Determination of quantitative peptide-binding motifs of four common equine MHC class I alleles, and identification of an equine herpesvirus type 1-derived cytotoxic T lymphocyte epitope

Dissertation zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

am Fachbereich Biologie, Chemie, Pharmazie

der Freien Universität Berlin

vorgelegt von

Tobias Bergmann

Berlin, Deutschland

Mai 2016

Diese Promotionsarbeit wurde im Zeitraum vom Januar 2012 bis zum Mai 2016 am Institut für Virologie der Freien Universität Berlin unter der Leitung von Prof. Dr. Nikolaus Osterrieder angefertigt.

1. Table of contents

2. List of figures and tables

3. Abbreviations

aa	amino acid
ABC	ATP-binding cassette
APC	antigen-presenting cell
ARB	average relative binding
ATP	adenosine triphosphat
BAC	bacterial artificial chromosome
BHV	bovine herpesvirus
BSA	bovine serum albumin
CD	cluster of differentiation
CNS	central nervous system
ConA	concanavalin A
cTEC	cortical thymic epithelia cell
CTL	cytotoxic T lymphocyte
ddH$_2$O	double distilled water
DMSO	dimethyl sulfoxide
DMP	dimethyl pimelimidate
DNA	deoxyribonucleic acid
DRiPs	defective ribosomal products
dsDNA	double-stranded DNA
E	early
EDTA	ethylenediaminetetraacetic acid
EHV	equine herpes virus
EIAV	equine infectious anemia virus
ELA	equine leukocyte antigen
ELISpot	enzyme-linked immunospot
ER	endoplasmic reticulum
ERAP	ER-resident aminopeptidases
Eqca	*Equus caballus*
FCS	Fetal calf serum
FT	flowthrough
gB	glycoprotein B
gC	glycoprotein C
gD	glycoprotein D
gE	glycoprotein E
gG	glycoprotein G
GFP	green fluorescent protein
gH/gL	glycoprotein H/glycoprotein L complex
HBC	hepatitis B virus
HCMV	human cytomegalovirus
HLA	human leukocyte antigen
HSV	herpes simplex virus

HVEM	herpesvirus entry mediator
ICP0	infected cell protein 0
ICP4	infected cell protein 4
ICP47	infected cell protein 47
IE	immediate early
IFN-γ	interferon-γ
IR	internal repeat
kbp	kilo base pair
KSHV	Kaposi sarcoma herpesvirus
L	late
LPS	lipopolysaccharide
MEM	minimal essential medium Eagle
MHC	major histocompatibility complex
MHC-I	MHC class I molecule
MLV	modified-live vaccine
NK cell	natural killer cell
O/N	over night
ORF	open reading frame
RNA	ribonucleic acid
rpm	rotations per minute
rt	room temperature
P	position
PAMPs	pathogen-associated molecular patterns
PBMC	peripheral blood mononuclear cell
PBS	phosphate buffered saline
PBST	PBS Tween
PCR	polymerase chain reaction
pi	post infection
PLC	peptide loading complex
PMSF	phenylmethylsulfonyl fluoride
PRR	pattern recognition receptors
PrV	pseudorabies virus
PSCL	positional scanning combinatorial libraries
P/S	penicillin/streptomycin
SD	standard deviation
SF	specificity factor
SFC	spot-forming cell
TAE	Tris-acetate-EDTA buffer
TAP	transporter associated with peptide processing
Th cell	T helper cell
TLM	translocation motif
TNF	tumor necrosis factor
TPP	tripeptidyl peptidase
U_L	unique long
U_S	unique short
VZV	varicella zoster virus

