tested in the optimised peptide exchange-based binding assay.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background (No rescue) in grey. Error bars depict standard error of the mean. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Peptide panel 3

RhCMV 68-1 vaccine-identified HIV Gag-derived optimal 15mers and overlapping 9mers

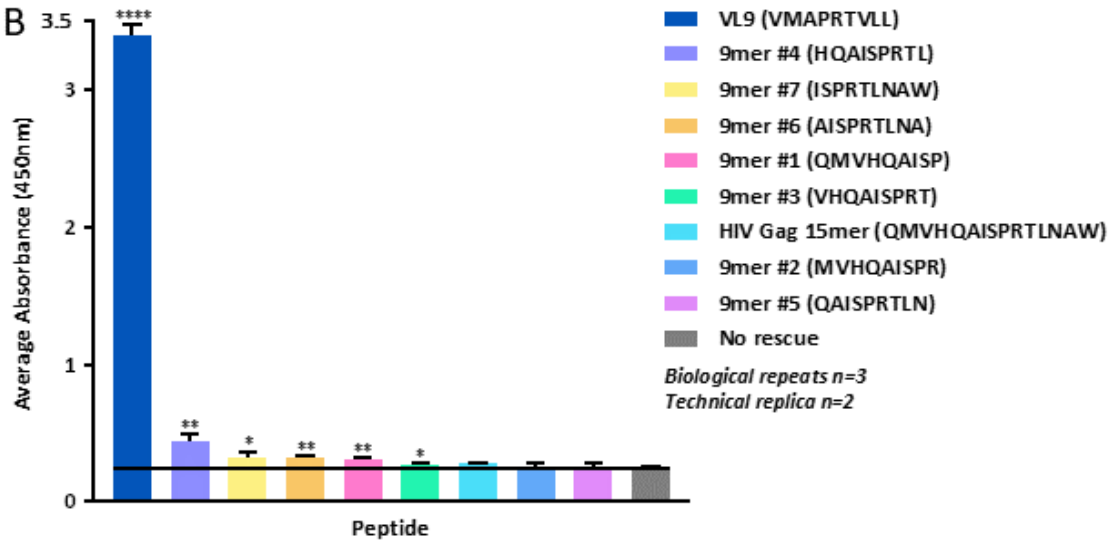
Optimal HIV Gag-derived 15mers from rhesus macaques vaccinated with RhCMV 68-1 vectors recombinant for HIV were identified in Mamu-E-restricted $CD8^+$ T cell stimulation assays by our collaborators in the Picker Lab (Louis Picker, unpublished data). We synthesised overlapping 9mers for each 15mer and tested these in the optimised peptide-exchange-based binding assay. As expected, due to suboptimal length for the closed MHC class I binding groove, no 15mer demonstrated statistically significant binding to HLA-E. However, the HIV Gag-derived 9mers also exhibited very low binding to HLA-E with many average absorbance readings only marginally above that of the no rescue background signal (Figure 18).

A

Peptide Panel 3

Peptide ID	Sequence	Aas	Organism	Protein	Ref
HIV Gag 15mer #1	QMVHQAISPRTLNAW	15			
HIV Gag 9mer #1.1	QMVHQAISP	9			
HIV Gag 9mer #1.2	MVHQAISPR	9			
HIV Gag 9mer #1.3	VHQAISPRT	9			
HIV Gag 9mer #1.4	HQAISPRTL	9			
HIV Gag 9mer #1.5	QAISPRTLN	9			
HIV Gag 9mer #1.6	AISPRTLNA	9			
HIV Gag 9mer #1.7	ISPRTLNAW	9			
HIV Gag 15mer #2	NTMLNTVGGHQAAMQ	15			
HIV Gag 9mer #2.1	NTMLNTVGG	9			
HIV Gag 9mer #2.2	TMLNTVGGH	9			
HIV Gag 9mer #2.3	MLNTVGGHQ	9			
HIV Gag 9mer #2.4	LNTVGGHQA	9			
HIV Gag 9mer #2.5	NTVGGHQAA	9			
HIV Gag 9mer #2.6	TVGGHQAAM	9			
HIV Gag 9mer #2.7	VGGHQAAMQ	9	HIV	Gag	Louis Picker, unpublished data
HIV Gag 15mer #3	STLQEQIGWMTNNPP	15			
HIV Gag 9mer #3.1	STLQEQIGW	9			
HIV Gag 9mer #3.2	TLQEQIGWM	9			
HIV Gag 9mer #3.3	LQEQIGWMT	9			
HIV Gag 9mer #3.4	QEQIGWMTN	9			
HIV Gag 9mer #3.5	EQIGWMTNN	9			
HIV Gag 9mer #3.6	QIGWMTNNP	9			
HIV Gag 9mer #3.7	IGWMTNNPP	9			
HIV Gag 15mer #4	QKQEPIDKELYPLAS	15			
HIV Gag 9mer #4.1	QKQEPIDKE	9			
HIV Gag 9mer #4.2	KQEPIDKEL	9			
HIV Gag 9mer #4.3	QEPIDKELY	9			
HIV Gag 9mer #4.4	EPIDKELYP	9			
HIV Gag 9mer #4.5	PIDKELYPL	9			
HIV Gag 9mer #4.6	IDKELYPLA	9			
HIV Gag 9mer #4.7	DKELYPLAS	9			

B



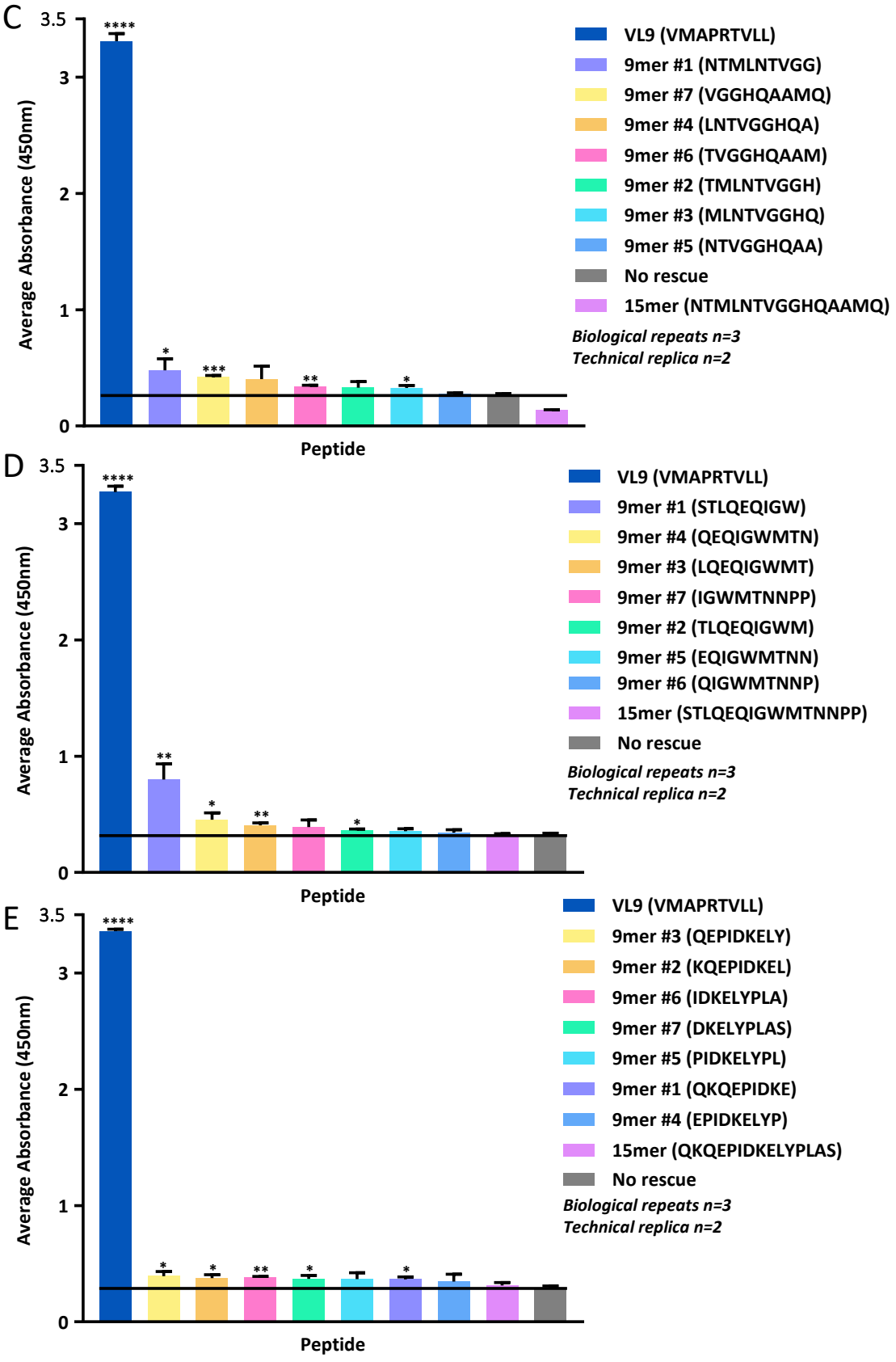

C
Average Absorbance (450nm)
3.5
3
2
1
0
****
*
***
**
*
Peptide
VL9 (VMAPRTVLL)
9mer #1 (NTMLNTVGG)
9mer #7 (VGGHQAAMQ)
9mer #4 (LNTVGGHQA)
9mer #6 (TVGGHQAAM)
9mer #2 (TMLNTVGGH)
9mer #3 (MLNTVGGHQ)
9mer #5 (NTVGGHQAA)
No rescue
15mer (NTMLNTVGGHQAAMQ)
Biological repeats n=3
Technical replica n=2
D
Average Absorbance (450nm)
3.5
3
2
1
0
****
**
*
**
*
Peptide
VL9 (VMAPRTVLL)
9mer #1 (STLQEQIGW)
9mer #4 (QEQIGWMTN)
9mer #3 (LQEQIGWMT)
9mer #7 (IGWMTNNPP)
9mer #2 (TLQEQIGWM)
9mer #5 (EQIGWMTNN)
9mer #6 (QIGWMTNNP)
15mer (STLQEQIGWMTNNPP)
No rescue
Biological repeats n=3
Technical replica n=2
E
Average Absorbance (450nm)
3.5
3
2
1
0
****
*
*
**
*
*
Peptide
VL9 (VMAPRTVLL)
9mer #3 (QEPIDKELY)
9mer #2 (KQEPIDKEL)
9mer #6 (IDKELYPLA)
9mer #7 (DKELYPLAS)
9mer #5 (PIDKELYPL)
9mer #1 (QKQEPIDKE)
9mer #4 (EPIDKELYP)
15mer (QKQEPIDKELYPLAS)
No rescue
Biological repeats n=3
Technical replica n=2

Figure 18: Screening of peptide panel 3 in the optimised peptide exchange-based binding assay

A. Table detailing peptides in panel 3. Peptide ID, sequence, length and origin are reported with the corresponding reference. **B-E.** Average absorbance readings are denoted on the Y-axis for RhCMV 68-1 vaccine-identified HIV Gag-derived peptides tested in the optimised peptide exchange-based binding assay.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background (No rescue) in grey. Error bars depict standard error of the mean. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Peptide panel 4

HIV Gag and Pol-derived peptides computationally predicted to bind HLA-E by Bette Korber's Group

A panel of HIV-derived peptides which were computationally-predicted to bind HLA-E were screened in the optimised binding assay (Bette Korber, unpublished data) (Figure 19, A). Although two peptides from panel 4 demonstrated significant binding, all peptides exhibited very low average absorbance signals approaching that of the negative control peptide-free background (Figure 19, B). Considering all such peptides contained the canonical position 2 Met anchor residue, the weak binding signals support the hypothesis that other key determinants play a role in dictating strength of binding to HLA-E.

A

Peptide Panel 4

Peptide ID	Sequence	Aas	Organism	Protein	Ref
MosG1	AMQMLKDTI	9	HIV	Gag	Bette Korber, unpublished
MosG2	AMQMLKETI				
MosG4	EMMTACQGV				
MosP1	SMNKELKKI			Pol	
MosP2	AMAGADCVA				
MosP3	PMGWKGSPA				
MosP4	SMNNELKKI				
MosP5	AMAGDDCVA				

B

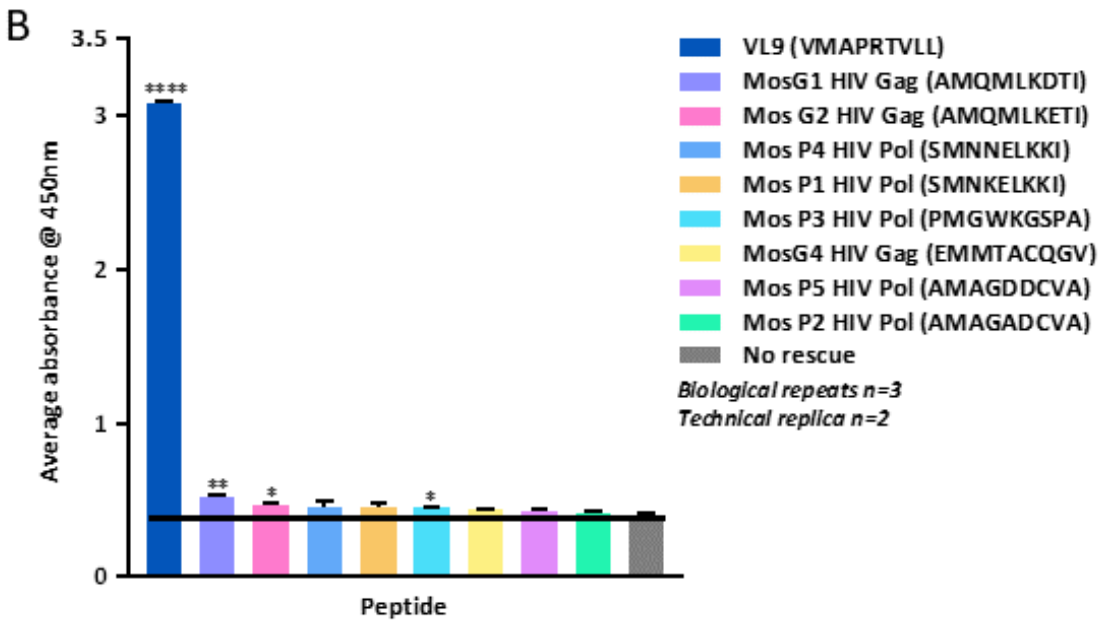


Figure 19: Screening of peptide panel 4 in the optimised peptide exchange-based binding assay

A. Table detailing peptides in panel 4. Peptide ID, sequence, length and origin are reported with corresponding reference. **B.** Average absorbance readings are denoted on the Y-axis for HIV Gag and Pol derived peptides computationally predicted to bind HLA-E by Bette Korber's group. Peptides are specified on the X-axis.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background (No rescue) in grey. Error bars depict standard error of the mean. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Peptide panel 5

HIV-derived NetMHC-predicted HLA-E binding peptides

A panel was constructed of HIV-derived peptides that were predicted in-house to bind HLA-E using the NetMHC 4.0 web-based server (Nielsen et al., 2003),(Andreatta and Nielsen, 2016) (Figure 20, A). The sequences in this panel were thus biased by computational predications based on previously established HLA-E binding motifs. 13 of the predicted peptides had the canonical Leu at position 9 and 11 contained a hydrophobic residue at position 2. A range of relative binding affinity was observed, with 13 of the 17 peptides demonstrating significant binding to HLA-E (Figure 20, B). Interestingly, the sixth strongest binder contained a non-canonical, charged Glu at position 2 and bound to HLA-E with a higher relative affinity than another peptide in the panel which contained the canonical anchor residues at both positions 2 and 9 (TMGAASMTL, derived from HIV Env).

A

Peptide Panel 5

Peptide ID	Sequence	Aas	Organism	Protein	Ref
NetMHC-HIV-1	TMGAASMTL			Env	
NetMHC-HIV-2	HQPISARTL			Gag/pol	
NetMHC-HIV-3	ENPYNTPVF			Gag/pol	
NetMHC-HIV-4	RGPGRAFVT			Env	
NetMHC-HIV-5	REPHNEWTL			Vpr	
NetMHC-HIV-6	RAMASDFNL			Gag/pol	
NetMHC-HIV-7	VIWGKTPKF			Gag/pol	
NetMHC-HIV-8	DIYKRWIIL			Gag/pol	Andreatta M & Nielsen M. Bioinformatics (2016)
NetMHC-HIV-9	CAPAGFAIL	9	HIV	Env	
NetMHC-HIV-10	QMHEDIISL			Env	
NetMHC-HIV-11	KHQKEPPFL			Gag/pol	
NetMHC-HIV-12	ILVESPTIL			Rev	
NetMHC-HIV-13	KLAGRWPVK			Gag/pol	
NetMHC-HIV-14	RLVNGSLAL			Env	
NetMHC-HIV-15	ILGQLQPSL			Gag/pol	
NetMHC-HIV-16	AIEAQQHLL			Env	
NetMHC-HIV-17	KQPTSQPRG			Tat	

B

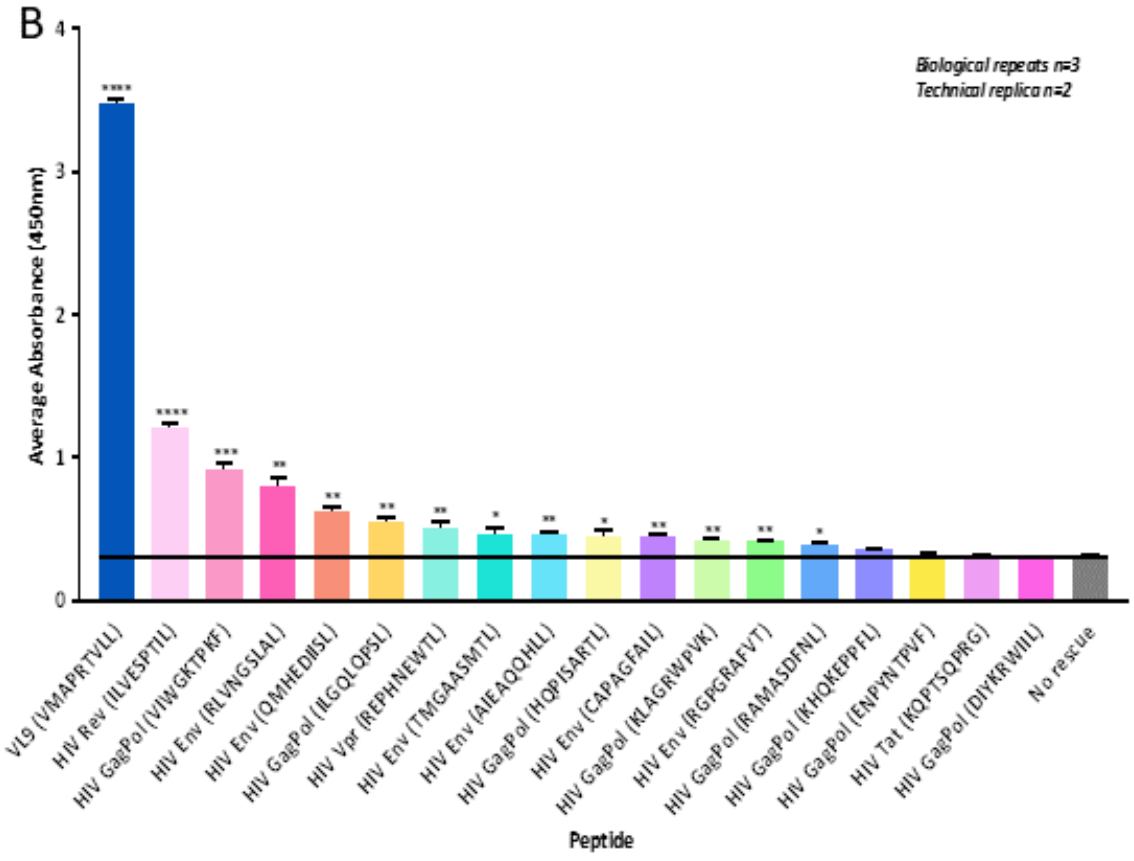


Figure 20: Screening of peptide panel 5 in the optimised peptide exchange-based binding assay

A. Table detailing peptides in panel 5. Peptide ID, sequence, length and origin are reported with the corresponding reference. **B.** Average absorbance readings are denoted on the Y-axis for HIV-derived peptides predicted to bind HLA-E by NetMHC. Peptides tested in the peptide exchange-based binding

assay are denoted on the X-axis.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background (No rescue) in grey. Error bars depict standard error of the mean. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Peptide panel 6

Mycobacterial peptides computationally predicted to bind HLA-E

A panel was constructed of previously published *Mycobacterial* peptides that were algorithmically predicted to bind HLA-E (Figure 21, A) - a number of which were subsequently shown to elicit HLA-E-restricted $CD8^+$ T cell responses in Tb-immune individuals (Joosten et al., 2010). Considering these peptides were initially identified *in silico* based on previous sequence motifs, they are largely canonical with either a hydrophobic Leu or Met at position 2 and a hydrophobic Leu or Ile at position 9. All seven peptides demonstrated significant binding to HLA-E in the optimised peptide exchange assay and exhibited a range of relative affinities (Figure 21, B). One peptide, Mtb44, exhibited strikingly strong binding, generating an average absorbance signal that rivalled that of the positive control VL9 peptide. Mtb44 was the first non-VL9 peptide screened in the binding assay that generated an average absorbance reading within the standard error of the mean of the positive VL9 control.

A

Peptide Panel 6

Peptide ID	Sequence	Aas	Organism	Protein	Ref
Mtb 14	RMAATAQVL	9	*Mycobacteria*	Rv2932	Joosten, S.A, et al. (2010)
Mtb 34	VMTTVLATL			Rv1734c	
Mtb 44	RLPAKAPLL			Rv1484	
Mtb 48	RLANLLPLI			Rv3189	
Mtb 54	FLLPRGLAI			Rv0056	
Mtb 55	VMATRRNVL			Rv1518	
Mtb 62	RMPPLGHEL			Rv2997	
Mtb 68	VLRPGGHFL			Rv1523	

B

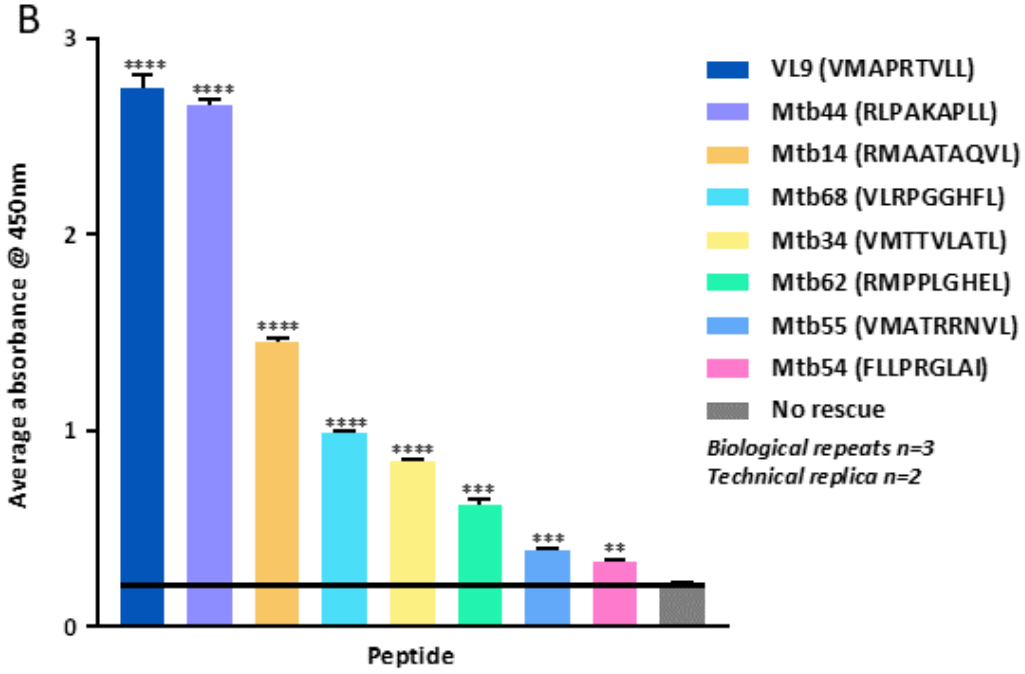


Figure 21: Screening of peptide panel 6 in the optimised peptide exchange-based binding assay

A. Table detailing peptides in panel 6. Peptide ID, sequence, length and origin are reported with the corresponding reference. **B.** Average absorbance readings from the optimised peptide binding assay are denoted on the Y-axis and the Mtb-derived peptides denoted on the X-axis.

*Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background (No rescue) in grey. Error bars depict standard error of the mean. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Peptide panel 7

HLA-E-restricted Mycobacterial peptides eluted from infected cells

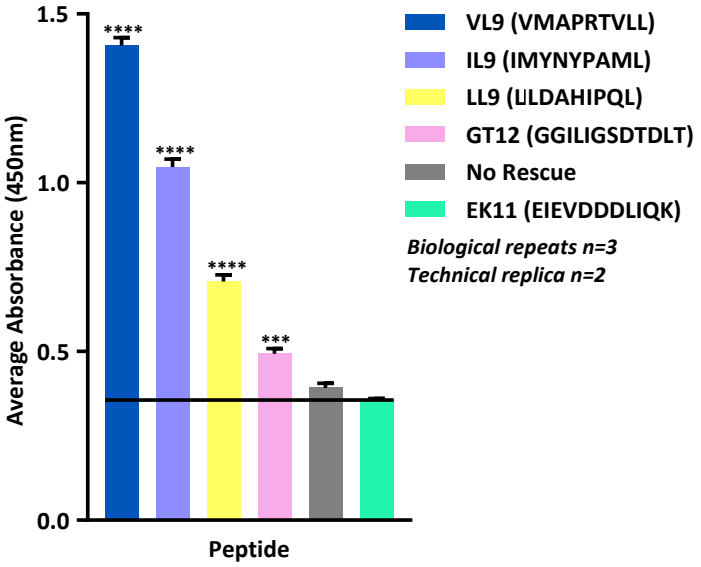
A second panel of *Mycobacterial* peptides was screened in the binding assay (Figure 22, A). Peptides in this panel not only deviated from the canonical HLA-E sequence motif but also from the common nonameric peptide length. This is likely since such peptides were eluted from Mtb-infected cells and were thus not subject to algorithmic biases for previously described HLA-E-restricted peptide sequence motifs (McMurtrey et al., 2017). Despite the closed nature of the MHC class I binding groove, which typically accommodates peptides of 8-10 amino acids, an Mtb-derived peptide comprising 12 amino acids, GT12, demonstrated significant binding to HLA-E in the peptide binding assay (Figure 22, B). To investigate whether this signal might be generated by fragmented nonamers within the peptide stock, we synthesised 2 internal nonameric sequences from within the GT12 peptide, one of which contained hydrophobic primary anchor residues (ILIGSDTDL). The GT12 9mer #2 peptide generated a comparable signal to the originally eluted 12mer peptide. Additionally, two nonameric Mtb peptides, IL9 and LL9, with hydrophobic residues at both primary anchor positions demonstrated binding to HLA-E with the greatest degree of statistical significance.

A

Peptide Panel 7

Peptide ID	Sequence	Aas	Organism	Protein	Ref
IL9	IMYNYPAML	9	Mycobacteria	$EsxH_{4-12}$	McMurtrey, C. et al. (2017)
LL9	LLDAHIPQL			$EsxG_{3-11}$	
GT12	GGILIGSDTDLT	12			
GT12 9mer #1	ILIGSDTDL	9		$Lpql_{89-100}$	
GT12 9mer #2	LIGSDTDLT				
EK11	EIEVDDDLIQK	11		$Rv0634A_{19-29}$	

B



C

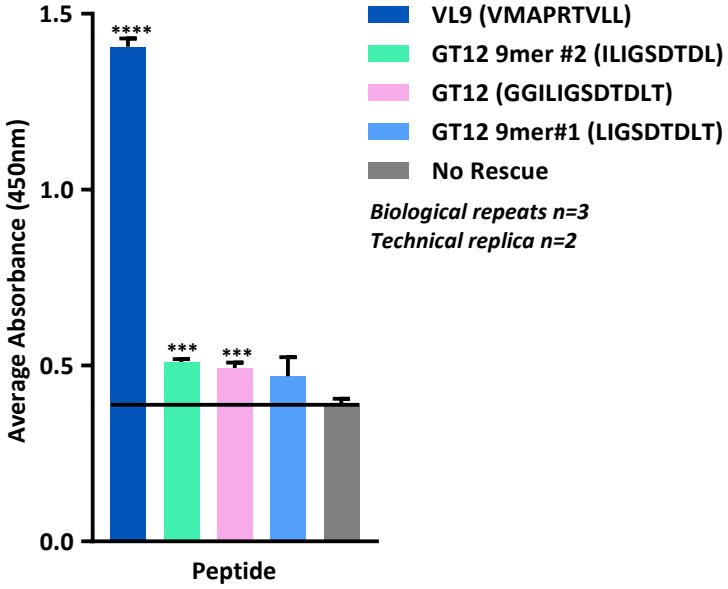


Figure 22: Screening of peptide panel 7 in the optimised peptide exchange-based binding assay

A. Table detailing peptides in panel 7. Peptide ID, sequence, length and origin are reported with the corresponding reference. **B.** Average absorbance readings from the optimised peptide binding assay are

denoted on the Y-axis and Mtb-derived peptides eluted from infected cells are denoted on the X-axis. *Positive control peptide VL9 (HLA-B7-derived leader sequence peptide VMAPRTVLL) shown in dark blue with peptide free background (No rescue) in grey. Error bars depict standard error of the mean. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

HLA-E sequence motifs based on screening of peptide panels in the peptide-exchange assay

Sequences of the 58 nonameric non-VL9 peptides that demonstrated statistically significant binding to HLA-E in optimised peptide exchange-based assays were inputted into the web-based Seq2Logo sequence motif generator (Thomsen and Nielsen, 2012). Longer HLA-E binding peptides such as the Hepatitis C virus Core protein-derived decamer and the Mtb-derived 12mer, GT12, were excluded according to Seq2Logo input requirements. Unsurprisingly, primary anchor positions 2 and 9 imposed the greatest degree of specificity, as indicated by the size and narrow-focus of residue stacks in the Seq2Logo motif (Figure 23, A). Further, the most frequently occurring residue at position 9, as indicated by font size, is consistent with canonical position 9 Leu. However, Leu was also the most frequently observed residue at position 2 of the Seq2Logo in place of canonical Met.

It is necessary to note such web-based motif generators are subject to bias based on the nature of the input sequences. It is possible canonical or biochemically-related primary anchor residues were overrepresented in the binding assay screens since certain panels comprise peptides that were originally identified *in silico* via algorithms based on established HLA-E binding motifs. To eliminate this potential bias, Seq2Logo motifs were exclusively generated for HLA-E binding peptides identified in a physiological capacity, such as from RhCMV 68-1

vaccination or elution from infected cells. Thus, peptide panels 4, 5 and 6 were omitted from further analysis.

Although Leu remained the most frequently observed residue for both primary anchor positions, position 2 Met was relegated from second to the fifth most frequent residue among physiologically-identified peptides, below that of charged Lys and Glu (Figure 23, B). Further, residue stack length increased for multiple peptide positions, such as the non-anchor position 4, where the stack height approached that of the primary anchor position 2. This indicates that sequence preferences at non-anchor positions feature in physiologically-identified HLA-E binding peptides. Further stratification was carried out according to peptide binding assay results, with a separate Seq2Logo being generated for physiologically-identified peptides which exhibited the greatest degree of statistical significance in the peptide exchange-based binding assay, as indicated by p values below or equal to 0.0001. Although Met reassumed its ranking as the second most observed residue at position 2, the stack heights for positions 4 and 7 remained extended (Figure 23, C). In summary, although such results are likely subject to biases relating to enrichment of certain residues in the input sequences, it appears that physiologically-identified HLA-E binding non-VL9 peptides do not tightly adhere to the previously established canonical binding motif for HLA-E, with a large number of permitted residues at both primary anchor positions.

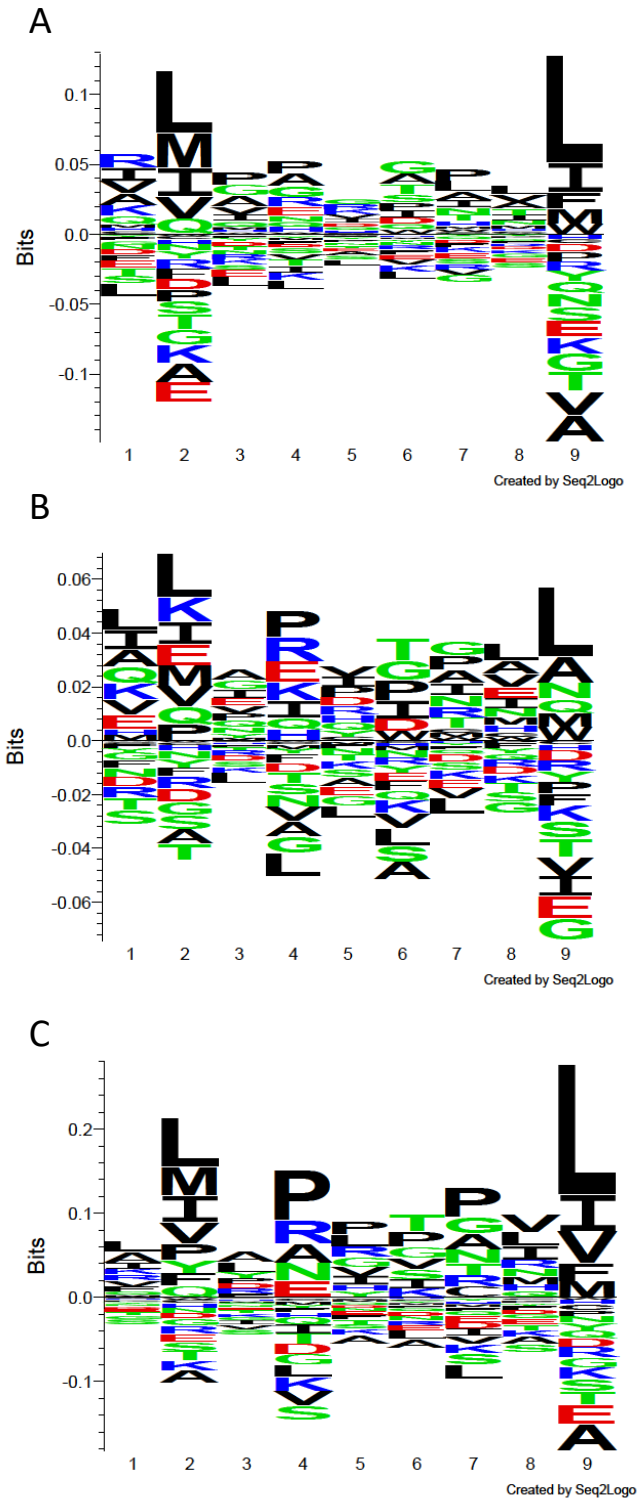


Figure 23: Seq2Logo-generated HLA-E peptide binding motifs

A. A Seq2Logo HLA-E peptide sequence motif based on statistically significant nonamers identified to bind HLA-E in optimised peptide exchange-based binding assay screens. **B.** Seq2Logo peptide sequence motif for physiologically-identified HLA-E binding peptides that generated statistically significant signals

in the optimised binding assay. Specifically, peptides from panels 4, 5 and 6 were excluded from this analysis. **C.** Seq2Logo peptide sequence motif for physiologically-identified HLA-E binding peptides that demonstrated the greatest degree of statistical significance in the optimised binding assay.
The Y axis denotes information in bits with peptide position plotted on the X-axis. Letters represent amino acids with font size indicating residue frequency. The length of residue stacks indicates the degree of conservation among inputted sequences at that position with small stack lengths reflecting increased residue variability.

Discussion:

In summary, the optimised peptide exchange-based binding assay has reliably demonstrated an HLA-E binding capacity for a range of pathogen- and self-derived peptides. Such results support an HLA-E peptide binding repertoire that is not limited to MHC class I-derived leader peptides. Despite this, our results are consistent with the original characterisation of the HLA-E peptidome, in which a strong preference for VL9 peptides was established (O'Callaghan et al., 1998) (Braud et al., 1997). Of the 91 screened peptides previously predicted to bind HLA-E, only one (Mtb44) demonstrated a relative binding affinity comparable to that of the positive VL9 peptide control. This is consistent with higher affinity VL9 peptides dominating the HLA-E ligandome in the steady state (McMurtrey et al., 2017). However, when MHC class I leader peptides are in short supply, following viral or intracellular bacterial immune evasion-targeting or cellular transformation, a diversification of the HLA-E peptide binding repertoire can potentially ensue, permitting broad sampling of lower affinity peptide ligands.

Peptide binding assay optimisation

Unusually high signals were observed for the peptide-free HLA-E heavy chain-β2M negative control in the micro-refold approach. Previously, minimal background signals were observed during the original application of the binding assay with HLA-A*0201 and HLA-A*1101 (Sylvester-Hvid et al., 2002). Further, reductions in peptide-free HLA-E-β2M background

signals did not ensue following the removal of potential contaminating peptide fragments via the addition of pre-purified HLA-E and β2M protein preparations to the micro-refold. This indicates that perhaps HLA-E possesses a superior affinity for β2M in the absence of a high affinity peptide ligand, relative to its classical MHC class I counterpart molecules. Elevated background signals are consistent with blue native gel analyses, conducted by Professor Geraldine Gillespie NDM Oxford University, in which peptide-free HLA-E forms a distinctive, diffuse band in which both the HLA-E heavy chain and β2M are present (Walters et al., 2018).

The identification of peptide exchange independent of photo-illumination by the dedicated UV cabinet during optimisation of the binding affinity assay raises the question of whether the photo-labile VL9-based peptide (VMAPJTLVL) is cleaved by lower wavelengths in the visible spectrum or by background UV radiation present in the lab. Alternatively, it is possible that full length, nonameric UV-labile peptide simply dissociates from the HLA-E binding groove overtime in the absence of rescue peptide, irrespective of exposure to electromagnetic radiation. Professor Geraldine Gillespie has previously shown, via blue native gel analysis, that a molar excess of low affinity peptide can facilitate peptide exchange with a pre-bound higher affinity peptide in the HLA-E binding groove (Walters et al., 2018). This demonstrates that peptide exchange from a higher to lower affinity peptide ligand is possible for HLA-E in the absence of a photo-cleavable peptide or UV irradiation when the lower affinity peptide is in molar excess.

Dissociation of intact UV-labile HLA-A2.1-restricted peptide was not originally observed following incubation in the absence of UV irradiation (Toebes et al., 2006), presumably due to the higher absolute affinity of HLA-A2.1-restricted UV-labile influenza A $M1_{58-66}$ epitopes relative to HLA-E-restricted UV-labile VL9. Thus, photo-cleavable VL9 peptides might be more susceptible to photo-unassisted dissociation from HLA-E. Additionally, the introduction of an unnatural beta amino acid at position 5 of the UV-labile VL9 peptide results in conformational

changes to the peptide main chain which can in turn impact binding affinity (Toebes et al., 2006). Therefore the modified VL9 peptide likely exhibits an increased propensity for dissociation relative to wild-type VL9. To definitively determine whether the peptide exchange observed in the binding assay derives from photo-cleavage or intact dissociation of the modified VL9, a VL9 peptide containing a UV-resistant beta amino acid building block lacking the nitrophenyl moiety could be synthesised. Since this peptide would be resistant to photo-cleavage, peptide exchange would presumably result from dissociation of the intact nonamer from the peptide binding groove.

Optimisation experiments determined that pH 8 optimally supported low affinity peptide binding to HLA-E in the peptide-exchange binding assay. By contrast, it was previously determined that a more acidic pH between 6 and 7 was favoured for peptide loading by six of the seven MHC class I alleles tested (Pedersen et al., 2001). This preference for a lower pH during peptide loading did, however, wane with increased peptide concentration, with optimal loading occurring across a wider pH range at higher peptide concentrations. Typically, classical MHC class I protein refolding is conducted in the presence of peptide at concentrations around 10µM (Garboczi et al., 1992). Due to the particularly low HLA-E binding affinities of many non-canonical pathogen-derived peptides, test peptides are supplied at higher concentrations of 100µM in the peptide exchange-based binding assay. Thus, it is possible these higher peptide concentrations mask an underlying preference for a lower pH during peptide loading. Further, given that pH can impact the chemical nature of peptide side chains - such as an uncharged His side chain which can become doubly protonated when the pH decreases from a neutral pH 7 to a more acidic pH 6 - it is probable that optimal loading into the HLA-E binding groove has a peptide sequence-dependent component, since differential pH-dependent side chain properties could in turn influence binding affinity to MHC class I molecules. Therefore, a more comprehensive elucidation of the optimal peptide loading pH for HLA-E would require sampling of a much larger panel of HLA-E-binding peptides.

Rhesus versus human MHC-E peptidomes

Deciphering the degree of functional similarity between the human and rhesus MHC-E homologues is critical in determining the potential efficacy of an HLA-E-targeted universal HIV vaccine in humans. Given the broad, motif-less array of non-canonical SIV- and HIV-derived epitopes presented by Mamu-E following RhCMV 68-1-SIV vaccination, it would be necessary for this binding breadth to extend to the human homologue, HLA-E. The breadth of SIV epitopes recognised by Mamu-E-restricted $CD8^+$ T cells following RhCMV 68-1 vaccination was superior to the recognition breadth of classically-restricted SIV-specific $CD8^+$ T cells elicited in more conventional vaccine approaches (Hansen et al., 2016). Thus, it is important to determine whether HLA-E is also capable of presenting this broad and diverse set of peptides as superior $CD8^+$ T cell recognition breadth likely constitutes a key contributing feature to the protective efficacy of RhCMV 68-1 vaccines, potentially suppressing the emergence of escape mutants early in HIV infection (van Deutekom et al., 2013).

Key differences have previously been identified regarding the human and rhesus MHC-E homologues. Not only are a greater number of Mamu-E alleles present at the population level, but additionally a greater number of alleles are expressed within individual rhesus macaques, relative to human HLA-E (Wu et al., 2018). Further, although the two human alleles HLA-E*0101 and HLA-E*0103 differ by a single residue at position 107, a greater degree of sequence diversity is observed between the more numerous Mamu-E alleles.

In an attempt to fully elucidate the extent of overlap in the HLA-E and Mamu-E peptide binding repertoires, the peptide binding assay was adapted to detect the rhesus homologue. Because the anti-HLA-E antibody 3D12 does not cross react with the rhesus homologue, a Mamu-E-specific antibody, 4D12, was trialled in the peptide binding assay in place of the 3D12 capture antibody. However, the 4D12-based assay generated a positive VL9 peptide control signal only marginally above that of the peptide-free background, indicative of a weak capacity

of 4D12 to discriminate between peptide-loaded and peptide-free Mamu-E forms. Due to the limited availability of alternative Mamu-E-specific antibodies, the rhesus-specific peptide binding assay was therefore not further pursued.

As an alternative, a comparative analysis of pocket lining residues was conducted for both HLA-E alleles in addition to the 22 published Mamu-E alleles (Figure 26). Despite the considerable degree of sequence divergence between the human and rhesus homologues, high sequence conservation was observed for the heavy chain-derived residues forming the B, D, C, E and F pockets of the MHC-E binding groove. Precisely, 19 of the 20 pocket-lining residues were conserved between HLA-E and 21 of the 22 Mamu-E alleles. The single sequence difference comprised an Ile to Thr substitution in Mamu-E at residue 73 of the C pocket. The Mamu-E*02:20 allele deviated from the HLA-E sequence at three additional positions, including a Ser to Phe substitution at residue 24 of the B pocket, a His to Tyr substitution at residue 99 of the D pocket and a Phe to Leu substitution at residue 74 of the C pocket.

It is well established that the specific combination of residues forming an anchor-receptive pocket in an MHC class I binding groove determines anchor residue specificity (Madden, 1995). Thus, given the stringent sequence conservation between pocket-forming residues of human and rhesus MHC-E, it is likely these homologues share a largely overlapping peptide binding repertoire. This sequence analysis is consistent with the functional data obtained by Wu, H,L. et al, whereby the human and rhesus MHC-E homologues were equally capable of presenting identical peptides and stimulating Mamu-E-restricted $CD8^+$ T cells from RhCMV 68-1 vaccinated macaques (Wu et al., 2018). The inability of Mamu-E-restricted $CD8^+$ T cells to discriminate between Mamu-E and HLA-E further adds to the functional equivalence of the two homologues, supporting the possibility of developing a broadly-targeted HLA-E-restricted $CD8^+$ T cell vaccine against HIV in humans.

	B-pocket				D-pocket			C-pocket				E-pocket				F-pocket				
	9	24	45	67	97	99	159	70	73	74	116	133	147	152	156	81	95	118	123	124
HLA-E	**H**	**S**	**M**	**A**	**W**	**H**	**Y**	**T**	**I**	**F**	**F**	**W**	**S**	**E**	**Q**	**L**	**L**	**Y**	**Y**	**L**
Mamu-E*02:01:02	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:02	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:03	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:04	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:05	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:07	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:09	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:10	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:11	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*0212:01	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*0212:02	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:13	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:14	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:15	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:16	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:17	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:18	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:19	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:20	H	F	M	A	W	Y	Y	T	T	L	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:21	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:22	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L
Mamu-E*02:23	H	S	M	A	W	H	Y	T	T	F	F	W	S	E	Q	L	L	Y	Y	L

Figure 26: Sequence analysis of pocket-lining residues in human and rhesus MHC-E

The amino acid residues present in the B, D, C, E and F pockets of the MHC-E peptide binding groove are specified for the homologous human and rhesus alleles. Mamu-E allele sequences were acquired from the IPD-MHC Database and aligned using the ClustalW2 web server. Residues which contribute to multiple pockets are listed only once.

Peptide binding assay screening

The peptide panel comprising optimal SIV-derived RhCMV 68-1 vaccine-identified epitopes yielded statistically significant HLA-E binding for seven of the eleven peptides, supporting diverse peptide sampling properties for HLA-E in addition to Mamu-E. Additionally, all eleven peptides yielded average absorbance signals above the no-rescue background, despite abundant biochemically dissimilar primary anchor residues. For example, the fourth strongest binding peptide, EK9, possesses a positively charged Lys at both primary anchor positions in place of the hydrophobic position 2 Met and 9 Leu. The hydrophobic HLA-E B pocket is not considered to be compatible with a positively charged Lys anchor side chain (Personal communication, Pamela Bjorkman). However, the presence of non-canonical anchor residues in HLA-E binding peptides does not necessarily indicate broadly receptive pockets in the HLA-E binding groove. Structural evidence would be required to elucidate at the atomic level whether non-canonical primary anchor side chains occupy their corresponding pockets. It is possible non-canonical peptides adopt alternative binding conformations within the HLA-E groove disrupting the classical pattern of pocket occupancy by specific anchor residues.

Further, ELISA screening data of SIV-derived epitopes concur with previous, computationally-derived negative free binding energies identified for all eleven optimal SIV epitopes, indicative of an HLA-E binding capacity (Hansen et al., 2016). Not only was a similar hierarchy of relative binding affinity observed for the eleven epitopes but also RL9SIV exhibited the strongest HLA-E binding in both the ELISA and computational predictions. Optimal SIV-derived epitopes, in addition to statistically significant HLA-E binding peptides identified in alterative physiological settings, not only contain non-canonical primary and secondary anchor residues but also lack a common dominant motif. However, a diverse motif-less HLA-E peptidome would prove particularly beneficial in an HIV vaccine context since the breadth of $CD8^+$ T cell-targeting early

in infection has been inversely correlated with the likelihood of the emergence of de novo escape mutants (van Deutekom et al., 2013).

Although the majority of screened peptides were nonameric, a Hepatitis C virus-derived decamer and an Mtb-derived 12mer, GT12, also exhibited statistically significant binding to HLA-E. It is possible internally-derived nonamers present in the 80% pure peptide stocks are responsible for these binding signals. Alternatively, this might reflect the accommodation of extended peptides by HLA-E. Unlike for MHC class II molecules, protrusion of peptides beyond the MHC class I binding groove is thought to be prohibited by narrowing of the groove ends to around 5 Å, in addition to blocking by the conserved bulky side chains of Tyr-84 and Trp-167 (Madden, 1995). However, despite these structural constraints multiple structures of MHC class I molecules have been solved in complex with elongated peptides exceeding the optimal range of 8-10 amino acids - such as the HLA-B7-restricted NY-ESO-1 tumour antigen which comprises 13 amino acids (Ebert et al., 2009). Such cases commonly exhibit increased bulking of the central peptide main chain yielding distinct solvent exposure profiles and in certain cases, including an EBV BZLF1-derived HLA-B*3501-restricted 11mer and HLA-B*3508-restricted 13mer, immunodominant $CD8^+$ T cell responses (Tynan et al., 2005)/(Green et al., 2004).

Results chapter 2

Structural analysis of pathogen-derived peptide binding to HLA-E

Introduction

Peptide binding assay data discussed in chapter 1 indicate that a broad array of pathogen-derived epitopes, a number of which contain non-canonical primary and secondary anchor residues, can bind HLA-E. However, these results do not provide insight into how such peptides are accommodated by the peptide binding groove. Importantly, it is unclear whether biochemically disparate, non-canonical side chains can be accommodated by corresponding pockets or whether stringent pocket specificity drives alternative binding motifs, vacant pockets and novel solvent exposure profiles. Moreover, even presupposing tolerance for certain non-canonical side chains, it is unclear whether structural rearrangements of HLA-E heavy chain pocket-lining side chains would ensue. Although structural rearrangement of pocket-forming side chains is well documented for classical MHC class I structures in the face of anchor residue substitution (Madden et al., 1993), only minor rotamer adjustments have previously been reported for HLA-E (Strong et al., 2003). Further, such slight adjustments, such as Ser-24 in the HLA-E B pocket arising from a position 2 Met to Thr substitution in the VL9-bound peptide, could simply be due to varying interpretations of mid-range resolution electron density maps (Strong et al., 2003).

Previously, HLA-E-β2M heterodimers have exclusively been crystallised in complex with highly conserved VL9 peptide variants derived from the leader sequences of a range of MHC class I alleles including HLA-B27 (VTAPRTLLL) (Strong et al., 2003), HLA-G*01 (VMAPRTLFL) (Hoare et al., 2008), HLA-Cw*07 (VMAPRALLL) (Hoare et al., 2008) and HLA-B7/8/14/39/42/48 (VMAPRTVLL) (O'Callaghan et al., 1998), (Strong et al., 2003). Additionally, a structure of HLA-E*0101 in complex with the HCMV UL40 protein-derived VL9 peptide mimic (VMGPRTLIL) has

been deposited in the protein Data Bank (PDB) and assigned the ID 3AM8**.** Similar to classical MHC class I-bound peptides, HLA-E-bound VL9 or HCMV UL40-derived mimic peptides adopt the prototypical kinked conformation which invariably runs parallel to the α1 helix in the peptide binding groove (Madden, 1995). Classically orientated VL9 peptides exhibit anchored amino and carboxy termini, with a more loosely constrained central region. A kink following the position 3 α-carbon projects the peptide main chain away from the floor of the binding groove yielding greater conformational flexibility and solvent exposure of central peptide residues.

Peptide binding motifs of classical MHC class I molecules typically involve the accommodation of two to three anchor residue side chains within respective groove pockets. However, particularly heavy binding pocket usage has previously been observed for VL9-bound HLA-E (O'Callaghan et al., 1998) (Strong et al., 2003) (Hoare et al., 2008). In previous VL9-bound structures of HLA-E, not only do the position 2 and 9 primary anchor side chains project into their respective B and F pockets, but also, positions 3, 6 and 7 constitute secondary anchor residues which project side chains into the D, C and E pockets, respectively. An array of specific, complex-stabilising contacts are formed between pocket-occupying anchor side chains and the MHC class I heavy chain. This imposes sequence restrictions at such anchor positions, reducing the overall magnitude of the potential peptide binding repertoire. In this way, anchor side chain-specific pocket-based interactions give rise to sequence binding motifs which impart peptide binding affinity in the balancing act with peptide binding breadth. Thus, highly promiscuous peptide binding achieved by classical MHC class I molecules is thought to be, in part, attributable to the limited number of anchor residues, thus relieving sequence restrictions at remaining positions along the peptide (Madden, 1995). It therefore seems plausible that the requirement to occupy a greater number of pockets, as observed in VL9-bound structures of HLA-E, would correlate with a less diverse peptidome.

Peptide sequence-restrictive structural features are consistent with the targeted functional role of HLA-E in presenting MHC class I-derived VL9 peptides to CD94/NKG2 class receptors to evade NK cell-mediated lysis (V. Braud et al., 1998) – a role which would presumably be more effectively-mediated in the absence of a broad peptidome. Thus, it is conceivable HLA-E evolved to accommodate a limited collection of peptides. Furthermore, analogous to VL9-regulated surface expression of HLA-E in humans (Veronique M. Braud et al., 1998), the murine homologue Qa-1, was also found to exhibit a strong binding preference for highly conserved murine MHC class I-derived signal peptides (AMAPRTLLL) (Kurepa et al., 1998). Thus, it was originally concluded that the HLA-E peptidome predominantly comprises MHC class I leader sequence VL9 peptides (O'Callaghan et al., 1998).

However, considering the emerging role of HLA-E as a $CD8^+$ T cell restriction element, primarily in Mtb infection and RhCMV-68-1 SIV/HIV vaccination, it is now important to understand how such pathogen-derived, non-canonical epitopes are accommodated by the HLA-E peptide binding groove. Elucidating the underlying structural basis of epitope binding to HLA-E could in turn provide insight regarding immunogenicity and recognition by unconventional, HLA-E-restricted $CD8^+$ T cells.

Crystal structures of Mtb-derived peptide-bound HLA-E*0103

Peptide affinity for MHC class I molecules and complex stability are key determinants of crystallographic outcome. Crystallisation typically requires a homogenous population of stable molecules which retain their folded structure for sufficient periods to permit super-saturation, nucleation and crystal growth. Instability may not only result in complex dissociation but also correlate with partial unfolding, increased flexibility and disorder, preventing formation of an ordered crystal lattice. Thus, there is currently no published structure of an unmodified, peptide-free MHC class I molecule and it is conceivable that crystallisation of HLA-E bound to

low affinity RhCMV 68-1 vaccine-identified SIV/HIV and Mtb-derived epitopes would be subject to similar limitations.

Additionally, low peptide affinity could also restrict protein yield obtained from macro-refolding and chromatographic purification. This would in turn limit the sampling of a wide range of buffer conditions in crystallisation screening reducing the probability of successfully identifying the precise conditions that facilitate nucleation and crystal growth. Furthermore, it was previously reported that diffracting crystals of the VL9 peptide-bound HLA-E structure, 1MHE, took 6 months to form, placing particularly rigorous demands on complex stability (O'Callaghan et al., 1998).

In order to maximise the likelihood of crystallisation, pathogen-derived epitopes which yielded the highest relative signals in the peptide-exchange-based binding assay were initially selected for structural studies. Specifically, the Mtb-derived epitope, Mtb44 (RLPAKAPLL) from panel 9, was the only peptide screened which rivalled binding signals of the positive control VL9 peptide. Crystallisation plates were also assembled for purified HLA-E-β2M complexes refolded with the second strongest Mtb-derived HLA-E-binding peptide from panel 9, Mtb14 (RMAATAQVL).

Mtb44 and Mtb14, derived from Rv1484 (Enoyl-[acyl-carrier-protein] reductase [NADH]) and Rv2932 (Malonyl CoA-acyl carrier protein transacylase), respectively within the Mtb genome, were originally predicted to bind HLA-E *in silico* (Joosten et al., 2010). Mtb14 was subsequently shown to elicit $CD8^+$ T cell responses in 4 purified protein derivative (PPD)-positive donors, but was not included in HLA-E restriction experiments performed in this study (Joosten et al., 2010). By contrast, both the antigenic nature and MHC-E restriction of Mtb44 was recently validated through the identification of Mamu-E restricted $CD8^+$ T cell responses in BCG-vaccinated rhesus macaques following peptide stimulation of peripheral blood mononuclear cells and VL9-peptide blocking (Walters et al., 2018).

Structural determination of HLA-E*0103 in complex with the Mtb44 peptide

*HLA-E*0103-Mtb44*

The extracellular domain of the HLA-E*0103 heavy chain comprising 274 residues, the β2M light chain comprising 100 residues and the nonameric Mtb44 peptide were refolded *in vitro* and purified material dispensed into sitting drop crystallisation sparse matrix screens. Crystallisation and refinement statistics are detailed in Figure 1. In summary, crystals grew in 2.4 M Ammonium sulphate, 0.1 M MES at pH 6, diffracted to 2.09 Å and packed in the P1 space group yielding unit cell dimensions a = 48.35 Å, b = 76.31 Å and c = 111.02 Å. The dataset was 89.1% complete overall, although this dropped to 47% in the highest resolution shell. However, given the $CC_{1/2}$ for the outer shell (0.668) exceeded the recommended lower threshold of 0.3, no reflections were excluded from subsequent analysis. 4 molecules were present in the asymmetric unit which had a solvent content of 46%, as predicted by the Matthew's coefficient. The structure was solved using coordinates from 1MHE - HLA-E*0101 in complex with the HLA-B8/7/14/39/42/48 leader peptide VMAPRTVLL (O'Callaghan et al., 1998). Molecule 1 from the asymmetric unit of 1MHE, with the hydrogen atoms and peptide coordinates removed, was used as a phasing model in molecular replacement. Refinement was subsequently carried out in Phenix yielding final R factors within the appropriate range given the resolution of the structure of 0.191 and 0.229 for R_{work} and R_{free}, respectively. Following geometric validation using MolProbity the coordinates were deposited in the worldwide RCSB protein data bank (PDB) and assigned the PDB ID, 6GH1. Density for a single Zinc ion was identified in the asymmetric unit and its presence later confirmed by fluorescence screening at Diamond Light Source (Performed by Dr. Karl Harlos of Oxford University). This was likely a contamination product from bacterial lysates during protein purification as Zinc was not present in subsequent refolding or crystallisation buffers.

	HLA-E*0103 in complex with peptide:	
Data collection	Mtb44	Mtb14
Crystallisation condition	2.4 M AS 0.1 M MES pH 6	3 M AS 0.1 M MES pH 6
Resolution Å	53.91-2.09 (2.13-2.09)	57.16-2.05 (2.124-2.05)
Space Group	P1	C2
Cell dimensions	a = 48.35 Å b = 76.31 Å c = 111.02 Å α = 99.26° β = 99.50° γ = 93.68°	a = 245.15 Å b = 47.93 Å c = 152.45 Å α = 90° β = 117.45° γ = 90°
Solvent content [%] (molecules per AU)	446(4)	44 (4)
Unique reflections	81176 (4505)	99611 (9754)
Completeness [%]	89.1 (47.0)	98.95 (96.24)
R_{merge} [I]	0.09 (0.67)	0.155 (0.828)
I/sigma	2.8 (1.1)	9.7 (1.8)
Multiplicity	1.7 (1.5)	12.7 (8.2)
$CC_{1/2}$	0.99 (0.67)	0.99 (0.60)
Refinement		
No. of non-hydrogen atoms	13388	13201
R_{factor} [%)]	0.191 (0.277)	0.2305
R_{free} [%]	0.229 (0.336)	0.2730
r.m.s.d. bonds [Å] §	0.002	0.008
r.m.s.d. angles [deg]	0.55	1.008
Ramachandran statistics		
Favoured [%]	98.21	98.72
Disallowed [%]	0	0

Figure 1: Crystallographic determination, refinement and validation of Mtb44/Mtb14-bound HLA-E*0103

Crystallographic data collection, structural determination, refinement and validation statistics for HLA-

E*0103 structures solved in complex with the Mtb-derived Mtb44 (RLPAKAPLL) and Mtb14 (RMAATAQVL) peptides.

AS: Ammonium sulphate. AU: asymmetric unit. § r.m.s.d.: root mean square deviation from ideal geometry. Statistics in parentheses comprise those deriving from the outer shell. R_{free}: R factor representing 5% of reflections omitted from refinement.

Overall structural architecture and alignment of Mtb44-bound HLA-E molecules

The Mtb44-bound HLA-E*0103 structure exhibited the classical MHC class I fold (Bjorkman, 1990), with a peptide binding groove guarded by 2 α-helices and an underlying β-sheet floor. Sufficient electron density for both secondary structure elements and flexible loops enabled modelling of every residue within the HLA-E heavy chain, invariant β2M light chain and peptide of all four molecules in the asymmetric unit.

Cα backbone alignment of the four Mtb44-HLA-E molecules revealed similarly positioned secondary structure elements of the peptide binding groove and β2M domains (Figure 2, A). Additionally, flexible loop regions connecting secondary structure elements within these domains, exhibited similar orientations. However, the α3-domain-situated flexible loop comprising residues 223-229, also called the '220's loop', exhibited distinct conformations for each molecule (Figure 2, B). Although variable positioning of the 220's loop was observed between molecules 1, 2, and 4, the most dramatically re-orientated loop was present in molecule 3. For example, the Gly-221 α-carbon of molecule 3 is situated 7.9 Å from that of molecule 4, 8.1 Å from that of molecule 2 and 9.2 Å from that of molecule 1 (Figure 2, C).

Furthermore, structural flexibility among molecules in the asymmetric unit of Mtb44-bound HLA-E was not limited to the 220's loop, with multiple α3-domain-derived β-strands exhibiting relative positional shifts. Superposition revealed that the β2M-distal β-sheet plane of the α3 domain of molecule 3 was substantially shifted relative to molecules 1, 2 and 4 (Figure 2, D).

For example, 2.7 Å separates the β-strand Val-270 α-carbon of molecules 3 and 4, and similarly sized movements are observed throughout the β-sheet plane.

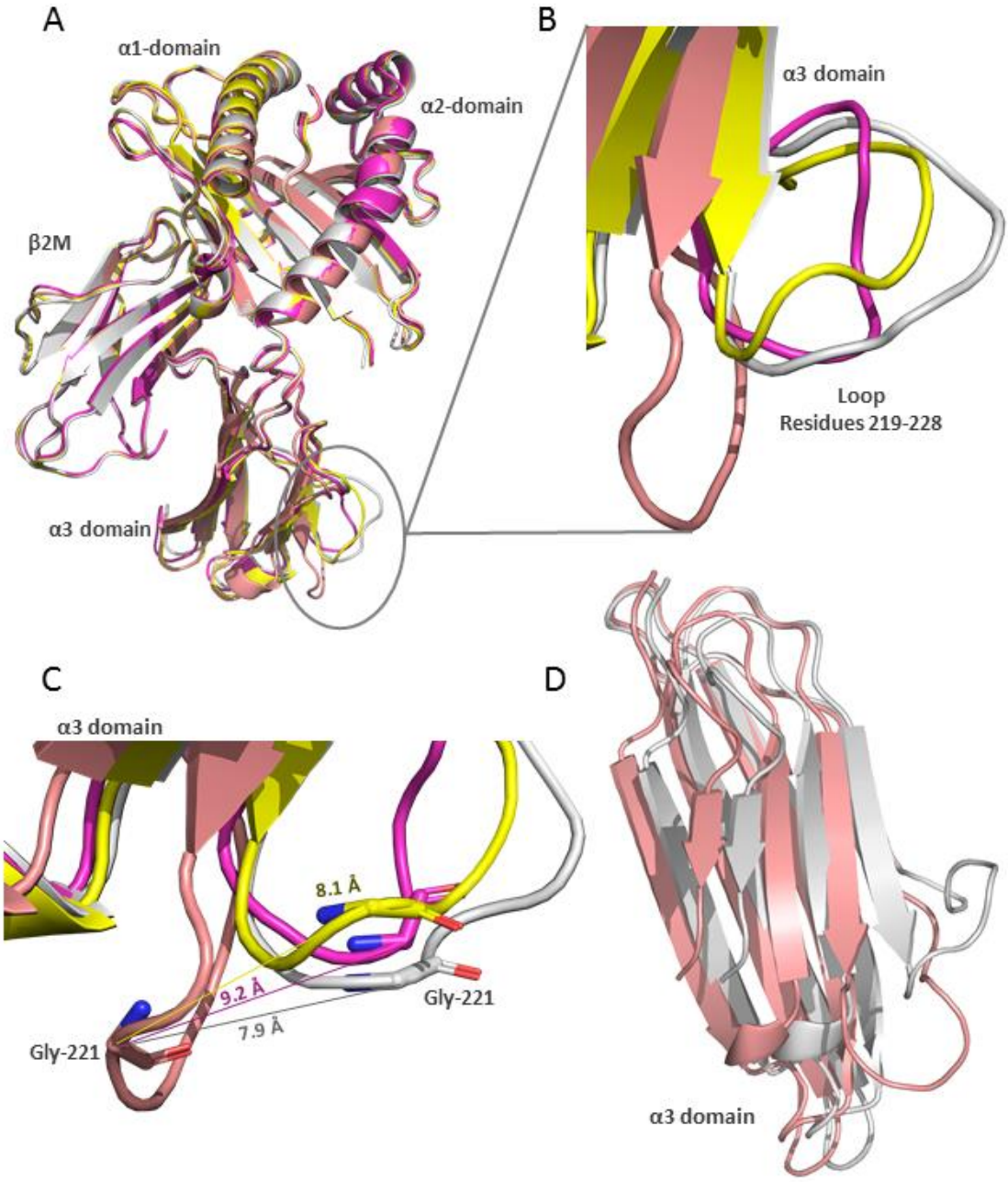


Figure 2: Alignment of the four molecules in the Mtb44-bound HLA-E*0103 structure

A. A cartoon visualisation of the 4 superposed molecules present in the asymmetric unit of the Mtb44-bound HLA-E*0103 structure. Molecule 1 is depicted in purple, 2 in yellow, 3 in light pink and 4 in grey. The flexible 220's loop region of the α3-domain is circled. **B.** Magnification of the 220's loop with colour coding of superposed molecules according to A. **C.** Visualisation of the flexible 220's loop, with position

221 Gly shown in stick-form. Distances between Gly-221 Cα atoms of the 4 molecules are depicted by lines and distances indicated in Å. The Gly-221 Cα of molecule 3 is situated 7.9 Å from that of molecule 4, 8.1 Å from that of molecule 2 and 9.2 Å from that of molecule 1. Colour coding of superposed molecules is according to A. **D.** Magnified α-3 domains from superposed Mtb44-HLA-E molecules 3 and 4, colour-coded according to A. Remaining heavy chain and β2M light chain residues omitted for clarity.

Root mean square deviation between molecules of the Mtb44-bound HLA-E*0103 structure

Multiple alignment in PDBeFOLD revealed molecules 3 and 4 deviated to the greatest extent. Root mean square deviation (rmsd) between matched α-carbons of molecule 3 and any of the remaining molecules was greater than that calculated between molecules 1, 2 or 4 (Figure 3, A). To demonstrate which regions were responsible for the plus 1 Å rmsd between molecules 3 and 4, various domains were sequentially removed from the coordinate file. In accordance with the conformational flexibility observed in PyMol Cα alignment, rmsd calculations between molecules 3 and 4 decreased from 1.026 to 0.957 following the exclusion of the 220's loop (Figure 3, B) and to 0.529 after the deletion of the entire α3 domain (Figure 3, C).

A

HLA-E*0103-Mtb44 structure RMSD (Å)

		Sym mol 1	Sym mol 2	Sym mol 3	Sym mol 4
Sym mol 1	Chain, A, B & P		0.497	0.745	0.642
Sym mol 2	Chains C, D & Q	0.497		0.753	0.554
Sym mol 3	Chains, E, F & R	0.745	0.753		1.026
Sym mol 4	Chain G, H & Z	0.642	0.554	1.026	

B

HLA-E*0103-Mtb44 structure (220's loop removed) RMSD (Å)

		Sym mol 1	Sym mol 2	Sym mol 3	Sym mol 4
Sym mol 1	Chain, A, B & P		0.478	0.635	0.619
Sym mol 2	Chains C, D & Q	0.478		0.672	0.532
Sym mol 3	Chains, E, F & R	0.635	0.672		0.957
Sym mol 4	Chain G, H & Z	0.619	0.532	0.957	

C

HLA-E*0103-Mtb44 structure α3-domain removed (residues 180-274) RMSD (Å)

		Sym mol 1	Sym mol 2	Sym mol 3	Sym mol 4
Sym mol 1	Chain, A, B & P		0.426	0.311	0.546
Sym mol 2	Chains C, D & Q	0.426		0.385	0.398
Sym mol 3	Chains, E, F & R	0.311	0.385		0.529
Sym mol 4	Chain G, H & Z	0.546	0.398	0.529	

Figure 3: Rmsd deviation between Mtb44-HLA-E*0103 molecules

A. Table of rmsd (Å) calculated by the PDBeFOLD server following alignment of Mtb44-bound HLA-E*0103 molecules in the asymmetric unit. **B.** Table of rmsd between molecules of the Mtb44-bound HLA-E*0103 structure in which heavy chain-derived residues of the 220's loop had been removed prior to alignment. **C.** Table of rmsd between molecules of the Mtb44-bound HLA-E*0103 structure in which the heavy chain-derived α-3 domain was deleted prior to alignment.

Abbreviations: Root mean square deviation (rmsd).

HLA-E heavy chain-derived alpha-3 domain and β2M interface analysis

The shifted α3 domain of molecule 3 in the Mtb44 HLA-E*0103 structure, raises the question of its impacts on the β2M interface. For example, such conformational readjustments might influence the inter-chain hydrogen bonding network between the α3 domain and β2M. An inter-chain interface analysis was therefore conducted in PDBePISA revealing 7 hydrogen bonds that connected β2M with the α3 domain in molecule 1, were lost in molecule 3 (Figure 4). Specifically, these 7 additional hydrogen bonds are formed between α3 domain-derived residues Thr-190, His-192, Glu-232, Arg-234, Pro-235 and Gly-237, and residues of the β2M light chain in molecule 1.

