Computational Mapping and In Vitro / In Vivo Evaluation of Immunogenic Epitopes for Lassa Fever Virus

By

Agnieszka Boesen

A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2005

This manuscript has been read and accepted for the Graduate Faculty in Biology in Satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

1/14/2005
Date

Chair of Examining Committee
Dr. Richard Coico, City College (Medical School)

1/19/05
Date

Executive Officer
Dr. Richard L. Chappell

Dr. Paul Gottlieb, City College (Medical School)

Dr. Mark Pezzano, City College

Dr. Laurel Eckhardt, Hunter College

Dr. John Coligan, National Institute of Allergy and Infectious Diseases/NIH

Supervising Committee

The City University of New York

Abstract

COMPUTATIONAL MAPPING AND IN VITRO / IN VIVO
EVALUATION OF IMMUNOGENIC EPITOPES FOR LASSA FEVER VIRUS

by

Agnieszka Boesen

Advisor: Professor Richard F. Coico

Lassa Fever Virus (LFV) causes severe hemorrhagic fever in humans with a high fatality rate of approximately 30%. LFV is endemic in rural West Africa and spreads to humans from a rodent host by contact with animal blood, urine or feces and between humans as a result of exposure to blood and body fluids. LFV is known to have been weaponized for potential use as a biowarfare agent. Unfortunately, there are currently no effective treatments for LFV infections or vaccines to prevent infection with this virus. Cellular responses mediated by cytotoxic T lymphocytes (CTL) play a primary role in neutralizing LFV during infection. The identification of the immunogenic epitopes expressed by LFV that contribute to the generation of class I-restricted effector cells may lead to new strategies for LFV-directed therapies, diagnostic tools, and, more importantly, protective vaccines.

In this study we performed computational analyses using BIMAS and SYFPEITHI algorithms to map potential immunogenic epitopes present within LFV glycoprotein (GP) and nucleoprotein (NP). Our analysis focused on the identification of

CTL epitopes with binding affinity for the MHC class I allele HLA*0201. We have also computationally identified T-helper cell and B-cell epitopes for LFV GP and NP using ProPred, SYFPEITHI and hydropathy, hydrobhobicity algorithms, respectively. Five nanomer peptides, candidates for CTL epitopes, derived form GP and two peptides derived from NP were identified and tested for their ability to stabilize class I MHC expression using HLA*0201^{+} T2 cells. Three of the five GP peptides significantly stabilized class I MHC expression on these cells whereas none of the NP peptides tested contributed to the stabilization of this class I allele.

Three peptides derived from LFV glycoprotein (GP258-266, GP441-449, and GP60-68) that were shown to stabilize class I MHC on T2 cells and one peptide (GP386-394) that did not stabilize the class I molecule *in vitro*, were chosen to be tested for their potential to generate CTLs *in vivo*. HLA*0201 transgenic mice were immunized with LFV peptides together with a T-helper peptide specific for I-A^{b} for Hepatitis B virus core protein 128-140 TPPAYRPPNAPIL. Splenocytes and the lymph nodes from the immunized mice were then tested for their cytolytic activity. We also enumerated IFN-γ-producing cells in an ELISPOT and intracellular IFN-γ staining assays. Three out of four *in vivo* tested peptides induced IFN-γ responses as well as CTL responses in immunized mice. LFV peptide-specific CD8 CTL responses generated *in vivo* by two of the tested peptides (GP258-266 and GP60-68) were essentially of the same magnitude as those seen in mice immunized with a positive control peptide. The third peptide (GP441-449) induced moderate responses *in vivo*. Finally, one peptide that did not stabilize class I molecules *in vitro* (GP386-394) failed to induce the CTL responses *in vivo*. We also enumerated the frequency of LFV peptide-specific CD8+ cells for one of the synthetic

peptides (GP258-266) using class I MHC tetramer staining. This analysis showed a small but significant population of peptide-binding $CD8^+$ cells in immunized mice as compared with control mice. Based on our results we conclude that computational prediction of virus-specific epitopes is a valid and useful tool for epitope mapping.

To

My parents and my husband

For their loving support and inspiration

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my mentor, Dr. Richard F. Coico. He has made my doctoral studies an intellectually fulfilling experience. I also want to thank my thesis committee members, Drs. John Coligan, Paul Gottlieb, Mark Pezzano and Laurel Eckhardt for their time and helpful suggestions.

I wish to thank Dr. Krishnan Sundar for sharing his knowledge during numerous scientific discussions. I also wish to express my special thanks to Dr. Viera Lima for her continued help and support. Finally, I want to thank Igor Toporovsky and many of my friends and colleagues for their support during this endeavor.

TABLE OF CONTENTS

APPROVAL PAGE..... ii
ABSTRACT..... iii
FOREWORD..... vi
ACKNOWLEDGEMENTS..... vii
TABLE OF CONTENTS..... viii
LIST OF FIGURES..... ix
LIST OF TABLES..... xi
ABBREVIATIONS..... xiii
INTRODUCTION..... 1
MATERIALS AND METHODS..... 42
RESULTS..... 61

PART I. Computational identification of CTL epitopes, T-helper cell and B-cell epitopes..... 61

1. Computational identification of LFV CTL epitopes..... 61

2. Computational identification of LFV T-helper cell epitopes..... 71

3. Computational identification of LFV B-cell epitopes..... 99

PART II. Validation of computational predictions of CTL epitopes..... 107

1. Standardization of MHC-stabilization assay..... 107

2. Evaluation of synthetic peptides as candidate CTL epitopes..... 112

PART III. In vivo investigation of immunogenicity of Lassa virus peptides 118

1. IFN-γ ELISPOT assay..... 119

2. Intracellular IFN-γ staining..... 121

3. Non-radioactive CTL assay..... 125

4. Tetramer staining..... 127

DISCUSSION..... 131
REFERENCES..... 154

LIST OF FIGURES

Figure 1. Crystal structure of two viral peptides in complex with murine MHC class I H-2Kb.....10

Figure 2. A peptide sitting in the groove of a class I MHC molecule.....11

Figure 3. A nonapeptide sitting in an MHC class I groove.....11

Figure 4. Polymorphic residues of HLA-A2.....12

Figure 5. TCR and its orientation in the TCR-MHC class I/peptide complex.....15

Figure 6. MHC-peptide-TCR interaction.....16

Figure 7. Lassa Fever Virus segment S, complete sequence Lassa Fever Virus segment L, complete sequence.....31

Figure 8 Construction of peptide-MHC tetramers.....53

Figure 9. B-cell epitope prediction: Graphical presentation of results generated using the ProtScale algorithm for Lassa Fever Virus GP.....102

Figure 10. B-cell epitope mapping.....103

Figure 11. B-cell epitope prediction: Graphical presentation of results generated using the ProtScale algorithm for Lassa Fever Virus NP.....105

Figure 12. B-cell epitope mapping.....106

Figure 13. Determination of anti-HLA-A2.1 antibody concentration.....109

Figure 14. Determination of antigenic peptide concentration.....110

Figure 15. MHC stabilization assay at varying temperatures.....111

Figure 16. Peptide binding to HLA-A2.1 as assessed using the MHC class I – stabilization.....115

Figure 17. MHC class I stabilization by LFV GP and NP peptides as measured by MFI.....116

Figure 18. Peptide binding to HLA-A2.1 as assessed using the MHC class I - stabilization.....117

Figure 19. Enumeration of IFN-γ producing cells.....120

Figure 20. ICCS for IFN-γ using lymphocytes stimulated with PMA and inomycin.....122

Figure 21. Enumeration of CD8-positive IFN-γ producing cells by ICCS.....123

Figure 22. Specific lysis of peptide-pulsed cells by peptide-primed effector cells from HLA-A2.1 transgenic mice.....126

Figure 23. Enumeration of Lassa LLG Peptide Binding CTLs by Tetramer Staining.....127

LIST OF TABLES

Table 1. Category A Priority Pathogens.....27

Table 2. Summary of algorithms available for CTL epitope mapping.....66

Table 3. BIMAS: HLA Coefficient Table for File "A_0201_ Standard".....67

Table 4. Computational prediction of LFV CTL epitopes.....68

Table 5. Computational prediction of LFV CTL epitopes using other available algorithms,GP.....69

Table 6. Computational predeiction of LFV CTL epitopes using other available algorithms,
NP.....70

Table 7. An example of a virtual matrix for HLA-DRB1*0101.....75

Table 8. Computational prediction of MHC class II-binding peptides for 51 different HLA-DR alleles for Lassa GP, allele cross reactivity.....76

Table 9. Comparison of HLADRB1*0101 epitope prediction by ProPred and SYFPEITHI algorithms for glycoprotein of Lassa Fever Virus.....85

Table 10. Lassa Fever Virus GP sequence with highlighted proposed cross-reactive MHC class II epitopes for 51 different HLA-DR alleles.....86

Table 11. Computational prediction of MHC class II binding peptides for 51 different HLA-DR alleles for Lassa NP, allele cross reactivity.....87

Table 12. Comparison of HLA-DRB1*0101 epitope prediction by ProPred and SYFPEITHI algorithms for nucleoprotein of Lassa Fever Virus.....98

Table 13. Lassa Fever Virus NP sequence with highlighted proposed cross reactive MHC class II epitopes for 51 different HLA-DR alleles.....99

Table 14. Amino acid scale values (conformational parameters for β-turns) and individual hydrophobicity values.....101

Table 15. Lassa Fever Virus GP sequence with highlighted
proposed B-cell epitopes.....104

Table 16. Lassa Fever Virus NP sequence with highlighted
proposed B-cell epitopes.....107

Table 17. Collective ranking of LFV peptides identified as immunogenic candidates.....128

Table 18. Summary of collective ranking of GP-derived LFV epitopes identified as immunogenic peptides.....129

Table 19. Influenza protein (CAA24282) containing positive control peptide, analyzed by BIMAS algorithm.....130

ABBREVIATIONS

aa residue(s) = amino acid(s)
APC antigen presenting cell
CFA complete Freund's adjuvant
CTL cytotoxic T lymphocyte
DMSO dimethyl sulfoxide
ELISPOT enzyme-linked immunosorbent assay
ER endoplasmic reticulum
FACS fluorescence-activated cell sorting
FBS fetal bovine serum
FITC fluorescein
HLA human leukocyte antigen
IFA incomplete Freund's adjuvant
IFN interferon
Ig immunoglobulin
IL interleukin
i.v. intravenous
LMP low-molecular mass polypeptide
MECL multicatalytic endoplasmatic complex like
MHC I/II major histocompatibility complex class I/II
NP nucleoprotein
PE phycoerythrin
PerCP peridinin-chlorophyll-protein complex
PMA phorbol 12-myristate-13-acetate
TAP transporter associated with antigen processing
TCR T cell receptor

Abbreviations for amino acid names:

1-letter code	3-letter code	Full name	Selected properties
A	Ala	Alanine	small
C	Cys	Cysteine	small, polar
D	Asp	Aspartic acid	negative charge
E	Glu	Glutamic acid	negative charge
F	Phe	Phenylalanine	bulky, hydrophobic
G	Gly	Glycine	small
H	His	Histidine	positive charge
I	Ile	Isoleucine	branched-chain, hydrophobic
K	Lys	Lysine	positive charge
L	Leu	Leucine	branched-chain, hydrophobic
M	Met	Methionine	hydrophobic
N	Asn	Asparagine	polar
P	Pro	Proline	kinked, polar
Q	Gln	Glutamine	polar
R	Arg	Arginine	positive charge
S	Ser	Serine	small, polar
T	Thr	Threonine	small, polar
V	Val	Valine	branched-chain, hydrophobic
W	Trp	Tryptophan	bulky, hydrophobic
Y	Tyr	Tyrosine	bulky, hydrophobic

INTRODUCTION

Epitope Identification and Computational Prediction Strategies

In the past 15 years, bioinformatics tools have been successfully exploited to help investigators answer many questions including epitope mapping. These tools greatly facilitate the ability of researchers to move from genome sequence to novel vaccine strategies that can be tested in animal models. The first step to applying bioinformatics to epitope mapping is to distinguish patterns that are potentially immunostimulatory from patterns that are not. Several bioinformatic programs, including the ones used in this study, are designed to identify pathogen-derived peptides that bind with high affinity to specific major histocompatibility complex (MHC) determinants thus providing a rapid means of honing-in on candidate epitopes that may induce protective T cell responses (De Groot *et. al.*, 2001).

Rötzschke and Falk (Rötzschke *et. al.*, 1991; Falk *et. al.*, 1991) were the first to describe the patterns of amino acids contained in peptides known to bind to MHC molecules on the surface of antigen presenting cells. They called these patterns *MHC binding motifs*. For example, one such motif would be, – L – – – – – V, where the letters represent amino acids and the dashes represent any amino acid. They further suggested that the patterns could be used to find peptides that would bind to MHC, engage T cell receptors, and stimulate T cell responses. The peptides that do so are called *T cell epitopes*. Such epitopes bind within the binding groove of MHC class I and class II molecules through interactions between their R group side chains and pockets located on the floor of the MHC. Different MHC molecules have different types of binding pocket-R group interactions, limiting the set of MHC-binding peptides that can be presented in the

context of any given MHC. The structural constraints on peptides binding the MHC groove result in MHC-linked "genetic restriction" – a phenomenon first described by Zinkernagel and Doherty (1997). The problem of selecting the peptides from each pathogen that will stimulate a protective immune response in the context of each individual's MHC has been considered a major stumbling block to the development of T cell epitope-based vaccines. This problem has been mitigated by the discovery of MHC allele *supertypes* described by Sette (1998, 1999) as will be discussed later in this proposal.

Several different computational approaches to the *in silica* identification of immunogenic epitopes have been described. The first MHC-binding-motif based algorithms were described by Rötzschke, Falk and Leighton (Rötzschke *et. al.*, 1991; Falk *et. al.*, 1991; Leighton *et. al.*, 1991), by Lipford (Lipford *et. al.*, 1993), by Parker *et. al.* (1994) and Meister *et al* (1995). Other bioinformatic approaches to predict T cell epitopes include artificial neural networks first described by Brusik *et. al.* (1994) and structural approaches, as described by Rosenfeld *et. al.* (1995). Prediction of peptides using MHC binding motif-based algorithms known as "extended MHC binding motifs" or "peptide side chain scanning" was described by Sette *et al* (1993) and Hammer *et al* (1994). Those methods allowed for the construction of a matrix of all possible amino acid side chain effects for a single MHC-binding motif.

Cytotoxic T Lymphocyte (CTL): Role in Control of Infectious Diseases

CTL-mediated cytotoxicity of target cells infected with intracellular pathogens is mediated by the release of cytolytic granules by CTLs, in the granule-exocytosis pathway

or by the ability of the CTLs to engage cell-surface death receptors, such as members of tumor necrosis factor receptor (TNFR) family, including FAS (CD95) (Lieberman 2003). The CTL response is initiated when CTL precursors are stimulated by binding of their T cell receptor (TCR) and CD8 to the antigenic peptide on class I MHC molecules. The area of apposition of the T-cell with its target assembles into a well-organized *immunological synapse.* In the central region of the T-cell synapse, the TCR-CD3 complex, CD8 and associated signaling molecules cluster and are surrounded by molecules such as CD2 and leukocyte function-associated antigen 1 (LFA1), that stabilize the synapse. The synapse forms within minutes of the initial interaction of the TCR with its antigen and lasts for more than an hour until the entire TCR complex is internalized and degraded (Russel *et. al.,* 2002).

The cytotoxic granules contain the pore-forming protein *perforin* and a family of serine proteases known as *granzymes* (for granule enzyme). Granzymes A and B are the most abundant enzymes in mice and humans. Granzyme A activates an endoplasmic reticulum-associated caspase-independent nuclease to induce single stranded DNA nicks and thus contributes to the target cell death (Beresford *et. al.*, 2001). Granzyme B activates the caspase-dependent as well as caspase-independent pathway of apoptosis (Metkar *et. al.,* 2002).

The release of granule contents accounts for most of the cytotoxic activity of CD8 effector CTLs but they can also execute their cytolytic function through a perforin-independent mechanism of cytotoxicity. That mechanism involves the binding of FAS (CD95) in the target cell membrane by the FAS ligand (CD95L), which is expressed in

the membranes of activated CTLs and T_H1 cells. Activation of FAS leads to target cell apoptosis (Fisher *et. al.,* 1995).

Another function of cytotoxic CD8 T cells in control of infectious diseases is the production of interferon- γ (IFN-γ), which can inhibit viral replication and induces MHC class I expression and macrophage activation (Fernandez *et. al.,* 1992).

Antigen Processing and Presentation by Class I MHC Molecules

Cytotoxic T cells recognize a complex formed between class I MHC molecules and antigenic peptides derived from intracellular processing of proteins (Germain *et. al.*, 1994). The peptide fragments that stimulate T cells are recognized only when bound to an appropriate, syngeneic MHC molecule (MHC restriction). The continuous presentation of non-self, antigen-derived peptides by MHC class I molecules to cytotoxic T lymphocytes initiates the adaptive immune response (York *et. al.*, 1996). At least three factors play a role in the induction of CTL responses to a given epitope. These include: (1) the efficiency with which a given epitope is generated through intracellular processing; (2) the binding affinity of the epitope to MHC class I molecules; and, (3) the existence of a suitable TCR repertoire. Since immunogens capable of stimulating CTL responses are typically designed to contain highly selected epitopes with high MHC binding affinity and defined immunogenicity, the efficiency of epitope processing may be one of the dominant variables effecting immunogenicity of multiepitope vaccine constructs (Livingston *et. al.*, 2001).

All viruses and some bacteria replicate in the cytosolic compartment of the cell. Processed viral antigens are then presented as peptides bound to MHC class I molecules

to CD8 T cells. Other bacteria and some parasites are engulfed into endosomes, usually by phagocytic cells such as macrophages, and are able to proliferate within the endocytic vesicles (endoplasmic reticulum, Golgi apparatus, endosomes, lysosomes). Extracellular pathogens may enter the vesicular system of the cells by binding to surface molecules (immunoglobulins) followed by endocytosis. The antigens derived form extracellular and intravesicular pathogens are presented by MHC class II molecules to CD4 T cells (Morrison *et. al.*, 1986). Since this study is focused on viral epitopoes, antigen presentation by MHC class I molecules will be the main consideration of this discussion.

The structural differences between class I and class II MHC molecules allow them to serve distinct functions in antigen presentation. MHC class I molecules consist of two polypeptide chains, an α or heavy chain encoded in the MHC, and a smaller non-covalently associated chain, β_2m, which is not encoded in the MHC. Only the class I α chain spans the cell membrane. The molecule has four domains, three formed from the MHC encoded α chain, and one contributed by β_2m. The α_1 and α_2 domains of class I heavy chains form a β-sheet platform, spanned by two α-helices that together form a peptide-binding cleft (Madden *et. al.*, 1993). Peptides that bind to MHC class I molecules are eight to eleven amino acids long and MHC molecules are unstable when peptides are not bound. The anchor residues, which are determined by the nature of the side chains in the binding pocket of a particular MHC class I allele-binding groove, are required for peptide binding, whereas the N- and C-terminal residues of peptides contribute significantly to binding affinity. The sum of requirements for allele-specific peptide-MHC class I interactions is believed to define the binding motif for a given MHC class I molecule (Wang *et. al.*, 2002).

The antigen fragments that bind to MHC class I molecules for presentation to CD8 T cells are typically derived from viruses that take over the cell's biosynthetic mechanisms to make their own proteins. Those proteins have to be degraded and antigen fragments have to be processed in order to be presented by the MHC class I molecules. A major part in cytosolic protein degradation is played by multicatalytic protease complex called the *proteasome*. The catalytic core of proteasomes is a cylindrical multi-subunit complex arranged as four axially stacked heptameric rings. The two outer rings are composed exclusively of seven homologous α subunits, whereas the two inner rings are composed of seven homologous β subunits. Proteasome α subunits are involved in proteasome structure whereas β subunits function in catalytic activities. Because the catalytic sites of proteasomes regulate each other allosterically, proteasomes degrade protein substrates by an ordered, cyclical bite-chew mechanism (Coux *et. al.,* 1996). In constitutive proteasomes, the three catalytic activities, chymotryptic-like, trypsin-like and peptidylglutamyl peptide hydrolase-like, are contributed by the β subunits X, Y and Z, respectively, and by their corresponding interferon–inducible LMP2, LMP7, and MECL 1. Upon infection, the interferon-inducible LMP2, LMP7 and MECL1 catalytic β subunits replace their catalytic β-subunit homologues, X, Y and Z, respectively and form newly assembled immunoproteasome that possesses altered catalytic specificities (Hwang *et. al.,* 2001). Cytotoxic T lymphocyte-mediated responses are mainly directed to immunoproteasome-dependent T-cell epitopes during the peak phase of viral elimination (Khan *et. al.,* 2001). Immunoproteasome-dependent processing of MHC class I-binding "antigens" appears to efficiently generate peptides that are optimal for translocation by

transporter associated with antigen processing (TAP) complexes and for binding to MHC class I heterodimers (Yang 2003).

TAP subunits *TAP1* and *TAP2, TAP1*-associated protein -*tapasin* together with proteasome catalytic subunits LMP2 and LMP7 are encoded within the MHC region and are upregulated during infection (York *et. al.,* 1996). MHC class I α chains assemble in the endoplasmic reticulum with a membrane-bound protein, calnexin. When this complex binds β_2m it is released from calnexin, and the partially folded MHC class I molecule then binds to the *TAP1* subunit of the TAP transporter by interacting with *TAP1*-associated protein tapasin. The prerequisite for releasing MHC class I molecules from MHC class I peptide-loading complexes is the binding of eight- to eleven-amino acid peptides, which completes the folding of the MHC class I molecules. Peptides generated by the degradation of proteins in the cytoplasm are transported into the lumen of the endoplasmic reticulum by the TAP transporter. Once the peptide has bound to the MHC molecule the trimolecular complex is transported through the Golgi to the cell surface. If MHC class I binding peptides are not optimal, MHC class I molecules are retrieved from the Golgi to the ER for another round of peptide loading (Ortmann *et. al.*, 1997; Pamer *et. al.,* 1998).

Structural Basis of MHC Class I Peptide Binding

As mentioned previously, MHC class I molecules consist of two polypeptide chains, an α or heavy chain encoded in the MHC, and a smaller non-covalently associated chain, β_2m, which is not encoded in the MHC. Only the class I α chain spans the cell membrane. The molecule has four domains, three formed from the MHC encoded α

chain, and one contributed by β_2m. The α_1 and α_2 domains of the MHC class I heavy chains form a β-sheet platform, spanned by two α-helices that together form a peptide binding cleft (Figure1) (Madden *et. al.*, 1993). The peptide-binding groove of MHC class I molecules tethers the amino and carboxyl termini of peptides through a network of hydrogen bonds between several conserved residues at each end of the groove and the peptide backbone (Figure 2) (Fremont *et. al.,* 1992). Peptides that bind to MHC class I molecules are linear and usually eight to eleven amino acids long. The binding of the peptide is stabilized at its two ends by contacts between atoms in the free amino and carboxy termini and invariant sites that are found at each end of the peptide-binding groove of MHC class I molecules (Figure 3). These contacts are thought to be the main stabilizing contacts for peptide:MHC class I complexes because synthetic peptide analogs lacking terminal amino and carboxyl groups fail to bind stably to the MHC class I molecules (Bouvier *et. al.,* 1994). The peptide lies in an elongated conformation along the groove. Variations in peptide length appear to be accommodated, in most cases, by a kinking in a peptide backbone.

A major characteristic of MHC class I heavy chains is their high degree of polymorphism within the α1 and α2 domains (Figure 4). In studies using Caucasoid populations, the MHC class I B gene was found to have 207 alleles, the gene C was found to have 50 alleles, and the gene A was found to have 95 alleles (Marsh, 1998). The polymorphic residues that line the peptide-binding groove determine the peptide-binding properties of the different MHC molecules. The interaction between the polymorphic amino acid side chains and the peptide are such that different allelic variants of MHC molecules bind different peptides preferentially. Peptides binding to a given allelic

variant of an MHC molecule have been shown to have the same or very similar amino acid residues at two or three specific positions along the peptide sequence. The amino acid side chains at these positions insert into pockets in the MHC molecule that are lined by the polymorphic amino acids. Because the binding of these side chains anchors the peptide to the MHC molecule, the peptide residues involved have been called *anchor residues*. Both the position and identity of these anchor residues can vary depending on the particular MHC class I allele that is binding the peptide. Most peptides that bind to MHC class I molecules have an anchor residue at the carboxy terminus that is usually hydrophobic or in some instances basic. Most synthetic peptides of suitable length that contain these anchor residues will bind the appropriate MHC class I molecule (Weiss *et. al.,* 1995). The set of anchor residues that allows binding to a given MHC class I molecule is called the *sequence motif.* These sequence motifs make it possible to identify peptides within a protein that can potentially bind the appropriate MHC molecule, which may be important in the design of peptides vaccines (Babbitt *et. al.*, 1985).

In the present study, the immunogenicity of T-cell peptide epitopes was investigated by focusing on their specific associations with the MHC class I allele used to predict such epitopes, namely, HLA-A2.1. Details of how we approached this final objective will be discussed in the sections that follow.

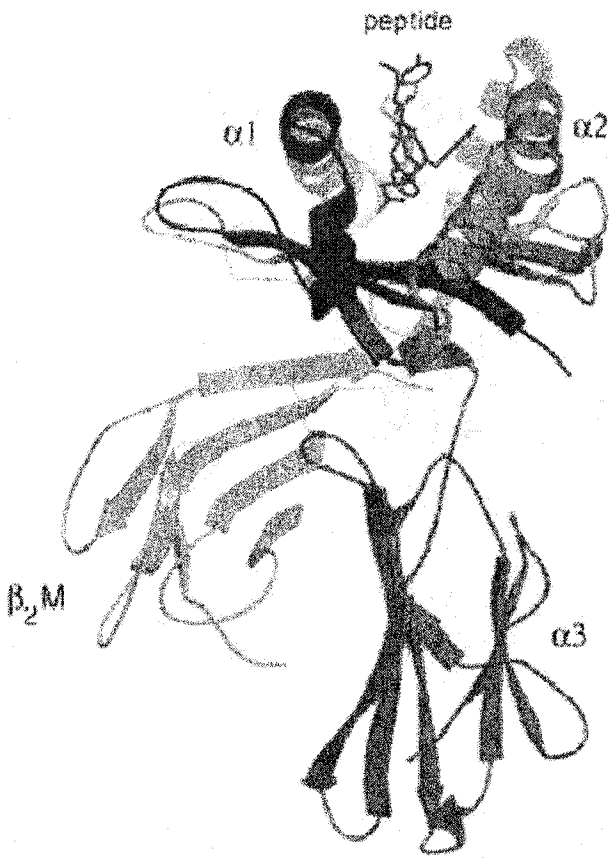


Figure 1. Crystal structure of two viral peptides in complex with murine MHC class I H-2Kb. The α1 and α2 domains of MHC class I heavy chain form a β-sheet platform spanned by two α-helices that together form a peptide-binding pocket and contact both the α3 domain of the MHC class I heavy chain and the light chain of β_2m (Fremont *et. al.*, 1992).

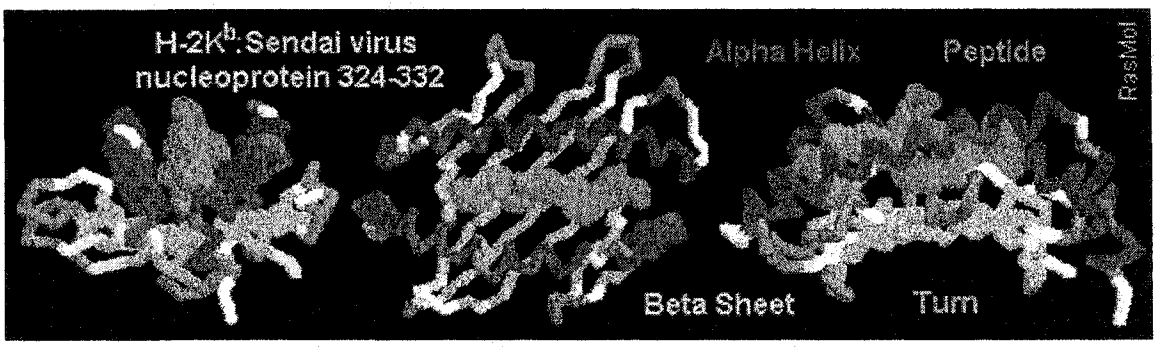


Figure 2. A peptide (green) sitting in the groove of a class I MHC molecule (mouse H-2Kb, alpha-1 and alpha-2 domains only). The center image, is the face to which the T cell antigen receptor would bind.
(1VAB.PDB; Rasmol http://www.umass.edu/microbio/rasmol/mhc-x3.htm)

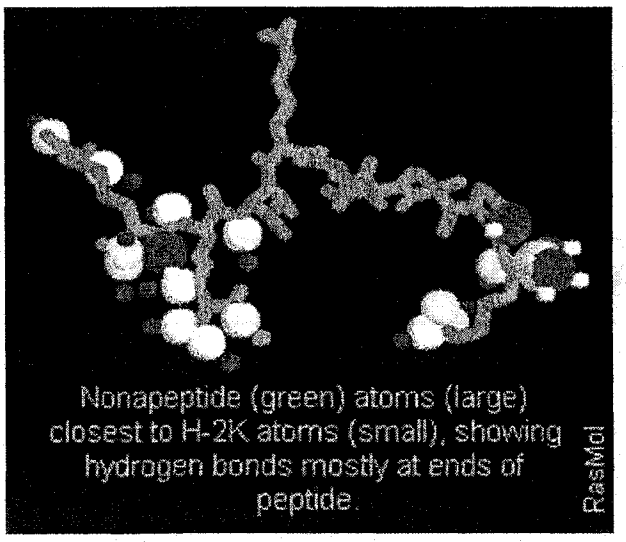


Figure 3. A nonapeptide sitting in an MHC class I groove. The nonapeptide is shown as a green stick model, including all amino acid side chains. The only peptide carbon atom close to the MHC is that of the carboxy terminus. Notice the absence of bonding to the center of the peptide binding.
(1ROG.PDB Rasmol http://www.umass.edu/microbio/rasmol/)

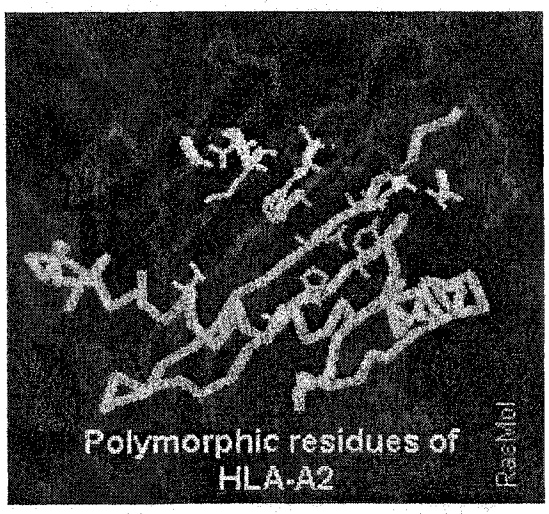


Figure 4. Polymorphic residues of HLA-A2. The α1 and α 2 domains of the human HLA-A2 molecule are shown as backbone representations. Amino acids, which vary between this and other HLA-A locus alleles, are shown as stick models, showing their side-chains. The point is that the residues that vary between alleles mostly point into the groove, thereby affecting peptide binding. (3HLA.PDB; Rasmol http://www.umass.edu/microbio/rasmol/).

Structural Basis for T-cell Antigen Receptor (TCR) Recognition of Peptide Complexed with MHC

As mentioned earlier, the specificity of T-cell responses is determined by recognition of an antigenic peptide bound to class I or class II MHC molecules by the antigen-specific TCR. This TCR interaction initiates signals that, together with non-antigen specific signals, trigger development of the T-cell repertoire, regulation of the immune response and activation of T cells in general.

The TCR consists of two transmembrane glycoprotein chains, α and β, bound to one another by a disulphide bond. The external portion of each chain consists of two domains, resembling immunoglobulin (Fab fragment) variable and constant region domains, respectively. Both chains have carbohydrate side chains attached to each domain. A short segment, analogous to an immunoglobulin hinge region, connects the

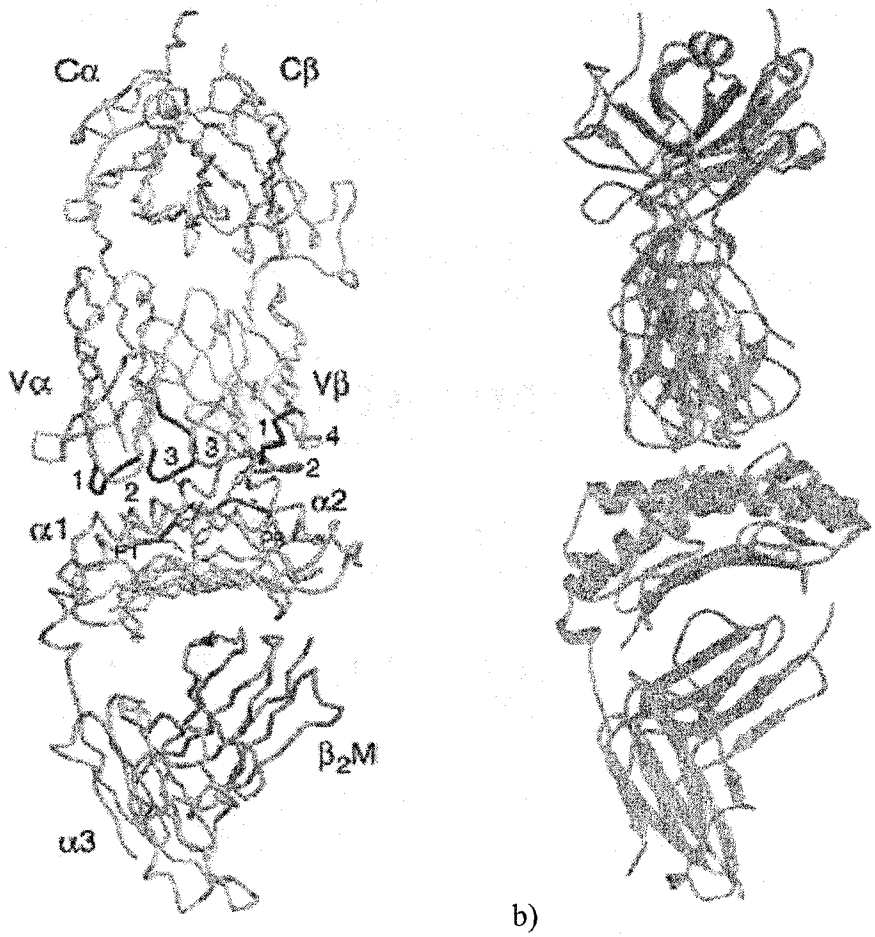
immunoglobulin-like domains to the membrane. There is an alternative type of TCR made up of different polypeptides designated γ and δ (Garboczi *et. al.*, 1996).

Like antibody molecules, αβ TCRs are encoded by a family of gene segments that rearrange to generate the diverse repertoire. Although TCRs and antibodies are both assembled by genetic recombination of V, D, J, and C segments, only antibodies undergo somatic mutation (i.e. affinity maturation) to increase their affinity for antigen. The diversity of TCRs is again generated during rearrangement and it is focused on the third complementarity-determining region (CDR3) (Kroneberg *et. al.*, 1986). The TCR/pMHC (peptide+MHC) affinities are substantially weaker than those of antibody/antigen complexes (Matsui *et. al.*, 1991).

In an antibody, the center of the antigen-binding site is formed by the CDR3 (hypervariable region) of the heavy and light chains. The structurally equivalent loops of the TCR α and β chains also form the center of the antigen-binding site on TCR, whereas the periphery of the site consists of the equivalent of the CDR1 and CDR2 loops (Garcia *et. al.*, 1998). Forty-two Vα and 46 Vβ genes encode the first two variable loops CDR1 and CDR2 of the human α – and β-chain; the third variable loops are encoded following the joining of V genes to one of a number of J-region and, in case of Vβ, D and J-region gene segments (Garboczi *et. al.*, 1996).

TCRs bind to MHC class I: peptide complexes with the long axis of the CDR diagonally across the peptide-binding groove of the MHC (Figure 5). In this orientation, the Vα domain makes contact primarily with the amino terminus of the bound peptide and the surrounding helices of the MHC class I molecule, whereas the Vβ domain contacts the carboxy terminus of the bound peptide and the surrounding helices of the

MHC class I molecule. The CDR3 loops of the TCR make the main contacts with the central residues of the peptide (Figure 6). The interface between Vα and Vβ CDR3 loops creates a deep pocket in which a side chain from the peptide binds (Ding *et. al.,* 1998). The CDR3 loops also make contacts with the MHC molecule as well as with the peptide. The TCR is not placed symmetrically over the MHC molecule. Whereas the Vα CDR1 and CDR2 loops are in close contact with the MHC:peptide complex at the amino terminus of the bound peptide, the β-chain CDR1 and CDR2 loops, which interact with the complex at the carboxy end of the bound peptide, appear to contribute much less to the binding (Teng *et. al.,* 1998). In most cases, once the TCR recognizes and binds to the MHC:peptide complex, the peptide will be buried in the TCR/MHC interface, and the peptide will be bound much more deeply in the MHC molecule than in the TCR. It is also noteworthy that the TCR does not bind to the MHC:peptide in isolation, but as part of multicomponent signaling complex, that includes the CD3 γ, δ, ε, and ζ chains, and coreceptors CD8 or CD4 (Janeway, 1992).



a) b)

Figure 5. TCR and its orientation in the TCR-MHC class I/peptide complex. a) A backbone tube representation of a T-cell receptor bound to an MHC class I molecule (bottom, α1-α-3 domains of heavy chain and β2m) and representing an octamer peptide with P1 residue positioned toward the left. The T-cell receptor covers the MHC class I-binding groove such that the Vα and Vβ complementary determining regions 1 and 2 (CDR1) are positioned over the N- and C-terminal regions of the bound peptide, respectively. The Vα and Vβ complementary-determining region 3 (CDR3) straddles the peptide between the helices around the central position of the peptide. β chain HV4 loop is visible (Garcia *et. al.,* 1996). b) Ribbon representation of HLA-A2 heavy chain, β2m, TCR α-chain and TCR β-chain (1OGA.PDB; Steward-Jones *et. al.*, 2003).