4. Introduction

Every organism on earth is target of constant attacks by microorganisms seeking to invade and procreate inside a host's body. An infection can be associated with disease or remain subclinical. The major defense against intruding pathogens is the immune system, which evolved in a constant arms race with microorganisms. The immune system is based on three tightly regulated columns. The first line of defense is the intrinsic immune system, which is always present and does not have to be induced. It includes physical barriers like the skin and mucosae as well as protein molecules with more specific functions. The second column represents the innate immune system composed of a variety of receptors, cytokines and immune cells. In contrast to intrinsic resistance, the innate immune system has to be activated upon contact with an antigen. Patrolling phagocytic cells internalize extracellular particles and recognize conserved patterns associated with pathogens (PAMPs: Pathogen associated molecular patterns), e.g. lipopolysaccharide of gram-positive bacteria, unmethylated CpG-DNA, or double-stranded RNA as an intermediary replication product of some viruses, with their diverse pattern recognition receptors (PRR). Recognition of PAMPs results in the production of pro-inflammatory cytokines like interleukins, type I and II interferons as well as tumor necrosis factor (TNF). That attracts the full complement of effector cells leading to activation of the adaptive immune system, which constitutes the third column. The adaptive immune system requires prolonged contact with the specific pathogen in order to mount a highly specific and effective immune response. Key players are B lymphocytes, which produce soluble epitope-specific immunoglobulins (Ig), and T lymphocytes, which specifically recognize intra- and extracellular antigens presented by infected or professional antigen presenting cells (APCs), respectively.

Once a cell has been identified as infected, $CD8^+$ cytotoxic T lymphocytes (CTL) destroy the particular cell. Key players in the process of recognition are major histocompatibility complex (MHC) molecules, which present pathogen-derived peptides on the surface of the cell. Of particular importance and interest are classical MHC class I and II (MHC-I, -II) molecules as they are directly involved in the recognition of cell-associated as well as cell-free pathogens. Both present peptide epitopes to T lymphocytes, which specifically recognize the complex via

their polymorphic T cell receptor (TCR). Class I molecules mainly associate with endogenous peptides, and in case of an intracellular infection, pathogen-derived peptides are presented to $CD8^+$ CTL, which upon binding proliferate and consequently destroy the infected cell. Thymic self-intolerance ensures that only epitopes of pathogen origin will effectively activate effector functions of CTLs. Class II molecules act in a similar way as they present exogenous peptides degraded in lysosomes to $CD4^+$ helper T lymphocytes (Th cells). However, upon activation Th cells do not directly destroy the activating cell but secrete a bandwidth of immunologically active cytokines and chemokines to orchestrate the immune response (summarized from [1]).

4.1 Herpesviruses

The order of Herpesvirales comprises a vast number of viruses capable of infecting virtually all vertebrate and even some invertebrate species. They are grouped into three families termed Herpesviridae, Alloherpesviridae and Malacoherpesviridae. Viruses infecting mammals, birds and reptiles fall into three subfamilies of the Herpesviridae, namely alpha-, beta- and gammaherpesvirinae, which all share structural features in their virion composition, envelopes and their genomes as well as their ability to cause latent infections [2, 3]. However, alphaherpesviruses have a fast replication cycle, a broad host range and establish latency in sensory ganglia whereas betaherpesviruses replicate slowly and are more restricted in their host range. The same is true for gammaherpesviruses, which mainly target immune cells for infection [4].
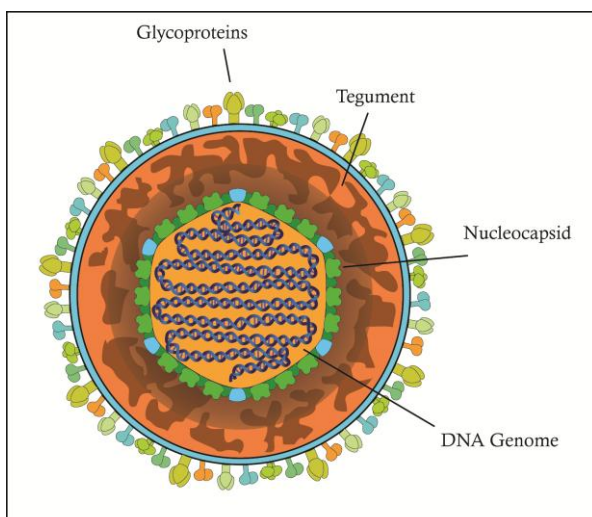
Figure 1: **A herpes simplex virus particle. The DNA genome is contained within the nucleocapsid, which is surrounded by the amorphous tegument. A lipid bilayer forms an envelope around it spiked with surface glycoproteins.**

All herpesviruses have a linear, double-stranded DNA genome with sizes ranging from $108 - 300$ kilo base pairs (kbp). For alphaherpesviruses, it is divided into a long and a short region, which harbor unique sequences (UL: unique long, US: unique short), flanked by internal and terminal repeats. Their large genome can sustain almost an entire DNA replication machinery enabling them to replicate independently from the cell cycle, and an array of immune modulatory genes, which help establish latent infections [3, 5].