A **Hydrogen Bonds Symmetry molecule 1**

	Chain A: Heavy chain	Distance (Å)	Chain B: β2M
1	A:ARG 234[NH1]	2.96	B:GLN 9[OE1]
2	A:GLN 242[NE2]	3.42	B:SER 12[O]
3	A:GLN 242[NE2]	2.84	B:ARG 13[O]
4	A:ARG 48[NH2]	3.17	B:ASP 54[OD1]
5	A:GLN 32[NE2]	2.57	B:ASP 54[OD2]
6	A:ARG 48[NE]	2.68	B:ASP 54[OD2]
7	A:GLN 96[NE2]	2.97	B:TRP 61[O]
8	A:TYR 27[OH]	3.35	B:TYR 64[OH]
9	A:ARG 202[NH1]	3.10	B:ASP 99[O]
10	A:ARG 202[NH1]	3.34	B:ASP 99[OD2]
11	A:HIS 192[NE2]	2.35	B:ASP 99[OD2]
12	A:THR 190[OG1]	2.90	B:ASP 99[OD2]
13	A:ARG 234[NH1]	2.71	B:MET 100[O]
14	A:TRP 244[NE1]	2.70	B:MET 100[O]
15	A:ARG 234[NH2]	3.18	B:MET 100[OXT]
16	A:GLN 96[OE1]	2.70	B:HIS 32[NE2]
17	A:ASP 119[O]	3.10	B:ILE 2[N]
18	A:ASP 122[OD1]	2.63	B:TRP 61[NE1]
19	A:GLU 232[O]	3.09	B:GLN 9[NE2]
20	A:GLU 232[OE2]	2.38	B:LYS 7[NZ]
21	A:PRO 235[O]	2.41	B:TYR 11[OH]
22	A:ALA 236[O]	3.39	B:ARG 13[NE]
23	A:ALA 236[O]	3.02	B:ASN 25[ND2]
24	A:GLY 237[O]	3.27	B:ARG 13[NH1]

B **Hydrogen Bonds Symmetry molecule 3**

	Chain E: Heavy chain	Distance (Å)	Chain F: β2M
1	E:ARG 234[NH1]	3.10	F:GLN 9[OE1]
2	E:GLN 242[NE2]	3.37	F:SER 12[O]
3	E:GLN 242[NE2]	2.68	F:ARG 13[O]
4	E:ARG 35[NE]	3.02	F:ASP 54[OD1]
5	E:ARG 48[NE]	2.60	F:ASP 54[OD2]
6	E:GLN 32[NE2]	2.89	F:ASP 54[OD2]
7	E:GLN 96[NE2]	2.91	F:TRP 61[O]
8	E:TYR 27[OH]	3.26	F:TYR 64[OH]
9	E:ARG 202[NH1]	2.89	F:ASP 99[O]
10	E:ARG 234[NH1]	2.78	F:MET 100[O]
11	E:ARG 234[NH2]	2.94	F:MET 100[O]
12	E:TRP 244[NE1]	3.04	F:MET 100[O]
13	E:GLN 96[OE1]	2.81	F:HIS 32[NE2]
14	E:ASP 119[O]	3.23	F:ILE 2[N]
15	E:GLU 232[O]	3.11	F:GLN 9[NE2]
16	E:ALA 236[O]	3.26	F:ARG 13[NE]
17	E:ALA 236[O]	2.97	F:ASN 25[ND2]

Figure 4: HLA-E heavy chain-β2M inter-chain hydrogen bonding in molecules 1 and 3 of the Mtb44-HLA-E*0103 structure

A. Table of hydrogen bonds formed between the heavy chain and β2M in molecule 1 of the Mtb44-HLA-E*0103 structure. **B.** Table of hydrogen bonds connecting heavy chain and β2M in molecule 3 of the Mtb44-HLA-E*0103 structure.

Hydrogen bonds preserved between molecules 1 and 3 are shaded blue, whereas those unique to molecule 1 are shaded yellow. Lower and upper thresholds for the donor/acceptor distance range is set to 2.5 Å and 3.5 Å, respectively.

Temperature factors

As expected, the highest B factors of the Mtb44-HLA-E structure were situated in flexible surface loops outside secondary structure elements or at heavy chain and β2M polypeptide termini (Figure 5). However distinct temperature factor profiles were observed between intra-structural molecules. Interestingly, the shifted α3 domain of molecule 3 exhibited the lowest B-factors relative to those deriving from the α3 domains of molecules 1, 2 and 4.

As previously mentioned, molecule 3 also contains the most dramatically reoriented 220's loop - this region is densely populated with orange or red residues in molecules 1, 2 and 4, indicating it constitutes one of most disordered regions in these molecules. However, an absence of orange or red residues within this region of molecule 3 suggests that the reoriented form constitutes the most ordered 220's loop conformation sampled in the asymmetric unit.

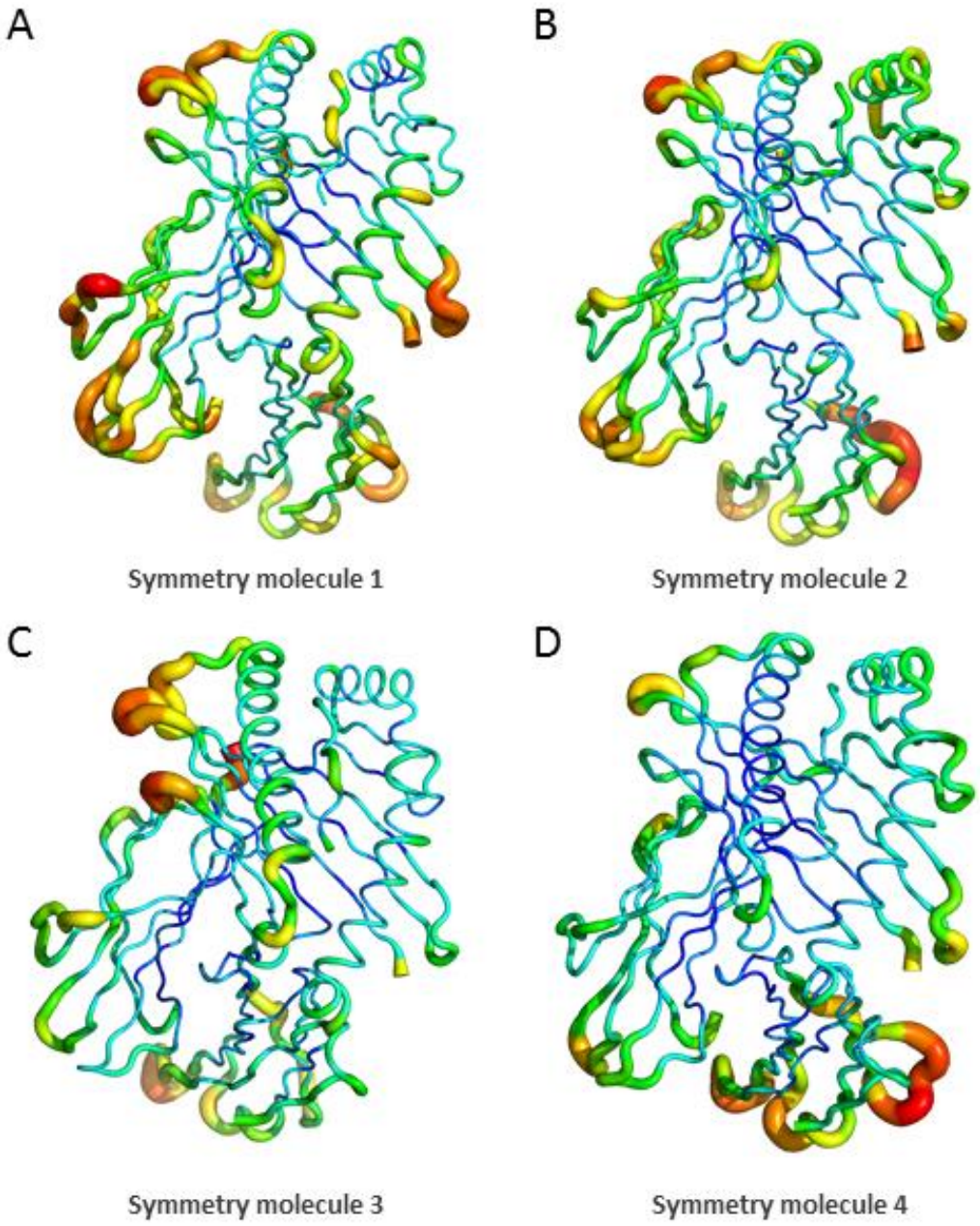


Figure 5: B factor analysis of Mtb44-HLA-E*0103 molecules in the asymmetric unit

PyMol-generated B factor putty representations for molecule 1 (**A**), molecule 2 (**B**), molecule 3 (**C**), and molecule 4 (**D**), of the Mtb44-HLA-E*0103 structure.

Cα atom B factors are depicted on a cool to warm colour scale with lowest relative B factors in blue and the highest in red.

Alignment of Mtb44 versus VL9-bound HLA-E molecules

Superposition of Mtb44 versus VL9 peptide-bound HLA-E molecules demonstrated similarly positioned secondary structure elements of the peptide binding groove and β2M (Figure 6, A). Further, the α3 domains of molecules 1, 2 and 4 of the Mtb44-HLA-E structure exhibited strong alignment to those modelled in published VL9-bound HLA-E structures (O'Callaghan et al., 1998) (Strong et al., 2003) (Hoare et al., 2008) (Figure 6, B), however the α3 domain of molecule 3 adopted a shifted orientation (Figure 6, C). Although a degree of conformational flexibility had previously been attributed to the 220's loop in VL9-bound structures of HLA-E (Hoare et al., 2008), this did not extend to the β-strands of the α3 domain, indicating such β-sheet repositioning is potentially a unique feature of HLA-E bound to non-VL9 peptides.

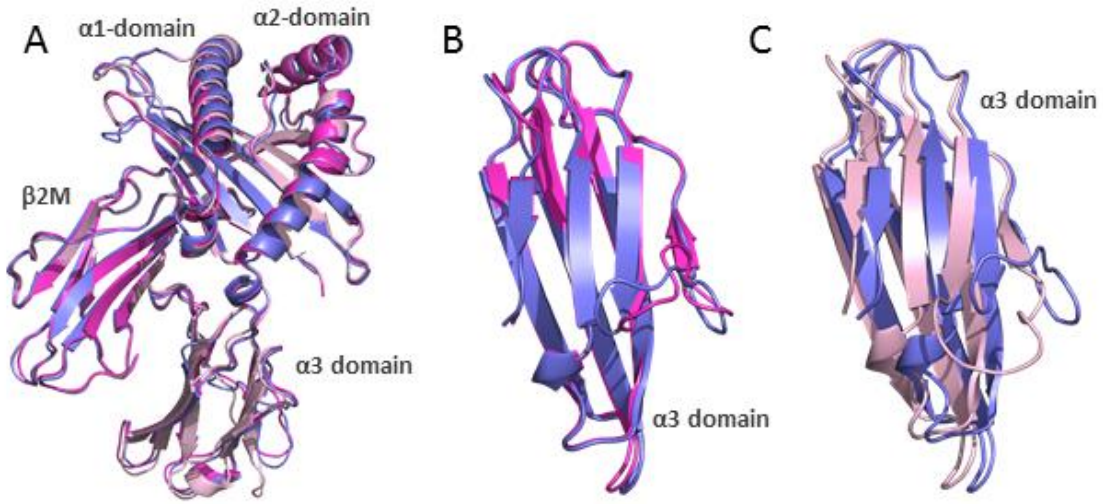


Figure 6: Superposition of Mtb44-HLA-E molecules with VL9 peptide-bound HLA-E

A. Superposition of Mtb44-bound HLA-E*0103 molecules 1 (magenta) and 3 (pink) with VL9-bound HLA-E*0101 molecule 1 of 1MHE (purple). **B.** Magnified α3 domain of superposed Mtb44-HLA-E molecule 1 (magenta) and VL9-HLA-E molecule 1 from 1MHE (purple). Remaining HLA-E heavy chain and β2M light chain residues omitted for clarity. **C.** Superposed α3 domain of Mtb44-HLA-E molecule 3 (light-pink) with VL9-HLA-E molecule 1 from 1MHE (purple).

Positioning of the Mtb44 peptide in the four molecules of the asymmetric unit

Superposition of the four Mtb44-HLA-E molecules revealed similarly orientated peptide backbones with minor α-carbon deviation not exceeding 1.2 Å (Figure 7). Unsurprisingly, the greatest displacement was observed at position 4 Ala between the Cα atoms of molecules 2 and 4. Position 4, in addition to position 5, lies within the kinked region of the peptide where the main chain projects out of the binding groove toward the solvent. Resultantly, this region forms fewer contacts with the HLA-E heavy chain and is subjected to fewer steric or pocket-based constraints in turn permitting the conformational variability observed between intra-structural molecules at this position. Furthermore, this region of the Mtb44 peptide might be subjected to even fewer constraints than the VL9 peptide in 1MHE which forms an additional hydrogen bond securing position 5 Arg of the VL9 peptide to the HLA-E heavy chain (O'Callaghan et al., 1998).

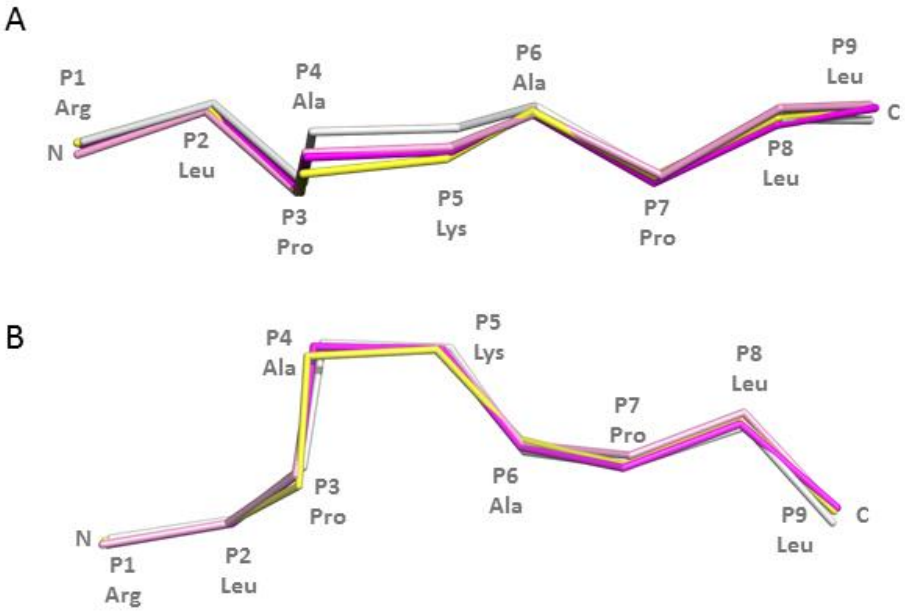


Figure 7: Intra-structural Mtb44 peptide alignment

A. Ariel view of the Mtb44 peptide backbone displayed in ribbon with the N and C terminals labelled in grey. Peptide residues are denoted with peptide position indicated as 'P'. The Mtb44-HLA-E*0103

molecule 1-derived peptide is depicted in magenta, molecule 2-derived peptide in yellow, molecule 3-derived peptide in light-pink and molecule 4-derived peptide in grey. **B.** Side-on visualisation of A.

A comparative B factor analysis of peptide main chain and side chain atoms from the four molecules in the asymmetric unit, revealed position 5 Lys had the highest all-atom average B factor in molecules 1, 2 and 3 and the second highest in molecule 4 (Figure 8). This is unsurprising given the kink in this central region of the peptide likely alleviates constraints imposed by pocket-anchor side chain interactions or hydrogen bonding at the peptide termini.

Given similar all-atom average B factor profiles for the Mtb44 peptides and similar positioning within the binding groove, as evidenced by intra-structural molecule superposition in PyMol, a single Mtb44 peptide from molecule 1 was selected for downstream analysis. Furthermore, the peptide from molecule 1 was chosen as it adopts an intermediate position in the unconstrained kinked region and lies between the peptide main chains present in molecules 2 and 4.

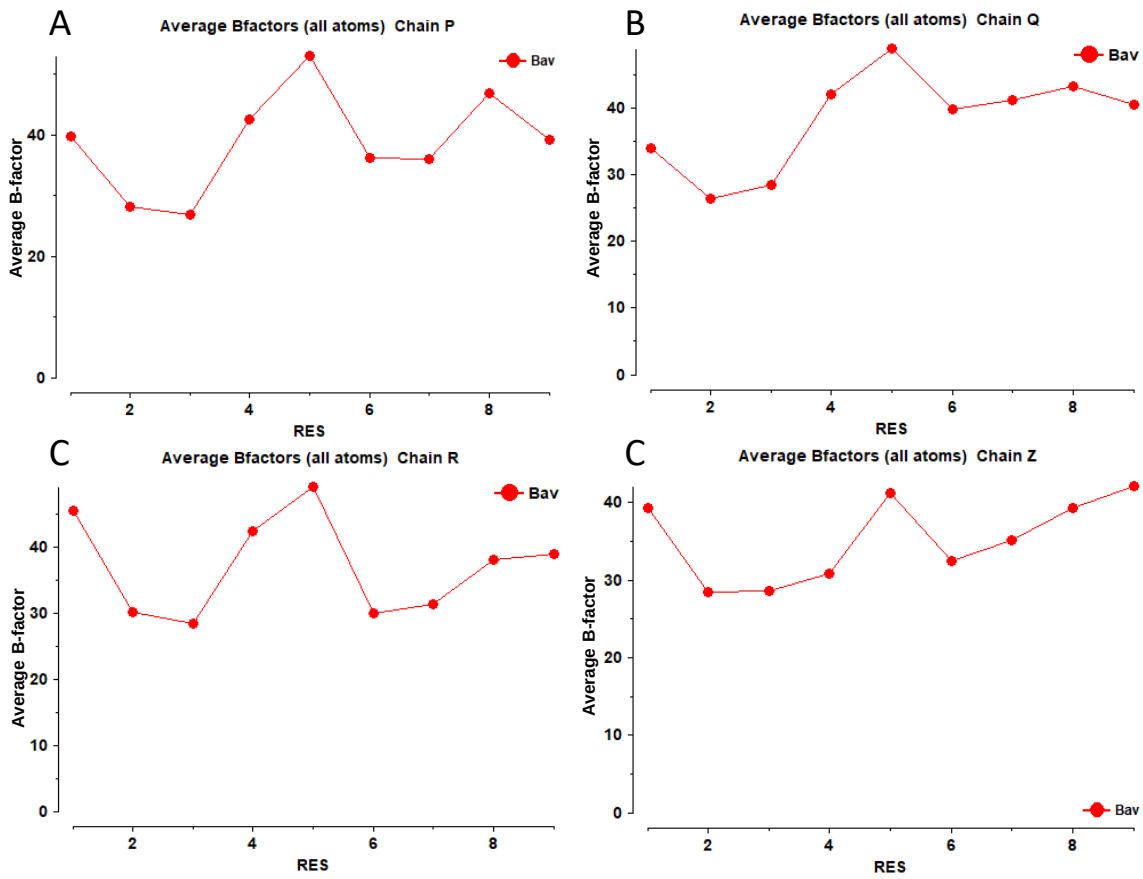


Figure 8: Comparative B factor analysis of Mtb44 peptides

All-atom average B factors (Å^2) for the Mtb44 peptide, Chain P, of molecule 1 (**A**), Chain Q of molecule 2 (**B**), Chain R of molecule 3 (**C**) and Chain Z of molecule 4 (**D**). The Y-axis denotes average B factor for main-chain and side chains atoms of each peptide residue, whereas peptide position is plotted on the X-axis.

Mtb44 peptide positioning within the HLA-E binding groove

Clear electron density for both the main chain and side chains of Mtb44 was present in the peptide binding groove (Figure 9, A). Position 2 and 9 primary anchor residues in the Mtb44 nonamer are biochemically similar to the canonical counterpart residues present in highly conserved VL9 peptides. Not only is the canonical position 9 Leu present in both VL9 and Mtb44 peptides, but the canonical position 2 Met is substituted for a biochemically similar, hydrophobic Leu in the Mtb44 peptide.

However, sequence differences exist at remaining positions, including secondary anchor residues, along the Mtb44 peptide relative to VL9. Notably, a positive charge is introduced at position 1 of the Mtb44 peptide, with Arg replacing the non-charged, canonical Val conserved among VL9 peptides. Furthermore, 2 Pro residues are substituted at positons 2 and 7 of the Mtb44 peptide in place of canonical Ala and Val, potentially introducing geometrical constraints on the peptide backbone. However, despite such sequence differences, superposition of Mtb44-HLA-E with VL9-bound HLA-E from 1MHE (O'Callaghan et al., 1998) revealed similar peptide positioning - Mtb44 conforms to the aforementioned classical peptide binding motif adopted by a large number of conventional MHC class I-restricted peptides and also by HLA-E-bound VL9 peptides (Figure 9, B and C).

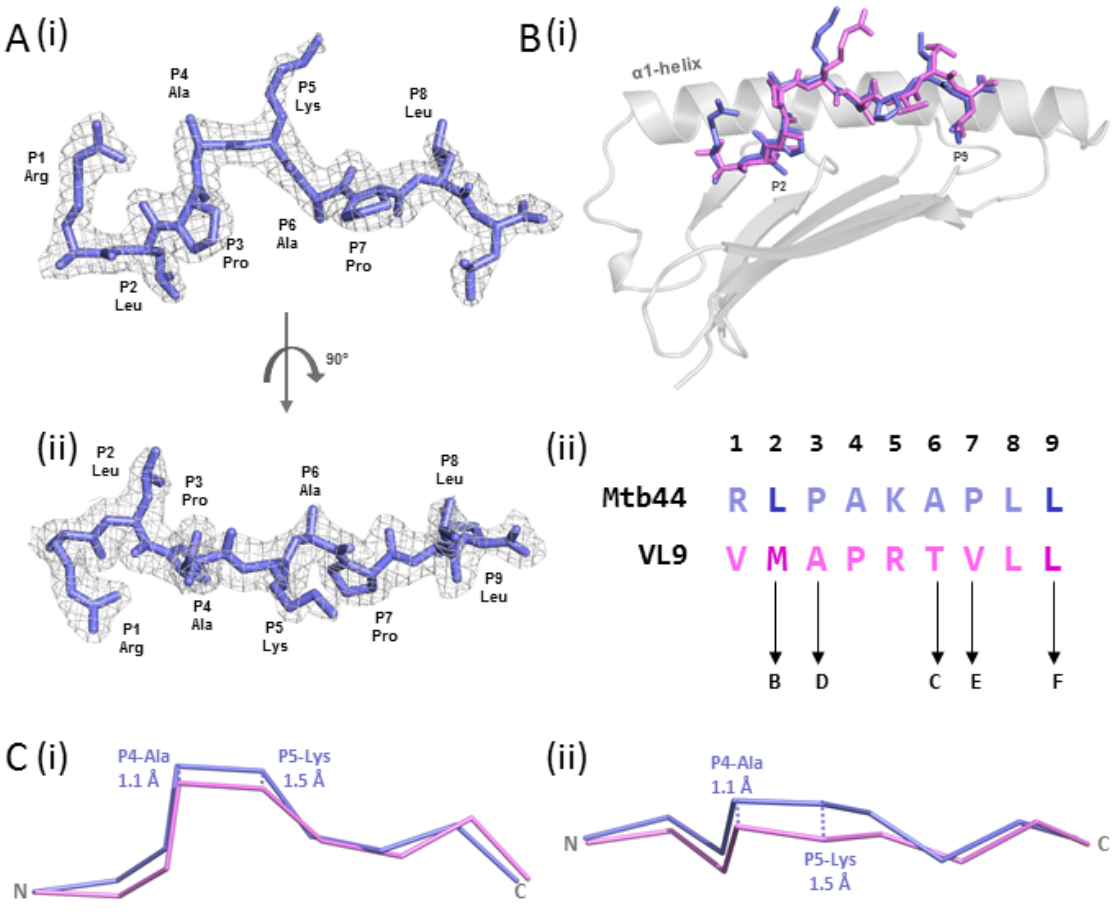


Figure 9: Analysis of Mtb44 peptide positioning

A. Mtb44 peptide shown in stick form, shaded purple and visualised from the side **(i)** and from an aerial

view. **(ii)** with an electron density map contoured to 1 sigma, shown in grey mesh overlaying the peptide. The HLA-E heavy chain and invariant β2M light chain are omitted for clarity. **B. (i)** The Mtb44 peptide (purple) is shown in stick form, superposed to the VL9 peptide (violet) from molecule 1 of 1MHE, with the α1 helix and β-sheet floor of the HLA-E peptide binding groove depicted in grey cartoon. The α-2 helix is omitted for clarity. **(ii)** Sequences for the Mtb44 and VL9 peptides illustrated in (i). **C.** Superposed backbone of the Mtb44 peptide (purple) and VL9 peptide (violet) from molecule 1 of 1MHE, side-on **(i)** and from above **(ii)**.

Pocket-binding in the HLA-E*0103-Mtb44 structure

The position 2-accommodating B pocket of the HLA-E binding groove is deep and hydrophobic in nature, with Met-45 conferring hydrophobicity and the small Ala-67 creating depth (O'Callaghan et al., 1998). Hence unsurprisingly, the position 2 Met to Leu substitution in the Mtb44 peptide is well tolerated. The Mtb44-derived position 2 Leu side chain projects down into the B pocket and aligns with the canonical Met side chain present in VL9 peptides (Figure 10, A). Furthermore, minimal side chain repositioning of B pocket-lining residues is observed in the Mtb44-bound HLA-E structure with B pocket-derived residues adopting similar rotamers to those present in VL9-bound structures of HLA-E. Such observations are consistent with previous structural and computational evidence that the HLA-E binding groove is unusually rigid compared to conventional MHC class I molecules (Strong et al., 2003) (Hansen et al., 2016). Regarding the F-pocket, canonical position 9 Leu, conserved among VL9 variants and also present in the Mtb44 peptide and is accommodated with minimal F pocket-derived side chain repositioning relative to VL9-bound HLA-E (Figure 10, B).

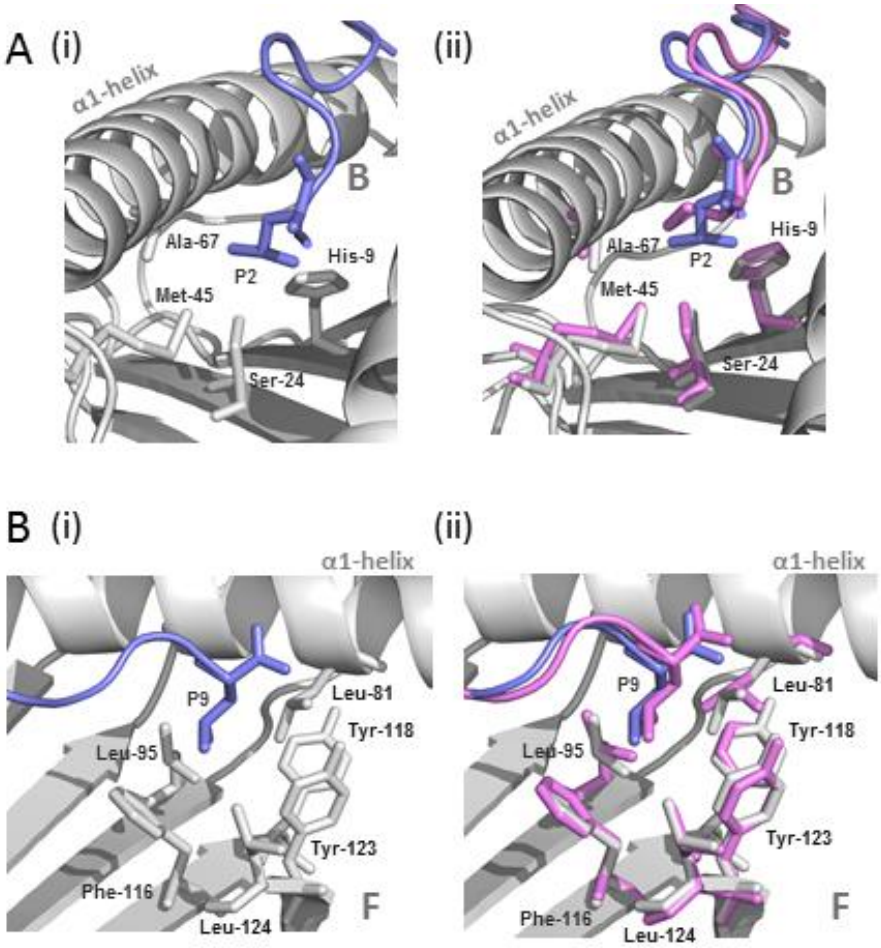


Figure 10: Primary B and F pocket occupancy by the Mtb44 peptide

A. (i) B pocket visualisation of Mtb44-bound HLA-E*0103 molecule 1. HLA-E B pocket forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as grey cartoon. The Mtb44 peptide is shown in purple cartoon with position 2 Leu in purple stick-form. **(ii)** Superposition of Mtb44-HLA-E molecule 1 with the VL9-bound molecule 1 of 1MHE. B pocket visualisation is colour-coded according to A (i), with superposed B pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 2 Met in stick form. **B. (i)** Visualisation of F pocket for Mtb44-bound HLA-E*0103 molecule 1. Heavy chain-derived F pocket-forming residues are depicted as grey sticks and the surrounding area of the peptide binding groove as grey cartoon. The Mtb44 peptide is shown in purple cartoon with position 9 Leu in purple stick-form. **(ii)** Superposition of the Mtb44-HLA-E molecule 1 with VL9-bound molecule 1 from 1MHE. F pocket visualisation is colour-coded according to B (i), with the

superposed F pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 9 Leu in violet stick-form.

Position 3 Pro of the Mtb44 peptide is accommodated by the D pocket with minimal side chain repositioning of pocket forming residues relative to VL9-bound HLA-E (Figure 11, A). Given the shallow nature of the D pocket, it is likely that canonical Ala to Pro substitution is well tolerated due to the small size of both such residues. Similarly, it is possible the canonical Val to Ala substitution present at position 6 of the Mtb44 peptide is permitted by the shallow C pocket due to the absence of a prohibitively large side chain (Figure 11, B).

The E pocket of HLA-E is deep and hydrophobic, and tolerates position 7 Pro of the Mtb44 peptide in place of a canonical Val or Leu present among VL9 peptides. E-pocket forming-residues exhibit the greatest structural readjustment between Mtb44 and VL9-bound HLA-E structures with Glu-152 and Ser-147 constituting the most distorted side chains (Figure 11, C). Alternate positioning of the Glu-152 side-chain is likely due to its participation in a hydrogen bond with position 5 Arg of the VL9 peptide, which is absent in the Mtb44-HLA-E structure.

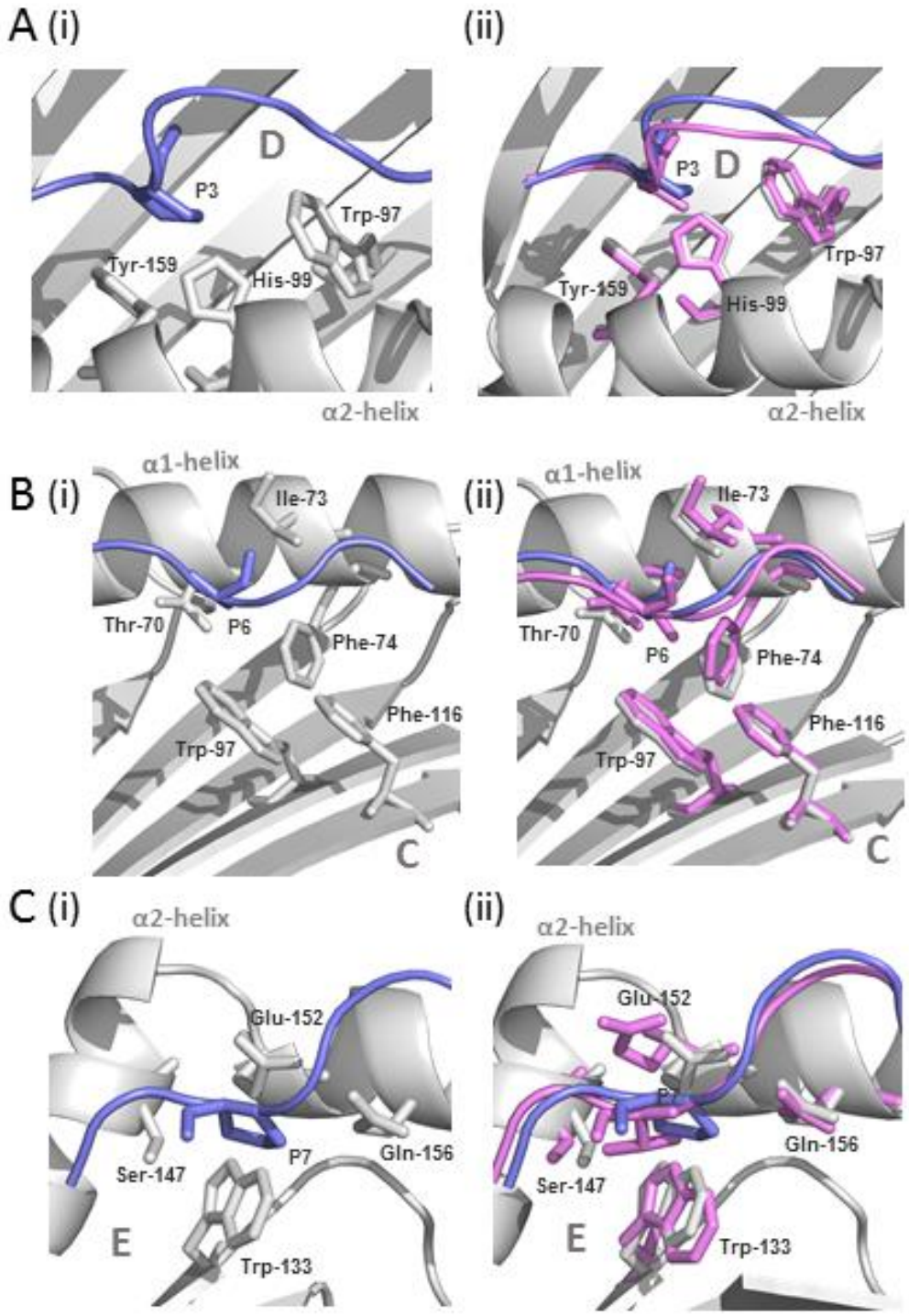


Figure 11: Secondary D, C and E pocket occupancy by the Mtb44 peptide

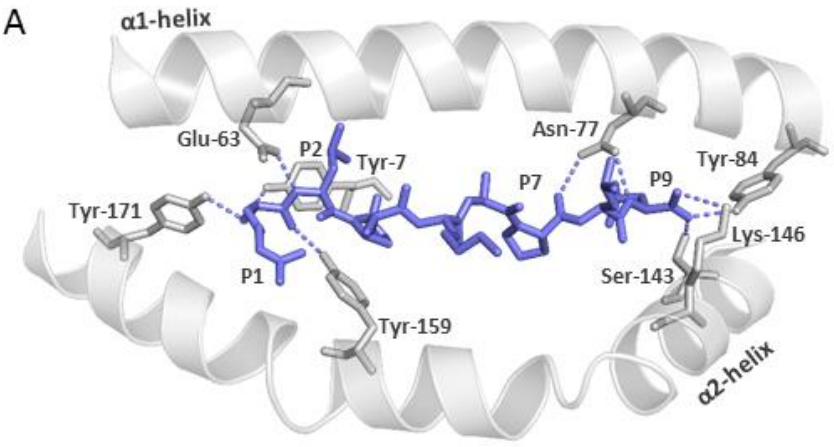
A. (i) D pocket visualisation of the Mtb44-bound HLA-E*0103 molecule 1. Heavy chain-derived D pocket-forming residues are depicted as grey sticks and the neighbouring area of the peptide binding groove as grey cartoon. The Mtb44 peptide is shown in purple cartoon with position 3 Pro in purple stick-form. **(ii)** Superposition of the Mtb44-HLA-E molecule 1 with the VL9-bound molecule 1 of 1MHE. D pocket

visualisation is colour-coded according to A (i), with superposed D pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 3 Ala in stick form. **B. (i)** Visualisation of the C pocket for Mtb44-bound HLA-E*0103 molecule 1. Heavy chain-derived C pocket-forming residues are depicted as grey sticks with the surrounding area of the peptide binding groove as grey cartoon. The Mtb44 peptide is shown in purple cartoon with position 6 Ala in purple stick-form. **(ii)** Superposition of the Mtb44-HLA-E molecule 1 with the VL9-bound molecule 1 of 1MHE. C pocket visualisation is colour-coded according to B (i) with the superposed C pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 6 Thr in stick-form. **C. (i)** E pocket visualisation of the Mtb44-bound HLA-E*0103 molecule 1. Heavy chain-derived E pocket-forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as grey cartoon. The Mtb44 peptide is shown in purple cartoon with position 7 Pro in purple stick-form. **(ii)** Superposition of the Mtb44-HLA-E molecule 1 with the VL9-bound molecule 1 of 1MHE. E pocket visualisation is colour-coded according to C (i) with superposed E pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 7 Val in stick-form.

Hydrogen bonding

Analyses of Mtb44-HLA-E inter-chain hydrogen bonding networks revealed a loss of the two hydrogen bonds securing position 5 Arg of the VL9 peptide to heavy chain-derived α2 helix-situated residues Glu-152 and Gln-156 in 1MHE (Figure 12). Consequently, the Mtb44 peptide main-chain is less constrained at position 5 and is orientated closer to the α1 helix relative to VL9. It has previously been shown that hydrogen bonds connecting VL9 peptides to the HLA-E heavy chain predominantly involve the N or C terminal peptide residues (O'Callaghan et al., 1998). Consistent with this, six of the nine peptide-heavy chain hydrogen bonds formed in the Mtb44-HLA-E structure involve terminal peptide residues position 1 Arg or position 9 Leu.

A



B

HLA-E*0103 - Mtb44

	Peptide	Dist. Å	Heavy chain
1	P1-Arg [N]	2.61	Tyr-171 [OH]
2	P1-Arg [O]	2.77	Tyr-159 [OH]
3	P2-Leu [N]	2.6	Glu-63 [OE2]
4	P2-Leu [N]	3.68	Tyr-7 [OH]
5	P7-Pro [O]	3.16	Asn-77 [ND2]
6	P9-Leu [N]	2.94	Asn-77 [OD1]
7	P9-Leu [O]	2.75	Tyr-84 [OH]
8	P9-Leu [O]	2.76	Ser-143 [OG]
9	P9-Leu [OXT]	3.18	Lys-146 [NZ]

Figure 12: Mtb44 peptide-HLA-E heavy chain hydrogen bonding network

A. Visualisation of the inter-chain hydrogen bonding network between the Mtb44 peptide and the HLA-E heavy chain. The Mtb44 peptide is shown in purple stick-form with heavy-chain-derived donor/acceptor residues depicted as grey sticks. Hydrogen bonds are shown as purple dashed lines. The α1 and 2 helices are visualised in grey cartoon with the remaining HLA-E heavy and light chain residues omitted for clarity. **B.** Table detailing hydrogen bonds formed between the Mtb44 peptide and the HLA-E heavy chain. Lower and upper thresholds for donor/acceptor distances are set to 2.5 Å and 3.5 Å, respectively.

Structural determination of HLA-E*0103 in complex with Mtb14

*HLA-E*0103-Mtb14*

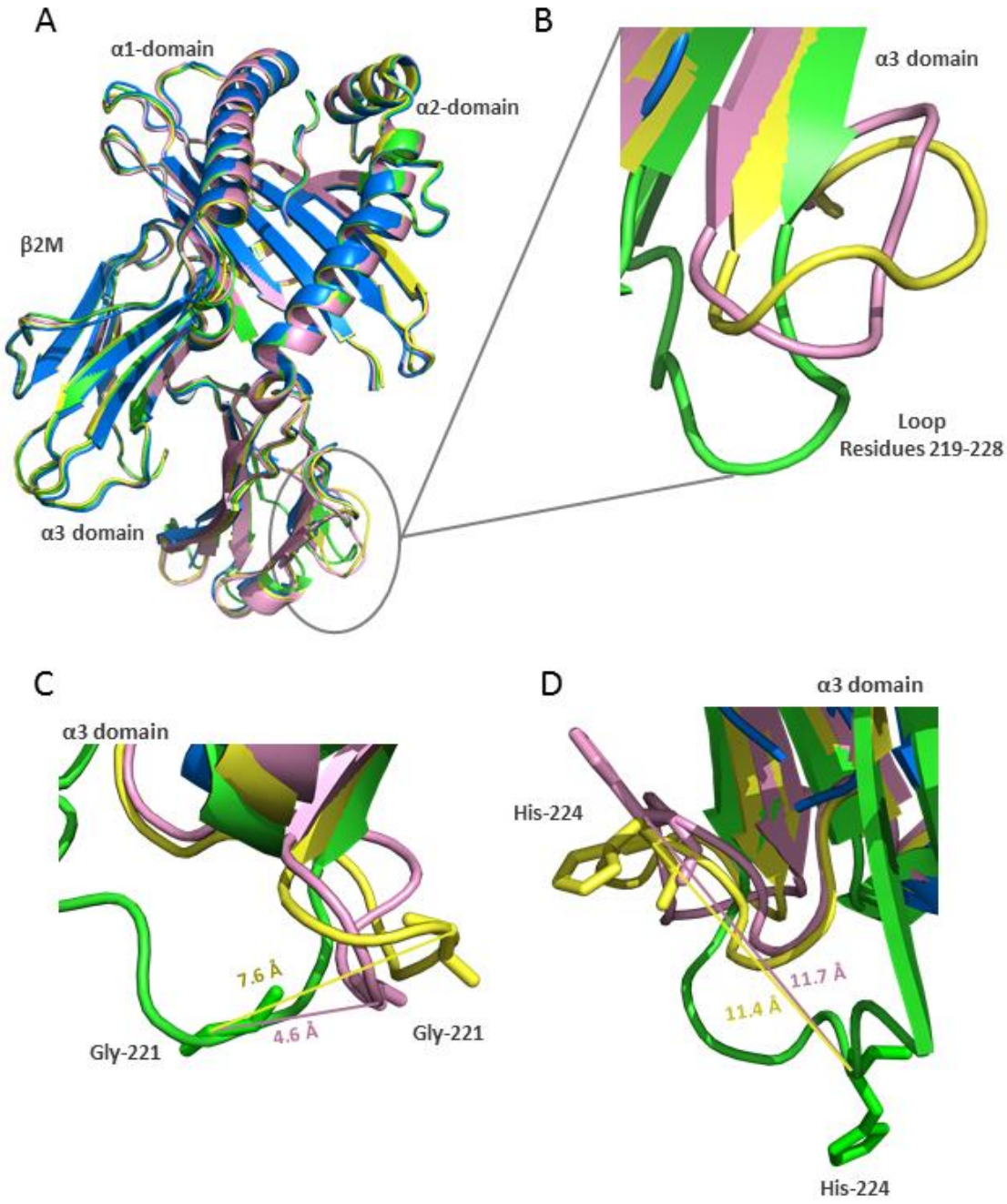
Crystallisation and refinement statistics for HLA-E*0103 in complex with the Mtb-derived Mtb14 (RMAATAQVL) are detailed in Figure 1. In summary, crystals grew in wells containing 3 M Ammonium sulphate and 0.1 M MES at pH 6, 78 days post assembly. Diffraction data were collected to 2.05 Å, however unlike 6GH1, packing was in the C2 space group. Initial Xtriage analysis of the Mtz file in Phenix revealed moderately anisotropic diffraction. However, anisotropy correction by the STARANISO server reduced the overall completeness of the dataset from 99% to 89%. Considering the anisotropic diffraction was not severe and also that Phaser incorporates automated anisotropy correction during molecular replacement, the original, uncorrected Mtz data file was selected for downstream analysis. Molecule 1 from the asymmetric unit of 6GH1 with the Mtb44 peptide coordinates removed, was used as a phasing model and similar to 6GH1, 4 copies were located in an asymmetric unit comprising 44% solvent. The final R factors for Mtb14 in complex with HLA-E*0103, following refinement and validation were 0.2305 and 0.2730 for R_{work} and R_{free}, respectively.

Mtb14-bound HLA-E*0103 molecule alignment

Sufficient electron density enabled modelling of both secondary structure elements and flexible loop regions for 3 of the 4 molecules in the asymmetric unit in their entirety. However, the HLA-E heavy chain of molecule 3, chain E, contained multiple regions of disorder which prohibited model building. Resultantly, residues 194-196, 218-227 and residue 248 were omitted from the model. All such residues are situated in the α3 domain of the heavy chain with positions 194-196 and 218-227 comprising surface loops and residue 248 forming part of a β-strand. Similar to the Mtb44 structure, electron density for Zinc ions was also present in the asymmetric unit which exhibited ordered coordination with Glu-183 and His-181 of the HLA-E heavy chain.

Similar to the Mtb44-bound HLA-E*0103 structure, Cα backbone superposition revealed that the α-helices, β-sheet floor and surface loops of the peptide binding groove, in addition to the β-sheets and flexible loops of the β2M light chain, exhibited strong alignment between Mtb14-HLA-E*0103 molecules (Figure 13, A). However, the 220's loop connecting β-sheet planes in the α3 domain exhibited the greatest conformational repositioning between molecules 1, 2 and 4 - this region is absent in molecule 3 (Figure 13, B). A distinct loop orientation was observed for these molecules with molecule 1 exhibiting the most dramatic conformational rearrangement. A distance of 4.6 Å was measured between Gly-221 Cα atoms of molecules 1 and 2 and 7.6 Å between those of molecules 1 and 4 (Figure 13, C). Furthermore, distances of 11.4 Å and 11.7 Å respectively separate the His-224 Cα atoms of molecules 1 and 4 and molecules 1 and 2 (Figure 13, D). Additionally, residues Gln-226-Val-231 rearrange to form an extended β-strand which is absent in molecule 3 and limited to residues Glu-229-Val-231 in molecules 4 and 2.

Consistent with Mtb44-bound HLA-E*0103 molecules, variable positioning was not limited to the α3-domain flexible 220's loop. The β2M-distal β-sheet plane of the α3-domain exhibits a spectrum of shifted profiles among Mtb14-HLA-E molecules (Figure 13, E). As such shifts were observed along a horizontal axis, vertical distances between the α3 domain and peptide binding groove were unaffected, similar to that observed in Mtb44 molecule 3. Mtb14-bound HLA-E*0103 molecules 1 and 4 adopted intermediate α3 domain profiles, whereas molecules 2 and 3 exhibited the most shifted α3 domain profiles with β-strand shifts in opposite directions along the horizontal axis. However, such movements were not as dramatic as those observed between shifted α3 domains of the Mtb44-HLA-E structure. For example, 2.7 Å separates Val-270 Cα atoms of molecules 3 and 4 in the Mtb44-bound HLA-E*0103 structure, whereas 1.7 Å separates those of molecules 2 and 3 in the Mtb14-bound HLA-E*0103 structure (Figure 13, F).


A
α1-domain
α2-domain
β2M
α3 domain
B
α3 domain
Loop
Residues 219-228
C
α3 domain
7.6 Å
4.6 Å
Gly-221
Gly-221
D
α3 domain
His-224
11.7 Å
11.4 Å
His-224

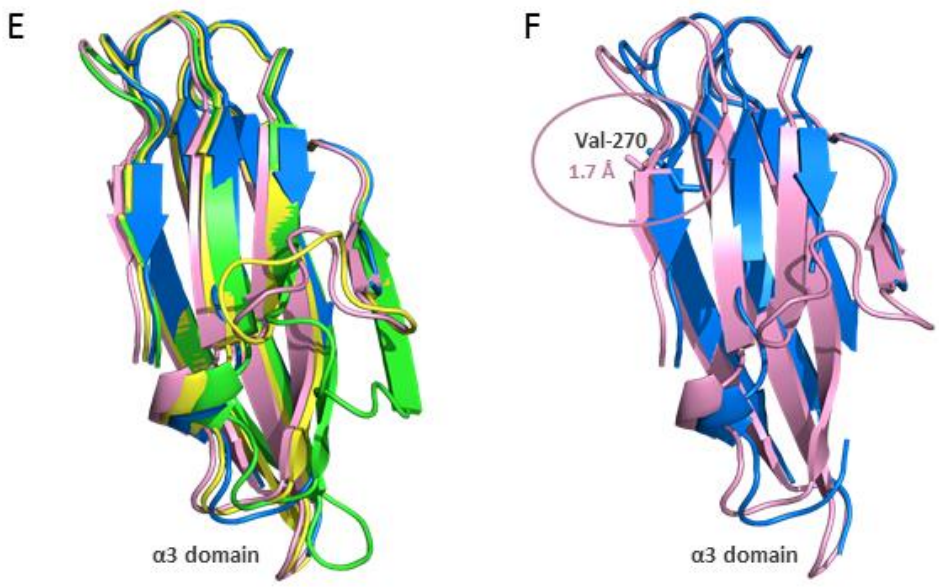


Figure 13: Analysis of Mtb14-bound HLA-E*0103 molecules in the asymmetric unit

A. A cartoon visualisation of the 4 superposed molecules present in the asymmetric unit of the Mtb14-bound HLA-E*0103 structure. Molecule 1 is depicted in green, 2 in pink, 3 in blue and 4 in yellow. The flexible 220's loop of the α3 domain is circled. **B.** Magnification of the 220's loop with colour coding of superposed molecules according to A. **C.** Visualisation of the flexible 220's loop, with position 221 Gly shown in stick-form. Distances between Gly-221 Cα atoms of the 4 molecules are depicted by lines and measured in Å. The Gly-221 Cα in molecule 1 is situated 4.6 Å from that of molecule 2 and 7.6 Å from that of molecule 4. Insufficient electron density prohibited modelling of Gly-221 in molecule 3. **D.** A visualisation of the flexible 220's loop with molecule colour coding according to A and His-224 depicted as sticks. Distances between His-224 Cα atoms are depicted by lines. The His-224 Cα in molecule 1 is situated 11.7 Å from that of molecule 2 and 11.4 Å from that of molecule 4. Insufficient electron density prohibited modelling of His-224 in molecule 3 **E.** Superposed α3 domains from the 4 Mtb14-HLA-E*0103 molecules, colour coded according to A. **F.** Superposed α3 domains of molecules 2 and 3, colour coded according to A, with Val-270 depicted in stick-form and the distance between Val-270 Cα atoms denoted in Å.

Temperature factor analysis

As expected, regions exhibiting the highest relative B factors were situated in flexible surface loops outside of secondary structure elements or at the polypeptide termini, for both the heavy and light chains in the Mtb14 HLA-E*0103 structure. However, distinct temperature factor profiles were observed between molecules in the asymmetric unit (Figure 14). Consistent with the presence of an extended β-strand, a high density of cool-coloured residues indicative of low α-carbon B factors, were observed in the α3-domain and 220's loop region of molecule 1. Furthermore, a high density of red and orange residues reflecting high α-carbon B factors, was consistent with the α3-domain of molecule 3 containing multiple regions of disorder.

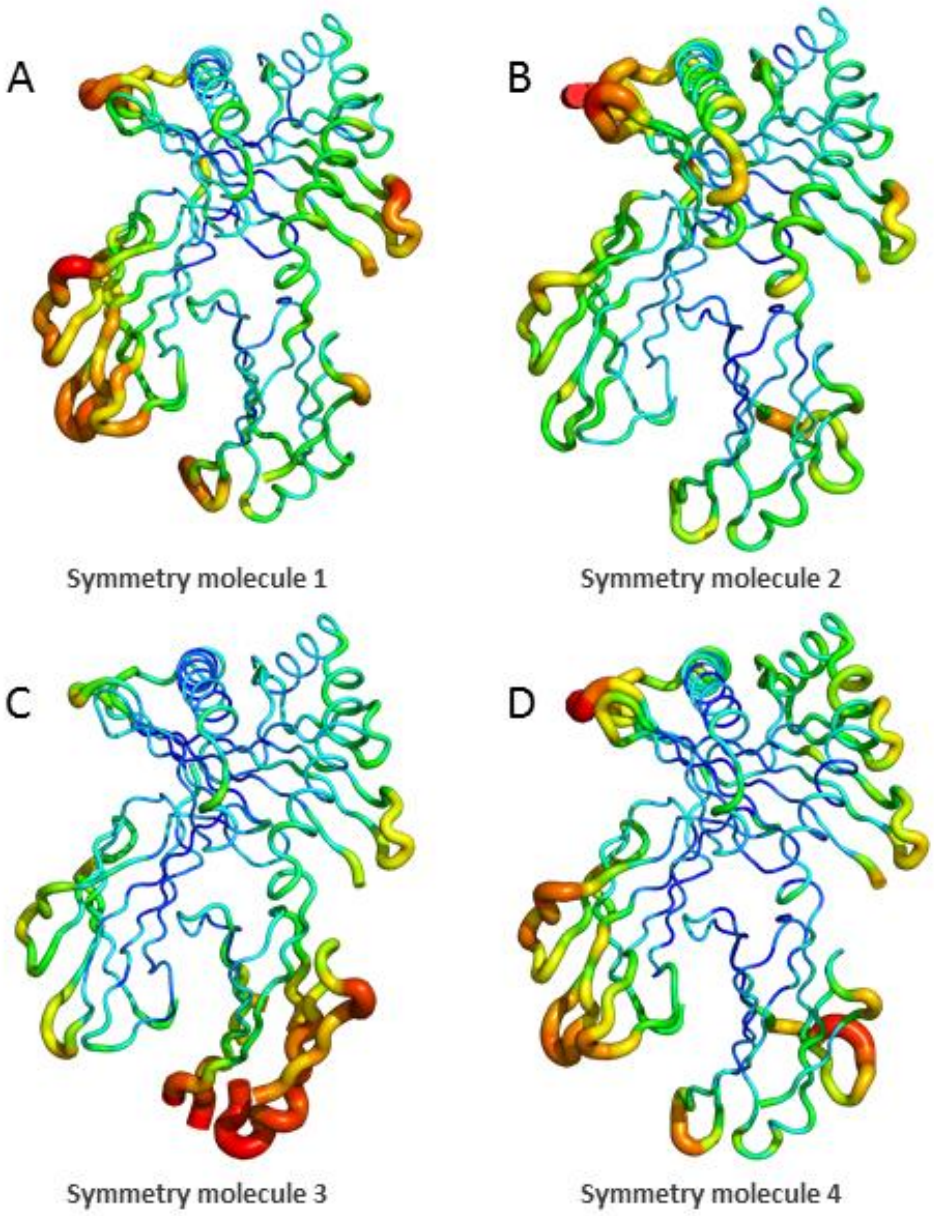


Figure 14: B factor analysis of Mtb14-HLA-E*0103 molecules

PyMol-generated B factor putty representations for molecule 1 (**A**), molecule 2 (**B**), molecule 3 (**C**), and molecule 4 (**D**), of the Mtb14-HLA-E*0103 structure.

Cα atom B factors are depicted on a cool to warm colour scale with lowest relative B factors in blue and the highest in red.

Superposition of Mtb14 versus VL9 peptide-bound HLA-E molecules demonstrated highly similar positioning of secondary structure elements in the vicinity of the peptide binding groove and β2M (Figure 15, A). However, the β2M-distal α3 domain-situated β-sheet plane of

molecules 2 and 3 of the Mtb14 structure were shifted in position relative to the superposed α3 domain of VL9-bound HLA-E (Figure 15, B-E). Furthermore, the extended β-strand present in molecule 1 of the Mtb14 structure, is absent in all VL9-bound molecules of HLA-E. Additionally, the flexible 220's loop orientation preceding the extended β-strand in molecule 1 exhibits the greatest deviation from that observed in the VL9 structure.

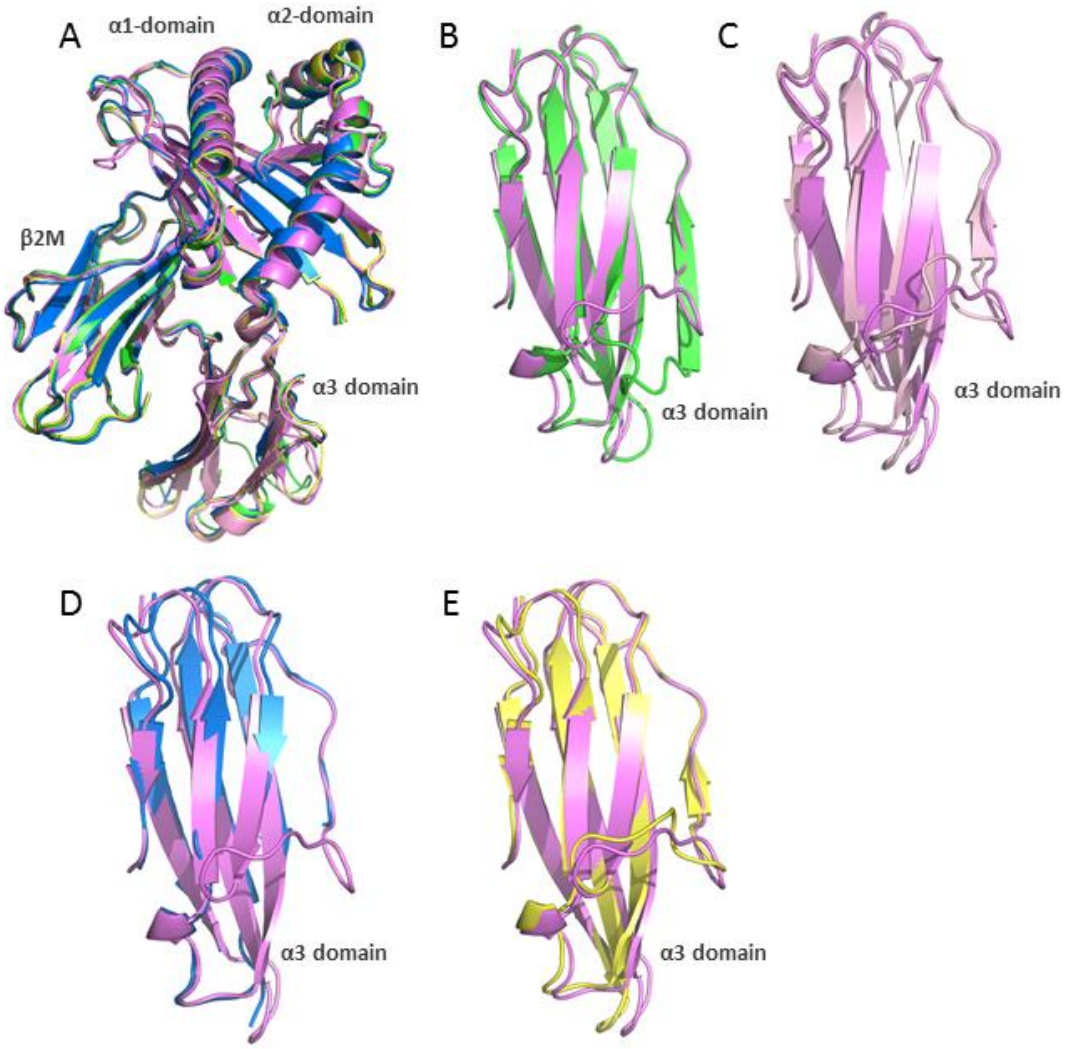


Figure 15: Superposition of Mtb14-HLA-E molecules with VL9 peptide-bound HLA-E

A. Superposition of Mtb14-bound HLA-E*0103 molecules 1 (green), 2 (light-pink), 3 (blue) and 4 (yellow) with VL9-bound HLA-E*0101 molecule 1 of 1MHE (violet). **B.** Magnified α3 domain of superposed Mtb14-HLA-E molecule 1 (green) and VL9-HLA-E molecule 1 from 1MHE (violet). Remaining HLA-E heavy chain and β2M light chain residues omitted for clarity. **C.** Superposed α3 domain of Mtb14-HLA-E

molecule 2 (light-pink) with VL9-HLA-E molecule 1 from 1MHE (violet). **D.** Superposed α3 domain of Mtb14-HLA-E molecule 3 (blue) with VL9-HLA-E molecule 1 from 1MHE (violet). **E.** Superposed α3 domain of Mtb14-HLA-E molecule 4 (yellow) with VL9-HLA-E molecule 1 from 1MHE (violet).

Superposition of Mtb14 and Mtb44 was conducted to investigate whether their shifted α3 domains (Mtb44 molecule 3 and Mtb14 molecules 2 and 3) aligned. As expected, superposed Mtb14 and Mtb44-bound molecules revealed strong alignment of the peptide binding groove and β2M (Figure 16, A and D). Similarly, the intermediately positioned α3 domains of Mtb14 molecules 1 and 4 aligned to the non-shifted Mtb44 α3 domain of molecule 1 (Figure 16, B). However, the shifted α3 domain of Mtb44 molecule 3 did not align to either of the α3 domain shifted profiles of Mtb14 molecules 2 and 3, but instead was re-positioned further along the same, horizontal axis (Figure 16, C-E). It could therefore be concluded that the β2M-distal β-sheet plane of the α3-domain of non-VL9-bound HLA-E can adopt a spectrum of horizontally shifted conformations.

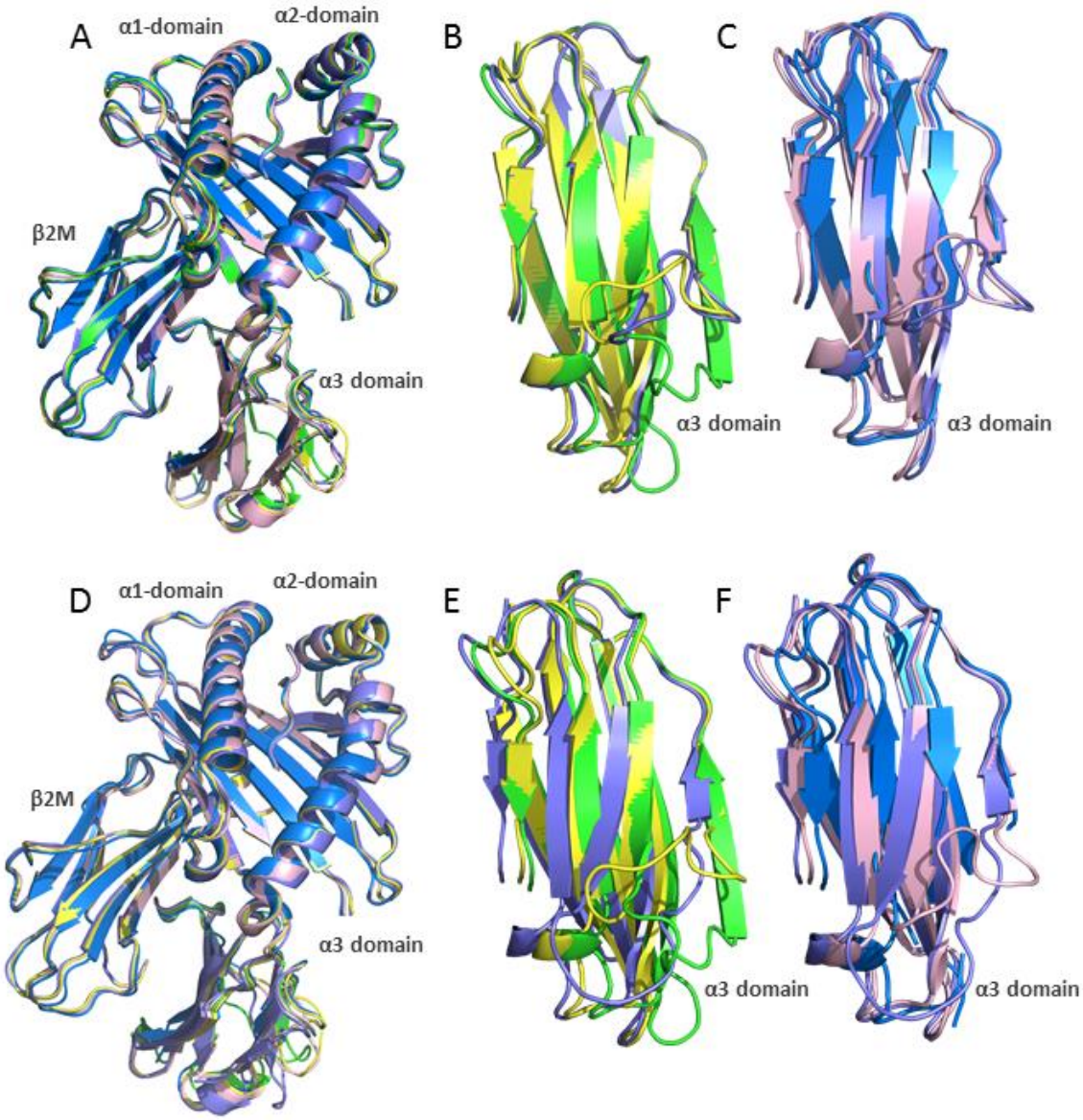


Figure 16: Superposition of Mtb14 and Mtb44 peptide-bound HLA-E molecules

A. Superposition of the Mtb14-bound HLA-E*0103 molecules 1 (green), 2 (light pink), 3 (blue) and 4 (yellow) with molecule 1 from the Mtb44-bound HLA-E*0103 structure (purple). **B.** Visualisation of the α3 domains of superposed molecules 1 and 4 from the Mtb14 structure and molecule 1 from the Mtb44 structure, colour-coded according to A. **C.** Visualisation of the α3 domains of superposed molecules 2 and 3 from the Mtb14 structure and molecule 1 from the Mtb44 structure, colour-coded according to A. **D.** Superposition of the Mtb14-bound HLA-E*0103 molecules 1 (green), 2 (light pink), 3 (blue) and 4 (yellow) with molecule 3 from the Mtb44-bound HLA-E*0103 structure (purple). **E.** Visualisation of the α3 domains of superposed molecules 1 and 4 from the Mtb14 structure and molecule 3 from the Mtb44

structure, colour-coded according to D. **C.** Visualisation of the α3 domains of superposed molecules 2 and 3 from the Mtb14 structure and molecule 3 from the Mtb44 structure, colour-coded according to D.

Mtb14 peptide positioning between molecules of the asymmetric unit

Superposition of the four molecules in the Mtb14-bound HLA-E*0103 structure, revealed similarly orientated peptide backbones with minor Cα atom repositioning, not exceeding 0.6 Å (Figure 17). Unsurprisingly, the greatest deviation was observed at position 4 Ala between the Cα atoms of molecules 1 and 2. As previously mentioned, this region is solvent-exposed and thus subjected to fewer constraints, in turn permitting conformational variability between molecules. Consistent with this, position 5 of the kinked peptide region exhibited the highest or second high relative B factor in the Mtb14 peptides of all four molecules in the asymmetric unit (Figure 18). Due to similar peptide positioning and B factor profiles, a single Mtb14 peptide of molecule 1 was selected for downstream analysis.

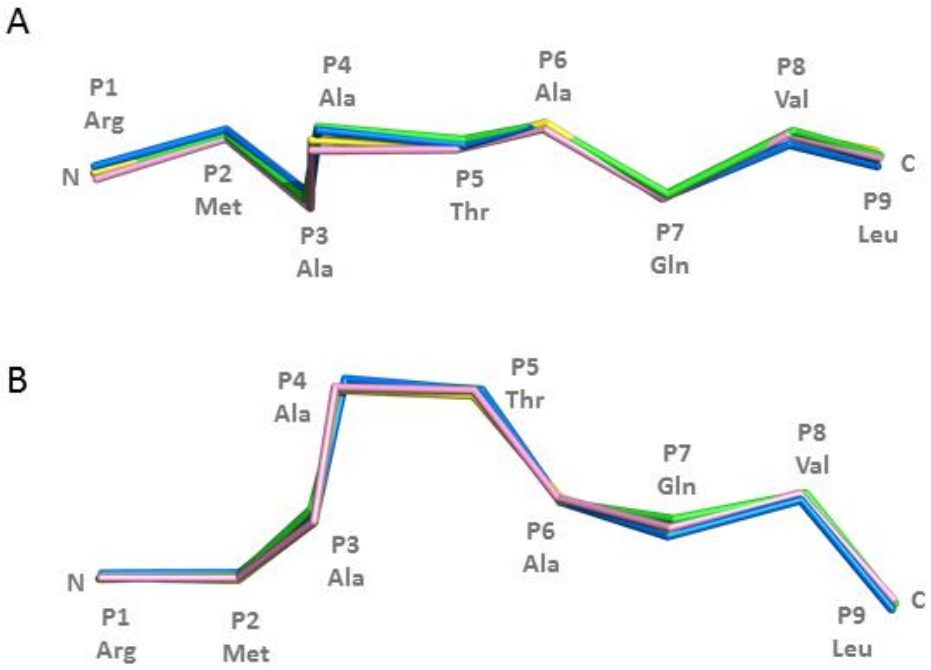


Figure 17: Comparison of Mtb14 peptide positioning

A. Visualisation from above of superposed Mtb14 peptide backbones from molecules 1 (green), 2 (pink),

3 (blue) and 4 (yellow), displayed in ribbon-form, with the N and C terminals denoted in grey and residue positions labelled along the nonamer. **B.** Side-on visualisation of A.

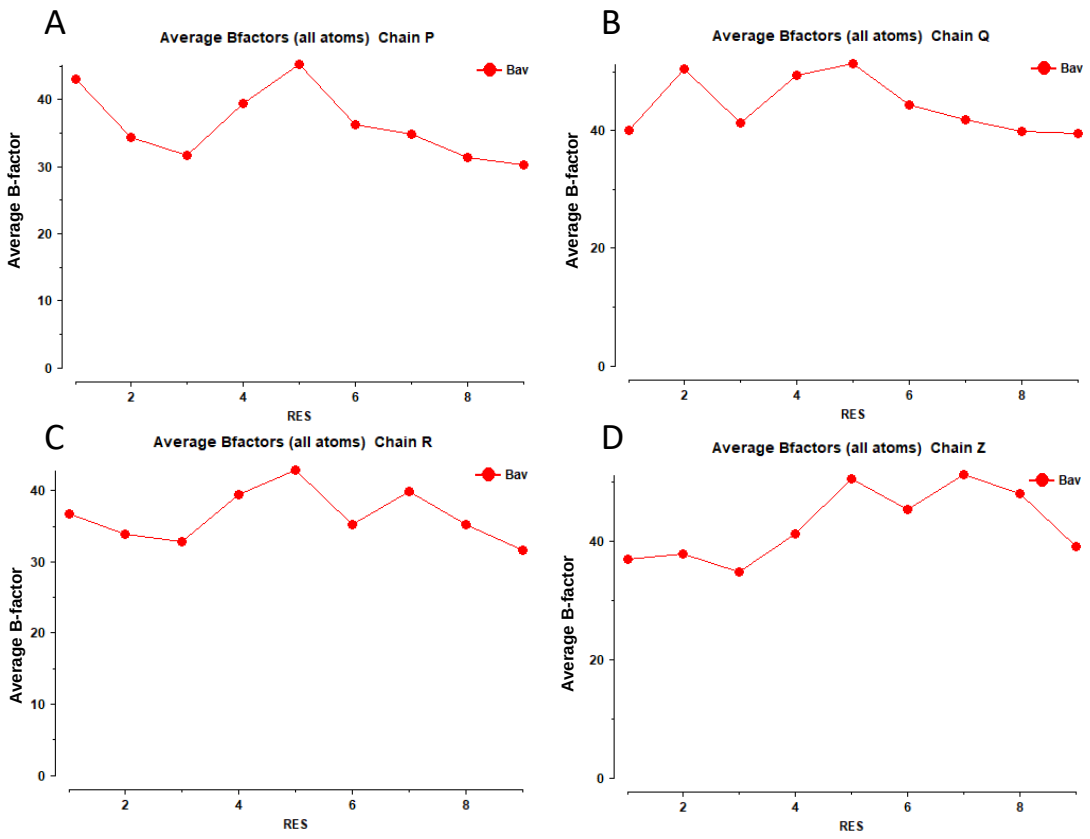


Figure 18: Comparative B factor analysis of Mtb14 peptides

All-atom average B factors (Å^2) for the Mtb14 peptide, Chain P, of molecule 1 (**A**), Chain Q of molecule 2 (**B**), Chain R of molecule 3 (**C**) and Chain Z of molecule 4 (**D**). The Y-axis denotes average B factor for main-chain and side chains atoms of each peptide residue, whereas peptide position is plotted on the X-axis.