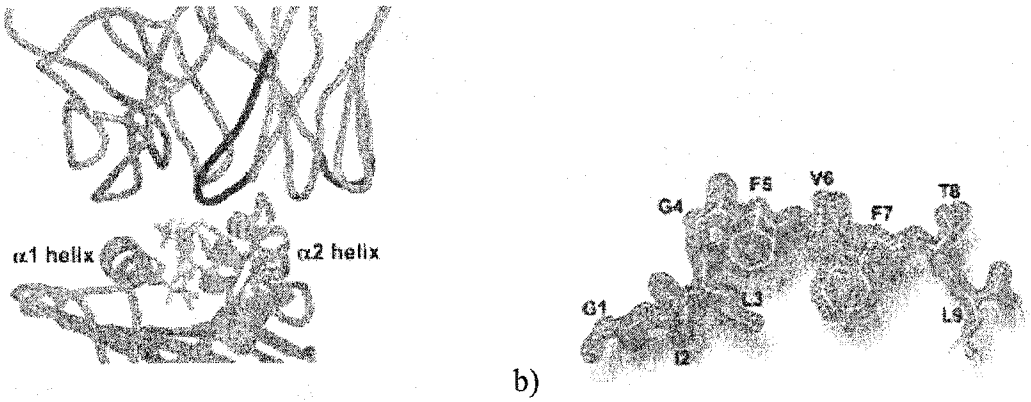


Figure 6. MHC-peptide-TCR interaction. a) Diagram highlighting the CDR loops of a TCR in complex with peptide (GILGFVFTL) and HLA-A2. b) GILGFVFTL peptide (1OGA.PDB; Steward-Jones *et. al.*, 2003).

Polymorphism of Human Leukocyte Antigens (HLA): Challenges for Vaccine Development

In humans, the molecules that constitute the MHC are called *human leukocyte antigens* (HLA). Consistent with other mammalian species, the protein products of human class I and class II MHC molecules are highly polymorphic. HLA class I molecules are cell surface glycoproteins encoded by a cluster of genes whose main biologic function is to bind antigenic peptides and present them to T cells (Germain *et. al.*, 1994). The HLA Nomenclature Committee report issued by the World Health Organization lists more than 500 different HLA class I and class II alleles (Bodmer *et. al.*, 1997). The frequency of different allelic variants of HLA molecules varies greatly among different ethnicities. Crystallographic studies have shown that HLA polymorphic regions correspond to the MHC binding pockets, which represent the variable parts of the MHC structure (Madden *et. al.*, 1995). Each allelic variant of the MHC molecule selects

peptides based on the complementary structure of the peptide and the polymorphic pocket within the peptide-binding groove. Thus, different HLA molecules are associated with different peptide binding patterns. The large degree of HLA polymorphism presents a significant challenge to investigators wishing to exploit specific epitopes identified as candidate immunogens because they must cover a wide range of MHC specificities. An even greater problem arises when it comes to coverage of different populations of individuals, where hundreds of different sets of epitopes would be required to generate an effective immune response. A proposed solution to this problem is based on the identification of peptides capable of binding to multiple HLA alleles.

Sidney and Sette (Sidney *et. al.*, 1995; 1996; 1996a; 2001) demonstrated that different HLA class I molecules can recognize similar peptide binding motifs, and that a significant overlap exists in their peptide-binding repertoires. They also observed that HLA molecules with overlapping peptide-binding repertoires can be grouped into families called HLA *supertypes*. A supertype is defined as a set of HLA molecules that have similar peptide binding motifs and overlapping peptide binding repertoires (Sette *et. al.*, 2002). Approximately 90% of all known HLA class I molecules can be classified in one of nine major HLA supertypes. Each of the supertypes is present in about 35% - 55% of the general population, irrespective of ethnicity (Sidney J *et. al.*, 1996b). The most prevalent HLA-A supertype is the A2 supertype, which was defined on the basis of observations by Fruci *et. al.* (1993) and from data obtained using live cell-binding assays (del Guercio *et. al.* 1995). The A2 supertype includes, A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*6802 and A*6901. These alleles recognize peptides of 9-10 residues in length which bear small or aliphatic hydrophobic residues L,

I, V, M, A, T, or Q at position 2 and L, I, V, M, A, or T at the peptide C-terminus (Sette *et. al.*, 1999). On the basis of structural analysis, other A*02 molecules for which peptide-binding motifs have not been experimentally determined are also expected to share the A2-supertype specificity. These include A*0208, A*0209, A*0210, A*0211, A*0212, A*0213, and A*0214.

Sidney et. al. (2001) have investigated the structural rules underlying peptide binding and cross-reactivity to A2-supertype molecules. To investigate the peptide-binding specificity of A2-supertype molecules, single substitution analog peptides and peptide libraries were used. They reported that over 70% of the peptides that bind to A*0201 with high affinity will also bind other A2 supertype molecules and thus the A2-supertype molecules share largely overlapping peptide binding repertoires. They also reported that the propensity to bind other A2-supertype alleles is highly correlated with A*0201 binding affinity and that A2-supertype molecules recognize similar features at primary (dominant) and secondary (auxiliary) anchor positions of their peptide ligands. As noted above, anchor residues are the amino acids side chains within the peptide that attach the peptide into the HLA class I peptide binding site by having a structure that is complementary to a pocket of the peptide-binding groove. Some of the other positions within the endogenous peptides are also enriched for specific amino acids and are defined as auxiliary anchor residues (Parker *et. al.*, 1994). The observation of cross-reactivity among A2 supertype molecules reflects the high structural similarity of those molecules. It has been shown that common A2-supertype alleles differ from each other by an average of 4.7 residues, considering MHC α1 and α2 domains (residues 1-182), which comprise the peptide-binding cleft (Fremont *et. al.*, 1992). By contrast, common A3- and

B7-supertype alleles differ from each other by an average of 9.1 and 13.1 residues, respectively. The existence of peptide-binding cross-reactivity to A2-supertype molecules may be exploitable with regard to vaccine development. Our choice of the HLA-A2.1 molecule was based upon the widespread frequency of this allele (average 43% coverage) in different ethnic populations in United States (Cao *et. al.*, 2001). Moreover, the availability of cell lines expressing HLA-A2.1 and of HLA-$A2^+$ transgenic mice allowed us to conduct studies designed to investigate and confirm the binding of predicted epitopes to this class I allele *in vitro* and, ultimately to test their immunogenicity *in vivo,* respectively.

Computational identification of CD4 T – cell, and B – cell epitopes for Lassa Fever Virus

The optimal protein- or peptide-based vaccine should evoke an immune response that would mimic the generation of protective immunity that occurs during natural infection. During Lassa Fever Virus (LFV) infection, both neutralizing antibody titers and high levels of cytokines associated with cellular immunity are detected, with cellular responses playing critical role in virus neutralization. Thus, a major objective of the studies described in this thesis was to identify LFV peptides capable of serving as immunogens to elicit both humoral and cellular immune responses.

Immunogenic epitopes expressed by pathogens that activate CD4 T cells are typically identified using T cell clones. Such clones are derived from peripheral blood mononuclear cells (PBMC) of previously infected individuals. The PBMC are stimulated with pools of overlapping peptides expressed by the pathogen and proliferative responses are studied. For example, in order to identify the immunogenic amine acid sequence

contained in a 500 amino acid protein, large numbers of overlapping peptides need to be synthesized. Computational algorithms have been designed to help investigators narrow down the number of candidate peptides. In contrast, B cell epitopes are commonly identified using sera from individuals (or animals) immunized with a pathogen-derived protein or peptides. Sera are then screened for antibodies that recognize these proteins or peptides. In our studies, in addition to the identification and experimental testing of CTL epitopes, we used computational algorithms to identify candidate immunogenic CD4 T cell epitopes and B-cell epitopes for LFV glycoprotein and nucleoprotein. We focused these studies on the identification of HLA-DR restricted peptides. While these studies aimed to identify these candidate epitopes, they can subsequently be experimentally tested for their binding to the class II molecule and for their ability to prime CD4 T cells, in order to generate specific immunological memory. As discussed in the Results section of this thesis, we also identified candidate immunogenic sequences that can be tested for binding to human sera from previously infected individuals and evaluated as candidate B-cell epitopes.

CD4 T-Cell and B-Cell Epitopes for LFV

Ideally, protein- or peptide-based vaccines should contain both T-cell epitopes, that will induce T helper and CTL responses, and B-cell epitopes that are important for the generation of protective antibody responses. Depending on the nature of the particular immune response to a given pathogen for which the vaccine is developed, the overall components of the vaccine must be designed to generate protective immunity to that pathogen. In the case of LFV, the CD8 T-cell population plays the most prominent role in

neutralization of the virus during infection. Since the T_H1 CD4 T-cell population plays a key role in cytotoxic T cell activation by the cytokines they release, and can also display cytotoxic activity themselves, these cells are critical for cellular immunity in general. It would therefore be expected that CD4 T cells play a role in protective immunity against LFV. Evidence in support of the important protective role of CD4 T-cells was first described by La Posta *et. al.* (1993). They demonstrated that vaccination of C3H/HeJ (H-*2k*) mice with recombinant vaccinia virus expressing the LFV envelope glycoprotein precursor induced cross-protective immunity mediated by CD4-positive CTLs. Specifically, they showed that these animals were protected from infection with closely related arenavirus -lymphocytic choriomeningitis virus (LCMV). Neither neutralizing antibodies nor CD8 CTLs specific for LCMV were present in these mice, yet they resisted a normally lethal LCMV challenge.

In another study, ter Meulen *et. al.* (2000) reported a strong memory CD4 T-cell response against the nucleoprotein of LFV in seropositive individuals from endemic areas. These individuals had not been reinfected with the virus for at least six years. In this study, peripheral blood mononuclear cells (PBMC) from the healthy individuals living in the Republic of Guinea, western Africa, were shown to proliferate in response to the nucleoprotein of the Josiah strain of LFV. It is noteworthy that this is the same strain for which CTL epitopes within a nucleoprotein and glycoprotein were identified in this thesis. The nucleoprotein was cloned, expressed and affinity-purified in *Escherichia coli.* T cell lines and the recombinant nucleoprotein-specific T cell clones were generated from PBMC, obtained from the individuals with high proliferative response profiles. T-cell epitopes were mapped for the generated T cell clones using overlapping synthetic

peptides derived from the nucleoprotein of LFV. They further demonstrated that these CD4 T cell clones proliferated and produced IFN-γ in response to a pool of LFV nucleoprotein peptides. They subsequently identified the following five LFV-specific CD4 T-cell epitopes: (1) NP176-188, (2) NP190-202, (3) NP288-300, (4) NP379-391, and (5) NP 498-510. The MHC-II alleles of the T cell clone donor were determined to be DRB1*0101 and -0301, DRB3*0301, DQA1*0101, and DQB1*02 and -0501. Because candidate LFV epitope vaccines should protect against different strains of the virus, the reactivity of the T cell clones with peptides derived from the homologous epitopes of the Nigeria and Mopeia strain was also tested. CD4 T-cell responses were found to be partly strain-specific and partly cross-reactive with other LFV strains.

The same group of investigators also addressed the issue of cross-reactivity with respect to the glycoprotein of Old (LFV -Josiah strain, Mopeia) and New World arenaviruses (North American: Bear Canyon and White-water Arroyo, and South American Junin, Machupo, Sabia and others) (ter Meulen *et. al.*, 2004). In order to study human T cell reactivity to the glycoprotein of LFV, T cell clones were generated from PBMC obtained from LFV antibody-positive individuals. T cell clones shown to proliferate in the presence of recombinant LFV glycoprotein were examined with regard to MHC restriction and T_H1/T_H2 phenotype. HLA-DR was shown to be the restricting MHC determinant for these cells that displayed cytokine profiles consistent with a T_H1-phenotype. Synthetic, overlapping peptide pools and then single peptides derived from a glycoprotein of LFV, were then used for mapping the T cell epitopes. Three LFV CD4 T cell epitopes were identified as follows: (1) GP282-294, (2) GP289-301, and (3) GP394-406. Each epitope was capable of inducing proliferative responses of PBMC from the T

cell clone donors. Using MHC class II -epitope prediction software, ProPred (discussed later), the investigators performed computational analysis of a LFV glycoprotein. They found a good correlation between experimentally observed and computationally predicted HLA-DR-restricted peptides. In the same study, one of the predicted T cell epitopes (GP289-301) was shown to be conserved in all strains of LFV and in LCMV -the Old World arenaviruses. The predicted epitope was also observed in several New World arenaviruses.

There is very little knowledge regarding the B-cell epitopes for LFV. Vladyko *et. al.* (1993) have reported that a nucleoprotein of LFV contains the following B-cell epitope regions: NP123-127, NP337-346 and NP518-527 and that a monoclonal antibody specific to LFV nucleoprotein reacts with NP123-127 synthetic peptide in solid phase ELISA.

B-Cell Epitope Mapping: Computer-Assisted Selection of Appropriate Antigenic Peptide Sequences

It has been shown that passive immunization with neutralizing antibodies is not a uniformly effective treatment for LFV infection in humans (Jahrling *et. al.*, 1985). Indeed, the role of neutralizing antibodies in protecting individuals from reinfection has not been fully elucidated. It is possible and, indeed likely, that the presence of neutralizing antibodies together with antigen-specific cytotoxic and helper T cells will confer more effective protection against primary infection than the presence of memory T cells alone. Therefore the identification of the antigenic peptides that contribute to the generation of neutralizing antibody is an important goal.

B-cell epitopes are defined as protein sequences to which the epitope-specific antibody will bind. In order for peptide-specific antibodies to recognize a protein containing the same sequence found in the peptide, the amino acids in the protein must be oriented to the antibody similar to that of the synthetic peptide. Thus, protein sequences need to be analyzed to ascertain that they meet the following criteria: (1) the stretch of the protein sequence to which the antibody will bind needs to be exposed to the solvent; and, (2) the sequence must be composed of amino acids without structures unrecognizable for the antibody, and the amino acids in the exposed sequence need to be hydrophilic containing polar or charged residues, particularly Arg and Lys (Grant 2003).

The general features of the protein structure that correspond to the criteria mentioned above are *turns* or *loop* structures. Such structural attributes are found on the surface of the secondary protein structure. The solvent-exposed areas of the secondary protein structure generally display relatively high levels of polar and charged residues that contribute to the hydrophilicity of those exposed stretches of proteins (Creighton, 1993).

Antigen Processing and Presentation by Class II MHC Molecules

MHC class II molecules present exogenous antigens to CD4-positive T-cells. Exogenous antigens typically derived from bacteria or viruses are taken up by antigen-presenting cells (APC) by endocytosis or phagocytosis. Intracellular vesicles carrying internalized proteins fuse with highly acidic endosomal and lysosomal vesicles, where the proteins get degraded. MHC class II molecules assemble in the endoplasmic reticulum (ER) as a nonameric complex consisting of an invariant chain trimer associated with three

class II αβ dimers (Cresswell 1996). The invariant chain blocks the binding of peptides and unfolded proteins in the endoplasmatic reticulum and during the transport of the class II molecule into acidified endocytic vesicles. Signals in the invariant chain cytoplasmic domain direct the complex into the endocytic pathway, where invariant chain degradation results in the transient formation of a class II αβ dimer with a residual fragment of the invariant chain, CLIP (class II-associated invariant chain peptides), in the peptide-binding groove (Riberdy *et. al.*, 1992). The interaction of the αβ-CLIP complex in the antigen-processing compartment, or MHC class II compartment (MIIC) with a second class II-like molecule, called HLA-DM in humans and H-2M in mice, induces CLIP dissociation (Denzin and Cresswell 1995). Association of empty αβ dimers with DM stabilizes them until high-affinity peptides derived from internalized proteins can bind. Mature αβ-peptide complexes then leave the endocytic pathway and are expressed on the cell surface.

LFV: A Potential Biowarefare Agent

Among the emerging infectious diseases, viral hemorrhagic fevers (VHF) represent a serious public health problem with recurrent outbreaks all over the world. In addition to infections occurring in endemic areas and those imported from endemic areas to Europe or the United States, a new concern has recently emerged due to development of bioweapons. The Center for Disease Control and Prevention (CDC) has grouped the potentially weaponizable agents into three categories, namely, A, B or C in accordance with their potential public health impact. Category A agents are those which possess the greatest potential for adverse public health impact due to expected mass causalities

associated with infections caused by these organisms. Agents in Category B have potential for large-scale dissemination with resultant illness but have lower medical and public health impact than agents category A. Finally, agents in Category C are not believed to present high bioterrorism risk to public health but could emerge as future threats as scientific understanding of these agents improves. Public health measures associated with biodefense policies include prophylactic vaccines against infectious disease agents, therapeutics consisting of antibiotics, and antibodies for passive protection.

There are four distinct families of viruses that cause VHF. These include the families *Arenaviridae, Buniyaviridae, Filoviridae* and *Flaviviridae.* Our laboratory is currently seeking to identify immunogenic CTL epitopes for two viruses that belong to *Filoviridae* family, namely, Ebola and Marburg, both of which are Category A viruses. The thesis research described in this report aimed to identify and evaluate CTL epitopes for LFV, a VHF that belongs to the *Arenaviridae* family. LFV is also categorized by CDC as a Category A agent (Table 1). It should be noted that another member of the *Arenaviridae* family of VHF is the well-studied and characterized virus called Lymphocytic Choriomeningitis Virus (LCMV). This importance of this virus to our understanding of how Lasa Fever Virus binds to and infects target cells will be discussed in more detail in subsequent sections of the thesis.

Table 1. Category A Priority Pathogens.
(http://www.niaid.nih.gov/biodefense/bandc_priority.htm)

- Bacillus anthracis (anthrax)
- Clostridium botulinum
- Yersinia pestis
- Variola major (smallpox) and other pox viruses
- Francisella tularensis (tularemia)
- Viral hemorrhagic fevers
    - Arenaviruses
        - LCM, Junin virus, Machupo virus, Guanarito virus
        - Lassa Fever
    - Bunyaviruses
        - Hantaviruses
        - Rift Valley Fever
    - Flaviruses
        - Dengue
    - Filoviruses
        - Ebola
        - Marburg

Lasa Fever Virus: Epidemiology, Concerns and World Health Needs

Lassa fever, caused by infection with Lasa Fever Virus is endemic in rural West Africa, Guinea, Sierra Leone, Liberia and in Nigeria where it was first described in the village of Lassa in 1969 (Monath *et. al.*, 1974). It has been estimated that LFV causes up to 300,000 infections/year and several thousand deaths due to hemorrhagic fever (McCormick *et. al.,* 1987) The fatality rate of hospitalized patients is about 17%, but in certain groups of patients, such as pregnant women, there is a > 30% mortality rate, and fetal and neonatal deaths approach 88% (Price *et. al.,* 1988; Fischer-Hoch *et. al.,* 2000). The principle vector for LFV transmission in humans is *Mastomys natalensis,* a ubiquitous and highly commensal rodent host. LFV infection of such rodents is persistent and asymptomatic. However, once infected with the virus, these animals exhibit profuse urinary excretion of LFV. Human infection is due to contact with rodents

or other LFV-infected people. The virus is transmitted from rodent to humans by contact with animal urine or feces or even blood that contains high virus titers. Alternatively, infection may be caused by human consumption of the rodent (ter Meulen *et. al.,* 1996). Transmission between humans has been reported as a result of exposure to blood, sexual contact, and breast feeding (McCormick *et. al.*, 1987a). The incubation period of Lassa fever may last up to three weeks. Therefore the virus may be imported into other regions of the world while infected individuals are asymptomatic or show early nonspecific signs of Lassa fever. So far, about 20 cases of imported Lassa fever have been reported worldwide (Drosten *et. al.*, 2003).

As noted above, LFV is a pathogen that is on the list of the potential weaponizable pathogens. In fact, it is known to have been weaponized (Hilleman, 2002). Hemorrhagic fevers are on the US military list of likely biowarefare agents for which prophylactic immunization or rapid diagnosis and treatment regimens can have a significant positive public health impact. A required infective aerosol dose of hemorrhagic fever complex, consisting of at least 50 highly virulent disease agents, which are members of 4 different families of HFV and which generally cause severe and lethal disease, is 1-10 microorganisms (Peters 2002). For comparison, the infective aerosol dose for smallpox virus or encephalitis virus is 10-100 microorganisms (Hileman, 2002). Among the pathogens that cause hemorrhagic fevers, there are only a few for which there is a vaccine under development as Investigative New Drugs (INDs). These include Rift valley fever, Argentinean and Bolivian hemorrhagic fever. Currently, the only effective therapy available for treatment of viruses causing hemorrhagic fevers is

Ribavirin. However, Ribavirin is only effective if administered within the first week of illness. Currently, there is no effective vaccine available for LFV (Hileman, 2002).

As previously mentioned, a rapid diagnosis of LFV infection can have a significant impact on the outcome in the situation where the bioweapons would be used or even in the situation when the virus is imported by a single infected individual from endemic areas. Thus, it is important to develop novel diagnostic methods for detection of LFV infection as well as methods that could allow for better monitoring and facilitate the study of immune responses in infected individuals. Currently, laboratory diagnosis for LFV infection is based upon three methods: RT-PCR, virus isolation in cell culture, and detection of specific IgM and IgG (Drosten *et. al.*, 2003)

LFV: Morphology, Genome Organization and Structure

As noted above, LFV is a member of the *Arenaviridae* family, genus *Arenavirus*. *Arenaviridae* are divided phylogenetically, serologically and geographically into two major complexes: (1) the Old World complex (e.g. LFV, LCMV); and, (2) the New World complex (e.g. Junin virus and Machupo virus). With the exception to the prototype Arenavirus LCMV , all other species show a specific geographic occurrence which can be explained by the geographical distribution of the respective natural host or reservoir species (rodents). Arenaviruses share major morphological characteristics. In electron microscopic photomicrographs, the interior of virion shows a typical granular pattern (Dresten *et. al.*, 2003) LFV is an enveloped virus with glycoprotein spikes on its surface. The viral genome consists of two single stranded RNAs, as indicated by a small (S) and large (L) RNA fragment, sizes 3.4 kb and 7 kb, respectively (Figure 7). Each segment (S

and L) contains two nonoverlapping genes arranged in ambisense orientation (in one region the RNA is negative sense and in another the RNA is a positive sense). The S RNA encodes the viral glycoprotein precursor protein (GPC) at the 5' end of the S RNA, and a nucleoprotein (NP), at the 3' end of the S RNA (Auperin *et. al.*, 1989). The L RNA encodes the viral polymerase gene at the 3' end of L RNA and a small, zinc-binding (Z) protein at 5' end of the L RNA (Djavani *et. al.*, 1997). The complete LFV L segment is 7279 nucleotides long and contains 60% AU and 40% GC (Genebank Accession # NC 004297).The L RNA exhibits an ambisense coding strategy. The S RNA is 3402 nucleotides long. The nucleotide base composition is 60% AU and 40% GC (Genebank Accession # NC 004296). The LFV glycoprotein is synthesized as a 76-kDa glycosylated precursor protein (GPC), which is posttranslationally cleaved into the N-terminal 44-kDa subunit (GP-1) and the membrane bound 36-kDa subunit (GP-2). The glycoprotein, GP-1 interacts with the cellular receptor α-dystroglycan, whereas GP-2 mediates pH-dependent fusion of the viral envelope with the cellular target membrane (Cao, *et. al.*, 1998). The posttranslational cleavage that triggers fusion activity is thus essential for virus entry into the host cell.

Antigenic differences have been detected between Nigerian LFV isolates and isolates originating in Sierra Leone, Liberia and Guinea (Jahrling *et. al.,* 1985) suggesting genetic diversity among these viruses. Phylogenetic analyses have shown four lineages for LFV (differing by approximately 12% in NP amino acids), three of which are found in Nigeria, and the fourth in Guinea, Liberia and Sierra Leone (Bowen *et. al.,* 2000).

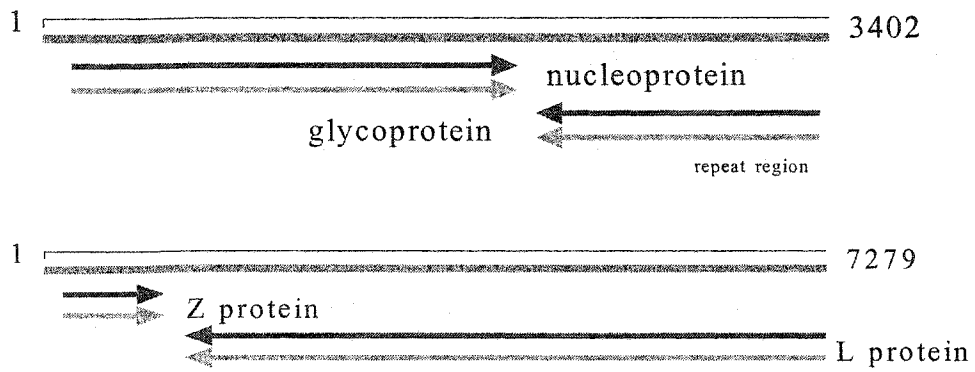


Figure 7. LFV segment S, complete sequence (upper) LFV segment L, complete sequence (lower) (http://www.ncbi.nlm.nih.gov/genomes).

LFV: Receptor and Life Cycle

Cellular receptors are key elements in determining the tropism and pathogenesis of virus infection. LFV and LCMV share the same receptor protein that allows for the attachment of the virus to the cell and subsequent infection. Infection is initiated when GP-1 anchors the virus to the cell surface through a proteinaceous receptor, found to be the α-dystroglycan (α-DG) (Cao *et. al.*, 1998). α-DG is an extracellular peripheral membrane protein that binds to the extracellular matrix and is noncovalently associated with β-DG, which is a transmembrane protein linked to the cytoskeleton (Henry *et. al.* 1996).

Replication and transcription of the genome takes place in the cytoplasm of an infected cell. During genome replication, full-length copies of genomic S and L RNAs are synthesized yielding the corresponding anti-genomic S and L RNAs. Expression of proteins requires transcription of mRNA from both the genomic and antigenomic RNA

within a ribonucleoprotein complex (Fields 1996). Glycosylated precursor protein (GPC) is directed posttranslationally into the endoplasmic reticulum. Cleavage of GPC is required for incorporation of glycoproteins into the virion envelope and thus for release of infectious LFV. The glycoproteins are transported to the cell membrane where budding and release of the virus takes place (Drosten *et. al.,* 2003).

LFV: Virulence Factors

The viral envelope and capsid proteins, core, matrix and nonstructural proteins, as well as noncoding regions of viral genome can influence LFV pathogenesis and virulence. The envelope glycoproteins and capsid surface proteins frequently function as viral cell recognition proteins. Changes in these proteins can alter viral tropism and pathogenicity both *in vitro* and *in vivo* (Fields *et. al.*, 1996). As mentioned earlier, GP-1 of LCMV and LFV interact with the cellular receptor α-dystroglycan, whereas GP-2 mediates pH-dependent fusion of the viral envelope with the cellular target membrane (Cao, *et. al.*, 1998). The spike glycoprotein complex in these virions is responsible for the initial interactions with the cell surface, and it has been shown that antibodies directed against GP–1 can neutralize virus infectivity or even prevent virus binding to cells (Burns *et. al.*, 1991; Parekh *et. al.*, 1986). It has been shown that mutations in the LCMV glycoprotein affect an early step in infection, possibly viral attachment and penetration (Fields *et. al.*, 1996). Thus, the glycoprotein of the LFV encoded by the S segment of the viral genome is believed to be an essential determinant of virulence and pathogenesis. The nucleoprotein (NP) encoded by the S RNA segment of LFV is associated with genomic RNAs both within infected cells and virons in the form of ribonucleoprotein

complexes (RNP). It is probable that the intracellular content of NP is influential in regulating the relative levels of transcription and replication within the infected cell. A phosphorylated derivative of NP has been detected in acute infection and at greater abundance in extracts of persistently infected cells (Bruns *et. al.*, 1986). This suggests that phosphorylation of NP may be involved in the attenuation of viral gene expression that marks the progression from acute to persistent infection.

A comparison of virulent and avirulent strains of LCMV using a guinea pig animal model suggests that the L RNA segment, which encodes the arenavirus polymerase, is an important determinant of LCMV virulence and, by extension, of Lassa virulence. Mutations in the L segment are involved in the generation of persistent infection in adult mice (Ahmed *et. al.*, 1998). According to a hypothesis by Lukasievich *et. al.* (2003), the virus must first replicate to a certain level before tropism and other factors begin to affect virulence. Thus, the replication level is likely to be the primary determinant of virulence. In addition, the L segment of the viral genome that carries the genes determining the level of virus replication is likely to play a significant role in virulence and pathogenesis.

LFV Infection: Clinical Manifestations and Pathogenesis

Although Lassa fever is classified as a hemorrhagic fever, clinical diagnosis is difficult because obvious bleeding is often absent even late in the course of illness (McCormick *et. al.*, 1987b). The virus enters the human body through the bloodstream, lymph vessels, respiratory tract, and/or digestive tract. It then multiplies in cells of the reticuloendothelial system. Virus replication in the reticuloendothelial cells causes

capillary lesions. These capillary lesions lead to erythrocyte and platelet loss, with mild to moderate thrombocytopenia and a tendency toward bleeding. Capillary lesions also cause increased vascular permeability and hemorrhage in various organs, such as the stomach, small intestine, kidneys, lungs, and brain. Lassa-infected individuals exhibit a wide range of non-specific clinical symptoms such as high fever, headache, sore throat and diarrhea. Differential diagnosis includes malaria, Ebola, Marburg, and Rift valley fever. Deaths caused by LFV infection are typically associated with multi-organ failure with or without hemorrhage and typically occur 14-16 days following infection. Neurological complications such as confusion, tremor, convulsion, and coma are frequent in critically ill patients who often die after the onset of these symptoms (Solbrig and McCormick 1991). Sensorineural deafness is a neurogical complication of the reconvalescence phase (Cummings *et. al.*, 1990). Decreased mortality is seen in patients when Ribavirin is administered within 6 days after onset of fever.

Two imported cases of Lassa fever were monitored in Germany in 2000 and a brief summary of these cases is provided here to provide some insight into a typical clinical course of infection (Schmitz *et. al.*, 2002). Serum samples were analyzed for virus load and cytokine levels. Both patients had non-specific clinical symptoms including high fever. Patient 1 developed multi-organ failure and died of hemorrhagic shock on day 15 of illness, while patient 2 died of respiratory failure due to aspiration without hemorrhage on day 16. Ribavirin was administered to both patients beginning on day 11 due to late diagnosis. Aspartate aminotransferase (AST) and alanine aminotransferase (ALT) levels, markers for hepatic necrosis, were elevated in both patients. A concomitant elevation of serumcreatine kinase (CK), an enzyme abundant in

muscle tissue, and serum lactate dehydrogenase (LDH) indicated rhabdomyolysis as the main cause of enzyme elevations. RT-PCR was positive for LFV in both patients. Neither IgM nor IgG specific for LFV were detected in the first of the two patients monitored. In this patient viremia increased steadily, while levels of proinflamatory cytokines IFN-γ and TNF-α, increased significantly, reaching extremely high levels shortly before death. The second patient had detectable levels of IgM antibodies on day 13 of the illness and both IgM and IgG antibodies to the Josiah strain were present on day 16. Virus load in the second patient decreased during the late stage of illness, which was paralleled by a decrease in the IFN-γ and TNF-α levels (Schmitz *et. al.*, 2002).

LFV: Immunopahogenesis and Immune Responses

Observations of humans and mice with genetic or acquired B and/or T cell deficiencies as well as experimentally manipulated animals with such immunodeficiencies clearly show that control of most acute viral infections is dependent upon the functional integrity of MHC class I-restricted CD8 CTLs (reviewed by Whitton *et. al.*) However, as discussed earlier, immune clearance of persistent viral infection depends not only on CD8 CTLs, but also requires the participation of CD4 T cells (Matloubian *et. al.*, 1994; Kalams *et. al.*, 1998; Tishon *et. al.*, 1995)

The findings reported by Schmitz *et. al.* (2002) suggest that high virus load and high levels of proinflamatory cytokines in the late stage of Lassa fever play an important role in the pathogenesis of hemorrhage, multi-organ failure, and shock in LFV-infected individuals. Mahanty *et. al.* (2001) investigated the role of cytokines in the pathogenesis of Lassa fever. The levels of pro- and anti-inflamatory cytokines and chemokines in

serum samples collected from hospitalized patients with fatal and nonfatal acute Lassa fever were compared with those from two control groups: (a) patients with other febrile illness; and, (b) uninfected individuals. Serum IL-8 and IFN-γ inducible protein-10 (IP-10) levels were significantly higher in patients with acute nonfatal Lassa fever than in control subjects. Levels of these cytokines were low or undetectable in patients with fatal Lassa fever. They concluded that low levels of IL-8 and IP-10 are associated with a fatal outcome. IL-8 is a chemokine that is secreted by many cell types including monocytes, neutrophils, and endothelial cells whose functions include chemoattraction for neutrophils and T cells at inflammatory sites. It also induces neutrophil degranulation and T-cell activation in Dengue virus infection (Juffrie *et. al.,* 2000). IP-10 is also a chemokine that is secreted by macrophages, endothelial cells, and T cells in response to a number of activating stimuli such as IFN-α, IFN-β, IFN-γ, TNF-α, and lipopolisaccharide (Kopydlowski *et. al.*, 1999). IP-10 is thought to participate in the pathogenesis of infections and shock by inhibition of endothelial cell function and chemoattraction of resting T cells and NK cells.

A study conducted using macaques confirmed the critical role of T cell responses in controlling virus replication and preventing the cascade of events that lead to death (Fisher-Hoch *et. al.*, 2000). Immunization of these nonhuman primates with LFV glycoprotein epitopes expressed in vaccinia virus resulted in cell-mediated protection against lethal challenge with the virus. The primary role of cytotoxic T lymphocytes in protection from Lassa fever was supported by low antibody responses observed in this study. The lack of response to treatment with high-titer human immune plasma (i.e. passive antibodies to LFV) in Lassa fever-infected humans as well as the minor role of

the detectable antibody titers confirms a minor role of humoral responses in this infection (McCormick *et. al.*,1986).

As discussed earlier, proinflamatory mediators play a critical role in the outcome of LFV infection (Mahanty *et. al.*, 2001). The lack of cellular responses in VHF infections, thus the absence of proinflammatory mediators, is thought to be caused by depletion of cytokine and chemokine-secreting cells by apoptosis. High levels of TNF-α were found to be responsible for apoptosis (Slowik *et. al.*, 1997). In addition, high levels of TNF-α have also been shown to contribute to trombocytopenia in these infections. In fact, in one study, fatal bleeding was accompanied by a sharp rise in the concentration of TNF-α (Schmitz *et. al.*, 2002). In patients with Argentine hemorrhagic fever, which is also caused by an Arenavirus, increased levels of TNF-α were directly related to severity of illness (Heller *et. al.*, 1992). Apoptosis of T cells has been reported to correlate with fatal outcome in Ebola hemorrhagic fever, while T cell responses with limited apoptosis correlate with survival and outcome is determined very early in the disease (Baize *et. al.*, 1999). Finally, lymphopenia and impaired lymphocyte proliferative responses are characteristic of severe Lassa fever. It may be that in Lassa fever, apoptosis is also a central feature of the failure to control virus replication and hence fatal outcome. Thus, severe illness caused by LFV can be a manifestation of aberrant host responses (e.g. overproduction of TNF-α) rather than lytic viral destruction of cells and, ultimately, organ failure (Fisher-Hoch *et. al.* 1988).

Prospective Protective LFV Epitopes

In light of the preceeding discussion, it is clear that CTLs, thus cell-mediated responses, play a primary role in the neutralization of LFV during infection. Hence, this study focused on CTL epitope mapping to identify and evaluate potential immunogenic epitopes that might be capable of generating protective immune responses. The processed viral peptides that will most likely be presented on HLA class I molecules to prime naïve cytotoxic T cells would ideally be derived from antigenic viral proteins known to stimulate protective immune responses. In the case of LFV, the G1 and G2 and the nucleoprotein encoded by the S RNA segment are the major antigens (virulence factors) thought to trigger the immune responses that contribute to viral clearance. Thus, in this study, the protein sequences of these two antigens were examined in an attempt to identify amino acid sequences (nanomers) with the potential to serve as CTL motifs. Computationally mapped peptides that are bonafide CTL epitopes would be expected to not only bind the selected HLA class I molecule *in vitro* and *in vivo* but also to be capable of generating CTL responses by virtue of their ability to activate T cells expressing epitope-specific TCRs. When these two requirements are fulfilled, the computationally predicted nanomer peptide is considered a CTL epitope.

The CD8 T cell response to LCMV infection have been shown to be specific for epitopes derived from the viral glycoprotein (GP) and nucleoprotein (NP) (reviewed in Wong and Pamer 2003). In C57BL/6 mice, three H-2D^{b}-restricted epitopes, GP33 (GP 33 -41), GP276 (GP276-286), and NP396 (NP396-404) are the major targets of CTL responses while in BALB/c mice, the single immunodominant epitope is NP118 (NP118-126) presented by H-2L^{d} which is discussed next in this paragraph.

Sequence comparisons have demonstrated that Old World and New World Arenaviruses share several amino acids with the nucleoprotein that consist of amino acids (NP 118-126) of LCMV that comprise the immunodominant cytotoxic T-lymphocyte (CTL) epitope for H-2^d mice (Whitton *et. al.*, 1989). This L^d restricted epitope was found to constitute >97% of the total clonal or total bulk primary CTL responses (Whitton *et. al.*, 1989). A single injection of recombinant vaccinia virus incorporating this epitope provided complete protection for H-2^d BALB/c mice later challenged with a lethal dose of LCMV. Further study of this CTL epitope indicated that one third of mice from several established haplotypes (H-2^d, H-2^u, and H-2^q) possessed MHC class I molecules capable of presenting the LCMV NP 118-126 peptide for recognition and lysis by virus-specific CTLs (Oldstone *et. al.,* 1992 and 1993). Oldstone *et. al.*, (2001) also characterized common antiviral cytotoxic T-lymphocyte epitopes for diverse Arenaviruses. In this study, three Old World Arenaviruses LCMV, LFV (LFV), and Mopeia virus, but not the new world virus Sabia virus, showed functional CTL cross-reactivity when their peptides were used to coat target cells. The use of site-specific amino acid exchanges in the NP CTL epitope among these Arenaviruses identified amino acids involved in MHC binding and CTL recognition. The outcome of these studies indicates that either common (i.e. cross-reactive) or exclusive (only for one virus, in our case LFV) CTL epitopes restricted to human molecules like HLA-A2.1 can be identified.