A virion of the prototypic herpes simplex virus (HSV) consists of an icosahedral nucleocapsid formed by 162 capsomers associated with genomic DNA [3] surrounded by an amorphous protein structure. This tegument forms an inner layer that is engulfed in a host cell-derived membrane harboring several glycoproteins essential for adsorption and penetration as well as immune modulation (Figure 1). Apart from having a mere scaffolding function, proteins of the tegument are widely important for initiation and implementation of the infection [3].

4.1.1 Replication cycle of equine herpesvirus type 1

In equid species, nine different herpesviruses have been identified, which can be classified into the subfamilies of alpha- and gammaherpesviruses. Horses are the natural hosts for alphaherpesviruses equine herpesvirus (EHV) -1, -3 and -4 as well as the gammaherpesviruses EHV-2 and -5 [3, 6-8]. In economical, clinical and epidemiological terms most important are EHV-1 and -4, which share a DNA sequence homology of up to 90 % [9]. EHV-1 is the causative agent of rhinopneumonitis, late-term abortion and neonatal foal disease as well as equine herpesvirus myeloencephalopathy (EHM). As typical for herpes viruses, EHV-1 has a high seroprevalence in domestic horse breeds ranging from 30 – 80 % [31].

EHV-1 is an alphaherpesvirus of the genus varicellovirus with a genome size of around 150 kbp, and harbors 80 open reading frames coding for at least 77 proteins [10, 11]. Entry into cells follows the typical mechanism that is known for most herpesviruses including the prototypic HSV-1. Initial attachment is mediated through a weak interaction of the glycoproteins gC and gB to cellular glucosaminoglycan and subsequent binding of gD to its specific cellular receptor [12-14]. For HSV-1, nectins, herpesviral entry mediator (HVEM) and heparan sulfate have been identified as gD-specific entry receptors [14]. EHV-1 can also use alternative receptors like MHC class I molecules preferably from alleles carrying alanine at position 172 in the α2 region [15-17]. However, as MHC class I molecules are constitutively expressed on all cell types, there may be other factors involved in cell tropism. Upon gD binding, it complexes with gB and gH/gL, probably inducing conformational changes in the fusogenic gB [14]. After attachment, the virus can enter a cell by either direct fusion with the cell membrane or from within endosomes at acidic pH. Glycoprotein H and its cellular interactor α4β1 integrin have been found to route the virion either way [18]. Once the nucleocapsid has been released into the cytosol, it is transported along microtubules and the DNA genome is released into the

nucleus where it circularizes [19]. During lytic replication, the whole set of viral genes is expressed allowing for a high production of progeny virus.

Herpes viral gene expression is temporally regulated and occurs in three phases. Immediate early (IE) genes are the first to be transcribed and they include other transcriptional regulators. The EHV-1 sole IE gene encodes an ICP4 (infected cell protein 4) homologue capable of binding several other viral proteins as well as cellular transcription factors like TFIIB [20-23]. The first viral proteins present in the cytoplasm, however, are tegument proteins including transactivators, which initiate immediate early gene expression, or modulating factors like the viral host shut-off protein (vhs) that degrades cellular mRNAs in order for the virus to take over cell metabolism [3, 24]. In the next step, proteins with early (E) kinetics are produced, which are mainly associated with DNA replication [3, 25]. Structural proteins that form the tegument, capsid and glycoproteins are produced as part of the late (L) phase [26]. One hallmark of herpes viruses is their ability to establish persistent infections that do not lead to the generation of progeny since gene expression is reduced to a minimum. In this latency stage, the viral genome is either present as a circular episome tethered to or for some herpesviruses as a provirus integrated into the cell chromosomes [27]. As almost no viral proteins are produced and thusly presented to T lymphocytes, herpesviruses can persist for a prolonged time in the host body without evoking an immune response. Primary cellular targets for latency are trigeminal ganglia for most alphaherpesviruses, or lymphatic tissues for beta- and gammaherpesviruses [3]. EHV-1 although being an alphaherpesvirus primarily latently infects $CD5^+/CD8^+$ T lymphocytes, but trigeminal ganglia have also been described as site of latent EHV-1 infections in horses [28, 29]. For the virus to infect a new host, however, it has to reactivate and enter the lytic replication cycle to produce infectious progeny [3].