Mtb14 peptide positioning within the HLA-E binding groove

Clear electron density was present in the peptide binding groove which enabled modelling of both the Mtb14 peptide main chain and side chains (Figure 19, A). Superposition of Mtb14 and VL9 peptides revealed similar positioning, which is perhaps unsurprising considering both position 2 Met and 9 Leu canonical anchor residues are present in the Mtb14 nonamer (Figure 19, B). However minor movements were measured between Mtb14 and VL9 Cα atoms, the greatest of which were found in the unconstrained kinked region between peptide positions 4 and 5 which deviated by 1.6 and 1.7 Å, respectively (Figure 19, C). Relative to VL9, Mtb14 is positioned higher in the groove, appears shifted horizontally toward the α1-helix and exhibits closer alignment to the Mtb44 peptide backbone. Such repositioning could be due to the loss of hydrogen bonds formed between position 5 of the VL9 peptide and the HLA-E heavy chain-derived α2-helix (Figure 20). Furthermore, the presence of a larger Gln side chain at position 7 of the Mtb14 peptide, which sits deeper in the secondary E pocket, could account for the elevated positioning of the position 7 Cα atom which is positioned 1.1 Å higher in the groove than its VL9 counterpart.

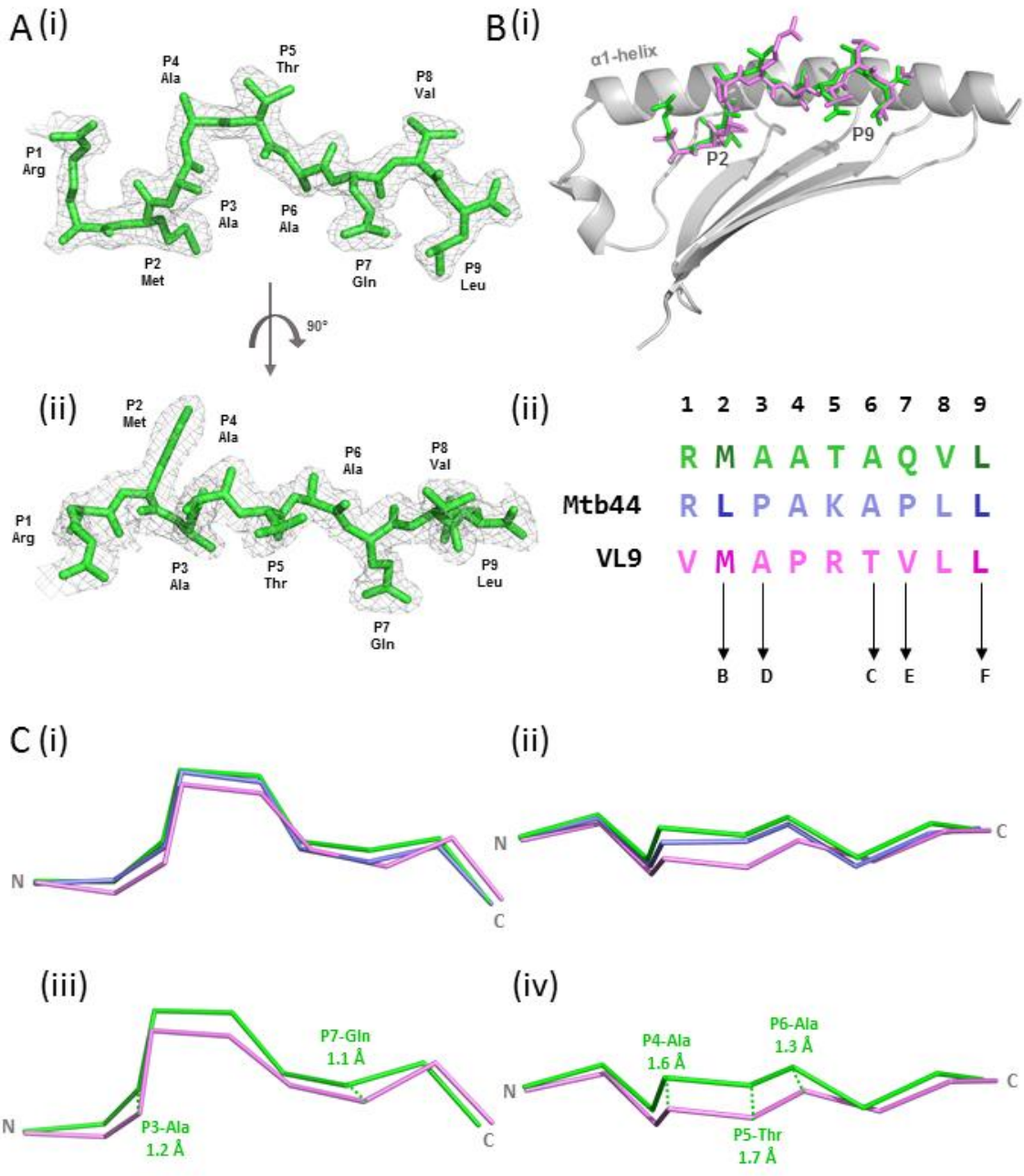
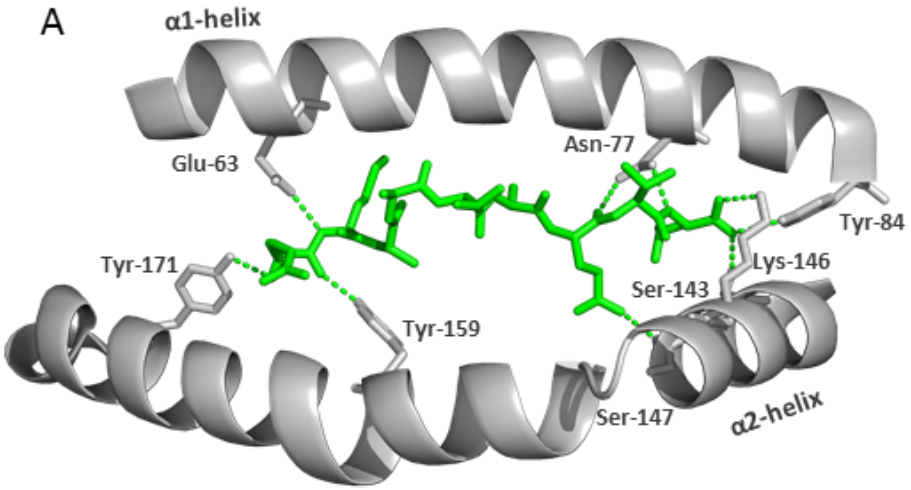


Figure 19: Analysis of Mtb14 peptide positioning

A. Mtb14 peptide shown in green stick-form, visualised from the side **(i)** and from above **(ii)** with superposed electron density map contoured to 1 sigma, in grey mesh. **B. (i)** The Mtb14 peptide (green) is shown in stick form, superposed to the VL9 peptide (violet) from molecule 1 of 1MHE. The α1-helix and β-sheet floor of the peptide binding groove are shown in grey cartoon with the α2 helix, α3 domain and β2M omitted for clarity. **(ii)** Sequences for the Mtb14, Mtb44 and VL9 peptides illustrated in C. **C.** Superposed backbones of the Mtb14 peptide (green) of molecule 1, Mtb44 peptide (purple) of molecule

1 and VL9 peptide (violet) of molecule 1 from 1MHE, side-on **(i)** and from above **(ii)**. Superposed Mtb14 (green) and VL9 (violet) peptide backbones with Cα distances indicated by green dashed lines, side on **(iii)** and from above **(iv)**. 1.2 Å separates the position 3 α-carbons, 1.6 Å the position 4 α-carbons, 1.7 Å the position 5 α-carbons, 1.3 Å the position 6 α-carbons and 1.1 the position 7 α-carbons. Only α-carbons deviating by over 1 Å are denoted.

A



B

HLA-E*0103 – Mtb14

	Peptide	Dist. Å	Heavy chain
1	P:ARG 1[N]	2.84	A:TYR 171[OH]
2	P:MET 2[N]	2.75	A:GLU 63[OE2]
3	P:GLN 7[NE2]	3.00	A:SER 147[OG]
4	P:LEU 9[N]	2.80	A:ASN 77[OD1]
5	P:ARG 1[O]	2.56	A:TYR 159[OH]
6	P:GLN 7[O]	3.16	A:ASN 77[ND2]
7	P:LEU 9[O]	2.73	A:TYR 84[OH]
8	P:LEU 9[O]	2.66	A:SER 143[OG]
9	P:LEU 9[OXT]	2.90	A:LYS 146[NZ]

Figure 20: Mtb14-HLA-E*0103 peptide-heavy chain hydrogen bonding network

A. Visualisation of the inter-chain hydrogen bonding network between the Mtb14 peptide and the HLA-E heavy chain. The Mtb14 peptide is shown in green stick-form with the heavy-chain-derived donor/acceptor residues depicted as grey sticks. Hydrogen bonds are shown as green dashed lines. The α-1 and 2 helices are visualised in grey with the remaining HLA-E heavy and light chain residues omitted

for clarity. **B.** Table detailing hydrogen bonds formed between the Mtb14 peptide and HLA-E heavy chain. The lower and upper thresholds for donor/acceptor distances are set to 2.5 Å and 3.5 Å, respectively.

As expected, the hydrophobic B pocket accommodated the canonical Met present at position 2 in Mtb14 with minimal side chain positioning relative to VL9 or Mtb44 HLA-E structures (Figure 21, A). Similarly, the presence of canonical Leu at position 9 of the Mtb14 peptide was accommodated by the F pocket with minimal repositioning of pocket forming residues relative to both VL9 and Mtb44 complexes (Figure 21, B).

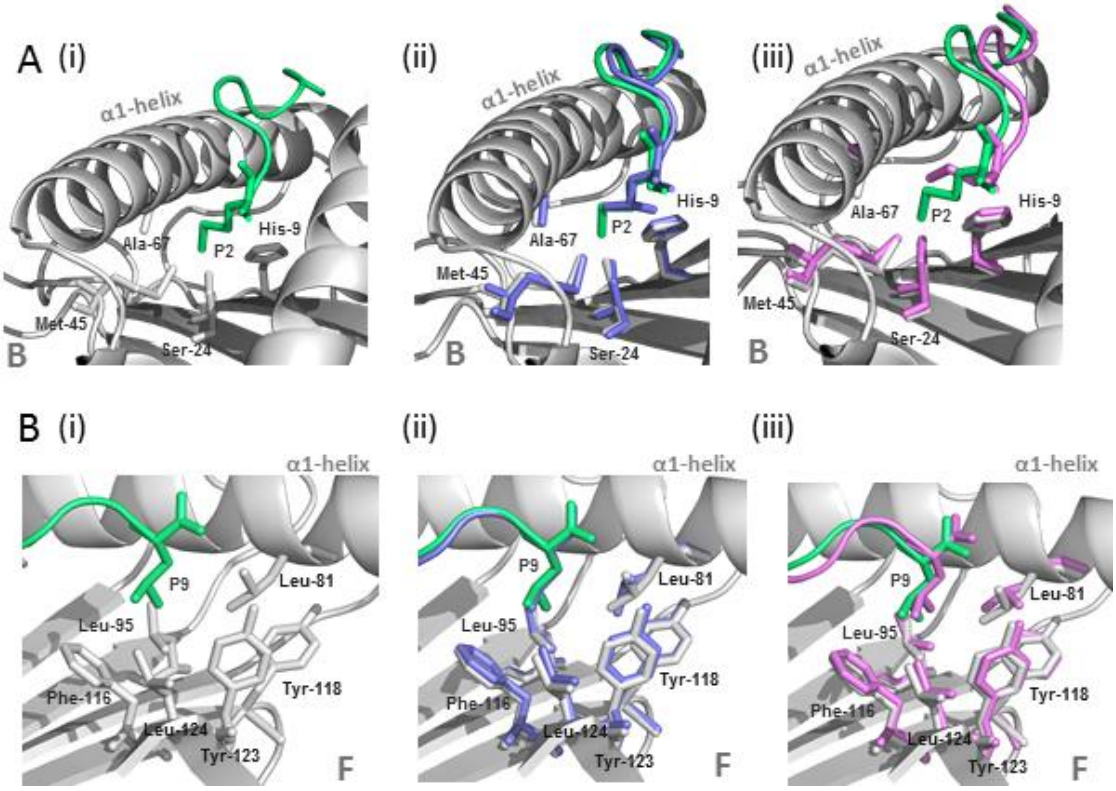


Figure 21: Primary B and F pocket occupancy by the Mtb14 peptide

A. (i) B pocket visualisation for Mtb14-bound HLA-E*0103 molecule 1. Heavy chain-derived B pocket-forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as grey cartoon. The Mtb14 peptide is shown in green cartoon with position 2 Met in green stick-form. **(ii)** B pocket visualisation of the superposed Mtb14-HLA-E molecule 1 and VL9-HLA-E from 1MHE. Colour-coding is according to A (i) with superposed B pocket-forming residues derived from the VL9-bound

HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 2 Met in stick-form. **B. (i)** Visualisation of the F pocket for Mtb14-bound HLA-E*0103 molecule 1. Heavy chain-derived F pocket forming-residues are depicted as grey sticks with the surrounding area of the peptide binding groove as grey cartoon. The Mtb14 peptide is shown in green cartoon with position 9 Leu in green stick-form. **(ii)** F pocket visualisation of the superposed Mtb14-HLA-E molecule 1 and VL9-HLA-E from 1MHE. Colour-coding is according to A (i) with superposed F pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 9 Leu in stick-form.

The canonical Ala present at position 3 of the Mtb14 peptide and also in the majority of VL9 variants, was accommodated by the shallow D pocket, showed strong alignment to VL9 position 3 and minimal rearrangement of pocket-forming residues (Figure 22, A). Despite the substitution of canonical Thr to a smaller Ala side chain at position 6 of the Mtb14 peptide, the C pocket forming residues exhibited minimal movement, supporting the theory of superior rigidity of the HLA-E binding groove despite the introduction of pocket cavities (Figure 22, B).

Typically Leu or Val is present at position 7 in VL9 peptides. However, in the Mtb14 peptide these hydrophobic residues are substituted for a polar Gln. The Gln side chain projects down into the deep hydrophobic E pocket and forms a hydrogen bond with the HLA-E heavy chain-derived Ser-147. Additionally, the position 7 Gln side chain forms a hydrogen bond with a water molecule which is present in the E pocket of all four molecules within the Mtb14-bound HLA-E*0103 asymmetric unit.

The E pocket-forming residues in the Mtb14 HLA-E structure align more closely with those of the Mtb44 rather than the VL9 structure (Figure 22, C). First, in the Mtb14 E pocket, the Trp-133 side chain tilts more acutely toward the groove floor, perhaps to accommodate for the larger Gln side chain which projects deeper into the E pocket than its Val counterpart in VL9. Second, the Glu-152 side chain adopts an alternative orientation projecting further toward the peptide C-terminus in the Mtb14 structure relative to VL9. Glu-152 repositioning could be

accounted for by differences in the inter-peptide-HLA-E-heavy-chain hydrogen bonding networks since it hydrogen bonds with position 5 Arg of the VL9 peptide but forms no detectable bonds with the Mtb14 peptide. Third, Ser-147 also adopts an alternative position in the Mtb14 E pocket relative to the VL9 structure, potentially due to the presence of a hydrogen bond between Ser-147 and position 7 of the Mtb14 peptide which is absent in the VL9 structure. In summary, E pocket-forming residues exhibit the greatest movement in the Mtb44 and Mtb14 structures relative to canonical VL9 and this is largely due to distinct peptide-heavy chain hydrogen bonding networks.

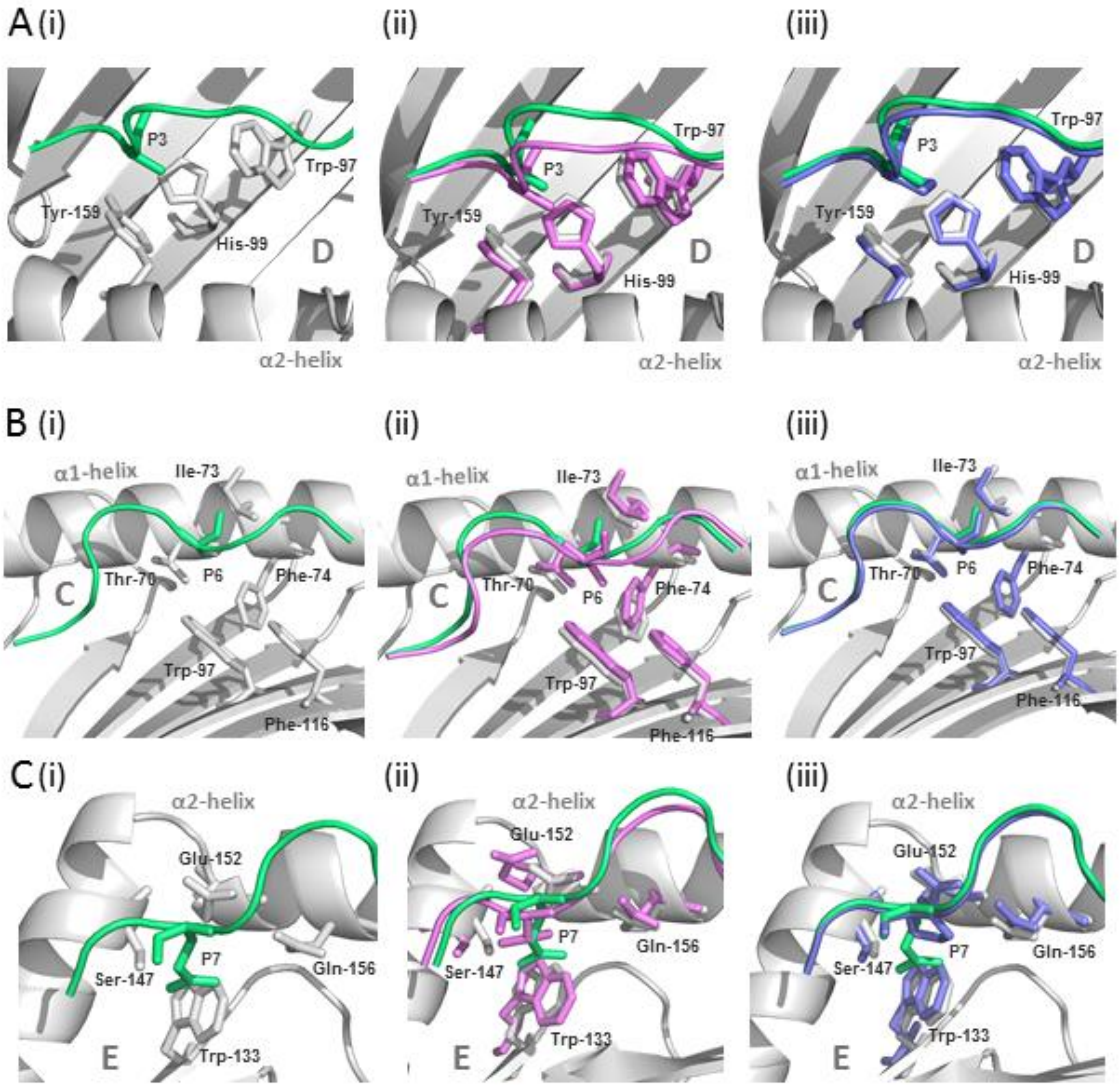


Figure 22: Secondary D, C and E pocket occupancy by the Mtb14 peptide

A. (i) D pocket visualisation for Mtb14-bound HLA-E*0103 molecule 1. Heavy chain-derived D pocket-forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as grey cartoon. The Mtb14 peptide is shown in green cartoon with position 3 Ala in green stick-form. **(ii)** D pocket visualisation of the superposed Mtb14-HLA-E molecule 1 and VL9-HLA-E from 1MHE. Colour-coding is according to A (i) with superposed D pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 3 Ala in stick-form. **B. (i)** Visualisation of the C pocket for Mtb14-bound HLA-E*0103 molecule 1. Heavy chain-derived C pocket forming-residues are depicted as grey sticks with the surrounding area of the peptide binding groove as grey cartoon. The Mtb14 peptide is shown in green cartoon with position 6 Ala in green stick-form. **(ii)** C pocket visualisation of the superposed Mtb14-HLA-E molecule 1 and VL9-HLA-E from 1MHE. Colour-coding is according to B (i) with superposed C pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 6 Thr in stick-form. **C. (i)** E pocket visualisation for Mtb14-bound HLA-E*0103 molecule 1. Heavy chain-derived E pocket-forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as grey cartoon. The Mtb14 peptide is shown in green cartoon with position 7 Gln in green stick-form. **(ii)** E pocket visualisation of the superposed Mtb14-HLA-E molecule 1 and molecule 1 of 1MHE. Colour-coding is according to C (i) with superposed E pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 7 Val in stick-form.

Combined structural and mutagenesis analyses of HLA-E*0103 in complex with Mtb44 position 2 and 9 variants

As previously discussed, it was originally concluded that the HLA-E peptide binding repertoire was restricted to VL9 peptides, in part due to the highly conserved nature of MHC class I leader sequence-derived variants, with the majority possessing Met at position 2 and Leu at position 9. Structural determination of VL9-bound HLA-E demonstrated that these hydrophobic anchor residues fit optimally within their corresponding B and F pockets. Further,

when canonical Met is substituted to Thr at position 2 of the HLA-B27-derived leader peptide, the lack of B pocket-forming residue rearrangement produces an unoccupied cavity - a feature that has previously been linked to loss of protein stability (Strong et al., 2003) (Xu et al., 1998). In summary, such observations support the theory that the HLA-E peptide binding groove is highly selective for VL9 peptides containing canonical primary anchor residues.

We therefore performed combined mutagenesis and structural studies to re-evaluate the primary anchor specificity of HLA-E and investigate whether the B and F pockets are capable of respectively, accommodating side chains other than the canonical Met and Leu. We synthesised a collection of VL9 and Mtb44 peptide variants in which a range of residues were substituted at positions 2 and 9. Such variants were initially screened in the optimised peptide binding assay and a subset were subsequently pursued in crystallographic studies.

Screening of Mtb44 and VL9 primary anchor variants in the peptide-exchange binding assay

A panel of Mtb44 peptide variants with a range of non-canonical primary anchor residues at positions 2 and 9 (Figure 23, A) were screened for binding to HLA-E in the optimised peptide-exchange-based binding assay (Results Chapter 1, Figures 6-15). This panel included Mtb44 variants with positively charged, negatively charged and polar side chains substituted in place of the standard hydrophobic side chains present at position 2 and 9 in VL9 peptides. As expected, the majority of charged residues including the negatively charged Glu at positions 2 or 9 and the positively charged Lys at position 9 reduced binding of Mtb44 peptide variants to HLA-E (Figure 23, B). However, Lys at position 2 yielded significant HLA-E binding, albeit producing a lower signal relative to the wild-type Mtb44 peptide.

In addition to wild-type Mtb44, 3 Mtb44 variants demonstrated HLA-E binding with the greatest degree of statistical significance in the peptide-exchange assay, including those with position 2 and 9 Phe and position 2 Gln substitutions. Given the prerequisite for stably

refolded material, which in the case of MHC proteins is afforded by high affinity peptides, crystallisation trials were set up for these three Mtb44 variants to elucidate at the atomic level, how such non-canonical primary anchor residues are accommodated in the HLA-E peptide binding groove.

A

Peptide ID	Sequence	Aas	Organism	Protein	Ref
Mtb44	RLPAKAPLL	9	*Mycobacteria*	Rv1484	Joosten, S.A, et al. (2010)
FP2	RFPAKAPLL				
FP9	RLPAKAPLF				
QP2	RQPAKAPLL				
KP2	RKPAKAPLL				
KP9	RLPAKAPLK				
EP2	REPAKAPLL				
EP9	RLPAKAPLE				

B

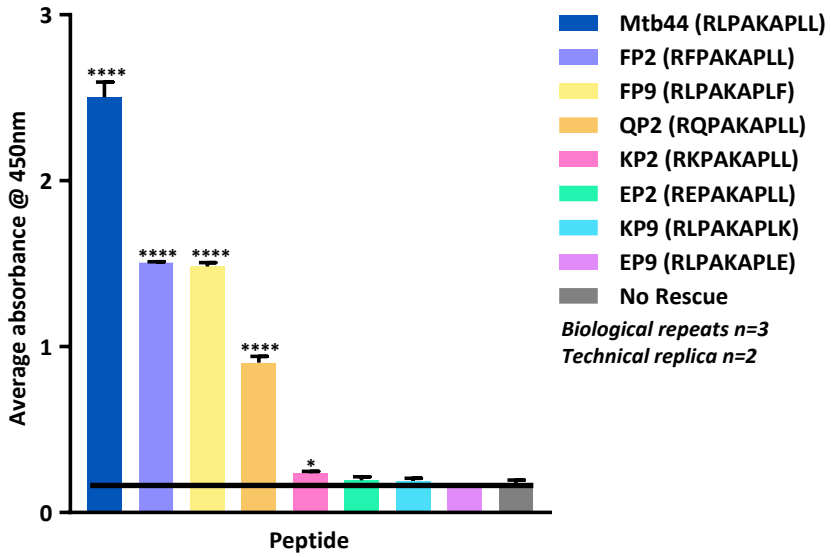


Figure 23: Peptide binding assay screening of Mtb44 position 2 and 9 variants

A. Table detailing the Mtb44 position 2 and 9 variant peptides screened in the optimised peptide-exchange-based binding assay. Peptide ID, sequence, length and origin are reported with the corresponding reference. **B.** Average absorbance readings from the peptide-exchange sandwich-ELISA-based binding assay are denoted on the Y-axis for Mtb44 primary anchor variant peptides. The positive control wild-type Mtb44 peptide (RLPAKAPLL) is shown in dark blue with the peptide-free background (No rescue) in grey.

Error bars depict standard error of the mean. Absorbance readings taken at 450nm. Biological repeats

*n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

VL9 peptide variants with substitutions at position 2 were also assembled and screened in the optimised peptide-exchange binding assay. First, a small panel based on the HLA-B7-derived leader peptide, VMAPRTVLL, was tested (Figure 24, A). Three of the four peptides in this panel, including a positively charged Lys and a polar Gln, demonstrated HLA-E binding, and generated average absorbance readings with the greatest degree of statistical significance (Figure 24, B).

Interestingly, the hierarchy of relative binding of position 2 anchor residue mutations differed on the Mtb44 and VL9 peptide panel backgrounds, indicating other positions along these nonamers may operate synergistically with position 2 substitutions to determine overall peptide binding affinity. For example, position 2 Lys generated very weak ELISA signals when substituted into the Mtb44 peptide backbone but yielded stronger binding when inserted into the VL9 nonamer. Furthermore, Gln at position 2 of the VL9 peptide yielded the strongest binding to HLA-E, compared to the position 2 Phe VL9 peptide variant. However, this trend was reversed for Mtb44 with the position 2 Phe variant demonstrating the strongest binding to HLA-E, above that of the position 2 Gln variant. Together, this suggests pocket specificity and resultant binding affinity is difficult to predict based solely on anchor residues and in isolation of the remaining peptide sequence.

A

Peptide ID	Sequence	Aas	Organism	Protein	Ref
VL9	VMAPRTVLL	9	Human	HLA-B7/8/14/39/42/48	O'Callaghan et al. (1998)
QP2	VQAPRTVLL				
FP2	VFAPRTVLL				
KP2	VKAPRTVLL				
EP2	VEAPRTVLL				

B

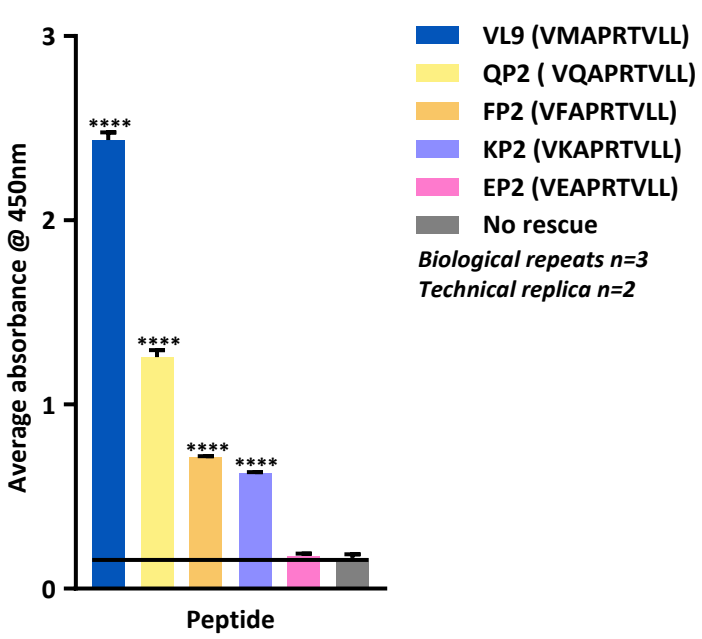


Figure 24: Peptide binding assay screening of VL9 (VMAPRTVLL) position 2 variants

A. Table detailing a panel of HLA-B7/8/14/39/42/48 leader sequence-derived VL9 peptide position 2 variants. Peptide ID, sequence, length and origin are reported with the corresponding reference. **B.** Average absorbance readings from the peptide-exchange sandwich-ELISA-based binding assay are denoted on the Y-axis for the VL9 position 2 variant peptide panel. The positive control VL9 peptide (VMAPRTVLL) is shown in dark blue with the peptide-free background (No rescue) in grey.

*Error bars depict standard error of the mean. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

A second panel of position 2 anchor residue variants based on an alternative VL9 peptide (VMAPRTLLL) derived from the HLA-A1 leader sequence was screened in the peptide exchange assay (Figure 25, A). Interestingly all 7 peptides in this panel demonstrated significant binding to HLA-E, including Glu at position 2 which did not yield significant binding when substituted into the HLA-E B7 leader sequence-derived VL9 peptide backbone (Figure 25, B). However, a polar position 2 Gln yielded the strongest binding signals for both the HLA-B7 and HLA-A1 leader peptides demonstrating a degree of consistency in hierarchies of relative binding affinity between VL9 peptide variants.

A

Peptide ID	Sequence	Aas	Organism	Protein	Ref
VL9	VMAPRTLLL	9	Human	HLA-A1	Strong, R,K. et al. (2003)
QP2	VQAPRTLLL				
IP2	VIAPRTLLL				
TP2	VTAPRTLLL				
WP2	VWAPRTLLL				
YP2	VYAPRTLLL				
EP2	VEAPRTLLL				
PP2	VPAPRTLLL				

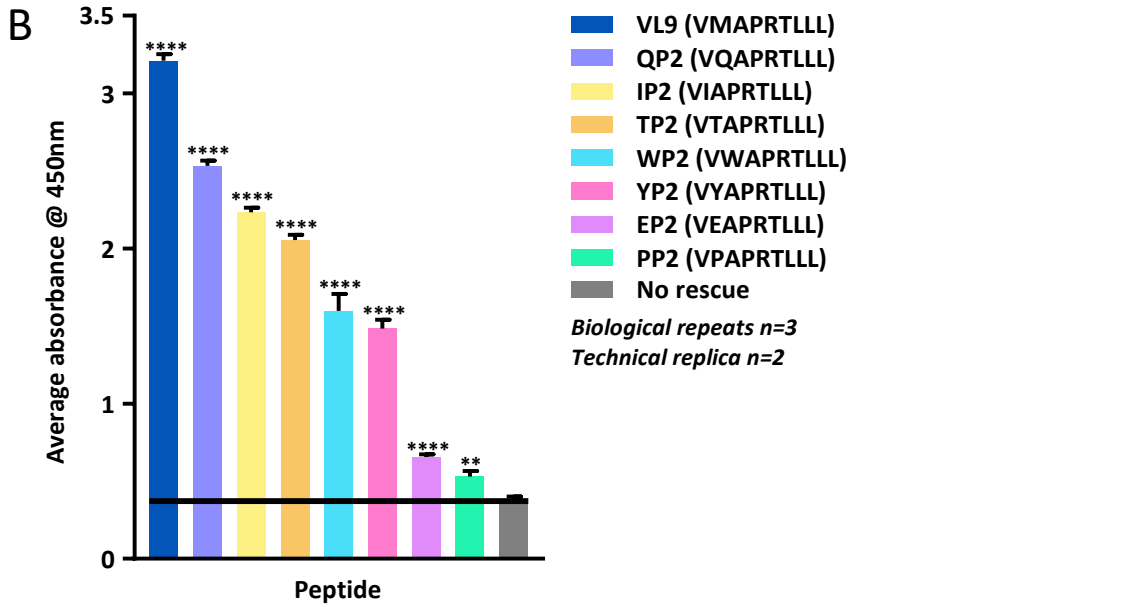


Figure 25: Peptide binding assay screening of VL9 (VMAPRTLLL) position 2 variants

A. Table detailing a panel of HLA-A1 leader sequence-derived VL9 peptide position 2 variants. Peptide ID, sequence, length and origin are reported with the corresponding reference. **B.** Average absorbance readings from the peptide-exchange sandwich-ELISA-based binding assay are denoted on the Y-axis for

the VL9 position 2 variant peptide panel. The positive control VL9 peptide (VMAPRTLLL) is shown in dark blue with the peptide free background (No rescue) in grey.

*Error bars depict standard error of the mean. Absorbance readings taken at 450nm. Biological repeats n=3, technical replica n=2. Stars above error bars reflect degree of significance in peptide binding by nonparametric t-tests: no star=P>0.05, *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****P=≤0.0001.*

Structural determination of HLA-E*0103 in complex with Mtb44 position 2 and 9 variants

*HLA-E*0103-Mtb44*P2-Phe*

HLA-E*0103 was crystallised in complex with the Mtb44 peptide variant in which position 2 Leu was substituted for a larger, hydrophobic Phe (Mtb44*P2-Phe). Crystallisation and refinement statistics are detailed in Figure 26. In summary, HLA-E*0103-Mtb44*P2-Phe crystals grew in 2.0 M Ammonium sulphate, 0.1 M MES at pH 6.5. 1mM Z_nSO_4 was also present in the crystallisation buffer and Zinc ions were later detected associated to molecules within the asymmetric unit. Diffraction data were collected from a number of isomorphous crystals that packed in the P1 space group yielding a 95.9% complete dataset to 2.7 Å following merging and integration by Xia2. Refinement and validation in Phenix resulted in low R_{work}/R_{free} values of 0.180 and 0.221, respectively.

*HLA-E*0103-Mtb44*P2-Gln*

HLA-E*0103 was crystallised in complex with the Mtb44 peptide variant in which the hydrophobic position 2 Leu was substituted for a polar Gln (Mtb44*P2-Gln). Crystallisation and refinement statistics are detailed in Figure 26. In summary, optimal crystallisation conditions for HLA-E*0103-Mtb44*P2-Gln constituted 2.4 M Ammonium sulphate and 0.1 M HEPES at pH 7. Similar to the wild-type Mtb44-bound HLA-E*0103 structure, crystals diffracted to 2.1 Å, packed in the P1 space group and 4 molecules were present in an asymmetric unit comprising 44% solvent. Following refinement and validation of this 98% complete dataset, the R_{work}

(0.202) and R_{free} (0.242) values fell within the accepted range for datasets collected to around 2 Å.

*HLA-E*0103-Mtb44*P9-Phe*

HLA-E*0103 was crystallised in complex with the Mtb44 peptide variant containing Phe at position 9 in place of the wild-type Leu. Crystallisation and refinement statistics are detailed in Figure 26. In summary, HLA-E*0103-Mtb44*P9-Phe crystals grew in 2.4 M Ammonium sulphate, 0.1 M HEPES at pH 7 and packed in the P1 space group. Considering the minimal symmetry of the P1 space group, reflections were collected over 180° to maximise completeness of the dataset. However, x-ray radiation damage over this large angular range compromised the completeness of datasets obtained from a single crystal. Therefore, a 97.7% complete dataset to 2.5 Å was obtained via merging diffraction data from multiple isomorphous crystals in Xia2. Two molecules were present in the asymmetric unit which had an estimated solvent content of 51%. Molecule 1 from the 6GH1 structure (HLA-E*0103 in complex with wild-type Mtb44) was implemented as a search model in molecular replacement. Following refinement and geometric validation, R factors within the acceptable range for the resolution limits were obtained – 0.218 and 0.264 for R_{work} and R_{free}, respectively.

	HlA-E*0103 in complex with peptide:		
Data collection	**Mtb44*P2-Gln**	**Mtb44*P2-Phe**	**Mtb44*P9-Phe**
Crystallisation condition	**2.4 M AS 0.1 M HEPES pH 7**	**2.0 M AS 0.1 M MES pH 6.5 1mM Z_nSO_4**	**2.4 M AS 0.1 M HEPES pH 7**
Resolution Å	**67.47 - 2.16 (2.20 - 2.16)**	**53.18 - 2.73 (2.78 - 2.73)**	**58.50 - 2.54 (2.59 - 2.54)**
Space Group	**P 1**	**P 1**	**P 1**
Cell dimensions	**a = 48.31 b = 76.41 c = 111.31 α = 99.39° β = 99.83° γ = 93.50°**	**a = 60.93 b = 66.04 c = 72.10 α = 102.30° β = 101.22° γ = 109.70°**	**a = 37.45 b = 64.68 c = 97.31 α = 107.03° β = 93.89° γ = 106.85°**
Solvent content [%] (molecules per AU)	**44 (4)**	**57 (2)**	**51 (2)**
Unique reflections	**81277 (3904)**	**25095 (882)**	**26488 (1053)**
Completeness [%]	**98.0 (92.9)**	**95.9 (67.6)**	**97.7 (77.5)**
R_{merge} [I]	**0.18 (0.85)**	**0.23 (0.49)**	**0.21 (0.78)**
I/sigma	**4.3 (1.2)**	**6.3 (2.6)**	**5.2 (1.0)**
Multiplicity	**3.3 (2.6)**	**6.1 (2.4)**	**4.9 (2.3)**
$CC_{1/2}$	**0.98 (0.50)**	**0.96 (0.63)**	**0.98 (0.57)**
Refinement			
No. of non-hydrogen atoms	**13288**	**6514**	**6423**
R_{factor} [%)]	**0.210 (0.322)**	**0.180 (0.265)**	**0.223 (0.322)**
R_{free} [%]	**0.239 (0.390)**	**0.221 (0.303)**	**0.266 (0.392)**
r.m.s.d. bonds [Å] §	**0.002**	**0.010**	**0.002**
r.m.s.d. angles [deg]	**0.544**	**0.692**	**0.545**
Ramachandran statistics			
Favoured [%]	**98.2**	**98.5**	**97.8**
Disallowed [%]	**0**	**0**	**0**

Figure 26: Crystallographic determination, refinement and validation of Mtb44 variant-bound HLA-E*0103

Crystallographic data collection, structural determination, refinement and validation statistics for the 3 HLA-E*0103 structures solved in complex with Mtb44 primary anchor variant peptides with position 2

Phe and Gln and position 9 Phe.

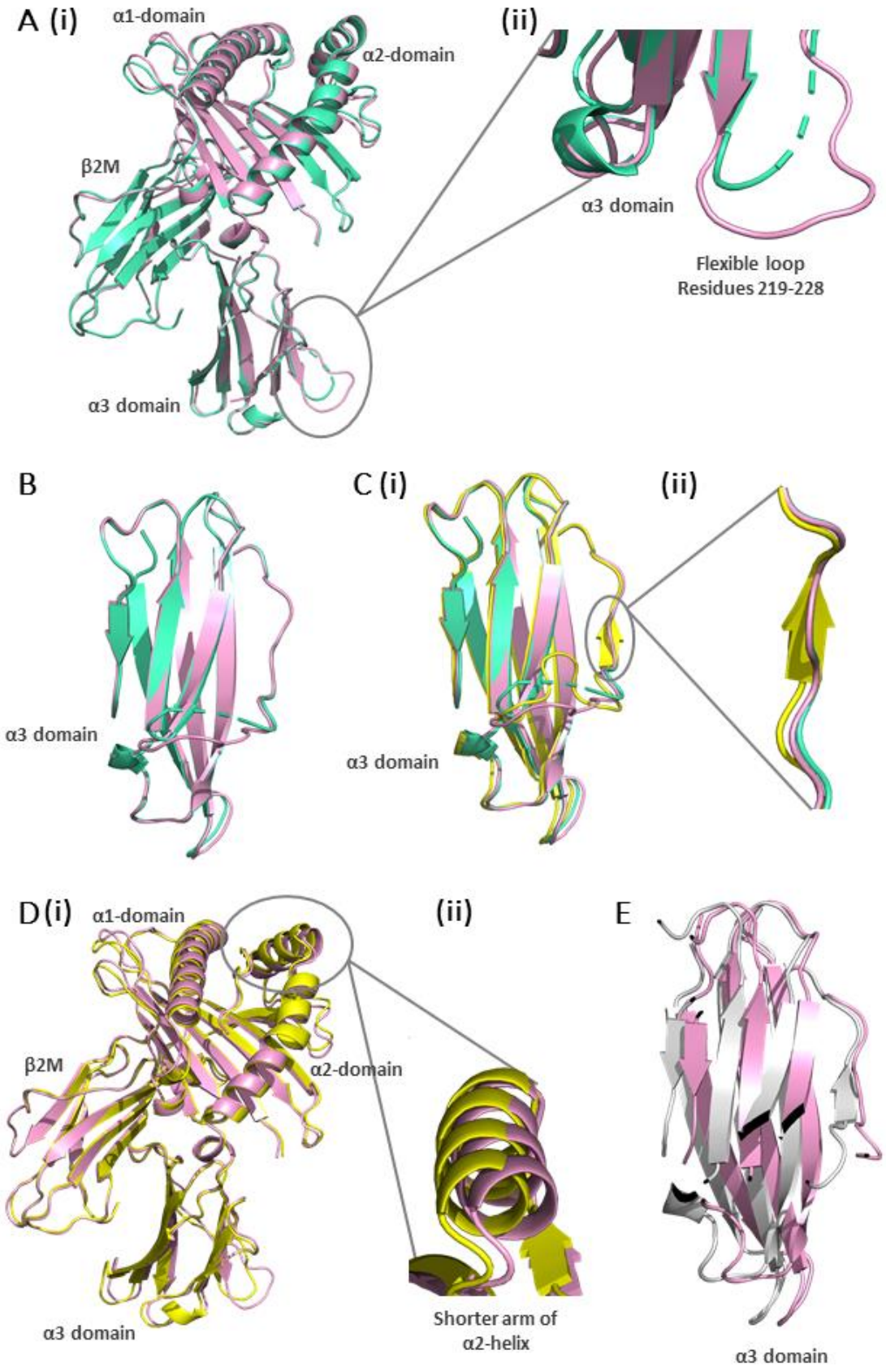
AS: Ammonium sulphate. AU: asymmetric unit. § r.m.s.d.: root mean square deviation from ideal geometry. Statistics in parentheses indicate those of the outer shell. R_{free}: R factor representing 5% of reflections omitted from refinement.

HLA-E*0103-Mtb44*P2-Phe structural interpretation

Sufficient electron density enabled modelling of a complete heavy and light chain for molecule 1 of the Mtb44*P2-Phe structure. However, disorder prohibited modelling a complete HLA-E heavy chain in molecule 2, resulting in residues 222-225 situated in the 220's loop, being omitted from the deposited coordinate file. Superposition of molecules in the asymmetric unit revealed strong alignment of secondary structure elements in both the heavy and light chains of the Mtb44*P2-Phe structure (Figure 27, A (i)). Similar to wild-type Mtb44-bound HLA-E*0103, the greatest positional difference was located in the 220's loop of the heavy chain-derived α3 domain (Figure 27, A (ii)).

Alignment of HLA-E*0103-Mtb44*P2-Phe molecules to molecule 1 of wild-type HLA-E*0103-Mtb44 revealed the greatest inter-structure deviation lay within the α3 domain and short arm of the α2 helix (Figure 27, B-E). For example, Ala-150 Cα atoms of the α2 helix short arm were separated by 1.7 Å in molecule 1 of HLA-E*0103-Mtb44*P2-Phe and molecule 1 of HLA-E*0103-Mtb44. Minor shifts were also observed between β-strands of the α3 domain in the HLA-E*0103-Mtb44*P2-Phe structure relative to wild-type HLA-E*0103-Mtb44. However, such positional differences did not exceed 1 Å and were substantially smaller than the intra-structure distances separating β-strands in the α3 domain between molecules of the wild-type HLA-E*0103-Mtb44 structure. Overall, HLA-E*0103-Mtb44*P2-Phe-derived α3 domains aligned more optimally with the α3 domain of HLA-E*0103-Mtb44 molecule 1 as opposed to the shifted α3 domain of Mtb44-HLA-E*0103 molecule 3.

In both HLA-E*0103-Mtb44*P2-Phe molecules, the short, α3 domain-located β-strand between positions 229-231, present in all 4 molecules of wild-type HLA-E*0103-Mtb44, is lost (Figure 27, C). Instead, no secondary structure is detected for these residues in the HLA-E*0103-Mtb44*P2-Phe structure and the flexible 220's loop is extended. This is consistent with the highest density of elevated B factors aggregating within the α3 domain of both HLA-E*0103-Mtb44*P2-Phe molecules (Figure 27, F and G).


A (i)
α1-domain
α2-domain
β2M
α3 domain
(ii)
α3 domain
Flexible loop
Residues 219-228
B
α3 domain
C (i)
(ii)
α3 domain
D (i)
α1-domain
β2M
α2-domain
α3 domain
(ii)
Shorter arm of
α2-helix
E
α3 domain

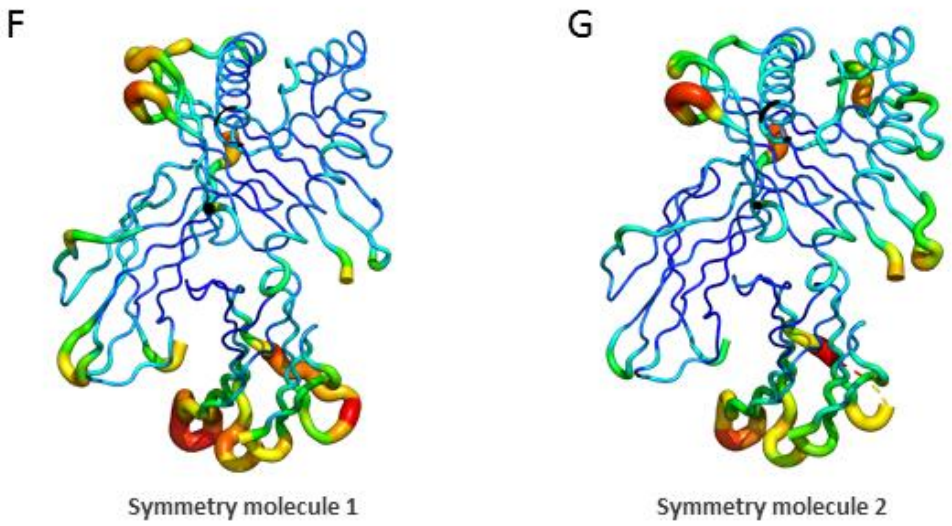


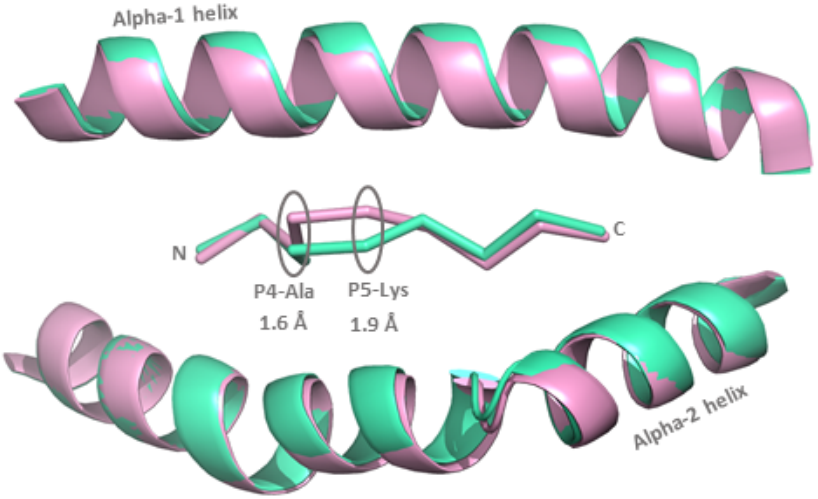
Figure 27: Analysis of molecules from the HLA-E*0103-Mtb44*P2-Phe structure

A. (i) Cartoon visualisation of the 2 superposed molecules present in the asymmetric unit of the Mtb44 position 2 Phe-bound HLA-E*0103 structure. Molecule 1 is depicted in light pink and molecule 2 in green-cyan. The flexible 220's loop of the α3-domain is circled. **(ii)** Magnification of the flexible 220's loop, colour-coded according to A. **B.** α3 domain of 2 superposed HLA-E*0103-Mtb44*P2-Phe molecules, colour-coded according to A. **C. (i)** Superposed α3 domains of HLA-E*0103-Mtb44*P2-Phe molecules colour-coded according to A and molecule 1 of the wild-type Mtb44 structure (yellow). **(ii)** Magnifications of the α3 domain-derived β-strand adjacent to the 220's loop. Molecules shaded according to C (i). **D. (i)** Superposition of Mtb44*P2-Phe molecule 1 (light pink) and molecule 1 from the wild-type Mtb44 structure (yellow). **(ii)** Magnification of the superposed α2 helix short arms from molecules 1 of the Mtb44*P2-Phe and wild-type Mtb44 structures, colour-coded according to D (i). **E.** Superposed α3 domains from molecule 1 of the Mtb44*P2-Phe structure (light pink) and molecule 3 of the original Mtb44 structure (grey). **F & G.** B factor putty visualisation of HLA-E*0103-Mtb44*P2-Phe molecules 1 and 2, respectively, colour-coded according to A. Cα backbone B factors are depicted on a cool to warm colour scale with the lowest relative B factors in blue and the highest in red.

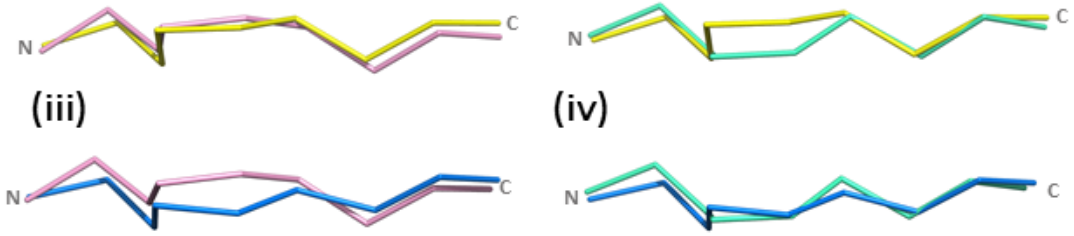
Mtb44*P2-Phe peptides modelled in each molecule exhibited a distinct orientation between positions 4 and 5 - inter-molecule Cα atoms of position 4 Ala were separated by 1.6 Å, and those of position 5 Lys by 1.9 Å (Figure 28, A). The central kinked region of the peptide, in which positions 4 and 5 are located, leans further towards the α1 helix in molecule 1 and aligns more optimally to wild-type Mtb44 peptide rather than VL9 peptide (Figure 28, B). By contrast, the Mtb44*P2-Phe peptide in molecule 2 leans further towards the α2 helix and exhibits stronger conformational alignment to the VL9 peptide of 1MHE, rather than wild-type Mtb44 peptide.

It is likely that the Mtb44*P2-Phe peptide from molecule 2 aligns more closely with the VL9 peptide of 1MHE due to conserved hydrogen bonding connecting position 5 of these two peptides to the α2 helix (Figure 28, C-E). Specifically, 2 hydrogen bonds connect position 5 Lys of the Mtb44*P2-Phe peptide in molecule 2 and position 5 Arg of the VL9 peptide to residues Glu-152 and Gln-156 of the α2-helix. However, such hydrogen bonds are not formed in wild-type Mtb44 peptides or in the Mtb44*P2-Phe peptide from molecule 1, and these peptides are instead positioned more closely to the α1 helix.

A



B (i) (ii)



(iii) (iv)

C (i) Hydrogen Bonds

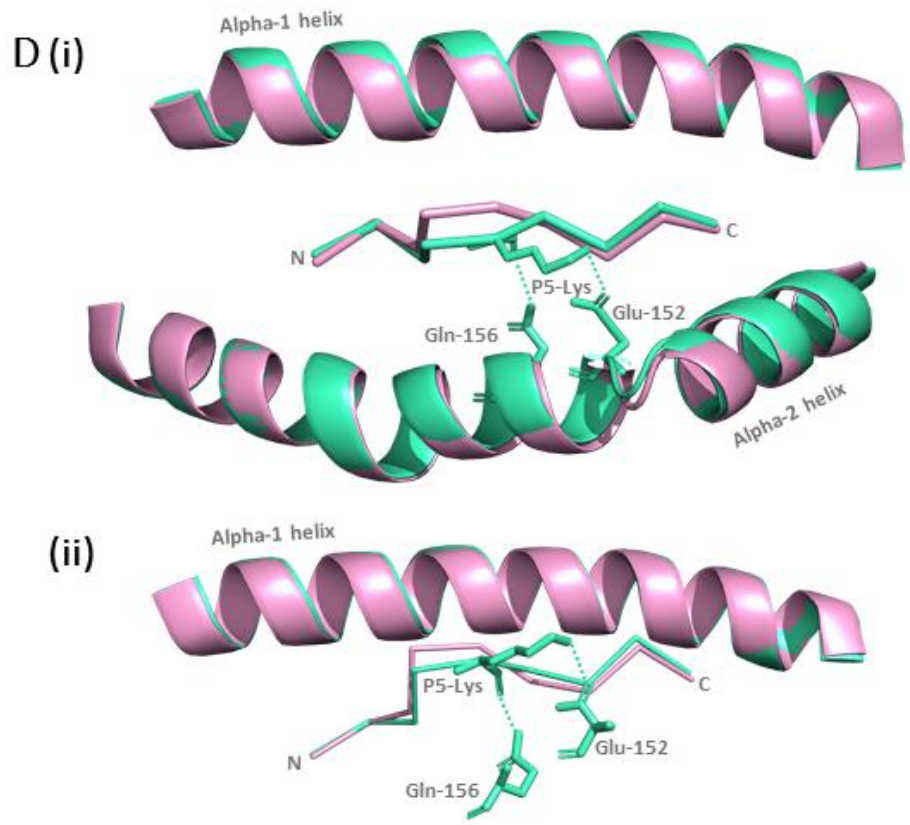
Symmetry molecule 1

	Chain P: Peptide	Distance (Å)	Chain A: Heavy chain
1	P:ARG 1[N]	3.33	A:TYR 171[OH]
2	P:ARG 1[N]	2.96	A:TYR 7[OH]
3	P:ARG 1[NE]	3.15	A:GLU 63[OE1]
4	P:PHE 2[N]	2.78	A:GLU 63[OE2]
5	P:PHE 2[N]	3.36	A:TYR 7[OH]
6	P:LEU 9[N]	3.07	A:ASN 77[OD1]
7	P:ARG 1[O]	2.67	A:TYR 159[OH]
8	P:PRO 7[O]	3.13	A:ASN 77[ND2]
9	P:LEU 9[O]	3.04	A:TYR 84[OH]
10	P:LEU 9[OXT]	3.56	A:LYS 146[NZ]

(ii) Hydrogen Bonds

Symmetry molecule 2

	Chain Q: Peptide	Distance (Å)	Chain C: Heavy chain
1	Q:ARG 1[N]	2.75	C:TYR 171[OH]
2	Q:ARG 1[N]	2.54	C:TYR 7[OH]
3	Q:ARG 1[NE]	3.30	C:GLU 63[OE1]
4	Q:PHE 2[N]	2.81	C:GLU 63[OE2]
5	Q:LYS 5[NZ]	2.79	C:GLU 152[OE1]
6	Q:LEU 9[N]	3.08	C:ASN 77[OD1]
7	Q:ARG 1[O]	2.88	C:TYR 159[OH]
8	Q:PHE 2[O]	3.27	C:ARG 62[NH1]
9	Q:LYS 5[O]	2.91	C:GLN 156[NE2]
10	Q:PRO 7[O]	3.21	C:ASN 77[ND2]
11	Q:LEU 9[O]	3.11	C:TYR 84[OH]
12	Q:LEU 9[OXT]	2.75	C:LYS 146[NZ]



E

Hydrogen Bonds
1MHE (HLA-E*0101-VMAPRTVLL)
Symmetry molecule 1

	Chain P: Peptide	Distance (Å)	Chain A: Heavy chain
1	P:VAL 1[N]	3.03	A:TYR 171[OH]
2	P:MET 2[N]	3.03	A:GLU 63[OE2]
3	P:MET 2[N]	3.20	A:TYR 7[OH]
4	P:ARG 5[NH1]	2.70	A:GLU 152[OE1]
5	P:LEU 9[N]	2.86	A:ASN 77[OD1]
6	P:VAL 1[O]	2.71	A:TYR 159[OH]
7	P:ARG 5[O]	2.67	A:GLN 156[NE2]
8	P:VAL 7[O]	2.87	A:ASN 77[ND2]
9	P:LEU 9[O]	2.77	A:TYR 84[OH]
10	P:LEU 9[O]	2.81	A:SER 143[OG]

Figure 28: Analysis of Mtb44*P2-Phe peptide positioning and hydrogen bonding

A. Cartoon of the peptide binding groove of superposed molecules 1 (light-pink) and 2 (green-cyan) from the HLA-E*0103-Mtb44*P2-Phe structure. The β-sheet floor is omitted for clarity and the superposed peptides are displayed in ribbon with the N and C terminals labelled in grey. Peptide Cα atoms at positions 4 and 5 are circled in grey with inter-molecule connecting distances denoted in Å. **B.**

The peptides from molecule 1 (light-pink) **(i)** and molecule 2 (green-cyan) **(ii)** of the HLA-E*0103-Mtb44*P2-Phe structure are superposed to the peptide from molecule 1 of the wild type Mtb44 structure. The peptides from molecule 1 (light-pink) **(iii)** or molecule 2 (green-cyan) **(iv)** of the HLA-E*0103-Mtb44*P2-Phe structure superposed to the peptide from molecule 1 of 1MHE (VL9-bound HLA-E*0101). Superposed peptide backbones are depicted in ribbon with the N and C-terminals labelled in grey. **C.** Table of inter-chain hydrogen bonds between the peptide and HLA-E heavy chain for molecule 1 **(i)** and molecule 2 **(ii)** of the HLA-E*0103-Mtb44*P2-Phe structure, generated by the PDBePISA server. The lower and upper thresholds for donor/acceptor distances are set to 2.5 Å and 3.5 Å, respectively. The 2 hydrogen bonds unique to the Mtb44*P2-Phe peptide of molecule 2 are highlighted in blue. **D.** Visualisation from above **(i)** and side-on **(ii)** of superposed molecules 1 (light-pink) and 2 (green-cyan) from the HLA-E*0103-Mtb44*P2-Phe structure in which the peptide binding groove is displayed in cartoon with the β-sheet floor omitted for clarity. Superposed peptide backbones are displayed in ribbon-form with the N and C terminals labelled in grey. Position 5 Lys of the Mtb44*P2-Phe peptide from molecule 2 is displayed in stick-form along with HLA-E heavy chain-derived residues Glu-152 and Gln 156. Inter-peptide-heavy-chain hydrogen bonds are depicted as dashed green lines. **E.** Inter-chain hydrogen bonding network between the VL9 peptide (VMAPRTVLL) (chain P) and HLA-E*0101 heavy chain (chain A) of molecule 1 from the 1MHE structure, for reference. The two hydrogen bonds involving position 5 of the VL9 peptide that are conserved with the Mtb44*P2-Phe peptide from molecule 2 are highlighted in blue. The lower and upper thresholds for donor/acceptor distances are set to 2.5 Å and 3.5 Å, respectively.

The 2 Mtb44*P2-Phe peptides show distinct B factor profiles with the peptide from molecule 2 exhibiting the highest B factors within the central kinked region that hydrogen bonds with the α2 helix (Figure 29). However, this peptide region does not comprise the highest relative B factors for the Mtb44*P2-Phe peptide from molecule 1. Overall, the Mtb44*P2-Phe peptide from molecule 1 exhibits lower B factors than that of molecule 2.

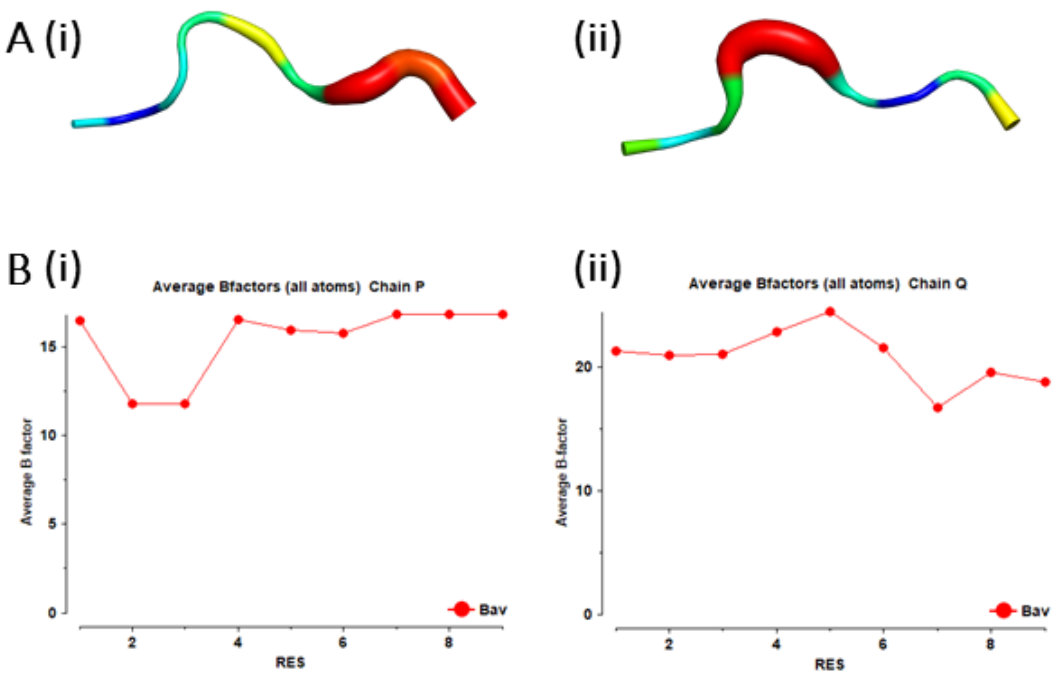


Figure 29: B factor analysis of Mtb44*P2-Phe peptides

A. B factor putty representation of the Mtb44*P2-Phe peptide from molecule 1 **(i)** and molecule 2 **(ii)**. Cα atom B factors are depicted on a cool to warm colour scale with the lowest relative B factors in blue and the highest in red. **B.** All-atom average B factors ($Å^2$) for the Mtb44*P2-Phe peptide, Chain P, of molecule 1 **(i)** and chain Q, of molecule 2 **(ii).** The Y-axis denotes average B factors for main chain and side chain atoms of each residue, whereas peptide position is plotted on the X-axis.

HLA-E*0103-Mtb44*P2-Gln structural interpretation

Electron density present in the asymmetric unit enabled model building of all residues within the peptide, heavy chain and β2M of all 4 HLA-E*0103-Mtb44*P2-Gln molecules, with the exception of a single heavy chain residue at position 227 in the 220's loop of molecule 3. Intra-structural molecule alignment revealed similar positioning of residues present in the peptide binding groove α-helices and β-sheet floor in addition to the β-sheets of the invariant light chain, β2M (Figure 30, A). However, marked differences were observed in the α3 domains between the four molecules in the asymmetric unit. Notably, the α3 domain-resident flexible 220's loop adopted distinct conformations in each molecule with that of molecule 4 exhibiting

the greatest deviation (Figure 30, B). For example, the Gly-221 Cα atom of molecule 4 sits 9.2 Å from that of molecule 1, 7.9 Å from that of molecule 2 and 8.3 Å from that of molecule 3 (Figure 30, C).

Furthermore, similar to observations made for molecule 3 of the wild-type HLA-E*0103-Mtb44 structure, molecule 4 of HLA-E*0103-Mtb44*P2-Gln also exhibits a shifted α3-domain profile, involving not only the 220's loop but also the β2M-distal β-sheet plane (Figure 30, D). For example, a distance of 2.8 Å is observed between Cα atoms of the α3 β-strand Val-270 in molecules 4 and 3 of the HLA-E*0103-Mtb44*P2-Gln structure, similar to the 2.7 Å distance observed between molecules 3 and 1 of the HLA-E*0103-Mtb44 structure. The α3 domain of HLA-E*0103-Mtb44*P2-Gln molecule 4 superposes to that of molecule 3 in the wild-type structure indicating both such molecules share a distinct shifted α3 domain profile. This is perhaps unsurprising considering both structures pack in the P1 space group and have similar unit cell dimensions.

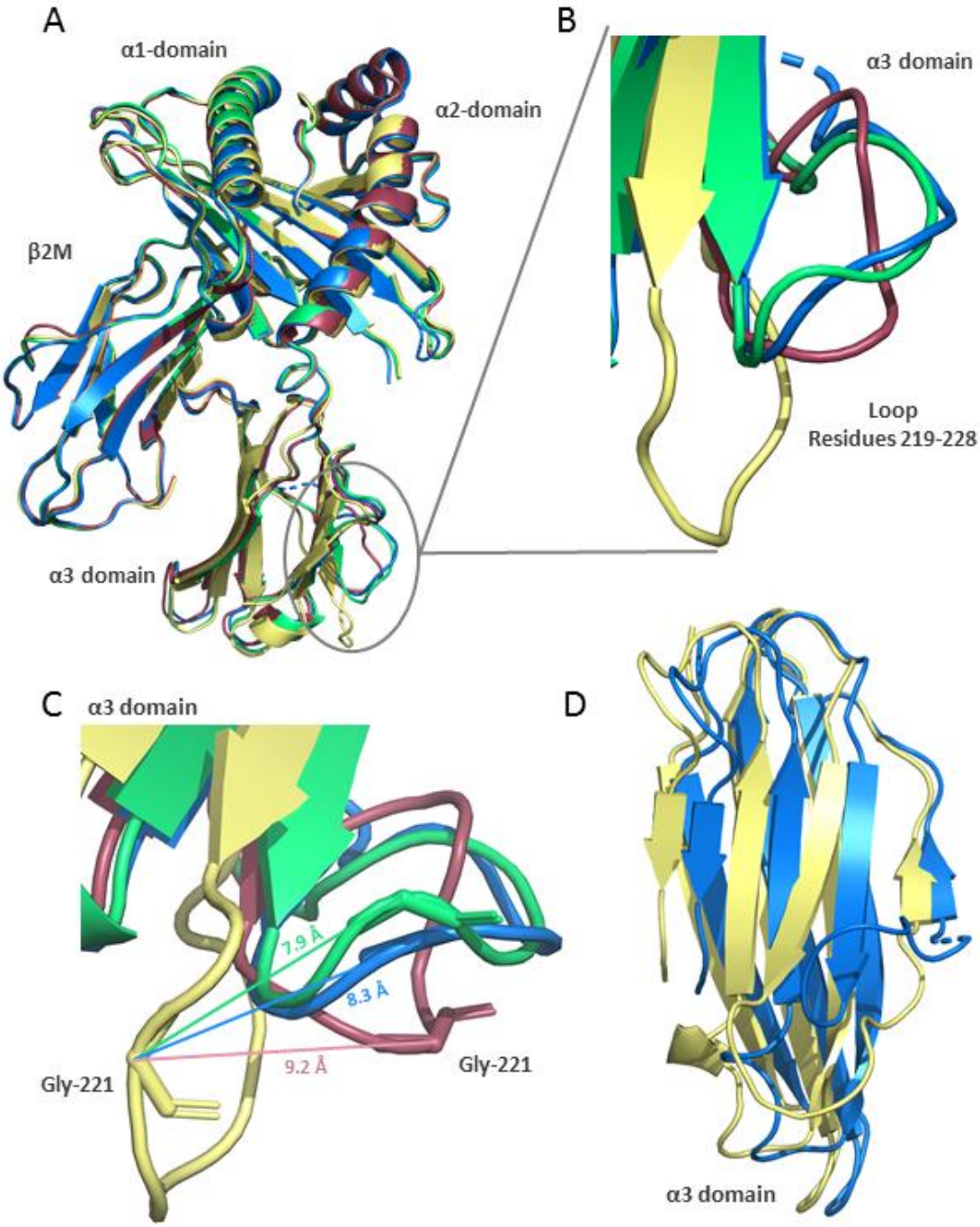


Figure 30: Analysis of molecules from the HLA-E*0103-Mtb44*P2-Gln structure

A. A cartoon visualisation of the 4 superposed molecules present in the asymmetric unit of the Mtb44*P2-Gln structure. Molecule 1 is depicted in maroon, molecule 2 in lime-green, molecule 3 in blue and molecule 4 in pale-yellow. The flexible 220's loop of the α3 domain is circled. **B.** Magnification of the flexible 200's loop with colour coding of superposed molecules according to A. **C.** Visualisation of the flexible 220's loop, with colour coding of molecules according to A. Gly-221 is depicted as sticks and

distances between Gly-221 Cα atoms of the 4 molecules are indicated by lines. The Gly-221 Cα in molecule 4 is situated 9.2 Å from that of molecule 1, 7.9 Å from that of molecule 2 and 8.3 Å from that of molecule 3. **D.** Superposed α3 domains of molecule 4 (pale-yellow) and 3 (blue) from the HLA-E*0103-Mtb44*P2-Gln structure, with remaining heavy chain and β2M-derived residues omitted for clarity.