Objectives

The studies carried out under this thesis project had several objectives summarized in the Aims below:

1) To identify promiscuous LFV-specific CTL, T-helper cell and B-cell epitopes using computational methods (prediction algorithms).

2) To evaluate the HLA A2.1-binding potential of mapped CTL epitopes *in vitro*. The mapped CTL epitopes predicted under Aim 1 were tested for MHC class I binding potential using flow cytometric analysis. TAP-negative, HLA A2.1-positive T cells were used to evaluate peptide binding by cytometrically measuring their ability to stabilize expression of this MHC class I allele. CTL epitopes were categorized as high, moderate and low binders, depending on the degree to which they stabilize MHC class I.

3) To evaluate mapped peptides *in vivo* for their ability to induce CTL responses using HLA transgenic mice. Synthetic peptides identified as high MHC class I binders under Aim 2 were used to immunize C57B/L transgenic mice expressing HLA A2.1. CTL responses were assessed by functional analysis using CTL assays, cytokine ELISPOT assays, cytokine intracellular staining and by flow cytometric analysis using fluorescent MHC/class I peptide complexes (i.e., tetramer staining).

Collectively, in our studies we have exploited the computational approach of epitope mapping in order to identify the immunodominant CTL, T-helper cell and B-cell epitopes for LFV. We have identified two peptides derived from the glycoprotein of the virus that have shown to induce HLA-A2 restricted CTL-mediated immunological

responses *in vivo*. Additionally, we computationally identified candidate T-helper cell and B-cell epitopes.

MATERIALS AND METHODS

Computational Mapping of CTL Epitopes

The protein sequences of LFV (Josiah strain, Genebank Accesion # J04324) nucleoprotein (NP) and glycprotein (GP) encoded by the S segment of the viral genome were obtained from the Entrez database. The 569 amino acid NP and 491 amino acid GP sequences were separately analyzed *in silica* using the BIMAS prediction server (http://bimas.dcrt.nih.gov/molbio/hla_bind). BIMAS is an algorithm which ranks potential 8-mer, 9-mer, or 10-mer peptides based on a predicted half-life of dissociation to a particular HLA class I molecule (Parker *et. al.*, 1994). The higher the score for a particular peptide, the higher the probability of its binding to the HLA class I molecule. The total protein sequence (NP or GP) was parsed into peptides, nine amino acids in length, each overlapping by eight amino acids. HLA-A2.1 (also referred to as HLA*0201) was chosen because of the widespread frequency of that allele in different ethnic populations within the United States (Cao *et. al.*, 2001). The computational scoring of the MHC class I binding peptides is the first step in determining whether a given region of the protein is likely to be recognized by primed T cells and thus defined as a CTL epitope. Binding of predicted peptides to MHC class I was confirmed using *in vitro* binding assays. The immunogenicity of the peptides was confirmed by testing the ability of CTLs to recognize the peptide in HLA A2.1-restricted cytotoxicity assays. Based on *in silica* analysis for LFV, viral GP and NP peptides predicted to be high binders were selected for subsequent analysis. A dissociation constant (BIMAS score) of ≥50 was used as the initial selection criterion. Selected peptides were then screened for similarity to the

human genome using the NIH Blast server (http://www.ncbi.nlm.nih.gov/blast/). Peptides showing homology with regard to the human proteome were eliminated for further study due to their potential to induce autoreactive responses when such self-peptides would be presented on MHC class I molecules.

In addition to the BIMAS algorithm, alternative prediction algorithms were used for identification of the promiscuous CTL epitopes. The alternative prediction algorithms were mainly used for the comparison purposes as a guide for choosing and ranking of peptides. These include: (1) Matrix-based which includes SYFPEITHI, ProPredI, RANKPEP and MAPPP; (2) Artificial Neural Network- based including nHLAPred; (3) Support Vector Machine-based including SVMHC; and (4) Structure- based, namely PREDEP (the URLs for these algorithms are included in the Reference list). Due to the voluminous nature of the data derived from these CTL epitope prediction algorithms, only the date form BIMAS is included in the thesis.

Computational Mapping of T-Helper Cell Epitopes

The protein sequences of LFV NP and GP were analyzed computationally to identify candidate T-helper cell epitopes. This was accomplished using the ProPred and SYFPEITHI algorithms to examine the 569 amino acid NP and the 491 amino acid GP sequences. Each sequence was analyzed against all available HLA-DR alleles simultaneously using the ProPred algorithm. Thus, the output results obtained from this algorithm allow for the comparison of predicted binders and their cross-reactivity against 51 available HLA-DR alleles. Predicted binders were displayed as blue-colored regions, with P1 identified as the starting residue of each predicted binding frame colored in red (Table 8 and 11). In additional analyses, the twenty highest scoring HLA-DRB1*0101

restricted peptides were ranked by the ProPred algorithm and compared to the twenty highest scoring DRB1*0101 restricted peptides ranked by the SYFPEITHI algorithm (Table 9 and 12).

The threshold value for the ProPred analysis was set to 3%. This threshold correlates with the peptide score and therefore with HLA-ligand interaction (Sturnilo *et. al.*, 1999). The peptide score was calculated after the multiplication of the partial values assigned to particular amino acids (see the Table 7). The threshold is defined as the percentage of the best scoring peptides. The lower the threshold (high stringency), the lower the false positive rate. For example a threshold of 1% would predict peptides in any given protein sequence that belongs to the 1% best scoring peptides. Such a high-stringency threshold (1%) results in fewer selected "positive" peptides, and might miss some potential binders. A comprehensive evaluation of these scoring results concluded that the most practical and optimal threshold level is 3% (Bian *et. al.*, 2003).

As with the analysis of candidate CTL epitopes, the LFV NP and GP sequences identified, as candidate T-helper epitopes were first screened for their similarity to the human genome using the NIH Blast server (http://www.ncbi.nlm.nih.gov/blast/). Peptides showing homology with regard to the human proteome were eliminated as possible candidates for T-cell epitopes.

Computational Mapping of B-cell Epitopes

The method employed for B-cell epitope mapping utilized the correlation between the hydrophilic character of peptide sequence and its propensity to form β-turn structures (Grant, 2003). In order to compute the hydropathy index and the tendency for β-turns, sequences of LFV-derived NP and GP in FASTA format were separately pasted in the

Kyte-Doolittle hydropathy and Chou-Fasman β-turn algorithms. The results were generated in both graphical and numerical form. A window size of 9 was used and equal weight was given to each amino acid. The window size determines the number of amino acids to be used in computing a value for the amino acid at the center of the window. For example, a window size of 9 includes 4 amino acids on each side of the central amino acid. The value computed for the central amino acid is the average of the values for each amino acid in the window.

Results of the two computational analyses (hydropathy and β-turn) were compared and the 3 areas of both NP and GP sequences that were high in turn tendency and high in hydrophilicity (low in hydrophobicity) were highlighted (Figure 9 and 11). These were chosen as antigenic stretches of the whole protein, and therefore, possible candidates for B-cell epitopes. The highlighted areas shown in Figures 10 and 12 correspond to positive peaks in Chou-Fasman analysis and negative peaks in the Kyte-Doolittle analysis.

As noted above, the immunogenic peptides identified as candidates were then screened for their similarity to the human genome using the NIH Blast server. Peptides showing homology with regard to the human proteome were eliminated as possible candidates.

In Vitro Evaluation of Binding Potential of Mapped CTL Epitopes

It is well documented that the immunogenic potential of peptides restricted by MHC class I molecules correlates with their ability to bind to and stabilize these MHC determinants (Sette *et. al.*, 1994). We have therefore used a binding/stabilization assay to

measure the interactions between peptides identified *in silica* as high binders and the HLA-A2.1 molecule. Peptides were tested for their ability to bind HLA A2.1 molecules *in vitro* using an MHC stabilization assay described by Nijman *et. al.* (1993). The peptides tested included 9-mers found to have BIMAS scores > 50 when analyzed for with respect to HLA A2.1 binding *in silica*. All peptides tested were custom synthesized (>95% purity) by Sigma-Genosys (The Woodlands, TX).

The TAP-defective T2 human lymphoma cell line expressing HLA-A2 was used as a standard cell line for these assays (Anderson *et. al.,* 1993). The T2 cell line *et. al.*has defects in peptide loading of class I, resulting in reduced cell surface expression of class I molecules due to a mutation or deletions in TAP genes. TAP proteins transport cytosolically-derived peptides into the ER or *cis*-Golgi, where the peptides can bind nascent class I molecules (Kleijmeer *et. al.*, 1992). This TAP-defective cell line exhibits the phenotype of low endogenous class I surface expression, and inability to process and present antigen to T cells (Townsend *et. al.*, 1990).

From studies using the fungal antibiotic drug, brefeldin A, which blocks the movement of proteins out of the ER, it has been shown that viral-derived peptides bind class I molecules in the endoplasmic reticulum (Nuchtern *et. al.*, 1989) In studies where the cells with peptide transport deficiencies were employed (e.g. T2 cells) the binding of the exogenous (exogenously added) peptides to live cells was shown to occur by their specific binding to surface class I molecules (Smith *et. al.*, 1992). One of the mechanisms by which exogenous peptides bind to MHC class I molecules expressed on live cells is by peptide exchange by class I molecules at the cell surface. Another mechanism by which

exogenous peptides can gain access to class I molecules is by retrograde transport of peptide from outside the cell into the ER lumen (Day *et. al.et. al.,* 1997).

MHC Stabilization Assay

The MHC stabilization assay relies on the ability of exogenously added peptides to increase the number of properly folded HLA-A2.1 molecules on the cell surface of TAP-deficient cells such as T2 cells (Nijman *et. al.et. al.,* 1993). The exogenously added peptides stabilize the class I/β_2m structure on these cells and this can be shown by immunofluorescence staining using a fluorochrome-conjugated anti-HLA-A2.1 antibody.

T2 cells were grown in IMDM medium containing L-glutamine, 25 mM HEPES buffer, 3,024 mg/L sodium bicarbonate and supplemented with 20% fetal bovine serum (FBS) (see Reagents and Solutions section for medium and buffers) overnight (18 hours) at 37°C. Cells were harvested and washed twice with AIM-V medium (Invitrogen, Grand Island, NY) containing FBS, L-glutamine, streptomycin sulfate at 50μg/ml and gentamycin sulfate at 50μg/ml. Cells (5 x 10^5/250μl) were then incubated with 50μM of single peptides overnight at 37°C in 2.5 ml borosilicate tubes with AIM-V medium. Following 18 hours incubation, the cells were washed twice with cold PBS (Invitrogen, Grand Island, NY) containing 3% FBS and incubated for 30 min at 4°C with fluorochrome-conjugated anti-HLA A2.1 antibody (Pharmingen, San Diego, CA) at the concentration of 500ng/5 x 10^5 cells. Following three washes with staining buffer (PBS + 3% FBS), HLA-A2.1 expression was measured by flow cytometry (Becton Dickinson LSR2 Biosciences, Mountain View, CA). The data was analyzed using FlowJo Software 4.5.8 (TreeStar, Ashland, OR).

The criterion used to identify a given peptide as a "high binder" was that it be capable of causing an increase in MFI values grater than 150% over no peptide control. High affinity binding peptides were selected for further study. The peptide SLYNTVATL, taken from the HIV molecular immunology database (http://hiv-web.lanl.gov/content/immunology/index.html/) or peptide GILGFVFTL derived from influenza matrix antigen (Micheletti *et. al.*, 2002) are both known to be high binding HLA-A2.1 epitopes and were used as positive controls (data shown only for the GILGFVFTL peptide). The peptide IAGNSAYEY, a known class I non-binder taken from the MHCBN database (http://www.imtech.res.in/raghava/mhcbn/) was used as a negative control.

In Vivo Evaluation of Mapped Epitopes: Induction of CTL Responses

The immunogenic potential of peptides and, thus their ability to elicit effector CTL responses, ultimately depends upon specific structural properties of these peptides including the presence of appropriate terminal MHC class I anchoring amino acids. The ability of peptides to bind to and stabilize MHC class I molecules correlates with the immunogenic potential of a given peptide (Minew *et. al.*, 2000). Thus, high affinity binding peptides identified in this study may also be immunogenic *in vivo.* To test this possibility, HLA-A2.1 transgenic mice (see below) were immunized with peptides with high affinity binding properties and responses generated were assessed using CTL assays, tetramer staining, IFN-γ ELISPOT and intracellular IFN-γ staining assay all of which are described below. In all experiments, three mice per peptide-stimulated group (or unstimulated controls) were employed. Experiments were repeated at least three times.

Experimental Animals

C57BL/6J transgenic mice expressing human MHC class I allele HLA-A2.1 were used to evaluate the ability of the synthetic peptides (candidate LFV CTL epitopes) identified *in vitro* as high affinity MHC class I binders to stimulate CTL responses *in vivo*. Six-to eight week-old female mice [strain name: C57BL/6-TgN (HLA-A2.1) 1Enge] were purchased from The Jackson Laboratory (Bar Harbor, ME). The choice of this transgenic mice model was based upon expression of the appropriate human HLA class I allele (HLA-A2.1) and their commercial availability. These transgenic mice were constructed by microinjection of the 7-kb EcoRI fragment containing the full length HLA-A2.1 gene into fertilized eggs from C57/BL/6 mice (Le *et. al.*, 1989).

Generation of CTLs in HLA-A2.1Transgenic Mice

Efficient induction of specific CTL responses by peptide immunization in incomplete Freund's adjuvant (IFA) has been described by several laboratories (Aichele *et. al.*, 1990; Gao *et. al.*, 1991; Okazaki *et. al.*, 2003). Eight week-old female HLA-A2.1 transgenic mice were immunized subcutaneously with a 100μl emulsion containing 1:1 Seppic adjuvant (Montanide ISA 70 Seppic, France) and PBS/ 5% DMSO solution (for hydrobhobic peptides) or PBS/ 5% water (for hydrophilic peptides) solution with 100 μg/mouse of the single mapped CTL epitopes and 140 μg/mouse of hepatitis B virus core128-140 helper epitope. Seppic adjuvant is physically equivalent to IFA except that in addition to mineral oil, it also contains mannitol-derived surfactants and it generates a low viscosity emulsion when mixed with antigen thus minimizing the physical trauma

caused by injection. Moreover, diffusion of the immunizing agent is also improved relative to the use of IFA, such that a less local inflammatory response is seen. Seven days following immunization, the animals were similarly boosted. Splenocytes and lymph nodes were harvested one week after boosting. The harvested cells were used to generate CTLs *in vitro* and subsequently tested for cytotoxic activity in the CTL assay described below. The same immunization protocol was used for cytokine ELISPOT, intracellular IFN-γ staining and tetramer staining assays.

In Vitro CTL Stimulation

Single-cell suspensions were prepared by pooling splenocytes and draining lymph node cells from immunized HLA-A2.1 transgenic mice. Red blood cells were lysed with lysing buffer (see Reagents and Solutions section). Immune spleen cells were stimulated in 24-well plates with autologous spleen cells and in the presence of 10μM of antigenic peptides in complete RPMI (10% FBS, 100 U/ml penicillin, 100 μg/ml streptomycin, 10 mM HEPES, and 50μM 2-mercaptoethanol) and 10% T-Stim (BD Labware, Bedford, MA) at 37°C. Human T-stim without PHA is an enriched mixture containing IL-2 and other cytokines. It is commonly used for growth and maintenance of IL-2-dependent murine T-cells. The cells were restimulated with 10μM peptide and T-Stim every day for 5 days. Separate cultures were set up for each peptide identified *in vitro* as high affinity peptide. Following a 5-day incubation period, non-adherent CTL cells were harvested and used as effector cells in a non-radioactive CTL assay described below.

CTL Assay

A non-radioactive CTL assay was used to assess the cytolytic activity of generated effector cells. For this purpose, we used the colorimetric Cyto Tox 96® Non-radioactive Cytotoxicity assay (Promega, Madison, WI). The assay quantitatively measures lactate dehydrogenase (LDH), a stable cytosolic enzyme that is released upon cell lysis, in much the same way as ^{51}Cr is released in a radioactive assay. Released LDH in culture supernatant is measured with a 30-minute coupled enzymatic assay, which results in conversion of a tetrazolium salt into a red formazan product. The amount of color formed is proportional to the number of lysed cells. Visible wavelength absorbance data are collected using a standard 96 well plate reader (Korzeniewski *et. al.*, 1983). The Cyto Tox96® Non-radioactive Cytotoxicity Assay has been compared to the ^{51}Cr release assay and was shown to be equally sensitive yielding virtually identical results (Brander *et. al.,* 1993). Several other studies have confirmed the validity of the non-radioactive cytotoxicity assay (Hernandez *et. al.,* 2003; Mizumoto *et. al.,* 2002; Kim *et. al.*, 2000; Finke *et. al.*, 1998; Even *et. al.*, 1995).

Effector cells generated in peptide-primed HLA-A2.1 transgenic mice were tested for their cytolytic activity by measuring their ability to kill peptide-pulsed T2 target cells expressing the HLA-A2.1 molecule. Target T2 cells (5×10^5/ml) were incubated overnight with 10μM of appropriate peptides in IMDM medium containing 10% FBS in 12-well plates at 37°C. The T2 cells were washed and peptide-pulsed and unpulsed (control) T2 cells were then added at 1×10^4 cells per well in 50μl to 96-well round-bottom plates. Effector cells were added to these target cells to achieve ratios ranging from 40:1 to 5:1, effector: target cells, respectively. Fifty microliter of supernatants were

removed from each well following a 4-hour incubation period and transferred to new flat-bottom 96-well plates for colorimetric assay performed according to the Promega Cyto Tox kit. The absorbance was recorded with Bio-Rad ELISA reader Bio-Rad, (Hercules, CA). The percentage of specific LDH release (cell lysis) was calculated using the following formula:

$$\frac{100 \times [(\text{experimental} - \text{effector spontaneous}) - \text{target spontaneous}]}{[\text{target maximum release} - \text{target spontaneous}]}.$$

Spontaneous release was determined from target cells incubated without effector cells and vice versa and maximum release was determined by adding 10µl of lysis solution (0,8% Triton® X-100) provided by the kit per 100µl culture medium.

Tetramer Staining

Traditionally, identification of antigen-specific T cells has been accomplished using functional assays such as cytotoxicity assays and *in vitro* proliferation studies. T cells recognize, through their TCRs, specific peptide epitopes complexed with MHC molecules on the surface of target cells serving as APCs. The interaction between TCR and the peptide-MHC complex is specific but is characterized by low affinity and fast off-rates (Altman *et. al.*, 1996). To detect T cells with specific TCRs, multimers of soluble peptide-MHC complexes have been constructed to increase the avidity of binding of such complexes to the specific TCR. Among the different strategies for constructing multimers, tetramers based on streptavidin-biotin (see below) are currently in use in

many laboratories. Tetramers can be made using either class I or class II MHC molecules, for detection of epitope-specific CD8 and CD4 positive T cells, respectively. MHC class I tetramers were the first to be successfully made and used for quantitating CTL effector cells (Altman *et. al.*, 1996). A biotinylatation signal sequence was engineered onto the C terminus of a recombinant MHC molecule expressed in *E. coli.* The recombinant proteins for the heavy and light chains of class I molecules were refolded with the epitope peptides to form peptide-MHC complex, followed by biotinylation with an enzyme. The biotinylated peptide-MHC complex is bound to avidin at a 4:1 ratio, resulting in the formation of tetrameric peptide-MHC complex, or tetramer. The avidin molecule is conjugated with a fluorochrome so that the specific T cells labeled with the tetramers can be readily detected and quantitated by flow cytometry (Figure 8) (Kita *et. al.*, 2003).

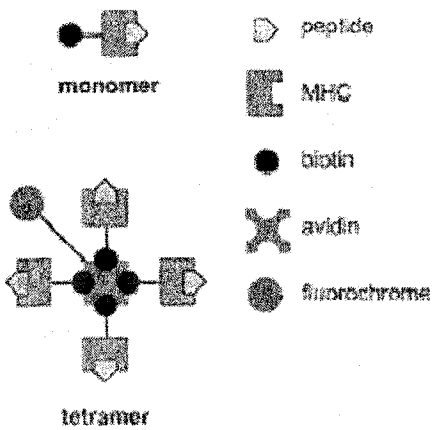


Figure 8. Construction of peptide-MHC tetramers. Recombinant MHC molecules are folded with peptide to form peptide-MHC complex, which is biotinylated with Bir A enzyme. The biotinylated peptide-MHC complex is mixed with avidin at ratio of 4:1 to form the tetramer (Kita *et. al.*, 2003).

Tetramers consist of monovalent class I molecules, which are produced in *E. coli* and refolded *in vitro* in the presence of a molar excess of antigenic peptide. To date, a large numbers of tetramers have been used to enumerate epitope-specific class I-restricted T cells in mice, humans and monkeys. Despite the success of this approach, several investigators have designed other strategies to produce multivalent peptide/MHC complexes (reviewed in Hugues *et. al.*, 2002). The first study describing the use of multivalent peptide/MHC molecules to quantitate antigen-specific T cells was published by Dal Porto *et. al.*, (1993). In this study a murine immunoglobulin was used as a molecular scaffold for MHC to produce dimeric MHC class I molecules. In short, a cDNA encoding the extracellular domains of a MHC class I was inserted at the 5' end of an immunoglobulin heavy chain IgG1 gene. The resulting plasmid was transfected into the murine plasmacytoma cell line. The expressed and purified MHC-Ig molecules were passively loaded with the purified synthetic peptides in the presence of molar excess of β2-microglobulin. MHC-IgG1 divalent molecules have successfully been used to stain CD8 T cells restricted by the human HLA-A2 (Greten *et. al.*, 1998) or the murine K^b (Dal Porto *et. al.*, 1993).

In another approach Choi *et. al.*, (2002) demonstrated that chimeric A2 tetramers containing mouse H-$2K^b$ (Kb) α3 domain ($A2K^b$ tetramers) can be used as staining reagents to monitor A2-restricted CTL responses in HLA-A2 transgenic mice. The increased ability of $A2K^b$ tetramers to stain mouse A2-restricted CTLs, as compared with A2 tetramers, correlated with their higher binding affinity for mouse A2-restricted CTLs. The researchers reasoned that the poor staining efficiency of A2 tetramers of splenocytes derived from A2 transgenic mice could be accounted for by lack of mouse CD8 binding

to human A2 molecules. It has been previously shown that CD8 binding to class I molecules is required to obtain high levels of staining with class I tetramers (Daniels and Jameson, 2000) which was supported by the finding that murine CD8 enhances affinity of TCR for class I molecules (Garcia *et. al.,* 1996).

In the present study, tetramer staining was performed as described by Singh *et. al.,* (2002). A tetramer specific for one of the selected LFV peptides (LLGTFTWTL) was prepared for us by the NIAID Tetramer Facility, Emory University, Atlanta (http://emory.edu/WHSC/TETRAMER/). Briefly, splenocytes and lymph node cells from immunized transgenic mice (the immunization protocol was described above) and adjuvant control, non-immunized mice were isolated. After lysing RBCs the lymphocytes were cultured overnight in compete RPMI medium with 10μM appropriate peptide (LLGTFTWTL) and T-Stim (BD Labware, Bedford, MA) at 37°C. Control cells (the lymphocytes from unimmunized mice) were cultured in T-Stim-containing medium alone. The cells were then harvested for tetramer staining. One million cells (in 100μl PBS containing 2% FBS) were stained with 100 ng of PE-conjugated MHC-I-peptide LLGTFTWTL-conjugated tetramer (the tetramer was titrated 1:100 to 1:1000) for 30 min. at room temperature in the dark, and subsequently with FITC-conjugated anti-CD3, and PerCP- conjugated anti-CD8 monoclonal antibodies (BD, Pharmingen, San Diego, CA) for 30 min. at 4°C in the dark. The cells were washed twice and were acquired with a FACS LSR II analyzer (Becton-Dickinson, San Diego, CA), the data was analyzed using FlowJo Software 4.5.8 (TreeStar, Ashland, OR).

ELISPOT Assay

The enzyme-linked immunospot (ELISPOT) assay is a powerful tool for detecting and enumerating individual cells that secrete a particular protein *in vitro*. In the present study, we useed an ELISPOT assay designed to measure IFN-γ production (Power *et. al.*, 1999). CTLs and T_H1 cells are known to secrete IFN-γ following their activation with antigenic epitopes. IFN-γ inhibits viral replication and induces MHC class I expression and macrophage activation *in vivo* (Fernandez *et. al.*, 1992). The ELISPOT assay takes advantage of the fact that T cells, in this case splenocytes from immunized mice, will not secrete effector cytokine IFN-γ until they encounter cognate antigen (antigenic peptide). Thus, when splenocytes from immunized mice are primed with the antigenic peptide *in vivo,* they will be activated when incubated with stimulator cells and antigenic peptide. Since IFN-γ production correlates with CTL responses, measuring IFN-γ secretion allowed us to assess the ability of mapped epitopes to stimulate CTL precursor cells.

ELISPOT assays were performed as described by Lycke and Coico (1996). The ELISPOT set (Beckton-Dickinson, San Diego, CA) containing ELISPOT plates, unlabeled capture antibody, biotynylated detection antibody, and enzyme conjugate was used. Wells of 96-well ELISPOT plates were incubated overnight at 4°C with 5 μg/ml of monoclonal anti-mouse IFN-γ antibody. Wells were then washed and blocked with RPMI medium supplemented with 10% FBS for 2 hr at room temperature. Splenocytes and the draining lymph node cells were prepared in complete RPMI, from immunized transgenic mice and 100μl/well of cell suspension at the concentration of 2×10^6 /ml together with the peptide concentrations ranging from 100 to 0.001μM (and without the peptide as a negative control and 5μg/ml Con A as positive control) were added to pre-coated plates.

The plates were incubated overnight at 37°C, in a 5% CO_2 humidified incubator. Following this incubation period, cell suspensions were aspirated and the wells were washed twice with deionized water and three times with wash buffer (1 x PBS containing 0.05% Tween-20). Biotynylated anti-mouse IFN-γ (detection antibody) was added to each well at the concentration of 2 μg/ml and the plated were incubated for 2 hr at room temperature. The plates were washed three times with wash buffer and enzyme conjugate, Streptavidin-HPR was added to each well and the plates were incubated for 1-hour at room temperature. The plates were washed and the 3-amino-9-ethylcarbazole (AEC) substrate solution with AEC chromogen (BD, Parmingen, San Diego, CA) was added. The reaction was stoped by addition of deionized water. ELISPOTS were enumerated using a dissecting microscope. Background (mean of wells without peptide) levels were subtracted from each well on the plate.

Intracellular IFN-γ staining

Another related method, which allows for quantitation of IFN-γ producing T-cells, involves intracellular cytokine staining (ICCS). In this method, cells containing cytosolic IFN-γ are measured by flow cytometry. The advantage of this method over the ELISPOT assay is that the flow cytometric analysis allows for direct determination of T_H1 versus T_H2 cells on the basis of staining with anti-CD8 and anti-CD4 monoclonal antibodies. Intracellular IFN-γ staining was performed in this study as described by Foster and Prussin (2002). The splenocytes and lymph node cells were polled and stimulated *in vitro* with 1 μM of appropriate peptides in complete IMDM medium containing 10% T-Stim (BD Labware, Bedford, MA) overnight at 37°C in a 5% CO_2

humidified incubator (20 ng/ml PMA and 1μM inomycin was added as a positive control and no peptide was added for the negative controls). The following day, 1μl/ml of monensin (Golgi Stop, BD, Pharmingen, San Diego, CA) was added to inhibit intracellular protein transport, and the cells were incubated for another 6 hr. The cells were then washed and stained with monoclonal α-CD8 PerCP (BD, Pharmingen, San Diego CA) in 1% FBS/PBS. The cells were then permeabilized with Cyofix/Cytoperm kit (BD, Biosciences, San Diego, CA), and stained with monoclonal α-IFN-γ (BD, Pharmingen, San Diego, CA). As controls, an isotype control antibody the FITC Rat IgG_1 (Pharmingen, San Diego, CA) was used to stain peptide stimulated cells. And monoclonal α-IFN-γ antibody was used to stain unstimulated cells. It is noteworthy that the Cytofix/Cytoperm buffer contains paraformaldehyde, thus the cells were fixed during the permeabilization step. The cells were acquired with a FACS LSR II analyzer (Becton-Dickinson, San Diego, CA), and the data was analyzed using FlowJo Software 4.5.8 (TreeStar, Ashland, OR) and IFN-γ-containing CD8 cells were enumerated.

Reagents and Solutions

Media

Complete IMDM (Iscove's modified Dulbecco's medium)

2Mm L-glutamine
25 mM HEPES buffer
3,024mg/L sodium bicarbonate
20% FBS
100 U/ml penicillin
100 ug/ml streptomycin sulfate
50 μm 2 meracptoethanol

Complete AIM-V medium

AlbuMAX® supplement (bovine serum albumine) substituted for human serum albumine
2 Mm L-glutamine
50μg/ml streptomycin sulfate
10μg/ml gentamycin sulfate

Complete RPMI 1640

2 mM L-glutamine
10% FBS
100 U/ml penicillin
100 ug/ml streptomycin sulfate
50μM 2-mercaptoethanol

Buffers

PBS

0.23 g NaH_2PO_4 (anhydrous; 1.9 mM)
1.15 g Na_2HPO_4 (anhydrous; 8.1 mM)
9.00 g NaCl (154 mM)
Add H_2O to 900 ml

Flow Cytometry Staining Buffer

PBS phosphate buffered saline + 3% FBS

ACK Lysing Buffer

8.29 g NH_4Cl (0.15 M)
1 g $KHCO_3$ (10.0 mM)
37.2 mg Na_2EDTA (0.1 mM)
Add 800 ml H_2O and adjust pH to 7.2-7.4 with 1 N HCl
Add H_2O to 1 liter

0,2 μm filter sterilized, RT stored

RESULTS

PART I. Computational identification of CTL epitopes, T-helper cell, and B-cell epitopes.

Computational algorithms were used to identify CTL, T-helper cell and B-cell epitopes contained in the glycoprotein and a nucleoprotein of LFV. The identification of CTL epitopes was done using BIMAS (Parker *et. al.*, 1994) and SYFPEITHI (Rammensee *et. al.*, 1999) algorithms; T-helper cell epitopes were identified using ProPred (Hammer *et. al.*, 1994). and SYFPEITHI algorithms, and; B-cell epitopes were identified using the algorithm for prediction of the hydropathy index (Kyte and Doolittle1982) and the algorithm used for prediction of the secondary structure for β-turns (Chou and Fasman 1974).

1. Computational identification of CTL epitopes.

In order to predict the MHC binding peptides starting from the primary sequence of viral proteins of interest, we used the BIMAS prediction algorithm and the SYFPEITHI algorithm (URLs: (http://bimas.dcrt.nih.gov/molbio/hla_bind/ and http://syfpeithi.bmi-heidelberg.com/Scripts/MHCServer.dll/EpPredict.htm, respectively). Both of these software applications are motif/matrix-based algorithms. The BioInformatics & Molecular Analysis Section (BIMAS) software allows one to locate and rank potential 8-mer, 9-mer, or 10-mer peptides based on a predicted half-time of dissociation to HLA class I molecules. The analysis is based on coefficient tables deduced from the literature published by Parker *et. al.* (1992a; 1992b; 1994). To calculate coefficients, these investigators measured the stability of a large number of HLA-A2

complexes containing distinct peptides, as assessed by measuring the rate of β_2 microglobulin (β_2m) dissociation. Peptides that form HLA-A2 complexes were distinguished from nonbinding peptides by use of gel filtration assay in which the ability of each peptide to promote incorporation of ^{125}I-β_2m into HLA-A2 complexes was assessed (Parker *et. al.*, 1992b). Parker and colleagues also compiled a list of peptides that were unable to make stable complexes. The β_2m dissociation data for each peptide was used to generate independent equations in which the measured half-life of β_2m dissociation was set equal to the product of the nine coefficients. In the case of HLA-A2, the BIMAS score corresponds to the estimated half time of dissociation of complexes containing the peptide at 37°C at pH 6.5. For other molecules, the estimate is based on the observed anchor residue preferences. The tables are also used for quantitative prediction about which amino acids in a given peptide are of primary importance for binding to HLA-A2. Parker et. al. (1994) tested the importance of every amino acid in the influenza A matrix peptide 58-66 (sequence GILGFVFTL) for binding to HLA-A2 using Ala-substituted and Lys-substituted peptides. They found that multiple positions were important for stable binding, including P2, P3, P5-P7, and P9 and they concluded that the P2 and p9 anchor residues are of prime importance for peptide binding to HLA-A2.

The idea behind coefficient tables is the assumption that each amino acid in the peptide contributes independently to the binding to the class I molecule. Thus binding affinity of any nonamer can be broken down into nine coefficients, each of which is dependent only on the identity of amino acid and its position within the peptide. A table containing 180 coefficients (20 amino acids x 9 positions) contains the information

necessary to calculate a probable binding affinity for any possible nonamer. An example of the coefficient table for a specific HLA class I specificity (A*0201) is shown in Table 3. Dominant anchor residues, which are critical for binding, have coefficients that are significantly greater than 1. Unfavorable (contributing to low predicted binding) amino acids have positive coefficients that are less than 1. Auxiliary anchor residues have coefficients that are different from 1 but smaller in magnitude than dominant anchor residues.

The score in the output table is obtained starting with the initial score that is set to 1.0. For each residue position, the program examines which amino acid is appearing at that position. The running score is then multiplied by the coefficient for that amino acid type, at that position, for the chosen HLA molecule. These coefficients have been pre-calculated and are stored for use by the scoring algorithm in a separate directory as a collection of HLA coefficient files. Using 9-mers, nine multiplications are performed. The resulting running score is multiplied by a final constant to yield an estimate of the half time of disassociation. This constant is stored at the end of the coefficient file for the HLA molecule. It can have a different value for each HLA molecule. The final multiplication yields the score reported in the output table.

The second algorithm used in this study (SYFPEITHI) for predicting which peptides bind to MHC molecules was developed by Hans-Georg Rammensee *et. al.* (1999). It should be noted that the name SYFPEITHI which indicates a single letter amino acid code for a peptide sequence was chosen by the developers to acknowledge the first MHC-eluted peptide that was sequenced (Falk *et. al.*, 1991). SYFPEITHI uses motif matrices deduced from refined motifs based on the pool sequence and single peptide

analysis exclusively of natural ligands. Potential binders for various MHC class I molecules are ranked according to the presence of primary and secondary anchor amino acids as we as favored and disfavored amino acids. The matrices in SYFPEITHI were adjusted manually, by assigning a high score of 10 for frequently occurring anchor residues, a score of 8 to amino acids that occur in a significant amount and a score of 6 to rarely occurring residues. Preferred amino acids in other positions have scores that range from 1 to 6 and amino acids regarded as unfavorable have scores ranging from –3 to –1.

Other matrix-based algorithms are RANKPEP (Reche et. al., 2002), ProPred (Singh and Raghava 2003) and SVMHC (Dönnes and Elofsson 2002). RANKPEP ranks peptides from an input protein sequence by their similarity to a set of peptides known to bind to a given MHC molecule. Support Vector Machine (SVMHC), a class of machine learning methods was used to predict peptides that bind to MHC class I molecule based on quantitative matrices. The algorithm contains prediction for 26 MHC class I types from MHCPEP (http://wehih.wehi.edu.au/mhcpep/) database and prediction for 6 MHC class I from the SYFPEITHI database. The ProPred algorithm implements matrices for 47 MHC class I alleles, as well as proteasomal and immunoproteasomal models. It identifies the promiscuous MHC class I binding regions in antigens. Finally, the nHLAPred algorithm described by Gulukota *et. al.,* (1997) employs artificial neural networks. This method allows prediction of MHC binding peptides or CTL epitopes of 67 MHC alleles. The predicted MHC binders can be filtered to further assess their potential as CTL epitopes by analyzing such motifs using proteasomal matrices. PREDEP the structure –based algorithm was developed by Altuvia *et. al.* (1997). The structural approach used for developing this algorithm relies on structural conservation observed in

crystallographically solved peptide-MHC complexes (this approach is limited to MHC types with a known structure). The peptide structure in the MHC groove is used as a template upon which peptide candidates are threaded to obtain a rough estimate of binding energy, which is based on interactions defined in the binding pocket of a particular MHC molecule. List of discussed freely available algorithms for CTL epitope mapping is shown in Table 2.

Although in the present study, we computationally analyzed NP and GP sequences using each of these algorithms, we focused our attention on BIMAS and SYFPEITHI in generating the final list of CTL epitope candidates to be tested *in vitro* and *in vivo*. The protein sequence of nucleoprotein NP and glycoprotein GP both encoded by S RNA segment of Lassa virus were obtained from Entrez database. The 569 amino acid NP sequence and 491 amino acid GP sequence were separately analyzed computationally using the BIMAS prediction server and SYFPEITHI algorithm. The total protein sequence (NP or GP) was parsed into peptides, nine amino acid in length, each overlapping by eight amino acids. A dissociation constant (BIMAS score) of ≥ 50 was used as the selection criteria. The ten highest-ranking peptides for GP and NP analyzed by BIMAS and SYFPEITHI are shown in Table 4.

Table 2. Summary of algorithms available for CTL epitope mapping.

Algorithm	URL	Based on[1]
BIMAS	http://bimas.dcrt.nih.gov/molbio/hla_bind/	Matrix
SYFPEITHI	http://syfpeithi.bmi-heidelberg.com/Scripts/MHCServer.dll/EpPredict.htm	Matrix
RANKPEP	http://mif.dfci.harvard.edu/Tools/rankpep.html	Matrix
ProPred	http://www.imtech.res.in/raghava/propred1/	ANN
nHLAPred	http://www.imtech.res.in/raghava/nhlapred/comp.html	Structure
PREDEP	http://bioinfo.md.huji.ac.il/marg/Teppred/mhc-bind/	Vector
SVMHC	http://www-bs.informatik.uni-tuebingen.de/Services/SVMHC	

[1] Matrix-based analyses utilizes quantitative matrices; Artificial neural networks (ANN) utilizes a particular information processing approach. Structure-based on structural conservation observed in crystallographically solved peptide-MHC complexes. Vector Machine utilized quantitative matrices with a particular machine learning approach. Additional information about these algorithms is provided in the Results section.