4.1.2 Pathogenesis of EHV-1 infections

The highly contagious EHV-1 is usually transmitted by direct contact through infected nasal discharge and in rare cases through aerosols or contaminated food or water [30]. Respiratory infection is characterized by shedding of virus through infected discharge, fever and anorexia [31]. EHV-1 primarily replicates in the epithelium of the upper respiratory tract (URT) and the nasal epithelium as early as 12 h post infection (pi). The virus cannot cross the basement membrane directly, but modulation of the extracellular matrix proteins facilitates shuttling to local lymph nodes via infected resident $CD172^+$ monocytes where it infects more leukocytes resulting in a cell-associated viraemia (Figure 2) [32, 33]. This is a prerequisite for the rapid

dissemination of the infection to secondary sites of replication like the endothelia of the endometrium and the central nervous system (CNS) [34-41]. An estimated 1 in 10^4 to 10^6 lymphocytes is infected leading to high levels of viraemia. Production of viral antigen indicative of viral replication is predominantly found in monocytes, $CD8^+$ and to a lesser extent in $CD4^+$ T cells [35, 41, 42]. Consequently, experimental infection of pathogen free foals (SPF) results in both a lymphopaenia and neutropaenia within the first week after infection [39].

Upregulation of adhesion molecules on the surface of infected lymphocytes facilitates adhesion to endothelia and spread of the virus through cell-to-cell contacts, which is the primary distribution route from peripheral blood mononuclear cells (PBMCs) to endothelial cells rather than by viral egress [43, 44]. In pregnant mares, infection of the endothelia in the endometrium of the uterus might result in severe vasculitis, thrombosis and subsequent fetal hypoxia, or focal transfer of virus-infected leukocytes in the microcotyledonic interface of mother and fetus by diapedesis, both leading to abortion of an uninfected or infected fetus in pregnant mares, respectively. Neonatal foal disease of newly born foals is the result of near-term transplacental infection [37, 45, 46]. Infection of the vascular endothelial cells in the CNS results in EHM. Increasing numbers of neurological disease have been observed in recent years in EHV-1 outbreaks prevalent for certain strains, e.g. Ab4 [47]. In 2006, Nugent et al. found a strong association between the neuropathogenic phenotype and a single mutation in the gene encoding the viral DNA polymerase (UL30) [48, 49]. The neuropathogenic strain was found

Figure 2: Pathogenesis of EHV-1 infections. The virus infects the epithelium of the URT and resident leukocytes (1), which carry the virus through the basement membrane to local lymph nodes (2). From there, it spreads via leukocyte-associated viraemia (3) to secondary sites of infection, e.g. the endothelia (4) of the endometrium and the CNS

to replicate more efficiently in horses leading to a prolonged viraemia [50]. However, around 25 % of horses with neurological disorder were infected by EHV-1 strains without this mutation suggesting more factors to be involved in EHM onset [47].

4.1.3 Immune response to EHV-1 infections

EHV-1 has an impressive capacity to evade the immune response of the host in order to stay under the immunological radar. Subversion of recognition and destruction of infected cells by antibody-dependent or CTL-mediated cell lysis as well as modulation of the host cytokine response are potent immune evasive strategies [51]. In the humoral immune response, antibodies specific for viral envelope proteins have two major functions. On the one hand, they recognize cell-free viral particles and associate with them, which by itself may neutralize viral infectious abilities, and greatly increase receptor-mediated phagocytosis by activated macrophages [1]. Antibody-dependent cell-mediated cytotoxicity (ADCC), on the other hand, relies on the presence of viral proteins on the surface of an infected cell. Natural killer cells (NK) display an Fc receptor (CD16) capable of binding IgG associated with viral antigen, which activates the NK cell to secrete type 2 interferon (IFN-γ), a potent antiviral immune modulator, as well as granzymes that lyse the target cell [1]. In a similar process, antibody-opsonization triggers the complement system to initiate its proteolytic cascade that results either in direct destruction of the infected cell by formation of pores in the cell membrane through the membrane-attack complex (MAC), or in enhanced phagocytosis [1].