Regions with the greatest density of elevated B factors comprised flexible loops and polypeptide termini in the HLA-E*0103-Mtb44*P2-Gln structure. However, similar to that observed in the wild-type Mtb44 structure, each molecule had a distinct B factor profile (Figure 31). In both Mtb44 and Mtb44*P2-Gln structures, the molecule containing the most dramatically reoriented 220's loop also exhibited the lowest B factors in the α3 domain.

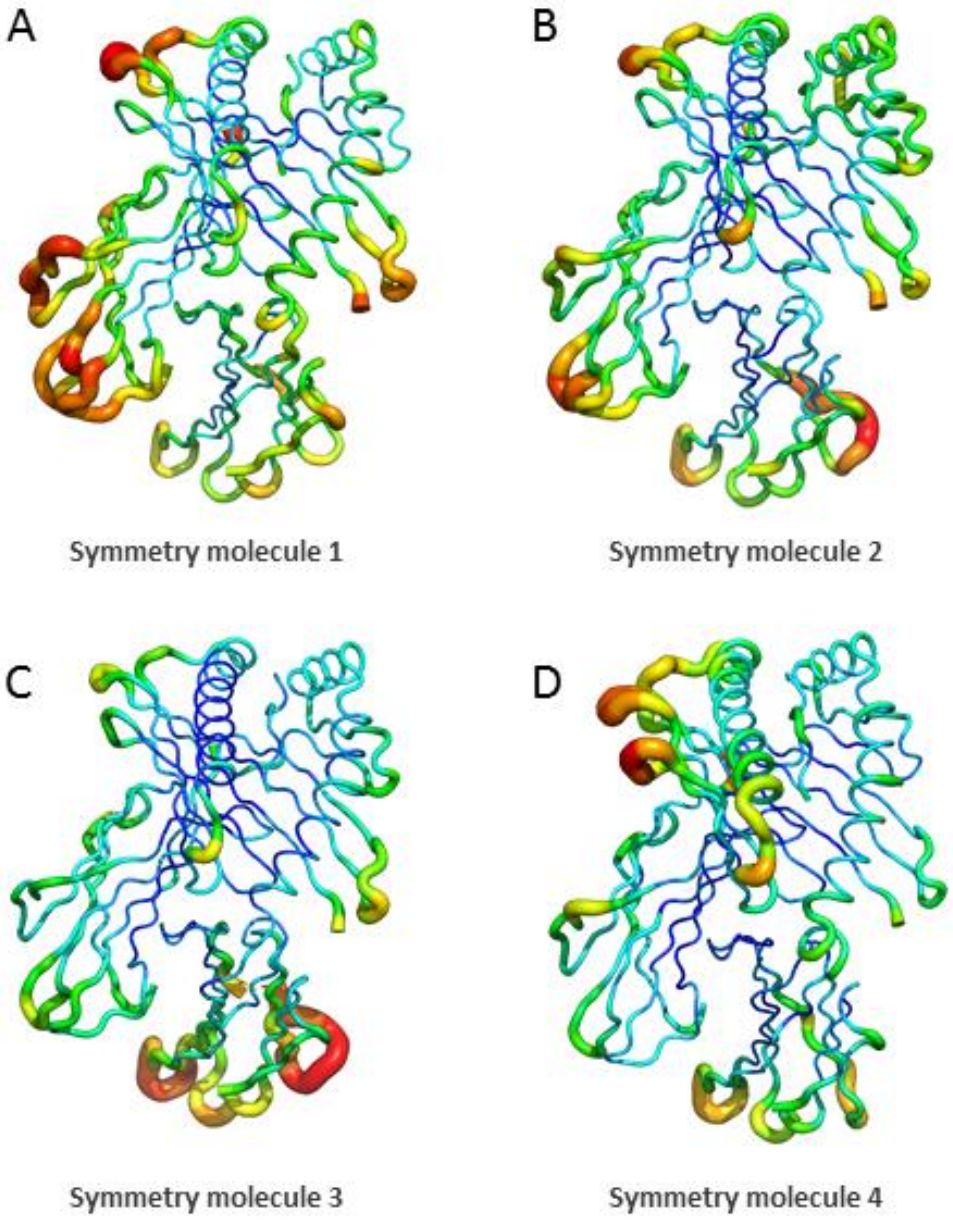


Figure 31: B factor analysis of Mtb44 position 2 Gln peptide-bound HLA-E molecules

B factor putty visualisation of HLA-E*0103-Mtb44*P2-Gln molecules 1 (**A**), 2 (**B**), 3 (**C**) and 4 (**D**). Cα backbone B factors are depicted on a cool to warm colour scale with the lowest relative B factors in blue and the highest in red.

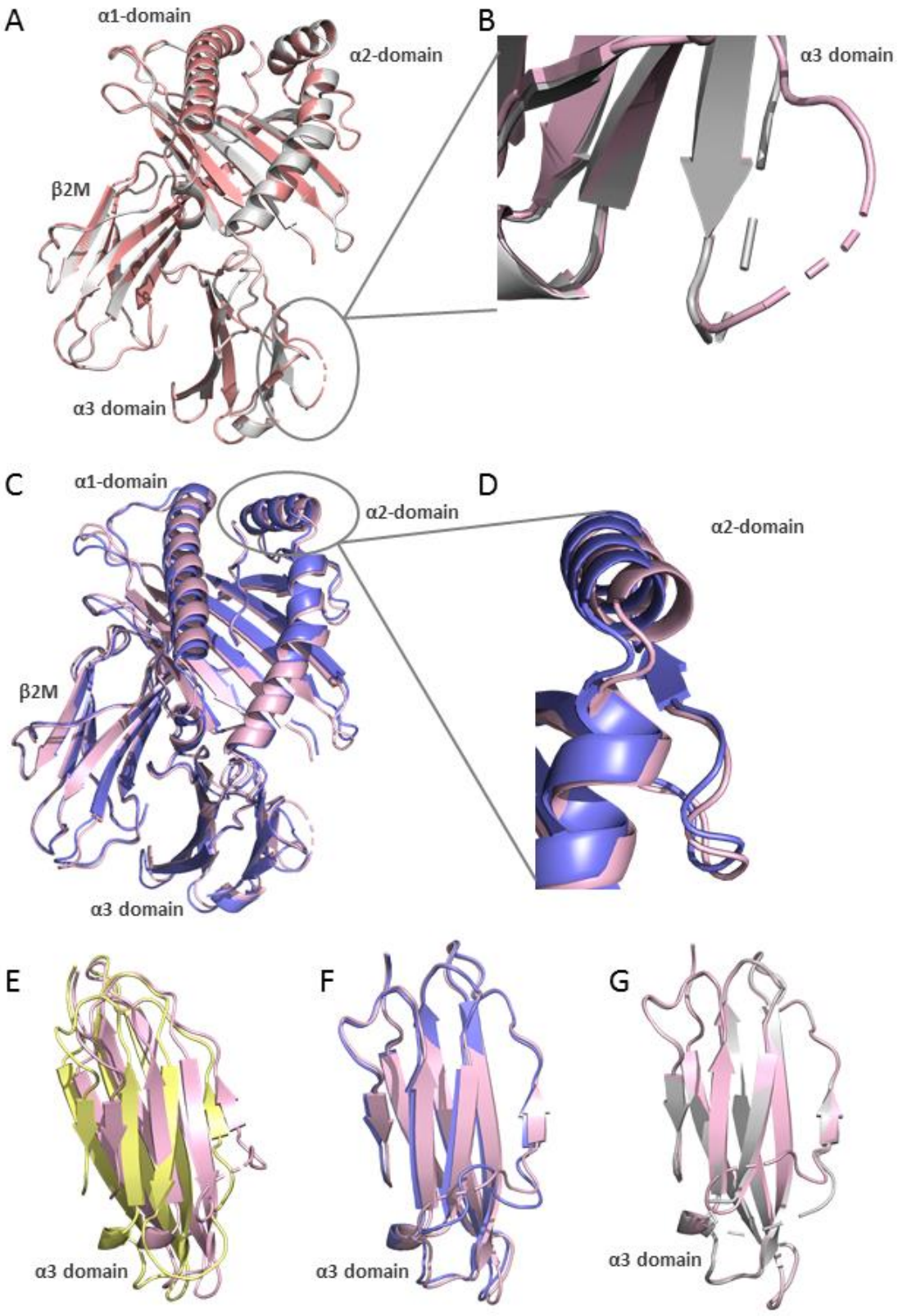
HLA-E*0103-Mtb44*P9-Phe structural interpretation

Insufficient electron density prevented the complete modelling of the HLA-E heavy chain for either molecule of the HLA-E*0103-Mtb44*P9-Phe structure. Consequently, residues 221-225 of molecule 1 and residues 222-223 of molecule 2 were omitted from the deposited coordinates – both such regions form part of the aforementioned flexible 220's loop.

Superposition of molecules 1 and 2 revealed similar orientations for all secondary structure elements and flexible loops, with the only exception comprising the heavy chain-derived flexible 220's loop of the α3 domain (Figure 32, A). Despite the absence of several residues from this region, similar to the wild-type HLA-E*0103-Mtb44 structure, variable loop positioning was still clearly detectable between molecules in the asymmetric unit (Figure 32, B). However, unlike the wild-type structure the α3 domain adopted similar orientations with no shifted α3 signatures in either molecule (Figure 32, E-G).

When HLA-E*0103-Mtb44*P9-Phe molecules were aligned to those from the wild-type Mtb44 structure, the most dramatic readjustments were present in the shorter arm of the α2 helix of the peptide binding groove (Figure 32, C and D). For example, the Cα atoms of Ala-150, situated in the kinked α2 helical region, are situated 1.9 Å apart between molecules 1 or 2 of the HLA-E*0103-Mtb44*P9-Phe structure compared to molecule 1 of the wild-type Mtb44 structure. As expected, Cα atom temperature factor analysis of Mtb44*P9-Phe molecules revealed the highest relative B factors were associated with residues situated in flexible loop regions or polypeptide termini (Figure 32, H and I). Further, the surface loops that displayed the greatest relative B factors were those connecting β-sheets in the α3-domain or β2M.

Considering the Mtb44*P9-Phe peptides were similarly positioned within the binding grooves of the two molecules in the asymmetric unit, with less than 1 Å movements between Cα atoms, a single peptide was selected for downstream analysis.


A
α1-domain
α2-domain
β2M
α3 domain
B
α3 domain
C
α1-domain
α2-domain
β2M
α3 domain
D
α2-domain
E
α3 domain
F
α3 domain
G
α3 domain

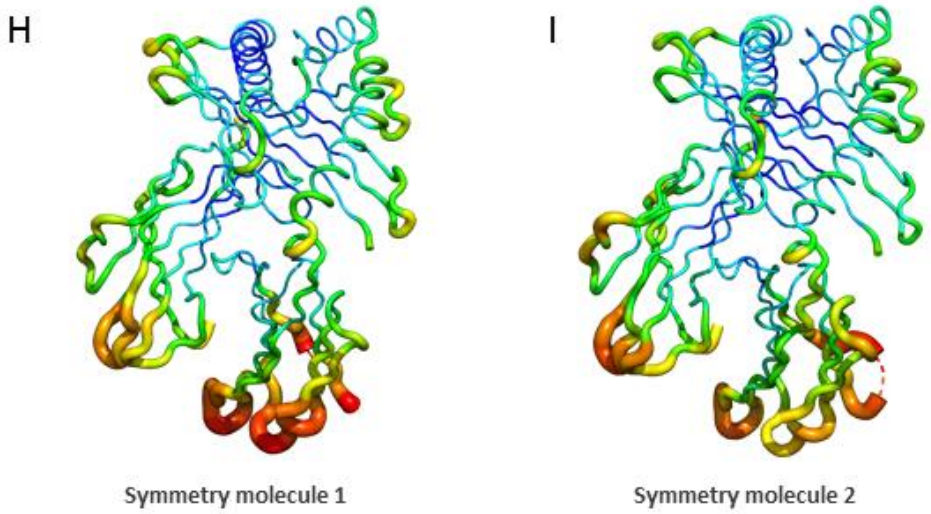


Figure 32: Analysis of molecules in the HLA-E*0103-Mtb44*P9-Phe structure

A. Superposition of molecules 1 (grey) and 2 (light pink) from the HLA-E*0103-Mtb44*P9-Phe structure. **B.** Magnification of the heavy-chain-derived flexible 220's loop of the alpha-3 domain. Mtb44*P9-Phe-bound molecules are superposed and colour coded according to A. **C.** Superposition of molecule 2 from the Mtb44*P9-Phe structure (light pink) and molecule 1 from the wild-type Mtb44 structure (purple). **D.** Magnification of the α2 helix short arm. Superposed molecules are colour-coded according to C. **E.** Superposed α3 domains from molecule 2 of the Mtb44*P9-Phe structure (light pink) and molecule 3 of the wild-type Mtb44 structure (yellow). **F.** Superposed α3 domains from molecule 2 of the Mtb44 position 9 Phe structure (light pink) and molecule 1 of the wild-type Mtb44 structure (purple). **G.** Superposed α3 domains from molecules 1 (grey) and 2 (light pink) of the Mtb44*P9-Phe structure. **H. & G.** B factor representation for Mtb44*P9-Phe molecules 1 and 2, respectively. Cα atom B factors are depicted on a cool to warm colour scale with the lowest relative B factors in blue and the highest in red.

Comparative analysis of peptide positioning between Mtb44 variant structures

Despite distinct orientations in the central kinked region of the Mtb44*P2-Phe peptides, strong alignment is observed at remaining positions including those of primary and secondary anchor residues. Thus, a single Mtb44*P2-Phe peptide from molecule 1 shall form the basis of subsequent comparative analysis to the Mtb44*P2-Gln and Mtb44*P9-Phe peptides.

Clear electron density was present in the binding groove for Mtb44*P2-Phe and Mtb44*P2-Gln modelling (Figure 33, A and B). Peptide superposition revealed that the Mtb44*P2-Gln peptide backbone aligned closely to the wild-type Mtb44 peptide, with all inter-peptide Cα positions deviating by less than 0.5 Å (Figure 33, C-E). Regarding the Mtb44*P2-Phe peptide, the Cα atoms of position 6 and 7 sit deeper in the binding groove by 0.9 Å and 0.8 Å respectively, relative to wild-type Mtb44. This in turn decreases the height of the main chain kink by 0.8 Å between positions 4 and 5 compared to the wild-type Mtb44 peptide.

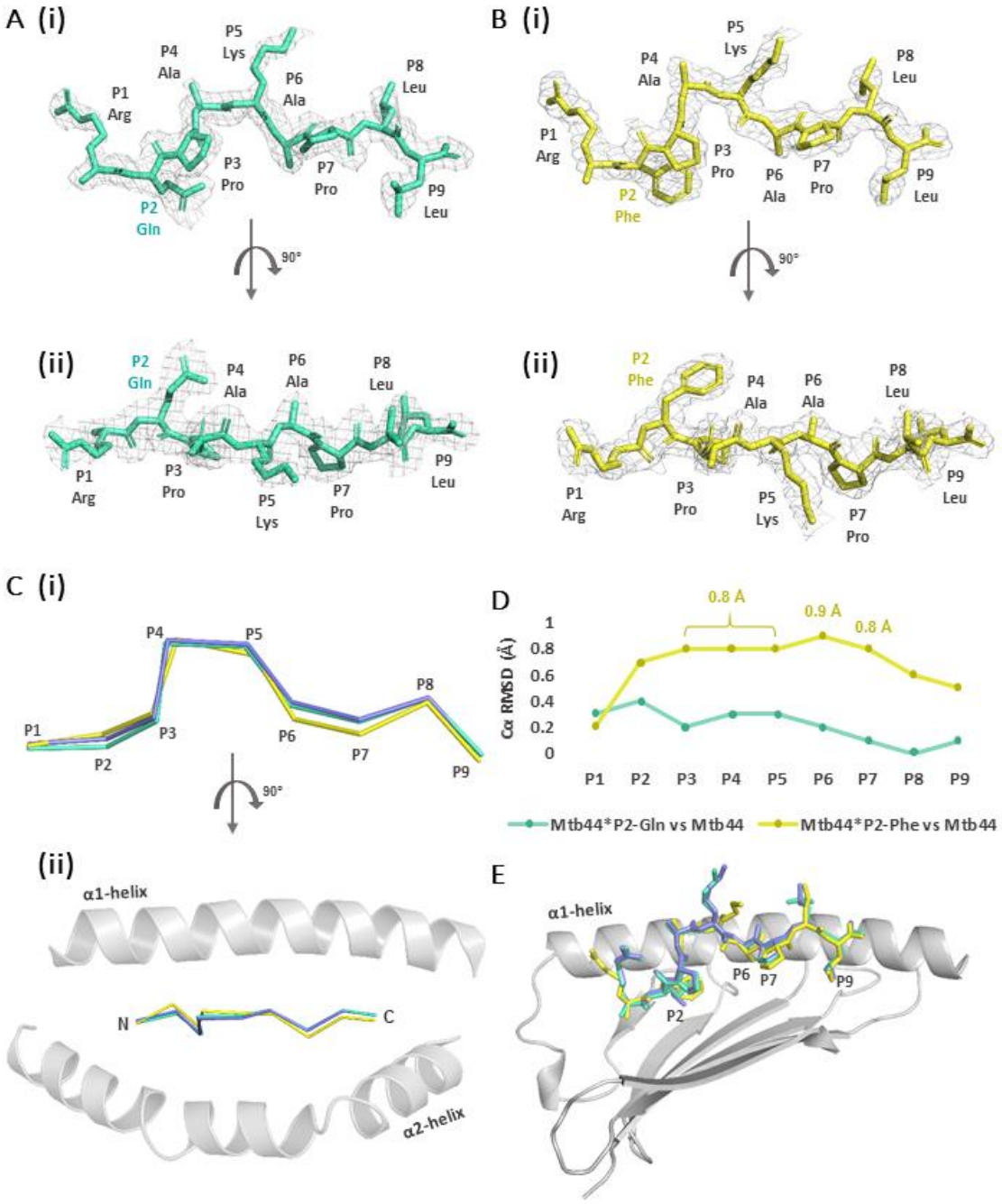


Figure 33: Mtb44 position 2 Gln and Phe variant peptide positioning relative to wild-type Mtb44

A. The Mtb44*P2-Gln peptide from molecule 1 is shown in lime-green stick-form and visualised from the side **(i)** and from above **(ii)** with an electron density map contoured to 1 sigma, shown in grey mesh overlaying the peptide. **B.** The Mtb44*P2-Phe peptide from molecule 1 is shown in yellow stick-form and visualised from the side **(i)** and from above **(ii)** with an electron density map contoured to 1 sigma, shown in grey mesh overlaying the peptide. **C.** Cα backbone superposition of the Mtb44*P2-Gln (lime-green) and Mtb44*P2-Phe (yellow) variant peptides with wild-type Mtb44 (purple). Peptides are

visualised side-on **(i)** or from above **(ii)**, in ribbon-form, with the α1 and 2 helices depicted in grey cartoon and the remaining HLA-E heavy and light chains omitted for clarity. **D.** Graph depicting deviation between Cα atoms of the Mtb44*P2-Gln (lime-green) and Mtb44*P2-Phe (yellow) variant peptides relative to wild-type Mtb44. Distances are specified in Å on the Y-axis while peptide position is plotted on the X-axis. **E.** Side-on visualisation of the Mtb44*P2-Gln (lime-green) and Mtb44*P2-Phe (yellow) peptides shown in stick-form superposed to the Mtb44 wild-type peptide (purple). The α1 helix and β-sheet floor of the peptide binding groove are depicted as grey cartoon with the α2 helix omitted for clarity.

Clear electron density was present for Mtb44*P9-Phe peptide main chain and side chains (Figure 34, A). All Cα atoms deviate by less than 1 Å in the Mtb44*P9-Phe variant relative to wild-type Mtb44, except position 5 which sits closer to the α1 helix in Mtb44*P9-Phe (Figure 34, B-D).

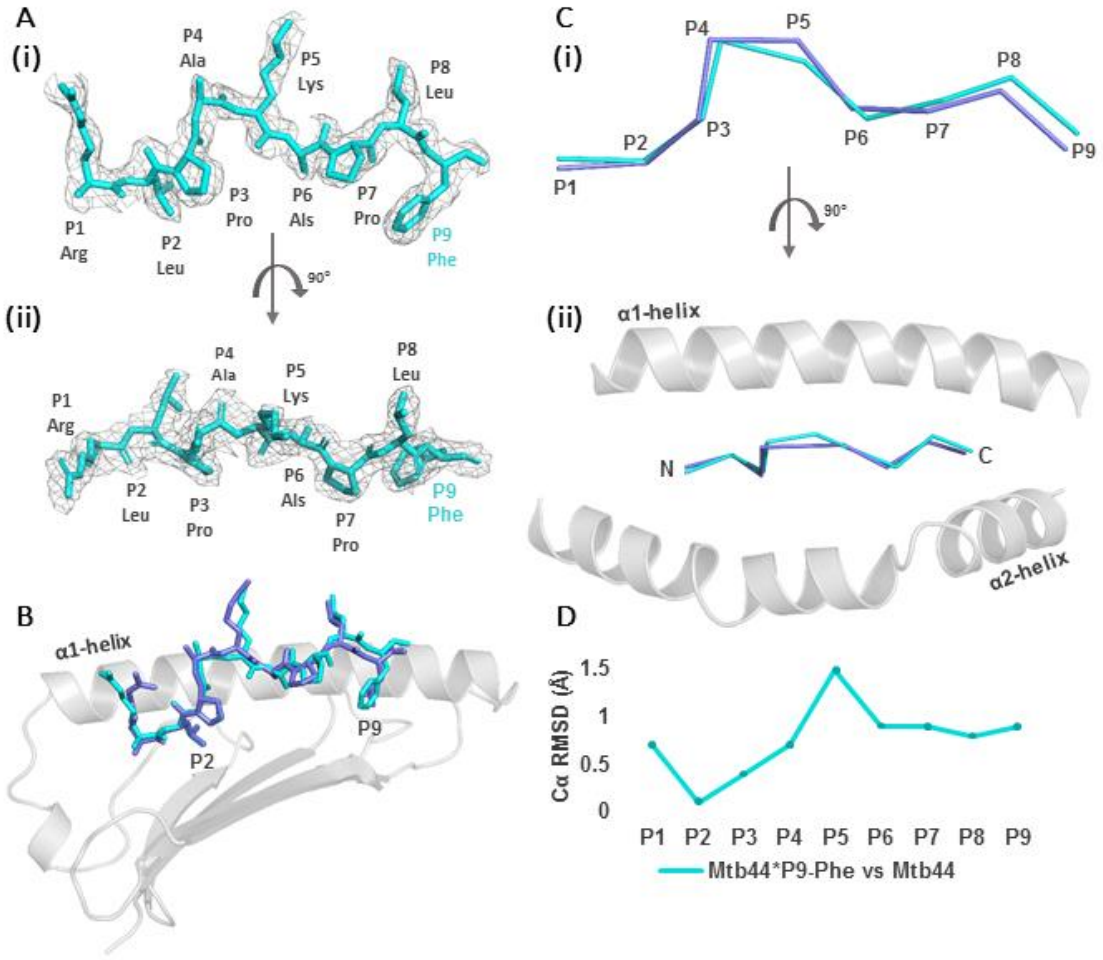


Figure 34: Positioning of the Mtb44*P9-Phe variant peptide relative to Mtb44

A. The Mtb44*P9-Phe peptide is shown in cyan stick-form from the side **(i)** and from above **(ii)** with an electron density map contoured to 1 sigma, shown in grey mesh overlaying the peptide. **B.** Side-on visualisation of the Mtb44*P9-Phe peptide (cyan) superposed to the Mtb44 wild-type peptide (purple). The α1 helix and β-sheet floor of the peptide binding groove are depicted as grey cartoon with the α2 helix omitted for clarity. **C.** Cα backbone superposition of the Mtb44*P9-Phe variant (cyan) with wild-type Mtb44 (purple). Peptides are visualised side-on **(i)** or from above **(ii)** in ribbon-form with the α1 and 2 helices depicted as grey cartoon. **D**. Graph depicting deviation between Cα atoms of the Mtb44*P9-Phe peptide relative to wild-type Mtb44. Distances are specified in Å on the Y-axis while peptide position is plotted on the X-axis.

Pockets

Substituted Gln and Phe side chains at position 2 of the Mtb44*P2-Gln and Mtb44*P2-Phe peptides projected into the B pocket toward the groove floor (Figure 35, A). Superposition of the position 2 Phe and Gln Mtb44 variants to wild-type Mtb44 revealed similar positioning of the peptide backbone at position 2 and resultantly, the larger Phe and Gln side chains project slightly deeper into the B pocket than the wild-type Leu. Furthermore, superposition of B pocket-forming residues revealed minimal repositioning in either of the position 2 Mtb44 variant structures relative to wild-type Mtb44.

Similarly, the side chain of the substituted position 9 Phe is accommodated by the F pocket and superposes to position 9 Leu of the wild-type Mtb44 peptide (Figure 35, B). Although both wild-type Leu and substituted Phe position 9 peptide side chains are orientated similarly in the F pocket, minor readjustments in both the F pocket-forming residues and the peptide backbone positioning are observed. First, the Cα atom of position 9 Phe in the Mtb44*P9-Phe peptide is positioned 1 Å above that of the wild-type Mtb44 position 9 Leu. Such repositioning likely enables accommodation of the larger Phe side chain in the F pocket. Additionally, the F-pocket forming residue, Phe-116 adopts a distinct orientation in the HLA-E*0103-Mtb44*P9-

Phe structure in which the side chain is positioned more acutely toward the floor of the groove. This side chain re-angling results in a slightly increased F pocket volume permitting the accommodation of the larger Phe side chain at position 9.

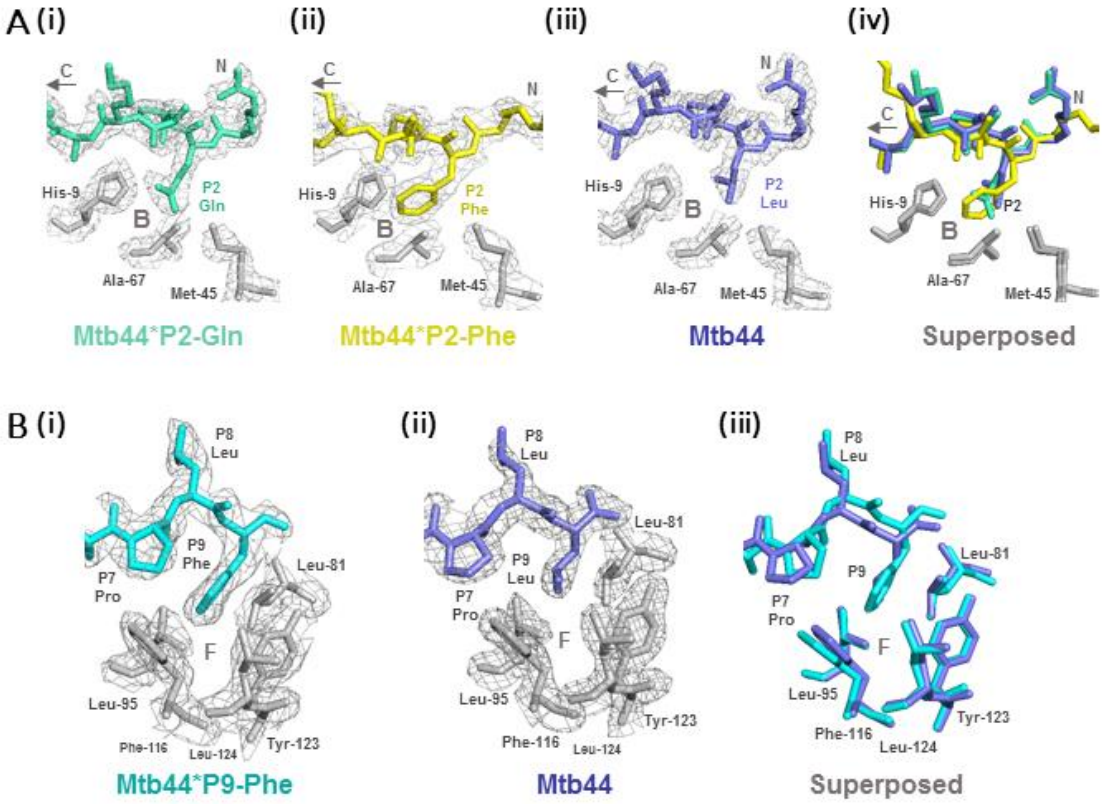


Figure 35: Primary pocket accommodation of Mtb44 position 2 and 9 primary anchor peptide variants

A. Visualisation of the B pocket accommodating the position 2 primary anchor residue side chain. B pocket-forming residues are depicted as grey sticks with overlaying electron density maps in grey mesh, contoured to 1 sigma. The Mtb44*P2-Gln peptide is depicted in green sticks **(i)**, the Mtb44*P2-Phe peptide in yellow **(ii)** and the wild-type Mtb44 in purple **(iii)**. **(iv).** Superposed B pocket-forming residues and Mtb44 position 2 variant and wild-type peptides, colour-coded according to A (i), (ii) and (iii). **B.** Visualisation of the F pocket accommodating the position 9 primary anchor residue side chain. F pocket-forming residues are depicted as grey sticks with overlaying electron density maps in grey mesh, contoured to 1 sigma. The Mtb44*P9-Phe peptide is depicted in cyan **(i)** and the wild-type Mtb44 in purple **(ii). (iii).** Superposed F pocket-forming residues and peptides from the Mtb44*P9-Phe and wild-type Mtb44 structures, colour-coded according to B (i) and (ii).

Inter-chain hydrogen bonding between Mtb44 variant peptides and the HLA-E heavy chain

Inter-chain hydrogen bonding networks between the peptide and HLA-E heavy chain for the Mtb44 position 2 and 9 variant structures (Figure 36) revealed that 3 hydrogen bonds are lost in the Mtb44*P9-Phe structure which connect peptide residues 1, 2 and 9 of wild-type Mtb44 to the HLA-E heavy chain. As previously discussed, due to distinct hydrogen bonding networks arising from conformational differences in Mtb44*P2-Phe peptide positioning, individual hydrogen bonding analyses were performed for each molecule in the asymmetric unit. Ten and twelve hydrogen bonds are formed between the Mtb44*P2-Phe peptide and HLA-E heavy chain in molecules 1 and 2, respectively. However, despite a discrepancy of at least 4 hydrogen bonds, both Mtb44*P2-Phe and Mtb44*P9-Phe peptides demonstrate similar relative binding strengths to HLA-E, as assessed in the peptide-exchange-based binding assay.

HLA-E*0103 - Mtb44*P2-Gln

	Peptide	Dist. Å	Heavy chain
1	P1-Arg [N]	2.74	Tyr-171 [OH]
2	P1-Arg [NH1]	3.61	Thr-163 [OG1]
3	P1-Arg [O]	2.7	Tyr-159 [OH]
4	P2-Gln [N]	2.87	Glu-63 [OE2]
5	P2-Gln [NE2]	3.38	Thr-70 [OG1]
6	P7-Pro [O]	3.1	Asn-77 [ND2]
7	P9-Leu [N]	2.93	Asn-77 [OD1]
8	P9-Leu [O]	2.89	Tyr-84 [OH]
9	P9-Leu [O]	2.9	Ser-143 [OG]
10	P9-Leu [OXT]	2.64	Lys-146 [NZ]

HLA-E*0103 - Mtb44*P9-Phe

	Peptide	Dist. Å	Heavy chain
1	P1-Arg [O]	2.73	Tyr-159 [OH]
2	P2-Leu [N]	2.58	Glu-63 [OE2]
3	P7-Pro [O]	3.11	Asn-77 [ND2]
4	P9-Phe [N]	2.74	Asn-77 [OD1]
5	P9-Phe [O]	3.02	Tyr-84 [OH]
6	P9-Phe [O]	3.38	Ser-143 [OG]

Figure 36: Peptide-heavy chain hydrogen bonding networks for Mtb44*P2-Gln and Mtb44*P9-Phe

A. Table detailing hydrogen bonds formed between Mtb44 variant peptides, Mtb44*P2-Gln and

Mtb44*P9-Phe, and the HLA-E heavy chain. Lower and upper thresholds for donor/acceptor distances are set to 2.5 Å and 3.5 Å, respectively.

Discussion

This chapter provides structural evidence for the HLA-E binding capacity of Mtb-derived epitopes and illustrates, at the atomic level, how such peptides are accommodated by the HLA-E binding groove. Further, it structurally demonstrates that the primary B and F pockets of HLA-E can tolerate a range of hydrophobic and polar residues.

Implications for T cell recognition and vaccine-design

First, we have shown that pathogen-derived non-leader peptides occupy the HLA-E peptide binding groove in a similar manner to VL9, despite sequence differences along the nonamers. Although Mtb44 and Mtb14 peptides adopt the classical kinked motif and superpose to the VL9 peptide backbone, sufficient differences in solvent exposed regions enable discrimination with self-derived VL9 by $CD8^+$ T cells. Position 8 of Mtb44 and Mtb14 (Leu and Val, respectively) constitute residues also present at position 8 of certain VL9 peptide variants. However, the position 5 side chain exhibits an even greater degree of solvent exposure than position 8 in the classical, kinked peptide binding motif. Position 5 of both Mtb44 and Mtb14 (Lys and Thr, respectively) comprise residues absent at position 5 in VL9 peptide variants and may thus contribute to $CD8^+$ T cell discrimination. Accordingly, as previously mentioned, Mtb44-specific Mamu-E-restricted $CD8^+$ T cell responses were identified in BCG-vaccinated rhesus macaques (Walters et al., 2018). Further, an excess of VL9 peptide abrogated such responses confirming their specificity and lack of self-reactivity.

Interestingly, such responses were resistant to position 2 primary anchor substitution in the Mtb44 backbone, with responses of a similar magnitude resulting from stimulation with Mtb44, Mtb44*P2-Phe or Mtb44*P2-Gln. A lack of T cell discrimination is unsurprising given

the strong positional alignment of the 3 Mtb44 variant backbones and the buried nature of the substituted position 2 side chain within the deep hydrophobic B pocket. This constitutes a clear example whereby primary anchor optimisation to fine-tune complex stability and immunogenicity without disrupting T cell recognition could be possible in a vaccine context. Given the well-established correlation between complex stability and immunogenicity (Harndahl et al., 2012), the inclusion of optimal anchor residues in candidate epitopes during immunogen design is predicted to yield more potent T cell-focused vaccines.

Superior immunogenicity deriving from enhanced complex stability via the modification of peptide anchors has previously been identified for HLA-A*0201-restrcited epitopes (Borbulevych et al., 2005). However, such results are uniquely peptide-dependent since the introduction of a specific side chain at a given position can result in steric interactions with remaining side chains along the peptide and cause twisting of the backbone in various directions to various degrees. Consequently, identical downstream side chains at the same peptide position have been shown to project in different directions when bound to the same MHC class I molecule (Madden et al., 1993). Thus, substitution at a single buried anchor position can impose distal conformational ramifications and influence peptide solvent exposure profiles, with unanticipated implications for *in vivo* antigenicity and T cell recognition. Equally, anchor residue substitution may indirectly influence T cell recognition through peptide-dependent conformational plasticity of the MHC class I binding groove (Hawse et al., 2013). Accordingly, sensitivity of H-2K^{b}-specific alloantibodies (Hogquist et al., 1993), in addition to differential T cell recognition (Chen, 1993), have both previously been reported in response to buried peptide anchor residue sequence variation. Thus, anchor-optimised candidate epitopes would likely require validation on an individual basis.

Conformational flexibility of the α3 domain

A shifted β-sheet plane in the α3 domain was identified in one of four molecules of the HLA-E*0103-Mtb44 and HLA-E*0103-Mtb44*P2-Gln structures. Considering multiple α3 domain configurations are observed within the same asymmetric unit, it is conceivable such movements are peptide-independent. Further, the shifted α3 domains in molecule 3 of HLA-E*0103-Mtb44 and molecule 4 of HLA-E*0103-Mtb44*P2-Gln, are superimposable. Given both structures packed in the P1 space group with almost identical unit cell dimensions, such α3 domain rearrangement is presumably a result of non-physiological crystal packing contacts specific to this space group and unit cell.

The stabilising interface connecting the α3 domain with the membrane-proximal plane of the peptide binding groove is reportedly weak relative to that of β2M (Madden, 1995). Such discrepancies in interface strength could render the α3 domain more susceptible to crystal packing forces, permitting the observed flexibility. Conversely, the more extensive network of contacts connecting β2M to the peptide binding groove is consistent with the rigid intra- and inter-structural disposition of this immunoglobulin-like domain, which exhibits strong alignment between classical and non-classical MHC class I structures (Madden, 1995).

Previously, distinct α3 domain β-sheet and 220's loop positioning has been observed among structures of classical MHC class I molecules, such as between HLA-A*0201 and HLA-A*6801 (Niu et al., 2013). Such movement was originally ascribed peptide-independence and attributed to the factors discussed above, such as crystal packing contacts in addition to weak tethering of the α3 domain to the rest of the protein complex (Madden, 1995). However, a recent meta-analysis focusing specifically on the α3 domain-situated 220's loop concluded that spurious and inconsistent crystal contacts formed within this region were insufficient to account for the substantial conformational flexibility observed among numerous MHC class I structures (Ayres et al., 2019). Further, similarly non-extensive crystal packing contacts were

formed at multiple alternate loop regions within the MHC heavy chain which exhibit minimal structural deviation across MHC class I structures relative to the 220's loop. Thus, physiological relevance and a possible peptide-dependence has since been postulated for α3 domain structural perturbations (Ayres et al., 2019).

Alignment of all 14 molecules from the asymmetric units of the 6 previously solved structures of VL9-bound HLA-E (1MHE, 1KTL, 1KPR, 3BZE, 3BZf and 3AM8) revealed minimal repositioning of the α3 domain β2M-distal β-sheet plane relative to that observed for Mtb44, Mtb44*P2-Gln and Mtb14-bound HLA-E molecules. For example, displacement of the β-sheet-resident Val-270 Cα atom does not exceed 1.1 Å between intra- and inter-structural VL9-bound HLA-E molecules. However, 2.8 Å, 2.7 Å and 1.7 Å separate intra-structural Val-270 Cα atoms of Mtb44, Mtb44*P2-Gln and Mtb14-bound HLA-E molecules, respectively. Further, considering VL9-bound HLA-E structures packed in a variety of space groups with different sized unit cells, re-aligned α3 domains might not constitute an artefact of crystal packing, but could instead represent a unique feature specific to non-VL9 peptide-bound HLA-E. Moreover, although shifted α3 domains were present in two structures that shared the same space group (P1) and similar unit cell dimensions (a = 48.35 Å, b = 76.31 Å and c = 111.02 Å for Mtb44-HLA-E and a = 48.31 Å, b = 76.41 Å and c = 111.31 Å for Mtb44*P2-Gln-HLA-E), α3 domain β-sheet re-positioning was also present in molecules of the Mtb14-bound HLA-E structure which packed in a different space group (C2) with a larger unit cell (a = 245.15 Å, b = 47.93 Å and c = 152.45 Å).

In further support of a physiologically relevant role, recent molecular dynamics simulation studies comprising 52 HLA-A*0201 complexes indicate bound peptides not only manipulate conformational sampling of proximal binding groove elements, but can also distally influence α3 domain dynamics through conserved networks via the β2M domain (Ayres et al., 2019). This process, termed 'dynamically driven allostery', is experimentally supported by peptide-

dependent differences derived from fluorescence anisotropy data for HLA-A2 position 220 Asp to Cys mutants (Ayres et al., 2019). It is therefore possible that the shifted α3 domains present in Mtb44-, Mtb44*P2-Gln- and Mtb14-bound structures of HLA-E, might reflect distinct, peptide-driven patterns of conformational sampling in solution. This could in turn impart indirect peptide-specificity to interactions with the multiple α3 domain-targeted binding partners including NK cell receptors, the CD8 co-receptor and chaperones.

Peptide-dependency was previously reported for murine H-2K^b recognition by the inhibitory NK cell receptor Ly49C (Franksson et al., 1999). However, subsequent co-complex structures revealed that bound peptide did not contribute to the H-2K^b-Ly49C binding interface, thus suggestive of an indirect mechanism for peptide-specificity (Dam et al., 2003). Considering 27% of H-2K^b interfacing residues are derived from the α3 domain, this represents an example whereby peptide-driven positional differences in the MHC class I α3 domain could be of physiological importance.

As predicted by mutagenesis analyses (Salter et al., 1990) and confirmed by an HLA-A*02-CD8αα co-complex structure (Gao et al., 1997), a critical part of the extended CD8 co-receptor binding site is situated between residues 216-230 of the MHC class I α3-domain. Distinct structural reorientation within the α3 domain 220's surface loop, arising from heavy chain polymorphism at position 245, has previously been linked to CD8 co-receptor binding affinity (Gao et al., 2000). For example, the α3 domain β2M-distal β-sheet plane is shifted in HLA-A*6801, and also presumably in HLA-B*4801, relative to the majority of MHC class I alleles such as HLA-A*0201, HLA-A*1101 and HLA-B*3501, resulting in 2 distinct 220's loop configurations. The HLA-A*6801 and HLA-B*4801 configuration disrupts CD8 co-receptor engagement and diminishes binding affinity, resulting in dissociation constants of around 1mM, as opposed to within the 90-220μM range reported for other MHC. Thus, this

represents a key example whereby α3 localised conformational reorganisation is potentially reflective of differential physiological functionality.

Considering the 220's loop of VL9-bound HLA-E molecules adopts a similar configuration to that formed by HLA-A*6801 and HLA-B*4801, it is unsurprising that HLA-E binds the CD8 co-receptor with comparably weak affinity ($K_d \geq 1mM$) (Gao et al., 2000). However, enhanced, intra-structural α3 domain flexibility which deviates from the reported HLA-A*6801 configuration was observed in multiple non-VL9-bound structures of HLA-E presented in this chapter. It is possible this could provide the 220's loop re-positioning required for stronger CD8 co-receptor interaction. Peptide-dependent increases in HLA-E-CD8 co-receptor binding affinity would be functionally consistent with the antigenic nature of non-VL9, Mtb-derived HLA-E-bound peptides and the alternative antigen presenting role of HLA-E in eliciting αβ $CD8^+$ T cell responses.

However, re-positioned α3 domain profiles and displaced 220's surface loops in non-VL9 structures of HLA-E do not align with those present in MHC class I alleles displaying higher CD8 binding affinity, such as HLA-A*0201. Instead, HLA-E α3 domain shifts are observed along a similar axis but in an opposite direction to that required for optimal superposition to the HLA-A*0201 α3 domain, with the HLA-E 220's loop assuming a distinct, third configuration. Thus, HLA-E-CD8 co-receptor binding affinity may alternatively remain unaffected by enhanced HLA-E α3 domain flexibility. Moreover, partial conversion of the 220's loop into an ordered, extended β-strand in molecule 1 of the Mtb14-bound HLA-E structure could render this region of the CD8 co-receptor binding site more inaccessible. Regardless of 220's loop distortion, deviation from the MHC class I consensus sequence at positions 219, 223 and 224 of the HLA-E heavy chain within the CD8 co-receptor binding interface could account entirely for the weak affinity interactions (Gao et al., 2000).

Despite the absence of a co-complex crystal structure, the α3 domain 220's surface loop has been identified as a tapasin-interacting region, in addition to residues 128-137 of the α2 domain (Suh et al., 1999). Tapasin, for which a spectrum of dependency has been established for different MHC alleles, imparts its chaperoning effect by stabilising high energy peptide-free MHC class I transition states (Veronique M. Braud et al., 1998). Subsequent tapasin-mediated peptide exchange yields more optimally bound peptide-MHC complexes. In addition, an overlapping binding interface, centring around MHC class I α3 domain-derived residues 229-230, is formed with the C-terminal domain of TAP-binding protein-related (TAPBPR) - a tapasin homolog with supposedly stringent peptide loading complex-independent proof reading functionality (Thomas and Tampé, 2017).

Tapasin-dependent HLA-E surface expression has previously been established, with dissociation of HLA-E from the endoplasmic reticulum-resident chaperone triggered by VL9 peptide binding (Veronique M. Braud et al., 1998). However, interactions between TAPBPR and HLA-E have not been documented. Although major α3-domain contact residues, Glu-229 and Lys-230, are conserved between HLA-E and the classical MHC class I murine allele, H2-D^b, present in the 3.3 Å TAPBPR co-complex structure, key α2 domain-derived interfacing residues are not. For example, residues Thr-143 and Trp-147 are highly conserved among MHC class I molecules and form key contacts with both the peptide C-terminus and TAPBPR-derived scoop loop. However, both Thr-143 and Trp-147 are replaced with two Ser residues in HLA-E.

Further, of the MHC α2 helix short arm-derived interfacing residues, 70% identified in the TAPBRP- H2-D^b co-complex structure are not conserved with HLA-E, including two critical Arg residues at positions 144 and 145, which form hydrogen bonds with TAPBRP-derived Gln-275 and Ser-107, respectively. The charge reversal introduced at position 144 of HLA-E by an Arg to Glu substitution could yield repulsive electrostatic forces with the proximal TAPBRP-derived Glu-105, disrupting the centre of the α2-domain amino terminal interface. Thus, although

peptide-driven plasticity of the HLA-E α3 domain and beyond might have implications for tapasin engagement, it is unclear whether such an effect could extend to the TAPBPR chaperone.

Additionally, from a biological perspective, it seems unlikely that this supposedly ruthless peptide editor would be present in the peptide loading or exchange compartments operating in RhCMV 68-1 vaccinated cells. Low affinity SIV-derived epitopes would conceivably offer insufficient binding energy to outcompete the stabilising TAPBRP scoop loop and N-terminal interfaces.

Conformational flexibility in the α2 helix short-arm

Superposition revealed positional differences in the α2 helix short arm (residues 138-149) between VL9-bound HLA-E relative to Mtb44, Mtb14 or Mtb44 primary anchor variant peptide-bound HLA-E. Shorter intra-structural distances observed between α2 helical Cα atoms relative to corresponding inter-structure distances, could indicate a peptide-dependent nature for the observed motions.

Previously, the MHC class I α2 helical amino terminal was identified as the most structurally flexible region of the peptide binding groove. A recent crystallographic meta-analysis of normalised Cα B factors in a collection of 297 non-receptor-bound MHC class I molecules, revealed the greatest B factor variance within the binding groove was located in the α2 helix short arm (Wieczorek et al., 2017). Originally, motional properties of the α2 helix short arm were associated with minor positional adjustments at the peptide carboxy-terminus - specifically a C-terminally displaced octameric peptide in complex with HLA-B*3501 with concomitant disruption of the peptide-heavy chain hydrogen bonding network (Smith et al., 1996b). However, subsequent studies, including a comparison of five HLA-A2 structures bound to different viral peptides, failed to correlate α2-helical structural flexibility with re-orientated

peptide C-termini (Madden et al., 1993). Thus, such movements were thereafter attributed to non-physiological packing contacts dependent on the crystalline environment (Madden, 1995).

However, the inability to correlate positional adjustments in bound peptide, does not preclude a possible peptide-dependency for the observed α2 helix amino terminal motions, although crystallographic coordinates alone are insufficient to establish causal patterns of protein dynamics. More recently, molecular dynamics studies have indicated that the short arm of the α2 helix not only exhibits peptide-dependent motional properties but constitutes the most sensitive and greatly influenced region within the binding groove (Ayres et al., 2019). Interestingly, such conformational flexibility correlated weakly with peptide affinity, the presence of canonical primary anchors or structural repositioning of the peptide within the MHC class I binding groove. Instead, α2 helical motional properties appear to be associated with dynamic, root mean square (rms) fluctuations of the peptide position in solution, as opposed to structural repositioning into stable, fixed conformations.

Although X-ray diffraction of crystal lattices does not provide insight regarding peptide motional fluctuations, 2 distinct peptide orientations with up to 1.9 Å displacement were observed in each molecule of the HLA-E*0103-Mtb44*P2-Phe structure. The presence of 2 distinct conformational motifs could reflect substantial rms fluctuation of the peptide in solution as it transitions between these 2 fixed configurations. Further, multiple Mtb44*P2-Phe peptide conformations correlate with α2 helix short arm displacement which broadens the neighbouring region of the peptide binding groove in both Mtb44*P2-Phe-bound HLA-E molecules, relative to wild-type Mtb44. Such observations are consistent with the aforementioned molecular dynamics study in which enhanced RMS fluctuation of HLA-A2-bound peptide invariably correlated with a widened groove (Ayres et al., 2019). Similarly, α2 helical displacement is also present in the Mtb44*P9-Phe structure with an increased groove width in the vicinity of the α2 helix short arm relative to wild-type Mtb44. It is conceivable that

such differential α2 helical geometry is in direct response to substitution of Leu to a larger Phe at position 9 of the Mtb44 peptide, since C-terminal increases in peptide volume were also correlated with broadened HLA-A2 binding grooves in molecular dynamics simulations (Ayres et al., 2019).

In summary, repositioning of the α2 helix short arm observed in numerous structures of classical and non-classical MHC molecules including HLA-E could reflect a peptide-driven extension of antigenicity with implications for T cell receptor recognition – given that the α2 helical amino terminal portion forms critical contacts in the T cell receptor binding interface. For HLA-E, such movements could play a key role in facilitating $CD8^+$ T cell receptor discrimination between self-derived VL9 nonamers and pathogen-derived peptide epitopes. Combined structural, biophysical and computational efforts will be required to delineate the possible relationship between peptide-dependent α2 helix motional properties and T cell recognition.

Pocket-forming residue rearrangement in response to anchor residue variation

Anchor residue variation in MHC class I-bound peptides frequently triggers structural rearrangement of heavy chain-derived pocket-lining residues – a phenomenon which is extensively documented for classical MHC class I alleles. For example, major rearrangements of 6 heavy chain peptide-interfacing residues with movements up to 6 Å, were observed in a comparison of 2 murine H-$2K^b$ structures bound to different viral peptides (Fremont et al., 1992). Further, repositioning of HLA-A2 heavy-chain-derived peptide-interfacing residues resulting from alternative peptide side chains is predominantly concentrated in the same α2 helical short arm region (residues 144-151), as in the H-$2K^b$ structures (Madden et al., 1993). Additionally, a coordinated structural rearrangement of two HLA-A2 pocket-forming residues, Arg-97 and Tyr-116, is observed in response to the absence of a downward-projecting side chain at position 7 of the peptide (Madden et al., 1993). Such rearrangements not only

facilitate optimal contacts between pocket-forming residues and the peptide anchor side chain, but also compensate for complex-destabilising cavities introduced by steric differences in accommodated peptide side chains (Xu et al., 1998), (Madden, 1995).

However, unlike classical MHC class I molecules, structural repositioning of pocket-forming residues in previous structures of VL9-bound HLA-E, have been highly limited. For example, despite a Met to smaller Thr substitution at position 2 of the HLA-B27-derived VL9 peptide variant and resultant expansion of a B pocket-resident cavity from 31-53 $Å^3$, no major positional rearrangements within the B pocket ensued (Strong et al., 2003). The lack of compensatory pocket-based reorganisation to fill this enlarged cavity correlates with lower thermal stability of HLA-E in complex with the HLA-B27 leader peptide. In turn, compromised stability is thought to account for reduced surface expression of HLA-B27 leader peptide-bound HLA-E (N Lee et al., 1998).

Consistent with structures of VL9-bound HLA-E, minimal reorganisation was observed in the primary B and F and secondary D and C pockets for both the Mtb44- and Mtb14-bound HLA-E structures. Minor rotamer readjustments within these pockets involve small (< 1 Å) movements which could arise from differential inter-structure refinement weights or electron density map interpretations. The only major pocket-lining residue perturbations observed for Mtb-derived peptide-bound HLA-E are concentrated in the same region in which the most dramatic classical MHC class I pocket reorganisation typically occurs - the α2 helix short arm (Madden et al., 1993). Glu-152 and Ser-147, situated in the E pocket of the Mtb44 and Mtb14-bound HLA-E structures, are displaced relative to their counterpart residues in VL9-bound HLA-E (1MHE).

Glu-152 repositioning is likely a direct result of peptide-specific interactions since it forms a hydrogen bond with position 5 Arg of the VL9 peptide but is not involved in the peptide-heavy chain hydrogen bonding network in either Mtb44 or Mtb14-bound structures of HLA-E.

Similarly, the reoriented Ser-147 side chain in Mtb14-bound HLA-E could stem from its unique involvement in the peptide-heavy chain hydrogen bonding network, in which it forms a novel hydrogen bond with Mtb14 position 7 Gln. Although Ser-147 does not participate in the peptide-heavy chain hydrogen bonding networks of Mtb44- or VL9-bound HLA-E, positional deviation of the Ser-147 side chain between these structures could arise from indirect peptide interactions. Specifically, in VL9-bound HLA-E the Ser-147 side chain might reorient in response to proximal Glu-152 repositioning to circumvent steric clashes. Thus, this could constitute an example of coordinated pocket-forming residue repositioning similar to that observed for Arg-97 and Tyr-116 in HLA-A2 (Madden et al., 1993).

In accordance with structurally limited pocket reorganisation, the HLA-E*0103 peptide binding groove exhibited superior rigidity relative to classical HLA-A*0201 in long time-scale molecular dynamics simulations (Hansen et al., 2016). In the absence of bound peptide the HLA-A*0201 binding groove collapses inward whereas the structural integrity of the HLA-E binding groove is better preserved. A rigid peptide binding groove in the absence of a high affinity, stabilising ligand is consistent with presentation of the large range of low affinity epitopes identified in RhCMV 68-1 SIV/HIV vaccine studies.

Despite limited pocket reorganisation in response to variable anchor residues, compensatory interactions with pocket-dwelling water molecules were formed in the HLA-E*0103-Mtb14 structure. Specifically, an E pocket resident water molecule, present in all four molecules of the HLA-E-Mtb14 asymmetric unit, formed hydrogen bonds with both the Mtb14 position 7 Gln side chain and the heavy chain-derived Ser-143 main chain. Such water-mediated hydrogen bonds indirectly anchor the peptide to the MHC heavy chain facilitating more optimal pocket interactions in addition to filling otherwise unoccupied pocket cavities - a phenomenon which, as previously discussed, has been linked to MHC class I complex instability (Strong et al., 2003) (Xu et al., 1998). Water molecules frequently operate as

hydrogen bonding intermediates connecting bound-peptide to various classical MHC class I heavy chains such as murine H-2Kb (Fremont et al., 2006) and human HLA-B53 (Smith et al., 1996a). This in turn yields more optimal pocket-based interactions and concomitantly permits the accommodation of a greater breadth of anchor residues (Smith et al., 1996a) - such as a polar position 7 Gln in the Mtb14 peptide in place of a canonical, hydrophobic Val or Leu.

Results chapter 3

HLA-E-restricted conformationally distinct peptide binding motifs and a biophysical analysis of suboptimal peptide binding to HLA-E

Introduction

VL9 peptide variants crystallised in complex with HLA-E adopt the previously described classical kinked conformation within the HLA-E binding groove (O'Callaghan et al., 1998) (Strong et al., 2003) (Hoare et al., 2008). Further, we have demonstrated that two HLA-E-binding Mtb-derived epitopes also assume this conformational motif, even when canonical position 2 Met and position 9 Leu primary anchor residues are substituted with other hydrophobic or polar residues (Walters et al., 2018). In this chapter we present crystal structures of HLA-E in complex with two additional pathogen-derived peptides: the Mtb-derived epitope, IL9 (IMYNYPAML), and the HIV-derived epitope, RL9HIV (RMYSPTSIL). We show that despite the presence of canonical primary anchor residues, each peptide adopts a distinct backbone configuration within the HLA-E binding groove with positional differences to the classical kinked motif adopted by VL9.

The IL9 peptide sequence comprises residues 4-12 of the Mtb-derived Low molecular weight protein antigen 7, EsxH. IL9 constitutes a naturally processed and presented epitope in Mtb infection with both unmodified and oxidised forms being identified in mass spectrometry analyses of the eluted HLA-E ligandome from Mtb-infected cells (McMurtrey et al., 2017). $CD8^+$ T cells from multiple donors latently infected with Mtb were subsequently shown to produce IFN-γ following IL9 peptide-pulsing of autologous dendritic cells. Although the HLA-E restriction of this epitope was not validated in the study, two of the three responding donors were negative for HLA-A2, which shares an overlapping peptide binding motif with HLA-E

(Parker et al., 1992). Additionally, IL9 yielded one of the greatest peptide signal to background ratios of any pathogen-derived peptide tested in the peptide exchange-based binding assay (Results Chapter 1, Figure 1.22). Thus, crystallisation trials were conducted for HLA-E refolded with the IL9 peptide.

In terms of HIV, the Gag-derived RL9HIV peptide is of particular physiological importance. First, RL9HIV elicits Mamu-E restricted $CD8^+$ T cell responses in rhesus macaques vaccinated with RhCMV 68-1 vectors recombinant for HIV Gag (Walters et al., 2018). Second, RL9HIV exhibits sequence alignment to RL9SIV (RMYNPTNIL) - one of two SIV-derived 'supertopes' recognised by Mamu-E restricted $CD8^+$ T cells in 100% of RhCMV 68-1-SIV vaccinated macaques (Hansen et al., 2016). Third, RL9HIV consistently demonstrated the strongest relative HLA-E binding of any HIV- or SIV-derived epitope tested in the optimised peptide exchange assay (Results Chapter 1, Figures 1.17a and 1.17b). Considering stronger HLA-E binding is not only typically correlated with the yield of correctly refolded material but also associated with complex stability, the RL9HIV peptide was selected for crystallographic analysis.

Alternative binding motif conformations assumed by IL9 and RL9HIV epitopes in the HLA-E peptide binding groove yielded disruptions in secondary pocket occupancy. Pocket-based cavities arising from displaced anchor side chains have previously been associated with compromised protein complex stability (Strong et al., 2003)(Xu et al., 1998), and might have contributed to the lower peptide binding assay signals relative to the canonical VL9 positive control (Results Chapter 1, Figures 1.22 and 1.17b). Surprisingly, RL9HIV-refolded and purified HLA-E complexes gave rise to two distinct bands, including a compact and a diffuse form, during blue native gel analyses performed by Professor Geraldine Gillespie, indicative of heterogeneity in the protein sample, whereas VL9-HLA-E complexes produced a single compact band likely reflecting a fully peptide-loaded homogeneous protein species (Walters et

al., 2018). Further, given the limited, lattice-constrained snapshot of protein structure afforded by crystallographic coordinates, an alternative biophysical approach was pursued to investigate sample heterogeneity and differential complex stability in solution when HLA-E is refolded with peptides of varying binding strength. Specifically, we used size exclusion chromatography-coupled small angle X-ray scattering to explore possible peptide affinity-related differences in the overall conformational ensemble of various HLA-E-β2M-peptide complexes.

Structural analysis

Structural determination of HLA-E*0103 in complex with the RL9HIV peptide

*HLA-E*0103-RL9HIV*

The extracellular domain of the HLA-E*0103 heavy chain and β2M light chain were crystallised in complex with the HIV-derived nonameric RL9HIV epitope. Crystallisation and refinement statistics are detailed in Figure 1. In summary, crystals grew in 2.4 M Ammonium sulphate, 0.1 M MES and 0.05% dichloromethane at pH 6, packed in the C2 space group and diffracted to 2.6 Å, yielding unit cell dimensions of a = 244.82 Å, b = 48.22 Å & c = 153.02. The dataset was 100% complete overall, with completeness only marginally reducing to 99.3% in the outer shell. Four molecules were present in the asymmetric unit which had a solvent content of 45%, as predicted by the Matthew's coefficient. Molecule 1 of the Mtb44-bound HLA-E structure discussed in Results Chapter 2 constituted a phasing model for molecular replacement. Refinement was subsequently carried out in Phenix yielding an R_{work} and R_{free} of 0.223 and 0.273, respectively. Following geometric validation using MolProbity the coordinates were deposited in the PDB and assigned the PDB ID, 6GL1.

The data set for HLA-E crystallised in complex with RL9HIV was formed by merging diffraction data collected at various rotations from multiple isomorphous crystals since individual crystals

were small in size and sensitive to radiation damage resulting in rapid deterioration of diffraction quality. Additionally, microscopic analysis of crystal morphology revealed an irregular appearance indicating that these were most likely mosaic, non-single crystals. Thus, potential irregularities within the crystal lattices may have contributed to a suboptimal R_{merge} value. The RL9HIV bound HLA-E data set also contained translational non-crystallographic symmetry (NCS) as indicated by Xtriage in Phenix (Adams et al., 2010), whereby molecules not related by crystallographic symmetry exhibit translational symmetry within the asymmetric unit. Due to resulting systematic changes to the diffraction pattern via constructive and deconstructive interference modulation, translational NCS violates assumptions made in structural determination and has hence been reported to complicate multiple stages of the crystallographic analysis including indexing and refinement (Read et al., 2013). Translational NCS is therefore likely a contributing factor which gave rise to the suboptimal crystallographic statistics for the RL9HIV HLA-E structure such as the slightly elevated final R_{work} and R_{free} values which exceed those obtained for structures with similar resolution limits such as Mtb44*P2-Phe and Mtb44*P9-Phe. With regard to potential indexing complications arising from translational NCS, an analysis was conducted in POINTLESS of the CCP4 suite to investigate the confidence of the assigned space group for RL9HIV-bound HLA-E (Evans, 2006) (Winn et al., 2011). This analysis yielded a probability of 1.0 for both space group and Laue group assignment and no probable alternative space groups were identified.

Although a low signal to noise ratio as indicated by I/sigma and a high R_{merge} was present in the outer shell for RL9HIV-bound HLA-E, the CC-half value exceeded the minimum threshold (>0.3) and thus the data set was not adjusted to exclude higher resolution shells. CC-half is a well-established and reproducible indicator of crystallographic data quality which is based on the Pearson correlation coefficient whereby reflection data are divided randomly into 2 subsets and the correlation coefficient between the average intensities of the subdivided data is calculated (Karplus and Diederichs, 2012). CC-half has superseded other indicators of data

quality such as I/sigma and R_{merge} which were previously used to determine optimal resolution limits. Traditionally, resolution limits were assigned by truncating shells for which the R_{merge} (or R_{meas} - a multiplicity-weighted version of R_{merge}) exceeded 0.6 or where the I/sigma value fell below 2 (Evans, n.d.). However, the relationship between optimal R_{merge} or R_{meas} values in the outer shells of a data set and ultimate model quality is poorly understood, whereas CC-half-guided resolution limit determination has been shown to correlate positively with model quality (Karplus and Diederichs, 2012). Further, it has been demonstrated that conservative data truncation according to the aforementioned R_{merge} and I/sigma thresholds can result in the elimination of useful data which would otherwise have contributed to superior model quality. For example, model quality was shown to improve following the inclusion of outer shell data even in cases where the R_{meas} was above 4.0 and the I/sigma was 0.3 – both such values considerably violate the commonly accepted cut-offs. Thus, CC-half was adopted as a determinant for data truncation as opposed to R_{merge} or I/sigma for all crystal structures presented in this thesis.

Disordered or absent electron density prohibited model building for 3 regions of the HLA-E heavy chain in molecule 4 (chain ID: G). As a result, heavy chain-derived residues 192-196, 217-224 and 247-258 were omitted from the deposited coordinates. Sufficiently clear electron density enabled model building for all remaining HLA-E heavy chain, β2M and peptide residues within the asymmetric unit.

Data collection	HLA-E*0103 in complex with peptide: RL9HIV (RMYSPTSIL)	IL9 (IMYNYPAML)
Crystallisation condition	2.4 M AS 0.1 M MES pH 6 0.05% dichloromethane	2.2 M AS 0.1 M MES pH 5.8
Resolution Å	135.87 - 2.62 (2.66 - 2.62)	114.03-1.72 (1.75-1.72)
Space Group	C2	C2
Cell dimensions	a = 244.82 b = 48.22 c = 153.02 α = 90.0° β = 117.38° γ = 90.0°	a =244.56 b = 48.53 c = 153.16 α = 90.0° β = 116.95° γ = 90.0°
Solvent content [%] (molecules per AU)	45 (4)	45 (4)
Unique reflections	48641 (2439)	171397 (8228)
Completeness [%]	100 (99.3)	99.8 (96.6)
R_{merge} [I]	0.31 (2.35)	0.12 (1.88)
I/sigma	5.2 (0.9)	8.1 (1.0)
Multiplicity	6.6 (6.7)	6.5 (5.5)
$CC_{1/2}$	0.98 (0.45)	1.0 (0.56)
Refinement		
No. of non-hydrogen atoms	12701	13404
R_{factor} [%)]	0.223 (0.303)	0.215 (0.332)
R_{free} [%]	0.273 (0.316)	0.241 (0.348)
r.m.s.d. bonds [Å] §	0.003	0.004
r.m.s.d. angles [deg]	0.62	0.75
Ramachandran statistics		
Favoured [%]	98.17	98.7
Disallowed [%]	0	0.2

Figure 1: Crystallographic determination, refinement and validation of RL9HIV and IL9-bound HLA-E*0103

Crystallographic data collection, determination, refinement and validation statistics for the HLA-E*0103 structures solved in complex with HIV-derived RL9HIV (RMYSPTSIL) and Mtb-derived IL9 (IMYNYPML) peptides.

AS: Ammonium sulphate. AU: asymmetric unit. § r.m.s.d.: root mean square deviation from ideal geometry. Statistics in parentheses indicate those deriving from the outer shell. R_{free}: The R factor representing 5% of reflections omitted from refinement.

HLA-E*0103-RL9HIV molecule alignment

Superposition of the four molecules from the asymmetric unit of the RL9HIV-bound HLA-E structure revealed strong alignment of peptide binding groove and β2M-derived secondary structure elements and surface loops (Figure 2, A (i)). However, multiple surface loops and both planes of β-sheets within the α3 domain exhibited considerable conformational variation. Considering the α3 domains of molecules 1 and 2 contain the greatest relative adjustments, with those of molecules 3 and 4 adopting intermediate positions, they shall form the basis of subsequent comparative analysis.

The α3 domain 220s loop exhibits dramatic re-orientation, similar to that observed between Mtb44- and Mtb14-bound molecules of HLA-E in Results Chapter 2 (Figure 2, A (ii)). The α-carbons of Glu-222, Gly-223 and His-224 in RL9HIV-bound HLA-E molecules 1 and 2 are displaced by 7.4 Å, 10.2 Å and 11.6 Å, respectively. Further, the conversion of residues Gln-226 to Val-231 into an extended β-strand is similar to that observed in molecule 1 of the Mtb14-bound HLA-E structure (Figure 2, B (ii)). This extended-form is absent in all molecules of the six previously crystallised VL9-bound structures of HLA-E, in which residues 226-228 are invariably incorporated into the flexible 220s loop (O'Callaghan et al., 1998) (Strong et al., 2003) (Hoare et al., 2008).

Structural repositioning is not limited to surface loops of the α3-domain. Distances exceeding 2 Å separate α carbons belonging to the β2M-distal α3-domain β-sheet plane. Specifically, 2.2

Å and 2.3 Å lie between α carbons of Tyr-257 and Gln-219, respectively (Figure 2, B (v)). Additionally, although the Val-270 α carbons of molecules 1 and 2 sit only 1.1 Å apart, those of Leu-272 are displaced by 1.5 Å (Figure 2, B (i)). Conformational shifts exceeding 3 Å are also present in the surface loop connecting the β2M-proximal and β2M-distal α3 domain β-sheet planes, such as 3.1 Å and 2.9 Å which separate α-carbons of Glu-253 and Gly-252, respectively (Figure 2, B (iv)). Furthermore, shifted β-strands are not exclusive to the β2M-distal plane, with both Val-249 and Pro-250 α carbons deviating by 2.3 Å in the β2M-proximal β-sheet plane of the α3 domain (Figure 2, B (iii)).

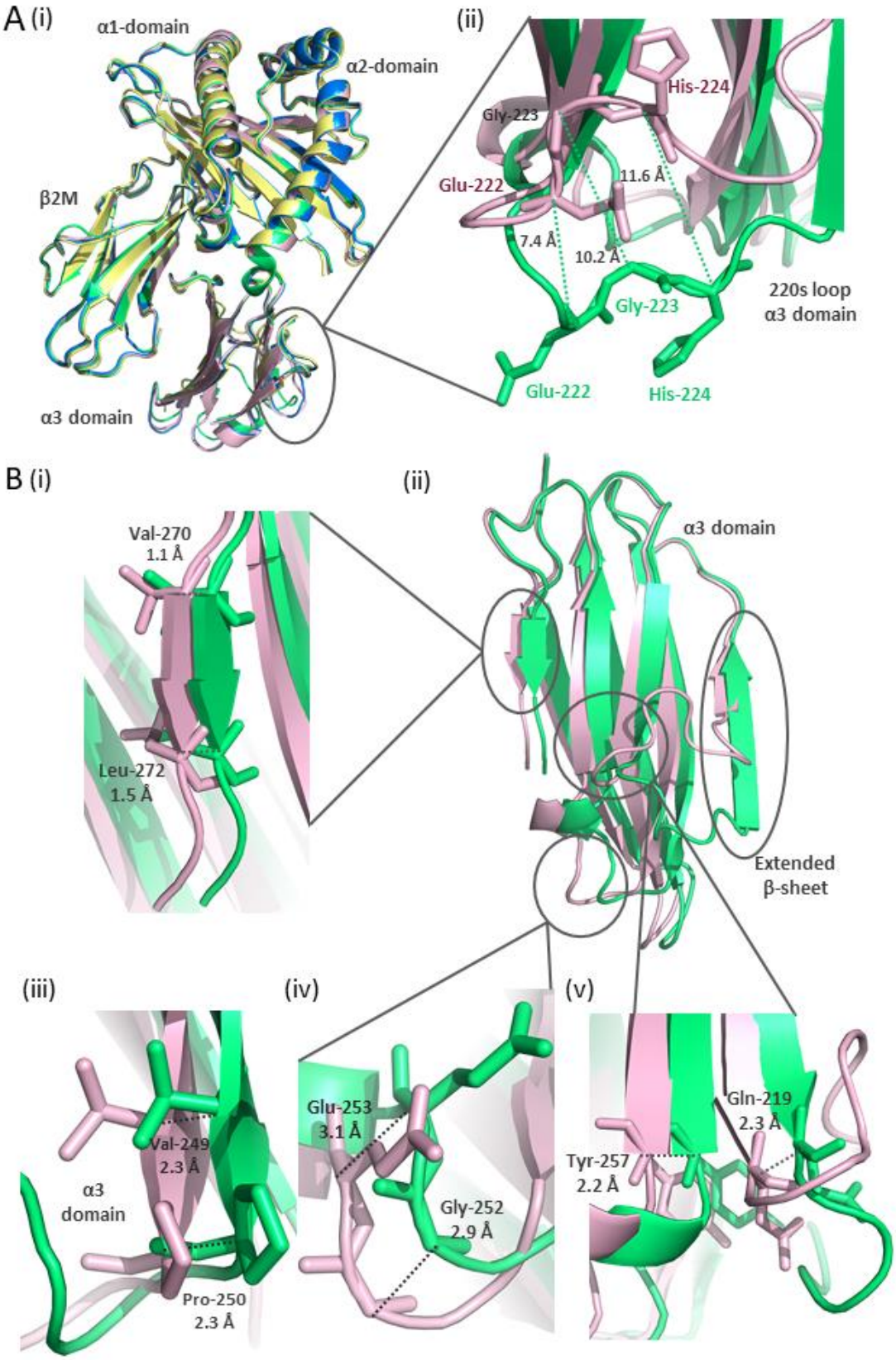


Figure 2: RL9HIV-bound HLA-E*0103 molecule alignment

A. (i) A cartoon visualisation of the 4 superimposed molecules present in the asymmetric unit of the

RL9HIV-bound HLA-E*0103 structure. Molecule 1 is depicted in lime-green, 2 in light-pink, 3 in blue and 4 in yellow. The flexible 220's loop of the α3-domain is circled. **(ii)** Magnification of the α3-domain-situated 220's loop for superimposed molecules 1 and 2, colour coded as per A (i). Loop-resident residues Glu-222, Gly-223 and His-224 are depicted as sticks, for which green dashed lines connect the α-carbons. Distances between α-carbons are specified in Å by grey labelling. **B. (i)** Magnification of a β2M-distal β-sheet region in the α3 domain of RL9HIV-bound HLA-E molecules 1 (lime-green) and 2 (light-pink). Residues Val-270 and Leu-272 are depicted as sticks, for which grey dashed lines connect the α-carbons and distances are denoted in Å. **(ii)** Visualisation of the α3-domain from superimposed molecules 1 and 2, colour coded according to B (i). Shifted regions are circled in grey in addition to the extended β-sheet situated in the 220's loop of molecule 1. **(iii)** Magnification of the terminus of a β-strand and surface loop in the β2M-proximal plane of the α3-domain, colour coded according to B (i). Val-249 and Pro-250 are presented as sticks and grey dashed lines separate the intra-structural α-carbons for which distances are denoted in Å. **(iv)** Magnified α3-domain-based surface loop connecting the β2M-proximal and -distal β-sheet planes, colour coded according to B (i). Gly-252 and Glu-253 are displayed as sticks and intra-structural α-carbons as grey lines, with distances labelled in Å. **(v)** Magnified region of the β2M-distal β-sheet plane of the α3-domain, colour coded as per B (i). The termini of 2 β-strands are displayed with residues Tyr-257 and Gln-219 displayed as sticks. Intra-structure α-carbon distances are denoted in Å by grey lines and adjacent labels.