Table 3. BIMAS: HLA Coefficient Table for File "A_0201_ Standard"

Amino Acid Type	Position								
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th
A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
C	1.000	0.470	1.000	1.000	1.000	1.000	1.000	1.000	1.000
D	0.075	0.100	0.400	4.100	1.000	1.000	0.490	1.000	0.003
E	0.075	1.400	0.064	4.100	1.000	1.000	0.490	1.000	0.003
F	4.600	0.050	3.700	1.000	3.800	1.900	5.800	5.500	0.015
G	1.000	0.470	1.000	1.000	1.000	1.000	0.130	1.000	0.015
H	0.034	0.050	1.000	1.000	1.000	1.000	1.000	1.000	0.015
I	1.700	9.900	1.000	1.000	1.000	2.300	1.000	0.410	2.100
K	3.500	0.100	0.035	1.000	1.000	1.000	1.000	1.000	0.003
L	1.700	**72.000**	3.700	1.000	1.000	2.300	1.000	1.000	4.300
M	1.700	52.000	3.700	1.000	1.000	2.300	1.000	1.000	1.000
N	1.000	0.470	1.000	1.000	1.000	1.000	1.000	1.000	0.015
P	0.022	0.470	1.000	1.000	1.000	1.000	1.000	1.000	0.003
Q	1.000	7.300	1.000	1.000	1.000	1.000	1.000	1.000	0.003
R	1.000	0.010	0.076	1.000	1.000	1.000	0.200	1.000	0.003
S	1.000	0.470	1.000	1.000	1.000	1.000	1.000	1.000	0.015
T	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.500
V	1.700	6.300	1.000	1.000	1.000	2.300	1.000	0.410	**14.000**
W	4.600	0.010	8.300	1.000	1.000	1.700	7.500	5.500	0.015
Y	4.600	0.010	3.200	1.000	1.000	1.500	1.000	5.500	0.015
final constant	0.069								

BIMAS				SYFPEITHI			
Rank	**Start Position**	**Peptide**	**Score**	**Rank**	**Start Position**	**Peptide**	**Score**
LFV-GP (Protein ID: NP_694870.1 – 491 AA)							
1	258	**LLGTFTWTL**	1,035.01	1	60	**SLYKGVYEL**	32
2	434	FVFSTSFYL	1,014.87	2	23	LIALSVLAV	28
3	386	WLVSNGSYL	540.47	3	42	**GLVGLVTFL**	28
4	257	RLLGTFTWT	288.13	4	258	**LLGTFTWTL**	27
5	314	RLFDFNKQA	286.39	5	336	LINKAVNAL	27
6	42	GLVGLVTFL	270.23	6	444	SIFLHLVKI	27
7	441	**YLISIFLHL**	186.71	7	442	LISIFLHLV	26
8	60	**SLYKGVYEL**	157.23	8	18	VMNIVLIAL	25
9	442	LISIFLHLV	142.09	9	441	**YLISIFLHL**	25
10	212	CIMTSYQYL	89.66	10	24	IALSVLAVL	24
LFV-NP (Protein ID: NP_694869.1– 569 AA)							
1	257	CMLDGGNML	234.05	1	101	ILAADLEK	30
2	497	AVWDQYKDL	95.41	2	277	GILKSILKV	28
3	277	GILKSILKV	81.39	3	194	DLNDAVQAL	26
4	288	ALGMFISDT	65.13	4	69	DLNQAVNNL	25
5	101	ILAADLEKL	40.93	5	98	DLLILAADL	24
6	531	LMDCIMFDA	32.55	6	257	CMLDGGNML	24
7	355	SLQSAGFTA	28.81	7	66	RLRDLNQAV	23
8	403	YQPSSGCYI	26.76	8	240	NISGYNFSL	23
9	370	LMTLKDAML	26.23	9	274	TMDGILKSI	23
10	66	RLRDLNQAV	21.67	10	365	LTYSQLMTL	23

Table 4. Computational prediction of LFV CTL epitopes by BIMAS and SYFPEITHI algorithms. The protein sequence of nucleoprotein NP and glycoprotein GP both encoded by S RNA segment of Lassa virus were obtained from Entrez database. The 569 amino acid NP sequence and 491 amino acid GP sequence were separately analyzed computationally using the BIMAS prediction server and SYFPEITHI algorithm. The total protein sequence (NP or GP) was parsed into peptides, nine amino acid in length, each overlapping by eight amino acids. The peptides in bold represent high affinity epitopes, able to stabilize HLA-A2.1molecules on T2 cells. The sequences in blue represent low affinity peptides as tested in the MHC-stabilization assay (Figure 16).

Table 5. Computational predeiction of LFV CTL epitopes using other available algorithms. The glycoprotein of LFV has been analysed.

Rank	PREDEP	At Position	ProPred	At Position	nHLAPred	At Position	RANKPEP	At Position
1	SLYKGVYEL	60	LLGTFTWTL	258	YLISIFLHL	441	SLYKGVYEL	60
2	ALMSIISTF	132	FVFSTSFYL	434	PLGLVDLFV	427	ALMSIISTF	132
3	LLSQRTRDI	244	WLVSNGSYL	386	WLVSNGSYL	386	LLSQRTRDI	244
4	LINKAVNAL	336	RLLGTFTWT	257	LLGTFTWTL	258	LINKAVNAL	336
5	LIALSVLAV	23	RLFDFNKQA	314	SLYKGVYEL	60	LIALSVLAV	23
6	YLISIFLHL	441	GLVGLVTFL	42	GLVGLVTFL	42	YLISIFLHL	441
7	GLVGLVTFL	42	YLISIFLHL	441	VMNIVLIAL	18	GLVGLVTFL	42
8	ELTLTNTSI	104	SLYKGVYEL	60	VLAVLKGLY	28	ELTLTNTSI	104
9	VIEEVMNIV	14	LISIFLHLV	442	IMGIPYCNY	358	VIEEVMNIV	14
10	IMTSYQYLI	213	CIMTSYQYL	212	MGICSCGLY	472	IMTSYQYLI	213
11	IISTFHLSI	136	QIVTFFQEV	3	RSCTTSLYK	55	IISTFHLSI	136
12	HVIEEVMNI	13	VMNIVLIAL	18	PSPIGYLGL	236	HVIEEVMNI	13
13	TLELNMETL	70	LLLCGRSCT	50	ATCGLVGLV	39	TLELNMETL	70
14	SIFLHLVKI	444	GLVTFLLLC	45	GTFTWTLSD	260	SIFLHLVKI	444
15	LLGTFTWTL	258	RMAWGGSYI	193	NGGKISVQY	158	LLGTFTWTL	258

Table 6. Computational prediction of LFV CTL epitopes using other available algorithms, NP. The nucleoprotein of LFV has been analysed.

Rank	PREDEP	At Position	ProPred	At Position	nHLAPred	At Position	RANKPEP	At Position
1	AVLPRDMVF	552	CMLDGGNML	257	CMLDGGNML	257	TLKDAMLQL	372
2	ALTDLGLIY	201	ILAADLEKL	101	NISGYNFSL	240	DLNDAVQAL	194
3	HCALMDCIM	528	NISGYNFSL	240	ILAADLEKL	101	ELSGYCSNI	19
4	TLKDAMLQL	372	DLNDAVQAL	194	LTSDDLLIL	94	GILKSILKV	277
5	TLTSDDLLI	93	SILKVKKAL	281	ALMDCIMFD	530	ILAADLEKL	101
6	TWMDIEGRP	386	DLLILAADL	98	SLNISGYNF	238	DLNQAVNNL	69
7	DIKLIDIAL	478	DLFATQPGL	437	DLLILAADL	98	TLTSDDLLI	93
8	DCIMFDAAV	533	ALMDCIMFD	530	DLNDAVQAL	194	RLRDLNQAV	66
9	DLFATQPGL	437	SQGRKDIKL	473	LTSAVIDAL	445	SILKVKKAL	281
10	KLIDIALSK	480	MIDTKKSSL	231	DLFATQPGL	437	TMDGILKSI	274
11	LMTLKDAML	370	LTSDDLLIL	94	AVKAGACML	251	RLTQSHPIL	221
12	AWENTVVDL	330	IMFDAAVSG	535	AGRDGVVRV	155	DLEKLKSKV	105
13	VVKDAQALL	31	QAVNNLVEL	72	ILRVGTLTS	88	GLTSAVIDA	444
14	FLWTQSLRR	10	FGTMPSLTL	176	SILKVKKAL	281	QLDQRRALL	133
15	DALPRNMVI	451	LTSAVIDAL	445	MDGILKSIL	275	KLKSKVIRT	108

2. Computational identification of LFV T-helper cell epitopes

Computational algorithms used for the mapping of epitopes presented on MHC class II molecules and recognized by CD4 positive T cells are similar to those used for mapping cytotoxic T cell epitopes. Such algorithms allow investigators to choose and rank potential class II binders, based on knowledge of critical anchor residues that facilitate the binding of peptides to particular MHC class II alleles. Two algorithms were used for the mapping of T-helper cell epitopes, namely, ProPred and SYFPEITHI. The ProPred algorithm (http://www.imtech.res.in/raghava/propred/) employs virtual MHC-class II matrices based on pocket profiles that provide a model in which the contribution to binding of each amino acid with each pocket/position (HLA binding cleft) is quantified. The virtual matrices were originally incorporated in TEPITOPE software (www.vaccinome.com), (Sturniolo *et. al.*, 1999). Pocket profile data were determined in competitive HLA class II binding experiments using sets of frame-fixed single-substituted peptides and purified HLA class II alleles (Hammer *et. al.*, 1994). The M13 phage-expressed random peptide library was screened against the DR molecules isolated from different cell lines. M13 phage/DR complexes were then purified and eluted and eluates were used to isolate and sequence individual M13 clones. Synthetic peptides labeled with ^{125}I and unlabeled competitor peptides were used in a competition assay with affinity-purified DR molecules (DRB1*0401, DRB1*0101 and DRB1*1101). Anchor residues for DRB1*0101 were determined to be an aromatic or aliphatic hydrophobic residue (Met, Ile, Val, Phe, Trp, and Tyr) at position P1, hydrobhobic residue (typically Met or Ala) at P4, Thr or Ser at P6 (Hammer *et. al.*, 1993) and Arg at P2, Ala or Glt at P3 and Leu at P7 (Hammer *et. al.*, 1994). Some of the anchor residues for DRB1*0401 were

shown to be shared by the DRB1*0101 and DRB1*1101 alleles. Once a pocket profile has been determined in vitro, it can be shared among other HLA-II alleles as long as the amino acid residues contributing to the pocket are identical. Thus, a relatively small number of profiles was used to build a large number of HLA-II matrices.

Virtual matrices (see Table 7) were formed by assigning and combining pocket-specific quantitative binding values derived from one HLA allele to other alleles via HLA sequence homology comparison (Bian *et. al.*, 2003). Because the virtual matrices concept addresses the problem of HLA polymorphism, it enables the prediction of peptide ligands for broad range of MHC class II alleles. Currently, the ProPred server includes matrices for 51 HLA-DR alleles that cover more than 90% of all MHC class-II molecules identified. Virtual matrix-based prediction models have been applied to predict class-II restricted T-cell epitopes in the context of infectious diseases (Panigada *et. al.*, 2002), oncology (Sturniolo *et. al.*, 1999), allergy (de Lalla *et. al.*, 1999), and autoimmune diseases (Hammer *et. al.*, 1997).

Using the SYFPEITHI algorithm, the prediction of class-II MHC restricted T-cell epitopes is available for a small number of alleles and the predicted peptide sequences are 15-amino acids in length. This algorithm was used here mainly for the comparison purposes as a guide for choosing and ranking of peptides.

Schultze *et. al.* (2003) have identified B- and T-cell epitopes within the Fibronectin-Binding domain of the SfbI Protein of *Streptococcus pyogenes*. One linear B-cell epitope was recognized by BALB/b and BALB/k mice, whereas two epitopes were recognized in BALC/c animals. A unique T-cell epitope was recognized by all three mouse strains. A motif within the peptide sequence that induced the proliferation of CD4

T cells from the previously immunized mice was also predicted to be recognized in the human system by the SYFPEITHI algorithm.

The protein sequences of LFV NP and GP were analyzed computationally to identify candidate T-helper cell epitopes. This was accomplished using the ProPred and SYFPEITHI algorithms to examine the 569 amino acid NP and the 491 amino acid GP sequences. Each sequence was analyzed against all available HLA-DR alleles simultaneously using the ProPred algorithm. Thus, the output results obtained from this algorithm allow for the comparison of predicted binders and their cross-reactivity against 51 available HLA-DR alleles. Predicted binders were displayed as blue-colored regions, with P1 identified as the starting residue of each predicted binding frame colored in red (see Table 8 for GP and Table 11 for NP). In an additional analyses, the twenty highest scoring HLA-DRB1*0101 restricted peptides were ranked by the ProPred algorithm and compared to the twenty highest scoring DRB1*0101 restricted peptides ranked by the SYFPEITHI algorithm for both GP and NP (Table 9 and 12). The LFV PG sequence with highlighted proposed cross reactive class II epitopes for 51 different HLA-DR alleles is shown in Table 10 and for NP in Table 13.

The threshold value for the ProPred analysis was set to 3%. This threshold correlates with the peptide score and therefore with HLA-ligand interaction (Sturnilo *et. al.*, 1999). The peptide score was calculated after the multiplication of the partial values assigned to particular amino acids (see Table 7). The threshold is defined as the percentage of the best scoring peptides. The lower the threshold (high stringency), the lower the false positive rate. For example a threshold of 1% would predict peptides in any given protein sequence that belongs to the 1% best scoring peptides. Such a high-

stringency threshold (1%) results in fewer selected "positive" peptides, and might miss some potential binders. Studies have shown that the most practical and optimal threshold level is 3% (Bian *et. al.*, 2003).

As with the analysis of candidate CTL epitopes, the LFV NP and GP sequences identified, as candidate T-helper epitopes were first screened for their similarity to the human genome using the NIH Blast server (http://www.ncbi.nlm.nih.gov/blast/). Peptides showing homology with regard to the human proteome were eliminated as possible candidates for T-cell epitopes.

Table 7. An example of a virtual matrix for HLA-DRB1*0101.

Amino acid/Position	P1	P2	P3	P4	P5	P6	P7	P8	P9
A:	-999	0	0	0	-	0	0	-	0
D:	-999	-1.3	-1.3	-2.4	-	-2.7	-2	-	-1.9
E:	-999	0.1	-1.2	-0.4	-	-2.4	-0.6	-	-1.9
F:	0	0.8	0.8	0.08	-	-2.1	0.3	-	-0.4
G:	-999	0.5	0.2	-0.7	-	-0.3	-1.1	-	-0.8
H:	-999	0.8	0.2	-0.7	-	-2.2	0.1	-	-1.1
I:	-1	1.1	1.5	0.5	-	-1.9	0.6	-	0.7
K:	-999	1.1	0	-2.1	-	-2	-0.2	-	-1.7
L:	-1	1	1	0.9	-	-2	0.3	-	0.5
M:	-1	1.1	1.4	0.8	-	-1.8	0.09	-	0.08
N:	-999	0.8	0.5	0.04	-	-1.1	0.1	-	-1.2
P:	-999	-0.5	0.3	-1.9	-	-0.2	0.07	-	-1.1
Q:	-999	1.2	0	0.1	-	-1.8	0.2	-	-1.6
R:	-999	2.2	0.7	-2.1	-	-1.8	0.09	-	-1
S:	-999	-0.3	0.2	-0.7	-	-0.6	-0.2	-	-0.3
T:	-999	0	0	-1	-	-1.2	0.09	-	-0.2
V:	-1	2.1	0.5	-0.05	-	-1.1	0.7	-	0.3
W:	0	-0.1	0	-1.8	-	-2.4	-0.08	-	-1.4
Y:	0	0.9	0.8	-1.1	-	-2	0.5	-	-0.9

Table 8. Computational prediction of MHC class II-binding peptides for 51 different HLA-DR alleles for Lassa GP, allele cross reactivity. The highlighted are the peptides that show cross- reactivity across most of the tested alleles.

```
                1----------10----------20----------30----------40----------50------55
DRB1_0101:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0102:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0301:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0305:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0306:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0307:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0308:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0309:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0311:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0401:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0402:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0404:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0405:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0408:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0410:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0421:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0423:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0426:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0701:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0703:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0801:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0802:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0804:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0806:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0813:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_0817:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1101:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1102:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1104:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1106:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1107:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1114:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1120:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1121:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1128:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1301:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1302:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1304:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1305:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1307:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1311:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1321:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1322:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1323:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1327:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1328:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1501:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1502:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB1_1506:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB5_0101:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
DRB5_0105:MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLLCGR
```

```
                56--------60----------70----------80----------90----------100-------108
DRB1_0101:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0102:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0301:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0305:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0306:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0307:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0308:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0309:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0311:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0401:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0402:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0404:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0405:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0408:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0410:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0421:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0423:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0426:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0701:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0703:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0801:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0802:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0804:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0806:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0813:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_0817:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1101:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1102:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1104:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1106:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1107:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1114:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1120:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1121:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1128:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1301:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1302:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1304:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1305:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1307:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1311:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1321:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1322:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1323:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1327:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1328:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1501:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1502:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB1_1506:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB5_0101:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
DRB5_0105:SCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETGLELTLT
```

```
                109----------120----------130----------140----------150----------160--163
DRB1_0101: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0102: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0301: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0305: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0306: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0307: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0308: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0309: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0311: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0401: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0402: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0404: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0405: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0408: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0410: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0421: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0423: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0426: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0701: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0703: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0801: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0802: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0804: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0806: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0813: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_0817: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1101: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1102: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1104: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1106: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1107: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1114: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1120: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1121: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1128: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1301: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1302: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1304: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1305: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1307: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1311: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1321: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1322: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1323: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1327: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1328: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1501: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1502: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB1_1506: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB5_0101: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
DRB5_0105: TNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSCDFNGGKIS
```

```
                164----------170----------180----------190----------200----------210----------215
DRB1_0101:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0102:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0301:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0305:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0306:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0307:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0308:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0309:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0311:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0401:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0402:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0404:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0405:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0408:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0410:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0421:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0423:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0426:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0701:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0703:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0801:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0802:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0804:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0806:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0813:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_0817:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1101:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1102:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1104:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1106:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1107:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1114:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1120:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1121:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1128:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1301:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1302:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1304:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1305:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1307:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1311:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1321:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1322:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIM
DRB1_1323:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1327:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1328:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1501:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1502:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB1_1506:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB5_0101:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
DRB5_0105:VQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDSGRGNWDCIMT
```

```
                 216----------230----------240----------250----------260----------270
DRB1_0101: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0102: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0301: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0305: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0306: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0307: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0308: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0309: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0311: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0401: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0402: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0404: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0405: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0408: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0410: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0421: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0423: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0426: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0701: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0703: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0801: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0802: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0804: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0806: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0813: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_0817: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1101: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1102: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1104: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1106: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1107: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1114: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1120: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1121: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1128: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1301: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1302: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1304: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1305: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1307: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1311: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1321: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1322: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1323: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1327: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1328: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1501: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1502: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB1_1506: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB5_0101: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
DRB5_0105: TSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRLLGTFTWTLSDSE
```

```
                271----------280----------290----------300----------310----------323
DRB1_0101:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0102:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0301:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0305:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0306:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0307:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0308:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0309:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0311:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0401:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0402:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0404:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0405:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0408:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0410:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0421:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0423:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0426:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0701:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0703:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0801:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0802:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0804:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0806:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0813:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_0817:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1101:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1102:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1104:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1106:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1107:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1114:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1120:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1121:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1128:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1301:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1302:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1304:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1305:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1307:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1311:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1321:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1322:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1323:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1327:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1328:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1501:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1502:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB1_1506:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB5_0101:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
DRB5_0105:GKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEFCDMLRLFDFNKQAI
```

```
                 324 --------330----------340----------350----------360----------370----------377
DRB1_0101:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0102:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0301:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0305:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0306:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0307:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0308:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0309:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0311:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0401:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0402:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0404:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0405:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0408:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0410:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0421:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0423:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0426:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0701:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0703:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0801:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0802:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0804:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0806:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0813:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_0817:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1101:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1102:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1104:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1106:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1107:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1114:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1120:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1121:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1128:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1301:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1302:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1304:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1305:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1307:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1311:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1321:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1322:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1323:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1327:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1328:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1501:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1502:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB1_1506:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB5_0101:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
DRB5_0105:QRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGIPYCNYSKYWYLNHTTT
```

```
                 378----------390----------400----------410----------420----------430
DRB1_0101:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0102:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0301:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0305:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0306:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0307:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0308:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0309:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0311:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0401:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0402:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0404:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0405:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0408:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0410:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0421:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0423:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0426:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0701:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0703:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0801:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0802:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0804:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0806:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0813:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_0817:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1101:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1102:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1104:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1106:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1107:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1114:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1120:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1121:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1128:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1301:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1302:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1304:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1305:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1307:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1311:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1321:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1322:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1323:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1327:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1328:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1501:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1502:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB1_1506:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB5_0101:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
DRB5_0105:GRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMITEMLQKEYMERQGKTPLGL
```

```
                 431----------440----------450----------460----------470----------480------486
DRB1_0101:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0102:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0301:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0305:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0306:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0307:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0308:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0309:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0311:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0401:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0402:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0404:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0405:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0408:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0410:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0421:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0423:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0426:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0701:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0703:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0801:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0802:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0804:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0806:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0813:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_0817:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1101:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1102:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1104:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1106:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1107:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1114:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1120:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1121:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1128:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1301:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1302:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1304:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1305:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1307:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1311:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1321:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1322:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1323:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1327:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1328:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1501:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1502:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB1_1506:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB5_0101:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
DRB5_0105:VDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPKPHRLNHMGICSCGLYKQPGVP
                                                                VKWKR
                                                                487---491
```

Table 9. Comparison of HLADRB1*0101 epitope prediction by ProPred and SYFPEITHI algorithms for glycoprotein of LFV. The highlighted regions are the sequences that are commonly shared by both algorithms. Peptide number 17 on ProPred algorithm and peptide number 11 on SYFPEITHI have been experimentally tested to be an HLA-DR restricted immunodominant epitope for Josiah strain of Lassa virus (ter Meulen *et. al.*, 2004).

ProPred				SYFPEITHI		
Rank	Sequence	At Position	Score	Sequence	At Position	Score
1	FLLLCGRSC	48	2.69	SCGLYKQPGVPVKWK	476	32
2	FVFSTSFYL	433	2.4	TCGLVGLVTFLLLCG	40	31
3	LVGLVTFLL	42	1.8	HALMSIISTFHLSIP	131	31
4	VLIALSVLA	21	1.6	YKGVYELQTLELNME	62	30
5	YLGLLSQRT	240	1.5	YLGLLSQRTRDIYIS	241	29
6	YQYLIIQNT	216	1	LVTFLLLCGRSCTTS	46	28
7	FFQEVPHVI	6	1	LFDFNKQAIQRLKAE	315	27
8	IVLIALSVL	20	0.9	QKEYMERQGKTPLGL	416	27
9	WLVSNGSYL	385	0.8	EEVMNIVLIALSVLA	16	26
10	YNFATCGLV	35	0.8	LQTFMRMAWGGSYIA	188	26
11	LYKQPGVPV	478	0.7	LTRWMLIEAELKCFG	280	26
12	YLIIQNTTW	218	0.59	EAQMSIQLINKAVNA	329	26
13	YLISIFLHL	440	0.5	PKCWLVSNGSYLNET	383	26
14	LIALSVLAV	22	0.5	VDLFVFSTSFYLISI	431	26
15	FLHLVKIPT	445	0.5	MNIVLIALSVLAVLK	19	25
16	MRMAWGGSY	191	0.3	ELNMETLNMTMPLSC	72	25
17	WMLIEAELK	282	0.3	NKAVNALINDQLIMK	338	25
18	LNMTMPLSC	77	0.3	IVTFFQEVPHVIEEV	4	24
19	MSIISTFHL	133	0.3	EVMNIVLIALSVLAV	17	24
20	YNLSHSYAG	165	0.2	VMNIVLIALSVLAVL	18	24

Table 10. LFV GP sequence with highlighted proposed cross-reactive MHC class II epitopes for 51 different HLA-DR alleles. The epitopes were chosen based on the results from the computational analysis of MHC class II binding peptides for 51 different HLA-DR alleles for Lassa GP, the epitopes showing the most cross reactivity are highlighted.

MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLL
CGRSCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETG
LELTLTNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSC
DFNGGKISVQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDS
GRGNWDCIMTSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRL
LGTFWTLSDSEGKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEF
CDMLRLFDFNKQAIQRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGI
PYCNYSKYWYLNHTTTGRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMIT
EMLQKEYMERQGKTPLGLVDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPK
PHRLNHMGICSCGLYKQPGVPVKWKR

Table 11. Computational prediction of MHC class II binding peptides for 51 different HLA-DR alleles for Lassa NP, allele cross reactivity. The highlighted are the peptides that show cross- reactivity across most of the tested alleles.

```
                1---------------10----------------20----------------30----------------40----------------50------55
DRB1_0101:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0102:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0301:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0305:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0306:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0307:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0308:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0309:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0311:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0401:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0402:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0404:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0405:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0408:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0410:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0421:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0423:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0426:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0701:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0703:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0801:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0802:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0804:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0806:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0813:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_0817:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1101:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1102:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1104:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1106:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1107:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1114:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1120:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1121:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1128:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1301:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1302:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1304:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1305:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1307:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1311:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1321:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1322:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1323:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1327:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1328:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1501:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1502:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB1_1506:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB5_0101:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
DRB5_0105:MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQRLM
```

```
                56--------60----------70----------80----------90----------100-------108
DRB1_0101:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0102:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0301:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0305:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0306:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0307:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0308:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0309:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0311:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0401:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0402:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0404:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0405:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0408:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0410:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0421:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0423:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0426:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0701:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0703:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0801:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0802:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0804:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0806:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0813:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_0817:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1101:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1102:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1104:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1106:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1107:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1114:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1120:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1121:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1128:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1301:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1302:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1304:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1305:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1307:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1311:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1321:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1322:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1323:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1327:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1328:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1501:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1502:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB1_1506:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB5_0101:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
DRB5_0105:RKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILAADLEK
```

```
                109----------120----------130----------140----------150----------161
DRB1_0101:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0102:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0301:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0305:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0306:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0307:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0308:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0309:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0311:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0401:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0402:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0404:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0405:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0408:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0410:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0421:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0423:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0426:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0701:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0703:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0801:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0802:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0804:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0806:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0813:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_0817:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1101:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1102:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1104:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1106:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1107:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1114:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1120:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1121:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1128:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1301:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1302:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1304:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1305:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1307:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1311:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1321:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1322:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1323:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1327:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1328:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1501:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1502:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB1_1506:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB5_0101:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
DRB5_0105:LKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGARAGRDGVV
```

```
          162---------170---------180---------190---------200---------210---------214
DRB1_0101:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0102:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0301:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0305:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0306:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0307:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0308:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0309:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0311:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0401:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0402:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0404:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0405:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0408:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0410:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0421:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0423:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0426:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0701:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0703:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0801:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0802:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0804:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0806:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0813:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_0817:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1101:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1102:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1104:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1106:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1107:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1114:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1120:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1121:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1128:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1301:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1302:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1304:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1305:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1307:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1311:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1321:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1322:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1323:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1327:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1328:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1501:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1502:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB1_1506:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB5_0101:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
DRB5_0105:RVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQALTDLGLIYTAKYP
```

```
                215----------230----------240----------250----------260----------268
DRB1_0101: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0102: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0301: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0305: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0306: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0307: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0308: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0309: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0311: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0401: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0402: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0404: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0405: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0408: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0410: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0421: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0423: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0426: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0701: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0703: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0801: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0802: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0804: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0806: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0813: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_0817: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1101: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1102: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1104: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1106: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1107: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1114: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1120: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1121: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1128: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1301: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1302: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1304: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1305: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1307: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1311: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1321: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1322: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1323: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1327: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1328: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1501: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1502: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB1_1506: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB5_0101: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
DRB5_0105: NTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAGACMLDGGNMLETI
```

```
                269----------280----------290----------300----------310----------322
DRB1_0101: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0102: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0301: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0305: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0306: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0307: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0308: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0309: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0311: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0401: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0402: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0404: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0405: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0408: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0410: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0421: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0423: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0426: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0701: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0703: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0801: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0802: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0804: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0806: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0813: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_0817: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1101: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1102: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1104: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1106: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1107: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1114: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1120: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1121: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1128: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1301: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1302: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1304: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1305: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1307: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1311: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1321: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1322: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1323: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1327: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1328: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1501: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1502: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB1_1506: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB5_0101: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
DRB5_0105: KVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENILYKICLSGDGWPYIAS
```

```
                323 -------330----------340----------350----------360----------370----------376
DRB1_0101:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0102:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0301:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0305:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0306:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0307:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0308:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0309:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0311:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0401:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0402:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0404:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0405:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0408:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0410:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0421:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0423:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0426:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0701:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0703:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0801:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0802:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0804:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0806:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0813:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_0817:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1101:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1102:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1104:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1106:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1107:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1114:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1120:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1121:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1128:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1301:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1302:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1304:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1305:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1307:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1311:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1321:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1322:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1323:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1327:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1328:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1501:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1502:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB1_1506:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB5_0101:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
DRB5_0105:RTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSAGFTAGLTYSQLMTLKDA
```

```
                 378---------------390---------------400---------------410---------------420---------------430
DRB1_0101: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0102: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0301: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0305: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0306: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0307: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0308: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0309: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0311: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0401: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0402: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0404: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0405: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0408: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0410: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0421: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0423: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0426: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0701: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0703: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0801: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0802: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0804: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0806: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0813: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_0817: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1101: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1102: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1104: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1106: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1107: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1114: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1120: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1121: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1128: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1301: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1302: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1304: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1305: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1307: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1311: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1321: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1322: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1323: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1327: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1328: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1501: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1502: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB1_1506: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB5_0101: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
DRB5_0105: MLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCYIHFFREPTDLKQFKQDAKYS
```

```
                431----------440----------450----------460----------470----------480-------486
DRB1_0101: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0102: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0301: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0305: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0306: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0307: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0308: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0309: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0311: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0401: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0402: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0404: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0405: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0408: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0410: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0421: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0423: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0426: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0701: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0703: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0801: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0802: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0804: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0806: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0813: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_0817: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1101: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1102: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1104: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1106: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1107: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1114: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1120: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1121: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1128: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1301: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1302: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1304: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1305: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1307: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1311: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1321: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1322: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1323: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1327: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1328: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1501: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1502: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB1_1506: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB5_0101: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
DRB5_0105: HGIDVTDLFATQPGLTSAVIDALPRNMVITCQGSDDIRKLLESQGRKDIKLIDI
```

```
                487----------500----------510----------520----------530----------538
DRB1_0101:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0102:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0301:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0305:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0306:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0307:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0308:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0309:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0311:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0401:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0402:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0404:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0405:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0408:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0410:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0421:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0423:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0426:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0701:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0703:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0801:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0802:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0804:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0806:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0813:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_0817:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1101:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1102:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1104:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1106:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1107:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1114:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1120:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1121:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1128:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1301:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1302:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1304:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1305:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1307:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1311:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1321:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1322:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1323:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1327:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1328:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1501:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1502:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB1_1506:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB5_0101:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
DRB5_0105:SKTDSRKYENAVWDQYKDLCHMHTGVVVEKKKRGGKEEITPHCALMDCIMFD
```

```
                539----------550----------560----------569
DRB1_0101: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0102: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0301: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0305: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0306: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0307: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0308: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0309: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0311: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0401: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0402: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0404: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0405: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0408: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0410: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0421: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0423: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0426: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0701: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0703: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0801: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0802: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0804: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0806: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0813: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_0817: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1101: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1102: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1104: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1106: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1107: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1114: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1120: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1121: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1128: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1301: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1302: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1304: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1305: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1307: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1311: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1321: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1322: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1323: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1327: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1328: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1501: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1502: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB1_1506: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB5_0101: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
DRB5_0105: AAVSGGLNTSVLRAVLPRDMVFRTSTPRVVL
```

Table 12. Comparison of HLA-DRB1*0101 epitope prediction by ProPred and SYFPEITHI algorithms for nucleoprotein of LFV. The highlighted regions are the sequences that are commonly shared by both algorithms.

ProPred				SYFPEITHI		
Rank	Sequence	At Position	Score	Sequence	At Position	Score
1	FRTSTPRVV	559	1.69	GIDVTDLFATQPGLT	432	34
2	FGTMPSLTL	175	1.5	AADLEKLKSKVIRTE	103	33
3	YQPSSGCYI	402	1.2	AELLNNQFGTMPSLT	169	32
4	LRAVLPRDM	549	1.12	HPILNMIDTKKSSLN	226	32
5	MGNLSSQQL	125	1	TDLFATQPGLTSAVI	436	32
6	YNFSLGAAV	243	0.9	GDGWPYIASRTSITG	314	31
7	IRTERPLSA	113	0.9	TSVLRAVLPRDMVFR	547	31
8	FSLGAAVKA	245	0.7	ILRVGTLTSDDLLIL	88	30
9	VKKALGMFI	284	0.59	GVYMGNLSSQQLDQR	123	30
10	LLNMIGMSG	139	0.29	NNQFGTMPSLTLACL	173	30
11	YKICLSGDG	307	0.1	NMVITCQGSDDIRKL	456	30
12	WPYIASRTS	316	-0.01	QKSILRVGTLTSDDL	85	29
13	IKLQVVKDA	26	-0.1	ISGYNFSLGAAVKAG	241	28
14	LEKLKSKVI	105	-0.1	GYNFSLGAAVKAGAC	243	28
15	LRRELSGYC	15	-0.2	YKDLCHMHTGVVVEK	502	28
16	FATQPGLTS	438	-0.2	SDDLLILAADLEKLK	96	27
17	ILRVGTLTS	87	-0.55	KSILKVKKALGMFIS	280	27
18	VDLNDAVQA	192	-0.56	ILKVKKALGMFISDT	282	27
19	IKLIDIALS	478	-0.6	QSLRRELSGYCSNIK	14	26
20	VQALTDLGL	198	-0.8	NQAVNNLVELKSTQQ	71	26

Table 13. LFV NP sequence with highlighted proposed cross reactive MHC class II epitopes for 51 different HLA-DR alleles. The epitopes were chosen based on the

results from computational analyses of MHC class II binding peptides for 51 different HLA-DR alleles for Lassa GP. The epitopes showing the most cross reactivity are highlighted.

**MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQR
LMRKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILA
ADLEKLKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGA
RAGRDGVVRVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQAL
TDLGLIYTAKYPNTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAG
ACMLDGGNMLETIKVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENIL
YKICLSGDGWPYIASRTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSA
GFTAGLTYSQLMTLKDAMLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCY
IHFFREPTDLKQFKQDAKYSHGIDVTDLFATQPGLTSAVIDALPRNMVITCQ
GSDDIRKLLESQGRKDIKLIDIALSKTDSRKYENAVWDQYKDLCHMHTGVV
VEKKKRGGKEEITPHCALMDCIMFDAAVSGGLNTSVLRAVLPRDMVFRTST
PRVVL**

3. Computational identification of LFV B-cell epitopes.

The algorithm used to predict of the hydropathy index was developed by Kyte and Doolittle (1982). The algorithm used for prediction of the secondary structure for β-turns was developed by Chou and Fasman (1974). Both algorithms are open access resources (ExPASy web site of the University of Geneva; http://expasy.org/tools). Results from β-turn analyses are generated based upon the existing amino acid scale values computed from 29 proteins (Table 14). The results for the hydrophobicity analysis are generated based on the individual values for 20 different amino acids (Table 14).

The method employed for B-cell epitope mapping utilized the correlation between the hydrophilic character of peptide sequence and its propensity to form β-turn structures (Grant, 2003). In order to compute the hydropathy index and the tendency for β-turns, sequences of LFV-derived NP and GP in FASTA format were separately pasted in the

Kyte-Doolittle hydropathy and Chou-Fasman β-turn algorithms. The results were generated in graphical form. A window size of 9 was used and equal weight was given to each amino acid. The window size determines the number of amino acids to be used in computing a value for the amino acid at the center of the window. For example, a window size of 9 includes 4 amino acids on each side of the central amino acid. The value computed for the central amino acid is the average of the values for each amino acid in the window.

Results (Figures 9 and 11) of the two computational analyses (hydropathy and β-turn) were compared and the 3 areas of both NP and GP sequences that were high in turn tendency and high in hydrophilicity (low in hydrophobicity) were highlighted and chosen as antigenic stretches of the protein, and therefore, possible candidates for B-cell epitopes (Tables 15 and 16). The highlighted areas shown (Figures 10 and 12) correspond to positive peaks in Chou-Fasman analysis and negative peaks in the Kyte-Doolittle analysis.

As noted above, the immunogenic peptides identified as candidates were then screened for their similarity to the human genome using the NIH Blast server. Peptides showing homology with regard to the human proteome were eliminated as possible candidates.

Table 14. Amino acid scale values (conformational parameters for β-turns) and individual hydrophobicity values.

Amino acid	Amino acid scale values	Hydrophobicity values
Ala	0.660	1.800
Arg	0.950	-4.500
Asn	1.560	-3.500
Asp	1.460	-3.500
Cys	1.190	2.500
Gln	0.980	-3.500
Glu	0.740	-3.500
Gly	1.560	-0.400
His	0.950	-3.200
Ile	0.470	4.500
Leu	0.590	3.800
Lys	1.010	-3.900
Met	0.600	1.900
Phe	0.600	2.800
Pro	1.520	-1.600
Ser	1.430	-0.800
Thr	0.960	-0.700
Trp	0.960	-0.900
Tyr	1.140	-1.300
Val	0.500	4.200

Figure 9. B-cell epitope prediction: Graphical presentation of results generated using the ProtScale algorithm for LFV GP. The protein sequence of the GP was analyzed by both beta–turn and hydrophobicity prediction algorithms.