Alignment of RL9HIV-HLA-E molecules to previous VL9-bound structures of HLA-E

Following superposition of RL9HIV-HLA-E molecules 1 and 2 with VL9-bound HLA-E molecule 1 from 1MHE (O'Callaghan et al., 1998), the region which exhibited the greatest structural repositioning comprised the α3 domain (Figure 3). The α3 domain β2M-distal β-sheet plane of the RL9HIV-bound HLA-E molecule 2 was shifted relative to that of VL9-bound HLA-E. Although this plane in the RL9HIV-bound HLA-E molecule 1 aligned more closely to that of VL9-bound HLA-E, the 220's loop between these two molecules exhibited greater readjustment. The RL9HIV-HLA-E 220's loop in molecule 1 is positioned closer to the membrane relative to that of VL9-bound HLA-E. In addition, the extended β-strand incorporating residues that typically form

part of the 220's loop in RL9HIV-HLA-E molecule 1, is absent in VL9-bound HLA-E and shorter in RL9HIV-HLA-E molecule 2.

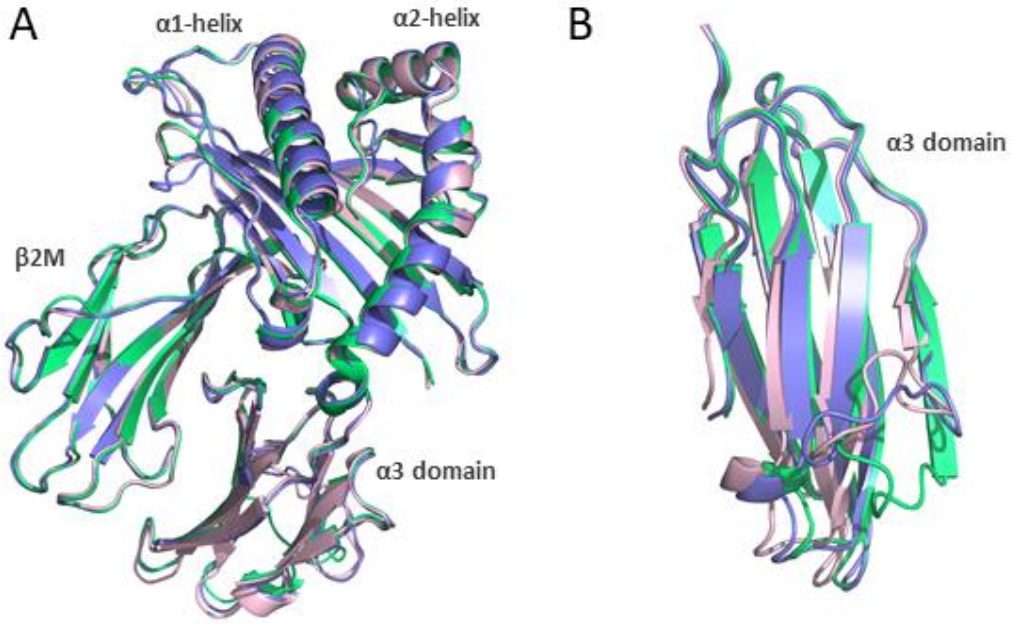


Figure 3: Superposition of RL9HIV-HLA-E molecules with VL9 peptide-bound HLA-E

A. Superposition of RL9HIV-bound HLA-E*0103 molecules 1 (green) and 2 (light-pink) with VL9-bound HLA-E*0101 molecule 1 of 1MHE (purple). **B.** Magnified α3 domain of superimposed RL9HIV-HLA-E molecules 1 (green) and 2 (light-pink) with VL9-HLA-E molecule 1 from 1MHE (purple).

RL9HIV peptide positioning within the HLA-E binding groove

Superimposed RL9HIV peptides from the four molecules of the asymmetric unit revealed minimal intra-structural perturbations with peptide Cα atoms deviating by less than 1 Å (Figure 4, A). Thus, a single RL9HIV peptide from molecule 1 was selected for downstream interpretation.

Electron density present in the HLA-E peptide binding groove enabled modelling of the RL9HIV peptide main chain and side chains (Figure 4, B). RL9HIV all-atom peptide B factors were considerably higher than those present in Mtb44 and Mtb14-bound HLA-E structures presented in Results Chapter 2 (Figure 4, E). All-atom peptide B factors of residues 4 and 5

rarely exceeded 50 $Å^2$ in Mtb44 and Mtb14, however the position 4 temperature factor of RL9HIV is around 10 $Å^2$ higher at 60 $Å^2$, indicative of greater disorder.

Cα backbone alignment to canonical VL9 peptide revealed striking conformational rearrangement of the RL9HIV peptide backbone within the HLA-E binding groove (Figure 4, C, D and F). Distortion of the classical kinked motif assumed by VL9 peptides in all previously determined crystal structures of HLA-E, is observed for RL9HIV. Although the amino and carboxy termini of RL9HIV superpose with those of VL9, RL9HIV exhibits C-terminally-shifted backbone arching. Classically, positions 4 and 5 of the peptide backbone arch away from the groove floor toward the solvent. However, α-carbons of positions 4 and 5 of the RL9HIV peptide are buried deeper in the groove by 2.8 Å and 2.6 Å, respectively, relative to canonical VL9. Alternatively, positions 6 and 7 of the RL9HIV peptide arch out of the groove and are concomitantly more solvent-exposed relative to the same region of the VL9 peptide (Figure 4, C-G). Specifically, the position 6 and 7 Cα atoms of RL9HIV are displaced by 4.1 Å and 3.4 Å respectively, relative to their more buried VL9 counterpart residues. Resultant perturbations in the RL9HIV solvent-exposure profile relative to canonical VL9 likely have implications for immunogenicity and T cell recognition.

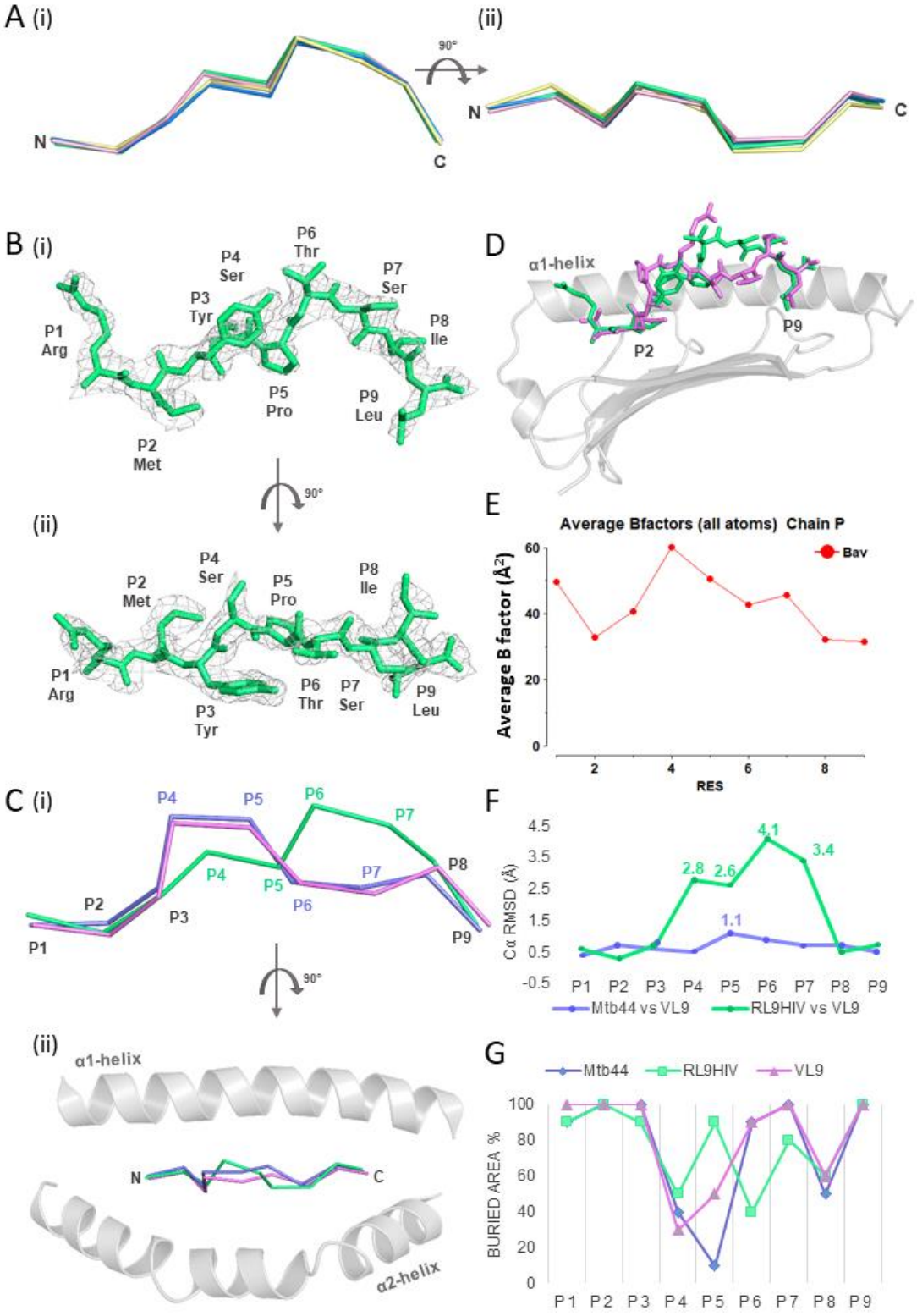


Figure 4: Peptide positioning in the RL9HIV peptide-bound HLA-E*0103 structure

A. Superimposed ribbon representation of the RL9HIV peptide Cα backbone from the 4 molecules of the asymmetric unit shown side-on **(i)** and from an aerial view, looking down on the peptide binding groove **(ii)**. Molecule 1 is shaded lime-green, molecule 2 light-pink, molecule 3 blue and molecule 4 yellow. **B.**

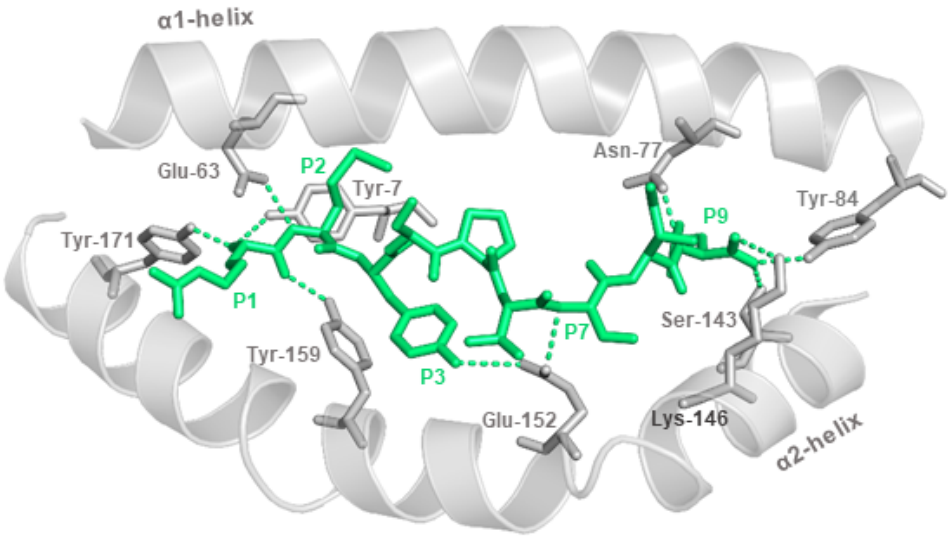
RL9HIV peptide from molecule 1 is depicted stick-form, shaded lime-green and visualised from the side **(i)** and from above **(ii).** An electron density map contoured to 1 sigma is displayed in grey mesh overlaying the peptide. **C.** Ribbon representation of the RL9HIV peptide Cα backbone from molecule 1 (lime-green) aligned to that of molecule 1 from the VL9-bound HLA-E structure, 1MHE (violet) and molecule 1 of the Mtb44 peptide-bound HLA-E structure, 6GH1 (purple). Visualisations are shown side-on **(i)** and from above, looking down on the peptide binding groove **(ii)**, with the α1 and 2 helices displayed in grey cartoon. **D.** The RL9HIV peptide from molecule 1 (lime-green) is shown in stick form, superimposed to the VL9 peptide (violet) from molecule 1 of 1MHE, with the α1 helix and β-sheet floor of the HLA-E peptide binding groove depicted in grey cartoon. **E.** All atom average B factors ($Å^2$) for the RL9HIV peptide, Chain P, of molecule 1. The Y-axis denotes average B factor including main chain and sidechain atoms for each peptide residue, whereas peptide position is plotted on the X-axis. **F.** Graph depicting α-carbon deviation between the RL9HIV and VL9 peptides (green) and Mtb44 and VL9 peptides (purple), for reference. The Y-axis denotes RMSD in Å between peptide Cα atoms whereas peptide position is plotted along the X-axis. **G.** Graph depicting solvent-accessible surface of the RL9HIV, VL9 (PDB ID: 1MHE) and Mtb44 (PDB ID: 6GH1) peptides. The percentage of buried area is plotted on the Y-axis with peptide position along the X-axis.

RES: Resolution. Bav: Average all-atom B factor.

Hydrogen bonding network between the RL9HIV peptide and HLA-E heavy chain

The alternative backbone configuration of the RL9HIV peptide results in a distinct peptide-heavy chain hydrogen bonding network depicted in Figure 5. In structures of VL9-bound HLA-E, the heavy chain-derived Glu-152 forms a hydrogen bond with position 5 Arg of the VL9 peptide. However, in RL9HIV-bound HLA-E Glu-152 forms alternative hydrogen bonds with both the position 3 Tyr side chain and position 7 Ser main chain of the RL9HIV peptide. Similar to the VL9-HLA-E inter-chain hydrogen bonding network, which is dominated by hydrogen bonds at the peptide termini, 7 of the 10 hydrogen bonds formed between the RL9HIV peptide and the HLA-E heavy chain involve position 1 Arg or position 9 Leu.

A



B

	RL9HIV Peptide	Distance (Å)	HLA-E Heavy chain
1	P1-Arg [N]	2.5	Tyr-171 [OH]
2	P1-Arg [N]	2.96	Tyr-7 [OH]
3	P1-Arg [O]	2.72	Tyr-159 [OH]
4	P2-Met [N]	2.82	Glu-63 [OE2]
5	P3-Tyr [OH]	3.46	Glu-152 [OE1]
6	P7-Ser [N]	3.28	Glu-152 [OE2]
7	P9-Leu [N]	2.88	Asn-77 [OD1]
8	P9-Leu [O]	2.73	Tyr-84 [OH]
9	P9-Leu [O]	2.84	Ser-143 [OG]
10	P9-Leu [OXT]	2.98	Lys-146 [NZ]

Figure 5: Peptide-heavy chain hydrogen bonding in the RL9HIV-HLA-E structure

A. Visualisation of the inter-chain hydrogen bonding network between the RL9HIV peptide and the HLA-E heavy chain. The RL9HIV peptide is shown as lime-green sticks with the heavy-chain-derived donor/acceptor residues as grey sticks. Hydrogen bonds are depicted as green-dashed lines. The α1 and 2 helices are visualised as grey cartoon with the remaining HLA-E heavy and light chain residues omitted for clarity. **B.** Table of hydrogen bonds formed between the RL9HIV peptide and HLA-E heavy chain. The lower and upper thresholds for the donor/acceptor distance range is set to 2.5 Å and 3.5 Å, respectively.

Pocket occupancy in the RL9HIV-HLA-E*0103 structure

Unsurprisingly, the canonical position 2 Met and position 9 Leu primary anchor side chains present in the RL9HIV peptide are well tolerated by the B and F pockets, respectively (Figure 6). Superposition of B and F pocket-forming residues from VL9- and RL9HIV-bound HLA-E structures revealed minimal side chain repositioning.

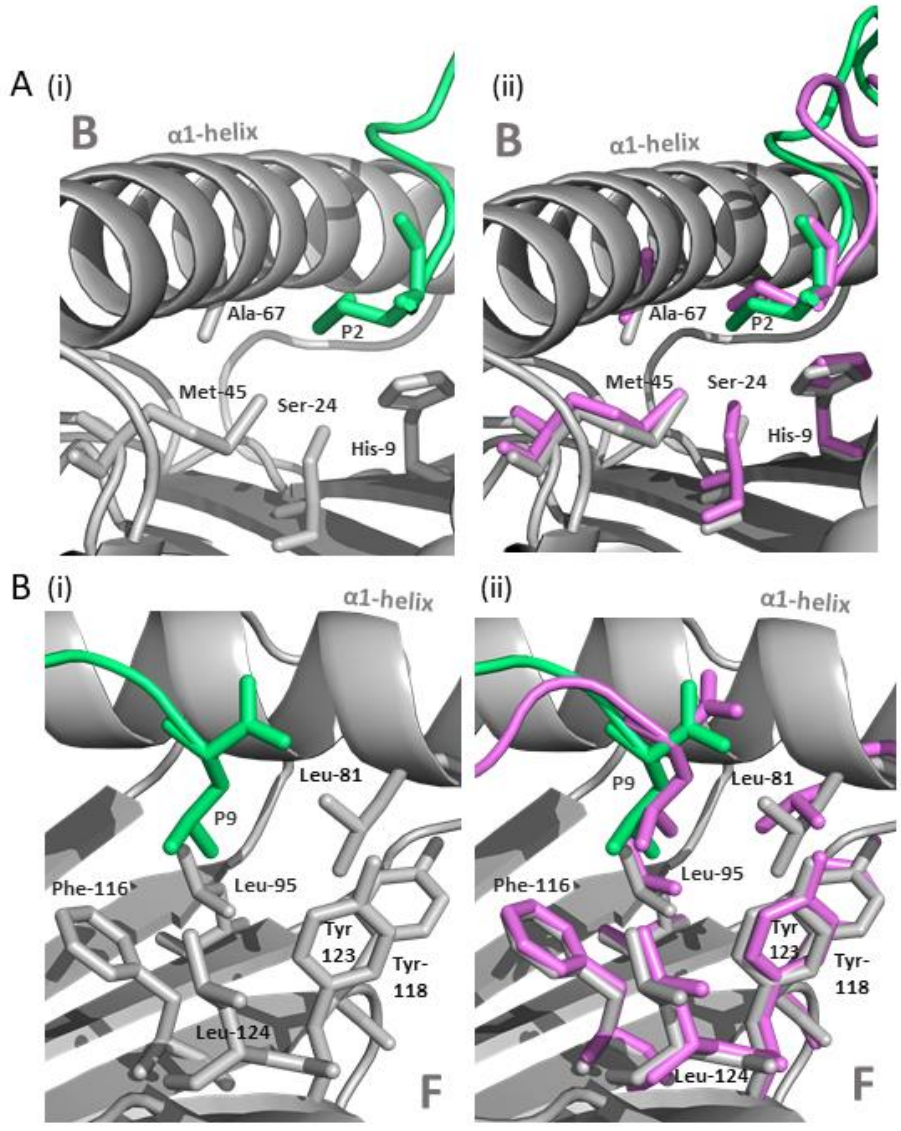


Figure 6: Primary B and F pocket occupancy by the RL9HIV peptide

A. (i) B pocket visualisation for RL9HIV-bound HLA-E*0103 molecule 1. Heavy chain-derived B pocket-forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as grey cartoon. The RL9HIV peptide is shown in green cartoon with position 2 Met in green stick-form. **(ii)**

B pocket visualisation of the superimposed RL9HIV-HLA-E molecule 1 and VL9-HLA-E molecule 1 from 1MHE. Colour-coding is according to A (i) with superimposed B pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 2 Met in stick-form. **B. (i)** Visualisation of the F pocket for RL9HIV-bound HLA-E*0103 molecule 1. Heavy chain-derived F pocket forming-residues are depicted as grey sticks with the surrounding area of the peptide binding groove as grey cartoon. The RL9HIV peptide is shown in green cartoon with position 9 Leu in green stick-form. **(ii)** F pocket visualisation of the superimposed RL9HIV-HLA-E molecule 1 and VL9-HLA-E molecule 1 from 1MHE. Colour-coding is according to A (i) with superimposed F pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 9 Leu in stick-form.

Although minimal reorganisation was observed between C pocket-forming residues of the VL9- and RL9HIV-bound HLA-E structures, the D pocket-resident His-99 adopted an alternative position in RL9HIV-bound HLA-E (Figure 7, A and B). Such repositioning likely derives from distinct intra-chain hydrogen bonding in RL9HIV-bound relative to VL9-bound HLA-E. Projection of the His-99 side chain further in the C-terminal direction of the RL9HIV peptide positions it closer to Gln-156 facilitating hydrogen bonding. By contrast, no such hydrogen bond is formed between heavy chain-derived His-99 and Gln-156 side chains in either molecule of the VL9-bound HLA-E structure, 1MHE (O'Callaghan et al., 1998).

Structural repositioning of E pocket forming residues, Ser-147 and Glu-152, is observed in the RL9HIV-bound structure relative to VL9-bound HLA-E (Figure 7, C). Both Ser-147 and Glu-152 side chains in the RL9HIV-HLA-E structure superpose to those from the Mtb14- and Mtb44-bound structures of HLA-E. This further supports the possibility of coordinated Ser-147 and Glu-152 rearrangement to facilitate hydrogen bonding between HLA-E heavy chain-derived Glu-152 and position 5 Arg of the VL9 peptide, as discussed in Results Chapter 2 (O'Callaghan et al., 1998). Further, the alternative involvement of Glu-152 in the peptide-heavy chain

hydrogen bonding network, in which it forms a novel bond with the position 3 Tyr side chain of the RL9HIV peptide, might also constitute a driving factor in Glu-152 structural repositioning.

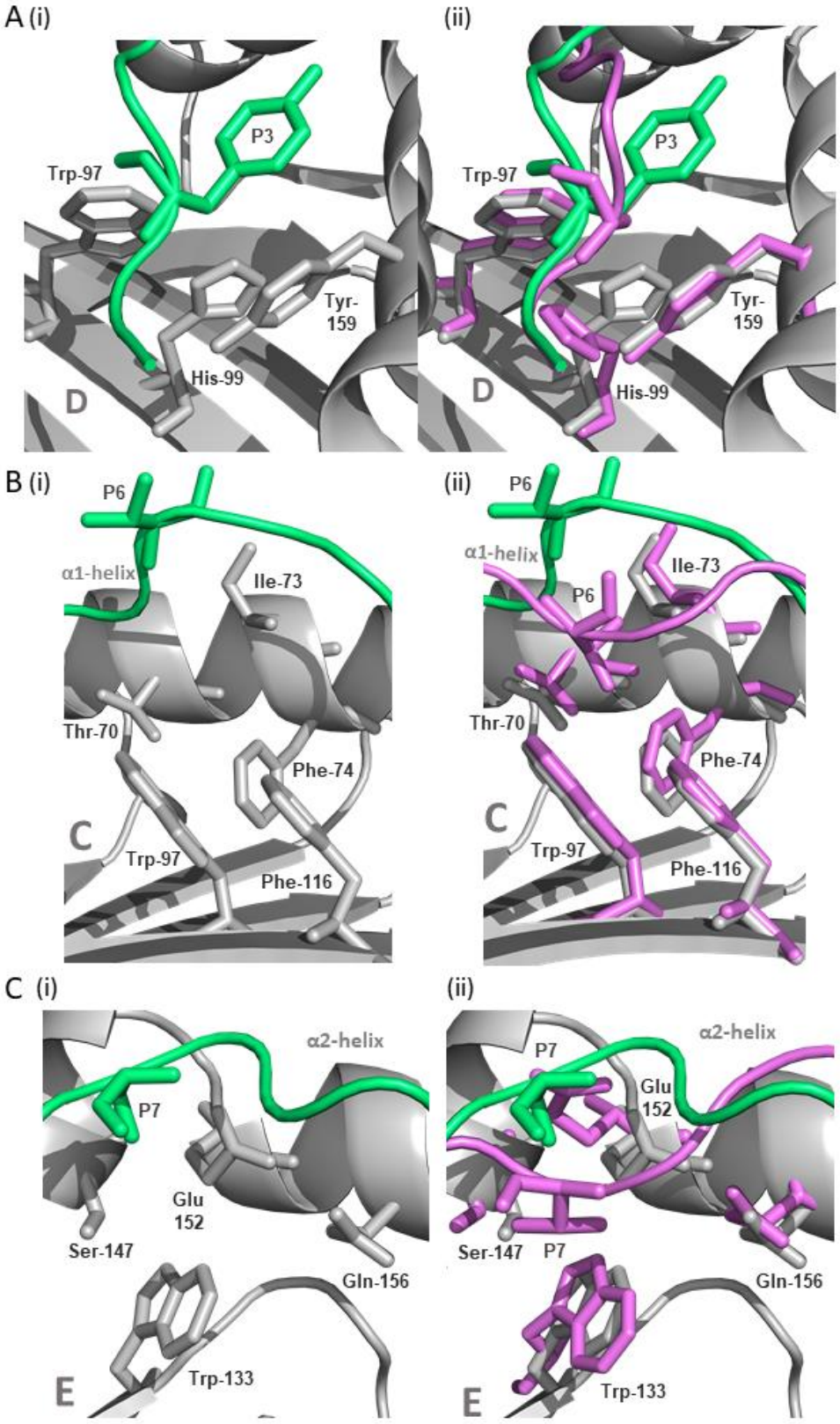


Figure 7: Secondary D, C and E pocket occupancy by the RL9HIV peptide

A. (i) D pocket visualisation for RL9HIV-bound HLA-E*0103 molecule 1. Heavy chain-derived D pocket-forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as

grey cartoon. The RL9HIV peptide is shown in green cartoon with position 3 Tyr in green stick-form. **(ii)** D pocket visualisation of the superimposed RL9HIV-HLA-E molecule 1 and VL9-HLA-E molecule 1 from 1MHE. Colour-coding is according to A (i) with superimposed D pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 3 Ala in stick-form. **B. (i)** Visualisation of the C pocket for RL9HIV-bound HLA-E*0103 molecule 1. Heavy chain-derived C pocket forming-residues are depicted as grey sticks with the surrounding area of the peptide binding groove as grey cartoon. The RL9HIV peptide is shown in green cartoon with position 6 Thr in green stick-form. **(ii)** C pocket visualisation of the superimposed RL9HIV-HLA-E molecule 1 and VL9-HLA-E molecule 1 from 1MHE. Colour-coding is according to B (i) with superimposed C pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 6 Thr in stick-form. **C. (i)** E pocket visualisation for RL9HIV-bound HLA-E*0103 molecule 1. Heavy chain-derived E pocket-forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as grey cartoon. The RL9HIV peptide is shown in green cartoon with position 7 Ser in green stick-form. **(ii)** E pocket visualisation of the superimposed RL9HIV-HLA-E molecule 1 and molecule 1 of 1MHE. Colour-coding is according to C (i) with superimposed E pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 7 Val in stick-form.

Structural determination of HLA-E*0103 in complex with the IL9 peptide

*HLA-E*0103-IL9*

The extracellular domain of the HLA-E*0103 heavy chain and β2M light chain were crystallised in complex with the Mtb-derived nonameric IL9 epitope. Crystallisation and refinement statistics are detailed in Figure 1. In summary, crystals grew in 2.2 M Ammonium sulphate, 0.1 M MES at pH 5.8 and diffracted to 1.7 Å. Molecules packed in the C2 space group and yielded a unit cell with similar dimensions to that observed for crystals of RL9HIV-bound HLA-E (a = 244.56 Å, b = 48.53 Å & c = 153.16). Accordingly, similar to the RL9HIV-HLA-E structure, four molecules were present in the asymmetric unit, which comprised 45% solvent. The dataset

was 99.8% complete overall, with marginal reductions to 96.6% in the outer shell. Molecule 1 of the Mtb44-bound HLA-E structure discussed in Results Chapter 2 constituted a phasing model for molecular replacement. Refinement was subsequently carried out in Phenix yielding an R_{work} and R_{free} of 0.215 and 0.241, respectively.

Clear electron density enabled model building of every heavy chain, β2M and peptide-derived residue in the four molecules present in the asymmetric unit except the α3-domain of molecule 3. Disordered electron density resulted in the omission of residues 191-200, 218-228 and 247-254 from chain E – such regions exhibit substantial overlap with residues also absent in molecule 3 of the RL9HIV-HLA-E structure. This is perhaps unsurprising given both structures packed in the same space group and yielded similar unit cell dimensions. Additionally, density for two Zinc ions was present in the asymmetric unit which exhibited structural coordination with HLA-E heavy-chain-derived Glu-183 and His-181 side chains of molecules 3 and 4, in a similar fashion to that observed for the Mtb44-HLA-E structure.

Alignment of HLA-E*0103-IL9 molecules from the asymmetric unit

Peptide binding groove and β2M-derived secondary structure elements and surface loops exhibited strong alignment between IL9-bound HLA-E molecules (Figure 8, A). However, multiple surface loops and both planes of β-sheets within the α3 domain exhibited conformational variation. Similar to the RL9HIV-HLA-E structure, the greatest adjustment was observed between the α3 domains of molecules 1 and 2, with intermediate positions seen for molecules 3 and 4. Molecules 1 and 2 shall therefore constitute the focus of subsequent discussion.

The α3 domain 220s loop is dramatically re-orientated in molecule 1, relative to the remaining molecules in the asymmetric unit. For example, the α-carbons of Glu-222, Gly-223 and His-224 in molecule 1 are respectively displaced by 7.8 Å, 10.3 Å and 11.3 Å, relative to molecule 2 (Figure 8, A, (ii)). Similar to molecule 1 of both the RL9HIV-HLA-E and Mtb14-HLA-E structure,

residues Gln-226 to Val-231 of IL9-bound molecule 1 assemble into an elongated β-strand (Figure 8, B (ii)).

Structural repositioning extends beyond surface loops to both the β2M-proximal and -distal β-sheet planes of the α3 domain. For example, 2.5 Å and 2.3 Å respectively displace α carbons of Tyr-257 and Gln-219, which lie at β-strand termini in the β2M-distal β-sheet plane (Figure 8, B (v)). Additionally, greater distances (1.5 Å and 2.0 Å, respectively) separate Val-270 and Leu-272 α-carbons of molecules 1 and 2 in the IL9-HLA-E structure relative to those in the RL9HIV structure (Figure 8, B (i)). β-strand shifts situated on the β-sheet plane in closest proximity to β2M approach 3 Å, with both Val-249 and Pro-250 α-carbons deviating by 2.9 Å and 2.7 Å, respectively (Figure 8, B (iii)). Conformational shifts measuring 3.5 Å are also present in the surface loop connecting the β2M-proximal and -distal β-sheet planes of the α3 domain, such as Glu-253 and Gly-252 α-carbons which are separated by 3.5 Å and 3.4 Å, respectively (Figure 8, B, (iv)).

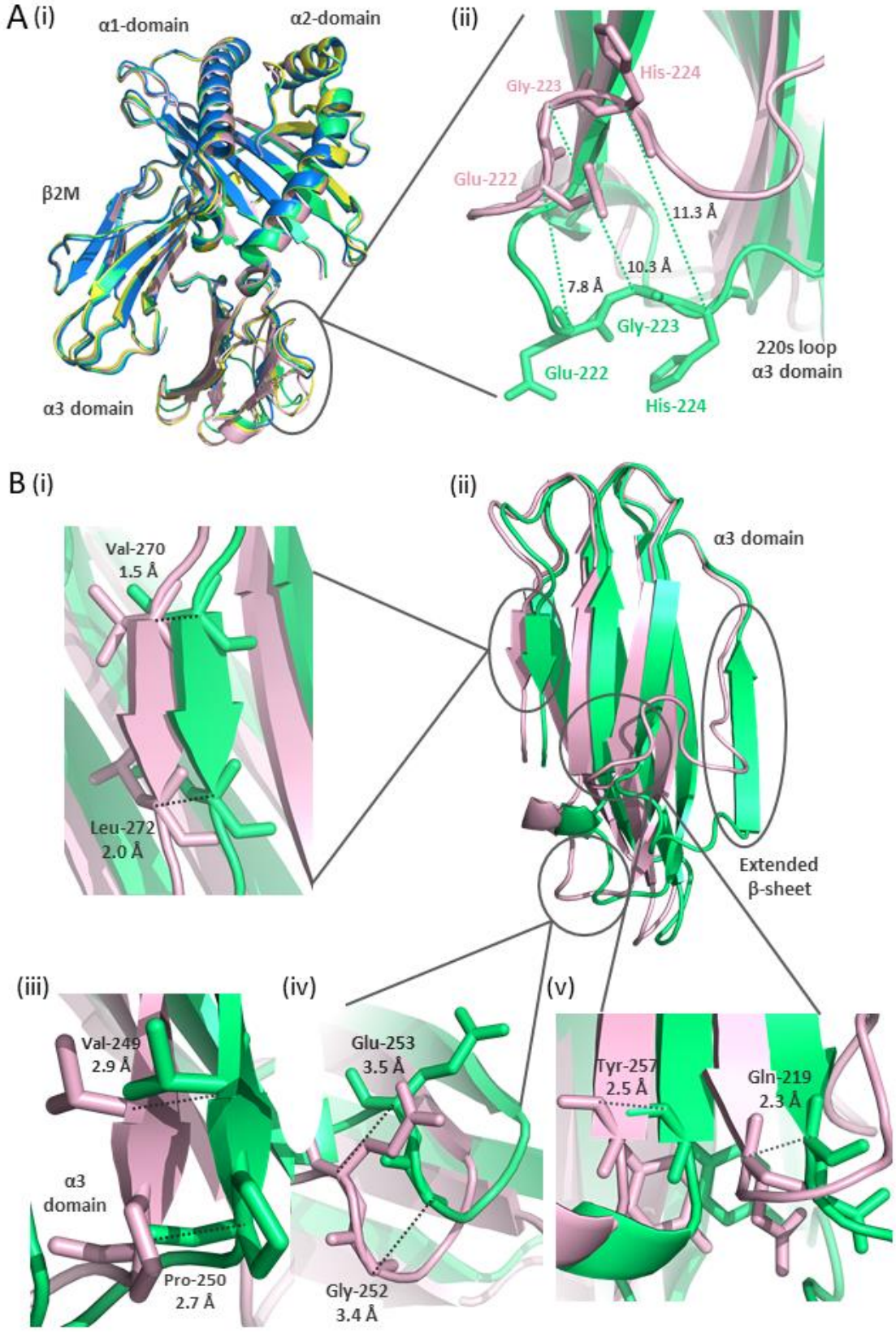


Figure 8: IL9-bound HLA-E*0103 molecule alignment

A. (i) A cartoon visualisation of the 4 superimposed molecules present in the asymmetric unit of the IL9-

bound HLA-E*0103 structure. Molecule 1 is depicted in lime-green, 2 in light-pink, 3 in blue and 4 in yellow. The flexible 220's loop of the α3-domain is circled. **(ii)** Magnification of the α3-domain-situated 220's loop for superimposed molecules 1 and 2, colour coded as per A (i). Loop-resident residues Glu-222, Gly-223 and His-224 are depicted as sticks, for which green dashed lines connect the α-carbons. Distances between α-carbons are specified in Å by grey labelling. **B. (i)** Magnification of a β2M-distal β-sheet region in the α3 domain of IL9-bound HLA-E molecules 1 (lime-green) and 2 (light-pink). Residues Val-270 and Leu-272 are depicted as sticks, for which grey dashed lines connect the α-carbons and distances are denoted in Å. **(ii)** Visualisation of the α3-domain from superimposed molecules 1 and 2, colour coded according to B (i). Shifted regions are circled in grey in addition to the extended β-sheet situated in the 220's loop of molecule 1. **(iii)** Magnification of a deviant β-sheet region in the β2M-proximal plane of the α3-domain, colour coded according to B (i). Val-249 and Pro-250 are presented as sticks and grey dashed lines separate the inter-molecule α-carbons for which distances are denoted in Å. **(iv)** Magnified α3-domain-based surface loop connecting the β2M proximal and distal β-sheet planes, colour coded according to B (i). Gly-252 and Glu-253 are displayed as sticks with inter-molecule α-carbon distances shown as grey lines and measured in Å. **(v)** Magnified region of the β2M-distal β-sheet plane of the α3-domain, colour coded as per B (i). The termini of 2 β-strands are displayed with residues Tyr-257 and Gln-219 displayed as sticks. Inter-molecule α-carbon distances are denoted in Å by grey lines and adjacent labels.

Alignment of IL9-HLA-E molecules to previous VL9-bound structures of HLA-E

Following superposition of IL9-HLA-E molecules 1 and 2 with VL9-bound HLA-E from 1MHE (O'Callaghan et al., 1998), the regions which exhibited the greatest structural deviation comprised the α3 domain and the central region of the α2 helix (Figure 9). The β2M-distal β-sheet plane of the IL9 molecule 2 α3 domain was shifted relative to that of VL9-bound HLA-E (Figure 9, B). Although this plane in IL9-HLA-E molecule 1 aligned more closely to that of VL9-bound HLA-E, the 220's loop between these two molecules exhibited greater repositioning. The IL9 220's loop in molecule 1 is positioned closer to the membrane relative to that of VL9-bound HLA-E. In addition, an extended β-strand incorporating residues which typically

comprise part of the 220's loop is formed in IL9 molecule 1; this is absent in both VL9-bound HLA-E and IL9-HLA-E molecule 2.

The central region of the α2 helix is shifted in both IL9 molecules 1 and 2 relative to VL9-bound HLA-E (Figure 9, C). For example, 1.6 Å and 1.5 Å separate the Ala-153 and Glu-154 Cα atoms between IL9-bound and VL9-bound HLA-E, respectively. Resultantly the central portion of the peptide binding groove is slightly wider in IL9-bound HLA-E relative to VL9-bound HLA-E.

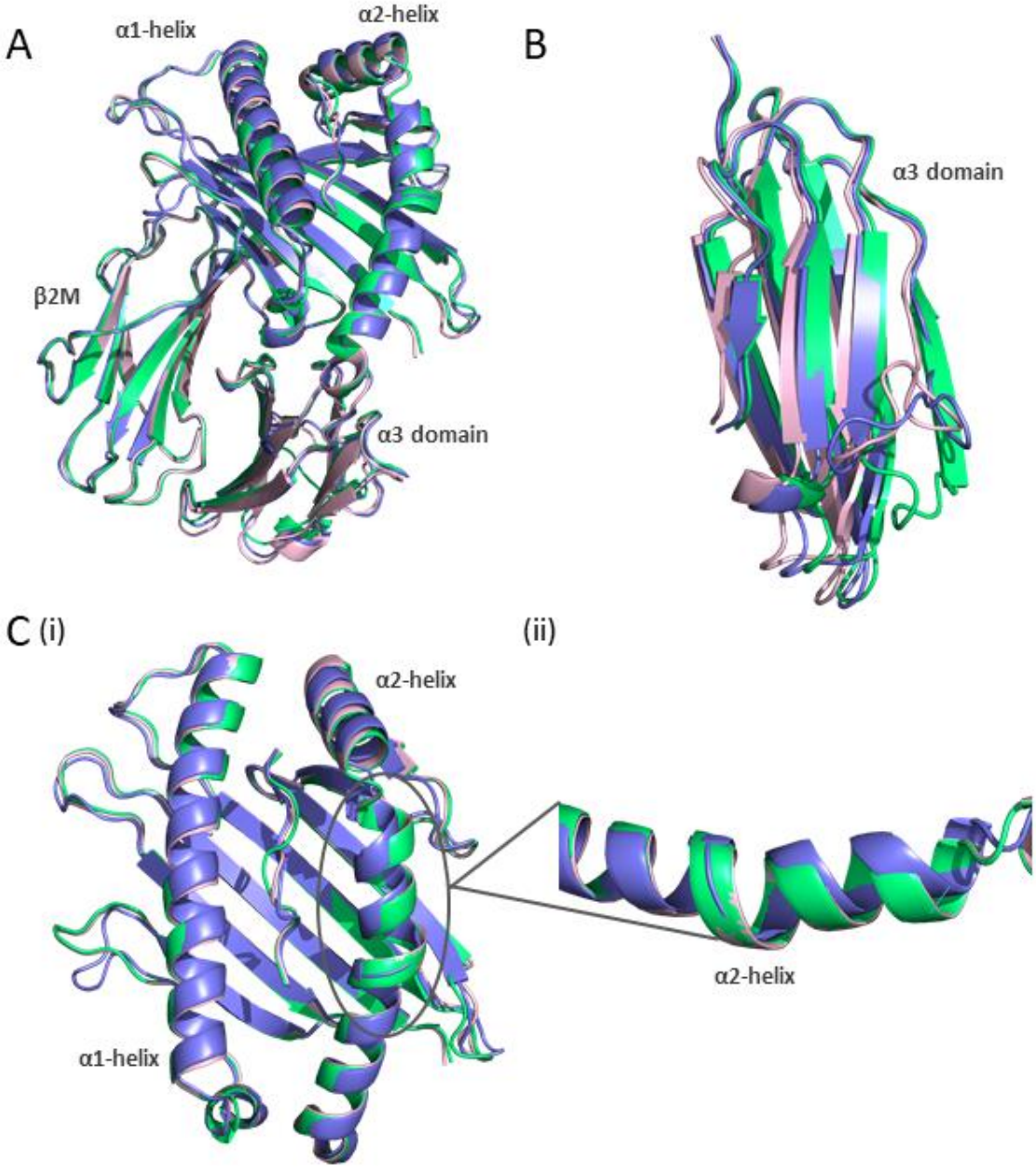


Figure 9: Superposition of IL9-HLA-E molecules with VL9 peptide-bound HLA-E

A. Superposition of IL9-bound HLA-E*0103 molecules 1 (green) and 2 (light-pink) with VL9-bound HLA-E*0101 molecule 1 of 1MHE (purple). **B.** Magnified α3 domain of superimposed IL9-HLA-E molecules 1 (green) and 2 (light-pink) with VL9-HLA-E molecule 1 from 1MHE (purple). **C. (i)** Peptide binding groove of superimposed IL9-HLA-E molecules 1 (green) and 2 (light-pink) with VL9-HLA-E molecule 1 from 1MHE (purple). **(ii)** Magnified central region of the α2 helix colour coded according to C (i).

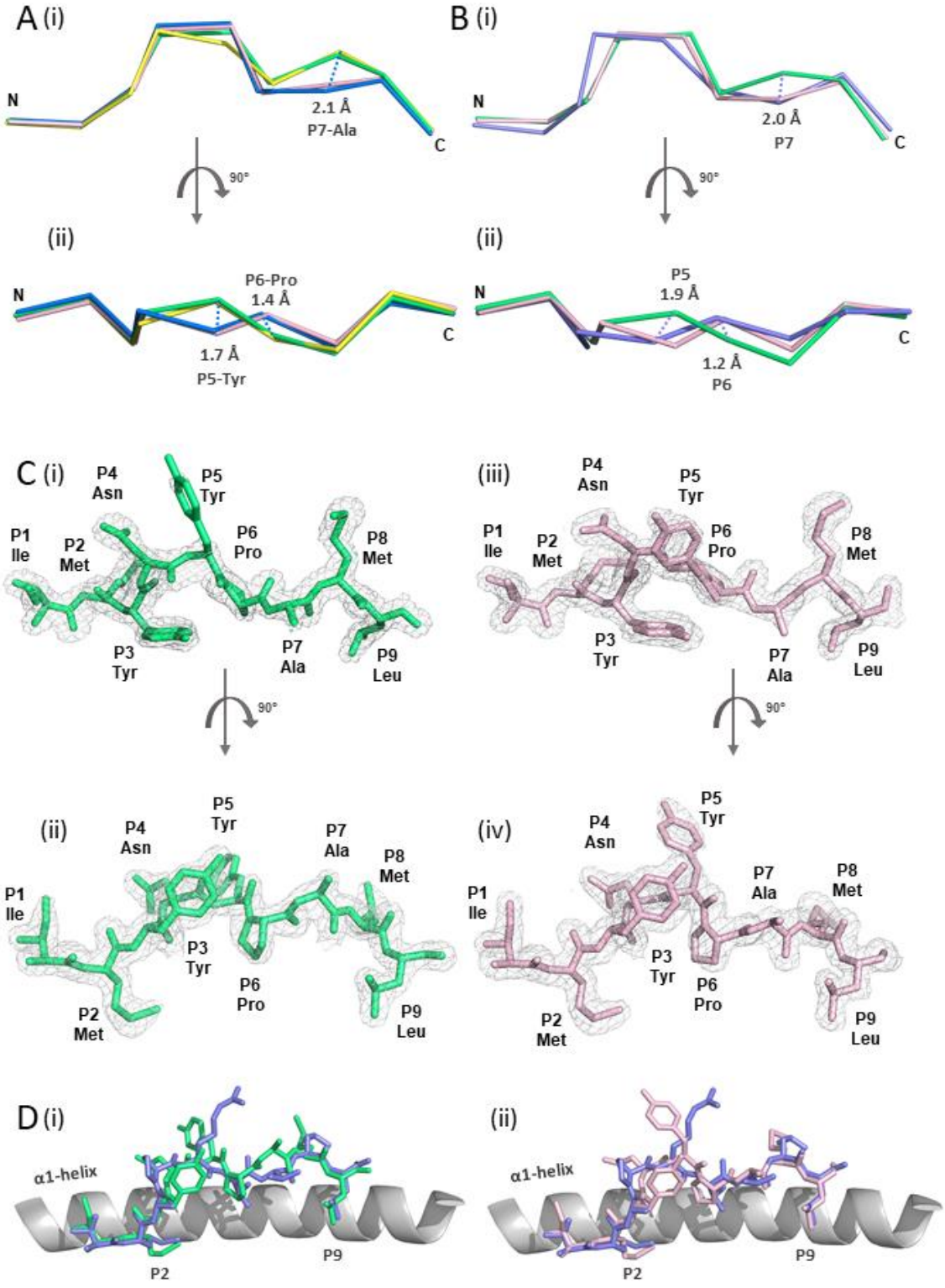
IL9 peptide positioning within the HLA-E binding groove

Superposition of the four molecules within the HLA-E*0103-IL9 asymmetric unit revealed two distinct peptide configurations (Figure 10, A). The IL9 peptide of molecules 2 and 3 exhibited strong alignment and adopted a conformation similar to the classical binding motif of VL9 peptides (Figure 10, B). However, the IL9 peptides of molecules 1 and 4 adopted an alternative motif, with up to 2.0 Å main chain deviation from the classical VL9 configuration. Thus, IL9 is capable of adopting two distinct orientations within the HLA-E peptide binding groove. Clear electron density was present in the groove for both such peptide configurations (Figure 10, C).

One IL9 peptide configuration resembles an intermediate motif situated between the RL9HIV and VL9 backbones encompassing peptide residues 6 to 8 (Figure 10, B). Where the RL9HIV main chain lies 3.4 Å above the VL9 backbone at position 7, the IL9 position 7 α-carbon is elevated, to a lesser degree, by 2.0 Å relative to that of VL9. However, contrary to the RL9HIV peptide binding motif, in which the RL9HIV main chain lies 4.1 Å above VL9 at position 6, the IL9 position 6 α-carbon converges with that of VL9.

Given the strong positional alignment of IL9 peptides from molecules 1 and 4 and from molecules 2 and 3, a single peptide representing each distinct configuration was selected for downstream analysis. Namely, IL9 peptides from molecules 1 (representative of configuration 1) and 2 (representative of configuration 2) shall constitute the basis of further discussion. Overall, the IL9 configuration 1 exhibited higher all-atom B factors, with 5 residues exceeding

40 Å^2 (Figure 10, E). By contrast, a single residue, in the kinked region at position 5, exceeded 40 Å^2 in the all-atom B factor analysis of IL9 peptide configuration 2.



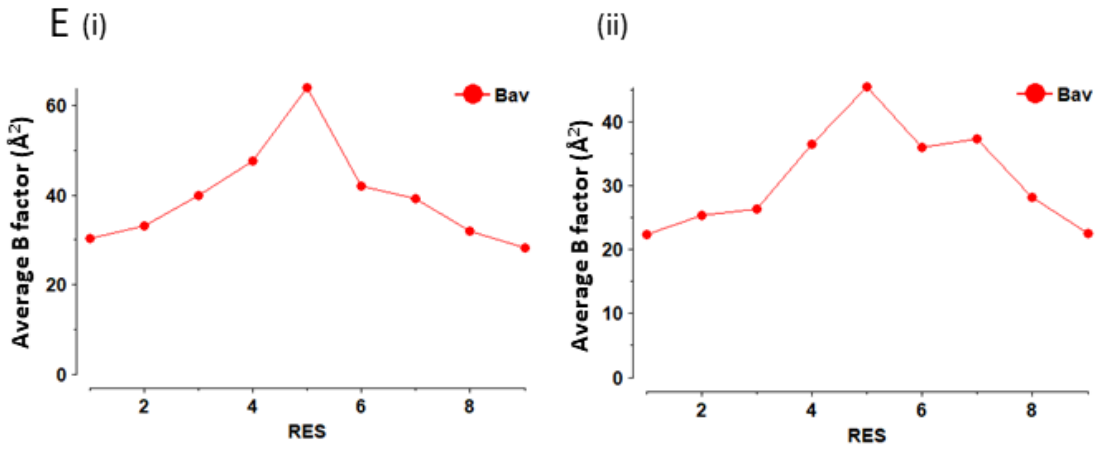


Figure 10: Peptide positioning in the IL9 peptide-bound HLA-E*0103 structure

A. Superimposed ribbon representation of the IL9 peptide Cα backbone from the 4 molecules of the asymmetric unit shown side-on **(i)** and from above, looking down on the peptide binding groove **(ii)**. The greatest α-carbon deviation between the 4 peptides at positions 5, 6, and 7 is indicated by blue dashed lines and denoted in Å. Molecule 1 is shaded lime-green, molecule 2 light-pink, molecule 3 blue and molecule 4 in yellow. **B.** Ribbon representation of the IL9 peptide Cα backbone from molecules 1 (lime-green) and 2 (light-pink) aligned to that of molecule 1 from the VL9-bound HLA-E structure, 1MHE (purple). Visualisations are shown side-on **(i)** and from above, looking down on the peptide binding groove **(ii)**. The greatest α-carbon deviation between the IL9 and VL9 peptide backbones at positions 5, 6, and 7 is indicated by purple dashed lines and denoted in Å. **C.** IL9 peptide from molecule 1 is depicted stick-form, shaded lime-green and visualised from the side **(i)** and from above **(ii).** An electron density map contoured to 1 sigma is displayed in grey mesh overlaying the peptide. IL9 peptide from molecule 2 is depicted stick-form, shaded light-pink and visualised from the side **(iii)** and from above **(iv).** An electron density map contoured to 1 sigma is displayed in grey mesh overlaying the peptide. **D. (i)** The IL9 peptide from molecule 1 (lime-green) is shown in stick form, superimposed to the VL9 peptide (violet) from molecule 1 of 1MHE, with the α1 helix of the HLA-E heavy chain depicted in grey cartoon. The alpha-2 helix is omitted for clarity. **(ii)** The IL9 peptide from molecule 2 (light-pink) is shown in stick form, superimposed to the VL9 peptide (violet) from molecule 1 of 1MHE, with the α1 helix of the HLA-E heavy chain depicted in grey cartoon. **E (i)** All atom B factors (Å^2) for the IL9 peptide, Chain P, of molecule 1. The Y-axis denotes average B factor including main chain and sidechain atoms for each peptide residue, whereas peptide position is plotted on the X-axis. **(ii)** All atom B factors (Å^2) for the IL9

peptide, Chain Q, of molecule 2. The Y-axis denotes average B factor, whereas peptide position is plotted on the X-axis.

RES: Resolution. Bav: All-atom average B factor.

The molecules present in the asymmetric unit of IL9-bound HLA-E packed in a manner which placed the peptide binding groove of one molecule facing the groove on an adjacent non-crystallographic symmetry-related molecule. Peptide packing contacts were therefore investigated as such crystal contacts could potentially contribute to the conformational orientation of interfacing residues belonging to the peptide. For IL9-bound HLA-E, it is possible such contacts may favour peptide conformational dimorphism in which two peptides within the groove adopt a certain configuration with the remaining two adopting a different orientation.

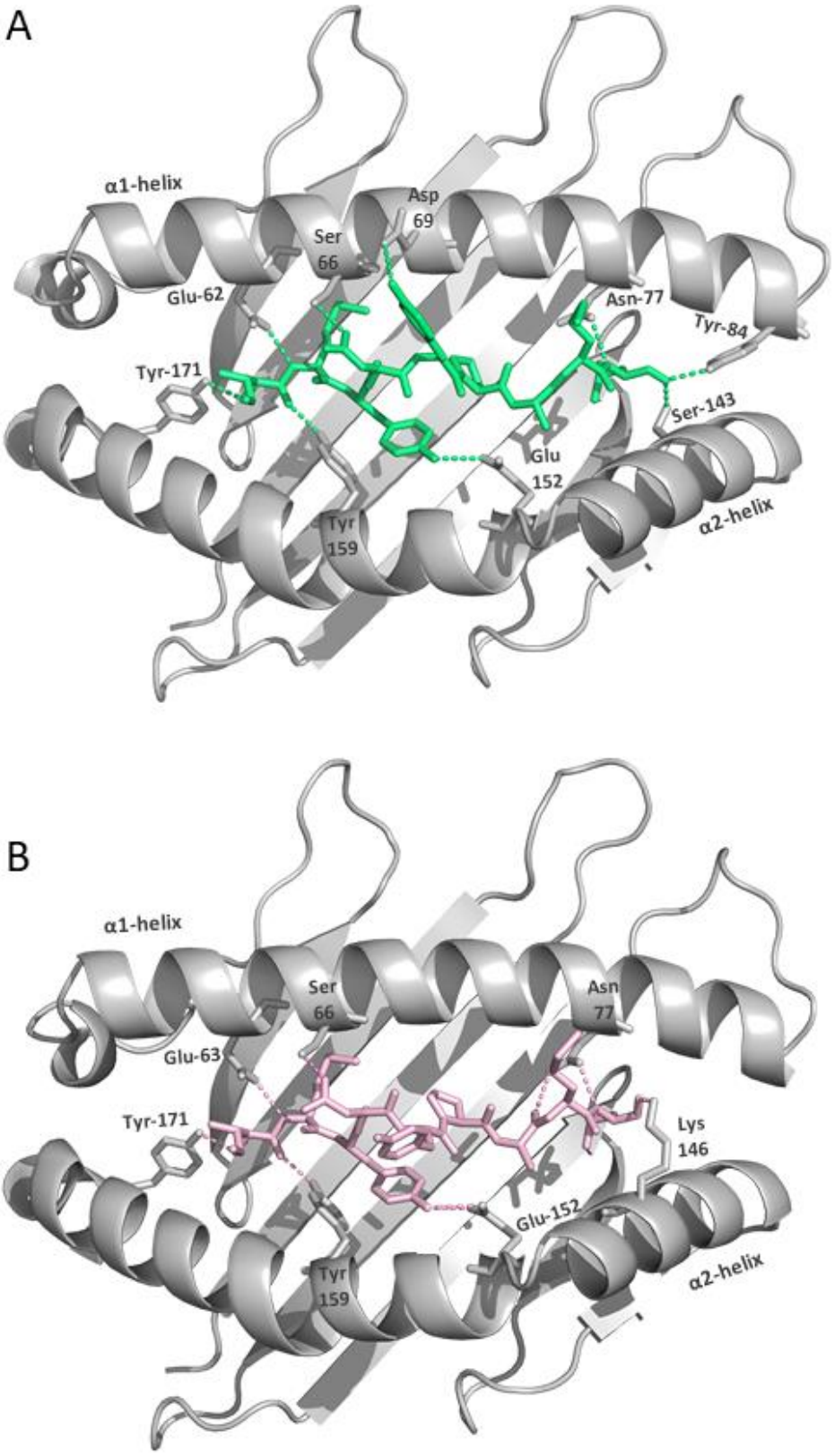
On the other hand, this mode of molecular assembly within the asymmetric unit is particularly common among structures of MHC class I proteins. For example, the two molecules in the asymmetric unit of HLA-E bound to VL9 (PDB ID: 1MHE) pack in an identical fashion with the peptide binding grooves directly facing one another. A subsequent structure of the HLA-E*0103 variant in complex with a VL9 peptide (PDB ID: 1KTL) also packed in this way in addition to the other HLA-E*0103 structures presented in this thesis bound to non-VL9 peptides such as Mtb44 and RL9HIV. Beyond HLA-E, many MHC class I structures exhibit the same mode of packing in which the peptide is involved in the crystal packing interface between adjacent symmetry molecules or additional molecules within the asymmetric unit. Examples of MHC class I allelic variants that assemble within the crystal lattice in a fashion which results in interfacing peptide binding grooves include HLA-A3*01 (PDB ID: 6O9B), HLA-A*0201 (PDB ID: 6NCA), HLA-C*0602 (PDB ID: 5W69) and HLA-A24 (PDB ID: 4WU5). In summary, this mode of molecular assembly is not unique to IL9-bound HLA-E and thus it is unlikely that crystal packing contacts are of particular relevance to this structure relative to other published structures of HLA-E and classical MHC class I proteins.

Further, packing interfaces involving the peptide in both structures of HLA-E bound to VL9 (1MHE) and IL9 are similar in size and strength. For both the two VL9 peptides present in 1MHE and for the four IL9 peptides present within the asymmetric unit, either 4 or 5 peptide-derived amino acids are classified as 'interfacing residues' according to an interface analysis in PDBePISA (Krissinel and Henrick, 2007). Further, a single hydrogen bond is formed between interfacing IL9 or VL9 peptides in the two structures. Therefore, although crystal contacts may play a role in peptide positioning and influence the orientation of solvent exposed peptide-derived side chains when molecules pack in this manner, the impact of such crystal contacts is likely similar for IL9-bound HLA-E compared to previously published HLA-E structures in complex with VL9 variant peptides.

Hydrogen bonding network between the IL9 peptide and HLA-E heavy chain

Distinct peptide-heavy chain hydrogen bonding networks are observed for the two alternative IL9 peptide binding configurations (Figure 11, A-C). Although the same number of inter-chain hydrogen bonds formed between IL9 and the HLA-E heavy chain in both configurations, only five such bonds are conserved in molecules 1 and 2 of the asymmetric unit.

A hydrogen bond formed between the IL9 position 3 Tyr side chain and the α2 helix-situated residue, Glu-152, is not only conserved between the two IL9 peptide binding configurations, but also with the RL9HIV peptide (Figure 11, D). The presence of a shallow, peptide position 3-accommodating D pocket indicates a preference for small residues at this secondary anchor position in HLA-E-restricted peptides. However, the formation of a novel hydrogen bond between the large position 3 Tyr side chain and Glu-152 of the α2 helix, which is conserved between IL9-bound and also the RL9HIV-bound HLA-E structure, redefines an alternative mode of how certain peptides can be stably accommodated at this position of the HLA-E binding groove.


A
α1-helix
Asp
69
Ser
66
Glu-62
Asn-77
Tyr-84
Tyr-171
Ser-143
Glu
152
Tyr
159
α2-helix
B
α1-helix
Ser
66
Asn
77
Glu-63
Tyr-171
Lys
146
Glu-152
Tyr
159
α2-helix

C (i)

Molecule 1

	Chain P: Peptide	Distance (Å)	Chain A: Heavy chain
1	P:ILE 1[N]	2.77	A:TYR 171[OH]
2	P:MET 2[N]	2.81	A:GLU 63[OE2]
3	P:TYR 3[OH]	3.35	A:GLU 152[OE2]
4	P:ASN 4[ND2]	2.78	A:SER 66[OG]
5	P:TYR 5[OH]	3.45	A:ASP 69[OD2]
6	P:LEU 9[N]	3.03	A:ASN 77[OD1]
7	P:ILE 1[O]	2.71	A:TYR 159[OH]
8	P:LEU 9[O]	2.72	A:TYR 84[OH]
9	P:LEU 9[O]	2.61	A:SER 143[OG]

(ii)

Molecule 2

	Chain Q: Peptide	Distance (Å)	Chain C: Heavy chain
1	Q:ILE 1[N]	2.82	C:TYR 171[OH]
2	Q:MET 2[N]	2.86	C:GLU 63[OE2]
3	Q:TYR 3[OH]	3.20	C:GLU 152[OE2]
4	Q:LEU 9[N]	2.95	C:ASN 77[OD1]
5	Q:ILE 1[O]	2.59	C:TYR 159[OH]
6	Q:TYR 3[OH]	3.07	C:GLU 152[OE1]
7	Q:ASN 4[OD1]	2.84	C:SER 66[OG]
8	Q:ALA 7[O]	3.38	C:ASN 77[ND2]
9	Q:LEU 9[O]	2.71	C:LYS 146[NZ]

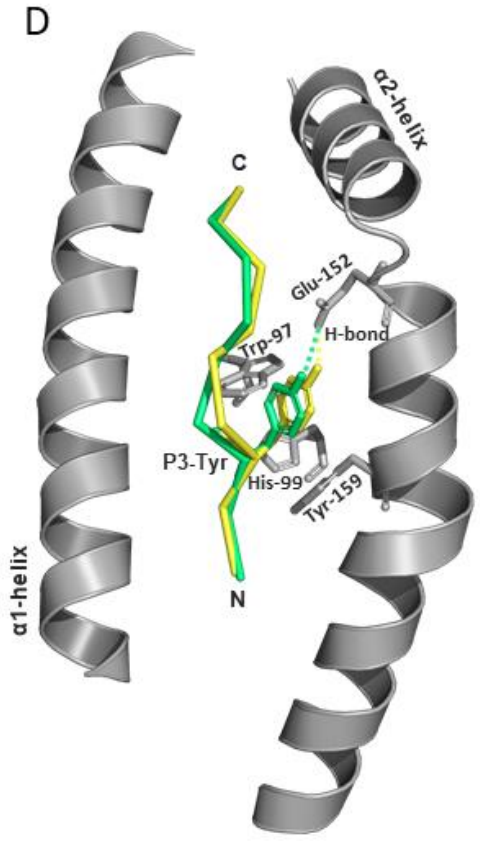


Figure 11: Peptide-heavy chain hydrogen bonding in the IL9-bound HLA-E structure

A. Visualisation of the inter-chain hydrogen bonding network between the IL9 peptide of molecule 1 and the HLA-E heavy chain. The IL9 peptide is shown as lime-green sticks with the heavy-chain-derived donor/acceptor residues as grey sticks. Hydrogen bonds are depicted as green-dashed lines. The α1 and 2 helices are visualised as grey cartoon with the remaining HLA-E heavy and light chain residues omitted for clarity. **B.** Visualisation of the inter-peptide-heavy-chain hydrogen bonding network for molecule 2.

The IL9 peptide is shown in light pink with the heavy-chain-derived donor/acceptor residues as grey sticks. Hydrogen bonds are depicted as pink-dashed lines and the α1 and 2 helices are visualised as grey cartoon. **C.** Table of hydrogen bonds formed between the IL9 peptide and HLA-E heavy chain of molecules 1 **(i)** and 2 **(ii)**. The lower and upper thresholds for the donor/acceptor distance range is set to 2.5 Å and 3.5 Å, respectively. **D.** Superposition of the IL9 (yellow) and RL9HIV peptides (green) displayed as ribbon with the position 3-Tyr in stick-form. The conserved hydrogen bonds connecting Tyr-3 to Glu-152 of the HLA-E heavy chain are displayed as dashed lines – a green dashed line connects position 3 of RL9HIV to Glu152 whereas a yellow-dashed line connects position 3 of IL9 to heavy-chain-derived Glu-152. The α1 and α2 helices are depicted in grey cartoon with Glu-152 and D pocket forming residues, His-99, Trp-97 and Tyr-159 in grey stick-form.

Pocket binding in the HLA-E*0103-IL9 structure

Both primary anchor residues are conserved between IL9 and VL9 peptides; the canonical position 2 Met and position 9 Leu are accommodated by the B and F pockets, respectively, in the IL9-bound HLA-E structure, with minimal rearrangement of pocket-forming residues relative to VL9-bound HLA-E*0101 (1MHE) (Figures 12 and 13) (O'Callaghan et al., 1998). Further, both IL9 peptide binding configurations exhibit strong convergence at such positions.

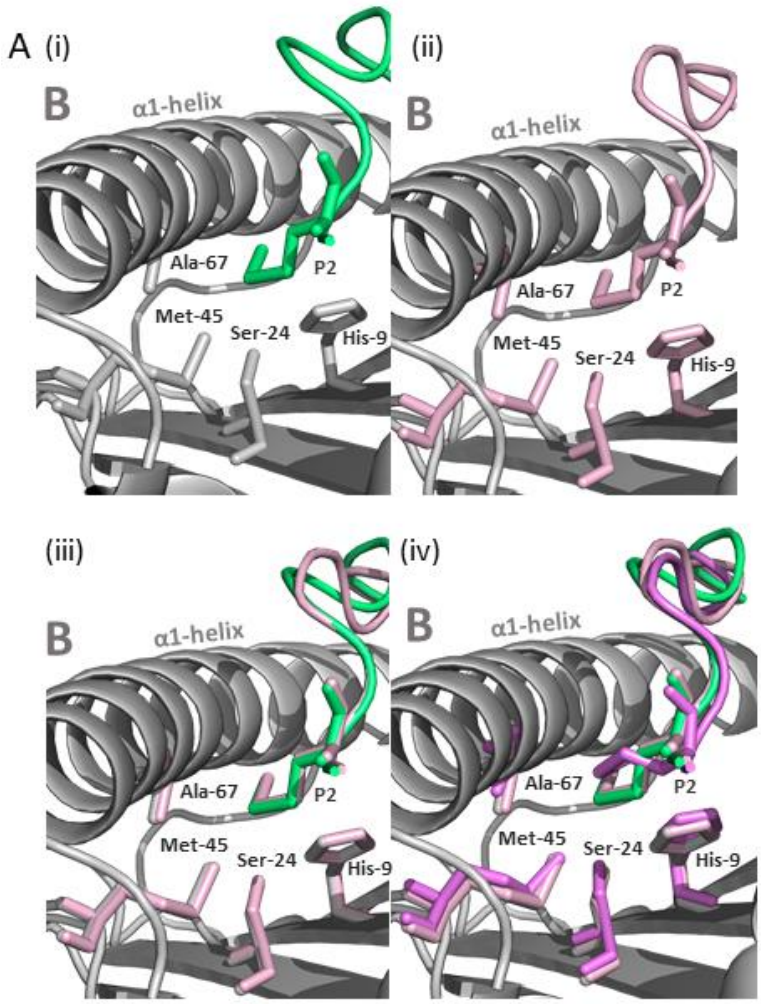


Figure 12: Primary B pocket occupancy by the IL9 peptide

A. (i) B pocket visualisation for IL9-bound HLA-E*0103 molecule 1. Heavy chain-derived B pocket-forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as grey cartoon. The IL9 peptide is shown in green cartoon with position 2 Met in green stick-form. **(ii)** B pocket visualisation for IL9-bound HLA-E*0103 molecule 2. Heavy chain-derived B pocket-forming residues are depicted as pink sticks with the neighbouring area of the peptide binding groove as grey cartoon. The IL9 peptide is shown in pink cartoon with position 2 Met in pink stick-form. **(iii)** Superimposed IL9 peptides and B pocket-forming residues of molecules 1 and 2, colour-coded according to A (i) and (ii). **(iv)** B pocket visualisation of the superimposed IL9-HLA-E molecules 1 and 2 with VL9-HLA-E from 1MHE. Colour-coding is according to A (i) and (ii) with superimposed B pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 2 Met in stick-form.