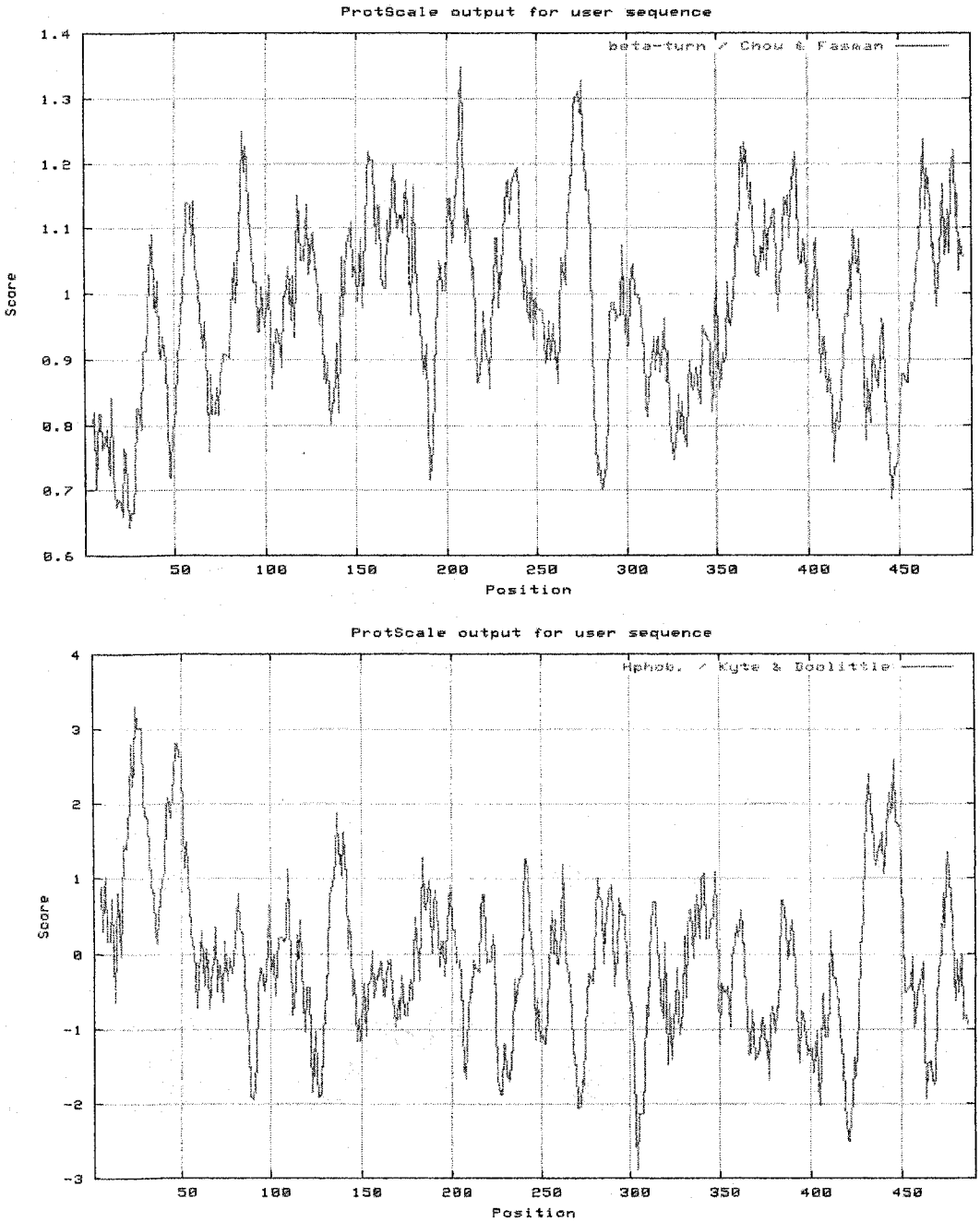
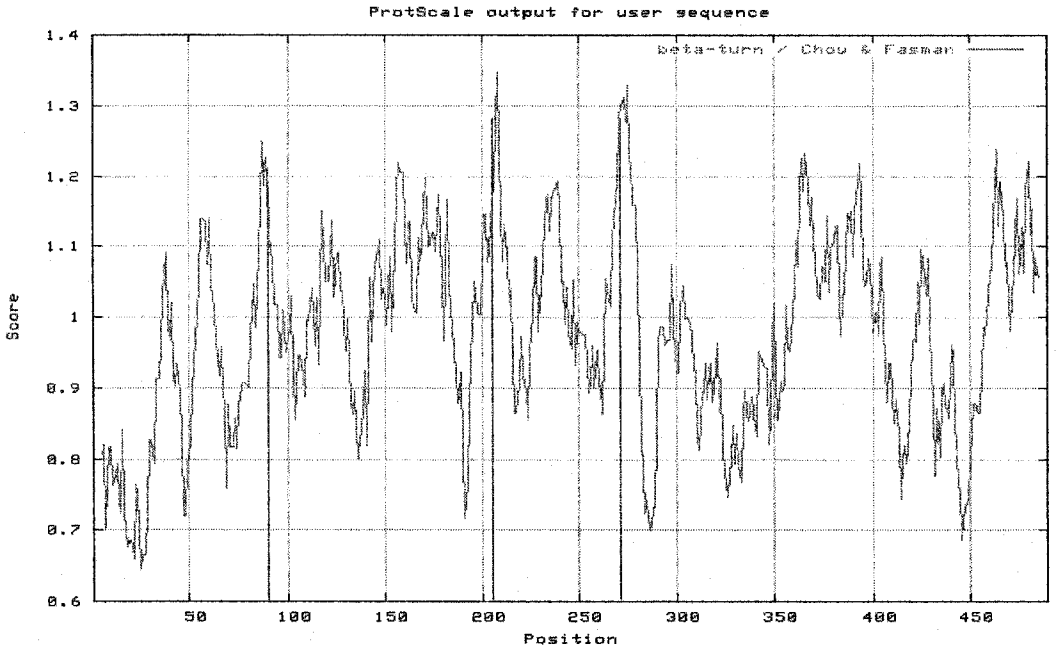


Figure 10. B-cell epitope mapping. Graphical presentation of the results generated using the ProtScale algorithm for t LFV GP showing proposed B-cell epitopes: GP 75 – 101,

198 – 218, 275 – 286. The peaks indicating beta-turns corresponding hydrophilicity stretches of the protein are identified in red.



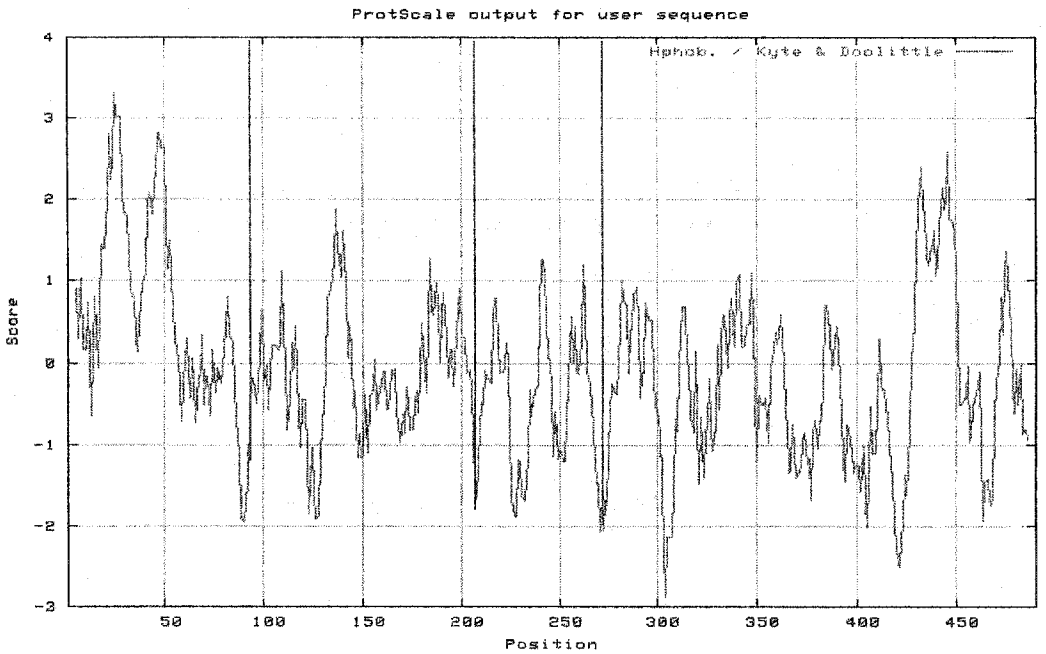


Table 15. LFV GP sequence with highlighted proposed B-cell epitopes. The beta-turns containing stretches of the protein that were high in hydrophilicity were chosen as prospective B- cell epitopes.

**MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLKGLYNFATCGLVGLVTFLLL
CGRSCTTSLYKGVYELQTLELNMETLNMTMPLSCTKNNSHHYIMVGNETG
LELTLTNTSIINHKFCNLSDAHKKNLYDHALMSIISTFHLSIPNFNQYEAMSC
DFNGGKISVQYNLSHSYAGDAANHCGTVANGVLQTFMRMAWGGSYIALDS
GRGNWDCIMTSYQYLIIQNTTWEDHCQFSRPSPIGYLGLLSQRTRDIYISRRL
LGTFWTLSDSEGKDTPGGYCLTRWMLIEAELKCFGNTAVAKCNEKHDEEF
CDMLRLFDFNKQAIQRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDIMGI
PYCNYSKYWYLNHTTTGRTSLPKCWLVSNGSYLNETHFSDDIEQQADNMIT
EMLQKEYMERQGKTPLGLVDLFVFSTSFYLISIFLHLVKIPTHRHIVGKSCPK
PHRLNHMGICSCGLYKQPGVPVKWKR**

Figure 11. B-cell epitope prediction: Graphical presentation of results generated using the ProtScale algorithm for LFV NP. The protein sequence of the NP was analyzed by both beta–turn and hydrophobicity prediction algorithms.

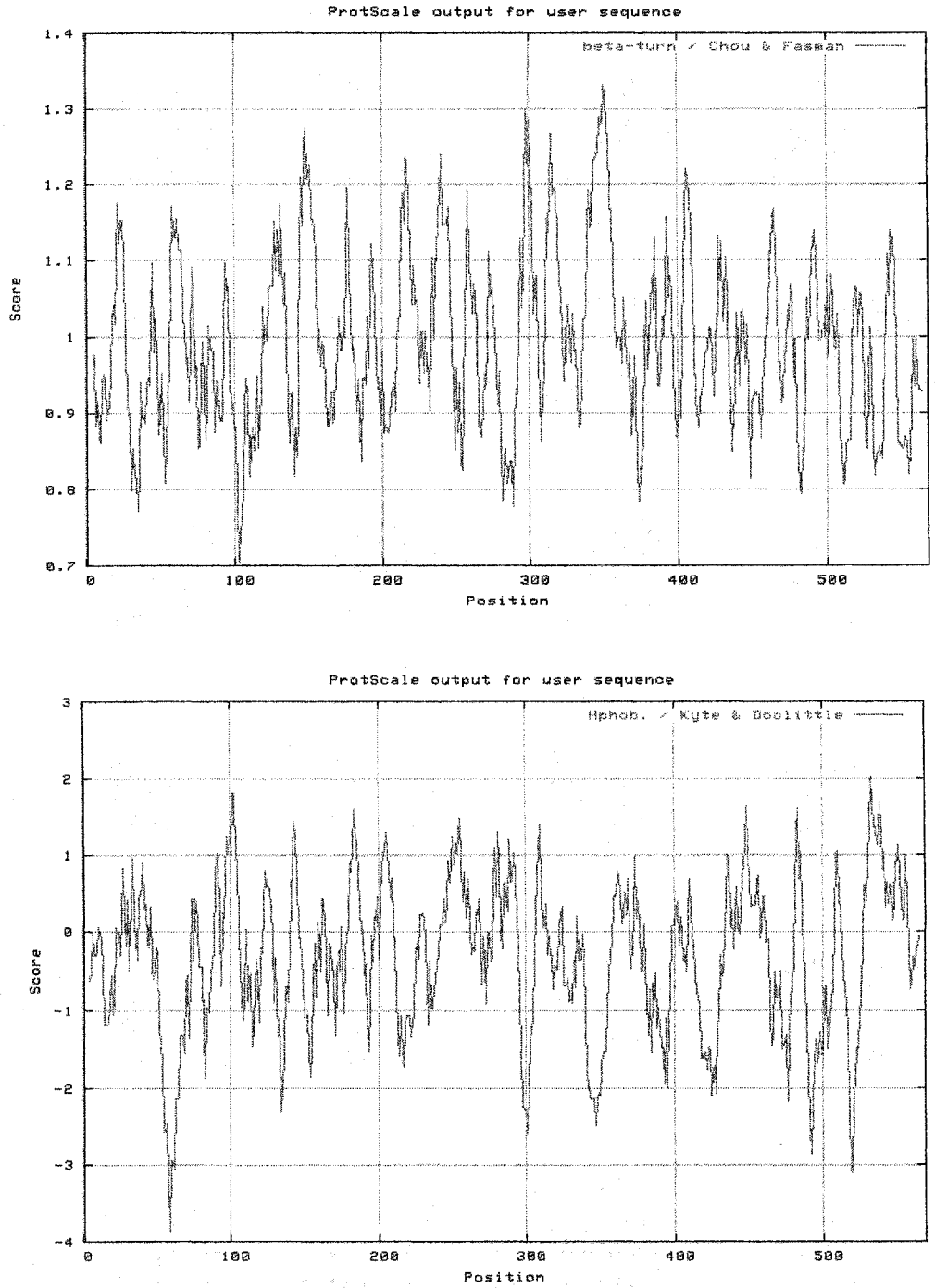


Figure 12. B-cell epitope mapping. Graphical presentation of the results generated by ProtScale algorithm for the NP of Lassa virus, with proposed B-cell epitopes: NP 143 – 158, 292 – 308, 338 – 359. The peaks indicating the beta-turn and corresponding hydrophilicity stretches of the protein are identified in red.

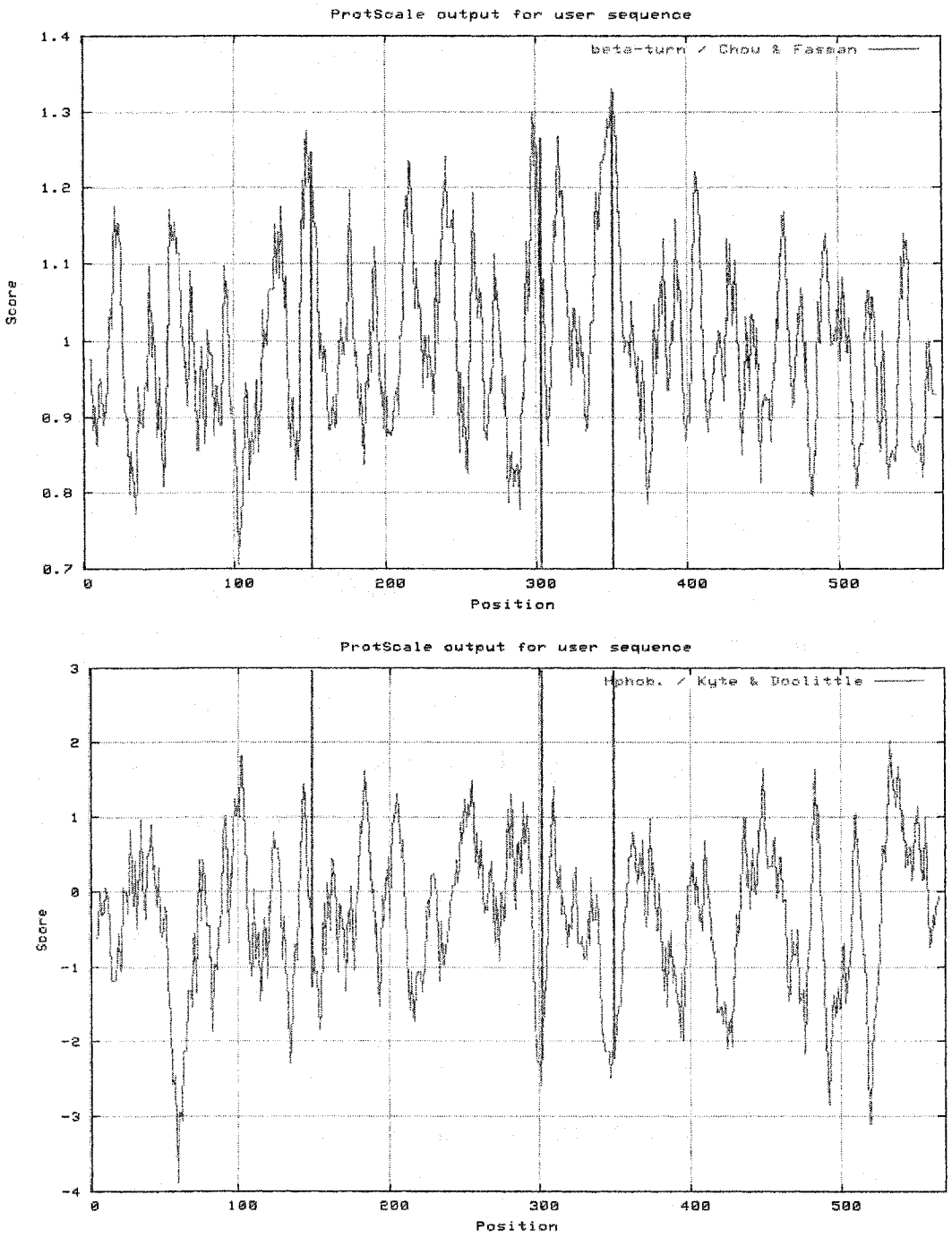


Table 16. Lassa virus NP sequence with highlighted proposed B-cell epitopes. The beta-turn containing stretches of the protein that were high in hydrophilicity were chosen as prospective B- cell epitopes.

**MSASKEIKSFLWTQSLRRELSGYCSNIKLQVVKDAQALLHGLDFSEVSNVQR
LMRKERRDDNDLKRLRDLNQAVNNLVELKSTQQKSILRVGTLTSDDLLILA
ADLEKLKSKVIRTERPLSAGVYMGNLSSQQLDQRRALLNMIGMSGGNQGA
RAGRDGVVRVWDVKNAELLNNQFGTMPSLTLACLTKQGQVDLNDAVQAL
TDLGLIYTAKYPNTSDLDRLTQSHPILNMIDTKKSSLNISGYNFSLGAAVKAG
ACMLDGGNMLETIKVSPQTMDGILKSILKVKKALGMFISDTPGERNPYENIL
YKICLSGDGWPYIASRTSITGRAWENTVVDLESDGKPQKADSNNSSKSLQSA
GFTAGLTYSQLMTLKDAMLQLDPNAKTWMDIEGRPEDPVEIALYQPSSGCY
IHFFREPTDLKQFKQDAKYSHGIDVTDLFATQPGLTSAVIDALPRNMVITCQ
GSDDIRKLLESQGRKDIKLIDIALSKTDSRKYENAVWDQYKDLCHMHTGVV
VEKKKRGGKEEITPHCALMDCIMFDAAVSGGLNTSVLRAVLPRDMVFRTST
PRVVL**

PART II. Validation of computational predictions of CTL epitopes

1. Standardization of MHC-stabilization assay

The MHC stabilization assay relies on the ability of exogenously added peptides to increase the number of properly folded HLA-A2.1 molecules on the cell surface of TAP-deficient cells such as T2 cells (Nijman *et. al.*, 1993). The exogenously added peptides stabilize the class I/β_2m structure on these cells and this can be shown by immunofluorescence staining using a fluorochrome-conjugated anti-HLA-A2.1 antibody.

Before the testing of the computationally identified LFV derived peptides for their ability to stabilize the class I molecules *in vitro*, the MHC stabilization assay was standardized. The optimal concentrations of both the tested peptides and the HLA-A2

specific antibody used in the assay as well as the optimal temperature conditions were determined. A positive control peptide GILGFVFTL derived from influenza matrix antigen (Micheletti *et. al.*, 2002) known to be high binding HLA-A2.1 epitope and a negative control peptide IAGNSAYEY, a known class I non-binder taken from the MHCBN database (http://www.imtech.res.in/raghava/mhcbn/) were used in the optimization assays.

The optimal antibody concentration was determined to be 500ng/5 x 10^5 cells and the optimal peptide concentration was 50μM. The Peptides had a better ability to stabilize the class I molecules in vitro when incubated with the T2 cells at 37°C as compared to the cells incubated with the peptides in the room temperature with or without shaking (Figures 13, 14, and 15).

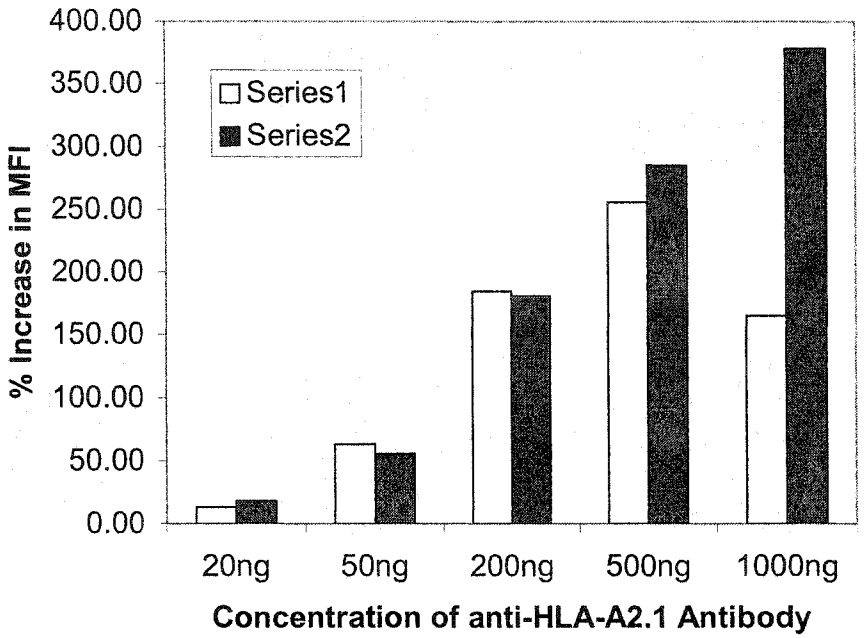


Figure 13. Determination of anti-HLA-A2.1 antibody concentration. T2 cells were grown in IMDM medium (see Materials and Methods) overnight (18 hours) at 37°C. Cells were harvested and washed twice with AIM-V medium. Cells (5 x 10^5/250µl) were then incubated with 100µM of single peptides overnight at 37°C in 2.5 ml borosilicate tubes with AIM-V medium (in duplicate: Series 1 and Series 2). Following 18 hours incubation, the cells were washed twice with cold PBS containing 3% FBS and incubated for 30 min at 4°C with different concentrations (see figure) of fluorochrome-conjugated anti-HLA A2.1 antibody. Following three washes with staining buffer, HLA-A2.1 expression was measured by flow cytometry.

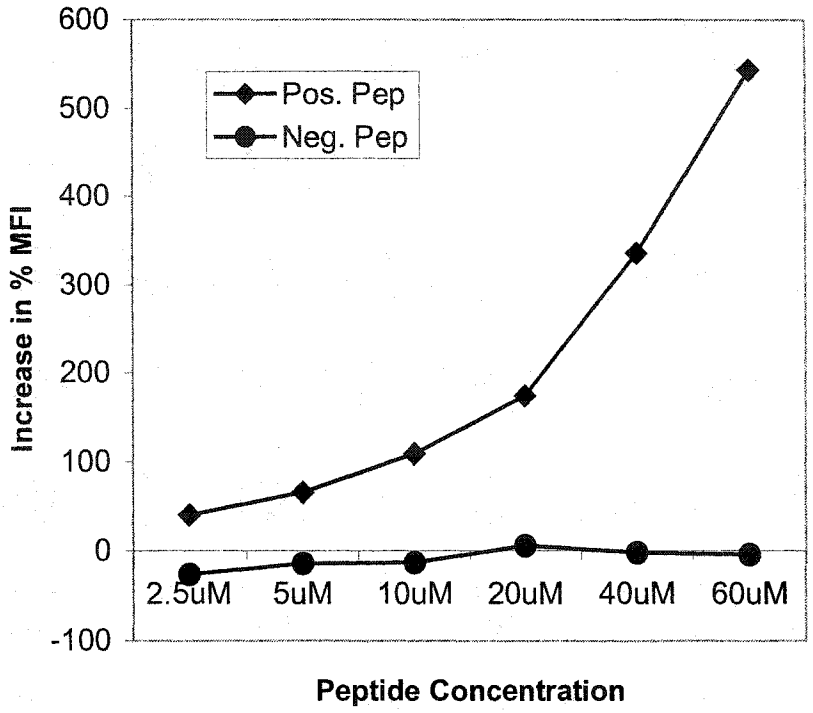


Figure 14. Determination of antigenic peptide concentration. T2 cells were grown in IMDM medium (see Materials and Methods) overnight (18 hours) at 37°C. Cells were harvested and washed twice with AIM-V medium. Cells (5 x 10^5/250µl) were then incubated with different concentrations (see figure) of single peptides overnight at 37°C in 2.5 ml borosilicate tubes with AIM-V medium (in duplicate: Series 1 and Series 2). Following 18 hours incubation, the cells were washed twice with cold PBS containing 3% FBS and incubated for 30 min at 4°C with of fluorochrome-conjugated anti-HLA A2.1 antibody at concentration of 500ng/5 x 10^5 cells. Following three washes with staining buffer, HLA-A2.1 expression was measured by flow cytometry.

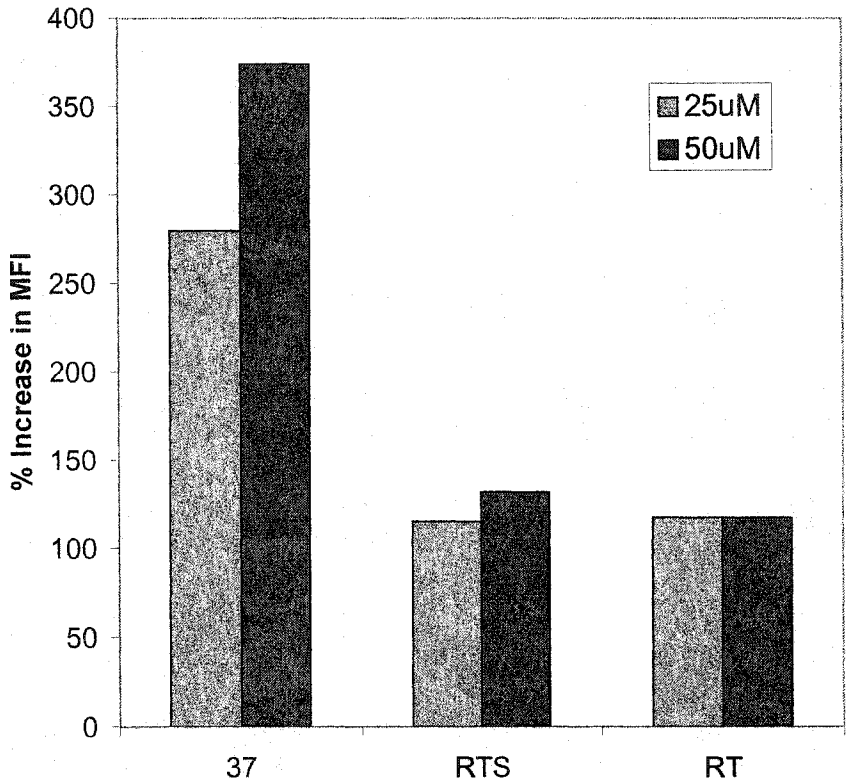


Figure 15. MHC stabilization assay at varying temperatures. T2 cells were grown in IMDM medium overnight (18 hours) at 37°C. Cells were harvested and washed twice with AIM-V medium. Cells (5 x 10^5/250µl) were then incubated with 25 and 50µM of single peptides overnight at 37°C (37), at room temperature with (RTS) and without (RT) shaking in 2.5 ml borosilicate tubes with AIM-V medium. Following 18 hours incubation, the cells were washed twice with cold PBS containing 3% FBS and incubated for 30 min at 4°C with fluorochrome-conjugated anti-HLA A2.1 antibody at the concentration of 500ng/5 x 10^5 cells. Following three washes with staining buffer (PBS + 3% FBS), HLA-A2.1 expression was measured by flow cytometry.

2. Evaluation of synthetic peptides as candidate CTL epitopes

The dominant antigenic peptides in HLA-A2-restricted immune responses are expected to be present among those peptides that bind to HLA-A2 *in vitro*. The algorithms used for computational analysis of a protein sequence are designed so that the highest scoring peptides will most likely bind to the HLA-A2 molecule *in vitro* and should thus be immunogenic when tested *in vivo*. Nevertheless, despite their use as valid tools for epitope mapping (Parker *et. al.,* 1994; Garcia *et. al.,* 2003; Martin *et. al.*, 2003), there will be instances where such screening generates high scoring peptides that fail to bind optimally to MHC class I molecules when tested *in vitro*. Moreover, there are discrepancies among the different algorithms used for the computational mapping of epitopes. This can be seen not only with regard to ranking profiles of tested peptides but also with respect to the hierocracy of the peptides scoring in the top ten predicted MHC class I binders. In order to confirm the validity of these predictions, it is crucial that the peptides selected are tested *in vitro* to ascertain their ability to bind a particular MHC class I allele. In our study, the LFV glycoprotein-derived peptides synthesized and tested *in vitro* were selected based on ranking scores observed in both BIMAS and SYFPEITHI algorithms. As noted above, the protein sequence for LFV GP was also analyzed using the other available epitope prediction algorithms (SVMHC, RANKPEP, nHLAPred and PREDEP, see Table 2. High scoring peptides were compared to the high scoring peptides on BIMAS and SYFPEITHI algorithm. The peptides selected for testing in MHC stabilization assays were as follows: LLGTFTWTL, WLVSNGSYL, GLVGLVTFL YLISIFLHL, and SLYKGVYEL for the LFV GP and CMLDGGNML and GILKSILKV for LFV NP. The GP-derived peptides were selected from the group of ten highest

scoring peptides appearing in both BIMAS and SYFPEITHI algorithms analyses. Note that hereafter, peptides are named using the first three letters appearing in that peptide unless otherwise stated. The LLG peptide was chosen because it scored number 1 on BIMAS algorithm and it also scored in the high positions on most of the other algorithms including SYFPEITHI. The GLV peptide was chosen because it scored high on most of the algorithms, in some instances higher than LLG peptide, and it scored number 6 on the BIMAS algorithm. The WLV peptide was chosen because it's ranking was number 3 i.e. between LLG and GLV peptide, and WLV peptide did not appear on SYFPEITHI algorithm in the 10 highest ranking peptides. It would be thus interesting to evaluate if this particular peptide will be binding to the HLA-A2 *in vitro*. The YLI and SLY peptides were chosen because both scored among the ten highest scoring peptides on both BIMAS and SYFPEITHI algorithms. The SLY peptide scored number 1 on SYFPEITHI algorithm. The candidates for the CTL epitopes derived from the nucleoprotein the CML and GILK peptides were selected randomly among the ten highest scoring peptides on both BIMAS and SYFPEITHI algorithms.

Three out of the five selected GP-derived peptides (LLG, YLI, and SLY) that were identified among the top ten highest scoring peptides on both BIMAS and SYFPEITHI algorithms stabilized expression of HLA-A2 on T2 cells as evidenced by their MFI profiles which were compared with the no peptide control (Figure 16, 17). Similarly, a negative control peptide (IAG) had no measurable effect on HLA-A2 expression on T2 cells (Figure 18). The LLG and SLY peptides caused a three-fold MFI shift and the YLI peptide caused a two-fold MFI shift as compared with the no peptide control (Figure 16). Thus, based on the criterion used to identify peptides as high-binders

(see Materials and Methods), we judged these three peptides to be candidates for further analysis. In contrast, the WLV peptide which was selected based upon computational criteria using the BIMAS algorithm failed to show any measurable stabilization of HLA-A2 on T2 cells (Figure 16). In fact, this peptide caused an MFI shift to the left relative to the no peptide control. This can be attributed to the fact that this peptide (WLVSNGSYL) is highly hydrophobic and thus insoluble in solvents like DMSO or ethanol. In fact different solubility conditions have been tested (different concentrations of DMSO/water and DMSO 100%, also 100% and < % ethanol) and none has helped to solubilize the peptide. It is noteworthy that the BIMAS score for the WLV peptide was higher, ranking #3, as compared to the high-binding YLI and SLY peptides which ranked #7 and #8, respectively). Moreover, this nonbinding peptide did not appear among the highest scoring peptides observed using the SYFPEITHI algorithm. Given this disparity, it was thus interesting to further evaluate this nonbonding peptide *in vivo* (see Part III).

Neither of the computationally selected NP-derived peptides (CML and GILK) were able to stabilize HLA-A2 expression on T2 cells (Figure 16). Thus, they were excluded as candidates for further study.

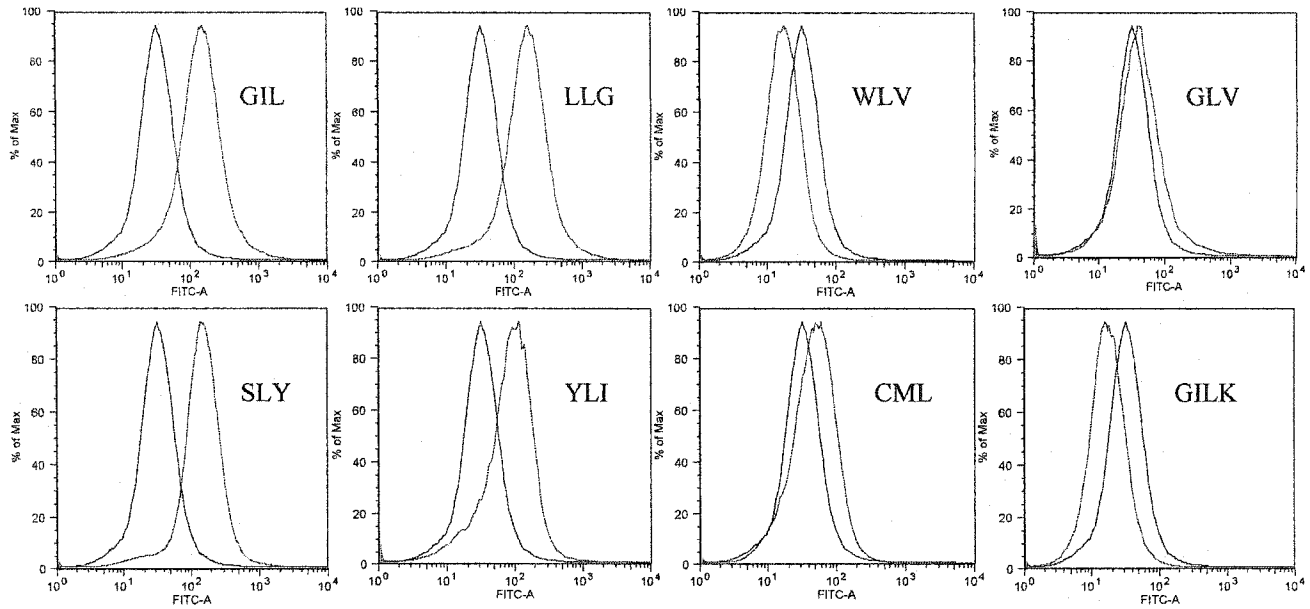


Figure 16. Peptide binding to HLA-A2 as assessed by MHC class I –stabilization. The ability of peptides to stabilize HLA-A2 on the surface of T2 cells was assessed by flow cytometric analysis after staining peptide-treated cells with anti-HLA-A.2 antibody. A positive control peptide, GILGFVFTL (GIL), was tested along with LV-GP peptides LLGTFTWTL (LLG), WLVSNGSYL (WLV), GLVGLVTFL (GLV), SLYKGVYEL (SLY), YLISIFLHL (YLI), and LV-NP peptides CMLDGGNML (CML) and GILKSILKV (GILK). Each peptide was used at 50μM concentration. A positive shift (to the right) at varying degrees was observed with all the peptides except WLV, where a negative shift (to the left) was observed.

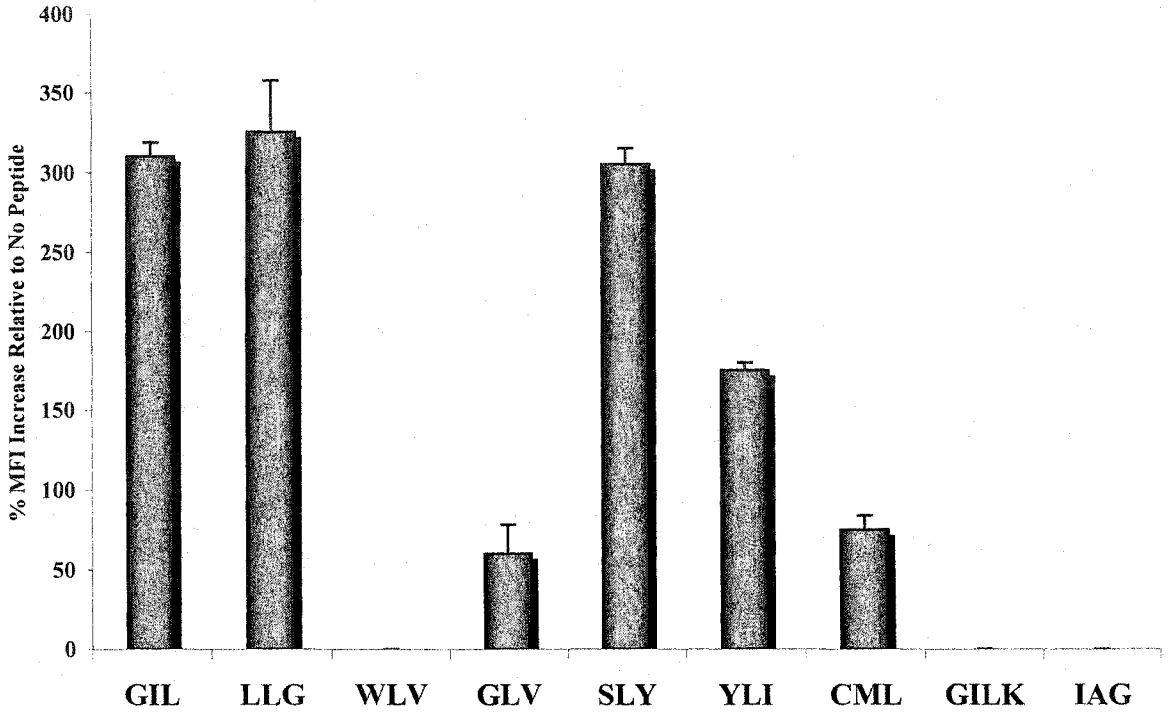


Figure 17. MHC class I stabilization by LFV GP and NP psptides as measured by MFI. Peptides were tested for their ability to stabilize MHC class I as described in Figure 8 and Materials and Methods. The percent MFI shown is the average percent increase relative to no peptide from three experiments. GILGFVFTL (GIL) and IAGNSAYEY (IAG) represent positive and negative controls, respectively. LFV GP peptides tested were as follows: LLGTFTWTL (LLG), WLVSNGSYL (WLV). GLVGLVTFL (GLV), SLYKGVYEL (SLY), YLISIFLHL (YLI), and LFV NP peptides tested were CMLDGGNML (CML) and GILKSILKV (GILK).