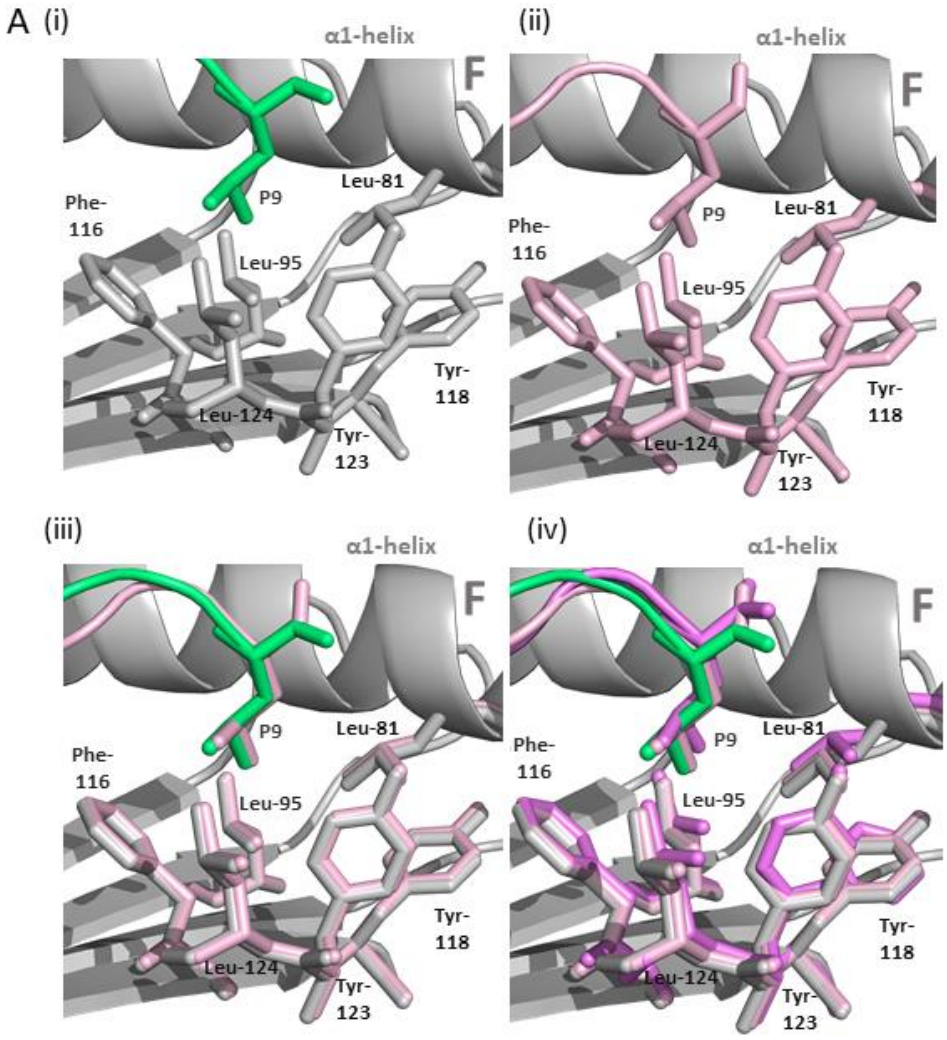


Figure 13: Primary F pocket occupancy by the IL9 peptide

A. (i) F pocket visualisation for IL9-bound HLA-E*0103 molecule 1. Heavy chain-derived F pocket-forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as grey cartoon. The IL9 peptide is shown in green cartoon with position 9 Leu in green stick-form. **(ii)** F pocket visualisation for IL9-bound HLA-E*0103 molecule 2. Heavy chain-derived F pocket-forming residues are depicted as pink sticks with the neighbouring area of the peptide binding groove as grey cartoon. The IL9 peptide is shown in pink cartoon with position 9 Leu in pink stick-form. **(iii)** Superimposed IL9 peptides and F pocket-forming residues of molecules 1 and 2, colour-coded according to A (i) and (ii). **(iv)** F pocket visualisation of the superimposed IL9-HLA-E molecules 1 and 2 with VL9-HLA-E from 1MHE. Colour-coding is according to A (i) and (ii) with superimposed F pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 9 Leu in stick-form.

Regarding secondary pocket binding, the side chain of position 3 Tyr appears prohibitively large to project downward into the shallow D pocket and instead projects towards the α2 helix forming a novel hydrogen bond with the HLA-E heavy chain (Figure 14). Minimal repositioning of pocket-forming residues is observed for both the D and C pockets of molecules 1 and 2 of IL9-bound HLA-E and molecule 1 of VL9-bound HLA-E*0101 (1MHE) (Figure 14 and 15) (O'Callaghan et al., 1998). Further, Cα backbone convergence of the two distinct IL9 peptide binding configurations is observed at positions 3 and 6. Substitution of canonical Thr to Pro in the IL9 peptide is well tolerated by the C pocket, presumably due to the small size of the substituted residue and resultant lack of steric incompatibility with this shallow pocket.

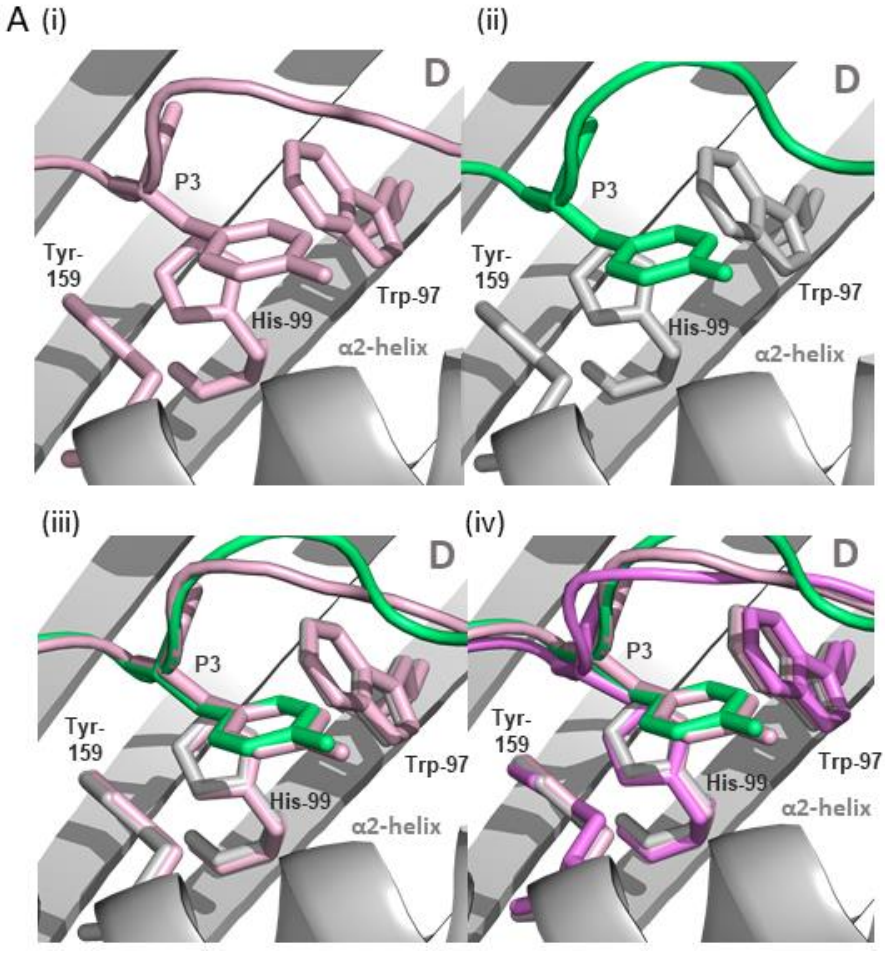


Figure 14: Secondary D pocket occupancy by the IL9 peptide

A. (i) D pocket visualisation for IL9-bound HLA-E*0103 molecule 1. Heavy chain-derived D pocket-forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as

grey cartoon. The IL9 peptide is shown in green cartoon with position 3 Tyr in green stick-form. **(ii)** D pocket visualisation for IL9-bound HLA-E*0103 molecule 2. Heavy chain-derived D pocket-forming residues are depicted as pink sticks with the neighbouring area of the peptide binding groove as grey cartoon. The IL9 peptide is shown in pink cartoon with position 3 Tyr in pink stick-form. **(iii)** Superimposed IL9 peptides and D pocket-forming residues of molecules 1 and 2, colour-coded according to A (i) and (ii). **(iv)** D pocket visualisation of the superimposed IL9-HLA-E molecules 1 and 2 with VL9-HLA-E from 1MHE. Colour-coding is according to A (i) and (ii) with superimposed D pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 3 Ala in stick-form.

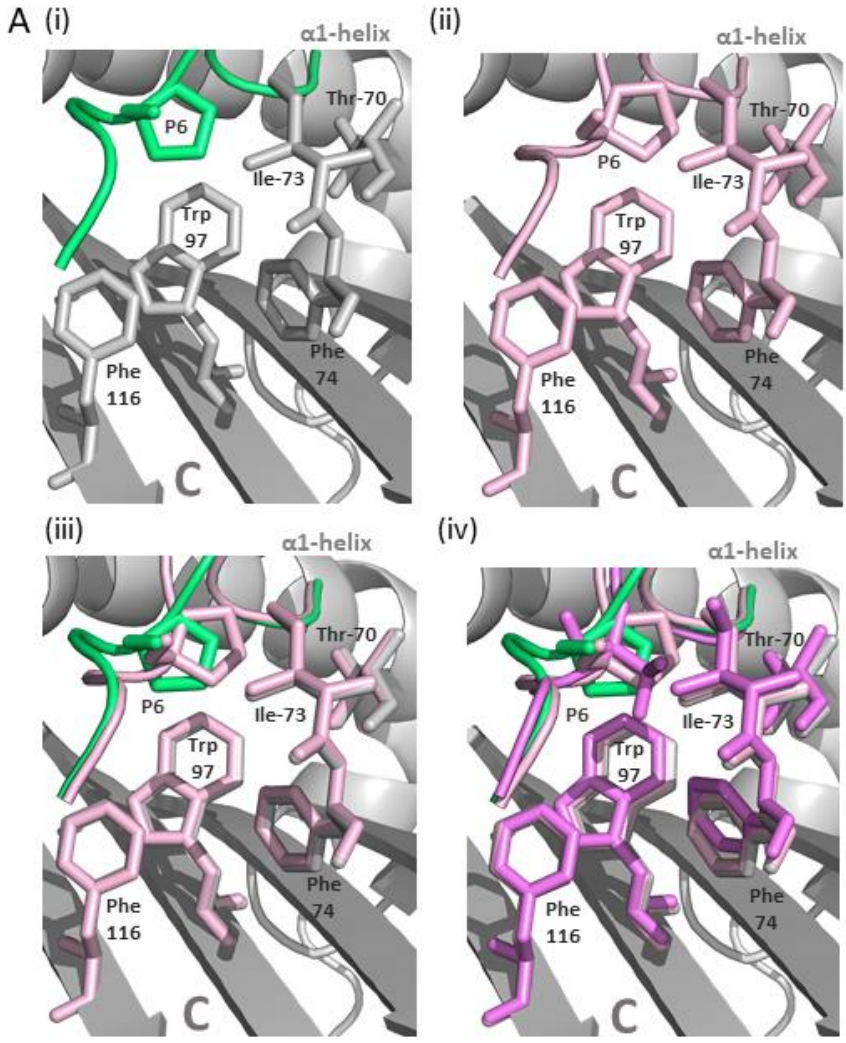


Figure 15: Secondary C pocket occupancy by the IL9 peptide

A. (i) C pocket visualisation for IL9-bound HLA-E*0103 molecule 1. Heavy chain-derived C pocket-

forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as grey cartoon. The IL9 peptide is shown in green cartoon with position 6 Pro in green stick-form. **(ii)** C pocket visualisation for IL9-bound HLA-E*0103 molecule 2. Heavy chain-derived C pocket-forming residues are depicted as pink sticks with the neighbouring area of the peptide binding groove as grey cartoon. The IL9 peptide is shown in pink cartoon with position 6 Pro in pink stick-form. **(iii)** Superimposed IL9 peptides and C pocket-forming residues of molecules 1 and 2, colour-coded according to A (i) and (ii). **(iv)** C pocket visualisation of the superimposed IL9-HLA-E molecules 1 and 2 with VL9-HLA-E from 1MHE. Colour-coding is according to A (i) and (ii) with superimposed C pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 6 Thr in stick-form.

For both configurations, the greatest structural readjustment of pocket-forming residues and peptide backbones is observed in the position 7-accommodating E pocket (Figure 16). Although the position 7 α-carbon of IL9 configuration 2 aligns to that of VL9 (1MHE) (O'Callaghan et al., 1998), the peptide is elevated 2.0 Å in the groove at position 7 of IL9 configuration 1. Further, the small position 7 Ala side chain in configuration 1 projects upward toward the solvent, disrupting E pocket-anchor side chain interactions. Despite these differences in E pocket binding, minimal pocket-forming residue repositioning is observed between the IL9 configurations. Conversely, the E pocket-forming Glu-152 is reoriented in the HLA-E*0103-IL9 structure relative to HLA-E*0101-VL9 (1MHE) (O'Callaghan et al., 1998). This is likely due to its alternative involvement in the peptide-heavy chain hydrogen bonding networks of these two structures. In HLA-E*0101-VL9 (1MHE) Glu-152 forms a hydrogen bond with position 5 Arg of the VL9 peptide. In contrast, a novel hydrogen bond between Glu-152 and the IL9 position 3 Tyr side chain – similar to that observed in the RL9HIV-HLA-E structure – re-orientates this heavy chain residue further toward the peptide N-terminus.

Side chain repositioning is also observed for heavy chain-derived Gln-156 in the HLA-E*0103-IL9 structure relative to HLA-E*0101-VL9 (O'Callaghan et al., 1998). Gln-156 reorients deeper

in the E pocket in IL9-bound HLA-E relative to canonically bound VL9. In IL9-bound HLA-E, the side chain of Gln-156 forms intra-chain hydrogen bonds with the side chains of Glu-152 of the α2 helix and Glu-114, positioned in the β-sheet groove floor. Neither intra-heavy chain hydrogen bonds are present in the VL9-bound HLA-E structure 1MHE (O'Callaghan et al., 1998). Although both hydrogen bonds are present in RL9HIV-bound HLA-E, Gln-156 also forms a hydrogen bond with the side chain of the heavy chain residue, His-99, which might account for its positional adjustment relative to IL9-bound HLA-E.

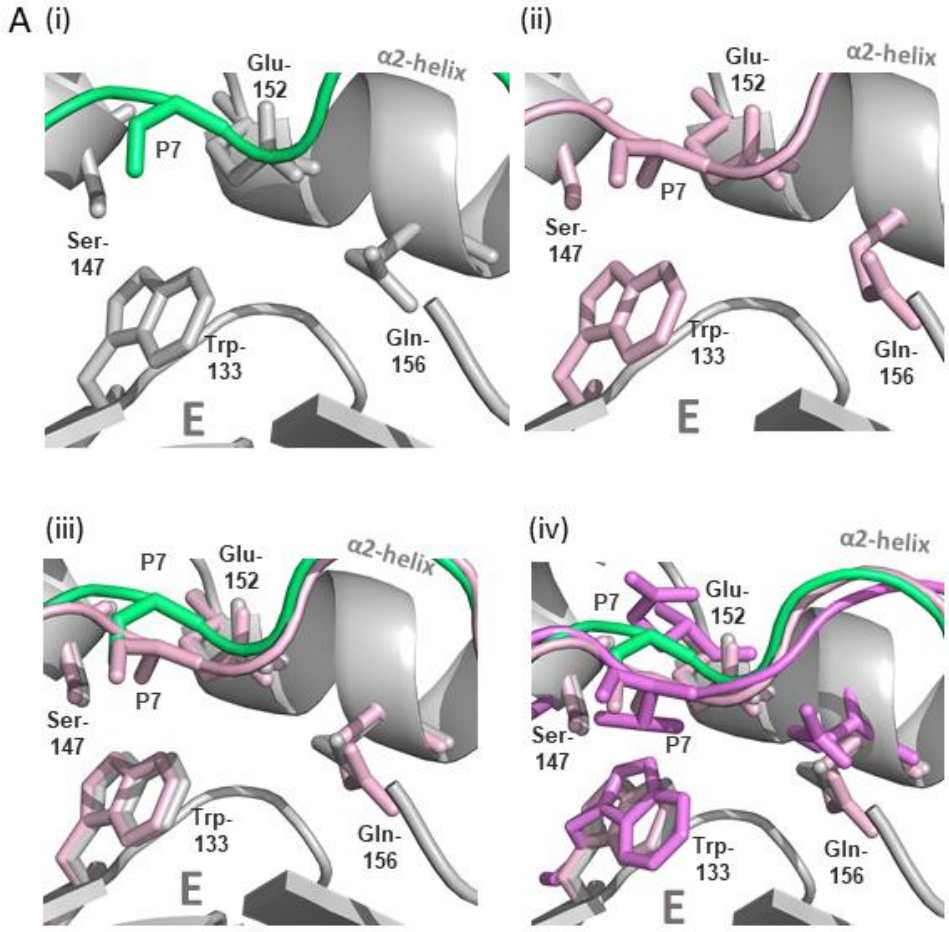


Figure 16: Primary E pocket occupancy by the IL9 peptide

A. (i) E pocket visualisation for IL9-bound HLA-E*0103 molecule 1. Heavy chain-derived E pocket-forming residues are depicted as grey sticks with the neighbouring area of the peptide binding groove as grey cartoon. The IL9 peptide is shown in green cartoon with position 7 Ala in green stick-form. **(ii)** E pocket visualisation for IL9-bound HLA-E*0103 molecule 2. Heavy chain-derived E pocket-forming

residues are depicted as pink sticks with the neighbouring area of the peptide binding groove as grey cartoon. The IL9 peptide is shown in pink cartoon with position 7 Ala in pink stick-form. **(iii)** Superimposed IL9 peptides and E pocket-forming residues of molecules 1 and 2, colour-coded according to A (i) and (ii). **(iv)** E pocket visualisation of the superimposed IL9-HLA-E molecules 1 and 2 with VL9-HLA-E from 1MHE. Colour-coding is according to A (i) and (ii) with superimposed E pocket-forming residues derived from the VL9-bound HLA-E*0101 structure depicted as violet sticks. The VL9 peptide is displayed as violet cartoon with position 7 Val in stick-form.

Size exclusion chromatography-coupled small angle X-ray scattering analysis of HLA-E in complex with peptides of differential binding affinity

HLA-E*0103 refolds were assembled in an identical fashion for a range of HLA-E-restricted peptides with a spectrum of relative binding strengths established in the peptide exchange-based binding assay (Results Chapter 1, Figures 16-22). The positive control VL9 peptide (VMAPRTVLL) was selected in addition to Mtb44 (RLPAKAPLL) to represent the highest HLA-E-binding bracket (Results Chapter 1, Figure 21). Six peptides that yielded intermediate to low relative HLA-E binding affinity were also tested including RL9HIV (RMYSPTSIL), IL9 (IMYNYPAML), Mtb14 (RMAATAQVL), RL9SIV (RMYNPTNIL), BZLF (SQAPLPCVL) and EK9 (EKQRESREK) (Results Chapter 1, Figures 17a, 17b, 21 and 22). A size exclusion chromatography-coupled small angle X-ray scattering (SEC-SAXS) experiment for HLA-E refolded in the absence of peptide was also conducted.

SEC-SAXS is a biophysical technique that provides low resolution insight regarding the overall conformational ensemble of the protein sample in solution alleviating the rate-limiting requirement of x-ray crystallography for protein crystal formation (Skou et al., 2014). Scattering patterns from x-ray-exposed material are collected during high-performance liquid chromatography (HPLC) elution, from which it is possible to calculate the mass distribution (radius of gyration R_g), maximal average dimensions (d_{max}) and degree of folding in the eluted protein sample following subtraction of buffer-induced x-ray scattering (Kikhney and Svergun,

2015). Eluted material from the HPLC column was exposed to X-rays every three seconds in SEC-SAXS experiments, generating over 600 frames over the entire HPLC elution trace. Following buffer subtraction and manual Guinier fitting in the ATSAS-supported SAXS-dedicated software, ScÅtter (Franke et al., 2017), the HLA-E refold containing VL9 peptide yielded a radius of gyration (R_g) of 25.55 Å and a maximal surface dimension (d_{max}) of 87 Å (Figure 17). Notably, the SEC-SAXS calculated d_{max} is higher than the maximum measured dimension across the crystallographic coordinates of VL9-bound HLA-E in 1MHE (O'Callaghan et al., 1998) - 70.8 Å separates the peak of the α2 helix kink with the most membrane-proximal loop of the α3 domain. However, this 16.2 Å discrepancy is likely due to lattice-imposed crystallographic constraints prohibiting conformational sampling that would otherwise occur within the protein ensemble in solution. Thus, it is not surprising that the maximum dimension of a peptide-bound MHC class I complex is, on average, slightly greater in solution. Additionally, such dimensional discrepancies between crystallographic and SEC-SAXS data might in part be due to the low resolution limits of SEC-SAXS (~10 Å) (Skou et al., 2014).

A

Peptide:	R_g (Å)	MW (Da)	d_{max} (Å)
VL9	25.55	5.20E+04	87
Mtb44	25.91	5.40E+04	97
IL9	27.91	5.60E+04	93
EBV BZLF1	34.51	7.10E+04	108
Mtb14	36.9	8.00E+04	120
RL9HIV	39.15	9.10E+04	127
RL9SIV	46.05	1.00E+05	155
EK9	50.07	3.50E+05	162
Peptide-free	64.22	1.60E+05	229

Figure 17: Size exclusion chromatography-coupled small angle X-ray scattering analysis of HLA-E

A. Table detailing the radius of gyration (R_g) in Å, molecular weight (MW) in Daltons and the maximum

dimension (d_{max}) in Å calculated by the SEC-SAXS-dedicated analysis software, ScÅtter, in association with ATSAS for HLA-E refolded with a range of specified peptides. Statistics were generated following buffer subtraction across the entire HPLC-eluted protein peak and subsequent manual Guinier fitting.

The molecular weight prediction based on the SEC-SAXS analysis of VL9-refolded HLA-E is within 15% of the actual molecular weight of the peptide-β2M-HLA-E complex (46kDa). Further, the DAMMINF *ab initio* bead model, which represents the average conformation in solution, superposes to the PDB coordinates of a single VL9-bound HLA-E molecule in 1MHE (Figure 18, A) (Kozin and Svergun, 2001) (O'Callaghan et al., 1998). Together, this indicates that a similar overall conformation is adopted in the crystal lattice and in solution for VL9-refolded HLA-E.

The crystallographic coordinates of VL9-bound and RL9HIV-bound HLA-E superpose and exhibit the classical overall architecture of an MHC class I refolded complex (O'Callaghan et al., 1998) (Walters et al., 2018). Although marked differences in peptide positioning and α3 domains shifts are detected in lattice-constrained RL9HIV-bound HLA-E relative to VL9-bound HLA-E, maximal dimensions across the protein complexes are conserved and unaffected by such localised structural repositioning. However unexpectedly, the SEC-SAXS-generated d_{max} for RL9HIV-bound HLA-E in solution (127 Å) was 40 Å greater than the SEC-SAXS-measured d_{max} of VL9-bound HLA-E (87 Å) (Figure 17). Further, a greater R_g was calculated for HLA-E refolded with RL9HIV (39.15 Å) relative to VL9 (25.55 Å), indicative of distinct mass distributions within the protein complexes reflected by differential average tumbling trajectories in solution. The SEC-SAXS molecular weight prediction for HLA-E refolded with RL9HIV (9.10×10^4 Da) was far greater than the known mass of the protein complex (4.6×10^4 Da). This overestimation in molecular weight likely reflects a degree of protein complex unfolding since the increased volume occupied by unfolded domains confounds molecular weight calculations. Alternatively, a greater molecular weight prediction could indicate dimerization, although this seems unlikely considering the HPLC elution peak for RL9HIV-refolded HLA-E protein complexes is

situated in an identical position on the HPLC elution trace relative to VL9-refolded HLA-E, indicative of a similar molecular weight.

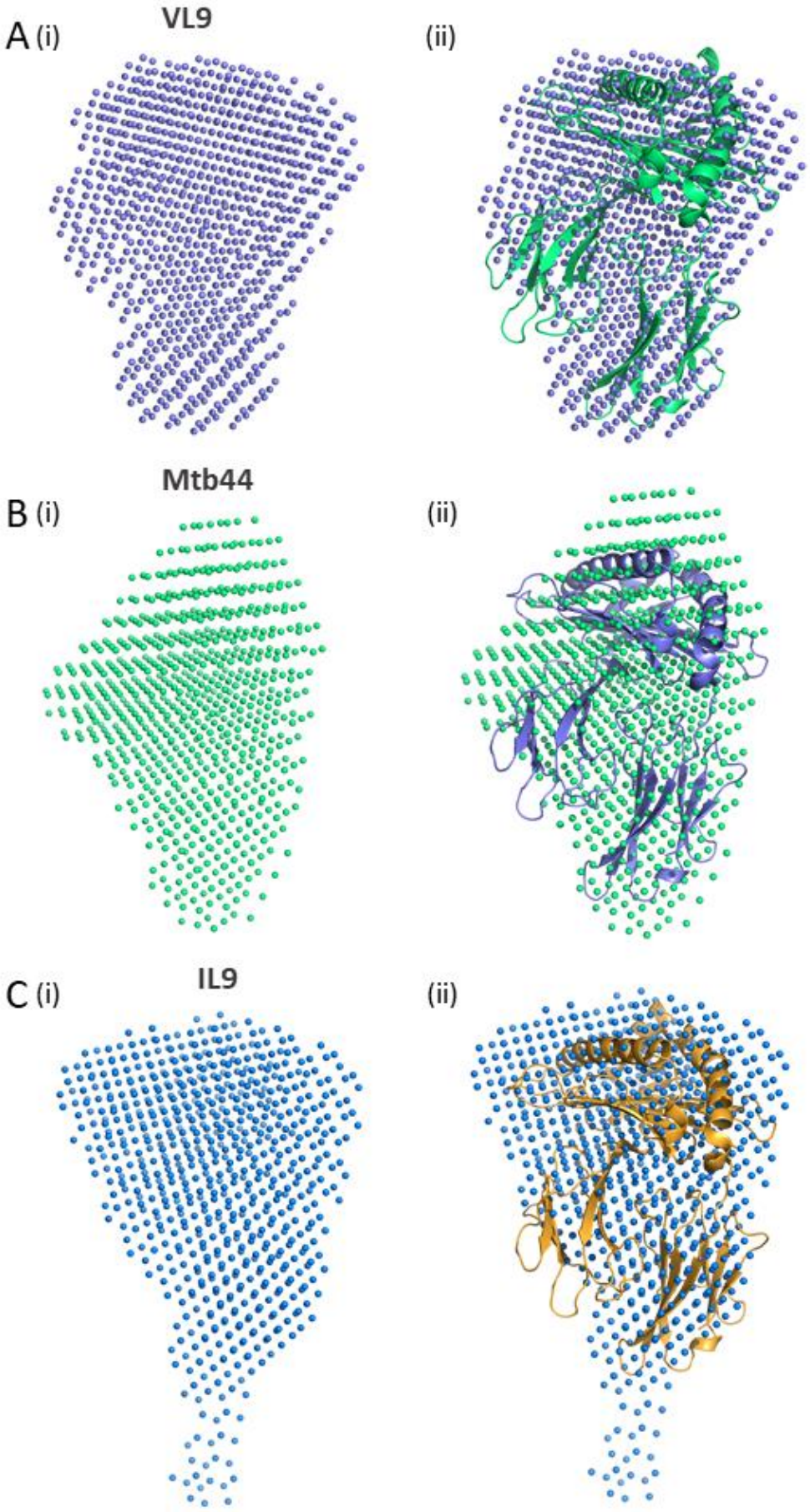
The DAMMINF bead model for RL9HIV-refolded HLA-E appeared extended relative to that of VL9-bound HLA-E (Figure 18, F). Similarly, superposition using the Supcomb package of ATSAS revealed an elongated DAMMINF bead model relative to the structural coordinates of RL9HIV-bound HLA-E (Kozin and Svergun, 2001). It is possible the larger average volume occupied by RL9HIV-refolded HLA-E in solution is generated by partially unfolded transitions states of HLA-E weakly bound to or dissociated from the RL9HIV peptide. Presumably, the more compact crystallographic coordinates of RL9HIV-bound HLA-E reflect a subpopulation in solution that was selectively incorporated into the crystal lattice due to reduced conformational disorder and flexibility.

A similar trend including a larger R_g, d_{max} and molecular weight in addition to an extended bead model was observed in SEC-SAXS runs for all remaining peptides refolded with HLA-E, except for Mtb44 (Figure 18, B) which closely resembled the SEC-SAXS analysis of VL9-bound HLA-E. HLA-E refolded with VL9 and Mtb44 produced compact bead models to which the structural coordinates of HLA-E superposed. HLA-E refolded with IL9 generated a bead model which appeared extended along the vertical axis with the point ensemble protruding below the superimposed HLA-E coordinates (Figure 18, C). The vertical elongation was more pronounced for HLA-E refolded with the EBV-derived BZLF1 peptide in addition to Mtb14 and RL9HIV which were previously crystallised in complex with HLA-E (Figure 18, D-F). The point ensembles projected both below and above the superimposed HLA-E coordinates for BZLF1-Mtb14- and RL9HIV-refolded HLA-E. Such results are consistent with higher d_{max} values of 108 Å, 120 Å and 127 Å, respectively, for HLA-E refolded with BZLF1, Mtb14 and RL9HIV, relative to the lower d_{max} values calculated for HLA-E refolded with VL9 (87 Å), Mtb44 (97 Å) and IL9 (93 Å).

Similar to RL9HIV-HLA-E, extended bead model representations of Mtb14- and IL9-refolded HLA-E do not superpose to the Mtb14- and IL9-bound HLA-E structural coordinates, respectively. This likely indicates that HLA-E refolded with Mtb14, IL9 or RL9HIV, does not give rise to a homogeneous population of compactly-refolded protein complexes, but instead the compact form constitutes a subpopulation which is incorporated into the crystal lattice.

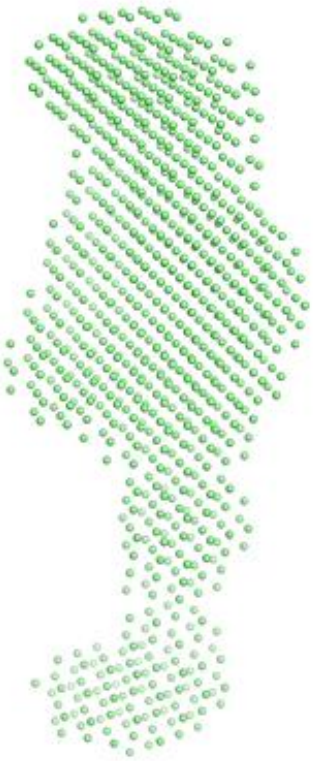
The DAMMINF bead model generated for HLA-E refolded with RL9SIV adopted a distinct conformation and appeared elongated along both the vertical and horizontal axes (Figure 18, G). A similar yet more exaggerated form was observed for both HLA-E refolded in the presence of EK9 and in the absence of peptide (Figure 18, H and I). In total, three overall bead model configurations were observed: the compact-form adopted by VL9- and Mtb44-refolded HLA-E, the vertically elongated-form adopted by IL9-, Mtb14-, BZLF- and RL9HIV-refolded HLA-E and the configuration elongated along multiple axes adopted by RL9SIV-, EK9- and peptide free-refolded HLA-E. The presence of multiple distinct bead model configurations, each containing a spectrum of elongation along various axes, for HLA-E refolded with different peptides, perhaps reflects affinity-related peptide-dependent dynamic sampling which affects the overall conformational ensemble in solution.

EK9-refolded HLA-E produced the largest d_{max} of 162 Å and greatest R_g of 50.07 Å of any peptide-HLA-E refold tested. However, the R_g and d_{max} of EK9-refolded HLA-E were slightly lower relative to that of HLA-E refolded in the absence of peptide (R_g=64.22 Å and d_{max}=229 Å). Although DAMMINF bead models of EK9 and peptide-free HLA-E appear extended along multiple axes, both such analyses should be treated with limited confidence due to small HPLC-eluted protein complex peaks resulting in weaker scattering intensities, thus compromising the quality and reliability of downstream data processing.

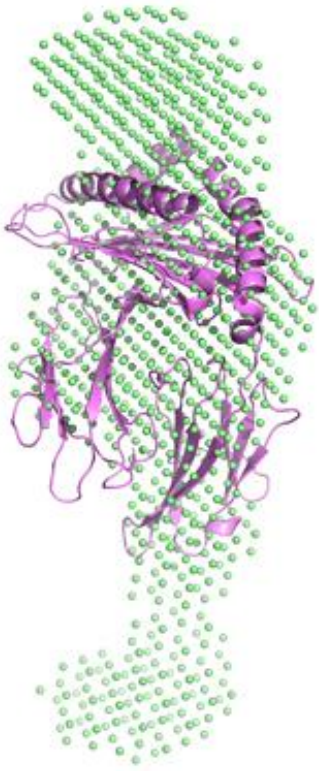

A (i)
VL9
(ii)
B (i)
Mtb44
(ii)
C (i)
IL9
(ii)

D (i)

BZLF



(ii)

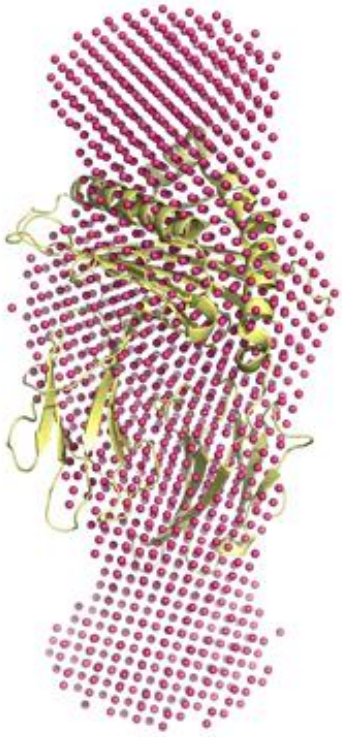


E (i)

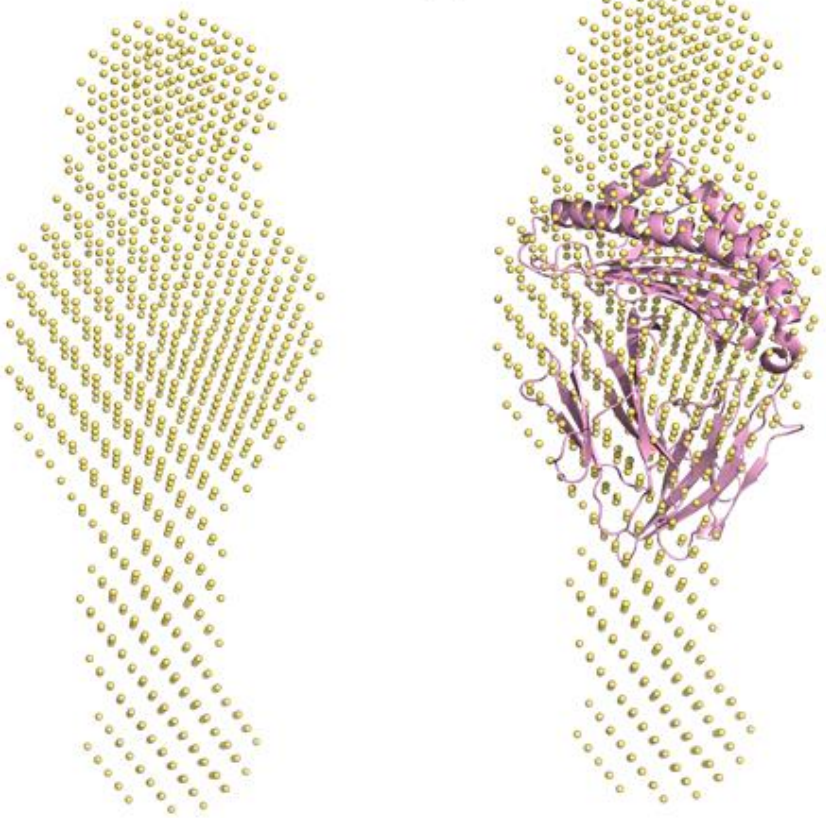
Mtb14



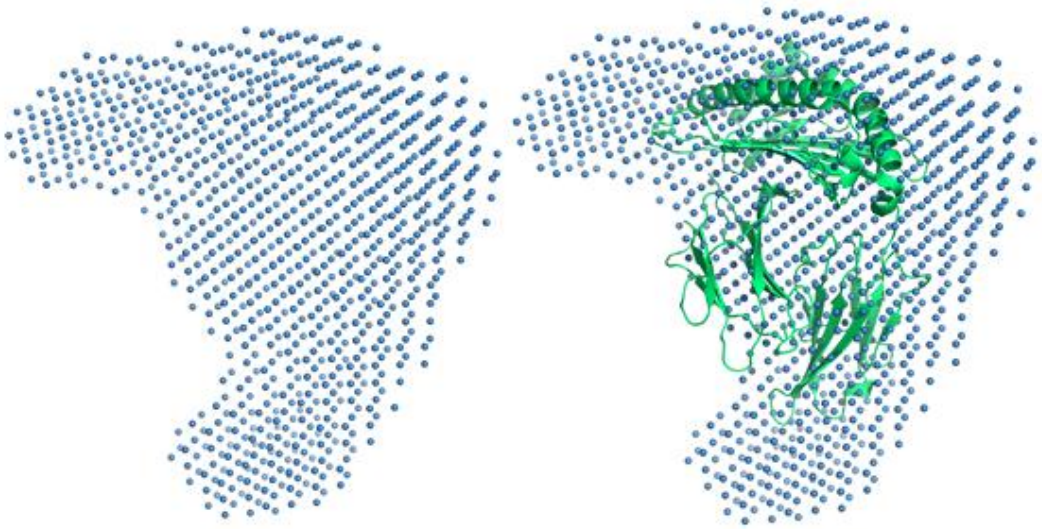
(ii)



F (i) RL9HIV (ii)



G (i) RL9SIV (ii)



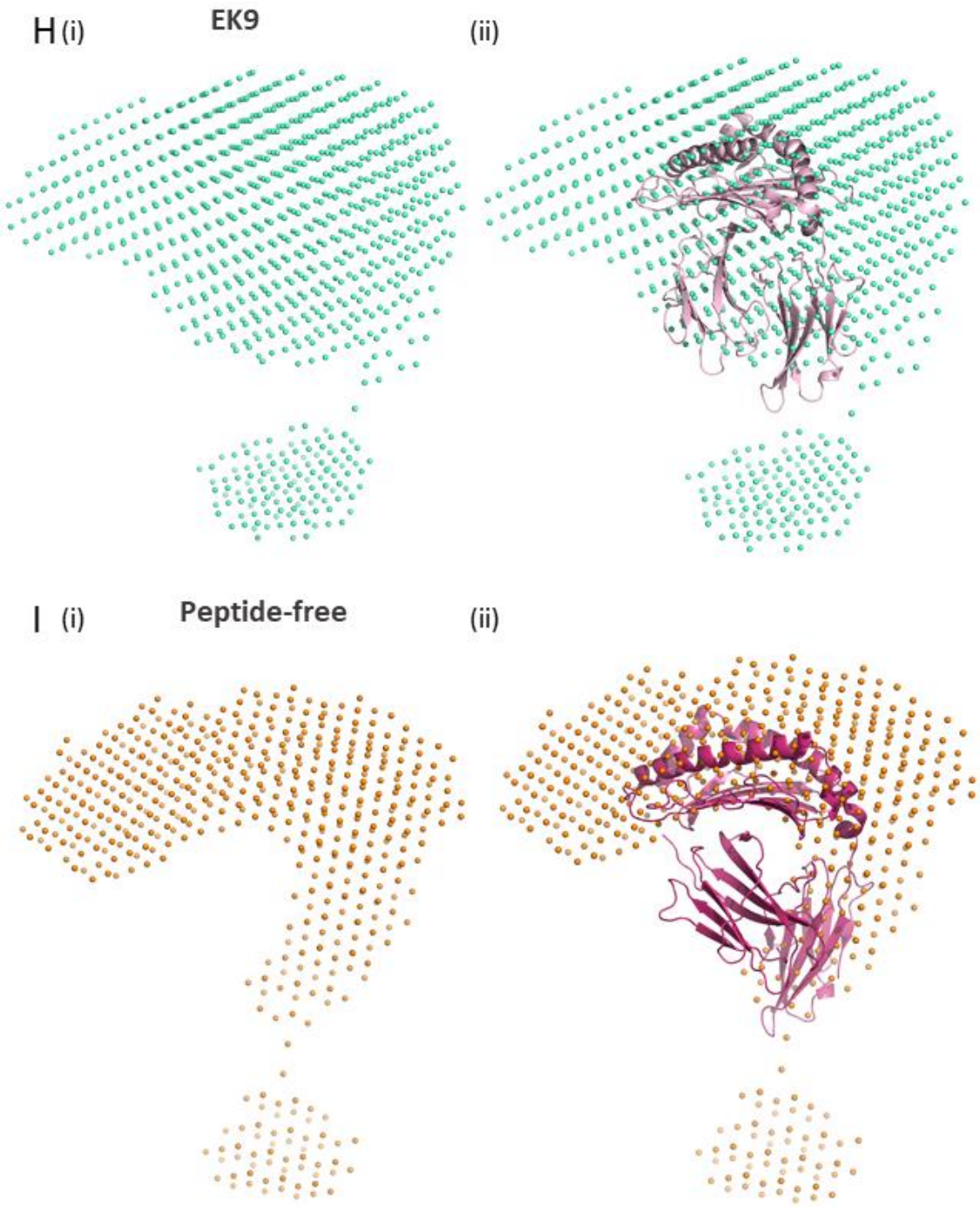


Figure 18: DAMMINF SEC-SAXS bead models of HLA-E refolded with high, medium and low affinity peptide

DAMMINF *ab initio* bead models based on real-space-refined SEC-SAXS analyses in ScÅtter in association with ATSAS. Bead model coordinates reflect the average conformation in solution for HLA-E refolded with VL9 **A. (i)**, Mtb44 **B. (i)**, IL9 **C (i)**, BZLF **D (i)**, Mtb14 **E (i)**, RL9HIV **F (i)**, RL9SIV **G (i)** and EK9 **H (i)**. **I. (i)** DAMMINF bead model of HLA-E refolded in the absence of added peptide. **A-I. (ii)** Superposition

of low resolution DAMMINF SEC-SAXS-based bead models (represented as spheres in PyMol) to high resolution crystallographic coordinates of HLA-E (depicted as cartoon) using the Supcomb package in ATSAS for HLA-E refolded in the presence of VL9, Mtb44, IL9, BZLF, Mtb14, RL9HIV, RL9SIV, EK9 and no-peptide, respectively.

Intra-sample conformational heterogeneity

A constant R_g in each X-ray-exposed frame across the HPLC-eluted HLA-E protein complex peak is indicative of a homogenous sample with similar tumbling trajectories in solution. Accordingly, the R_g plotted across the HPLC-eluted protein complex peak for VL9- and Mtb44-refolded HLA-E assumed a straight line reflecting sample homogeneity. However, for IL9-, RL9HIV-, Mtb14- and BZLF1-refolded HLA-E a gradient of decreasing R_g was observed across frames 300-350 on the signal plot which corresponds to the HPLC elution trace. This is indicative of conformational differences among intra-sample protein complexes. Thus, a stratified analysis was conducted on the HPLC elution peaks of RL9HIV- and IL9-refolded HLA-E, in addition to VL9-refolded HLA-E as a positive control. Specifically, the protein elution peaks were subdivided in two and a downstream analysis was independently carried out for both the leading and tailing peak fractions.

Consistent with a gradient of differential R_g values across the RL9HIV-refolded HLA-E elution peak, d_{max} values that deviated by 43 Å were generated in the separate analyses of the leading (152 Å) and tailing (109 Å) peak fractions (Figure 19, A). Additionally, different degrees of bead model elongation were observed for the two RL9HIV-refolded HLA-E peak fractions (Figure 19, B). The bead model representing the average conformation within the leading peak fraction appeared more elongated than the bead model representing the entire protein peak. By contrast, the bead model for the tailing peak fraction was more compact than that generated for the entire protein peak – although the tailing peak point ensemble still exceeded the vertical dimensions of the superimposed crystal structure of HLA-E. In summary, the gradient

of varying R_g values over the protein peak in addition to differentially elongated bead models for RL9HIV-refolded HLA-E likely reflect a spectrum of partially unfolded protein complexes within the protein sample.

A

Peptide:	R_g (Å)	MW (Da)	d_{max} (Å)
RL9HIV	39.15	9.10E+04	127
RL9HIV leading peak fraction	44.22	1.10E+05	152
RL9HIV tailing peak fraction	34.41	7.40E+04	109

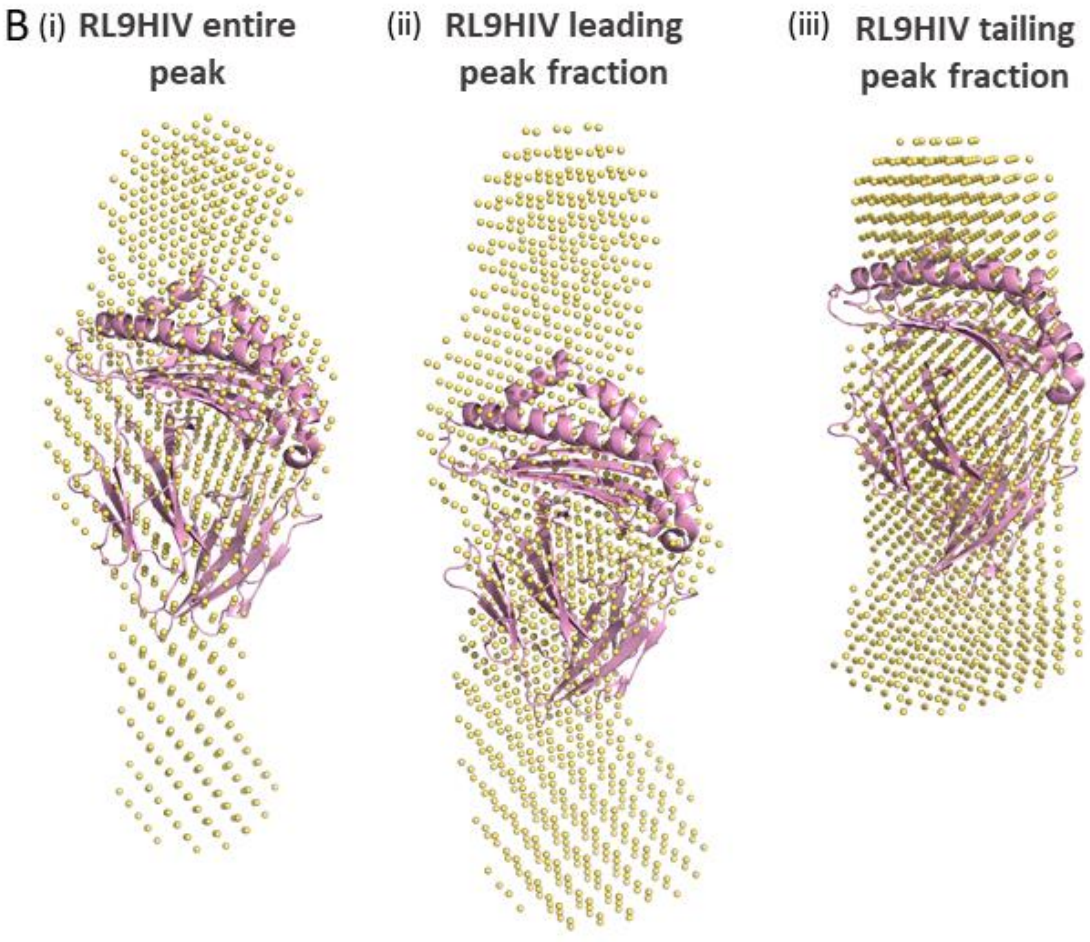


Figure 19: Stratified SEC-SAXS peak analysis for RL9HIV-refolded HLA-E*0103

A. Table detailing the radius of gyration (R_g) in Å, molecular weight (MW) in Daltons and the maximum

dimension (d_{max}) in Å calculated by the SEC-SAXS-dedicated analysis software, ScÅtter, in association with ATSAS for HLA-E refolded with RL9HIV. Such statistics were generated following buffer subtraction across the entire HPLC-eluted protein peak in addition to the leading and tailing peak fractions. **B.** Superposition of low resolution DAMMINF SEC-SAXS-based bead models (yellow spheres) to high resolution crystallographic coordinates of HLA-E (pink cartoon) using the Supcomb package in ATSAS. Bead models are based on buffer-subtracted real-space-refined SEC-SAXS data across the entire HPLC-eluted protein peak **(i)**, the leading peak fraction **(ii)**, and the tailing peak fraction **(iii)** for HLA-E refolded in the presence RL9HIV.

Distinct d_{max} and R_g values were also calculated for the leading and tailing peak fractions of IL9-refolded HLA-E (Figure 20, A). A similar trend to the peak-partitioned RL9HIV-refolded HLA-E analysis was observed for IL9 with the leading peak fraction exhibiting a larger d_{max} of 113 Å relative to the 93 Å d_{max} of the entire peak analysis. Accordingly, the bead model for the leading peak fraction of IL9-refolded HLA-E exhibited greater elongation relative to that of the entire peak, with the point ensemble protruding both above and below the superimposed structural coordinates (Figure 20, B). By contrast, the bead model for the tailing peak section was more compact and protruded beyond the superimposed structural coordinates to a lesser degree than that of the entire peak.

A

Peptide:	R_g (Å)	MW (Da)	d_{max} (Å)
IL9	27.91	5.60E+04	93
IL9 leading peak fraction	32.4	6.30E+04	113
IL9 tailing peak fraction	28.54	5.30E+04	100

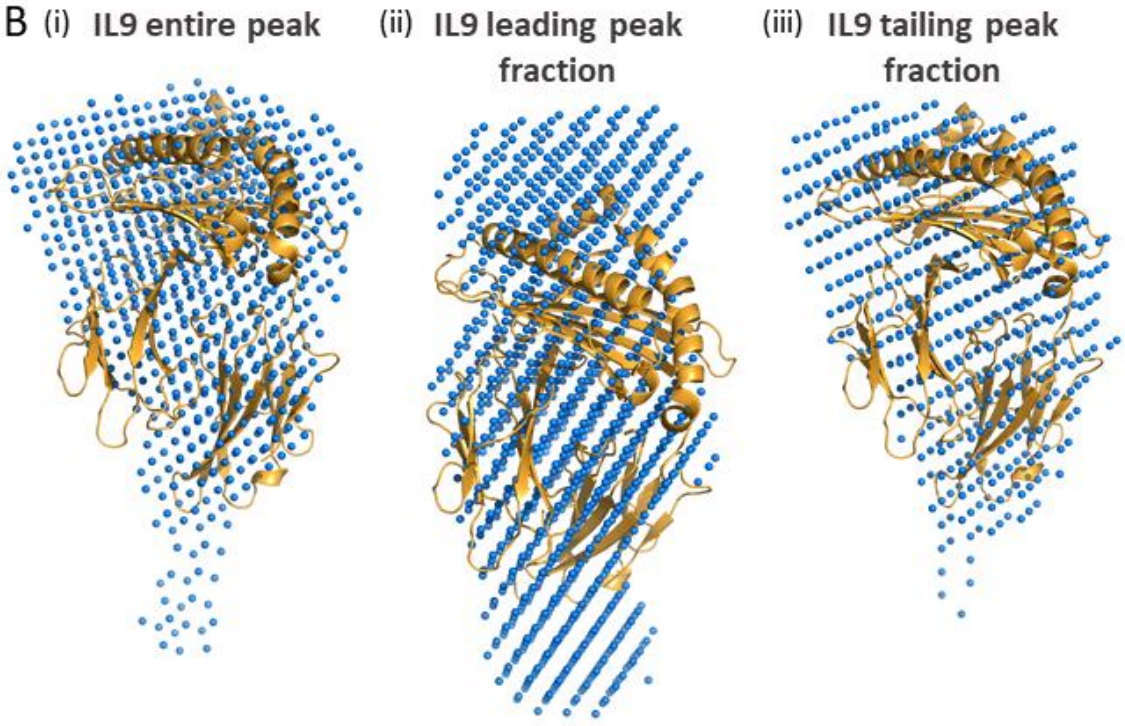


Figure 20: Stratified SEC-SAXS peak analysis for IL9-refolded HLA-E*0103

A. Table detailing the radius of gyration (R_g) in Å, molecular weight (MW) in Daltons and the maximum dimension (d_{max}) in Å calculated by the SEC-SAXS-dedicated analysis software, ScÅtter, in association with ATSAS for HLA-E refolded with IL9. Such statistics were generated following buffer subtraction across the entire HPLC-eluted protein peak in addition to the leading and tailing peak fractions. **B.** Superposition of low resolution DAMMINF SEC-SAXS-based bead models (blue spheres) to high resolution crystallographic coordinates of HLA-E (orange cartoon) using the Supcomb package in ATSAS. Bead models are based on buffer-subtracted real-space-refined SEC-SAXS data across the entire HPLC-

eluted protein peak **(i)**, the leading peak fraction **(ii)**, and the tailing peak fraction **(iii)** for HLA-E refolded in the presence IL9.

HPLC-elution peak partitioning was conducted for VL9-refolded HLA-E despite the presence of a constant R_g across the protein complex HPLC elution peak, indicative of a homogenous conformational ensemble. As predicted, bead models for the entire peak in addition to the leading and tailing peak fractions adopted highly similar, compact forms (Figure 21). Unlike RL9HIV- and IL9-refolded HLA-E the bead model for the leading peak fraction of VL9-refolded HLA-E did not substantially protrude below the superimposable structural coordinates. However, a trend of increasing R_g and d_{max} calculations from the leading peak fraction to the entire peak to the tailing peak fraction was consistent for RL9HIV- and VL9-refolded HLA-E, albeit less pronounced for VL9. Further, although the bead models for VL9-refolded HLA-E elution peak subsections adopted a similar overall configuration, slight differences were detectable. For example, consistent with a smaller d_{max}, the bead model generated for the tailing peak fraction of VL9-refolded HLA-E appeared more compact than those generated for the entire peak and leading peak fraction. Further, the tailing peak bead model was more compact in certain regions relative to the superimposed structural coordinates, with the α2 helical kink and a portion of the β2M domain protruding beyond the point ensemble.

A

Peptide:	R_g (Å)	MW (Da)	d_{max} (Å)
VL9	25.55	5.20E+04	87
VL9 leading peak fraction	26.04	5.30E+04	96
VL9 tailing peak fraction	25.12	5.50E+04	80

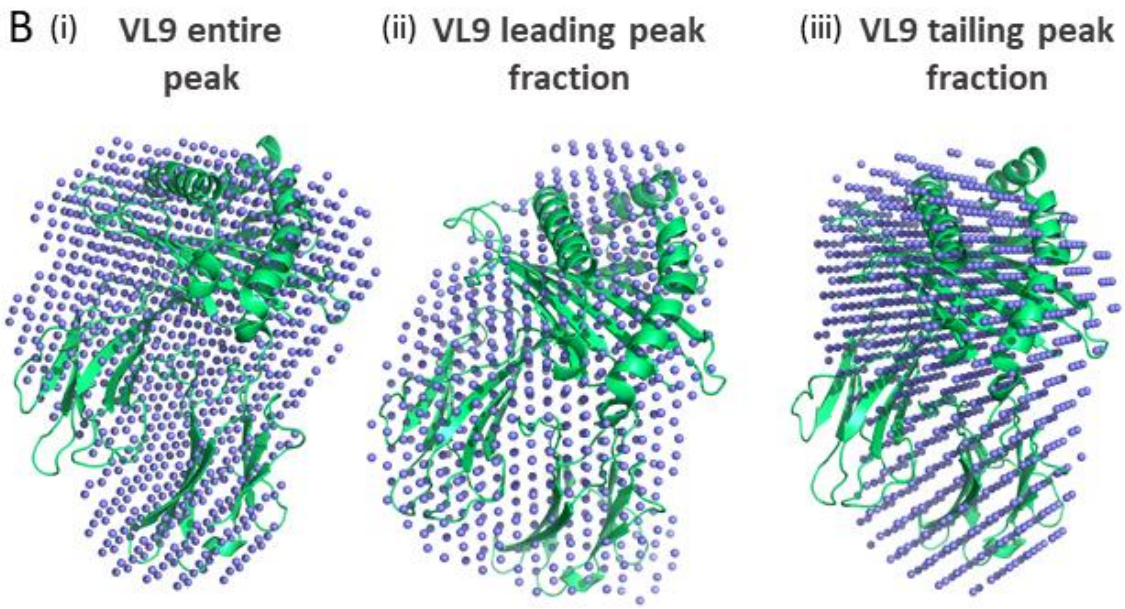


Figure 21: Stratified SEC-SAXS peak analysis for VL9-refolded HLA-E*0103

A. Table detailing the radius of gyration (R_g) in Å, molecular weight (MW) in Daltons and the maximum dimension (d_{max}) in Å calculated by the SEC-SAXS-dedicated analysis software, ScÅtter, in association with ATSAS for HLA-E refolded with VL9. Such statistics were generated following buffer subtraction across the entire HPLC-eluted protein peak in addition to the leading and tailing peak fractions. **B.** Superposition of low resolution DAMMINF SEC-SAXS-based bead models (blue spheres) to high resolution crystallographic coordinates of HLA-E (green cartoon) using the Supcomb package in ATSAS. Bead models are based on buffer-subtracted real-space-refined SEC-SAXS data across the entire HPLC-eluted protein peak **(i)**, the leading peak fraction **(ii)**, and the tailing peak fraction **(iii)** for HLA-E refolded in the presence VL9.

A different buffer system trialled in the SEC-SAXS analysis of HLA-E

An alternate buffer system was tested in the SEC-SAXS analysis of HLA-E to identify whether the results obtained in the L-Arginine-Tris pH 8.0 buffer were reproducible in different conditions. The Lutrol-Tris-maleate buffer system originally used in the micro-refolding sandwich ELISA (Sylvester-Hvid et al., 2002) was trialled in SEC-SAXS experiments for HLA-E refolded with a range of peptides including VL9, IL9 and Mtb14. However, this buffer system dramatically perturbed the classical HPLC elution trace, reflected by the signal plot (Figure 22). In the L-Arginine-Tris-based buffer HLA-E protein complexes are eluted between X-ray-exposed frames 300-350. However, no discrete peak was observed between such frames when HLA-E refolds were buffered by Lutrol-Tris-maleate.

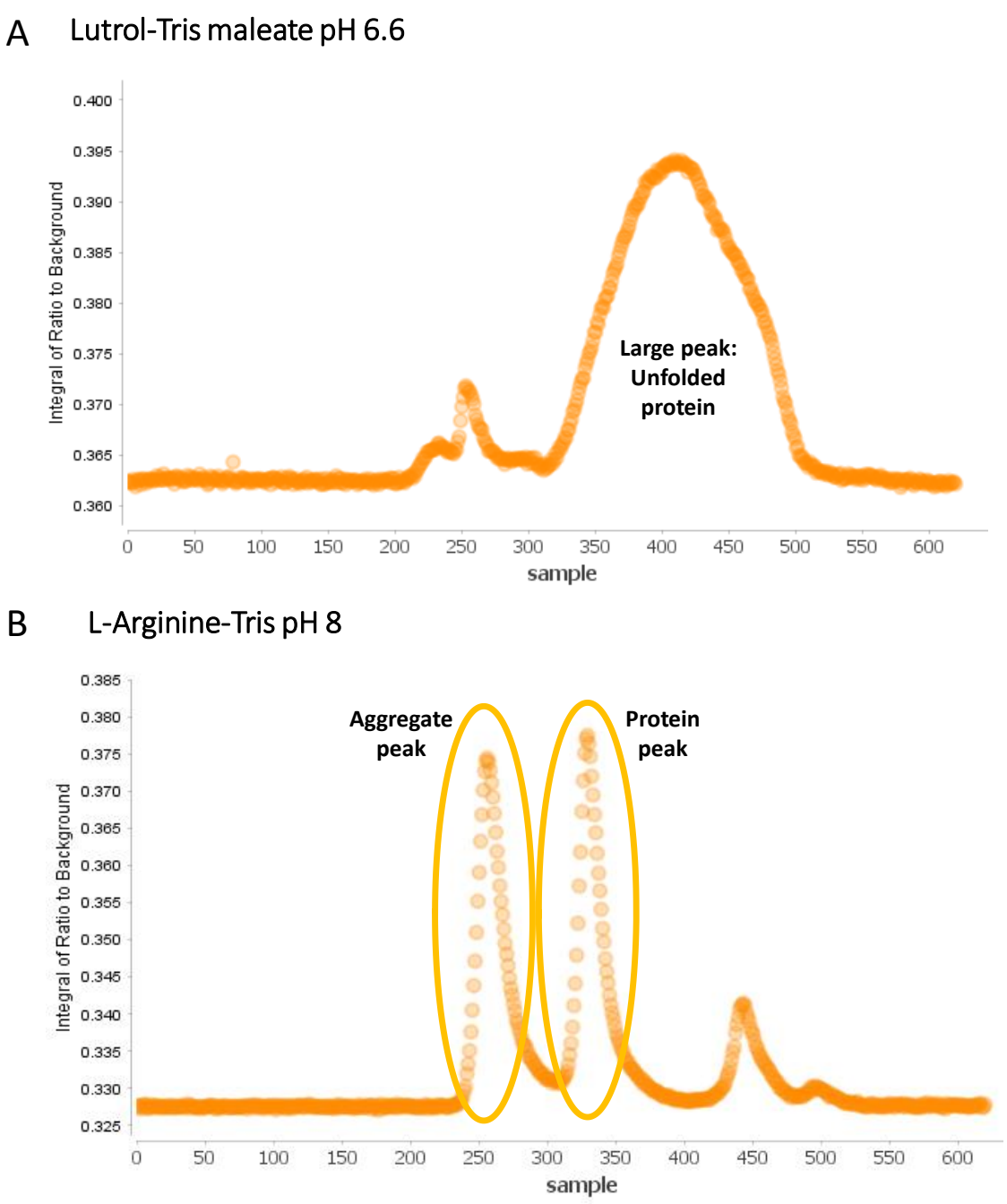


Figure 22: SEC-SAXS buffer optimisation for VL9-refolded HLA-E*0103

A. SEC-SAXS signal plot which corresponds to the HPLC elution trace for VL9-refolded HLA-E*0103 in the Lutrol-Tris-maleate pH 6.6 buffer system. The X-ray-exposed sample number which corresponds to the elution time on the HPLC run is plotted on the X-axis with the integral of ratio to background for the scattering signal on the Y-axis. **B.** SEC-SAXS signal plot which corresponds to the HPLC elution trace for VL9-refolded HLA-E*0103 in the L-Arginine-Tris pH 8.0 buffer system. The aggregate and HLA-E protein complex elution peaks are circled in orange. The X-ray-exposed sample number which corresponds to

the elution time on the HPLC run is plotted on the X-axis with the integral of ratio to background for the scattering signal on the Y-axis.

Buffer subtraction was attempted in ScÅtter for each peak present on the signal plot following injection of HLA-E refolded with VL9 peptide in the Lutrol-Tris-maleate buffer. ScÅtter was unable to subtract the background for the peaks between frames 225 and 275. Following buffer subtraction for the large peak situated between frames 325 and 500, several statistics indicated that this peak comprised fully unfolded material. First, in addition to distinct intensity plots, the curve on the Kratky Plot for Lutrol-Tris-maleate-buffered VL9-refolded HLA-E did not adopt a Gaussian profile, but instead the curve continued to increase into the larger q (q represents the scattering angle) range, indicative of elongated, unfolded protein (Figure 23 A and B) (Kikhney and Svergun, 2015). The analogous curve for HLA-E refolded with VL9 in the L-Arginine-Tris buffer yields a bell-shaped curve, consistent with globular, folded macromolecules. Second, a plateau is observed for Lutrol-Tris-maleate-buffered VL9-refolded HLA-E in the Kratky-Debye Plot, again indicative of unfolded protein (Figure 23, C). Third, L-Arginine-Tris-buffered VL9-refolded HLA-E exhibits a plateau in the Porod-Debye-Plot whereas the curve for Lutrol-Tris-maleate-buffered VL9-refolded HLA-E continues to increase. Together these curvatures for Lutrol-Tris-maleate-buffered VL9-refolded HLA-E are indicative of elongated, unfolded protein forms (Rambo and Tainer, 2011). Thus, further analysis of the large peak was not pursued since it does not likely comprise folded HLA-E complexes. The signal plots for HLA-E refolded with Mtb14 and IL9 in the Lutrol-Tris-maleate buffer exhibited similar perturbations to those observed for HLA-E refolded with VL9. Hence, further analysis of these SEC-SAXS runs was not pursued.

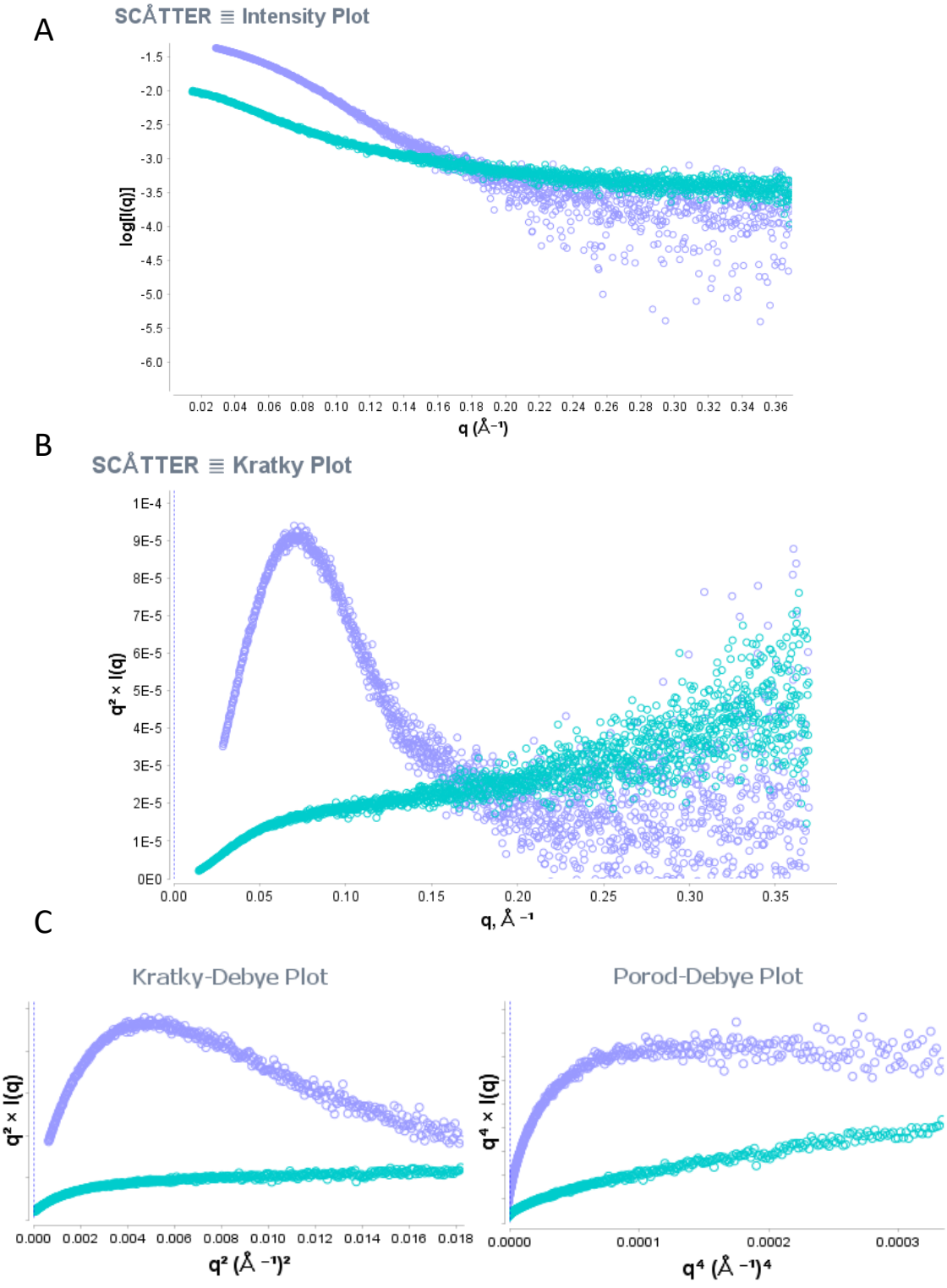


Figure 23: SEC-SAXS comparative analysis for VL9-refolded HLA-E*0103 in two different buffer systems

A. SAXS intensity plot generated in ScÅtter following buffer subtraction for HLA-E*0103 refolded with VL9 peptide in the L-Arginine-Tris pH 8 buffer (purple) and the Lutrol-Tris-maleate pH 6.6 buffer (blue).

Plotted on the X-axis is the scattering vector, q, measured in $Å^{-1}$ which for small angles is proportional to the scattering angle θ. The scattered intensity, I(q), is plotted on the Y-axis with a log scale. **B.** Kratky plot for HLA-E*0103 refolded with VL9 peptide in the L-Arginine-Tris pH 8 buffer (purple) and the Lutrol-Tris-maleate pH 6.6 buffer (blue). Plotted on the X-axis is the scattering vector, q, measured in $Å^{-1}$ which for small angles is proportional to the scattering angle θ. The Y-axis denotes the square of the scattering vector (q) multiplied by the scattered intensity (I(q)). **C.** The Kratky-Debye and Porod-Debye Plots generated by the flexibility analysis in ScÅtter for HLA-E*0103 refolded with VL9 peptide in the L-Arginine-Tris pH 8 buffer (purple) and the Lutrol-Tris-maleate pH 6.6 buffer (blue). For the Kratky-Debye Plot the square of the scattering vector (q) is plotted on the X-axis while the square of the scattering vector (q) multiplied by the scattered intensity (I(q)) is plotted on the Y-axis. For the Porod-Debye Plot the scattering vector (q) is plotted to the power of 4 on the X-axis while the scattering vector (q) to the fourth power multiplied by the scattering intensity (I(q)) is plotted on the Y-axis.

Excess peptide in the HPLC elution buffer

As previously demonstrated, HLA-E refolded with RL9HIV, RL9SIV, IL9 and Mtb14 generate elongated bead models relative to VL9. Additionally these peptides exhibit lower relative binding affinity to HLA-E relative to VL9, as determined by the peptide exchange-based binding assay developed in Results Chapter 1. Further, crystal structures of RL9HIV- and IL9-bound HLA-E revealed disrupted secondary anchor side chain interactions with corresponding secondary pockets in the peptide binding groove. Thus, it could be predicted that varying degrees of elongation observed in DAMMINF bead models arise from differential HLA-E peptide binding affinity. To investigate this, 60µM or 120µM excess peptide was added to the HPLC elution buffer, according to peptide stock and B21 beam-time availability at Diamond Light Source, in a subsequent round of SEC-SAXS experiments - an approach adopted following helpful communications with Professor Sebastian Springer, Jacobs University Bremen.