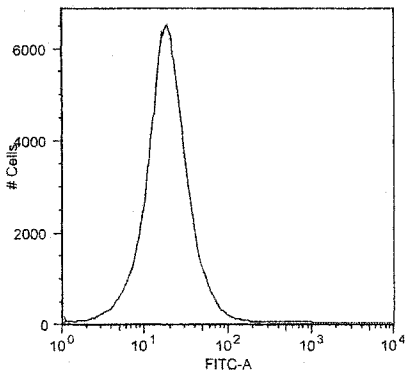


Figure 18. Peptide binding to HLA-A2.1 as assessed using the MHC class I – stabilization for negative control peptide. A negative control peptide IAGNSAYEY, was tested in this assay.

PART III. *In vivo* investigation of immunogenicity of LFV GP and NP peptides

We then asked whether the peptides shown to stabilize MHC class I *in vitro* would have immunogenic properties when tested in HLA-A2-expressing transgenic mice. Mice were immunized with the following selected single peptides: LLGTFTWTL, WLVSNGSYL, SLYKGVYEL, and YLISIFLHL. All peptides selected for the *in vivo* studies were derived form the LFV glycoprotein. None of the NP derived peptides was considered in the *in vivo* studies because both peptides tested in the *in vitro* studies failed to stabilize MHC class I molecules. Immune responses were assessed in several functional assays including the CTL assay, IFN-γ ELISPOT assay, intracellular IFN-γ staining and MHC class I tetramer staining.

As discussed below, using this mouse strain, we were able to show that all but one of the LFV GP CTL epitopes identified by *in silica* and *in vitro* analyses were, in fact, immunogenic. Mice were immunized as described in Materials and Methods.

1. IFN-γ ELISPOT assay

The IFN-γ ELISPOT assay was used as first step in verification of chosen peptides as antigenic epitopes. In this assay, two LFV GP peptides (LLGTFTWTL and SLYKGVYEL) were shown to contribute to the generation of peptide specific CTLs *in vivo* as measured by detection of the IFN-γ producing cells. Responses of primed lymphocytes to the corresponding peptides was comparable to responses generated following immunization with the positive control peptide (GILGFVFTL). IFN-γ production was found to be dependent on the concentration of the stimulating peptide. When the positive control stimulating peptide was used at 100μM, the number of IFN-γ-

producing cells per 2 x 10^5 cells was 51. Hence, the frequency of IFN-γ-producing cells was $255/10^6$. Similarly, the number of IFN-γ-producing cells for the LLGTFTWTL and SLYKGVYEL at the 100μM concentration was $225/10^6$ and $235/10^6$ respectively. The frequency of IFN-γ-producing cells decreased with decreasing peptide concentrations as shown in Figure 19. The responses of the primed lymphocytes to the corresponding YLISIFLHL LFV peptide was significantly lower then the response generated by the other two peptides. The number of IFN-γ producing cells generated in response to YLISIFLHL (100μM) was $155/10^6$ cells. This response decreased with the decreasing peptide concentration showing a dose-dependent profile. The WLVSNGSYL peptide failed to prime or stimulate the splenocytes for the IFN-γ production (see Figure19).

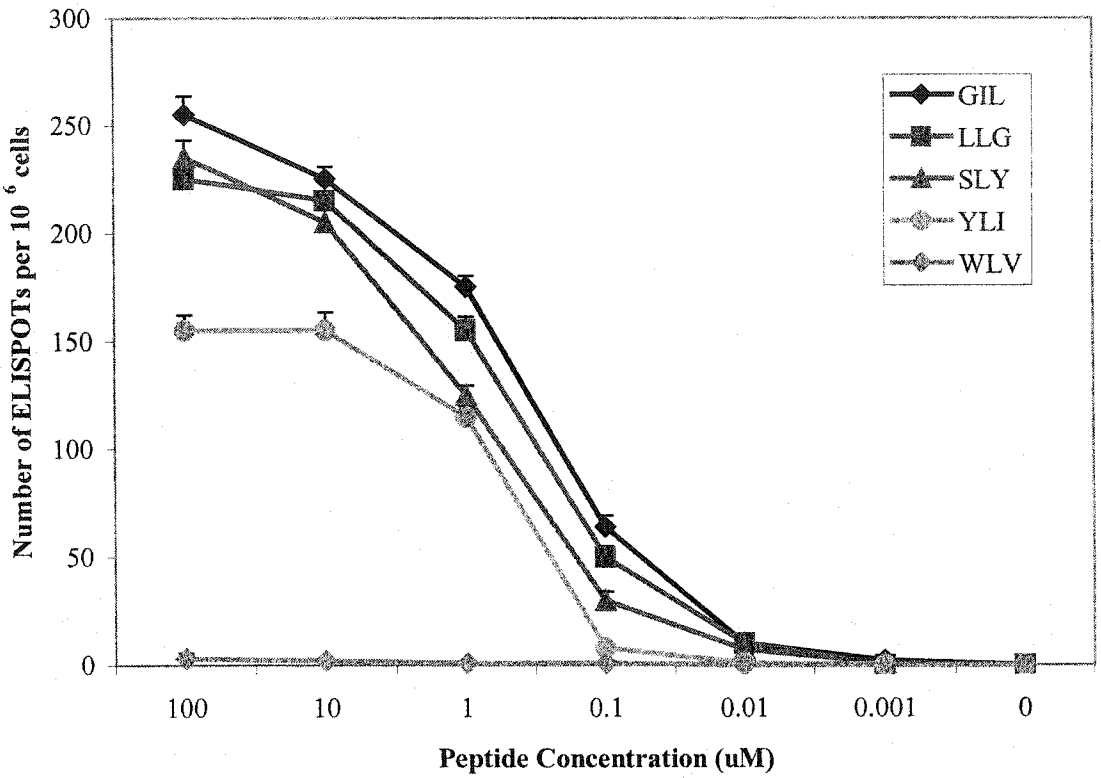


Figure 19. Enumeration of IFN-γ producing cells. The IFN-γ ELISPOT assay was performed as described in Materials and Methods using peptide-stimulated mice (n=3 per peptide tested). The following peptides were tested: LLGTFTWTL (LLG), WLVSNGSYL (WLV), SLYKGVYEL (SLY), and YLISIFLHL (YLI). All assays included positive control peptide GILGFVFTL (GIL). ELISPOTs (triplicates per data point) were enumerated using a dissecting microscope. Peptide-specific spot numbers reflect total spots minus background spots observed in the absence of peptide. Data points shown represent mean IFN-γ ELISPOTS/10^6 ± SD from a representative experiment that was repeated at least three times.

2. Intracellular IFN-γ staining

IFN-γ production was also measured intracellularly using an intracellular cytokine staining (ICCS) assay. When the lymphocytes were polyclonally stimulated with the PMA and inomycin, the percentage of IFN-γ-producing CD8 positive cells was found to range from 0.9 to 1.05%. The results shown are from a representative experiment showing the 1.05% value, which corresponds to 10.3% of CD8-positive cells. In each of these experiments, we observed a population of CD8-negative cells that were also producing IFN-γ. In the representative experiment shown, 1.04% of the CD8-negative cells were producing IFN-γ (Figure 20). When *in vivo*-primed lymphocytes from HLA-A2 mice previously immunized with GILGFVFTL were then stimulated *in vitro* with this positive control peptide, 0.1% of all the cells, and 0.6% of CD8-positive cells showed intracellular IFN-γ (Figure 21A). Background values for IFN-γ production among unstimulated cells were found to be 0.06% in the experiment shown (Figure 21A, 3). In contrast, LFV LLGTFTWTL peptide-stimulated cells resulted in the generation of 0.6% CD8-positive cells containing intracellular IFN-γ (Figure 21B). Similarly, 0.8% of CD8-positive cells IFN-γ-positive when stimulated with SLYKGVYEL peptide (Figure 21C) and 0.31% of the cells produced IFN-γ when stimulated with YLISIFLHL peptide (Figure 21D).

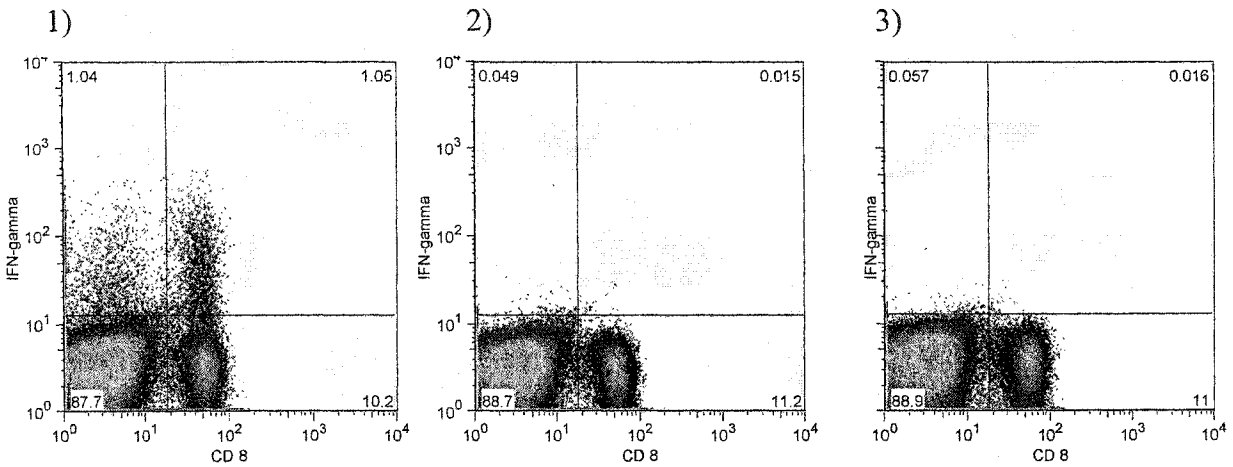


Figure 20. ICCS for IFN-γ using lymphocytes stimulated with PMA and inomycin. 1) This assay was performed as described in Materials and Methods. Briefly, PMA was used at 20 ng/ml and ionomycin was used at 1μM to polyclonally activate lymphocytes. Cells were subjected to flow cytometric analysis following ICCS. 2) As controls, an isotype control antibody the FITC Rat IgG_1 was used to stain peptide stimulated cells and 3) monoclonal α-IFN-γ antibody was used to stain unstimulated cells.

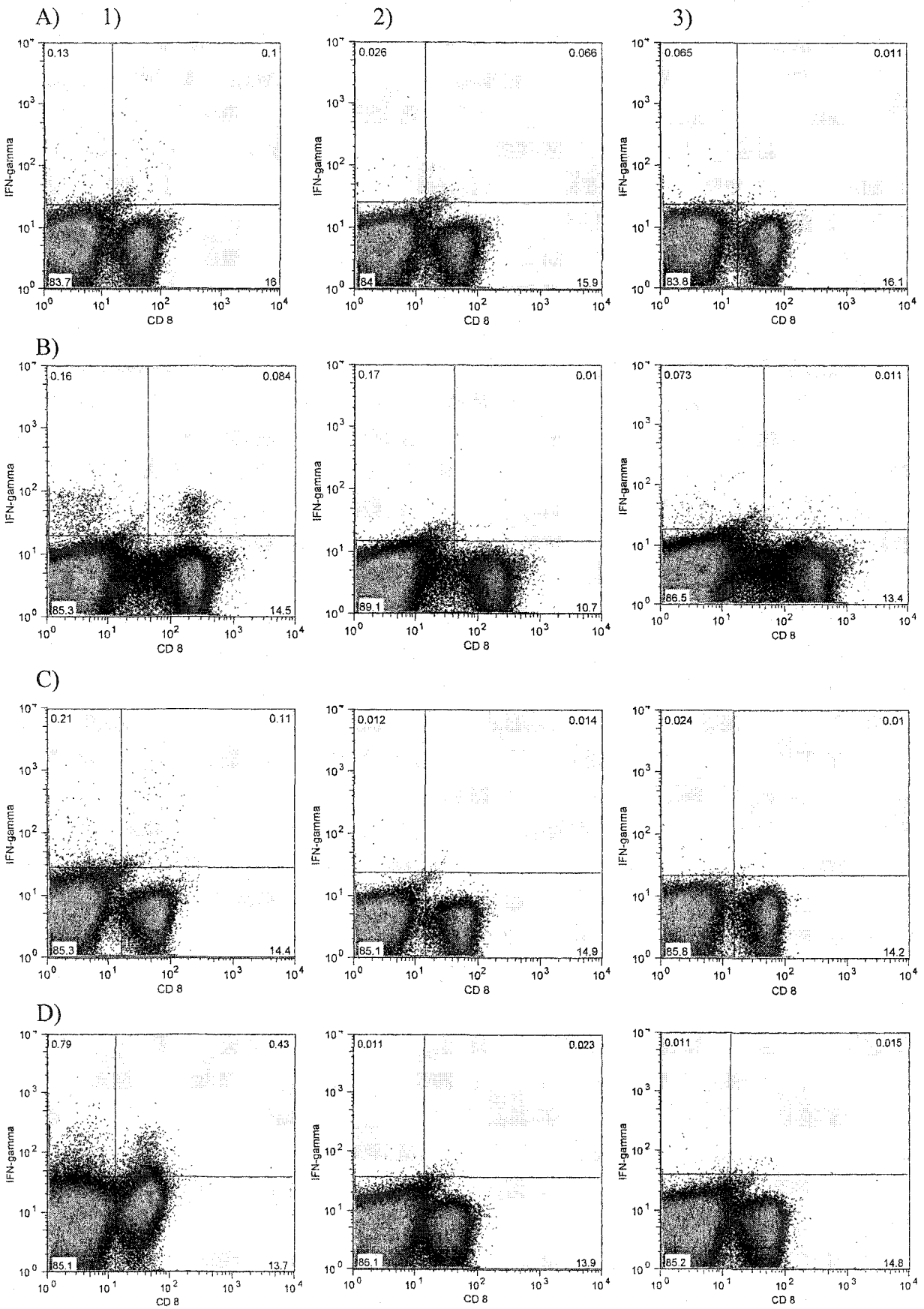


Figure 21. Enumeration of CD8-positive IFN-γ producing cells by ICCS. Splenocytes and lymph node cells were pooled from peptide-stimulated mice (n=3 per peptide tested) and then stimulated *in vitro* with 10 μM of corresponding peptides. (A) GILGFVFTL: positive control, (B) LLGTFTWTL, (C) SLYKGVYEL, and (D) YLISIFLHL. Cells were cultured in complete IMDM medium containing 10% T-Stim overnight at 37°C As described in Materials and Methods. The following day, 1μl/ml of monensin was added and the cells were incubated for another 6 hr, washed, and stained with monoclonal α-CD8 PerCP in 1% FBS/PBS. The cells were then permeabilized and stained with monoclonal α-IFN-γ antibody and analyzed flow cytometrically. As controls, an isotype control antibody the FITC Rat IgG_1 was used to stain peptide stimulated cells. And monoclonal α-IFN-γ antibody was used to stain unstimulated cells.

3. Non-radioactive CTL assay

Effector cells generated in peptide-primed HLA-A2.1 transgenic mice were tested for their cytolytic activity by measuring their ability to kill peptide-pulsed T2 target cells expressing the HLA-A2.1 molecule. Only peptides that passed the MHC class I stabilization criterion were tested for immunogenicity. These included LLGTFTWTL, SLYKGVYEL, and YLISIFLHL all derived from the glycoprotein of LFV. The influenza peptide (GILGFVFTL) used in our *in vitro* assays was also tested as a positive control immunogen. As shown in Figure 22, the effector cells generated after 5 days of *in vitro* stimulation of the *in vivo-peptide*-primed splenocytes and lymph node cells killed peptide-pulsed target cells in a dose-dependent fashion using various effector: target ratios. Cells stimulated with LLGTFTWTL peptide showed a specific lysis of 30% at the 40:1 effector:target ratio using LLGTFTWTL pulsed target cells. In case of the SLYKGVYEL peptide, effector cell specific lysis at this effector: target ratio was found to be 17%. Finally, effector cell specific lysis using YLISIFLHL-primed cells was found to be 9% at the 40:1 effector: target ratio. The effector cell specific lysis using cells stimulated with the positive control peptide (GILGFVFTL) was 37% at this effector: target ratio (Figure 22).

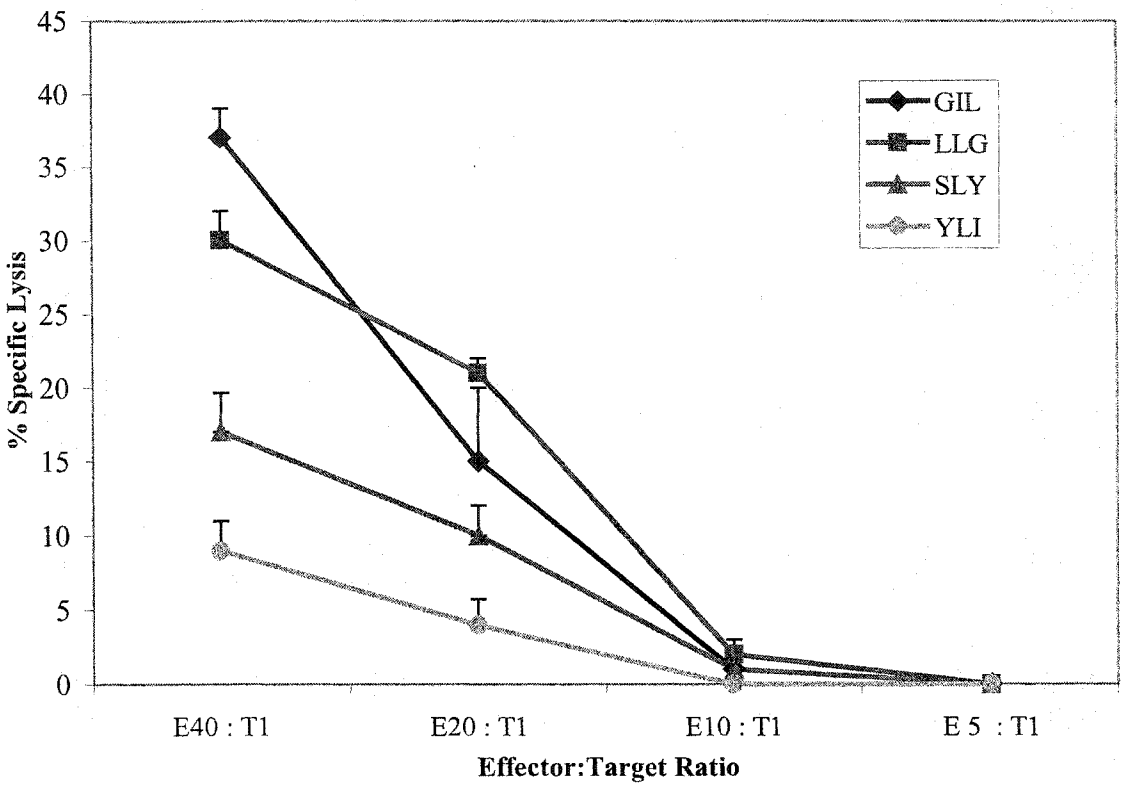


Figure 22. Specific lysis of peptide-pulsed cells by peptide-primed effector cells from HLA-A2.1 transgenic mice. A non-radioactive CTL assay was performed after 5 days of *in vitro* stimulation using spleen and draining lymph node cells from HLA-A2.1 transgenic mice (n=3 per peptide tested) immunized with various LFV peptides as described in Materials and Methods. The following peptides were tested: LLGTFTWTL (LLG), SLYKGVYEL (SLY), and YLISIFLHL (YLI). A positive control peptide GILGFVFTL (GIL) was included in all assays. Briefly, target T2 cells (5 x 10^5/ml) were incubated overnight with 10μM of appropriate peptides in IMDM medium containing 10% FBS in 12-well plates. The cells were washed, and then peptide-pulsed and non-pulsed (control) T2 cells were added at 1 x 10^4 cells per well. Effector cells were added to these target cells to achieve effector: target ratios ranging from 40:1 to 5:1, respectively. Following a four-hour incubation, supernatants from triplicate wells for each effector: target ratio were transferred to separate wells of a flat-bottom 96-well plates for colorimetric assay. The absorbance was recorded using a Bio-Rad ELISA reader and the percentage of specific LDH release (cell lysis) was calculated. Data points represent mean percent specific lysis ± SD from a representative experiment that was repeated at least three times. The average background lysis for all the peptides was 10%.

4. Tetramer staining

The frequency of LLGTFTWTL-specific CTLs in peptide-immunized mice was analyzed by tetramer staining using PE-conjugated HLA-A2.1- LLGTFTWTL tetramers. Tetramer staining of pooled spleen and lymph node cells revealed a small, but significant population of tetramer-positive cells in the LLGTFTWTL-primed mice as compared with unimmunized mice (Figure 23).

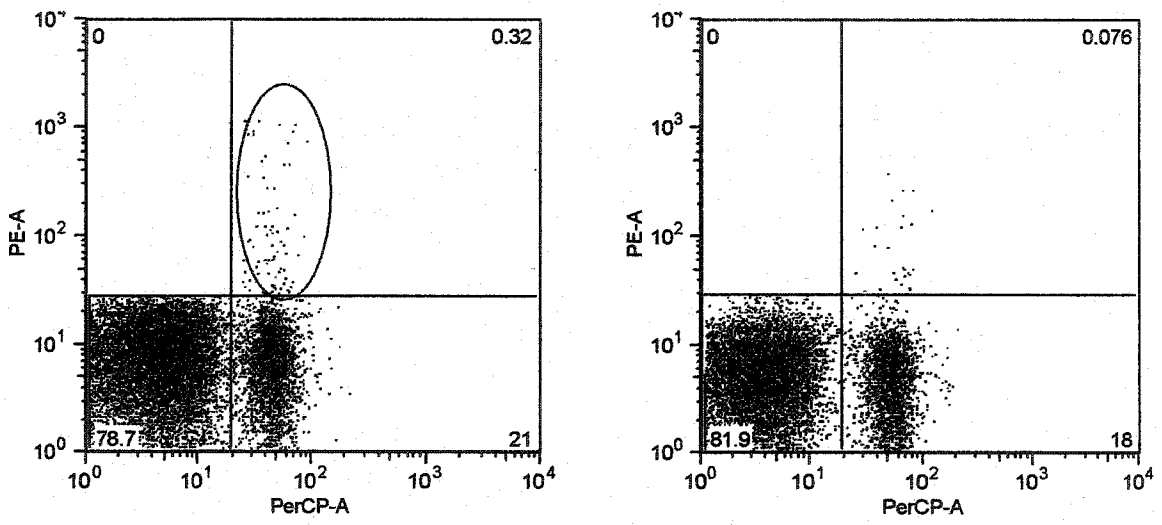


Figure 23. Enumeration of Lassa LLG Peptide Binding CTLs by Tetramer Staining Tetramer-positive antigen-specific T lymphocytes were analyzed in spleen and draining lymph nodes from Lassa (LLG) peptide-immunized and unimmunized-control HLA-A2.1 transgenic mice. Prior to staining, LLG-primed cells were stimulated overnight with 10μM LLG peptide in T-Stim-containing medium. Control cells were cultured in T-Stim-containing medium alone. Lymphocytes were stained with the LLG HLA-A2.1 PE-conjugated tetramer (1:800), and PerCP- conjugated anti-CD8 monoclonal antibodies.

Table 17. Collective ranking of LFV peptides identified as immunogenic candidates.

Control or LFV peptide	BIMAS Rank[1]	Class I binding[2]	# IFN-γ ELISPOTs[3]	IFN-γ ICCS[4]	CTL Responses [5]	Collective rank[6]
GILGFVFTL	9	305	225	6	37	582
LLGTFTWTL	10	325	215	6	30	586
WLVSNGSYL	8	0	0	Not tested	Not tested	
GLVGLVTFL	5	50	Not tested	Not tested	Not tested	
YLISIFLHL	4	175	155	31	9	374
SLYKGVYEL	3	300	205	8	17	533
CMLDGGNML	10	60	Not tested	Not tested	Not tested	
GILKSILKV	8	0	Not tested	Not tested	Not tested	

[1] The values for computational rank were inverted to achieve the values noted (peptide scoring as number 1 has value of 10). The positive control peptide ranked as number 2 on the BIMAS algorithm (see Table 19).

[2] MHC class I binding scores represent % MFI increase values.

[3] The number of IFN-γ producing cells is shown per one million cells at the 10μM peptide concentration.

[4] Intracellular staining, % of IFN-γ and CD8+ cells. The values for the % of IFN-γ producing cells were calculated as the percent of the IFN-γ producing cells among the CD8 positive cells only X 10.

[5] The CTL rersponse values are expressed as % specific lysis at 40:1 effector: target ratio

[6] The collective rank value is a sum of all the partial values from class I binding %MFI increase, IFN-γ ELISPOT, % specific lysis, and computational rank.

Table 18. Summary of collective ranking of GP-derived LFV epitopes identified as immunogenic peptides.[1]

	Relative ranking					
LFV CTL epitope	BIMAS	Class I binding	# IFN-γ ELISPOTs	IFN-γ ICCS	CTL Responses	Collective rank
LLGTFTWTL	1	1	1	3	1	1
SLYKGVYEL	3	2	2	2	2	2
YLISIFLHL	2	3	3	1	3	3

[1] Rankings shown are derived from the collective ranking scores calculated in Table 17.

Rank	Start Position	Subsequence Residue Listing	Score (Estimate of Half Time of Disassociation of a Molecule Containing This Subsequence)
1	3	LLTEVETYV	2666.276
2	58	GILGFVFTL	550.927
3	134	RMGAVTTEV	50.232
4	59	ILGFVFTLT	24.07
5	51	ILSPLTKGI	17.736
6	130	LIYNRMGAV	15.555
7	164	QMVTTTNPL	15.428
8	138	VTTEVAFGL	11.994
9	60	LGFVFTLTV	10.852
10	180	VLASTTAKA	8.446

Table 19. Influenza protein (CAA24282) containing positive control peptide, analyzed by BIMAS algorithm. The protein sequence of influenza protein was obtained from Entrez database. The protein sequence was analyzed computationally using the BIMAS prediction server. The total protein sequence was parsed into peptides, nine amino acid in length, each overlapping by eight amino acids. The highlighted peptide is the experimentally tested positive control peptide.

DISCUSSION

Experimental strategies that utilize computational methods combined with *in vitro/in vivo* studies have proven to very useful for the identification of immunogenic T-cell and B-cell epitopes. Such epitopes are commonly identified using defined antigens that are typically derived from pathogenic organisms, tumors-specific proteins, or self-antigens involved in autoimmune diseases. In the current investigation study, we employed the use of computational methods to map potential immunogenic CTL, T-helper and B-cell epitopes for the LFV. After the *in silica* identification of the CTL epitopes, we evaluated the immunogenic potential of the candidate peptides *in vitro* and *in vivo*. Using a class I MHC allele-specific analysis and the BIMAS and SYFPEITHI algorithms we were able to identify epitopes within the glycoprotein of LFV. Other algorithms were also used for comparison purposes and to aid in the selection of peptide sequences to be synthesized for subsequent *in vitro/in vivo* studies. We confirmed the ability of specific peptides to bind class I MHC using a class I MHC- stabilization assay. The peptides selected for testing in MHC stabilization assays were as follows: LLGTFTWTL, WLVSNGSYL, GLVGLVTFL YLISIFLHL, and SLYKGVYEL for the LFV GP and CMLDGGNML and GILKSILKV for LFV NP. Three out of the five selected GP-derived peptides (LLG, YLI, and SLY) that were identified among the top ten highest scoring peptides on both BIMAS and SYFPEITHI algorithms stabilized expression of HLA-A2 on T2 cells as evidenced by their MFI profiles which were compared with the no peptide control (Figure 16). Neither of the computationally selected NP-derived peptides (CML and GILK) were able to stabilize HLA-A2

expression on T2 cells (Figure 16). Based on the selection criteria used to define possible candidates for CTL epitopes (see Materials and Methods), peptides LLG, YLI, and SLY were selected for testing *in vivo*. These peptides demonstrated high class I MHC binding profiles in MHC stabilization assays. I*n vivo* studies showed that peptides LLG and SLY were highly immunogenic and peptide YLI was moderately immunogenic in HLA-A2 transgenic mice as demonstrated by IFN-γ ELISPOT, ICCS, and CTL assays as well as by class I tetramer staining. We conclude from these findings that the LFV peptide sequences that fulfilled the *in silica, in vitro*, and *in vivo* criteria that we defined to qualify them as immunogenic epitopes are, in fact, immunogenic using the HLA-A2 animal model. Collectively, these findings suggest that the mapped CTL epitopes have the potential to generate protective immune responses against LFV.

The algorithms we used to map these LFV CTL epitopes are open access computational resources accessible on the World Wide Web. They have been used with success in recent years for the identification of immunogenic epitopes from a wide range of antigens. As noted earlier in this thesis, BIMAS ranks potential HLA- binders according to predicted half-time dissociation rates of MHC-peptide complexes. While the motif matrix for HLA-A2 is based solely on binding studies, other matrices have been deduced from reported motifs, regardless of method used for their identification. SYFPEITHI uses motif matrices deduced from refined motifs based exclusively on sequence and single peptide analyses of natural ligands.

Huang *et. al.*, (2004) have identified novel HLA-A2.1-restricted CD8 T-cell epitopes on hepatitis delta virus (HDV). They selected these epitope peptides using the SYFPEITHI database and screened those peptides identified in the MHC-stabilization

assay and then immunized HLA-A2.1 transgenic mice (the same strain of mice we used in our studies) with recombinant DNA encoding the two selected peptides. Tetramer staining confirmed that 0.9% of CD8+ splenocytes were specific for the tested peptides. These results were confirmed by showing that peptide-specific effector cells generated CTL responses to the relevant peptides. The mapped CTL epitopes were also detected in two patients chronically infected with HDV in the absence of active HDV replication.

Using the same algorithm (SYFPEITHI) Turcanova and Höllsberg (2004) have identified novel immunodominant HLA-B7-restricted CTL epitopes derived from Epstein-Barr Virus helicase-primase-associated protein. In this study, the specificity of selected peptides for HLA-B7 was assessed by competitive binding assays and was found to correlate with scores assigned by the predictive algorithm. Bulk PBMC from HLA-B7-positive individuals previously infected with Epstein-Barr Virus Nuclear Antigen-1 (EBNA-1) were pulsed with appropriate peptides and used as APCs. CD4+ and CD8+ cells were used as responder cells in an IFN-γ ELISPOT assay and results confirmed the computational identification of immunodominant HLA-B7-restricted CTL epitopes..

Lu and Celis (2000) used BIMAS and SYFPEITHI prediction algorithms to identify HLA-B7-restricted CTL epitopes present in tumor-associated carcinoembryonic antigen (CEA). The candidate peptides were evaluated for their capacity to induce CTL responses *in vitro* using lymphocytes from HLA-B7-positive donors. The results showed that one of the peptides selected based upon computational prediction studies, i.e., CEA9 632-640 (IPQQHTQVL) efficiently induced primary CTL responses when dendritic cells were used as APCs *in vitro*. These CTLs were able to kill tumor cells expressing HLA-B7, as measured in the cytotoxicity assay. Other investigators have reported that

computational approaches to epitope mapping are not always successful. Andersen *et. al.*, (2000) measured class I MHC binding for 84 peptides selected on the basis of the presence of peptide binding motifs experimentally demonstrated for HLA-A2.1, HLA-B2705, HLA-B3501, HLA-B0702 and one murine molecule H-2K^{k}. The peptides tested were derived from viral proteins and oncogene products. Using the same class I binding assay as we used in our studies, they compared peptide MHC-stabilization with eptiope predictions obtained using BIMAS and SYFPEITHI algorithms. They concluded that there was no strong correlation between actual and predicted binding when using computer algorithms and that the peptide binding assay is an important step in the identification of CTL epitopes.

Fortunately, the number of reports demonstrating the utility of computational approaches in mapping potential immunogenic epitopes far exceeds those, which have not successfully validated these methods. Vonderheide *et. al.*, (1999) have successfully used BIMAS algorithm to identify HLA-A2.1 restricted CTL epitopes from the telomerase catalytic subunit (hTERT) a widely expressed tumor-associated antigen. Sun *at al.*, (2000) identified HLA-A2 –restricted T-cell epitope from the tyrosinase-related protein 2 (TRP2) melanoma antigen using the SYFPEITHI algorithm. In this study, six computationally identified peptides contained in TRP2 were synthesized and tested for binding affinity to HLA-A2.1 molecule using the class I MHC stabilization assay. Selected peptides were then pulsed onto autologous dendritic cells and used to stimulate peptide-specific CTLs in CD8-enriched T cells isolated from peripheral blood of HLA-A2.1 positive healthy donors or melanoma patients. Among the peptides tested, the one that scored the highest as a class I stabilizer, namely TRP-2 288-296 (SLDDYNHLV),

elicited specific CTLs in some of the patients and healthy donors. The CTLs generated in cells isolated from the patients and healthy donors efficiently recognized HLA-A2.1^{+} TRP-2^{+} cells, as measured by cytokine production and specific lysis (CTL assays).

Several investigators have addressed the validation of computational algorithms when they computationally mapped already identified CTL epitopes. Simmons *et. al.*, (2004) identified murine H-2^{d} and H-2^{b} restricted CTL epitopes in Ebola virus nucleoprotein. They immunized animals with recombinant vaccine vectors expressing Ebola proteins and screened for T-cell epitopes using overlapping peptide libraries. The epitopes identified stimulated the production of IFN-γ in CD8+ splenocytes from immunized mice. One of the experimentally predicted H-2^{d}-restricted epitopes identified in this study had a high-predicted binding affinity when analyzed using the SYFPEITHI algorithm. Two other experimentally identified epitopes had a low predicted binding affinity when analyzed by the same algorithm.

Beside the two most commonly used algorithms (BIMAS and SYFPEITHI), there are several other algorithms used for computational mapping of epitopes. However, minimal supportive data validating the efficacy of these algorithms exists. These include algorithms such as the artificial neural networks, vector machine, and structure-based algorithms. With the exception of structure-based algorithms, the other algorithms often utilize the quantitative matrices derived from BIMAS or the SYFPEITHI programs. They differ from BIMAS and SYFPEITHI with respect to the machine learning technique they utilize. Finally, commercial algorithms are also available which, in contrast with the others noted above, are not publicly available (e.g. EpiMatrix; De Groot *et. al.*, 1997, http://www.brown.edu/Research/TB-HIV_Lab/epimatrix.html).

It is noteworthy that while epitope mapping algorithms are quite useful, there are some drawbacks to using this approach to identify immunogenic peptides. One major drawback is a high failure rate. Even if the peptide in question binds well to the appropriate MHC molecule and peptide-stimulated T cells recognize target cells loaded with the synthetic peptide *in vitro*, there is no guarantee that the candidate peptide will be stimulating protective immune responses *in vivo*. Cytotoxic killing may occur *in vitro* but not in the *in vivo* system. One possible explanation for how this might happen is that the peptide in question may not be generated by the cell's processing machinery. Another possibility is that the peptide-specific T cells generated *in vivo* may have low affinity TCRs that thwart their ability to optimally recognize MHC/peptide complexes. Finally, another possible explanation is that the peptides may have low solubility properties *in vivo*.

To minimize the number of false positive candidate peptides identified by prediction algorithms, various groups have used computational strategies that integrate prediction formulations with those that consider proteasomal cleavage preferences. This approach helps to selectively fine-tune the potential epitopes mapped. The first published proteasomal prediction algorithm was based on the experimental cleavage data of only seven naturally occurring peptides (Holzhütter *et. al.*, 1999). The same group later published the proteasome model called FragPredict, which was based on kinetic parameters (Holzhütter and Kloetzel, 2000; http://www.mpiib-berlin.mpg.de/MAPP/). Kuttler *et. al.* (2000) created a proteasomal model based on 20S proteasome cleavages in unmodified enolase-1, a protein containing 436 amino acids. This model (PAProC) is available as a public access resource (Nussbaum *et. al.*, 2001;\ http://www.paproc.de).

One CTL epitope prediction algorithm termed nHLAPred incorporates proteasomal filters into the analysis software. We took advantage of PAProC when we initially set out to analyze the glycoprotein of LFV for candidate CTL epitopes. Thus, we enabled this proteasomal filter in the hope of facilitating the identification of immunogenic epitope peptides for LFV GP. The objective was to select only for peptide sequences that would be preserved following digestion within the proteasome. When we performed this analysis using PAProC, predicted cleavages occurred over one hundred times in a 491 amino acid long sequence of LFV GP thus eliminating all peptides >9 amino acids in length as candidate epitopes. It should be noted that proteasomal prediction algorithms are only accurate approximately 60% of the time. Thus, in their current form, their usefulness is dubious, at best since there is a great risk of overlooking potential immunogenic epitopes due to incorrect predictions in protein cleavages occurring in the proteasomes.

Because of the inherent caveats in the use of proteasomal prediction algorithms, we did not take into consideration proteasomal cleavage predictions in the identification of the CTL epitopes for LFV. Another argument for eliminating proteasomal cleavage prediction filters in our analysis of LFV GP and NP is that for CTL epitopes, it has been shown that potential proteasomal cleavage sites in an epitope-containing sequence do not automatically abolish presentation of that epitope (Stoltze *et. al.*, 2000). Furthermore, proteasomal enzymes do not always carry out the same cleavages in two identical amino acid chains, thus, they sometimes treat individual protein molecules differently. This is illustrated by the pattern of overlapping peptides produced in proteasomal digestion experiments (Kloetzel 2001).

Future comprehensive computational predictions

In addition to computational algorithms used for epitope or proteasomal cleavage predictions, computational software has been developed which simulates the dynamics of viral infection, proliferation and differentiation of naïve CD8+ T cells. Chao *et. al.* (2004) have designed software for a stochastic model of cytotoxic T cell responses (http://www.cs.unm.edu/~dlchao). This model incorporates antigen- and affinity-dependent stimulation of naïve cells as well as an antigen-dependent programmed proliferative response and differentiation into effector cells. It allows one to examine simulated responses at many levels, from the total T cells count to the number of responding cells present in each T-cell clone. The model takes into consideration the effects of TCR/peptide/MHC affinity and allows for qualitative and quantitative evaluation of cellular immune responses. This approach is based on "digit strings" and "strings match" rules to represent antigens and affinities. Each antigen in the simulation is associated with one or more digit strings representing its epitopes. Each epitope has a scalar value representing its surface expression density and each MHC allele is represented by a different random string. Finally, each CTL clone has a string representing its TCR. A string match rule defines the binding affinity between the TCR and the MHC-peptide complexes. The avidity of a TCR for an epitope is a function of its affinity, determined by the epitopes density on the cell surface.

Chao *et. al.,* (2004) also used their software to simulate cellular immune responses to the lymphocytic choriomeningitis virus (LCMV) in mice. It is noteworthy that LCMV is an Arenavirus closely related to LFV as discussed earlier in this thesis.

They found that their simulated results agreed with those generated experimentally. It had been previously shown that increasing the number of naïve cells by adoptive transfer results in earlier and larger response to infection with LCMV (Ehl *et. al.,* 1998). As the number of naïve cells was increased by 1000-fold, mice exposed to LCMV generated peak responses to viral infection occurred 1-2 days earlier and the viral load was reduced by about 2 logs. Similarly in their simulation model, when they increased the number of naïve cells from 50 to 50,000, peak simulated virus responses were virtually observed one day earlier with viral loads approximately 2 logs smaller as compared with control responses. Furthermore, their computer-simulated results also correlated with regard to the magnitude of the overall response to LCMV in studies designed to examine the consequences of removing the antigen upon the initiation of anti-viral responses. In yet another study, Mercado *et. al.,* (2002) used antibiotics to eliminate the *L. monocytogenes* infection in mice 24h after inoculation. Peak T-cell responses occurred at the same time in the antibiotic-treated mice and in the non-treated mice, and the elimination of the infectious agent only caused a small reduction in the magnitude of the response. In the study by Chao *et. al.,* (2004), LCMV was eliminated at 36h post infection. The magnitude of the response (as measured by the relative number of virus-specific CTLs) was almost unaffected by the removal of antigen and peak response occurred one day earlier in the mice in which LCMV was eliminated at 36h post infection.

Another useful feature of the immunological modeling software designed by Chao *et. al.,* (2004), is that the software can be used to predict the affinity between a TCR and an epitope. The clonal composition of T cell responses was studied with a single viral epitope and two T cell clones with different avidities. The peak of the high-avidity

clone's response was >1 log higher as compared with the low avidity clone. They concluded that such simulation software might be useful in the virtual evaluation of vaccine efficacy.

It is possible that in the future, methods that combine computational simulation models with epitope prediction algorithms will emerge. Such computational resources may allow us to predict the outcome of immunizations with particular antigens in individuals with specific MHC phenotypes.

Mapping of Class II MHC-binding Epitopes

Class I epitope prediction methods identify peptide motifs that comply with the "rules" of peptide-class I ligation. The optimal length of class I MHC ligands is typically nine amino acids. Well-characterized class I motifs contributed to the development of class I prediction algorithms before the development of algorithms designed to map class II MHC epitopes. In the latter, the elucidation of peptide motifs is more difficult to accomplish due to the variable length of the binding peptides and the more degenerate nature of anchor positions. While some alleles, such as HLA-DRB1*0405, show a strong preference for certain related amino acids in the anchor positions that is comparable to MHC class I motifs (Rammensee *et. al.*, 1999), other alleles make the definition of primary anchor positions virtually impossible. The first matrix-based prediction algorithm for class II MHC (HLA-DRB1*0401) was based on side-chain scanning experiments using polyalanine-based peptide libraries (Hammer *et. al.*, 1994) (discussed in Results section). Computational algorithms that were subsequently generated using that experimental data, serve as a useful tool in mapping of T-helper cell epitopes.

Like CD8 T-cell epitopes, CD4 T-cell epitopes are commonly identified in the context of infectious diseases, cancer, autoimmune diseases, and allergy. By combining the computational selection of candidate epitopes with *in vitro* and *in vivo* functional analyses, numerous HLA-DR restricted peptides have been identified. The advantages of preselecting candidate epitopes using computational predictions are primarily related to savings in cost and time. The ProPred algorithm, which utilizes the virtual matrices originally incorporated into the TEPITOPE software has been used with success to map class II MHC-binding epitopes. The validity of this algorithm has been demonstrated in a number of studies as described below.

Cochlovius *et. al.*, (2000) have shown *in vitro* and *in vivo* proliferation of CD4 T-cells primed with dendritic cells (DC) loaded with melanoma-associated glycoprotein 100 (gp100). Dendritic cells from nine donors with distinct HLA-DR haplotypes, including patients with malignant melanoma, were loaded with seven TEPITOPE-predicted peptides. PBMC from all nine donors generated CD4 T-cell responses to one of the DC-presented peptides. These CD4 T-cell responses were also confirmed *in vivo* by showing that tumor growth in the SCID mice was reduced following transfer of PBMC together with peptide-loaded DC (Cochlovius *et. al.*, 2000).

In another study, investigators applied TEPITOPE analysis to identify T-cell epitopes expressed by *Mycobacterium tuberculosis* (Panigada *et. al.*, 2002). Selected peptides induced proliferative responses of CD4 T-cells from patients infected with *M. tuberculosis*. These peptides had binding affinity for the twelve different HLA class II alleles expressed on the patients' antigen-presenting cells. One of the peptides tested was shown to elicit CD4 T-cell proliferation in 50% of the subjects tested.

In a recent study by ter Meulen *et. al.*, (2004), overlapping peptide pools as well as single peptides derived from LFV GP were screened to map CD4 T-cell epitopes. T-cell clones generated from PBMC obtained from LFV sero-positive individuals living in Western Africa were shown to proliferate in response to particular pools of peptides. Three class II LFV epitopes were identified as follows: (1) GP282-294, (2) GP289-301, and (3) GP394-406. Epitopes 1 and 2 overlap partially with each other. Each epitope was capable of inducing proliferative responses of PBMC from the donors used to prepare T-cell clones. Using the class II MHC epitope prediction software, ProPred, the investigators then performed computational analysis of LFV GP. They found a good correlation between experimentally observed and computationally predicted HLA-DR-restricted peptides in the case of the epitope contained within the GP282-301 amino acid sequence. It is noteworthy that we, too, identified this epitope as a candidate for promiscuous class II binding as shown in Table 10 and 13 in the Results section. Promiscuous binders are those that can bind to several HLA-DR alleles. They typically cross-react with approximately 50 different HLA-DR alleles, which accounts for > 90% of the class II MHC (HLA) molecules known to be expressed within the species. . Ultimately, such candidate epitopes can be tested for their ability to induce proliferative responses in PBMC from individuals previously infected with LFV. As noted above, in the study recently reported by ter Meulen *et.al.* (2004), the LFV epitope they identified, and which we confirmed computationally in the present study (GP282-301) did, indeed, have the ability to stimulate T cell responses in LFV-infected individuals. It is important to note that this study was performed using a limited number of PBMC donors, and thus the variability of the restricting MHC determinants was also limited. In order to

confirm that the T-cell epitopes identified in our study are, indeed, promiscuous with regard to their ability to bind a broad spectrum of class II HLA alleles, PBMC from infected individuals from non-endemic areas would be needed. Clearly, this is not an option given the rare incidence of Lassa fever outside of Western Africa.

Mapping of B-Cell Epitopes

Like CD4 T cell epitopes that bind to class II MHC molecules, B-cell epitopes are also commonly identified in the context of infectious diseases, autoimmune diseases, and allergy. B-cell epitopes may range in complexity from simple linear sequences of amino acids to complex quaternary structures. These linear or continuous epitopes may comprise a consecutive stretch of up to ten amino acids, sometimes referred to as the *primary epitope*. Secondary B-cell epitopes consist of three-dimentional structures, such as α-helices or β-turns and are commonly identified using computer-based structure analyses and computer algorithms (Blüthner *et. al.*, 2000; Blüthner *et. al.*, 2002). In the case of tertiary epitopes, the amino acids comprising the epitope can be distributed over distant parts of the protein's primary sequence. In quartenary epitopes, the subunits of a macromolecular complex interact to form a structure recognized by antibody (Mahler *et. al.*, 2003). All the epitopes mentioned may be hidden in the structure of the native antigen, thereby forming hidden or *cryptic epitopes* (Deshmukh *et. al.*, 1999). The degree of crypticity depends upon whether the binding site is hidden trough secondary, tertiary, or quarternary structure formation.

Methods for B-cell epitope identification and evaluation in which antigenic protein sequences are probed with human or animal antisera, include immunoblot,

ELISA, immunoprecipitation, and phage display. Schulze *et. al.*, (2003) have identified B-cell epitopes within the fibronectin-binding domain of the SfbI protein of *Streptococcus pyogenes*. One linear B-cell epitope was recognized by BALB/b and BALB/k mice and two epitopes were found in BALB/c animals. The animals were immunized intranasaly with a polypeptide encompassing the fibronectin-binding domain. ELISAs were carried out using streptavidin-coated microtiter plates, which were incubated with 16-mer biotinylated overlapping peptides spanning the first two fibronectin-binding repeats of the SfbI protein with an offset of 4 amino acids. Plates were then incubated with sera from immunized animals, and bound antibodies were detected. To confirm the results obtained from the ELISA, immunoblots were performed using cellulose membranes spotted with 15-mer peptides encompassing four fibronectin repeats. The membranes were incubated with sera from immunized animals and specific antibody binding was detected.

In our studies, we also used computational methods to identify candidate B-cell epitopes for LFV GP and NP. We applied prediction algorithms in order to map β-turns (Chou-Fasman) and hydrophobicity of the solvent-exposed sequences (Kyte-Doolittle) of the proteins analyzed. The results of the β-turn and hydrophobicity analyses were compared and the areas of high turn tendency and high hydrophilicity were chosen as the antigenic sequences of interest - possible B-cell epitope candidates. As discussed in the Results section, the maximum positive deflection (β-turns) and maximum negative deflection (hydrophilicity) of the chosen regions for both glycoprotein and nucleoprotein sequences correlate very well. Since computational algorithms are typically only 50% to 70% successful in identifying B-cell epitopes (Carter, 1994), only functional studies

measuring the immunogenicity and specificity of generated antibodies can identify actual antigenic sequences.

Experimental Verification of Mapped CTL Epitopes

Different approaches can be used to experimentally verify predicted CTL and T-helper epitopes, with regard to analysis of binding properties and analysis of T-cell recognition. Binding of synthetic peptides to MHC molecules is evaluated *in vitro* mainly using competitive binding assays (Sette *et. al.*, 1994; Olsen *et. al.*, 1994), reconstruction of MHC-peptide complexes on cell surface (Stuber *et. al.*, 1994), or, the method used in the current study - stabilization of MHC-peptide complexes on TAP deficient cells.

In competition assays, the peptides under investigation compete with a labeled reference peptide of intermediate affinity for binding to MHC molecules. Such assays are performed either using whole cells or purified MHC molecules in solution. In both cases, the amount of labeled MHC-bound peptide is measured. For MHC class II-binding assays, usually fluorescence dyes such as fluorescein or 7-amino-4methyl-coumarin-3-acetic acid (AMCA), are used for labeling the N-terminal amino acid of the MHC ligand. For MHC class I-binding studies, amino acid side chains have to be labeled with fluorescence dyes or with radioactive isotopes, such as ^{125}I. In order to purify the MHC molecules, the MHC-peptide complexes must be separated from unbound free peptide by gel filtration (Sette *at al.*, 1994) or high-performance size exclusion chromatography (Vogt *et. al.*, 1994). Competition-based binding assays result in IC_{50} values, which allow a comparison of binding affinities of different peptides. Competition assays using purified MHC molecules employ soluble MHC molecules, which might have a slightly

different structure compared to native molecules embedded in cell membrane. Whole cell assays overcome these problems by using MHC molecules on the surface of intact cells. Mild acid treatment of the cells is used to remove peptides from the MHC binding grove. Cells are then incubated with a fluorescence-labeled reference peptide together with different concentrations of the peptides of interest (Kessler *et. al.*, 2003). The amount of bound reference peptide can be measured flow cytometrically. Low fluorescence intensity values indicate stronger binding of the peptide of interest.

In reconstitution assays, cells expressing the appropriate HLA allele are stripped of all MHC-bound peptides by incubation at pH 3.2 for a short time, then the peptide of interest and a conformation-dependent monoclonal antibody are added to the cell suspension. The difference in mean fluorescence intensity (MFI) between cells incubated with and without peptide after staining with fluorescence-labeled antibody is used to measure peptide binding (Storkus *et. al.*, 1993).

We utilized class I MHC stabilization as our criterion for evaluating the binding of the synthetic peptides to class I molecules (see Materials and Methods). Following optimization of the assay conditions, this approach proved to be both consistent and efficacious in allowing us to distinguish the MHC-stabilizing from non-stabilizing peptides. We analyzed seven LFV-derived peptides for their binding capability to the HLA molecules on the TAP deficient T2. Two GP-derived peptides reproducibly stabilized class I MHC (HLA-A2.1) expressed on T2 cells as evidenced by the three-fold shift in the MFI when compared to the no peptide control as measured by flow cytometry. One GP-derived peptide caused a two-fold shift in the MFI over no peptide control. The remaining peptides tested failed to show any significant shift in the MFI values.

While we employed the use of human TAP-deficient T2 cells, a murine cell line (RMA-S) can also be used in these assays (Ljunggren and Karre, 1985). In fact, the transfection of virtually any class I-expressing human cell line with cDNA encoding the immediate early protein ICP47 of the Herpes simplex virus can be used to generate a TAP-deficient phenotype and thus, permits the extension of the stabilization assay to virtually any HLA allele (Gatfield *et. al.*, 1998).

The experimental verification of a predicted epitope is achieved by the identification of T-cells that specially recognize the naturally processed epitope in an HLA-restricted fashion. As mentioned earlier, a common approach to such validation employs the use of PBMC from healthy donors who are either already primed *in vivo* with the antigen of interest or which can be primed *in vitro*. The specific CD8+ or CD4+ T-cells can also be generated by immunizing HLA class I and class II transgenic mice (Faulkner *et. al.*, 1998; Sonderstrup *et. al.*, 1999). In our *in vivo* studies, evaluation of candidate epitopes was achieved using a transgenic mouse model. Six-to eight week-old female mice (C57BL/6-TgN [HLA-A2.1]) 1Enge) were purchased from Jackson Laboratory (Bar Harbor, ME). The choice of this transgenic mice model was based upon expression of the appropriate human HLA class I allele (HLA-A2.1) and their commercial availability. These transgenic mice were constructed by microinjection of the 7-kb EcoRI fragment containing the full length HLA-A2.1 gene into fertilized eggs from C57/BL/6 mice (Le *et. al.*, 1989).

Other HLA-A2.1 transgenic mouse models exist. One commonly used mouse model was developed by Pascolo *et. al.* (1997). These transgenic mice (HHD-2) have their β_2-microglobulin and H-2D^b genes knocked out. In addition, they are transgenic for

a chimeric human HLA-A2.1 expressing the α1 and α2 domains of HLA-A2.1 and a murine D^b-derived α3 domain to allow interaction with mouse CD8. Finally, they are also engineered to have HLA-A2.1 covalently linked human β_2-microglobulin to compensate for lack of any free murine β_2-microglobulin. Thus, these mice do not express mouse class I MHC molecules. This model is especially useful when the evaluating the potential for protective immune responses generated by mapped epitopes since positive responses must be attributed to the ability of the peptide under investigation to bind human class I molecules. Okazaki *et. al.* (2003) have successfully used this strain to demonstrate protective immune responses following immunization with reverse transcriptase peptides and subsequent challenge with vaccinia virus expressing reverse transcriptase protein of HIV.

In vitro T-cell responses are commonly measured using a variety of functional assays including proliferation and cytokine production and traditional CTL assays. Lytic activity of CTLs is typically analyzed using a so-called ^{51}Cr-release assay. In our studies, we used a non-radioactive cytotoxicity assay as described in Materials and Methods. While this assay may not be as sensitive as the ^{51}Cr-release assay, it is a safe, convenient, and quantitative method for measuring effector cell cytotoxicity. Positive CTL responses were shown for three of the LFV GP-derived peptides identified in this study.

Assays commonly used to examine cytokine production by peptide-stimulated T-cells include the IFN-γ ELISPOT assay (to measure IFN-γ secretion) and IFN-γ intracellular cytokine staining, both of which were employed in our studies. While one might expect to observe a 100% correlation between results obtained using these assays since they both measure IFN-γ, results of IFN-γ intracellular cytokine staining

experiments sometimes disagreed with the positive IFN-γ ELISPOT data. It is noteworthy that the IFN-γ-producing cells detected in ELISPOT assays detects IFN-γ produced by all splenocytes (both CD4+ and CD8+ cells). Another difference between these assays concerns the location of the target cytokine being measured. Optimal detection of intracellular IFN-γ by peptide-stimulated cells may require more stringent assay conditions. For example, it is possible that the window of time for detection of intracellular IFN-γ may be shorter than that associated with cell secretion of this cytokine. The IFN-γ ELISPOT assay was therefore very useful in our studies and the results obtained were consistent when performed several times with different peptides. Using this assay, three of the LFV GP-derived peptides stimulated IFN-γ production in the *in vitro* peptide-stimulated lymphocytes. The YLISIFLHL peptide, which induced IFN-γ production at a frequency of $155/10^6$ cells as measured by ELISPOT, induced IFN-γ production in 3.1% of CD8+ cells as measured by intracellular staining. Among the CD8+ T-cell population, 0.6% showed intracellular IFN-γ following *in vitro* stimulation with LFV GP peptide LLGTFTWTL. The same peptide induced IFN-γ production at a frequency of $215/10^6$ as measured in the ELISPOT assay. Finally, the LFV GP SLYKGVYEL peptide induced IFN-γ production at a frequency of $205/10^6$ cells and intracellular IFN-γ was detected in 0.8% of CD8+ cells. The IFN-γ responses obtained in the ELISPOT assay correlated well with the CTL responses of the respective peptides. Both the CTL responses and the IFN-γ production were dose-dependent with regard to effector: target cell ratios and peptide concentration, respectively. The highest CTL response was observed at the 40:1 effector: target ratio and it gradually decreased with decreasing effector: target ratios for each of the tested peptides. Similarly the production

of IFN-γ decreased with the decreasing stimulating peptide concentration. While functional assays such as the ones used in this study provide clear evidence for the immunogenic potential of given peptide motifs, they do not provide quantitative information related to the incidence of peptide-specific T cells in peptide-immunized mice. Enumeration of antigen-specific CTLs can be achieved flow cytometrically using class I MHC tetramers (Altman *et. al.*, 1996). We selected one LFV GP peptide (LLGTFTWTL) that fulfilled our criteria as a candidate immunogen and then commissioned the NIH Tetramer Facility to prepare class I MHC tetramers using this synthetic peptide. We did not have the option to test the other candidate epitopes in tetramer staining experiments due to our inability to have corresponding peptide-class I MHC complexes prepared by the NIH tetramer facility. The overwhelming number of tetramer preparation requests received by the NIH makes it necessary for them to limit the number of preparations to one per investigator. Following their successful preparation of the LLGTFTWTL/HLA-A2.1 tetramer, we performed flow cytometric analysis to measure the frequency of peptide-specific CD8+ cells. The Results illustrates representative data from our tetramer staining experiments and shows that lymphoid cells from mice primed *in vivo* with LLGTFTWTL contained peptide-specific CD8+ cells at a frequency of 0.32%. This frequency varied slightly from experiment to experiment but it was consistently significantly higher than the frequency of peptide-specific T-cells detected in splenocytes from unimmunized mice.

Following verification of peptide immunogenicity using peptide-immunized HLA-expressing mice, CTL epitopes ideally, should then be tested for their ability to generate immune responses in humans. The ultimate goal of epitope mapping is design of

protective or therapeutic vaccines. The advantages of using epitope-based vaccines include their safety, the opportunity to engineer epitopes for increased potency, and the ability to focus immune responses on conserved epitopes. Most advances in epitope-based vaccines have been seen in the area of cancer immunotherapy. Indeed, a variety of MHC class I-restricted epitopes for cancer-associated antigens have been identified to date and clinical studies have demonstrated promising results (Berzowsky *et. al.*, 2004). Several approaches to therapeutic tumor vaccines have been reported. They include adoptive transfer of DC loaded with peptides (Mackensen *et. al.*, 2000); vaccinations with synthetic peptides (Rosenberg *et. al.*, 1998); vaccinations with modified tumor cells (Simons *et. al.*, 1999); or vaccinations with recombinant viruses expressing tumor antigens (Marshall *et. al.*, 2000). Since peptides are normally not highly immunogenic, novel approaches to enhancing their immunogenicity have also been used. These include the use of cytokines, chemokines, and costimulatory molecules as adjuvants (Ahlers *et. al.*, 2002). Similarly, studies are underway to investigate the efficacy of mapped epitopes (peptides) as immunotherapeutic agents for the treatment of chronic infections such as those caused by HIV (McGaughey *et. al.*, 2004), HCV (Engler *et. al.*, 2004), and HPV (Fausch *et. al.*, 2003).

In summary, results of this thesis research indicate that the computational algorithms can serve as a valid and useful tool in CTL epitope prediction. Using BIMAS and SYFPEITHI algorithms, we have identified candidate CTL epitopes for LFV. Using a class I MHC stabilization assay to test selected synthesized peptides we showed that three out of five GP-derived LFV peptides (LLGTFTWTL, SLYKGVYEL, and YLISIFLHL) stabilize the molecule encoded by the HLA-A2 allele. The two NP-derived

LFV peptides were found to be nonbinders. The GP-derived nanomers stimulated CTL responses in transgenic mice expressing the HLA-A2.1. They also induced immune responses in HLA-A2.1 transgenic mice as measured in CTL, IFN-γ ELISPOT, and intracellular cytokine staining assays. Finally, we confirmed that one selected nanomer (LLGTFTWTL) have the ability to induce measurable peptide-specific CD8 T cell responses as shown by class I tetramer staining of lymphocytes from LLGTFTWTL-primed mice. We summarized our results in Tables 17 and 18 of the Results section. We generated a collective ranking scheme for each of the mapped LFV epitopes identified in this study. The outcome of this epitope ranking approach demonstrates that the mapped epitope with the highest immunogenic profile, based upon *in vitro* and *in vivo* studies reported herein, is the peptide that scored as the highest predicted HLA-A2-binding LVF GP peptide using the BIMAS algorithm (i.e., LLGTFTWTL).

All three immunogenic LFV epitopes identified in this study are contained within the LFV glycoprotein(GP). This finding is fortuitous and potentially clinically exploitable in light of the fact that the principle LFV virulence factors are believed to be associated with the GP moiety of this virus (discussed in the Introduction section of this thesis).

We have also computationally identified T-helper cell epitopes and B-cell epitopes for LFV. Candidate T-helper epitopes were chosen according to predictions assessed using the ProPred and SYFPEITHI algorithm. One of the predicted T-helper epitopes identified by us (LFV GP282-301) was recently identified and confirmed experimentally by ter Meulen *et. al.,* (2004). This epitope scored among the 20 highest-ranking peptides using both the ProPred and SYFPEITHI algorithms. Other LFV GP- and NP-derived peptides, which were found to be cross-reactive among 51 HLA-DR alleles,

have not been identified in other studies to date. A good correlation between experimental and computational data is needed to suggest that the T-helper cell (class II MHC) epitopes which we have identified are truly immunogenic epitopes.

Applying the β-turn and hydrophobicity prediction algorithms we also mapped candidate B-cell epitopes. We have identified three candidate B-cell epitope peptides from a glycoprotein (GP75-100, GP180-220, and GP255-270) and three from the nucleoprotein (NP145-160, NP285-305, and NP335-350) of LFV. The B-cell epitopes for LFV have not been previously identified.

Once identified, CTL, T-helper and B-cell epitopes can be combined as a subunit complex to investigate their possible synergistic effect in priming both arms of the immune system to generate protective immune responses against viral pathogens. Successful identification of such epitopes can also help overcome other vaccine-related challenges. For example, such information greatly facilitates the development of recombinant epitope-based vaccines. While it may be premature to offer an overall conclusion regarding the ultimate utility of mapped peptides in a clinical context, the combined advances in computational and structural biology, microbial pathogenesis, and translational research will likely yield the vaccines of the future. It is hoped that the present study has contributed to this goal with respect to our knowledge of potential strategies for developing prophylactic approaches to the prevention of infections caused by hemorrhagic LFV.

REFERENCES

Aichele P, Hengartner H, Zinkernagel R, Schultz M. (1990) Antiviral cytotoxic T cell response induced by in vivo priming with a free synthetic peptide. *J. Exp. Med.* **171**:1815.

Ahlers JD, Belyakov IM, Terabe M, Koka R, Donaldson DD, Thomas EK, Berzofsky JA. (2002)A push-pull approach to maximize vaccine efficacy: abrogating suppression with an IL-13 inhibitor while augmenting help with granulocyte/macrophage colony-stimulating factor and CD40L. *Proc. Natl. Acad. Sci.* **99**:13020-13025.

Ahmed R, Simon R, Matloubian M (1988) genetic analysis of in vivo selected viral variants causing chronic infection:importance of mutation in the l RNA segment of lymphocytic choriomeningitis virus. *J Virol* **62**:3301-3308.

Altuvia Y *et al* (1997) Structure-Based Algorithm to predict Potential Binding Peptides to MHC Molecules with Hydrophobic Binding Pockets. *Human Immunology* **58**:1-15.

Allsopp CE, Plebanski M, Gilbert S, Sinden RE, Harris S, Frankel G, Dougan G, Hioe C, Nixon D, Paoletti E, Layton G, Hill AV. (1996) Comparizon of numerous delivery systems for the induction of cytotoxic T lymphocytes by immunization. *Eur. J. Immunol.* **26**(8):1951-1959.

Altman JD, Moss PA, Goulder PJ, Barouch DH, McHeyzer-Williams MG, Bell Jl, McMichael AJ, Davis MM. (1996) Phenotypic analysis of antigen-specific T lymphocytes. *Science* **4**; 274(5284):94-96.

An L, and Whitton JL. (1997) A multivalent minigene vaccine, containing B-cell, Cytotoxic T-lymphocyte , and Th epitopes from several microbes, induces appropriate responses in vivo and confers protection against more than one pathogen. *J. Virol.* **71**: 2292.

Andersen MH, Tan L, Sondergaard I, Zeuthen J, Elliot T, Haurum JS. (2000) Poor correspondence between predicted and experimental binding of peptides to class I MHC molecules. *Tissue Antigens*. **55**: 519-531.

Anderson KS, Alexander J, Wei M, Cresswell P. (1993) Intracellular transport of class I MHC molecules in antigen processing mutant cell lines. *Journal of Immunol.* **151**: 3407-3419.

Appay V and Rowland-Jones SL. (2002) The assessment of antigen-specific CD8 T cells through the combination of MHC class I tetramer and intracellular staining. *Journal of Immunological Methods* **268**: 9-19.

Arnon R, Ben-Yedida T. (2003) Old and new vaccine approaches. *International Immunopharmacology* **3**:1195-1204.

Babbitt B, Allen PM, Matsueda G, Haber P, Unuanue ER. (1985) Binding of immunogenic peptides to histopatibility molecules. *Nature* **317**:359.

Baize S, Leroy EM, Georges-Courbot MC, Capron M, Lansoud-Soukate J, Debre P, Sisher-Hoch P, McCormick JB, GeorgesAJ. (1999) Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients. *Nat. Med.* **5**:423-426

Beresford PJ, Zhang D, Oh DY, Fan Z, Greer EL, Russo ML, Jaju M, Lieberman J. (2001) Granzyme A activates an endoplasmic reticulum-associated caspase-independent nuclease to induce single-stranded DNA nicks. *J Biol. Chem.* **276**:43285-43293.

Berzowsky JA, Masaki T, SangKon O, Belyakov I, Ahlers JD, Janik JE, Morris J. (2004) Progress on new vaccine strategies for the immunotherapy and prevention of cancer. The *Journal of Clinical Investigation.* 11, **113**:1515-1523.

Bian H, Reidhaar-Olson JF, Hammer J. (2002) The use of bioinformatics for identifying class II-restricted T-cell epitopes. *Methods* **29**: 299-309.

Bluthner M, Mahler M, Muller DB, Dunzl H, Bautz FA. (2003) Identification of an alpha-helical epitope region on the PM/Scl-100 autoantigen with structural homology to a region on the heterochromatin p25beta autoantigen using immobilized overlapping synthetic peptides. *J Mol Med.* **78**(1):47-54.

Bodmer JG, Marsh SGE, Albert ED, Bodmer WF, Bontrop RE, Charron DE, Dupont B, Erlich HA, Fauchet T, Mach B, Mayr WR, Parham P, Sasazuki T, Schreuder G, Bowen MD, Rolin PE, Ksiazek TG, et. al. Genetic diversity among LFV strains. *J Virol* **74**:6992-7004.

Bowen MD, Rollin PE, Ksiazek T, Hustad HL, Bausch DG, Nichol S. (2000) Genetic diversity among LFV strains. *Journal of Virology*, Vol **74**, 15, 6992-7004.

Bouvier M, Willey (1994) Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. *Science* **265**:398-402.

Brander, C., Wyss-Coray, T., Mauri, D., Bettens, F. and Pichler, W.J. (1993) Carrier-mediated uptake and presentation of a major histocompatibility complex class I-restricted peptide. *Eur. J. Immunol.* **23**(12), 3217-23.

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Storminger JL, Wiley DC. (1993) Three-dimentional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**: 33-39.

Bruns M, Zeller Wrohdewohl H, lehmann-Grube F. (1986) Lymphocytic choriomeningitis virus.IX. Properties of the nucleocapsid. *Virology* **151**:77-85.

Brusik V, Rudy G, Harrison LC. (1994) Prediction of MHC binding peptides using artificial neural networks. In: Stonier RJ, Yu XS, editors. Complex systems, mechanism of adaptation. *Amsterdam: IOS press*, 253-260.

Burns JW, Buchmeier MJ. (1991) Protein-protein interactions in lymphocytic choriomeningitis virus. *Virology* **183**:620-629.

Cao K, Hollenbach J, Shi X, Shi W, Chopek M, Fernandez-Vina MA, (2001) Analysis of the Frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United states reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Human Immunol.* **62**:1009-1030.

Cao W, Henry MD, Borrow P, Yamada H, Elder JH, et. al. (1998) Identification of α-Dystroglycan as receptorfor Lymphocytic Choriomeningitis Virus and LFV. *Science* **282**; 2079-2081.

Celis E, Tsai V, Crimi C, DeMars R, Wentworth PA, Chesnut RW, Grey HM, Sette A, Serra HM. (1993) Induction of anti-tumor cytotoxic T lymphocytes in normal humans using primary cultures and synthetic peptide epitopes. Proc Natl Acad Sci U S A. **91**(6):2105-2109.

Chao DL, Davenport MP, Forrest S, Perelson AS. (2004) A stochastic model of cytotoxic T cell responses. *J Theor Biol.* **21**;228(2):227-40.

Choi EM, Palmowski M, Chen J, Cerundolo V. (2002) The use of chimeric A2Kb tetramers to monitor HLA A2 immune response in HLA A2 transgenic mice. *Journal of Immunological Methods* **268**, 35-41.

Chou PY, Fasman GD. (1978) Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol Relat Areas Mol Biol.* **47**:45-148.

Cochlovius B, Stassar M, Christ O, Raddrizzani L, Hammer J, Mytilineos I, Zoller M. (2000) In vitro and in vivo induction of a Th cell response toward peptides of the melanoma-associated glycoprotein 100 protein selected by the TEPITOPE program. J Immunol. **165**(8):4731-41.

Colucci F, Caligiuri MA, Di SantoJP. (2003) What does it take to make a natural killer? *Nature Rev. Immunol.* **3**, 413-425.

Coux O, Tanaka K, Goldberg AL (1996) Structure and functions of the 20S and 26S proteasomes. *Annu. Rev. Biochem.* **65**, 801-848.
Craighton TE. (1993) Proteins: Structure and Molecular Properties, 2nd ed. WH Freeman, New York.

Cummins D, McCormick, Bennett D, Samba JA, Farrar B, Machin SJ, Fischer-Hoch SP. (1990) Acute sensorineural deafness in Lassa fever. *JAMA* **264**, 2093-2096.

Dal Porto J, Johansen TE, catipovic B, Parfiit DJ, Tuveson D, Gether U, Kozlowski S, Fearon DT, Schneck JP. (1993) A soluble divalent class I major histocompatibility complex molecule inhibits alloreactive T cells at nanomolar concentrations. *Proct. Natl. Acad. Sci.* **90**, 6671-6675.

Daniels MA, Jameson SC. (2000) Critical role for CD8 in T cell receptor binding and activation by peptide/major histocompatibility complex multimers. *J. Exp. Med.* **191**,335.

Day PM, Yewdell JW, Porgador A, germain RN, Bennink JR. (1997) Direct delivery of exogenous MHC class I molecule-binding oligopeptides to the endoplasmic reticulum of viable cells. *PNAS* **22**;94(15):8064-8069.

del Guercio M-F, Sidney J, hermanson G, Perez C, Grey HM, Kubo RT, Sette A. (1995) binding of a peptide antigen to multiply HLA alleles allows definition of an A2-like supertype. *J Immunol.* **154**:685.

De Groot AS, Aubin CS, Bosma A, Sbai H, Rayner J, Martin W. (2001a) Rapid Determination of HLA B*07 Ligands from the West Nile Virus NY99 Genome. *Emer. Infect. Dis.* **7**: 706-713.

De Groot AS, Bosma A, Chinai N, Frost J, Jesdale B, Gonzales M, martin W, Saint-Aubin. (2001b) From genome to vaccine: in silico predictions, ex vivo verification *Vaccine* **19**, 4385-4395.

Denzin LK, Sant'Angelo DB, Hammond C, Surman MJ, Cresswell P (1997) Negative regulation by HLA-DO of MHC class II-restricted antigen processing. *Science.* **278**(5335): 106-9.

Deres K, Schild H, Wiesmuller KH, Jung G, Rammensee HG. (1989) In vivo priming of virus-specific cytotoxic T lymphocytes with synthetic lipopeptide vaccine. *Nature* **342**:561-564.

Ding YH, Smith KJ, Garboczi DN, Utz U, Biddison WE, Wiley DC. (1998) Two human T cell receptors bind in a similar diagonal mode to the HLA-A2/tax peptide complex using different TCR amino acids. *Immunity* **8**: 403-411.

Dönnes and A Elofsson (2002)Prediction of MHC class I binding peptides, using SVMHC. *BMC Bioinformatics*; **3**:1471-2105

Drosten C, Kummerer B, Schmitz H, Gunther S. (2003) Molecular diagnostics of viral hemorrhagic fevers. *Antiviral research* **57**; 61-8.

Ehl S, Hombach J, Aichele P, Rulicke T, Odermatt B, Hengartner H, Zinkernagel R, Pircher H. (1998) Viral and bacterial infections interfere with peripheral tolerance induction and activate CD8+ T cells to cause immunopathology. *J Exp Med.* **2**;187(5):763-74

Engler OB, Schwendener RA, Dai WJ, Wolk B, Pichler W, Mordapour D, Brunner T, Cerny A. (2004) A liposomal peptide vaccine inducing CD8(+) T cells in HLA-A2.1 transgenic mice , which recognize human cells encoding hepatitis C virus(HCV) proteins. *Vaccine* **15**:23(1): 58-68.

Even, C., Rowland, R.R. and Plagemann, P.G. (1995) Cytotoxic T cells are elicited during acute infection of mice with lactate dehydrogenase-elevating virus but disappear during the chronic phase of infection. *J. Virol.* **69**(9), 5666-76.

Ezekowitz RAB, Williams DJ, Koziel H, Amstrong MYK, Warner A, Richards FF, Rose RM. (1991) Uptake of *Pneumocystis carini* mediated by the macrophage mannose receptor. *Nature* **351**:155-158.

Falk K, Rötzschke O, Stevanovic S, Jung G, Rammensee H-G (1991) Allele-specificmotifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* **351**:290-296.

Faulkner L, Borysiewicz LK, Man S. (1998) The use of human leukocyte antigen class I transgenic mice to investigate human immune function. *Journal of Immunological Methods.* **221**:1-16.

Fausch SC, Da Sliva DM, Eiben GL, Le Poole IC, Kast WM. (2003). HPV protein/peptide vaccines: from animal models to clinical trials. *Front Biosci.* **8**:81-91.

Fields BN, Knipe DM, Howley et. al., (1996) Fundamental virology, Third Edition. Lippincott – Raven Publishers, Philadelphia.

Finke, S., Trojaneck, B., Lefterova, P., Csipai, M., Wagner, E., Kircheis, R., Neubauer, A., Huhn, D., Wittig, B., and Schmidt-Wolf, I.G.H. (1998) Increase of proliferation rate and enhancement of antitumor cytotoxicity of expanded human CD3+ CD56+ immunologic effector cells by receptor-mediated transfection with the interleukin-7 gene. *Gene Therapy* **5**(1), 31-39.