As predicted, a spectrum of elongation was observed for RL9HIV-refolded HLA-E bead models in the absence and presence of 60µM and 120µM excess peptide in the elution buffer, with

the most extended bead model representing an absence of excess peptide and the most compact bead model representing the highest peptide concentration tested (120µM) (Figure 24). The bead model for RL9HIV-refolded HLA-E in the presence of 120µM excess peptide closely resembled the compact point ensemble configuration observed when HLA-E is refolded with VL9 in the absence of excess peptide. Further, the d_{max} for conditions with both 60µM and 120µM excess peptide decreased by over 35 Å relative to that calculated for RL9HIV-refolded HLA-E in the absence of excess peptide.

A

Peptide:	R_g (Å)	MW (Da)	d_{max} (Å)
RL9HIV	39.15	9.10E+04	127
RL9HIV 60µM excess peptide	26.83	4.80E+04	90
RL9HIV 120µM excess peptide	26.3	5.40E+04	91

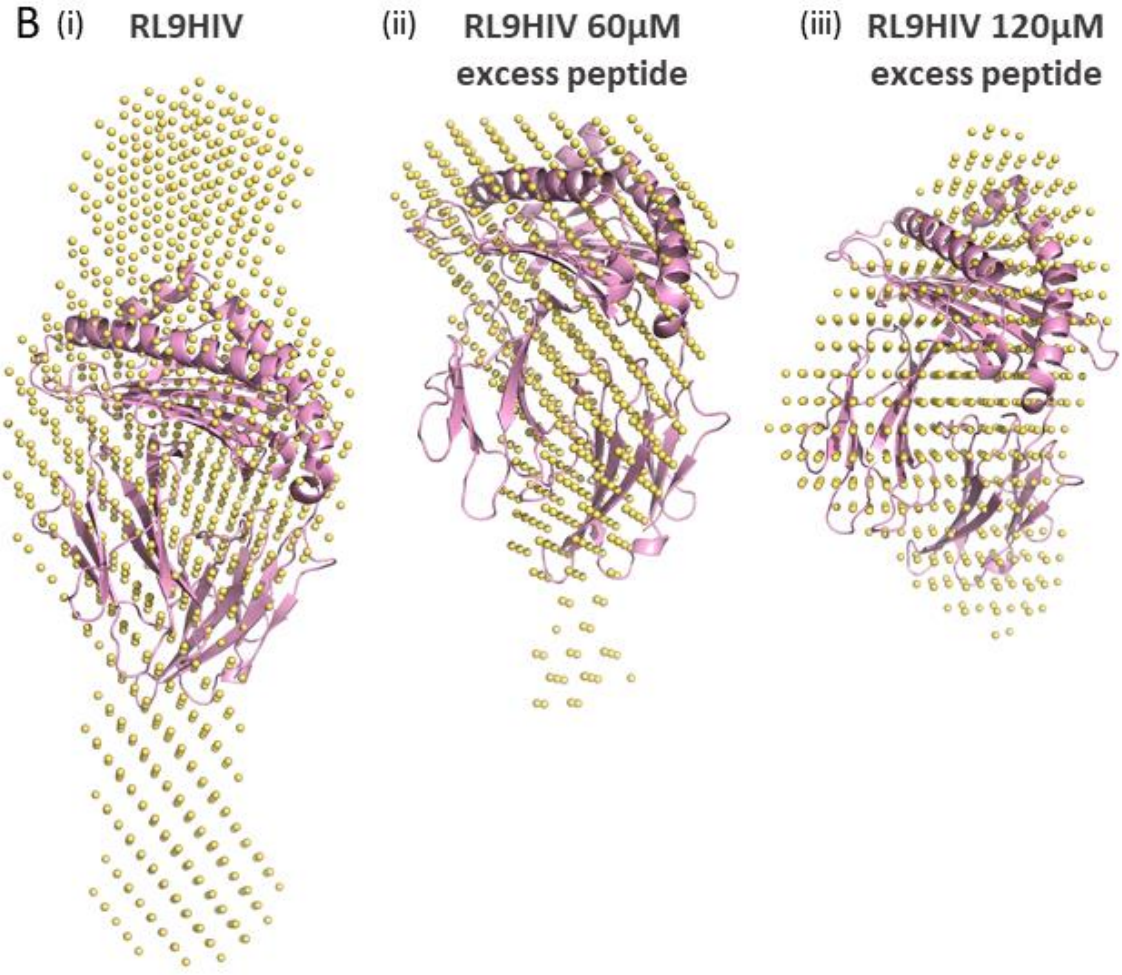


Figure 24: SEC-SAXS analysis for RL9HIV-refolded HLA-E*0103 with HPLC elution buffer excess peptide

A. Table detailing the radius of gyration (R_g) in Å, molecular weight (MW) in Daltons and the maximum dimension (d_{max}) in Å calculated by the SEC-SAXS-dedicated analysis software, ScÅtter, in association with ATSAS for HLA-E refolded with RL9HIV. Such statistics were generated following buffer subtraction across the entire HPLC-eluted protein peak for RL9HIV-refolded HLA-E in the absence and presence of 60µM and 120µM excess peptide in the HPLC elution buffer. **B.** Superposition of low resolution

DAMMINF SEC-SAXS-based bead models (yellow spheres) to high resolution crystallographic coordinates of HLA-E (pink cartoon) using the Supcomb package in ATSAS. Bead models are based on buffer-subtracted real-space-refined SEC-SAXS data across the entire HPLC-eluted protein peak for RL9HIV-refolded HLA-E in the absence **(i)** and presence of 60µM **(ii)** and 120µM **(iii)** excess peptide in the HPLC elution buffer.

A similar trend was observed for the SIV-derived equivalent peptide when excess peptide was added to the HPLC elution buffer in the SEC-SAXS run (Figure 25). Due to limited peptide stock and beamline availability, a single concentration of 60µM excess peptide was tested for RL9SIV-refolded HLA-E. The d_{max} decreased from 155 Å to 104 Å following the addition of 60µM excess peptide. Further, the distinct bead model configuration observed for RL9SIV-refolded HLA-E contracted in the presence of excess peptide.

A

Peptide:	R_g (Å)	MW (Da)	d_{max} (Å)
RL9SIV	46.05	1.00E+05	155
RL9SIV 60µM excess peptide	32.26	6.80E+04	104

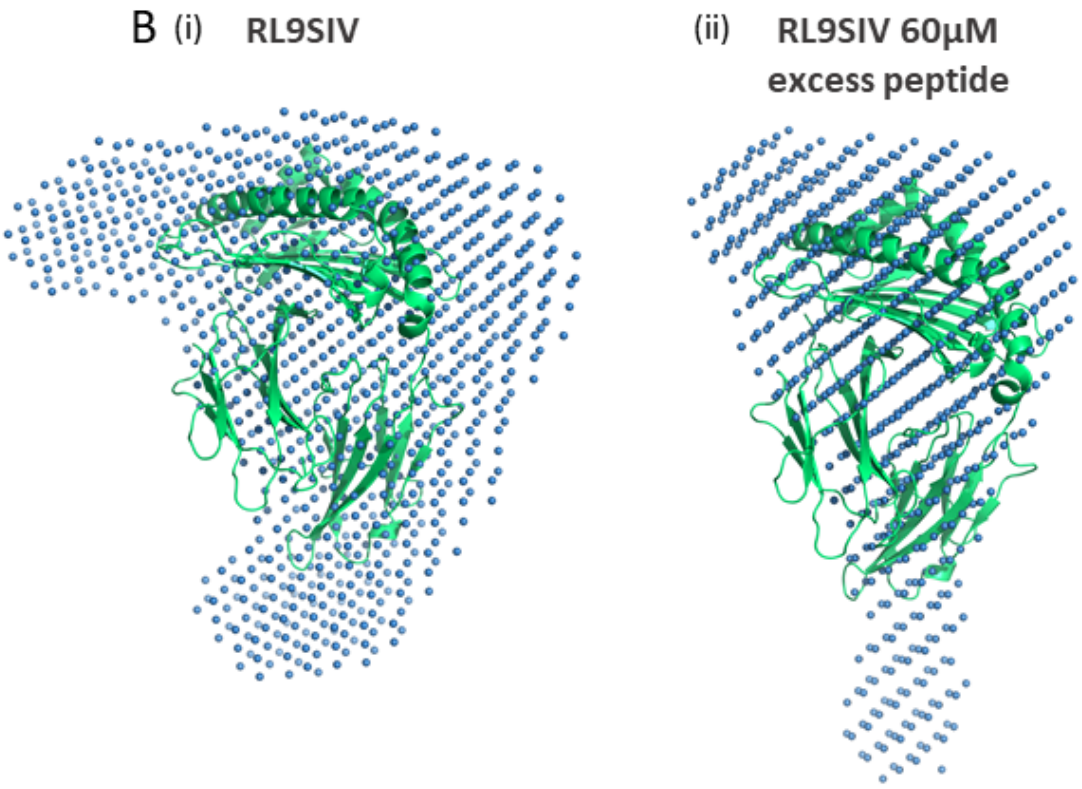


Figure 25: SEC-SAXS analysis for RL9SIV-refolded HLA-E*0103 with HPLC elution buffer excess peptide

A. Table detailing the radius of gyration (R_g) in Å, molecular weight (MW) in Daltons and the maximum dimension (d_{max}) in Å calculated by the SEC-SAXS-dedicated analysis software, ScÅtter, in association with ATSAS for HLA-E refolded with RL9SIV. Such statistics were generated following buffer subtraction across the entire HPLC-eluted protein peak for RL9SIV-refolded HLA-E in the absence and presence of 60µM excess peptide in the HPLC elution buffer. **B.** Superposition of low resolution DAMMINF SEC-SAXS-based bead models (blue spheres) to high resolution crystallographic coordinates of HLA-E (green cartoon) using the Supcomb package in ATSAS. Bead models are based on buffer-subtracted real-space-refined SEC-SAXS data across the entire HPLC-eluted protein peak for RL9SIV-refolded HLA-E in the absence **(i)** and presence of 60µM **(ii)** excess peptide in the HPLC elution buffer.

The protruding region below the superimposed HLA-E structural coordinates in the IL9-refolded HLA-E bead model was considerably reduced following the addition of 120μM excess peptide to the HPLC elution buffer (Figure 26). Further, the presence of excess peptide resulted in a bead model with a compact configuration, similar to that observed for Mtb44-refolded HLA-E in the absence of excess peptide. Bead models for both Mtb44- and IL9-refolded HLA-E in the absence and presence of 120μM excess peptide, respectively, exhibit a raised peak in the point ensemble above the kinked region of the α2 helix of superimposed HLA-E structural coordinates, possibly reflecting common conformational sampling in this region. This ensemble protrusion is absent in the bead model generated for IL9-refolded HLA-E in the absence of excess peptide.

A

Peptide:	R_g (Å)	MW (Da)	d_{max} (Å)
IL9	27.91	5.60E+04	93
IL9 120µM excess peptide	25.38	5.40E+04	91

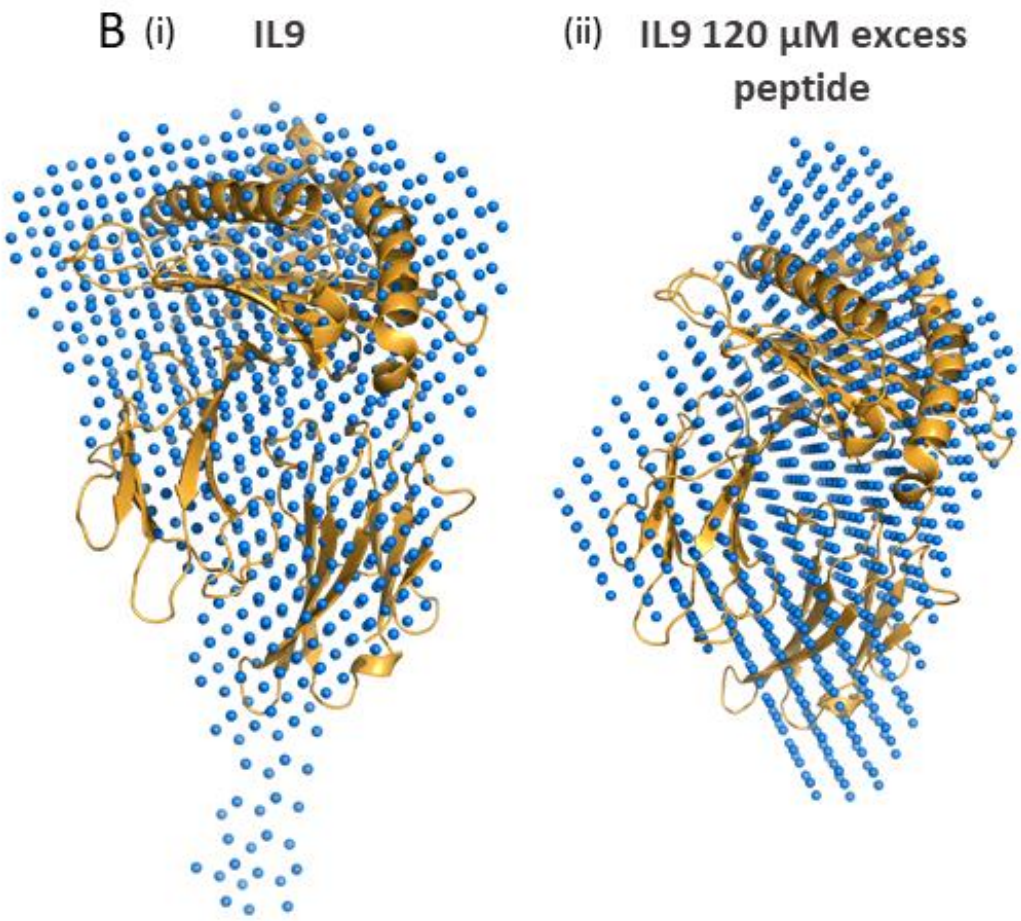


Figure 26: SEC-SAXS analysis for IL9-refolded HLA-E*0103 with HPLC elution buffer excess peptide

A. Table detailing the radius of gyration (R_g) in Å, molecular weight (MW) in Daltons and the maximum dimension (d_{max}) in Å calculated by the SEC-SAXS-dedicated analysis software, ScÅtter, in association with ATSAS for HLA-E refolded with IL9. Such statistics were generated following buffer subtraction across the entire HPLC-eluted protein peak for IL9-refolded HLA-E in the absence and presence of 120µM excess peptide in the HPLC elution buffer. **B.** Superposition of low resolution DAMMINF SEC-

SAXS-based bead models (blue spheres) to high resolution crystallographic coordinates of HLA-E (orange cartoon) using the Supcomb package in ATSAS. Bead models are based on buffer-subtracted real-space-refined SEC-SAXS data across the entire HPLC-eluted protein peak for IL9-refolded HLA-E in the absence **(i)** and presence of 120µM **(ii)** excess peptide in the HPLC elution buffer.

The elongated bead model generated for Mtb14-refolded HLA-E dramatically contracted following the addition of 120µM excess peptide to the HPLC elution buffer, in a similar manner to that observed for RL9HIV-refolded HLA-E (Figure 27). In the presence of 120µM excess peptide the bead model for Mtb14-refolded HLA-E does not protrude above the superimposed structural coordinates of HLA-E. Further, the protruding region of the bead model point ensemble below the superimposed HLA-E crystal structure is greatly reduced following the addition of 120µM excess peptide. Such qualitative observations are consistent with a reduced d_{max} by over 20 Å and a reduced R_g by over 10 Å for Mtb14-refolded HLA-E in the presence of excess peptide.

A

Peptide:	R_g (Å)	MW (Da)	d_{max} (Å)
Mtb14	36.9	8.00E+04	120
Mtb14 120µM excess peptide	25.96	5.40E+04	99

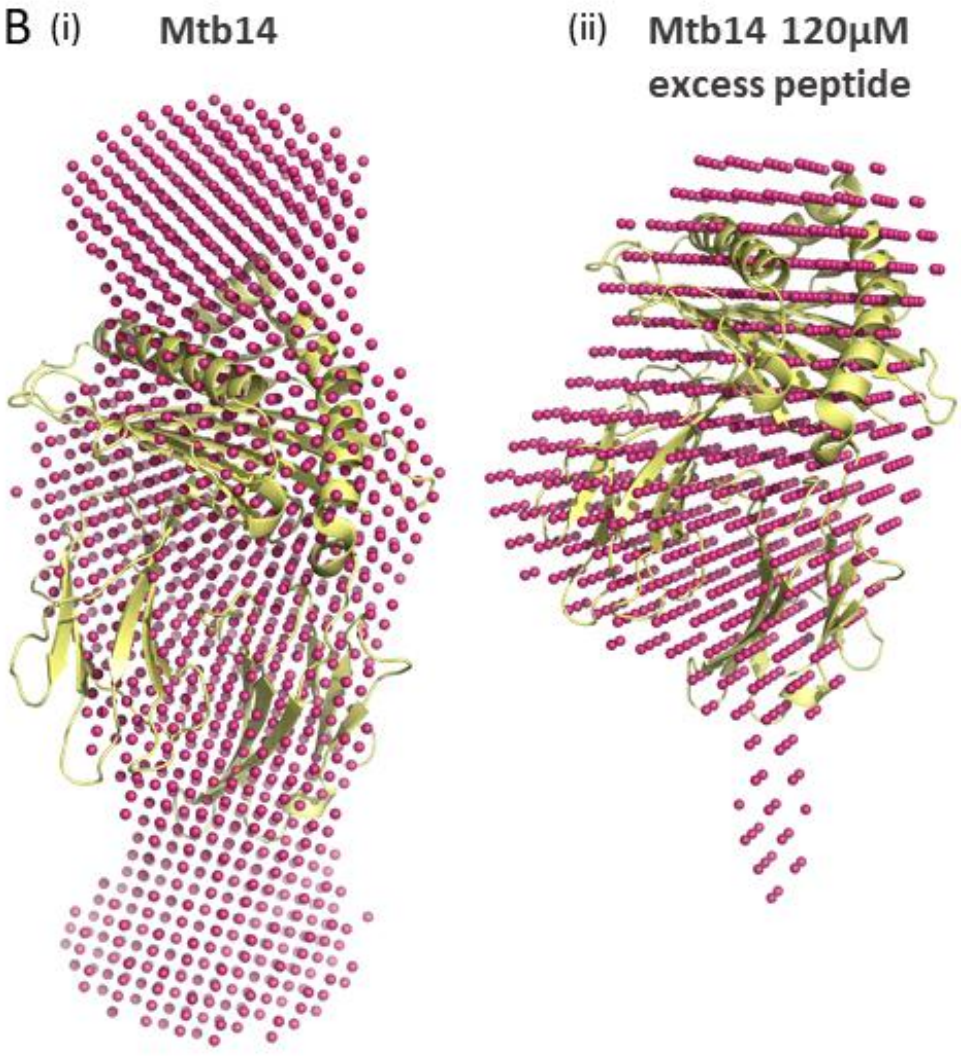


Figure 27: SEC-SAXS analysis for Mtb14-refolded HLA-E*0103 with HPLC elution buffer excess peptide

A. Table detailing the radius of gyration (R_g) in Å, molecular weight (MW) in Daltons and the maximum dimension (d_{max}) in Å calculated by the SEC-SAXS-dedicated analysis software, ScÅtter, in association with ATSAS for HLA-E refolded with Mtb14. Such statistics were generated following buffer subtraction

across the entire HPLC-eluted protein peak for Mtb14-refolded HLA-E in the absence and presence of 120μM excess peptide in the HPLC elution buffer. **B.** Superposition of low resolution DAMMINF SEC-SAXS-based bead models (magenta spheres) to high resolution crystallographic coordinates of HLA-E (yellow cartoon) using the Supcomb package in ATSAS. Bead models are based on buffer-subtracted real-space-refined SEC-SAXS data across the entire HPLC-eluted protein peak for Mtb14-refolded HLA-E in the absence **(i)** and presence of 120μM **(ii)** excess peptide in the HPLC elution buffer.

Mtb44-refolded HLA-E, sampled in the presence or absence of excess peptide in the HPLC elution buffer, yields DAMMINF bead models with compact overall configurations in addition to similar R_g values which vary by less than 1 Å and d_{max} values that differ by less than 5 Å (Figure 28). However, slight differences are detectable in the bead model forms. Specifically, the bead model point ensemble projects above the α2 helical kink of superimposed crystallographic coordinates for Mtb44-refolded HLA-E in the absence of excess peptide, whereas this protrusion is lost following the addition of excess peptide.

A

Peptide:	R_g (Å)	MW (Da)	d_{max} (Å)
Mtb44	25.91	5.40E+04	97
Mtb44 120µM excess peptide	26.02	5.40E+04	93

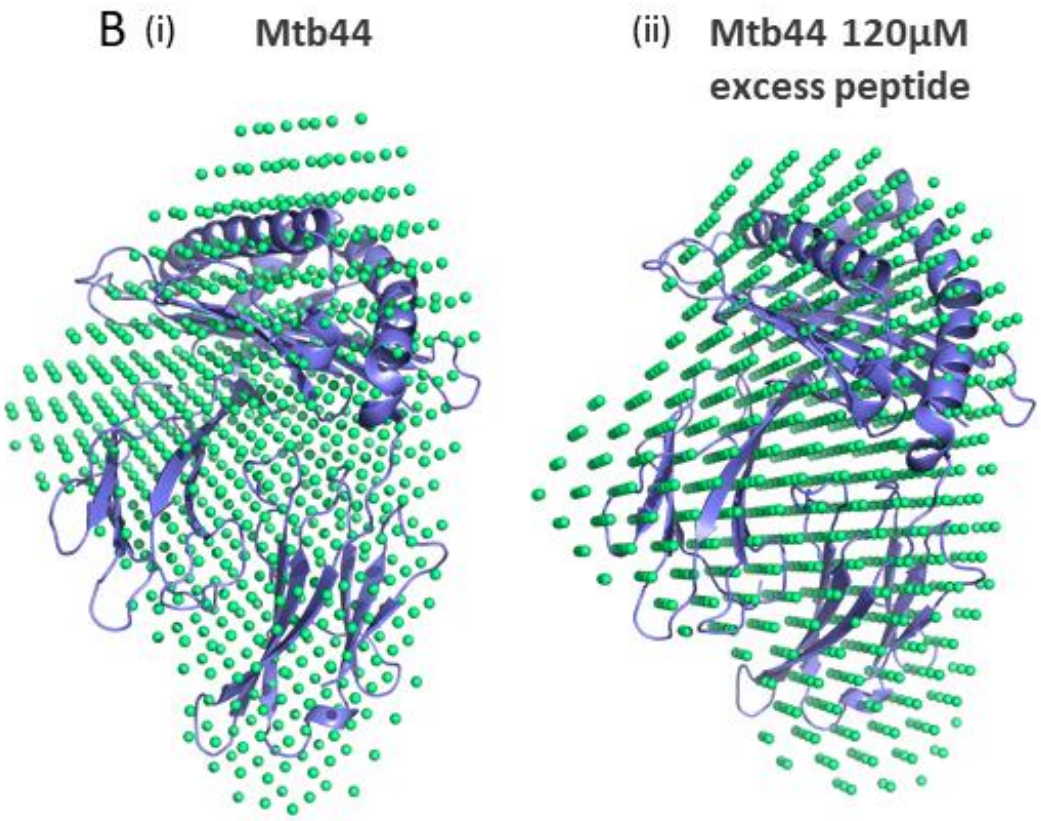


Figure 28: SEC-SAXS analysis for Mtb44-refolded HLA-E*0103 with HPLC elution buffer excess peptide

A. Table detailing the radius of gyration (R_g) in Å, molecular weight (MW) in Daltons and the maximum dimension (d_{max}) in Å calculated by the SEC-SAXS-dedicated analysis software, ScÅtter, in association with ATSAS for HLA-E refolded with Mtb44. Such statistics were generated following buffer subtraction across the entire HPLC-eluted protein peak for Mtb44-refolded HLA-E in the absence and presence of 120µM excess peptide in the HPLC elution buffer. **B.** Superposition of low resolution DAMMINF SEC-SAXS-based bead models (green spheres) to high resolution crystallographic coordinates of HLA-E (purple cartoon) using the Supcomb package in ATSAS. Bead models are based on buffer-subtracted

real-space-refined SEC-SAXS data across the entire HPLC-eluted protein peak for Mtb44-refolded HLA-E in the absence **(i)** and presence of 120μM **(ii)** excess peptide in the HPLC elution buffer.

The average maximum dimension across the protein complex in solution only marginally increased by 3 Å for VL9-refolded HLA-E when 120μM excess peptide was added to the HPLC elution buffer (Figure 29, A). Equally, the R_g values were highly similar (25.55 Å and 25.25 Å) regardless of excess peptide, indicative of similar mass distributions within the macromolecules. Although the corresponding bead models both adopted the compact configuration, slight differences were observed (Figure 29, B). For example, multiple β2M-based surface loops in the structural coordinates of HLA-E protruded beyond the perimeter of the point ensemble generated for VL9-refolded HLA-E in the absence of excess peptide on the HPLC column. By contrast, such surface loops were almost totally enveloped by the point ensemble generated for VL9-refolded HLA-E in the presence of 120 μM excess peptide. Alternatively, a surface loop connecting β-strands of the peptide binding groove in the structural coordinates of HLA-E extends beyond the bead model for VL9-refolded HLA-E with 120μM excess peptide. However, it is possible such minor perturbations in low resolution bead model point ensembles simply reflect experimental or analytical error and thus repeated SEC-SAXS runs will be necessary to validate potential biophysical relevance.

A

Peptide:	R_g (Å)	MW (Da)	d_{max} (Å)
VL9	25.55	5.20E+04	87
VL9 120µM excess peptide	25.25	5.40E+04	90

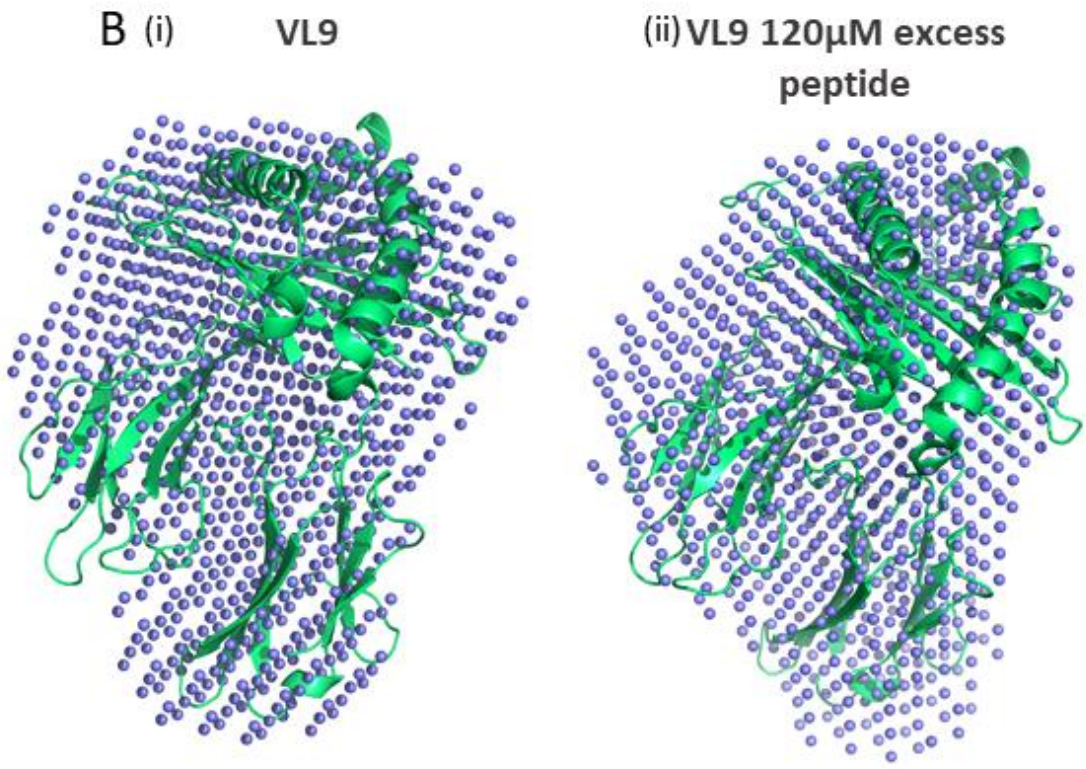


Figure 29: SEC-SAXS analysis for VL9-refolded HLA-E*0103 with HPLC elution buffer excess peptide

A. Table detailing the radius of gyration (R_g) in Å, molecular weight (MW) in Daltons and the maximum dimension (d_{max}) in Å calculated by the SEC-SAXS-dedicated analysis software, ScÅtter, in association with ATSAS for HLA-E refolded with VL9. Such statistics were generated following buffer subtraction across the entire HPLC-eluted protein peak for VL9-refolded HLA-E in the absence and presence of 120µM excess peptide in the HPLC elution buffer. **B.** Superposition of low resolution DAMMINF SEC-SAXS-based bead models (purple spheres) to high resolution crystallographic coordinates of HLA-E (green cartoon) using the Supcomb package in ATSAS. Bead models are based on buffer-subtracted real-

space-refined SEC-SAXS data across the entire HPLC-eluted protein peak for VL9-refolded HLA-E in the absence **(i)** and presence of 120μM **(ii)** excess peptide in the HPLC elution buffer.

Structural modelling of SAXS intensity profiles

The next stage of our SEC-SAXS study on HLA-E shall constitute structural modelling of SAXS intensities to yield atomistic coordinate files reflecting sampled conformations in solution (Schneidman-Duhovny et al., 2013). Through the application of various sampling algorithms to the structural coordinates of HLA-E, an array of coordinate files are generated, each containing a distinct conformation. Hypothetical SAXS intensity profiles are subsequently calculated for the coordinate files and aligned to the experimentally-obtained SAXS intensity. The coordinate file which corresponds to the hypothetical SAXS intensity with the best fit to the experimentally derived SAXS intensity, is selected as a model for downstream interpretation. Although preliminary work using the AllosMod-FoXs web-based server has been conducted for SAXS intensities of HLA-E refolded with VL9, IL9, Mtb14 and RL9HIV, chi-squared values, which indicate the quality of fit between hypothetical and real SAXS intensities, were suboptimal (>2) (Schneidman-Duhovny et al., 2010) (Schneidman-Duhovny et al., 2016) . We are therefore working in collaboration with Professor Robert Rambo at Diamond Light Source who has recently developed an alternative, currently unpublished software for the structural modelling of SAXS intensities. Such modelling might shed light on the structural conformations underlying the elongated bead models of, for example, RL9HIV-bound HLA-E (Figure 18), and might provide structural insight regarding the peptide-dependent conformational adjustments observed between compact bead model configurations generated for peptide-refolded HLA-E in the presence of excess peptide in the HPLC elution buffer (Figures 24-29).

Discussion

Conformationally distinct HLA-E-restricted peptide binding motifs

Despite the presence of canonical anchor residues at positions 2 (Met) and 9 (Leu), the conformational binding motifs adopted by the HIV-derived RL9HIV and Mtb-derived IL9 peptides deviated from the classical kinked orientation displayed by VL9 (O'Callaghan et al., 1998). The alternate peptide binding configuration assumed by RL9HIV yields a distinct solvent-exposure profile whereby the position 6 and 7 α carbons respectively project 4.1 Å and 3.4 Å further toward the solvent relative to VL9. Secondary anchor side chains are typically buried within the C and E pockets at positions 6 and 7 of VL9 peptides, respectively. However, both such side chains assume more solvent-exposed positions in RL9HIV. Marked differences in solvent-accessible regions might facilitate $CD8^+$ T cell discrimination between VL9 and RL9HIV *in vivo* - RL9HIV immunogenicity was previously demonstrated in rhesus macaques vaccinated with RhCMV 68-1 vectors recombinant for HIV Gag (Walters et al., 2018). Given that seven of the nine residues are preserved between RL9HIV and RL9SIV including the buried primary anchor residues, it could be predicted that similar configurations are adopted by both peptides in the binding groove. Thus, with the distinct solvent-exposure profile possibly extending to RL9SIV, this could account for its superior immunogenicity in rhesus macaques vaccinated with RhCMV 68-1 vectors recombinant for SIV Gag - RL9SIV constitutes one of two 'supertopes' which are recognised by 100% of vaccinated macaques (Hansen et al., 2016).

Similarly, the conformational motif adopted by IL9 in molecules 1 and 4 of the asymmetric unit exhibited main chain readjustment at position 7 of up to 2 Å, with implications for stability, owing to the absence of secondary anchor side chain-E pocket interactions. Resultant discrepancies in solvent-exposure profiles within this region of the IL9 main chain relative to VL9 could also account for the immunogenic nature of the Mtb-derived epitope; $CD8^+$ T cell

responses were stimulated by IL9 in multiple HLA-A2-negative donors latently infected with Mtb (McMurtrey et al., 2017).

As previously discussed in Results Chapter 2, minimal pocket reorganisation was observed for HLA-E when bound to peptides containing biochemically distinct amino acids at the same position – such as substitution of the canonical position 2 Met to Leu, Phe or a polar Gln residue. For the IL9- and RL9HIV-bound HLA-E structures presented in this chapter, minimal pocket reorganisation was observed not only following substitution but also when secondary anchor side chains projected away from their corresponding pockets toward the solvent. Such repositioning is thought to disrupt pocket-anchor side chain interactions generating pocket-based cavities, previously linked to complex instability (Strong et al., 2003) (Xu et al., 1998). The observed structural rigidity of HLA-E, in the face of disrupted secondary anchor accommodation, further supports molecular dynamics simulations which identified an unusually inflexible HLA-E peptide binding groove with enhanced structural integrity relative to classical HLA-A2 molecules in the absence of bound peptide (Hansen et al., 2016).

Both RL9HIV and IL9 contain a Tyr at position 3 in place of the canonical Ala typically present in VL9 peptide variants. Given the shallow nature of the position 3-accommodating D pocket in the HLA-E binding groove, the larger Tyr side chain does not assume a downward-pointing position which might otherwise result in steric clashes with D pocket-forming residues. Instead, an alternative, conserved position is adopted in both RL9HIV and IL9 with superposable Tyr side chains projecting toward the α2-helix with which a novel hydrogen bond, not present in VL9-bound structures of HLA-E, is formed. Specifically, the OH group of the Tyr side chain hydrogen bonds with the side chain of Glu-152, triggering this heavy chain-derived residue to reorient and project further toward the peptide N-terminus in both IL9- and RL9HIV-bound HLA-E relative to VL9-bound HLA-E. Such conserved positioning and hydrogen

bonding of this large side chain over the shallow D pocket in both RL9HIV- and IL9-bound HLA-E redefines what can be stably accommodated at position 3 in HLA-E-restricted peptides.

A superposable position 3 Tyr with conserved inter-chain hydrogen bonding belonging to an octameric vesicular stomatitis virus nucleoprotein-derived peptide was detected in the peptide binding groove of the murine MHC class I molecule, H2-K^b (Fremont et al., 1992). Previously, position 3 Tyr selectivity had been identified as part of the sequence motif for H2-K^b-restricted peptides (Van Bleek and Nathenson, 1991). Interestingly, the same study identified that the H2-K^{bm1} mutant containing a Glu to Ala substitution at position 152 loses its preferential binding to peptides with Tyr at position 3. By contrast, numerous MHC class I molecules contain an Ala or Val at position 152 - such as HLA-C*07:02 (Ala-152), HLA-B35 (Val-152), HLA-A2 (Val-152) and HLA-A68 (Val-152) - instead of Glu, which is present at position 152 of both H2-K^b and HLA-E. Considering the side chain of Glu-152 is capable of forming conserved hydrogen bonds with position 3 Tyr of both HLA-E- and H2-K^b-bound peptides, it possibly constitutes a contributing factor in rendering the HLA-E and H2-K^b grooves permissive to position 3 Tyr binding.

Notably, the H2-K^{bm1} mutant also possesses an Arg to Tyr mutation at position 155 of the murine heavy chain (Van Bleek and Nathenson, 1991). H2-K^b-derived Arg-155 was also subsequently shown to hydrogen bond with the side chain of peptide position 3 Tyr, in addition to Glu-152 (Fremont et al., 1992). It is possible the combined effects of Arg-155 and Glu-152 are responsible for position 3 Tyr secondary anchor selectivity in the H2-K^b peptide motif. Although Glu-152 is conserved between H2-K^b and HLA-E, position 155 differs via an Arg to His substitution in the HLA-E heavy chain relative to H-2K^b. Consequently, this may result in weaker binding preferences for peptide position 3 Tyr by HLA-E since hydrogen bonds are not formed between His-155 and position 3 Tyr of RL9HIV or IL9. In summary, the presence of large, sterically incompatible side chains over shallow pockets in the MHC class I binding

groove may disrupt stabilising pocket-based interactions with otherwise smaller, downward-projecting anchor side chains. However, preserved inter-structural positioning of position 3 Tyr in HLA-E- and H2-K^b-bound peptides provides an example of stable, ordered secondary anchor side chain positioning in which novel inter-chain hydrogens bonds may compensate for a lack of pocket-anchor side chain shape complementarity.

Differential structural positioning of the IL9 peptide main chain

Two distinct IL9 peptide configurations were observed in the asymmetric unit. Specifically, in addition to the IL9 position 7 side chain projecting in different directions, main chain readjustments exceed 2.0 Å. Variable main chain and side chain positioning within the central peptide region commonly features in peptide-bound structures of the same MHC class I molecule (Madden et al., 1993) (Madden, 1995). Further, analogous to IL9, peptide conformational dimorphism has previously been reported for classical MHC class I-bound peptides, such as the HLA-B*2705-restricted glucagon receptor-derived self-peptide, pGR, which adopts two distinct binding configurations within the HLA-B*2705 groove (Rückert et al., 2006). However, only subtle inter- or intra-structural differences in VL9 variant positioning have previously been reported for HLA-E (O'Callaghan et al., 1998) (Strong et al., 2003) (Hoare et al., 2008).

It is possible that the multiple, redundant, non-lattice-symmetry-related views of IL9-bound HLA-E afforded by the presence of four molecules in the asymmetric unit provide insight into variable peptide positioning in solution. IL9 may fluctuate between the two structurally-observed conformations, with the presence of multiple lattice-constrained configurations potentially reflecting increased peptide dynamics. As discussed in Results Chapter 2, a broadened MHC class I binding groove has previously been correlated with enhanced rms peptide fluctuation in molecular dynamics simulations of 52 HLA-A2 molecules (Ayres et al., 2019). Accordingly, the central portion of the α2 helix in all four molecules of IL9-bound HLA-E

is shifted by up to 1.6 Å relative to that of VL9-bound HLA-E. Resultant localised increases in IL9-bound HLA-E groove width would be consistent with a highly fluctuating peptide. Interestingly, peptide-dependent MHC class I binding groove contraction has recently been correlated with inferior immunogenicity in a molecular dynamics study simulating HLA-A*02:01 in complex with a collection of known $CD8^+$ T cell epitopes (Pang et al., 2018). Thus, it is possible that broadening of the IL9-bound HLA-E groove relative to VL9 contributes to its immunogenicity *in vivo*, given the identification of multiple IL9-responding HLA-A2-negative donors latently infected with Mtb (McMurtrey et al., 2017).

In addition to variable positioning of the α2 helix between IL9-bound and VL9-bound HLA-E, intra-structural readjustments between molecules of the asymmetric unit – involving shifted β-strand signatures and 220's loop positioning in the α3 domain - are observed for both IL9- and RL9HIV-bound HLA-E. Such distinct, intra-structural positioning of the α3 domain is also observed for HLA-E bound to Mtb44 and Mtb14 (Results Chapter 2, Figures 2 and 13). Potential physiological implications for these movements are discussed at length in the discussion section of Results Chapter 2 and shall therefore not be further elaborated upon here.

Insights from SEC-SAXS analyses of HLA-E refolds

SEC-SAXS analyses of both RL9SIV- and EK9-refolded HLA-E yielded elongated bead models along multiple axes; the point ensemble of the bead model not only projects above and below superimposed structural coordinates of HLA-E, but also protrudes horizontally to the side of the peptide binding groove. By contrast, an alternate bead model configuration is observed for the three HLA-E-refolded pathogen-derived peptides, Mtb14, RL9HIV and IL9, for which corresponding crystal structures are presented in Results Chapters 2 and 3. Specifically, Mtb14-, RL9HIV- and IL9-refolded HLA-E yield DAMMINF point ensembles which appear to elongate along a single vertical axis. Contracted bead models following the addition of excess

peptide to the HPLC elution buffer implicates weak peptide binding affinity as the key contributing factor to the extended average conformations observed in solution. The majority of SIV- and HIV-derived RhCMV 68-1 vaccine-identified epitopes exhibit low HLA-E binding affinity, as evidenced by binding assays studies (Figures 1.17a, 1.17b and 1.18 Results Chapter 1). It is therefore possible that increased HLA-E complex plasticity or disorder imparted by low affinity ligands, relative to the more compact VL9-bound HLA-E configuration, is of physiological relevance.

RhCMV 68-1 SIV/HIV vaccine-identified TCRs exhibit unusually extensive cross-reactivity across the broad array of motif-less SIV/HIV epitopes (Unpublished data, Louis Picker and Jonah Sacha Labs). TCR cross-reactivity comprises a critical property presumably selected for, over the course of almost 400 million years of peptide-MHC-TCR interface coevolution, to compensate for superior diversity of presented antigenic peptide relative to the available TCR repertoire (Garcia and Adams, 2005). However, the level of degeneracy observed among RhCMV 68-1-stimulated TCRs is unusually high, with a small number of TCRs covering the entire breadth of vaccine-identified SIV-derived epitopes (Unpublished data, Louis Picker and Jonah Sacha Labs). One possible explanation for such unprecedented cross-reactivity constitutes indirectly mediated peptide specificity whereby RhCMV 68-1 SIV/HIV-identified Mamu-E-restricted TCRs exclusively recognise and interact with peptide-induced conformational differences in the MHC class I protein.

Differential, peptide-dependent conformational sampling of MHC class I proteins has been demonstrated both experimentally and via molecular dynamics simulations (Ayres et al., 2019). A crystallographically-determined example constitutes peptide-dependent rearrangement of the HLA-A2-derived α2 helix upon A6 TCR interaction (Borbulevych et al., 2009). Subsequent biophysical analyses revealed that peptide-dependent motional properties extend to other MHC class I alleles such as murine H-2K^{b}, are not restricted to specific peptide

ligands and encompass the α1 helix, α2 helix and β-sheet floor of the peptide binding groove in addition to the α3 domain (Hawse et al., 2013) (Ayres et al., 2019). This phenomenon of peptide-driven differential MHC class I conformational sampling has been predicted to constitute a presently largely undetected layer of 'altered self' antigenicity (Wieczorek et al., 2017).

In summary, differential dynamics of MHC class I proteins governed by bound peptide have been shown to impart important physiological consequences. Peptides influence the energy barriers dictating MHC class I dynamic sampling resulting in differential sub-state stability and peptide-dependent entropic costs of TCR engagement. The strikingly cross-reactive RhCMV 68-1-stimulated Mamu-E-restricted $CD8^+$ T cell subtype could represent a more innate-like phenotype in which TCRs recognise conserved peptide-dependent conformational patterns in the MHC-E heavy chain that presumably arise following non-VL9 suboptimal peptide loading. Although distinct bead model configurations for peptide-refolded HLA-E might reflect relevant differences in the conformational ensemble, higher resolution, dynamically-informative techniques such as NMR would be required to correlate specific, peptide-induced conformational motifs with patterns of TCR cross-reactivity.

Obtaining absolute affinity measurements is technically challenging for suboptimal, low affinity MHC class I-restricted peptide given the propensity of free heavy chain to aggregate with peptides that provide insufficient binding energy (Hawse et al., 2013). However, marked differences in bead model configurations for RL9HIV-, Mtb14-, IL9- and RL9SIV-refolded HLA-E in the absence or presence of excess peptide in the HPLC elution buffer are indicative of fast off rates and weak peptide binding affinity - given that all peptide-HLA-E refolds in SEC-SAXS analyses contained excess peptide in the HPLC injection buffer. Previously, biophysical analyses involving hydrogen-deuterium exchange mass spectrometry and fluorescence anisotropy of classically-restricted HLA-A2-bound peptides were unaffected by the presence or

absence of excess peptide (Hawse et al., 2013). Such results are consistent with lower peptide binding affinities of HLA-E-restricted pathogen-derived peptides. Even the strongest binding peptide identified in peptide exchange-based binding assays, Mtb44, generated a slightly contracted bead model in the presence of excess peptide relative to the sample with no excess peptide in the HPLC elution buffer. Differences in the overall conformational ensemble of Mtb44-, IL9-, RL9HIV-, and Mtb14-bound HLA-E might have been overlooked in the absence of molecular or biophysical analyses since crystallographic coordinates of all such peptide-bound complexes yielded superposable global architectures of HLA-E with similar dimensions. The low-energy native state of HLA-E-β2M-peptide complexes may have been selected for incorporation into the crystal lattice due to reduced plasticity and disorder. However, elongated bead model configurations suggest corresponding protein structures of RL9HIV-Mtb14- and IL9-bound HLA-E likely represent a subpopulation of the conformational ensemble in solution.

SEC-SAXS data are consistent with blue-native-polyacrylamide gel electrophoresis (PAGE) and thermal melt analyses of purified, peptide-refolded HLA-E, conducted by Professor Geraldine Gillespie, NDM, University of Oxford (Walters et al., 2018). VL9- and Mtb44-refolded HLA-E yield compact blue-native gel bands in addition to compact SEC-SAXS bead model configurations and constant R_g values across the HPLC elution peak. Such results are indicative of sample uniformity and likely represent peptide-loaded optimally-folded material. However, RL9HIV-refolded HLA-E generated a diffuse band on the blue native gel in addition to a faint compact band - the former was hypothesised to reflect partially peptide-bound or peptide-dissociated flexible transition states since both HLA-E heavy chain and β2M could be detected when the diffuse band was tested in a second dimension via denaturing SDS-PAGE. Further, immediate incorporation of fluorescent dye not only prohibited the identification of a reliable Tm in thermal melt analyses of RL9HIV-refolded HLA-E, but also indicated the presence of a degree of protein unfolding and sample non-uniformity. Thus, elongated SEC-SAXS bead

models, also indicative of partially unfolded material, in addition to differentially elongated bead models following HPLC protein peak partitioning, which indicate conformational heterogeneity among RL9HIV-refolded HLA-E complexes, are consistent with diffuse blue-native gel bands and irreproducible thermal melt profiles. Equally, the faint, compact blue-native gel band likely represents the fully-folded, peptide-loaded subpopulation of RL9HIV-bound HLA-E, which exists in solution, dominates the conformational ensemble in the presence of 120µM excess peptide in the HPLC elution buffer and was exclusively incorporated into the crystal lattice.

Unsuccessful attempts to crystallise low affinity peptide-bound HLA-E

Crystallisation plates were set up for SIV- and HIV-derived epitopes that yielded low relative HLA-E binding affinity in peptide-exchange-based assays presented in Results Chapter 1. Specifically, sparse matrix screens were assembled for HLA-E-refolded with the SIV-derived peptide AE9 and HIV-derived peptide AA9, however crystal growth did not ensue. It is possible this was due to the dissociation of sub-optimally-bound peptides from the binding groove resulting in protein complex disintegration. Alternatively, weakly peptide-bound unstable or partially disordered transition states may have prohibited the formation of a uniform, ordered lattice, given that protein stability is a critical factor permitting crystal formation (Deller et al., 2016). Elongated SEC-SAXS DAMMINF bead models for HLA-E refolded with a range of intermediate affinity binding peptides (such as RL9HIV, Mtb14, BZLF1, IL9 and RL9SIV) support such a hypothesis indicating that the conformational ensemble is dominated by partially unfolded and disordered complexes. Further, stratified analyses of SAXS intensities according to the HPLC elution peak demonstrated intra-sample conformational heterogeneity - a phenomenon also commonly thought to impede crystal lattice formation (Deller et al., 2016). Considering elongated configurations and conformational heterogeneity were observed in protein samples containing HLA-E refolded with crystallisable peptide ligands such as RL9HIV,

Mtb14 and IL9, it is likely such prohibitive features are more pronounced in HLA-E refolds containing lower affinity peptides. Protein complexes incorporated into crystal lattices of IL9-Mtb14- and RL9HIV-bound HLA-E likely comprise the native sub-state sampled in solution. However, the formation or proportions of such crystallisable sub-states may be compromised for purified HLA-E pre-refolded with lower affinity peptides.

Overall discussion and concluding remarks

The peptide binding assay data presented in this thesis demonstrate that a diverse range of peptides reproducibly stabilise HLA-E to varying degrees above peptide-free background levels. A number of these HLA-E-binding peptides contain non-hydrophobic, chemically disparate primary and secondary anchor residues, such as positively charged Lys or negatively charged Glu in place of the canonical primary anchor position 2 Met or position 9 Leu. The structural basis for non-canonical primary anchor accommodation by the HLA-E peptide binding groove was subsequently determined for a subset ELISA-identified alternative anchor residues including position 2 Gln and Phe and position 9 Phe - although on the background of an optimal HLA-E binding peptide, Mtb44, with suitable remaining anchor residues in place (Figures 26-36, Results Chapter 2) (Walters et al., 2018). Early experiments delineating the HLA-E binding motif applied Ala/Gly scanning substitutions along the VL9 peptide (Braud et al., 1997) or analysed acid eluted peptide from *in vitro* refolded HLA-E with VL9 anchor residue variant peptide libraries via mass spectrometry (Miller et al., 2003). Such experiments successfully established an optimal peptide binding motif for HLA-E comprising primary anchor position 2 Met and position 9 Leu with additional selectivity for secondary anchor positions 3, 6 and 7. However, the capacity of the HLA-E binding groove to accommodate suboptimal peptides and the sequence identity of potential non-VL9 HLA-E-binding peptides was not fully elucidated. For example, the sensitivity of peptide elution from *in vitro* refolded HLA-E would have been compromised by the co-presence of an excess of almost optimal VL9 variant peptide in the pool - such as those with position 2 Met to Leu or Ile substitutions which were shown to only marginally reduce binding affinity to HLA-E (Miller et al., 2003). Peptides with lower relative binding affinities would likely have been outcompeted in these settings and thus not detected. By contrast, the peptide binding assay presented in Results Chapter 1 afforded superior sensitivity as the binding capacity of a large range of suboptimal HLA-E binding

peptides was tested in the absence of competition from higher affinity peptide. Overall, our peptide binding data indicate that the HLA-E binding groove is highly promiscuous and capable of accommodating a diverse set of suboptimal sequences which lack a common motif and contain chemically dissimilar non-canonical anchor side chains. However, our results are also consistent with the conclusions of previous work that HLA-E has evolved to preferentially accommodate MHC class I leader peptides (Braud et al., 1997) (Miller et al., 2003) (O'Callaghan et al., 1998) – as only a single pathogen-derived peptide, Mtb44 (RLPAKAPLL), rivalled the binding strength of positive control VL9 peptides in the peptide binding assay (Figure 21, Results Chapter 1) and generated compact average bead models comparable to those produced for VL9-refolded HLA-E in SEC-SAXS experiments (Figure 18, Results Chapter 3).

Binding of suboptimal peptides by HLA-E is largely masked under normal physiological conditions with a combination of factors operating in concert to ensure the HLA-E peptidome is dominated by CD94/NKG2A-interacting MHC class I leader peptide (McMurtrey et al., 2017) to protect healthy cells from NK cell-mediated lysis (V. M. Braud et al., 1998). The evolutionarily fine-tuned complementarity of the HLA-E binding groove for its optimal VL9 peptide ligand in combination with HLA-E/VL9 co-localisation in the ER and the sheer abundance of VL9 peptides - with expression of certain MHC class I molecules reaching 20-fold greater levels than that of HLA-E - allows VL9 variants to dominate the HLA-E binding repertoire (Veronique M. Braud et al., 1998) (N Lee et al., 1998). However, the alternative functional role of HLA-E appears to emerge when normal physiological conditions are disrupted, for example when the MHC class I antigen presentation pathway is besieged in transformed or infected cells (Oliveira et al., 2010) (Hansen et al., 2016). Virally- or intracellular bacterially-mediated downregulation of classical MHC class I molecules not only reduces the supply of HLA-E binding VL9 peptide ligand but it also decreases competition for β2M association - the latter of which may be particularly crucial given that classical MHC class I

molecules are reported to exhibit higher affinities for β2M than HLA-E in the presence of natural peptide ligands (Ulbrecht et al., 1999). Additionally, an exogenous supply of excess β2M has previously been shown to stabilise peptide-free MHC class I molecules on the cell surface suggesting that increased availability of free β2M can relieve the requirement to bind and present optimal peptide ligands (Townsend et al., 1990). It is therefore possible that reduced competition for β2M may contribute to the upregulation of HLA-E surface expression detected in multiple contexts where classical MHC class I molecules are partially downregulated such as in Hepatitis C Virus infection (Schulte et al., 2009), HIV infection (Nattermann et al., 2005b) or in tumour cells of diverse origin (Marín et al., 2003).

Although HLA-E has previously been reported to associate with β2M at lower affinities relative to classical MHC class I molecules (Ulbrecht et al., 1999), several lines of evidence suggest that both HLA-E and its murine counterpart Qa-1 are able to remain associated to β2M in the absence of optimal or suboptimal peptide ligand (Kambayashi et al., 2004) (Walters et al., 2018). First, unusually high peptide-free background signals were observed in the micro-refolding-based peptide binding assay for HLA-E (Figures 1-3, Results Chapter 1) (Walters et al., 2018) relative to HLA-A*0201 and HLA-A*1101 (Sylvester-Hvid et al., 2002) despite identical molar concentrations of MHC heavy chain and β2M, indicating that HLA-E association with β2M is less dependent on the presence of a suitable peptide ligand. Second, identically positioned protein peaks were present on the HPLC elution trace for VL9 or peptide-free refolded HLA-E during SEC-SAXS experiments presented in Results Chapter 3. Considering superposable HPLC elution peaks reflect material of similar molecular weight, it is likely that the protein population in the elution peak from the peptide-free HLA-E refold was β2M-bound, although further exploration using higher resolution HPLC columns will be required to confirm this. Third, Professor Geraldine Gillespie has also demonstrated β2M-association by HLA-E in the absence of peptide ligand via blue native gel analysis, where peptide-free refolded and purified HLA-E complexes formed a diffuse gel band from which individual HLA-E

heavy chain and β2M components could be detected via subsequent 2D gel separation (Walters et al., 2018). Importantly, it was also shown that such peptide-free HLA-E complexes were peptide-receptive and could be converted into peptide-loaded material, as evidenced by the formation of compact blue native gel bands following the addition of excess peptide. Finally, such results are supported by the detection of unusually stable peptide-free Qa-1 complexes on the cell surface which are thought to arise due to rapid Qdm dissociation kinetics in addition to an intrinsic capacity of Qa-1 to remain β2M-associated in the absence of a peptide ligand (Kambayashi et al., 2004).

The detection of HLA-E in peripheral intracellular compartments such as in Mtb-infected phagosomes (Grotzke et al., 2009) or autophagosomal-lysosomal compartments in newly differentiated macrophage (Camilli et al., 2016), suggests certain stimuli such as disrupted antigen processing might also lead to the re-routing or differential recycling of HLA-E-β2M complexes. Such peripherally-sequestered HLA-E would be exposed to alternative pools of exogenous antigen which, in combination with reduced competition from higher affinity VL9 peptides in settings of downregulated classical MHC class I molecules such as Mtb infection, could enable a diversification of the peptide repertoire presented by HLA-E. Further, given the rapid dissociation kinetics observed for both VL9 and Qdm peptides from HLA-E and Qa-1, respectively, relative to classical MHC class I molecules and the propensity of unusually stable peptide-free Qa-1 complexes to accumulate at the cell surface (Kambayashi et al., 2004), it is likely that cell surface expressed HLA-E would have lost its VL9 ligand prior to recycling into early endosomal compartments permitting re-loading with lower affinity, diverse antigen. It is also possible that the alternative, endosomally-based MHC class II-like trafficking pathway might favour the presentation of low affinity HLA-E-binding peptides due to the absence of MHC class I chaperones, such as tapasin, in peripheral intracellular compartments; chaperoning effects might otherwise enforce stringent quality control over peptide loading resulting in the retention of sub-optimally loaded MHC class I complexes (Wearsch and

Cresswell, 2008). It is unlikely that low affinity Mamu-E-restricted SIV epitopes identified in RhCMV 68-1-SIV vaccination studies (Hansen et al., 2016) would provide sufficient binding energy to outcompete the stabilising interfaces formed between the MHC class I-dedicated chaperone, tapasin, and HLA-E (A. P. Williams et al., 2002) (Grandea and Van Kaer, 2001) (Veronique M. Braud et al., 1998). The remarkable sensitivity of TCR recognition might allow for the detection of sub-optimally-bound HLA-E complexes with short half-lives at the cell surface, analogous to low affinity Mamu-E-restricted SIV epitope recognition by Mamu-E-restricted $CD8^+$ T cells following RhCMV 68-1-SIV vaccination (Hansen et al., 2016). In conclusion, abnormal physiological conditions might promote the cooperative elimination of stressed, transformed and infected cells by HLA-E-restricted $CD8^+$ T cells in addition to NK cells - since inhibitory CD94/NKG2A interactions would not be preserved when presented with a diversified HLA-E peptidome largely comprising peptides with incompatible combinations of CD94-interacting residues at positions 5-8 (Kaiser et al., 2008) (Miller et al., 2003).

Peptide binding assay signals for certain SIV epitopes identified in RhCMV 68-1-SIV vaccination studies only marginally exceeded those generated by the peptide-free negative control (Figure 17a, Results Chapter 1). Further, SEC-SAXS-based analyses of the SIV supertope (RL9SIV) that yielded the strongest signal in the binding assay were suggestive of very low affinity binding with the average conformation in solution changing dramatically in the presence or absence of excess peptide supplied in the HPLC elution buffer (Figure 25, Results Chapter 3). By contrast, VL9-refolded HLA-E appeared largely insensitive to the addition of excess peptide (Figure 29, Results Chapter 3). It could therefore alternatively be argued that despite the highly sensitive nature of $CD8^+$ T cell recognition, exceptionally low HLA-E binding affinity of RhCMV 68-1-SIV vaccine-identified SIV epitopes would result in premature peptide dissociation prohibiting intracellular trafficking and cell surface presentation to $CD8^+$ T cells. How such remarkably low affinity SIV-derived epitopes generate robust, protective MHC-E-restricted $CD8^+$ T cell responses in RhCMV 68-1-vaccintaed rhesus macaques, remains enigmatic. It is possible that

the conditions of our *in vitro* assays poorly reflect the intracellular environments that support MHC-E peptide loading in RhCMV 68-1 vaccinated cells. It is equally possible that the conventional mode of peptide binding might not apply in such settings and instead a currently undetected mechanism facilitates suboptimal peptide presentation by MHC-E. For example, it could be speculated that an unorthodox mode of peptide presentation might be operating in such settings which serves to stabilise low affinity peptide binding to MHC-E. Abacavir, an anti-retroviral nucleoside analog reverse transcriptase inhibitor used to treat HIV-infected individuals, was structurally demonstrated to specifically bind to the F pocket of HLA-B*5701, in turn altering the presented peptide repertoire and resulting in self-reactivity (Illing et al., 2012). It is possible that, analogous to this drug-modified repertoire, a currently unidentified F pocket or additional pocket stabilising interaction might exclusively occur in contexts such as RhCMV 68-1-SIV vaccination or Mtb infection, which could either sufficiently stabilise HLA-E complexes to permit the presentation of low affinity non-canonical peptide or alter the landscape of the HLA-E binding groove in turn transforming its peptide binding preferences. For example, it could be possible that a small, highly abundant metabolite upregulated in such settings or a small peptide fragment available in excess in HLA-E-enriched peripheral compartments, might optimally occupy a pocket in the HLA-E binding groove. Accordingly, dipeptide fragments with F pocket complementarity have previously been shown to stabilise folding of classical MHC class I molecules, such as the dipeptide Gly-Leu which supported complex formation of HLA-A*0201 and the murine H-2K^{b} (Saini et al., 2013). Also, and unusually, crystal structures of the classical MHC class I molecule, HLA-A*1101, were determined in complex with a combination of non-contiguous peptide fragments, and these multi-peptide-bound complexes were subsequently shown to elicit CD8^{+} T cell responses (Xiao et al., 2017). Such results support the possibility of dual peptide occupancy in the HLA-E peptide binding groove, for example by short stabilising F pocket-directed fragments perhaps

in combination with low affinity pathogen-derived epitopes that exhibit poor complementarity for HLA-E.

In addition to HLA-E-restricted $CD8^+$ T cell responses directed against pathogen-derived peptide (van Meijgaarden et al., 2015), this unconventionally restricted T cell subset has also been found to recognise self-peptides that reflect cellular stress. The three self-proteins Fam49B, Prdx5 and Hsp60, whose expression levels are differentially regulated in normal or stressed conditions, were found to encode HLA-E-binding peptides that elicited Qa-1- or HLA-E-restricted $CD8^+$ T cell responses (Michaëlsson et al., 2002) (Jiang et al., 2010) (Sensi et al., 2009) (Nagarajan et al., 2012). Prdx5 expression is specifically induced by oxidative stress and its catalytic activity serves to reduce cellular levels of alkyl hydro-peroxidases and peroxynitrite (Rhee et al., 2005). Hsp60 expression is also upregulated during the stress response and helps to prevent heat-induced denaturation of cellular protein (Martin et al., 1992). The functional role and expression pattern of Fam49B have remained more elusive, although a recent study demonstrated that silenced Fam49B expression resulted in a deregulated mitochondrial redox and metabolic state (Chattaragada et al., 2018). Although such proteins are thought to be upregulated by cellular stress, basal expression levels are detected in normal physiological conditions (Sensi et al., 2009) (Chattaragada et al., 2018). This raises the potential issue that such self-peptide might become incorporated into the HLA-E presented peptidome in normal conditions which could in turn result in the elimination of healthy cells by HLA-E-restricted $CD8^+$ T cells. However, an additional, currently undetected mechanism could conceivably be operating in certain contexts to ensure stress-related self-peptides are recognised exclusively in stressed cells. Analogous to N-formylation which consistently yields 10,000 fold increases in peptide binding affinity to H-2M3 in part due to specific interactions formed between the peptide formyl group and His-9 of the peptide binding groove (Wang et al., 1995) (Lenz et al., 1996), it is possible that certain modifications to peptides such as that derived from Fam49B could confer superior binding affinity to HLA-E. Further, it could be speculated that

deregulated intracellular conditions, such as oxidative stress, that arise in abnormal physiological settings might provide a conducive environment for the generation of such modified peptides. Thus, although stress-related self-peptide may be generated in the steady state, such peptides may lack crucial modifications that otherwise enable efficient HLA-E binding and presentation to HLA-E-restricted $CD8^+$ T cells.

RhCMV 68-1-SIV-stimulated Mamu-E-restricted $CD8^+$ T cell responses exhibit an unusual degree of cross-reactivity and a small number of TCRs provide cross-recognition of all vaccine-identified SIV epitopes (Unpublished data, Louis Picker). Given the uncharacteristically high degree of TCR cross-reactivity and the lack of biochemically similar moieties between SIV epitopes recognised by the same TCR, it could be hypothesised that Mamu-E-restricted $CD8^+$ T cell responses exhibit a more innate-like nature; although TCR recognition might be dependent on the presence of a range of biochemically unrelated peptides, the TCRs might not interact directly with such peptides. Instead, peptide-induced conformational changes in the peptide binding groove or distal α3 domain might constitute sites of Mamu-E-restricted TCR recognition – the phenomenon of peptide-induced differential conformational sampling of MHC class I heavy chains has previously been recognised as an additional potential layer of largely undetected antigenicity (Hawse et al., 2013) (van Hateren et al., 2017). It is possible that particular subsets of non-VL9 peptides induce characteristic patterns of conformational sampling at proximal or distal sites in the MHC-E heavy chain which are in turn recognised by highly cross-reactive innate-like Mamu-E-restricted TCRs. Although SEC-SAXS analyses are limited by low resolution it does appear that HLA-E refolded with VL9 peptide produces conformationally distinct, more compact bead models when compared to lower affinity pathogen-derived HLA-E-binding epitopes, which appear to permit greater degrees of conformational plasticity and structural elongation (Figure 18, Results Chapter 3). Further, enhanced structural flexibility of the α3 domain was observed between molecules belonging to the same asymmetric unit in structures of HLA-E bound to pathogen-derived peptides

including Mtb44, Mtb14, IL9 and RL9HIV (Results Chapters 2 & 3) (Walters et al., 2018), but similar movements were not observed in previously determined VL9-bound HLA-E structures (O'Callaghan et al., 1998) (Strong et al., 2003) (Hoare et al., 2008). In summary, it is possible that conformational changes in the HLA-E heavy chain could arise following suboptimal non-VL9 peptide docking, which could in turn be recognised by innate-like HLA-E-restricted $CD8^+$ T cells. In this way, both CD94/NKG2A and HLA-E-restricted $CD8^+$ T cells might have evolved to recognise absence of the HLA-E-presented VL9 peptide - both such receptors might sense the loss of VL9-dependent conformations within the HLA-E heavy chain, with CD94/NKG2A also recognising a loss of direct, VL9 peptide-specific contacts.

In summary we have provided molecular, biophysical and structural evidence for promiscuous peptide sampling by HLA-E. Overall, our results suggest HLA-E is capable of binding a range of low affinity, suboptimal peptide ligands and that the unusual rigidity of the HLA-E peptide binding groove, previously predicted by molecular dynamics simulations (Hansen et al., 2016), persists following primary and secondary anchor substitution or when peptide backbone repositioning yields unoccupied pockets (Walters et al., 2018). It is possible that the superior structural integrity of the HLA-E peptide binding groove in addition to an unusual capacity of HLA-E to remain β2M-associated in the absence of a peptide ligand, contribute to its ability to present a diverse motif-less array of suboptimal epitopes to HLA-E-restricted $CD8^+$ T cells under atypical physiological conditions. Together, our results support the possibility of developing a universally-effective HLA-E-targeted vaccine against HIV in humans, and possibly also for other human pathogens or malignancies. Given that a range of pathogens and tumours upregulate surface expressed HLA-E, presumably to preserve CD94/NKG2A interaction and thus avoid NK cell-mediated lysis, such cells would likely constitute particularly vulnerable targets for vaccine-induced HLA-E-restricted $CD8^+$ T cell responses (Marín et al., 2003) (Nattermann et al., 2005b). To expand upon the results presented in this thesis future experiments shall include attempts to crystallise co-complexes of SIV epitope-loaded Mamu-E

bound to cognate RhCMV 68-1-SIV vaccine-identified TCRs, to elucidate the underlying structural basis of MHC-E-restricted $CD8^+$ T cell recognition of low affinity pathogen-derived peptide.

References

Abd Hamid, M., Wang, R.-Z., Yao, X., Fan, P., Li, X., Chang, X.-M., Feng, Y., Jones, S., Maldonado-Perez, D., Waugh, C., Verrill, C., Simmons, A., Cerundolo, V., McMichael, A., Conlon, C., Wang, X., Peng, Y., Dong, T., 2019. Enriched HLA-E and CD94/NKG2A Interaction Limits Antitumor CD8 + Tumor-Infiltrating T Lymphocyte Responses . Cancer Immunol. Res. 7, 1293–1306.

Adams, P.D., Afonine, P. V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., IUCr, 2010. *PHENIX* : a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 213–221.

Afonine, P. V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D., 2012. Towards automated crystallographic structure refinement with *phenix.refine*. Acta Crystallogr. Sect. D Biol. Crystallogr. 68, 352–367.

Agostini, C., Semenzato, G., Lieberman, J., Shankar, P., Swamy, M., Andersson, J., 2002. Why antiviral CD8 T lymphocytes fail to prevent progressive immunodeficiency in HIV-1 infection [3] (multiple letters), Blood.

Aldrich, C.J., DeCloux, A., Woods, A.S., Cotter, R.J., Soloski, M.J., Forman, J., 1994. Identification of a tap-dependent leader peptide recognized by alloreactive T cells specific for a class Ib antigen. Cell 79, 649–658.

André, P., Denis, C., Soulas, C., Bourbon-Caillet, C., Lopez, J., Arnoux, T., Bléry, M., Bonnafous, C., Gauthier, L., Morel, A., Rossi, B., Remark, R., Breso, V., Bonnet, E., Habif, G., Guia, S., Lalanne, A.I., Hoffmann, C., Lantz, O., Fayette, J., Boyer-Chammard, A., Zerbib, R., Dodion,

P., Ghadially, H., Jure-Kunkel, M., Morel, Y., Herbst, R., Narni-Mancinelli, E., Cohen, R.B., Vivier, E., 2018. Anti-NKG2A mAb Is a Checkpoint Inhibitor that Promotes Anti-tumor Immunity by Unleashing Both T and NK Cells. Cell 175, 1731-1743.e13.

Andreatta, M., Nielsen, M., 2016. Gapped sequence alignment using artificial neural networks: Application to the MHC class i system. Bioinformatics 32, 511–517.

Ayres, C.M., Abualrous, E.T., Bailey, A., Abraham, C., Hellman, L.M., Corcelli, S.A., Noé, F., Elliott, T., Baker, B.M., 2019. Dynamically driven allostery in MHC proteins: Peptide-dependent tuning of class I MHC global flexibility. Front. Immunol. 10, 966.

Badrinath, S., Saunders, P., Huyton, T., Aufderbeck, S., Hiller, O., Blasczyk, R., Bade-Doeding, C., 2012. Position 156 influences the peptide repertoire and tapasin dependency of human leukocyte antigen B*44 allotypes. Haematologica 97, 98–106.

Bakker, A.H., Hoppes, R., Linnemann, C., Toebes, M., Rodenko, B., Berkers, C.R., Hadrup, S.R., van Esch, W.J.E., Heemskerk, M.H.M., Ovaa, H., Schumacher, T.N.M., 2008. Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7, Proceedings of the National Academy of Sciences.

Barnes, P.D., Grundy, J.E., 1992. Down-regulation of the class I HLA heterodimer and beta2-microglobulin on the surface of cells infected with cytomegalovirus. J. Gen. Virol. 73, 2395–2403.

Barre-Sinoussi, F., Chermann, J., Rey, F., Nugeyre, M., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., Montagnier, L., 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science (80-.). 220, 868–871.