Firat H, Garcia-Pons, Tourdot S, Pascolo S, Scardino A, Garcia Z, Michel ML, Jack RW, Jung G, Kosmatopoulos, Mateo l, Suhrbier A, Lemonnier FA, Langlad A, Demoyen P. (1999) H-2 class I knockout, HLA-A2.1-transgenic mice:a versatile animal model for preclinical evaluation of antitumor immunoterapeutic strategies. *Eur J Immunol.* **10**:3112-21.

Fisher GH, Rosenberg EJ, Starus SE, Dale JK, Middleton LA, Lin AY, Strober W, Leonardo MJ, Puck JM. (1995) Dominant interfering Fas gene mutations impair apoptosis in a human autoimmune lymphoproliferative syndrome. *Cell* **81**:935-946.

Fisher-Hoch SP, Hutwagner L, Brown B, *et. al.* (2000) Effective vaccine for Lassa fever. *J Virol* **74**:6777-6783.

Fisher-Hoch SP, McCormick JB, Sasso D, Craven RB. (1988) Hematologic disfunction in Lassa fever. *J. Med. Virol.* **26**:127-135.

Foster B. and Prussin C. Detection of intracellular cytokines by flow cytometry. In *Current Protocols in Immunology* (2002) by John Wiley & Sons, Inc, New York. Pp 6.21.1 – 6.24.15.

Fremont DH, Matsumura M, Stura EA, Peterson PA, Wilson IA (1992) Crystal structures of two viral peptides in complex with murine MHC class I H-2 kb. *Science* **257**:919-927.

Fruci D, Rovero P, Falasca G, Chersi A, Sorrentino R, Butler R, Tanigaki N, Tosi R. (1993) Anchor residue motifs of HLA class-I-binding peptides analyzed by the direct binding of synthetic peptides to HLA class I α chains. *Human Immunol.* **38**:187.

Gao X, Zheng B, Liew S, Brett, Tite J. (1991) priming of influenza virus-specific cytotoxic T lymphocytes in vivo by short synthetic peptides. *J Immunol.* **147**: 3268.

Garboczi DN, Ghosh P, Utz U, Fan QR, Biddison WE, Wiley DC. (1996) Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* **484**: 134-141.

Garcia F, SepulvedaP, Liegeard P, Gregoire J, Hermann E, Lemonnier F, Langlade-Demoyen P, Hontebeyrie M, Chun Lone Y. (2003) Identification of HLA-A2.1-restricted cytotoxic T-cell epitopes of *Trypanosoma cruzi* TcP2β protein in HLA-transgenic mice and patients. *Microbes and Infection* **5**; 351-359.

Garcia KC (1999) Structural basis of T cell recognition. *Annu. Rev. Immunol.* **17**: 369-397.

Garcia KC, Degano M, Pease LR, Huang M, Peterson PA, Leyton L, Wilson IA. (1998) Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science* **279**: 1166-1172.

Garcia KC, Scott CA, Brunmark A, carbone FR, Peterson PA, Wilson IA, Teylon L. (1996) CD8 enhances formation of stable T-cell receptor/ MHC class I molecule complexes. *Nature* **384**, 577.

Germain RN. (1994) MHC-dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation. *Cell* **76**:287.

Grant G. (2003) Current Protocols in Immunology by John Wiley & Sons, Inc, New York. Pp 9.2.1-9.2.19

Greten TF, Slansky JE, Kubota R, Soldan SS, Jaffee EM, Leist TP, Pardol DM, Jacobson S, Schneck JP. (1998) Direct visualization of antigen-specific T cells: HTLV-1 Tax11-19- specific CD8(+) T cells are activated in peripheral blood and accumulate in cerebrospinal fluid from HAM/TSP patients. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7568-7573.

Gulukota K,Sidney A, Sette A and Delisi C (1997) Two complementary methods for predicting peptides binding major histocompatibility complex molecules. *J Mol Biol* **267**, 1258-1267.

Hammer J, Belunis C, Bolin D, Papadopoulos J, Walsky R, Higelin J, Danho W, Sinigaglia F, Nagy ZA. (1994) High-affinity binding of short peptides to major histocompatibility complex class II molecules by anchor combinations. *Proc Natl Acad Sci U S A.***10**;91(10):4456-60.

Hammer J, Valsasnini P, Tolba K, Bolin D, Higelin J, Takacs B, Sinigaglia F (1993) Promiscuous and allele-specific anchors in HLA-DR-binding peptides. Cell. **16**;74(1):197-203.

Hammer J, Bono E, Gallazi F, Belunis C, Nagy Z, Sinigaglia (1994) Precise prediction of major histocompatibility complex class II-peptide interaction based on peptide side chain scanning. *J Exp Med* **180**:2353.

Heller MV, Saavedra MC, Falcoff R, Maiztegui JI, Molinas FC. (1992) Dystroglycan: an extracellular matrix receptor linked to the cytoskeleton *J. Infect. Dis.* **166,** 1203-1204.

Henry M and Campbell, (1996) *Curr. Opin. Cell. Biol.* **8**, 625.

Hernandez JM, Bui MHT, Han K, Nguyen DG, Belldegrun AS. (2003) Novel kidney cancer immunotherapy based on granulocyte-macrophage colony-stimulating factor and carbonic anhydrase IX fusion gene. *Clin. Can. Res.* **9**, 1906-1916.

Hioe CE, Qiu H, Chend PD, Bian Z, Li J, Singh M, Kuebler P, McGee P, O'Hagan D, Zamb T, Koff W, Allsopp C, Wang CY, Nixon DF. (1996) Comparison of adjuvant formulations for cytotoxic T cell induction using synthetic peptides. *Vaccine* **14**(5):412-418.

Hilleman MR. (2002) Overview: cause and prevention in biowarfare and bioterrorism. *Vaccine.* **19**;20(25-26):3055-67.

Huang YH, Tao MH, Hu CP, Syu WJ, Wu JC (2004) Identification of novel HLA-A2.1-restricted CD8+ T-cell epitopes on hepatitis delta virus. *J Gen Virol.* **85**(Pt 10):3089-98.

Hugues S, Malherbe L, Filippi C, Glaichenhaus N. (2003) Generation and use of alternative multimers of peptide/MHC complexes. *Journal of Immunological Methods* **268**, 83-92.

Hwang LY, Lieu PT, Peterson Y, PA, Yang Y. (2001) Functional regulation of immunoproteasomes and transporter associated with antigen processing. *Immunol. Res.* **24**, 245-272.

Ishioka GY, Fikes J, Hermanson G, Livingston B, Chestnut RW, Sette A. (1999) Utilization of MHC class I transgenic mice for development of minigene DNA vaccines encoding multiple HLA-restricted CTL epitopes. *J. Immunology* **162**: 3915-3925.

Janeway CA, Bottomly K. (1994) Signals and signs for lymphocyte responses. *Cell* **76**:275-285.

Jesdale BM, Deocampo G, Meisell, Beall J, Marinello MJ, Chicz RM, De Groot A (1997) Matrix-based prediction of MHC binding peptides: The EpiMatrix algorithm, reagent for HIV research. Vaccines '97. Cold Spring Harbor, NY: Cold Spring Harbor Press, 1997.

Juffrie M, van Der Meer GM, Hack CE, Haasnoot K, Sutaryo, Veerman AJ, Thijs LG (2000) Inflamatory mediators in Dengue virus infection in children:interleukin -8 and its relationship to neutrophil degranulation. *Infect. Immun.* **68**: 702-707.

Kalams SA and Walker BD. (1998) The critical need for CD4 help in maintaining effective cytotoxic T lymphocyte responses. *J. Exp. Med.* **188**:2199-2204.

Kessler JH, Beekman NJ, Bres-Vloemans SA, Verdijk P, van Veelen PA, Kloosterman-Joosten AM, Vissers DC, ten Bosch GJ, Kester MG, Sijts A, Wouter Drijfhout J, Ossendorp F, Offringa R, Melief CJ. (2001) Efficient identification of novel HLA-A(*)0201-presented cytotoxic T lymphocyte epitopes in the widely expressed tumor antigen PRAME by proteasome-mediated digestion analysis. *J Exp Med.***193**(1):73-88.

Kessler JH, Mommaas B, Mutis T, Huijbers I, Vissers D, Benckhuijsen WE, Schreuder GM, Offringa R, Goulmy E, Melief CJ, van der Burg SH, Drijfhout JW. (2003) Competition-based cellular peptide binding assays for 13 prevalent HLA class I alleles using fluorescein-labeled synthetic peptides. *Hum Immunol.* **64**(2):245-55.

Kita H, HeXS, Gershwin E. (2003) Application of tetramer technology in studies on autoimmune diseases. *Autoimminity Rev.* **2**; 43-49.

Kleijmeer MJ, Kelly A, geuze HJ, Slot JW, Townsend A, Trowsdale J. (1992) Location of MHC encoded transporters in the endoplasmic reticulum and *cis*-Golgi. *Nature* **357**:342.

Kloetzel PM. (2001) Antigen processing by the proteasome *Nat Rev Mol Cell Biol.* **2**(3):179-87.

Khan S, Van der Broek M, Schwarz K, De Giuli R, Diener PA, Groettrup M. (2001) *J. Immunol.* **167**, 6859-6868.

Kim, H.S., Park, C.H., Cha, S.H., Lee, J.H., Lee, S., Kim, Y., Rah, J.C., Jeong, S.J. and Suh, Y.H. (2000) Carboxyl-terminal fragment of Alzheimer's APP destabilizes calcium homeostasis and renders neuronal cells vulnerable to excitotoxicity. *FASEB J.* **14**, 1508-17.

Kopydlowski KM, Salkowski CA, Cody MJ et. al. (1999) Regulation of macrophage chemokine expression by lipopolysaccharide in vitro and in vivo. *J Jmmunol.* **163**: 1537-1544.

Korzeniewski C, Callewaert DM. (1983) An enzyme-release assay for natural cytotoxicity. *J. Immunol. Meth.* **64**, 313-320.

Kroneberg M, Siu G, Hood LE, Shastri N. (1986) The molecular genetics of the T-cell antigen receptor and T-cell antigen recognition. *Annu. Rev. Immunol.* **4**:529-591.

Kruisbeek A and Shevach E. (1996) Blocking cellular division of accessory/stimulator cells. In: *Current Protocols in Immunology* Vol **1**, J. Coligan et. al. (Eds.), John Wiley & Sons, Inc, New York. Pp 3.12.18.

Kyte J, Doolittle RF. (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* **5**;157(1):105-32.

La Posta V, Auperin DD, Kamin-Lewis R, Cole G. (1993) Cross-protection against Lymphocytyc Choriomeningitis Virus Mediated by a CD4+ T-cell clone specific for an envelope glycoprotein epitope of Lassa Virus. *Journal of Virology.* **67**, 6:3497-3506.

Le AX, Bernhard EJ, Holterman M, Strub S, Parham P, Lacy E, Engelhard VH. (1989) Cytotoxic T cell responses in HLA-A2.1 transgenic mice. Recognition of HLA alloantigens and utilization of HLA-A2 as a restriction element. *J Immunology* 142; 1366-1371.

Leighton J, Sette A, Sidney J, Apella E, Ehrhardt C, Fuchs S (1991) Comparison of structural requirements for interaction of the same peptide with I-Ek and I-Ed molecules in the activation of the MHC class II-restricted T cells. *J Immunol* **147**(1):198-204.

Lieberman J. (2003) The ABCs of granule-granule mediated cytotoxicity: new weapons in the arsenal. *Nature rev. Immunol.* **3**:361-370.

Lipford GB, Hoffman M, Wagner H, Heeg K (1993) Primary in vivo responses to ovalbumin. Probing The predictive value of the Kb binding motif. *J Immunol* **150**:1212-1216.

Livingston BD, Newman M, Crimi C, McKinney D, Chestnut R, Alessandro Sette (2001) Optimization of epitope processing enhances immunogenicity of multiepitope DNA vaccines. *Vaccine* **19**, 4652-4660.

Ljunggren HG, Karre K. (1985) Host resistance directed selectively against H-2-deficient lymphoma variants. Analysis of the mechanism. *J Exp Med.*162(6):1745-59.

Lu J, Celis E. (2000) Use of two predictive algorithms of the World Wide Web for the identification of tumor-reactive T-cell epitopes. *Cancer research.* **60**:5223-5227.

Lucey DR, Clerici M, Shearer GM. (1996) type I and type 2 cytokine dysregulation in human infectious, neoplastic and inflammatory diseases. *Clin. Microbiol. Rev.* **9**(4):532-562.

Lukasievich IS, Tikhonov I, Rodas JD, Zapata JC, Yang Y, Djavani M, Salvato MS. (2003) Arenavirus-mediated liver pathology:acute lymphocytic choriomeningitis virus infection of Rjasus macaques is characterized by high-level Interleukin-6 expression and hepatocyte proliferation. *Journal of Virology* **77**; 1727-1737.

Lycke NY and Coico RF. (1996) ELISPOT Assay for measurement of antigen-specific and polyclonal antibody responses. In: *Current Protocols in immunology* Vol. **2**, D. margulies (Ed.), John Wiley & Sons, Inc, New York. Pp 7.14.1 – 7.14.7

Madden DR, Garboczi DN, Wiley DC. (1993). The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides presented by HLA-A2. *Cell* **75** pp. 693

Mahanty S, Bausch D.G, Thomas R.L, Goba A, Bah A, PetersC.G, Rollin P.E (2001) Low levels of interleukin 8 and interferon-inducible protein-10 in serum are associated with fatal infections in acute Lassa fever. *I. Inf. Dis.* **183**; 1713-1721.

Mahler M, Bluthner M, Pollard KM (2003) Advances in B-cell epitope analysis of autoantigens in connective tissue diseases. *Clin Immunol.* **107**(2):65-79.

Mackensen A, Herbst B, Chen JL, Kohler G, Noppen C, Herr W, Spagnoli GC, Cerundolo V, Lindemann A. (2000) Phase I study in melanoma patients of a vaccine with peptide-pulsed dendritic cells generated in vitro from CD34(+) hematopoietic progenitor cells. *Int. J. Cancer.* **86**:385-392.

Marshall JL, Hoyer RJ, Toomey MA, Faraguna K, Chang P, Richmond E, Pedicano JE, Gehan E, Peck RA, Arlen P, Tsang KY, Schlom J. (2000) Phase I study in advanced cancer patients of a diversified prime-and-boost vaccination protocol using recombinant

vaccinia virus and recombinant nonreplicating avipox virus to elicit anti-carcinoembryonic antigen immune responses. *J. Clin. Oncol.* **18**:3964-3973.

Marsh SG. (1998) Nomenclature for factors of the HLA system. WHO nomenclature committee for factors of the HLA system. *Hum. Immunol.* **59**:256-257.

Martin W, Sbai H, De Groot AS. (2003) Bioinformatics tools for identifying class I-restricted epitopes. *Methods* **29**: 289-298.

Matloubian M, Concepcion J, Ahmed R. (1994) CD4 T cells are required to sustain CD8 cytotoxic T cell responses during chronic viral infection. *J. Virol.* **68**:8056-8063.

Matsui K, Boniface JJ, Reay PA, Schild H, Fazekas de St Groth, Davis MM. (1991) Low affinity intereaction of peptide-MHC complexes with T cell receptors. *Science* **254**:1788-1791.

McCormick JB, Webb PA, Krebs KM, Johnson KM, Smith ES. (1987a) A prospective study of the epidemiology and ecology of Lassa fever. *J. Infect. Dis.* **155**:437-444.

McCormick JB, King IJ, Webb PA, Johnson KM. (1987b) A case-control study of the clinical diagnosis and course of Lassa fever. *J Infect. Dis.* **155**; 445-455.

McCormick JB, King IJ, Webb PA, Scribner CL, Craven RB, Johnson , Elliott, Williams RB. (1986) Lassa fever. Effective therapy with ribavirin. *N. Engl. J. Med.* **314**: 20-26.

McGaughey GB, Barbato G, Binachi E, Freidinger RM, Garsky VM, Hurni WM, Joyc JG, Liang X, Miller MD, Pessi A, Shiver JW, Bogusky MJ. (2004) Progress towards the development of HIV-1gp41-directed vaccine. *Curr. HIV Res.* **2** (2): 193-204.

Medzhitov R, Janeway Jr C. (2000) Innate immune recognition: mechanisms and pathways. *Immunol. Rev.* **173**:89-97.

Meister GE, Roberts CG, Berzowsky JA, De Groot AS. (1995) Two novel T cell prediction algorithms based on MHC-binding motifs; comparison of predicted and published epitopes from *Mycobacterium tuberculosis* and HIV protein sequences. *Vaccine* **13**:581-586.

Metkar SS, wang B, Aguilar-Santelises M, Raja SM, Uhlin-Hansen L, Podack E, Trapani JA, Froelich CJ. (2002) Cytotoxic cell granule-mediated apoptosis: perforin delivers granzyme B-serglycine complexes into target cells without plasma membrane pore formation. *Immunity* **16**:417-428.

Micheletti F, Monini P, Fortini C, Rimessi P, Bazzaro M, Andreoni M, Giuliani M, traniello S, Ensoli B, gaviolo R. (2002) Identification of cytotoxic T lymphocyte epitopes of human herpesvirus 8. *Immunology* **106**; 395-403.

Minev BR, Chavez FL, Dudouet BM, Mitchell MS, (2000) synthetic insertion signal sequence MHC class I presentation of a peptide from the melanoma antigen MART-1. *Eur. J Immunol.* **30**:2115-2124.

Mizumoto N, kumamoto T, Robson SC, matsue H, takashima A. (2002) CD39 is the dominant Largenhans cell-associated ecto-NTPDase: Modulatory roles in inflammation and responsiveness. *Nat. Med.* **8**: 358-365.

Monath TP, Newhouse VF, Kemp GE, Setzer HW, Cacciapuoti. (1974) LFV isolation from Mastomys natalensis rodents during an epidemic in Sierra Leone. *Science* **185**:263-265.

Morrison LA, Lukacher AE, Braciale VL, Fan DP, Braciale TJ. (1986) Difference in antigen presenation to MHC class I- and class II-restricted influenza virus-specific cytolytic T-lymphocyte clones. *J. Exp. Med.* **163**:903-315.

Mossman TR, Coffman RL. (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. *Annu. Rev. Immunol.* **7**:145-173.

Nijman HW, houbiers JG, Vierboom MP, van der Burg SH, Drijfhout JW, D'Amaro J, Kenemans P, Melief CJ, Kast WM. (1993) *Eur J Immunol.* **23**(6):1215-1219.

Nuchtern JG, Bonifacio JS, Biddison WE, Klausner RD. (1989) Brefeldin A implicates egress from endoplasmic reticulum in class I restricted antigen presentation. *Nature* **18**; 339(6221):223-6.

Nussbaum AK, Kuttler C, Tenzer S, Schild H. (2003) Using World Wide Web for predicting CTL epitopes. *Current Opinion in Immunology.* **15**:69-74.

Nussbaum AK, Kuttler C, Hadeler KP, Rammensee HG, Schild H. (2001) PAProC: a prediction algorithm for proteasomal cleavages available on the WWW. *Immunogenetics.* **53**:87-94.

Oldstone MBA, Lewicki H, Homann D, Nguyen C, Julien S, Gairin JD. (2001) Common antiviral cytotoxic T-lymphocyte epitope for diverse arenaviruses. *Journal of Virology* Vol **75**, No. 14, 6273-6278.

Oldstone MBA, Tishon A, Eddleston MJ, de la Torre C, McKee T, Whitton JL. (1993) Vaccination to prevent persistent viral infection. *J. Virol.* **67**:4372-4378.

Oldstone MBA, Tishon A, Geckeler R, Lewicki H, Whitton JL. (1992) A common antiviral cytotoxic T-lymphocyte epitope for diverse major histocompatibility haplotypes: implications for vaccination. *Proc. Natl. Acad. Sci. USA* **89**:2752-2755.

Okazaki T, Pendleton D, Lemonnier F, Berzowsky Jay. (2003) Epitope-enhanced conserved HIV-1 peptide protects HLA-A2-transgenic mice against virus expressing HIV-1 antigen. *J Immunol.* **171**(5):2548-55

Ortmann B, Copeman J, Lehner PJ, Sadasivan B, Herberg JA, Grandea AG, Riddell SR, tampe R, Spies T, Trowsdale J, Cresswell P. (1997) A critical role of tapasin in the assembly and function of multimeric MHC class I-TAP complexes. *Science* **277**: 1306-1309.

Pamer E, Cresswell P. (1998) Mechanism of MHC class –restricted antigen processing. *Annu. Rev. Immunol.* **16**:323-358.

Panigada M, Sturniolo T, Besozzi G, Boccieri MG, Sinigaglia F, Grassi GG, Gras (2002) Identification of a promiscuous T-cell epitope in Mycobacterium tuberculosis Mce proteins. *Infect Immun.* **70**(1):79-85.

Parker KC, Bednarek MA, Coligan JE. (1994) Scheme for ranking potential HLA-A2 Binding peptides based on independent binding of individual peptide side-chains. *J Immunology* **152**:163-175.

Parker KC, Bednarek MA, Hull LK, Utz U, Cunningham B, Zweerink HJ, Biddison WE, Coligan JE. (1992a). Sequence motifs important for peptide binding to the human MHC class I molecule, HLA-A2. *J Immunology* **1**; 149(11):3580-3587.

Parker KC, Dibrino M, Hull L, Coligan JE. (1992b) The β2-microglobilin dissociation rate is an accurate measure of the stability of MHC class I heterotrimers and depends on which peptide is bound. *J. Immunology* **149**: 1896.

Parker DC. (1993) T cell-dependent B cell activation. *Annu. Rev. Immunol.* **11**:331-340.

Parekh BS, Buchmeier MJ. (1986) Proteins of lymphocytic choriomeningitis virus: antigenic topography of the viral glycoproteins. *Virology* **153**:168-178.

Pascolo S, Bervas N, Ure JM, Smith AG, Lemonnier FA, Perarnau B. (1997) HLA-A2.1-restricted education and cytolytic activity of CD8(+) T lymphocytes from beta2 microglobulin (beta2m) HLA-A2.1 monochain transgenic H-2Db beta2m double knockout mice. *J Exp Med.* **185**(12):2043-51.

Peters CJ. (2002) Meny viruses are potential agents of bioterrorism. *ASM News* **68**: 168-173.

Power C, Grand C, Ismail N, Peters N, Yurkowski D, Bretscher. (1999) A valid ELISPOT assay for enumeration of ex-vivo antigen-specific IFN-γ producing T cell. J. *Immunol. Meth.* **227**:99.

Price ME, Fischer-Hoch SP, Craven RB, McKormick. (1988) A prospective study of material and fetal outcome in acute Lassa fever infection during pregnancy. *Br. Med. J.* **279**:584-587.

Rammensee H-G, Bachman J, Emmerich NP, Bachor OA, Stevanovic S. (1999) SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* **50**:213-219.
Reche PA, Gluttig JP and Reinerz EL (2002) Prediction of MHC Class I Binding Peptides Using Profile Motifs *Human Immunology*; 701-709

Robenstock E, Laster M. (2002) Viral hemorrhagic fevers a biological weapon. *Harefuah.* **141**, 96-99, 119.

Rosenberg SA, Yang JC, Schwartzentruber DJ, Hwu P, Marincola FM, Topalian SL, Restifo NP, Dudley ME, Schwarz SL, Spiess PJ, Wunderlich JR, Parkhurst MR, Kawakami Y, Seipp CA, Einhorn JH, White DE. (1998) Immunologic and therapeutic evaluation of a synthetic peptide vaccine for the treatment of patients with metastatic melanoma. *Nat Med.* **4**:321-327.

Rötzschke O, Falk K, Stevanovic S, Jung G, Walden P, Rammensee H-G. (1991) Exact prediction of natural T cell epitope. *Eur J Immunol* **21**(11):2891-2894.

Russel JH, Ley TJ. (2002) Lymphocyte-mediated cytotoxicity. *Annu. Rev. Immunol.* **20**:323-370.

Schmitz H, Kohler B, Laue T, et. al. (2002) Monitoring of clinical and laboratory data in two cases of imported Lassa fever. *Microbes and Infection* **4**; 43-50.

Schulze K, Medina E, Chhatwal GS, Guzman CA. (2003) Identification of B- and T-cell epitopes within the fibronectin-binding domain of the SfbI protein of Streptococcus pyogenes. Infect Immun. **71**(12):7197-201.

Sette A, Fikes J. (2003) Epitope-based vaccines: an update on epitope identification, vaccine design and delivery. *Curr Opin Immunol.* **15**(4):461-70.

Sette A, Newman M, Livingston B, McKinney, Sidney J, Ishioka G, Tangri S, Alexander J, Fikes J, Chestnut R. (2002) Optimizing vaccine design for cellular processing, MHC binding and TCR recognition. *Tissue Antigens.* **59**:443-451.

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.

Sette A, Sidney J. (1998) HLA supertypes and supermotifs: a functional perspective on HLA polymorphism. *Curr Opinion Immunol* **10**:478-82.

Sette A, Vitielo A, Raherman B, Fowle P, Nayersina R, Kast WM, Melief CJ, Osteroff C, yuan L, Ruppert J. (1994) The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes. *J Immunol.* **153**:5586-5592.

Sette A, Sidney J, Oserof C, del guercio MF, Southwood S, arrhenious T, et. al. (1993) HLA DR4w4-binding motifs illustrate the biochemical basis of degeneracy and specificity in peptide-DR interactions. *J Immunol* **151**(6):3163-3170.

Sette A, Sidney J, del Guercio MF, Southwood S, Ruppert J, Dahlberg C, Grey HM, Kubo RT. (1994) Peptide binding to the most frequent HLA-A class I alleles measured by quantitative molecular binding assays. *Mol Immunol.* **31**(11):813-22.

Shirai M, Arichi T, Nishioka M, Nomura T, Ikeda K, Kawanishi K, Engelhard VH, Feinstone S, Berzowsky JA. (1995) CTL responses of HLA-A2.1-transgenic mice specific for hepatitis C viral peptides predict epitopes for CTL humans carrying HLA-A2.1. J Immunology **154**: 2733-2742.

Schirle M, Weinschenk T, Stevanovic S (2001) Combining computer algorithms with experimental approaches permits the rapid and accurate identification of T cell epitopes from defined antigens. *Journal of Immunological Methods.* **257**:1-16.

Sidney J, Southwood S, Mann DL, Newman MJ, Sette A. (2001) Majority of peptides Binding HLA-A2.1 with high affinity crossreact with other A2-supertype molecules. *Human immunol.* **62**, 1200-1216.

Sidney J, Grey HM, Southwood S, Celis E, Wentworth PA, del Guercio M-F, Kubo RT Chestnut R, Sette A. (1996) Definition of HLA-A3-like supermotif demonstrates the overlapping peptide binding repertoires of common HLA molecules. *Human immunology.* **45**:79.

Sidney J, Southwood S, del Guercio M-F, Grey HM, Chestnut R, Kubo RT, SetteA (1996a) Specificity and degeneracy in peptide binding to HLA-B7-like class I molecules. *J Immunol* **157**:3468.

Sidney J, Grey HM, Kubo RT, Sette A. (1996b) Practical, biochemical and evolutionary implications of the discovery of HLA class I supermotifs. *Immunol Today* **17**: 261.

Sidney J, del Guercio M-F, Southwood S, Engelhard VH, Apella E, Rammensee HG, falk K, Rotzschke O, Takiguchi M, Kubo RT, Grey HM, Sette A. (1995) Several HLA alleles share overlapping peptide specificities. *J Immunol* **154**:247.

Simmons G, Lee A, Rennekamp AJ, Fan X, Bates P, Shen H. (2004) Identification of murine T-cell epitopes in Ebola virus nucleoprotein. *Virology.* **5**;318(1):224-230.

Simons JW, Mikhak B, Chang JF, DeMarzo AM, Carducci MA, Lim M, Weber CE, Baccala AA, Goemann MA, Clift SM, Ando DG, Levitsky HI, Cohen LK, Sanda MG,

Mulligan RC, Partin AW, Carter HB, Piantadosi S, Marshall FF, Nelson WG. (1999) Induction of immunity to prostate cancer antigens: results of a clinical trial of vaccination with irradiated autologous prostate tumor cells engineered to secrete granulocyte-macrophage colony-stimulating factor using ex vivo gene transfer. *Cancer Res.***59**:5160-5168.

Singh RA, Wu L, Barry MA. (2002) Generation of genome-wide CD8 T cell responses in HLA-A2.1 transgenic mice by an HIV-1 ubiquitin expression library immunization vaccine. *J. Immunol.* **168**:379-391.

Singh H and Raghava (2003) ProPred1: Prediction of promiscuous MHC class I binding sites. *Bioinformatics* **19**: 1015-1018

Slowik MR, Min W, Ardito T, Karsan M et. al. (1997) Evidence that tumor necrosis factor triggers apoptosis in human endothelial cells by inteleukin-1 converting enzyme-like protease-dependent and -independent pathways. *Lab. Investig.***77**: 257-267.

Smith JD, Lie WR, Gorka J, Myers N, Hansen TH. (1992) Extensive peptide ligand exchange by surface class I major histocompatibility complex molecules independent of exogenous β2-microglobulin. *Immunology* **89**: 7767-7771.

Solbrig VM, McKormick JB. (1991) Lassa fever: central nervous system manifestations. *J. Trop. Geograph. Neurol.* **1**, 23-30.

Sonderstrup G, Cope AP, Patel S, Congia M, Hain N, Hall FC, Parry SL, Fugger LH, Michie S, McDevitt HO. (1999) HLA class II transgenic mice: models of the human CD4+ T-cell immune response. *Immunol Rev.***172**:335-43.

Stoltze L, Nussbaum AK, Sijts AJ, Emmerich NPN, Kloetzel PM, Schild H. (2000) The function of proteasome system in MHC class I antigen processing. *Immunology Today.* **21**:317-319.

Strominger JL, Terasaki PI. (1997) Nomenclature for factors of the HLA system. *Human Immunol* **53**:98.

Stuber G, Leder GH, Storkus WT, Lotze MT, Modrow S, Szekely L, Wolf H, Klein E, Karre K, Klein G. (1994) Identification of wild-type and mutant p53 peptides binding to HLA-A2 assessed by a peptide loading-deficient cell line assay and a novel major histocompatibility complex class I peptide binding assay. *Eur J Immunol.* **24**(3): 765-768.

Sturniolo T, Bono E, Ding J, Raddrizzani L, Tuereci O, Sahin U, Braxenthaler M, Gallazzi F, Protti MP, Sinigaglia F, Hammer J. (1999) Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat Biotechnol.* **17**(6):555-561.

Sun Y, Song M, Stevanovic S, Rammensee HG. (2000) Identification of a new HLA-A2.1-restricted T-cell epitope from the tyrosinase-related protein 2 (TRP2) melanoma antigen. *Int. J. Cancer.* **87**:399-404.

terMeulen J, Badusche M, Satoguina J, Strecker T, Lenz O, Loeliger C, Sakho M, Koulemou K, Koivogui L, Hoerauf A. (2004) Old and New World arenaviruses share a highly conserved epitope in the fusion domain of the glycoprotein 2, which is recognized by Lassa virus-specific human CD4+ T-cell clones. *Virology* **321**:134-143.

ter Meulen J, Badusche M, Kuhnt K, Doetze A, Satoguina J, Marti T, Loeliger C, Koulemou K, Koivogui L, Schmitz H, Fleischer B, Hoerauf A. (2000) Characterization of human CD4(+) T-cell clones recognizing conserved and variable epitopes of the Lassa virus nucleoprotein. *Journal of Vorology.* **75**, 5:2186-2192.

ter Meulen, Lukashevich IS, Sidibe K, Inapogui AP, Marx M, Dorleman A, Yansane ML, Koulemou K, Chang-Claude J, Schmitz H. (1996) Hunting of peridomestic rodents and consumption of their meat as possible risk factors for rodent-to-human transmission of LFV in the Republic of Guinea. *Am. J. Trop. Med. Hyg.* **55**:661-666.

Teng MK, Smolyar A, Tse AGD, Liu JH, Liu J, Hussey RE, Nathenson SG, Chang HC, Reinherz EL, Wang JH. (1998) Identification of a common docking topology with a substantial variation among different TCR-MHC-peptide complexes. *Curr. Biol* . **8**: 409-412.

Thomson SA, Elliot L, Sheritt MA, et. al., (1996) recombinant polyepitope vaccines for the delivery of multiple CD8 cytotoxic T cell epitopes. *J. Immunol.* **157**: 822.

Tishon A, Lewicki H, van Binnendijk, Osterhaus and Oldstone MBA. (2000) Transgenic mice expressing human HLA and CD8 molecules generate HLA-restricted measles virus cytotoxic T lymphocytes of the same specificity as humans with natural measles virus infection . *Virology* **275**:286-293.

Toes RE, Hoeben RC, et. al., (1997) Protective anti-tumor immunity induced by vaccination with recombinant adenoviruses encoding multiple tumor-associated T lymphocyte epitopes in a string-of –beads fashion. *Proc. Natl. Acad. Sci. USA* **94**: 14660.

Townstead A, Elliott T, Cerundolo V, Foster L, barber B, Tse A. (1990) Assembly of MHC class I molecules analyzed in vitro. *Cell* 62:285.

Tsunoda I, Sette A, Fujinami RS, Oseroff C, Ruppert J, Dahlberg C, Southwood S, Arrhenius T, Kuang LQ, Kubo RT, Chestnut RW, Ishioka GY. (1999) Lipopeptide particles as the immunologically active component of CTL inducing vaccines. *Vaccine* **26**; 17(7-8):675-685.

Turcanova V, Hollsberg P. (2004) Sustained CD8+ T-cell immune response to a novel immunodominant HLA-B*0702-associated epitope derived from an Epstein-Barr virus helicase-primase-associated protein. *Journal of Medical Virology*. 72:635-645.

van der Burg SH, Ras E, Drijfhout JW, Benckhuijsen WE, Bremers AJ, Melief CJ, Kast WM. (1995) An HLA class I peptide-binding assay based on competition for binding to class I molecules on intact human B cells. Identification of conserved HIV-1 polymerase peptides binding to HLA-A*0301. *Hum Immunol.* **44**(4):189-198.

Vitiello A, Ishioka G, Grey HM, Rose R, Farness P, LaFond R, Yuan L, Chisari FV, Furze J (1995) Development of lipopeptide-based therapeutic vaccine to treat chronic HBV infection. Induction of a primary cytotoxic T lymphocyte respoponse in humans. *J Clin. Invest.* **95**(1):341-9.

Vonderheide RH, Hahn WC, Schultze JL, Nadler LM. (1999) The telomerase catalytic subunit is a widely expressed tumor-associated antigen recognized by cytotoxic T lymphocytes. *Immunity*.**10**(6):673-679.

Wang JH, Reinherz. (2002) Structural basis of T cell recognition of peptides bound to MHC molecules. *Mil. Immunol.* **38**, 1039-1049.

Weinschenk T, Gouttefangeas C, Schirle M, Obermayr F, Walter S, Schoor O, Kurek R, Loeser W, Bichler KH, Wernet D, Stevanovic S, Rammensee HG. (2002) Integrated functional genomics approach for the design of patient-individual antitumor vaccines. *Cancer Res.* **62**(20):5818-27.

Weiss GA, Collins EJ, Garboczi DN, Wiley DC, Schreiber SL. (1995) *Chem . Biol.* **2**:401-407.

Wong P and Pamer EG (2003) CD8 T cell responses to infectious pathogens. *Annu. Rev. Immunol.* **21** : 29-70.

Whitton JL, Sheng N, Oldstone MBA, Mc Kee TA. (1993) A "string-of-beads" vaccine, comprising linked minigenes, confers protection from lethal dose virus challenge. *J. Virol.* **67**: 348.

Whitton JL, Tishon A, Lewicki H, Gebhard J, Cook T, salvato M, Oldstone MBA. (1989) Molecular analyses of a five amino acid cytotoxic T lymphocyte (CTL) epitope: an immunodominant region which induces nanreciprocal CTL cross-reactivity. *J. Virol.* **63**:2527-2531.

Whitton JL and Oldstone MBA. The immune response to viruses. Fields virology, 4th edition. Lippincott Williams & Wilkins, Philadelphia, PA.
Yang Y (2003) Generation of major histocompatibility complex class I antigens. *Microbes and Infection* **5**, 39-47.

York IA, Rock Kl. (1996) Antigen processing and presentation by the class I major histocompatibility complex. *Annu. Rev. Immunol.* **14**, 369-396.

Zinkernagel RM, Doherty PC. (1997) The discovery of MHC restriction. *Immunol Today* **18**(1):14-17.

Zinkernagel RM. (2003) On natural and artificial vaccinations. *Annu. Rev. Immunol.* **21**: 515-546.

BIMAS prediction server (September 10, 2003) http://bimas.dcrt.nih.gov/molbio/hla_bind.

Category A Priority Pathogens (September 12, 2003) http://www.niaid.nih.gov/biodefense/bandc_priority.htm

HIV molecular immunology database (September 15, 2003) http://hiv-web.lanl.gov/content/immunology/index.html/

MAPPP prediction server (October 17, 2003) http://www.mpiib-berlin.mpg.de/MAPPP/binding.html

MHCBN database (September 20, 2003) http://www.imtech.res.in/raghava/mhcbn/

NIH Blast server (September 15, 2003) http://www.ncbi.nlm.nih.gov/blast/.

nHLAPred prediction server (October 17, 2003) http://www.imtech.res.in/raghava/nhlapred/

ProPred-I prediction server (October 17, 2003) http://www.imtech.res.in/raghava/propred1/

PREDEP prediction server (October 17, 2003) http://bioinfo.md.huji.ac.il/marg/Teppred/mhc-bind/mhc form

RANKPEP prediction server (October 17, 2003) http://mif.dfci.harvard.edu/Tools/rankpep.html

SYFPEITHI prediction server (September 10, 2003) http://syfpeithi.bmi-heidelberg.com/Scripts/MHCServer.dll/EpPredict.htm

SVMHC prediction server (September 10, 2003) http://www.sbc.su.se/svmhc/

Tetramer facility, Emory University, Atlanta (September 15, 2003) http://emory.edu/WHSC/TETRAMER/