Beckman, E.M., Porcelli, S.A., Morita, C.T., Behar, S.M., Furlong, S.T., Brenner, M.B., 1994. Recognition of a lipid antigen by CD1-restricted αβ+ T cells. Nature 372, 691–694.

Betts, R.F., Hanshaw, J.B., 1977. Cytomegalovirus (CMV) in the Compromised Host(s). Annu. Rev. Med. 28, 103–110.

Bjorkman, P., 1990. Structure, Function, And Diversity Of Class I Major Histocompatibility Complex Molecules, Annual Review of Biochemistry.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., Wiley, D.C., 1987a. Structure of the human class I histocompatibility antigen, HLA-A2. Nature 329, 506–512.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., Wiley, D.C., 1987b. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. Nature 329, 512–518.

Bonsignori, M., Pollara, J., Moody, M.A., Alpert, M.D., Chen, X., Hwang, K.-K., Gilbert, P.B., Huang, Y., Gurley, T.C., Kozink, D.M., Marshall, D.J., Whitesides, J.F., Tsao, C.-Y., Kaewkungwal, J., Nitayaphan, S., Pitisuttithum, P., Rerks-Ngarm, S., Kim, J.H., Michael, N.L., Tomaras, G.D., Montefiori, D.C., Lewis, G.K., DeVico, A., Evans, D.T., Ferrari, G., Liao, H.-X., Haynes, B.F., 2012. Antibody-dependent cellular cytotoxicity-mediating antibodies from an HIV-1 vaccine efficacy trial target multiple epitopes and preferentially use the VH1 gene family. J. Virol. 86, 11521–32.

Borbulevych, O.Y., Baxter, T.K., Yu, Z., Restifo, N.P., Baker, B.M., 2005. Increased Immunogenicity of an Anchor-Modified Tumor-Associated Antigen Is Due to the Enhanced Stability of the Peptide/MHC Complex: Implications for Vaccine Design. J. Immunol. 174, 4812–4820.

Borbulevych, O.Y., Piepenbrink, K.H., Gloor, B.E., Scott, D.R., Sommese, R.F., Cole, D.K., Sewell, A.K., Baker, B.M., 2009. T Cell Receptor Cross-reactivity Directed by Antigen-Dependent Tuning of Peptide-MHC Molecular Flexibility. Immunity 31, 885–896.

Borrego, F., Ulbrecht, M., Weiss, E.H., Coligan, J.E., Brooks, A.G., 1998. Recognition of human

histocompatibility leukocyte antigen (HLA)-E complexed with HLA class I signal sequence-derived peptides by CD94/NKG2 confers protection from natural killer cell-mediated lysis. J. Exp. Med. 187, 813–8.

Brackenridge, S., Evans, E.J., Toebes, M., Goonetilleke, N., Liu, M.K.P., Di Gleria, K., Schumacher, T.N., Davis, S.J., Mcmichael, A.J., Gillespie, G.M., 2011. An Early HIV Mutation within an HLA-B*57-Restricted T Cell Epitope Abrogates Binding to the Killer Inhibitory Receptor 3DL1 †. J. Virol. 85, 5415–5422.

Bracq, L., Xie, M., Benichou, S., Bouchet, J., 2018. Mechanisms for cell-to-cell transmission of HIV-1. Front. Immunol. 9, 1.

Braud, V., Calan, C.A.O., So, K., Andrea, A.D., Ogg, G.S., Lazetic, S., Young, N.T., Bell, J.I., Phillips, J.H., Lanier, L.L., Mcmichael, A.J., Hospital, J.R., Ox, O., 1998. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C, Nature.

Braud, V., Jones, E.Y., McMichael, A., 1997. The human major histocompatibility complex class Ib molecule HLA-E binds signal sequence-derived peptides with primary anchor residues at positions 2 and 9, European Journal of Immunology.

Braud, V. M., Allan, D.S.J., O'Callaghan, C.A., Soderstrom, K., D'Andrea, A., Ogg, G.S., Lazetic, S., Young, N.T., Bell, J.I., Phillips, J.H., Lanler, L.L., McMichael, A.J., 1998. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. Nature 391, 795–799.

Braud, Veronique M., Allan, D.S.J., Wilson, D., McMichael, A.J., 1998. TAP- and tapasin-dependent HLA-E surface expression correlates with the binding of an MHC class I leader peptide, Current Biology.

Brooks, A.G., Borrego, F., Posch, P.E., Patamawenu, A., Scorzelli, C.J., Ulbrecht, M., Weiss, E.H., Coligan, J.E., 1999. Specific recognition of HLA-E, but not classical, HLA class I molecules by soluble CD94/NKG2A and NK cells. J. Immunol. 162, 305–30513.

Burian, A., Wang, K.L., Finton, K.A.K., Lee, N., Ishitani, A., Strong, R.K., Geraghty, D.E., 2016. HLA-F and MHC-I Open Conformers Bind Natural Killer Cell Ig-Like Receptor KIR3DS1. PLoS One 11, e0163297.

Burton, D.R., Hangartner, L., 2016. Broadly Neutralizing Antibodies to HIV and Their Role in Vaccine Design. Annu. Rev. Immunol. 34, 635–659.

Bushman, F.D., Fujiwara, T., Craigie, R., 1990. Retroviral DNA integration directed by HIV integration protein in vitro. Science (80-.). 249, 1555–1558.

Camilli, G., Cassotta, A., Battella, S., Palmieri, G., Santoni, A., Paladini, F., Fiorillo, M.T., Sorrentino, R., 2016. Regulation and trafficking of the HLA-E molecules during monocyte-macrophage differentiation. J. Leukoc. Biol. 99, 121–130.

Carrington, M., O'Brien, S.J., 2003. The Influence of HLA Genotype on AIDS . Annu. Rev. Med. 54, 535–551.

Chattaragada, M.S., Riganti, C., Sassoe, M., Principe, M., Santamorena, M.M., Roux, C., Curcio, C., Evangelista, A., Allavena, P., Salvia, R., Rusev, B., Scarpa, A., Cappello, P., Novelli, F., 2018. FAM49B, a novel regulator of mitochondrial function and integrity that suppresses tumor metastasis. Oncogene 37, 697–709.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. *MolProbity* : all-atom structure validation for macromolecular crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 12–21.

Chen, W., 1993. Changes at peptide residues buried in the major histocompatibility complex (MHC) class I binding cleft influence T cell recognition: a possible role for indirect conformational alterations in the MHC class I or bound peptide in determining T cell recognit, Journal of Experimental Medicine.

Chen, W., Zhang, L., Liang, B., Saenger, Y., Li, J., Chess, L., Jiang, H., 2007. Perceiving the avidity

of T cell activation can be translated into peripheral T cell regulation. Proc. Natl. Acad. Sci. U. S. A. 104, 20472–20477.

Cohen, G.B., Gandhi, R.T., Davis, D.M., Mandelboim, O., Chen, B.K., Strominger, J.L., Baltimore, D., 1999. The selective downregulation of class I major histocompatibility complex proteins by HIV-1 protects HIV-infected cells from NK cells. Immunity 10, 661–671.

Collins, K.L., Chen, B.K., Kalams, S.A., Walker, B.D., Baltimore, D., 1998. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. Nature 391, 397–401.

Connick, E., Mattila, T., Folkvord, J.M., Schlichtemeier, R., Meditz, A.L., Ray, M.G., McCarter, M.D., MaWhinney, S., Hage, A., White, C., Skinner, P.J., 2007. CTL Fail to Accumulate at Sites of HIV-1 Replication in Lymphoid Tissue. J. Immunol. 178, 6975–6983.

Corey, L., Gilbert, P.B., Tomaras, G.D., Haynes, B.F., Pantaleo, G., Fauci, A.S., 2015. Immune correlates of vaccine protection against HIV-1 acquisition. Sci. Transl. Med. 7, 310rv7.

Corey, L., Gray, G.E., 2017. Preventing acquisition of HIV is the only path to an AIDS-free generation. Proc. Natl. Acad. Sci. U. S. A. 114, 3798–3800.

Correale, J., Villa, A., 2008. Isolation and characterization of CD8+ regulatory T cells in multiple sclerosis. J. Neuroimmunol. 195, 121–134.

Cotterill, L.A., Stauss, H.J., Millrain, M.M., Pappin, D.J.C., Rahman, D., Canas, B., Chandler, P., Stackpoole, A., Simpson, E., Robinson, P.J., Dyson, P.J., 1997. Qa-1 interaction and T cell recognition of the Qa-1 determinant modifier peptide. Eur. J. Immunol. 27, 2123–2132.

Dam, J., Guan, R., Natarajan, K., Dimasi, N., Chlewicki, L.K., Kranz, D.M., Schuck, P., Margulies, D.H., Mariuzza, R.A., 2003. Variable MHC class I engagement by Ly49 natural killer cell receptors demonstrated by the crystal structure of Ly49C bound to H-2K b. Nat. Immunol. 4.

Daniel, M.D., Kirchhoff, F., Czajak, S.C., Sehgal, P.K., Desrosiers, R.C., 1992. Protective effects of a live attenuated SIV vaccine with a deletion in the nef gene. Science (80-.). 258, 1938–1941.

Davies, A., Lopez-Briones, S., Ong, H., O'Neil-Marshall, C., Lemonnier, F.A., Nagaraju, K., Metcalf, E.S., Soloski, M.J., 2004. Infection-Induced Expansion of a MHC Class Ib-Dependent Intestinal Intraepithelial γδ T Cell Subset. J. Immunol. 172, 6828–6837.

de Kruijf, E.M., Sajet, A., van Nes, J.G.H., Natanov, R., Putter, H., Smit, V.T.H.B.M., Liefers, G.J., van den Elsen, P.J., van de Velde, C.J.H., Kuppen, P.J.K., 2010. HLA-E and HLA-G Expression in Classical HLA Class I-Negative Tumors Is of Prognostic Value for Clinical Outcome of Early Breast Cancer Patients. J. Immunol. 185, 7452–7459.

DeCloux, A., Woods, A.S., Cotter, R.J., Soloski, M.J., Forman, J., 1997. Dominance of a single peptide bound to the class I(B) molecule, Qa-1b. J. Immunol. 158, 2183–91.

Deller, M.C., Kong, L., Rupp, B., 2016. Protein stability: A crystallographer's perspective. Acta Crystallogr. Sect. Struct. Biol. Commun. 72, 72–95.

Ebert, L.M., Yu, C.L., Clements, C.S., Robson, N.C., Jackson, H.M., Markby, J.L., Dimopoulos, N., Bee, S.T., Luescher, I.F., Davis, I.D., Rossjohn, J., Cebon, J., Purcell, A.W., Chen, W., 2009. A long, naturally presented immunodominant epitope from NY-ESO-1 tumor antigen: Implications for cancer vaccine design. Cancer Res. 69, 1046–1054.

Elliott, T., Cerundolo, V., Elvin, J., Townsend, A., 1991. Peptide-induced conformational change of the class I heavy chain. Nature 351, 402–406.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of *Coot*. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 486–501.

Evans, P., 2006. Scaling and assessment of data quality. In: Acta Crystallographica Section D: Biological Crystallography. pp. 72–82.

Evans, P.R., n.d. Biological Crystallography An introduction to data reduction: space-group determination, scaling and intensity statistics.

Evans, P.R., Murshudov, G.N., 2013. How good are my data and what is the resolution? Acta Crystallogr. Sect. D Biol. Crystallogr. 69, 1204–1214.

Fahnestock, M.L., Tamir, I., Narhi, L., Bjorkman, P.J., 1992. Thermal stability comparison of purified empty and peptide-filled forms of a class I MHC molecule. Science (80-.). 258, 1658–1662.

Fauci, A.S., Marston, H.D., 2014. Ending AIDS — Is an HIV Vaccine Necessary? N. Engl. J. Med. 370, 495–498.

Fenouillet, E., Gluckman, J.C., Jones, I.M., 1994. Functions of HIV envelope glycans, Trends in Biochemical Sciences.

Ferreira, L.M.R., Meissner, T.B., Tilburgs, T., Strominger, J.L., 2017. HLA-G: At the Interface of Maternal–Fetal Tolerance. Trends Immunol. 38, 272–286.

Franke, D., Petoukhov, M. V., Konarev, P. V., Panjkovich, A., Tuukkanen, A., Mertens, H.D.T., Kikhney, A.G., Hajizadeh, N.R., Franklin, J.M., Jeffries, C.M., Svergun, D.I., 2017. ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. J. Appl. Crystallogr. 50, 1212–1225.

Franksson, L., Sundbäck, J., Achour, A., Bernlind, J., Glas, R., Kärre, K., 1999. Peptide dependency and selectivity of the NK cell inhibitory receptor Ly-49C. Eur. J. Immunol. 29, 2748–2758.

Fremont, D.H., Matsumura, M., Stura, E.A., Peterson, P.A., Wilson, I.A., 1992. Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb. Science (80-.). 257, 919–927.

Fremont, D.H., Stura, E.A., Matsumura, M., Peterson, P.A., Wilson, I.A., 2006. Crystal structure of an H-2Kb-ovalbumin peptide complex reveals the interplay of primary and secondary anchor positions in the major histocompatibility complex binding groove., Proceedings of the National Academy of Sciences.

Friedrich, T.C., Valentine, L.E., Yant, L.J., Rakasz, E.G., Piaskowski, S.M., Furlott, J.R., Weisgrau, K.L., Burwitz, B., May, G.E., León, E.J., Soma, T., Napoe, G., Capuano Iii, S. V, Wilson, N.A., Watkins, D.I., 2007. Subdominant CD8 T-Cell Responses Are Involved in Durable Control of AIDS Virus Replication. J. Virol. 81, 3465–3476.

Früh, K., Ahn, K., Djaballah, H., Sempé, P., van Endert, P.M., Tampé, R., Peterson, P.A., Yang, Y., 1995. A viral inhibitor of peptide transporters for antigen presentation. Nature 375, 415–418.

Fukazawa, Y., Lum, R., Okoye, A.A., Park, H., Matsuda, K., Bae, J.Y., Hagen, S.I., Shoemaker, R., Deleage, C., Lucero, C., Morcock, D., Swanson, T., Legasse, A.W., Axthelm, M.K., Hesselgesser, J., Geleziunas, R., Hirsch, V.M., Edlefsen, P.T., Piatak, M., Estes, J.D., Lifson, J.D., Picker, L.J., 2015. B cell follicle sanctuary permits persistent productive simian immunodeficiency virus infection in elite controllers. Nat. Med. 21, 132–139.

Galli-Stampino, L., Meinjohanns, E., Frische, K., Meldal, M., Jensen, T., Werdelin, O., Mouritsen, S., 1997. T-cell recognition of tumor-associated carbohydrates: the nature of the glycan moiety plays a decisive role in determining glycopeptide immunogenicity. Cancer Res. 57, 3214–22.

Gao, G.F., Tormo, J., Gerth, U.C., Wyer, J.R., McMichael, A.J., Stuart, D.I., Bell, J.I., Jones, E.Y., Jakobsen, B.K., 1997. Crystal structure of the complex between CD8αα human and HLA-A2. Nature 387, 630–634.

Gao, G.F., Willcox, B.E., Wyer, J.R., Boulter, J.M., O'Callaghan, C.A., Maenaka, K., Stuart, D.I.,

Jones, E.Y., Van Der Merwe, P.A., Bell, J.I., Jakobsen, B.K., 2000. Classical and nonclassical class I major histocompatibility complex molecules exhibit subtle conformational differences that affect binding to CD8αα, Journal of Biological Chemistry.

Garboczi, D.N., Hung, D.T., Wiley, D.C., 1992. HLA-A2-peptide complexes: Refolding and crystallization of molecules expressed in Eschericlhia coli and complexed with single antigenic peptides, Immunology.

Garcia, K.C., Adams, E.J., 2005. Minireview How the T Cell Receptor Sees Antigen-A Structural View. Cell 122, 333–336.

Genescà, M., McChesney, M.B., Miller, C.J., 2009. Antiviral CD8+ T cells in the genital tract control viral replication and delay progression to AIDS after vaginal SIV challenge in rhesus macaques immunized with virulence attenuated SHIV 89.6. J. Intern. Med. 265, 67–77.

Geraghty, D.E., Stockschleader, M., Ishitani, A., Hansen, J.A., 1992. Polymorphism at the HLA-E locus predates most HLA-A and -B polymorphism. Hum. Immunol. 33, 174–184.

Gilbert, P.B., Berger, J.O., Stablein, D., Becker, S., Essex, M., Hammer, S.M., Kim, J.H., DeGruttola, V.G., 2011. Statistical interpretation of the RV144 HIV vaccine efficacy trial in Thailand: A case study for statistical issues in efficacy trials. J. Infect. Dis. 203, 969–975.

Gooden, M., Lampen, M., Jordanova, E.S., Leffers, N., Trimbos, J.B., van der Burg, S.H., Nijman, H., van Hall, T., 2011. HLA-E expression by gynecological cancers restrains tumor-infiltrating $CD8^+$ T lymphocytes. Proc. Natl. Acad. Sci. U. S. A. 108, 10656–61.

Grandea, A.G., Van Kaer, L., 2001. Tapasin: an ER chaperone that controls MHC class I assembly with peptide. Trends Immunol. 22, 194–199.

Gray, G., Buchbinder, S., Duerr, A., 2010. Overview of STEP and Phambili trial results: Two phase IIb test-of-concept studies investigating the efficacy of MRK adenovirus type 5 gag/pol/nef subtype B HIV vaccine. Curr. Opin. HIV AIDS 5, 357–361.

Green, K.J., Miles, J.J., Tellam, J., van Zuylen, W.J.M., Connolly, G., Burrows, S.R., 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.

Grimsley, C., Kawasaki, A., Gassner, C., Sageshima, N., Nose, Y., Hatake, K., Geraghty, D.E., Ishitani, A., 2002. Definitive high resolution typing of HLA-E allelic polymorphisms: Identifying potential errors in existing allele data. Tissue Antigens 60, 206–212.

Grimsley, C., Ober, C., 1997. Population genetic studies of HLA-E: Evidence for selection. Hum. Immunol. 52, 33–40.

Grotzke, J.E., Harriff, M.J., Siler, A.C., Nolt, D., Delepine, J., Lewinsohn, D.A., Lewinsohn, D.M., 2009. The Mycobacterium tuberculosis Phagosome Is a HLA-I Processing Competent Organelle. PLoS Pathog. 5, 1000374.

Haase, A.T., 2010. Early Events in Sexual Transmission of HIV and SIV and Opportunities for Interventions.

Hansen, S.G., Ford, J.C., Lewis, M.S., Ventura, A.B., Hughes, C.M., Coyne-Johnson, L., Whizin, N., Oswald, K., Shoemaker, R., Swanson, T., Legasse, A.W., Chiuchiolo, M.J., Parks, C.L., Axthelm, M.K., Nelson, J.A., Jarvis, M.A., Piatak, M., Lifson, J.D., Picker, L.J., 2011. Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine. Nature 473.

Hansen, Scott G, Piatak, M., Ventura, A.B., Hughes, C.M., Gilbride, R.M., Ford, J.C., Oswald, K., Shoemaker, R., Li, Y., Lewis, M.S., Gilliam, A.N., Xu, G., Whizin, N., Burwitz, B.J., Planer, S.L., Turner, J.M., Legasse, A.W., Axthelm, M.K., Nelson, J.A., Früh, K., Sacha, J.B., Estes, J.D., Keele, B.F., Edlefsen, P.T., Lifson, J.D., Picker, L.J., 2013. Immune clearance of highly pathogenic SIV infection. Nature 502.

Hansen, Scott G., Sacha, J.B., Hughes, C.M., Ford, J.C., Burwitz, B.J., Scholz, I., Gilbride, R.M., Lewis, M.S., Gilliam, A.N., Ventura, A.B., Malouli, D., Xu, G., Richards, R., Whizin, N., Reed,

J.S., Hammond, K.B., Fischer, M., Turner, J.M., Legasse, A.W., Axthelm, M.K., Edlefsen, P.T., Nelson, J.A., Lifson, J.D., Früh, K., Picker, L.J., 2013. Cytomegalovirus vectors violate CD8+ T cell epitope recognition paradigms. Science (80-.). 340.

Hansen, S.G., Vieville, C., Whizin, N., Coyne-Johnson, L., Siess, D.C., Drummond, D.D., Legasse, A.W., Axthelm, M.K., Oswald, K., Trubey, C.M., Piatak, M., Lifson, J.D., Nelson, J.A., Jarvis, M.A., Picker, L.J., 2009. Effector memory T cell responses are associated with protection of rhesus monkeys from mucosal simian immunodeficiency virus challenge. Nat. Med. 15, 293–299.

Hansen, S.G., Wu, H.L., Burwitz, B.J., Hughes, C.M., Hammond, K.B., Ventura, A.B., Reed, J.S., Gilbride, R.M., Ainslie, E., Morrow, D.W., Ford, J.C., Selseth, A.N., Pathak, R., Malouli, D., Legasse, A.W., Axthelm, M.K., Nelson, J.A., Gillespie, G.M., Walters, L.C., Brackenridge, S., Sharpe, H.R., López, C.A., Früh, K., Korber, B.T., McMichael, A.J., Gnanakaran, S., Sacha, J.B., Picker, L.J., 2016. Broadly targeted CD8+ T cell responses restricted by major histocompatibility complex E. Science (80-.). 351, 714–720.

Harndahl, M., Rasmussen, M., Roder, G., Dalgaard Pedersen, I., Sørensen, M., Nielsen, M., Buus, S., 2012. Peptide-MHC class I stability is a better predictor than peptide affinity of CTL immunogenicity. Eur. J. Immunol. 42, 1405–1416.

Harriff, M.J., Wolfe, L.M., Swarbrick, G., Null, M., Cansler, M.E., Canfield, E.T., Vogt, T., Toren, K.G., Li, W., Jackson, M., Lewinsohn, D.A., Dobos, K.M., Lewinsohn, D.M., 2017. HLA-E Presents Glycopeptides from the Mycobacterium tuberculosis Protein MPT32 to Human CD8 + T cells. Sci. Rep. 7, 4622.

Haurum, J.S., Høier, I.B., Arsequell, G., Neisig, A., Valencia, G., Zeuthen, J., Neefjes, J., Elliott, T., 1999. Presentation of cytosolic glycosylated peptides by human class I major histocompatibility complex molecules in vivo. J. Exp. Med. 190, 145–50.

Hawse, W.F., Gloor, B.E., Ayres, C.M., Kho, K., Nuter, E., Baker, B.M., 2013. Peptide modulation of class I major histocompatibility complex protein molecular flexibility and the implications for immune recognition. J. Biol. Chem. 288, 24372–24381.

Haynes, B.F., Gilbert, P.B., McElrath, M.J., Zolla-Pazner, S., Tomaras, G.D., Alam, S.M., Evans, D.T., Montefiori, D.C., Karnasuta, C., Sutthent, R., Liao, H.-X., DeVico, A.L., Lewis, G.K., Williams, C., Pinter, A., Fong, Y., Janes, H., DeCamp, A., Huang, Y., Rao, M., Billings, E., Karasavvas, N., Robb, M.L., Ngauy, V., de Souza, M.S., Paris, R., Ferrari, G., Bailer, R.T., Soderberg, K.A., Andrews, C., Berman, P.W., Frahm, N., De Rosa, S.C., Alpert, M.D., Yates, N.L., Shen, X., Koup, R.A., Pitisuttithum, P., Kaewkungwal, J., Nitayaphan, S., Rerks-Ngarm, S., Michael, N.L., Kim, J.H., 2012. Immune-Correlates Analysis of an HIV-1 Vaccine Efficacy Trial. N. Engl. J. Med. 366, 1275–1286.

Heinzel, A.S., Grotzke, J.E., Lines, R.A., Lewinsohn, D.A., McNabb, A.L., Streblow, D.N., Brand, V.M., Grieser, H.J., Belisle, J.T., Lewinsohn, D.M., 2002. HLA-E-dependent presentation of Mtb-derived antigen to human CD8+ T cells. J. Exp. Med. 196, 1473–1481.

Hoare, H.L., Sullivan, L.C., Clements, C.S., Ely, L.K., Beddoe, T., Henderson, K.N., Lin, J., Reid, H.H., Brooks, A.G., Rossjohn, J., 2008. Subtle Changes in Peptide Conformation Profoundly Affect Recognition of the Non-Classical MHC Class I Molecule HLA-E by the CD94–NKG2 Natural Killer Cell Receptors. J. Mol. Biol. 377, 1297–1303.

Hoare, H.L., Sullivan, L.C., Pietra, G., Clements, C.S., Lee, E.J., Ely, L.K., Beddoe, T., Falco, M., Kjer-Nielsen, L., Reid, H.H., McCluskey, J., Moretta, L., Rossjohn, J., Brooks, A.G., 2006. Structural basis for a major histocompatibility complex class Ib-restricted T cell response. Nat. Immunol. 7, 256–264.

Hogquist, K.A., Grandea, A.G., Bevan, M.J., 1993. Peptide variants reveal how antibodies recognize major histocompatibility complex class I. Eur. J. Immunol. 23, 3028–3036.

Hughes, A.L., Nei, M., 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335, 167–170.

Iaccino, E., Schiavone, M., Fiume, G., Quinto, I., Scala, G., 2008. The aftermath of the Merck's HIV vaccine trial. Retrovirology 5.

Illing, P.T., Vivian, J.P., Dudek, N.L., Kostenko, L., Chen, Z., Bharadwaj, M., Miles, J.J., Kjer-Nielsen, L., Gras, S., Williamson, N.A., Burrows, S.R., Purcell, A.W., Rossjohn, J., McCluskey, J., 2012. Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. Nature 486, 554–558.

Jennes, W., Verheyden, S., Demanet, C., Adjé-Touré, C.A., Vuylsteke, B., Nkengasong, J.N., Kestens, L., 2006. Cutting Edge: Resistance to HIV-1 Infection among African Female Sex Workers Is Associated with Inhibitory KIR in the Absence of Their HLA Ligands. J. Immunol. 177, 6588–6592.

Jiang, H., Canfield, S.M., Gallagher, M.P., Jiang, H.H., Jiang, Y., Zheng, Z., Chess, L., 2010. HLA-E-restricted regulatory CD8+ T cells are involved in development and control of human autoimmune type 1 diabetes. J. Clin. Invest. 120, 3641–3650.

Johansen, T.E., McCullough, K., Catipovic, B., Su, X.M., Amzel, M., Schneck, J.P., 1997. Peptide binding to MHC class I is determined by individual pockets in the binding groove. Scand. J. Immunol. 46, 137–146.

Johnson, J.A., Barouch, D.H., Baden, L.R., 2013. Nonreplicating vectors in HIV vaccines. Curr. Opin. HIV AIDS 8, 412–420.

Johnson, R.P., Desrosiers, R.C., 1998. Protective immunity induced by live attenuated simian immunodeficiency virus. Curr. Opin. Immunol. 10, 436–43.

Joint United Nations Programme on HIV/AIDS, 2016. Global AIDS UP DATE 2016, Unaids.

Joosten, S.A., Sullivan, L.C., Ottenhoff, T.H.M., 2016. Characteristics of HLA-E Restricted T-Cell Responses and Their Role in Infectious Diseases. J. Immunol. Res. 2016, 1–11.

Joosten, S.A., Van Meijgaarden, K.E., Van Weeren, P.C., Kazi, F., Geluk, A., Savage, N.D.L., Drijfhout, J.W., Flower, D.R., Hanekom, W.A., Klein, M.R., Ottenhoff, T.H.M., 2010. Mycobacterium tuberculosis peptides presented by HLA-E molecules are targets for human CD8+ T-cells with cytotoxic as well as regulatory activity. PLoS Pathog. 6.

Kabsch, W., 2010. XDS. Acta Crystallogr. D. Biol. Crystallogr. 66, 125–32.

Kaiser, B.K., Barahmand-pour, F., Paulsene, W., Medley, S., Geraghty, D.E., Strong, R.K., 2005. Interactions between NKG2x Immunoreceptors and HLA-E Ligands Display Overlapping Affinities and Thermodynamics. J. Immunol. 174, 2878–2884.

Kaiser, B.K., Pizarro, J.C., Kerns, J., Strong, R.K., 2008. Structural basis for NKG2A/CD94 recognition of HLA-E. Proc. Natl. Acad. Sci. U. S. A. 105, 6696–6701.

Kambayashi, T., Kraft-Leavy, J.R., Dauner, J.G., Sullivan, B.A., Laur, O., Jensen, P.E., 2004. The nonclassical MHC class I molecule Qa-1 forms unstable peptide complexes. J. Immunol. 172, 1661–9.

Karplus, P.A., Diederichs, K., 2012. Linking crystallographic model and data quality. Science (80-.). 336, 1030–1033.

Kati, W.M., Johnson, K.A., Jerva, L.F., Anderson, K.S., 1992. Mechanism and fidelity of HIV reverse transcriptase, Journal of Biological Chemistry.

Kikhney, A.G., Svergun, D.I., 2015. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. FEBS Lett. 589, 2570–2577.

Kim, H.J., Verbinnen, B., Tang, X., Lu, L., Cantor, H., 2010. Inhibition of follicular T-helper cells by CD8 + regulatory T cells is essential for self tolerance. Nature 467, 328–332.

Kirchhoff, F., 2010. Immune evasion and counteraction of restriction factors by HIV-1 and other primate lentiviruses. Cell Host Microbe 8, 55–67.

Kjer-Nielsen, L., Clements, C.S., Brooks, A.G., Purcell, A.W., Fontes, M.R., McCluskey, J., Rossjohn, J., 2002. The structure of HLA-B8 complexed to an immunodominant viral determinant: peptide-induced conformational changes and a mode of MHC class I dimerization. J. Immunol. 169, 5153–60.

Kozin, M.B., Svergun, D.I., 2001. Automated matching of high- and low-resolution structural models. J. Appl. Crystallogr. 34, 33–41.

Kraft, J.R., Vance, R.E., Pohl, J., Martin, A.M., Raulet, D.H., Jensen, P.E., 2000. Analysis of Qa-1(b) peptide binding specificity and the capacity of CD94/NKG2A to discriminate between Qa-1-peptide complexes. J. Exp. Med. 192, 613–623.

Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 2256–2268.

Krissinel, E., Henrick, K., 2007. Inference of Macromolecular Assemblies from Crystalline State. J. Mol. Biol. 372, 774–797.

Kurepa, Z., Hasemann, C.A., Forman, J., 1998. Qa-1 b Binds Conserved Class I Leader Peptides Derived from Several Mammalian Species . J. Exp. Med. 188, 973–978.

Leavenworth, J.W., Tang, X., Kim, H.J., Wang, X., Cantor, H., 2013. Amelioration of arthritis through mobilization of peptide-specific CD8 + regulatory T cells. J. Clin. Invest. 123, 1382–1389.

Lee, N, Goodlett, D.R., Ishitani, A., Marquardt, H., Geraghty, D.E., 1998. HLA-E surface expression depends on binding of TAP-dependent peptides derived from certain HLA class I signal sequences. J. Immunol. 160, 4951–60.

Lee, Ni, Llano, M., Carretero, M., Akiko-Ishitani, Navarro, F., López-Botet, M., Geraghty, D.E., 1998. HLA-E is a major ligand for the natural killer inhibitory receptor CD94/NKG2A. Proc. Natl. Acad. Sci. U. S. A. 95, 5199–5204.

Lehner, P.J., Karttunen, J.T., Wilkinson, G.W., Cresswell, P., 1997. The human cytomegalovirus US6 glycoprotein inhibits transporter associated with antigen processing-dependent peptide translocation. Proc. Natl. Acad. Sci. U. S. A. 94, 6904–9.

Lenz, L.L., Dere, B., Bevan, M.J., 1996. Identification of an H2-M3-restricted Listeria epitope: implications for antigen presentation by M3. Immunity 5, 63–72.

Lewinsohn, D.M., Alderson, M.R., Briden, A.L., Riddell, S.R., Reed, S.G., Grabstein, K.H., 1998. Characterization of human CD8+ T cells reactive with Mycobacterium tuberculosis-infected antigen-presenting cells. J. Exp. Med. 187, 1633–1640.

Lilja, A.E., Shenk, T., 2008. Efficient replication of rhesus cytomegalovirus variants in multiple rhesus and human cell types, Proceedings of the National Academy of Sciences of the United States of America.

Lin, A., Yan, W.-H., 2019. The Emerging Roles of Human Leukocyte Antigen-F in Immune Modulation and Viral Infection. Front. Immunol. 10, 964.

Llano, M., Lee, N., Navarro, F., García, P., Albar, J.P., Geraghty, D.E., López-Botet, M., 1998. HLA-E-bound peptides influence recognition by inhibitory and triggering CD94/NKG2 receptors: preferential response to an HLA-G-derived nonamer. Eur. J. Immunol. 28, 2854–2863.

Lo, W.-F., Woods, A.S., DeCloux, A., Cotter, R.J., Metcalf, E.S., Soloski, M.J., 2000. Molecular mimicry mediated by MHC class Ib molecules after infection with Gram-negative pathogens. Nat. Med. 6, 215–218.

Lu, L., Kim, H.J., Werneck, M.B.F., Cantor, H., 2008. Regulation of CD8+ regulatory T cells: Interruption of the NKG2A-Qa-1 interaction allows robust suppressive activity and

resolution of autoimmune disease. Proc. Natl. Acad. Sci. U. S. A. 105, 19420–19425.

Macdonald, W.A., Purcell, A.W., Mifsud, N.A., Ely, L.K., Williams, D.S., Chang, L., Gorman, J.J., Clements, C.S., Kjer-Nielsen, L., Koelle, D.M., Burrows, S.R., Tait, B.D., Holdsworth, R., Brooks, A.G., Lovrecz, G.O., Lu, L., Rossjohn, J., McCluskey, J., 2003. A naturally selected dimorphism within the HLA-B44 supertype alters class I structure, peptide repertoire, and T cell recognition. J. Exp. Med. 198, 679–91.

Madden, D.R., 1995. THE THREE-DIMENSIONAL STRUCTURE OF PEPTIDE-MHC COMPLEXES, Annu. Rev. Immunol.

Madden, D.R., Garboczi, D.N., Wiley, D.C., 1993. The antigenic identity of peptide-MHC complexes: A comparison of the conformations of five viral peptides presented by HLA-A2. Cell 75, 693–708.

Marín, R., Ruiz-Cabello, F., Pedrinaci, S., Méndez, R., Jiménez, P., Geraghty, D.E., Garrido, F., 2003. Analysis of HLA-E expression in human tumors. Immunogenetics 54, 767–75.

Martin, J., Horwich, A.L., Hartl, F.U., 1992. Prevention of protein denaturation under heat stress by the chaperonin Hsp60. Science 258, 995–8.

Martini, F., Agrati, C., D'Offizi, G., Poccia, F., 2005. HLA-E Up-Regulation Induced by HIV Infection May Directly Contribute to CD94-Mediated Impairment of NK Cells. Int. J. Immunopathol. Pharmacol. 18, 269–276.

Mazzarino, P., Pietra, G., Vacca, P., Falco, M., Colau, D., Coulie, P., Moretta, L., Mingari, M.C., 2005. Identification of effector-memory CMV-specific T lymphocytes that kill CMV-infected target cells in an HLA-E-restricted fashion. Eur. J. Immunol. 35, 3240–3247.

McMurtrey, C., Harriff, M.J., Swarbrick, G.M., Duncan, A., Cansler, M., Null, M., Bardet, W., Jackson, K.W., Lewinsohn, D.A., Hildebrand, W., Lewinsohn, D.M., 2017. T cell recognition of Mycobacterium tuberculosis peptides presented by HLA-E derived from infected human

cells. PLoS One 12.

Medlock, J., Pandey, A., Parpia, A.S., Tang, A., Skrip, L.A., Galvani, A.P., 2017. Effectiveness of UNAIDS targets and HIV vaccination across 127 countries. Proc. Natl. Acad. Sci. U. S. A. 114, 4017–4022.

Michaëlsson, J., Teixeira de Matos, C., Achour, A., Lanier, L.L., Kärre, K., Söderström, K., 2002. A Signal Peptide Derived from hsp60 Binds HLA-E and Interferes with CD94/NKG2A Recognition. J. Exp. Med. 196, 1403–1414.

Migueles, S.A., Sabbaghian, M.S., Shupert, W.L., Bettinotti, M.P., Marincola, F.M., Martino, L., Hallahan, C.W., Selig, S.M., Schwartz, D., Sullivan, J., Connors, M., 2000. HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. Proc. Natl. Acad. Sci. 97, 2709–2714.

Miles, B., Miller, S.M., Folkvord, J.M., Levy, D.N., Rakasz, E.G., Skinner, P.J., Connick, E., 2016. Follicular Regulatory CD8 T Cells Impair the Germinal Center Response in SIV and Ex Vivo HIV Infection. PLoS Pathog. 12.

Miller, J.D., Weber, D.A., Ibegbu, C., Pohl, J., Altman, J.D., Jensen, P.E., 2003. Analysis of HLA-E Peptide-Binding Specificity and Contact Residues in Bound Peptide Required for Recognition by CD94/NKG2. J. Immunol. 171, 1369–1375.

Monaco, E. Lo, Tremante, E., Cerboni, C., Melucci, E., Sibilio, L., Zingoni, A., Nicotra, M.R., Natali, P.G., Giacomini, P., 2011. Human Leukocyte Antigen E Contributes to Protect Tumor Cells from Lysis by Natural Killer Cells. Neoplasia 13, 822-IN14.

Morgan, C.S., Holton, J.M., Olafson, B.D., Bjorkman, P.J., Mayo, S.L., 1997. Circular dichroism determination of class I MHC-peptide equilibrium dissociation constants. Protein Sci. 6, 1771–3.

Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn,

M.D., Long, F., Vagin, A.A., 2011. *REFMAC* 5 for the refinement of macromolecular crystal structures. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 355–367.

Nagarajan, N.A., Gonzalez, F., Shastri, N., 2012. Nonclassical MHC class Ib-restricted cytotoxic T cells monitor antigen processing in the endoplasmic reticulum. Nat. Immunol. 13, 579–586.

Nattermann, J., Nischalke, H.D., Hofmeister, V., Ahlenstiel, G., Zimmermann, H., Leifeld, L., Weiss, E.H., Sauerbruch, T., Spengler, U., 2005a. The HLA-A2 restricted T cell epitope HCV core35-44 stabilizes HLA-E expression and inhibits cytolysis mediated by natural killer cells, American Journal of Pathology.

Nattermann, J., Nischalke, H.D., Hofmeister, V., Kupfer, B., Ahlenstiel, G., Feldmann, G., Rockstroh, J., Weiss, E.H., Sauerbruch, T., Spengler, U., 2005b. HIV-1 infection leads to increased HLA-E expression resulting in impaired function of natural killer cells. Antivir. Ther. 10, 95–107.

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

Nielsen, M., Lundegaard, C., Worning, P., Lauemøller, S.L., Lamberth, K., Buus, S., Brunak, S., Lund, O., 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci. 12, 1007–1017.

Niu, L., Cheng, H., Zhang, S., Tan, S., Zhang, Y., Qi, J., Liu, J., Gao, G.F., 2013. Structural basis for the differential classification of HLA-A*6802 and HLA-A*6801 into the A2 and A3 supertypes. Mol. Immunol. 55, 381–392.

Noble, A., Zhao, Z.S., Cantor, H., 1998. Suppression of immune responses by CD8 cells. II. Qa-1 on activated B cells stimulates CD8 cell suppression of T helper 2 responses. J. Immunol. 160, 566–71.

O'Callaghan, C.A., Tormo, J., Willcox, B.E., Braud, V.M., Jakobsen, B.K., Stuart, D.I., McMichael, A.J., Bell, J.I., Jones, E.Y., 1998. Structural features impose tight peptide binding specificity in the nonclassical MHC molecule HLA-E. Mol. Cell 1, 531–541.

Odyniec, A.N., Barral, D.C., Garg, S., Tatituri, R. V., Besra, G.S., Brenner, M.B., 2010. Regulation of CD1 antigen-presenting complex stability. J. Biol. Chem. 285, 11937–11947.

Okulicz, J.F., Marconi, V.C., Landrum, M.L., Wegner, S., Weintrob, A., Ganesan, A., Hale, B., Crum-Cianflone, N., Delmar, J., Barthel, V., Quinnan, G., Agan, B.K., Dolan, M.J., Sam Houston, F., Jackson, H.M., 2009. Clinical Outcomes of Elite Controllers, Viremic Controllers, and Long-Term Nonprogressors in the US Department of Defense HIV Natural History Study 200.

Oliveira, C.C., van Veelen, P.A., Querido, B., de Ru, A., Sluijter, M., Laban, S., Drijfhout, J.W., van der Burg, S.H., Offringa, R., van Hall, T., 2010. The nonpolymorphic MHC Qa-1b mediates CD8+ T cell surveillance of antigen-processing defects. J. Exp. Med. 207, 207–21.

Pang, Y.-P., Elsbernd, L.R., Block, M.S., Markovic, S.N., 2018. Peptide-Binding Groove Contraction Linked to the Lack of T Cell Response: Using Complex Structure and Energy To Identify Neoantigens. ImmunoHorizons 2, 216–225.

Panoutsakopoulou, V., Huster, K.M., McCarty, N., Feinberg, E., Wang, R., Wucherpfennig, K.W., Cantor, H., 2004. Suppression of autoimmune disease after vaccination with autoreactive T cells that express Qa-1 peptide complexes. J. Clin. Invest. 113, 1218–24.

Parham, P., Lomen, C.E., Lawlor, D.A., Ways, J.P., Holmes, N., Coppin, H.L., Salter, R.D., Wan, A.M., Ennis, P.D., 1988. Nature of polymorphism in HLA-A, -B, and -C molecules. Proc. Natl. Acad. Sci. U. S. A. 85, 4005–9.

Parker, K.C., Bednarek, M.A., Hull, L.K., Utz, U., Cunningham, B., Zweerink, H.J., Biddison, W.E., Coligan, J.E., 1992. Sequence motifs important for peptide binding to the human MHC

class I molecule, HLA-A2. J. Immunol. 149, 3580–7.

Pedersen, L.O., Nissen, M.H., Hansen, N.J.V., Nielsen, L.L., Lauenmøller, S.L., Blicher, T., Nansen, A., Sylvester-Hvid, C., Thromsen, A.R., Buus, S., 2001. Efficient assembly of recombinant major histocompatibility cmplex class I molecules with performed disulfide bonds. Eur. J. Immunol. 31, 2986–2996.

Petersen, J.L., Morris, C.R., Solheim, J.C., Brodsky, F.M., 2003. Virus evasion of MHC class I molecule presentation. J. Immunol. 171, 4473–8.

Petrie, E.J., Clements, C.S., Lin, J., Sullivan, L.C., Johnson, D., Huyton, T., Heroux, A., Hoare, H.L., Beddoe, T., Reid, H.H., Wilce, M.C.J., Brooks, A.G., Rossjohn, J., 2008. CD94-NKG2A recognition of human leukocyte antigen (HLA)-E bound to an HLA class I leader sequence. J. Exp. Med. 205, 725–735.

Pietra, G., Romagnani, C., Manzini, C., Moretta, L., Mingari, M.C., 2010. The Emerging Role of HLA-E-Restricted CD8 + T Lymphocytes in the Adaptive Immune Response to Pathogens and Tumors. J. Biomed. Biotechnol. 2010.

Pitisuttithum, P., Gilbert, P., Gurwith, M., Heyward, W., Martin, M., van Griensven, F., Hu, D., Tappero, J.W., Choopanya, K., 2006. Randomized, Double-Blind, Placebo-Controlled Efficacy Trial of a Bivalent Recombinant Glycoprotein 120 HIV-1 Vaccine among Injection Drug Users in Bangkok, Thailand. J. Infect. Dis. 194, 1661–1671.

Pontesilli, O., Klein, M.R., Kerkhof-Garde, S.R., Pakker, N.G., de Wolf, F., Schuitemaker, H., Miedema, F., 1998. Longitudinal Analysis of Human Immunodeficiency Virus Type 1-Specific Cytotoxic T Lymphocyte Responses: A Predominant Gag-Specific Response Is Associated with Nonprogressive Infection, The Journal of Infectious Diseases.

Poropatich, K., Sullivan, D.J., 2011. Human immunodeficiency virus type 1 long-term non-progressors: The viral, genetic and immunological basis for disease non-progression. J.

Gen. Virol. 92, 247–268.

Prezzemolo, T., van Meijgaarden, K.E., Franken, K.L.M.C., Caccamo, N., Dieli, F., Ottenhoff, T.H.M., Joosten, S.A., 2018. Detailed characterization of human Mycobacterium tuberculosis specific HLA-E restricted CD8+ T cells. Eur. J. Immunol. 48, 293–305.

Rajagopalan, S., Long, E.O., 1999. A human histocompatibility leukocyte antigen (HLA)-G-specific receptor expressed on all natural killer cells. J. Exp. Med. 189, 1093–100.

Rambo, R.P., Tainer, J.A., 2011. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. Biopolymers 95, 559–571.

Ravindranath, M.H., Pham, T., El-Awar, N., Kaneku, H., Terasaki, P.I., 2011. Anti-HLA-E mAb 3D12 mimics MEM-E/02 in binding to HLA-B and HLA-C alleles: Web-tools validate the immunogenic epitopes of HLA-E recognized by the antibodies. Mol. Immunol. 48, 423–430.

Read, R.J., Adams, P.D., McCoy, A.J., 2013. Intensity statistics in the presence of translational noncrystallographic symmetry. Acta Crystallogr. Sect. D Biol. Crystallogr. 69, 176–183.

Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., Kaewkungwal, J., Chiu, J., Paris, R., Premsri, N., Namwat, C., De Souza, M., Adams, E., Benenson, M., Gurunathan, S., Tartaglia, J., McNeil, J.G., Francis, D.P., Stablein, D., Birx, D.L., Chunsuttiwat, S., Khamboonruang, C., Thongcharoen, P., Robb, M.L., Michael, N.L., Kunasol, P., Kim, J.H., 2009. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N. Engl. J. Med. 361, 2209–2220.

Rhee, S.G., Chae, H.Z., Kim, K., 2005. Peroxiredoxins: A historical overview and speculative preview of novel mechanisms and emerging concepts in cell signaling. Free Radic. Biol. Med. 38, 1543–1552.

Robert-Guroff, M., 2007. Replicating and non-replicating viral vectors for vaccine development,

Current Opinion in Biotechnology.

Rodenko, B., Toebes, M., Hadrup, S.R., van Esch, W.J.E., Molenaar, A.M., Schumacher, T.N.M., Ovaa, H., 2006. Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. Nat. Protoc. 1, 1120–1132.

Rodgers, J.R., Cook, R.G., 2005. MHC class IB molecules bridge innate and acquired immunity. Nat. Rev. Immunol. 5, 459–471.

Rolph, M.S., Raupach, B., Köbernick, H.H.C., Collins, H.L., Pérarnau, B., Lemonnier, F.A., Kaufmann, S.H.E., 2001. MHC class Ia-restricted T cells partially account for β2-microglobulin-dependent resistance toMycobacterium tuberculosis. Eur. J. Immunol. 31, 1944–1949.

Rückert, C., Fiorillo, M.T., Loll, B., Moretti, R., Biesiadka, J., Saenger, W., Ziegler, A., Sorrentino, R., Uchanska-Ziegler, B., 2006. Conformational dimorphism of self-peptides and molecular mimicry in a disease-associated HLA-B27 subtype. J. Biol. Chem. 281, 2306–16.

Saini, S.K., Ostermeir, K., Ramnarayan, V.R., Schuster, H., Zacharias, M., Springer, S., 2013. Dipeptides promote folding and peptide binding of MHC class I molecules. Proc. Natl. Acad. Sci. 110, 15383–15388.

Salerno-Gonçalves, R., Fernandez-Viña, M., Lewinsohn, D.M., Sztein, M.B., 2004. Identification of a Human HLA-E-Restricted CD8 + T Cell Subset in Volunteers Immunized with Salmonella enterica Serovar Typhi Strain Ty21a Typhoid Vaccine . J. Immunol. 173, 5852–5862.

Sallusto, F., Geginat, J., Lanzavecchia, A., 2004. C entral M emory and E ffector M emory T C ell S ubsets : Function, Generation, and Maintenance . Annu. Rev. Immunol. 22, 745–763.

Salter, R.D., Benjamin, R.J., Wesley, P.K., Buxton, S.E., Garrett, T.P.J., Clayberger, C., Krensky, A.M., Norment, A.M., Littman, D.R., Parham, P., 1990. A binding site for the T-cell co-

receptor CD8 on the α3 domain of HLA-A2. Nature 345, 41–46.

Sasaki, T., Kanaseki, T., Shionoya, Y., Tokita, S., Miyamoto, S., Saka, E., Kochin, V., Takasawa, A., Hirohashi, Y., Tamura, Y., Miyazaki, A., Torigoe, T., Hiratsuka, H., Sato, N., 2016. Microenvironmental stresses induce HLA-E/Qa-1 surface expression and thereby reduce CD8 + T-cell recognition of stressed cells. Eur. J. Immunol. 46, 929–940.

Schacker, T., Collier, A.C., Hughes, J., Shea, T., Corey, L., 1996. Clinical and Epidemiologic Features of Primary HIV Infection. Ann. Intern. Med. 125, 257.

Schmitz, J.E., Kuroda, M.J., Santra, S., Sasseville, V.G., Simon, M.A., Lifton, M.A., Racz, P., Tenner-Racz, K., Dalesandro, M., Scallon, B.J., Ghrayeb, J., Forman, M.A., Montefiori, D.C., Rieber, E.P., Letvin, N.L., Reimann, K.A., 1999. Control of Viremia in Simian Immunodeficiency Virus Infection by CD8+ Lymphocytes. Science (80-.). 283, 857–860.

Schneidman-Duhovny, D., Hammel, M., Sali, A., 2010. FoXS: a web server for rapid computation and fitting of SAXS profiles. Nucleic Acids Res. 38.

Schneidman-Duhovny, D., Hammel, M., Tainer, J.A., Sali, A., 2013. Accurate SAXS profile computation and its assessment by contrast variation experiments. Biophys. J. 105, 962–974.

Schneidman-Duhovny, D., Hammel, M., Tainer, J.A., Sali, A., 2016. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. Nucleic Acids Res. 44, W424–W429.

Schulte, D., Vogel, M., Langhans, B., Krämer, B., Körner, C., Nischalke, H.D., Steinberg, V., Michalk, M., Berg, T., Rockstroh, J.K., Sauerbruch, T., Spengler, U., Nattermann, J., 2009. The HLA-E R /HLA-E R Genotype Affects the Natural Course of Hepatitis C Virus (HCV) Infection and Is Associated with HLA-E–Restricted Recognition of an HCV-Derived Peptide by Interferon-γ–Secreting Human CD8 + T Cells . J. Infect. Dis. 200, 1397–1401.

Schwartz, O., Riviere, Y., Heard, J.-M., Danos, O., 1993. Reduced Cell Surface Expression of Processed Human Immunodeficiency Virus Type 1 Envelope Glycoprotein in the Presence of Nef, JOURNAL OF VIROLOGY.

Sekaly, R.P., 2008. The failed HIV Merck vaccine study: A step back or a launching point for future vaccine development? J. Exp. Med. 205, 7–12.

Seliger, B., Jasinski-Bergner, S., Quandt, D., Stoehr, C., Bukur, J., Wach, S., Legal, W., Taubert, H., Wullich, B., Hartmann, A., 2016. HLA-E expression and its clinical relevance in human renal cell carcinoma. Oncotarget 7, 67360–67372.

Sensi, M., Pietra, G., Molla, A., Nicolini, G., Vegetti, C., Bersani, I., Millo, E., Weiss, E., Moretta, L., Mingari, M.C., Anichini, A., 2009. Peptides with dual binding specificity for HLA-A2 and HLA-E are encoded by alternatively spliced isoforms of the antioxidant enzyme peroxiredoxin 5. Int. Immunol. 21, 257–268.

Sette, A., Sidney, J., 1999. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics 50, 201–212.

Shibata, R., Igarashi, T., Haigwood, N., Buckler-White, A., Ogert, R., Ross, W., Willey, R., Cho, M.W., Martin, M.A., 1999. Neutralizing antibody directed against the HIV-1 envelope glycoprotein can completely block HIV-1/SIV chimetic virus infections of macaque monkeys. Nat. Med. 5, 204–210.

Shiver, J.W., Fu, T.M., Chen, L., Casimiro, D.R., Davies, M.E., Evans, R.K., Zhang, Z.Q., Simon, A.J., Trigona, W.L., Dubey, S.A., Huang, L., Harris, V.A., Long, R.S., Liang, X., Handt, L., Schleif, W.A., Zhu, L., Freed, D.C., Persaud, N. V., Guan, L., Punt, K.S., Tang, A., Chen, M., Wilson, K.A., Collins, K.B., Heidecker, G.J., Fernandez, V.R., Perry, H.C., Joyce, J.G., Grimm, K.M., Cook, J.C., Keller, P.M., Kresock, D.S., Mach, H., Troutman, R.D., Isopi, L.A., Williams, D.M., Xu, Z., Bohannon, K.E., Volkin, D.B., Montefiori, D.C., Miura, A., Krivulka, G.R., Lifton, M.A.,

Kuroda, M.J., Schmitz, J.E., Letvin, N.L., Caulfield, M.J., Bett, A.J., Youil, R., Kaslow, D.C., Emini, E.A., 2002. Replication-incompetent adenoviral vaccine vector elicits effective anti-immunodeficiency-virus immunity, Nature.

Sidney, J., Peters, B., Frahm, N., Brander, C., Sette, A., 2008. HLA class I supertypes: a revised and updated classification. BMC Immunol. 9, 1.

Sieling, P., Chatterjee, D., Porcelli, S., Prigozy, T., Mazzaccaro, R., Soriano, T., Bloom, B., Brenner, M., Kronenberg, M., Brennan, P., et, al., 1995. CD1-restricted T cell recognition of microbial lipoglycan antigens. Science (80-.). 269, 227–230.

Skou, S., Gillilan, R.E., Ando, N., 2014. Synchrotron-based small-angle X-ray scattering of proteins in solution. Nat. Protoc. 9, 1727–1739.

Smith, K.J., Reid, S.W., Harlos, K., McMichael, A.J., Stuart, D.I., Bell, J.I., Jones, E.Y., 1996a. Bound water structure and polymorphic amino acids act together to allow the binding of different peptides to MHC class I HLA-B53, Immunity.

Smith, K.J., Reid, S.W., Stuart, D.I., McMichael, A.J., Jones, E.Y., Bell, J.I., 1996b. An altered position of the α2 helix of MHC class I is revealed by the crystal structure of HLA-B3501. Immunity 4, 203–213.

Stanton, T.H., Carbon, S., 1982. Gene(s) affecting the expression of Qa-1. Immunogenetics 16, 435–44.

Strong, R.K., Holmes, M.A., Li, P., Braun, L., Lee, N., Geraghty, D.E., 2003. HLA-E allelic variants: Correlating differential expression, peptide affinities, crystal structures, and thermal stabilities. J. Biol. Chem. 278, 5082–5090.

Suh, W.-K., Cohen-Doyle, M.F., Williams, D.B., Derby, M.A., Berzofsky, J.A., Schoenhals, G.J., Früh, K., 1999. Interaction of murine MHC class I molecules with tapasin and TAP enhances peptide loading and involves the heavy chain α3 domain, Journal of

Immunology.

Sullivan, L.C., Clements, C.S., Beddoe, T., Johnson, D., Hoare, H.L., Lin, J., Huyton, T., Hopkins, E.J., Reid, H.H., Wilce, M.C.J., Kabat, J., Borrego, F., Coligan, J.E., Rossjohn, J., Brooks, A.G., 2007. The Heterodimeric Assembly of the CD94-NKG2 Receptor Family and Implications for Human Leukocyte Antigen-E Recognition. Immunity 27, 900–911.

Sullivan, L.C., Walpole, N.G., Farenc, C., Pietra, G., Sum, M.J.W., Clements, C.S., Lee, E.J., Beddoe, T., Falco, M., Mingari, M.C., Moretta, L., Gras, S., Rossjohn, J., Brooks, A.G., 2017. A conserved energetic footprint underpins recognition of human leukocyte antigen-E by two distinct αβ T cell receptors. J. Biol. Chem. 292, 21149–21158.

Svergun, D.I., 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys. J. 76, 2879–86.

Sylvester-Hvid, C., Kristensen, N., Blicher, T., Ferré, H., Lauemøller, S.L., Wolf, X.A., Lamberth, K., Nissen, M.H., Pedersen, L.Ø., Buus, S., 2002. Establishment of a quantitative ELISA capable of determining peptide - MHC class I interaction. Tissue Antigens 59, 251–8.

Tang, X., Maricic, I., Purohit, N., Bakamjian, B., Reed-Loisel, L.M., Beeston, T., Jensen, P., Kumar, V., 2006. Regulation of Immunity by a Novel Population of Qa-1-Restricted CD8αα + TCRαβ + T Cells . J. Immunol. 177, 7645–7655.

Telesnitsky, A., Goff, S., 1997. Reverse Transcriptase and the Generation of Retroviral DNA, Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).

Thans, T. van S., Akko, J.I., Niehrs, A., Garcia-Beltran, W.F., Richert, L., Stürzel, C.M., Ford, C.T., Li, H., Ochsenbauer, C., Kappes, J.C., Hahn, B.H., Kirchhoff, F., Martrus, G., Sauter, D., Altfeld, M., Hölzemer, A., 2019. Primary HIV-1 strains use Nef to downmodulate HLA-E surface expression. J. Virol.

Thomas, C., Tampé, R., 2017. Structure of the TAPBPR-MHC I complex defines the mechanism

of peptide loading and editing. Science 358, 1060–1064.

Thomsen, M.C.F., Nielsen, M., 2012. Seq2Logo: A method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. Nucleic Acids Res. 40.

Tindall, B., Barker, S., Donovan, B., Barnes, T., Roberts, J., Kronenberg, C., Gold, J., Penny, R., Cooper, D., Adams, A.I., Berry, G., Cooper, D.A., Whyte, B., Johnson, M., McLaws, M., Perdices, M., Best, S., Blunt, J., Borrow, P., Chapman, G., Hurren, L., Imrie, A., Katakouzinos, C., Kehrer, S., McCabe, H., Malcolm, A., North, M., Painter, L., Schreurs, E., Stevens, M., Wilson, E., Wilson, R., Anns, M., Grant, D., Morlet, A., Philpot, R., Jones, R., Burcham, J., Barnes, T.G., Brooke, P., Finlayson, R., Haynes, P., Lovett, R., Michelmore, H.M., Mutimer, K., Price, R., Wolfe, C., 1988. Characterization of the Acute Clinical Illness Associated With Human Immunodeficiency Virus Infection. Arch. Intern. Med. 148, 945.

Tirosh, B., Iwakoshi, N.N., Lilley, B.N., Lee, A.-H., Glimcher, L.H., Ploegh, H.L., 2005. Human Cytomegalovirus Protein US11 Provokes an Unfolded Protein Response That May Facilitate the Degradation of Class I Major Histocompatibility Complex Products. J. Virol. 79, 2768–2779.

Toebes, M., Coccoris, M., Bins, A., Rodenko, B., Gomez, R., Nieuwkoop, N.J., Van De Kasteele, W., Rimmelzwaan, G.F., Haanen, J.B.A.G., Ovaa, H., Schumacher, T.N.M., 2006. Design and use of conditional MHC class I ligands. Nat. Med. 12, 246–251.

Tomasec, P., Braud, V.M., Rickards, C., Powell, M.B., McSharry, B.P., Gadola, S., Cerundolo, V., Borysiewicz, L.K., McMichae, A.J., Wilkinson, G.W.G., 2000. Surface expression of HLA-E, an inhibitor of natural killer cells, enhanced by human cytomegalovirus gpUL40, Science.

Townsend, A., Elliott, T., Cerundolo, V., Foster, L., Barber, B., Tse, A., 1990. Assembly of MHC

class I molecules analyzed in vitro. Cell 62, 285–295.

Townsend, A.R.M., Gotch, F.M., Davey, J., 1985. Cytotoxic T cells recognize fragments of the influenza nucleoprotein. Cell 42, 457–467.

Tremante, E., Lo Monaco, E., Ingegnere, T., Sampaoli, C., Fraioli, R., Giacomini, P., 2015. Monoclonal antibodies to HLA-E bind epitopes carried by unfolded β<inf>2</inf>m-free heavy chains. Eur. J. Immunol. 45, 2356–2364.

Trivedi, M., Laurence, J., Siahaan, T., 2009. The Role of Thiols and Disulfides on Protein Stability, Current Protein & Peptide Science.

Tynan, F.E., Borg, N.A., Miles, J.J., Beddoe, T., El-Hassen, D., Silins, S.L., Van Zuylen, W.J.M., Purcell, A.W., Kjer-Nielsen, L., McCluskey, J., Burrows, S.R., Rossjohn, J., 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.

Ulbrecht, M., Couturier, A., Martinozzi, S., Pla, M., Srivastava, R., Peterson, P.A., Weiss, E.H., 1999. Cell surface expression of HLA-E: interaction with human β2-microglobulin and allelic differences. Eur. J. Immunol. 29, 537–547.

Ulbrecht, M., Honka, T., Person, S., Johnson, J.P., Weiss, E.H., 1992. The HLA-E gene encodes two differentially regulated transcripts and a cell surface protein. J. Immunol. 149, 2945–53.

Ulbrecht, M., Martinozzi, S., Grzeschik, M., Hengel, H., Ellwart, J.W., Pla, M., Weiss, E.H., 2000. Cutting edge: the human cytomegalovirus UL40 gene product contains a ligand for HLA-E and prevents NK cell-mediated lysis. J. Immunol. 164, 5019–22.

Ulbrecht, M., Weiss, E.H., Modrow, S., Srivastava, R., Peterson, P.A., 1998. Interaction of HLA-E with peptides and the peptide transporter in vitro: Implications for its function in antigen

presentation, Journal of Immunology.

Vagin, A., Teplyakov, A., 2010. Molecular replacement with *MOLREP*. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 22–25.

Valés-Gómez, M., Reyburn, H.T., Erskine, R.A., López-Botet, M., Strominger, J.L., 1999. Kinetics and peptide dependency of the binding of the inhibitory NK receptor CD94/NIKG2-A and the activating receptor CD94/NKG2-C to HLA-E. EMBO J. 18, 4250–4260.

Van Bleek, G.M., Nathenson, S.G., 1991. The structure of the antigen-binding groove of major histocompatibility complex class I molecules determines specific selection of self-peptides, Proceedings of the National Academy of Sciences of the United States of America.

van Deutekom, H.W.M., Wijnker, G., de Boer, R.J., 2013. The Rate of Immune Escape Vanishes When Multiple Immune Responses Control an HIV Infection. J. Immunol. 191, 3277–3286.

Van Hall, T., Wolpert, E.Z., Van Veelen, P., Laban, S., Van Der Veer, M., Roseboom, M., Bres, S., Grufman, P., De Ru, A., Meiring, H., De Jong, A., Franken, K., Teixeira, A., Valentijn, R., Drijfhout, J.W., Koning, F., Camps, M., Ossendorp, F., Kärre, K., Ljunggren, H.G., Melief, C.J.M., Offringa, R., 2006. Selective cytotoxic T-lymphocyte targeting of tumor immune escape variants. Nat. Med. 12, 417–424.

van Hateren, A., Bailey, A., Elliott, T., 2017. Recent advances in Major Histocompatibility Complex (MHC) class I antigen presentation: Plastic MHC molecules and TAPBPR-mediated quality control. F1000Research 6, 158.

van Meijgaarden, K.E., Haks, M.C., Caccamo, N., Dieli, F., Ottenhoff, T.H.M., Joosten, S.A., 2015. Human CD8+ T-cells Recognizing Peptides from Mycobacterium tuberculosis (Mtb) Presented by HLA-E Have an Unorthodox Th2-like, Multifunctional, Mtb Inhibitory Phenotype and Represent a Novel Human T-cell Subset. PLOS Pathog. 11, e1004671.

Vinay, D.S., Ryan, E.P., Pawelec, G., Talib, W.H., Stagg, J., Elkord, E., Lichtor, T., Decker, W.K.,

Whelan, R.L., Kumara, H.M.C.S., Signori, E., Honoki, K., Georgakilas, A.G., Amin, A., Helferich, W.G., Boosani, C.S., Guha, G., Ciriolo, M.R., Chen, S., Mohammed, S.I., Azmi, A.S., Keith, W.N., Bilsland, A., Bhakta, D., Halicka, D., Fujii, H., Aquilano, K., Ashraf, S.S., Nowsheen, S., Yang, X., Choi, B.K., Kwon, B.S., 2015. Immune evasion in cancer: Mechanistic basis and therapeutic strategies. Semin. Cancer Biol. 35, S185–S198.

Vinuesa, C.G., Cyster, J.G., 2011. How T cells earn the follicular rite of passage. Immunity 35, 671–680.

Virgin, H.W., Wherry, E.J., Ahmed, R., 2009. Redefining Chronic Viral Infection. Cell 138, 30–50.

Walter, T.S., Diprose, J.M., Mayo, C.J., Siebold, C., Pickford, M.G., Carter, L., Sutton, G.C., Berrow, N.S., Brown, J., Berry, I.M., Stewart-Jones, G.B.E., Grimes, J.M., Stammers, D.K., Esnouf, R.M., Jones, E.Y., Owens, R.J., Stuart, D.I., Harlos, K., 2005. A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. Acta Crystallogr. Sect. D Biol. Crystallogr. 61, 651–657.

Walters, L.C., Harlos, K., Brackenridge, S., Rozbesky, D., Barrett, J.R., Jain, V., Walter, T.S., O'Callaghan, C.A., Borrow, P., Toebes, M., Hansen, S.G., Sacha, J., Abdulhaqq, S., Greene, J.M., Früh, K., Marshall, E., Picker, L.J., Jones, E.Y., McMichael, A.J., Gillespie, G.M., 2018. Pathogen-derived HLA-E bound epitopes reveal broad primary anchor pocket tolerability and conformationally malleable peptide binding, Nature Communications.

Wang, C.-R., Castan˜o, A.R., Peterson, P.A., Slaughter, C., Fischer Lindahl, K., Deisenhofer, J., 1995. Nonclassical binding of formylated peptide in crystal structure of the MHC class Ib molecule H2-M3. Cell 82, 655–664.

Wang, D., Shenk, T., 2005. Human cytomegalovirus virion protein complex required for epithelial and endothelial cell tropism, Proceedings of the National Academy of Sciences

of the United States of America.

Wang, E.C.Y., McSharry, B., Retiere, C., Tomasec, P., Williams, S., Borysiewicz, L.K., Braud, V.M., Wilkinson, G.W.G., 2002. UL40-mediated NK evasion during productive infection with human cytomegalovirus. Proc. Natl. Acad. Sci. U. S. A. 99, 7570–5.

Wearsch, P.A., Cresswell, P., 2008. The quality control of MHC class I peptide loading. Curr. Opin. Cell Biol. 20, 624–631.

Wearscha, P.A., Peapera, D.R., Cresswell, P., 2011. Essential glycan-dependent interactions optimize MHC class I peptide loading. Proc. Natl. Acad. Sci. U. S. A. 108, 4950–4955.

Weintrob, A.C., Giner, J., Menezes, P., Patrick, E., Benjamin, D.K., Lennox, J., Pilcher, C.D., Eron, J.J., Hicks, C.B., 2003. Infrequent Diagnosis of Primary Human Immunodeficiency Virus Infection. Arch. Intern. Med. 163, 2097.

Wieczorek, M., Abualrous, E.T., Sticht, J., Álvaro-Benito, M., Stolzenberg, S., Noé, F., Freund, C., 2017. Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation. Front. Immunol. 8, 292.

Williams, A.P., Peh, C.A., Purcell, A.W., McCluskey, J., Elliott, T., 2002. Optimization of the MHC Class I Peptide Cargo Is Dependent on Tapasin. Immunity 16, 509–520.

Williams, M., Roeth, J.F., Kasper, M.R., Fleis, R.I., Przybycin, C.G., Collins, K.L., 2002. Direct Binding of Human Immunodeficiency Virus Type 1 Nef to the Major Histocompatibility Complex Class I (MHC-I) Cytoplasmic Tail Disrupts MHC-I Trafficking. J. Virol. 76, 12173–12184.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the *CCP* 4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 235–242.

World Health Organization (WHO/ CDS/HIV), 2019. Progress report on HIV, viral hepatitis and sexually transmitted infections 2019: Accountability for the global health sector strategies, 2016–2021.

Wu, H.L., Wiseman, R.W., Hughes, C.M., Webb, G.M., Abdulhaqq, S.A., Bimber, B.N., Hammond, K.B., Reed, J.S., Gao, L., Burwitz, B.J., Greene, J.M., Ferrer, F., Legasse, A.W., Axthelm, M.K., Park, B.S., Brackenridge, S., Maness, N.J., McMichael, A.J., Picker, L.J., O'Connor, D.H., Hansen, S.G., Sacha, J.B., 2018. The Role of MHC-E in T Cell Immunity Is Conserved among Humans, Rhesus Macaques, and Cynomolgus Macaques. J. Immunol. 200, 49–60.

Xiao, Z., Ye, Z., Tadwal, V.S., Shen, M., Ren, E.C., 2017. Dual non-contiguous peptide occupancy of HLA class i evoke antiviral human CD8 T cell response and form neo-epitopes with self-antigens /631/250/21/324/1509 /631/326/596/1553 /13/31 /13/106 /13/109 /145 /82/83 article. Sci. Rep. 7.

Xu, J., Baase, W.A., Baldwin, E., Matthews, B.W., 1998. The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect. Protein Sci. 7, 158–177.

Yagi, R., Zhu, J., Paul, W.E., 2011. An updated view on transcription factor GATA3-mediated regulation of Th1 and Th2 cell differentiation. Int. Immunol. 23, 415–20.

Ying, G., Wang, J., Kumar, V., Zajonc, D.M., 2017. Crystal structure of Qa-1a with bound Qa-1 determinant modifier peptide. PLoS One 12, e0182296.

Zeng, L., Sullivan, L.C., Vivian, J.P., Walpole, N.G., Harpur, C.M., Rossjohn, J., Clements, C.S., Brooks, A.G., 2012. A Structural Basis for Antigen Presentation by the MHC Class Ib Molecule, Qa-1 b . J. Immunol. 188, 302–310.

Zhang, C., Anderson, A., DeLisi, C., 1998. Structural principles that govern the peptide-binding motifs of class I MHC molecules. J. Mol. Biol. 281, 929–947.

Zuhair, M., Smit, G.S.A., Wallis, G., Jabbar, F., Smith, C., Devleesschauwer, B., Griffiths, P., 2019. Estimation of the worldwide seroprevalence of cytomegalovirus: A systematic review and meta-analysis. Rev. Med. Virol. 29, e2034.