As mentioned before, EHV-1 infection results in a leukocyte-associated viraemia, which shields viral particles from antibody recognition. Additionally, no viral envelope proteins can be detected on infected peripheral blood mononuclear cells (PBMCs), which could circumvent ADCC [52]. Several viral envelope proteins can function as decoy receptors. Glycoprotein C (gC) binds to the complement factor C3 interfering with the activation of the complement cascade, and gG associates with a broad range of chemokines inhibiting their immunomodulatory functions [53, 54]. Protective immunity against infection of horses with respiratory pathogens is mainly attributed to virus neutralizing mucosal IgA antibodies [55, 56]. Consequently, challenge infection of horses with virulent EHV-1 strains results in a mucosal humoral immune response with neutralizing IgA being the predominant antibody isotype [57]. Inactivated vaccines, however, fail to induce protective IgA titers [57]. This might explain why circulating antibodies alone do not protect against infection and why mucosal as well as cellular immunity is required [58, 59]. Cellular immunity, namely CD8[+] CTL activity, is dependent on viral epitopes being presented on MHC class I molecules of an infected cell. As outlined later on, viral epitopes are generated in a multistep process, and in herpesvirus infections any step can be tackled by viral evasive strategies. EHV-1 codes for two proteins, pUL49.5 and pUL56, which interfere with the antigen processing pathway [60, 61].

Nevertheless, despite all those evasive strategies some degree of protection and reduction in viral shedding can be achieved through vaccination reducing the number of abortion storms and overall veterinary burden [59, 62]. Several inactivated vaccines are available, which can induce virus neutralizing (VN) as well as complement-fixing (CF) antibody titers reducing the severity and duration of infection, but their efficacy and their potential to prevent viraemia and consequently abortion or EHM is very limited [58, 59, 63]. Frequent re-immunizations are necessary since acquired immunity is particularly short-lived and re-infections can occur within 3 – 6 months [59].

Increasing evidence suggests that cell-mediated immunity plays a major role in protection from and control of EHV-1 infections and associated diseases. Virus-specific CTL frequencies in peripheral blood of uninfected or with inactivated vaccine immunized horses are undetectable or very low and those animals develop typical respiratory symptoms and viraemia, and almost all pregnant mares abort upon challenge infection [64, 65]. Only multiply infected horses are completely protected with no detectable viraemia but significantly increased virus-specific precursor CTL frequencies [65]. This strong association suggests an EHV-1 infection to be controlled predominantly by virus-specific cytotoxic T lymphocyte activity. Additionally, CD4$^+$ T cells also seem to play a role in protection, which might be attributed to direct antiviral effects of IFN-γ [66, 67]. In these studies, the majority of IFN-γ-producing lymphocytes were CD8$^+$, but also CD4$^+$ T cells were shown to contribute to IFN-γ production, which increases after multiple EHV-1 infections. NK cells seem not to play a big role in the IFN-γ response [68].

Vaccination strategies that target the induction of cellular immunity classically rely on attenuated virus strains [1]. One modified-live vaccine (MLV) based on the attenuated strain RacH is currently available for EHV-1 [69]. In regard to the virulent strain Ab4, it harbors several deletions in its genome including ORF1, which codes for pUL56, and ORF2, glycoprotein 2 and IR6 [70]. Although it confers protection against respiratory disease, its influence on viraemia, abortion and EHM incidence is still controversial [59]. Several studies have shown that it induces high titers of neutralizing antibodies but its effect on viraemia is very limited [71-73]. Those limitations, in addition to the safety concerns associated with live vaccine applications particularly in pregnant mares, make the MLV for EHV-1 less than an ideal vaccine.

To better understand the poor efficacy of vaccines, a closer look at the CTL response was taken. For equine infectious anemia virus (EIAV), few CTL epitopes have been identified that are restricted primarily by loci of the MHC class I haplotypes ELA-A5 and -A9 [74-76]. For

EHV-1, the ICP4 (IE) protein encoded by gene 64 has been shown to contain CTL epitopes restricted by MHC class I molecules of the ELA-A3 and −A2 haplotypes, however, a precise peptide epitope remains to be defined [77-79]. Using a vaccinia virus-based vector vaccine expressing the ICP4 gene (NYVAC-IE), or alternative approaches using particle-mediated DNA vaccines coding for gC, gB, gD, ICP4 or an ICP27 homologue pUL5 failed to induce sufficient protection against EHV-1 challenge infection [80, 81]. However, animals used in these studies were not MHC-I haplotyped. A later study showed partial protection of ELA-A3.1$^+$ (Eqca-1*00101) ponies vaccinated with NYVAC-IE against EHV-1infection. The vaccinated ponies showed increased IFN-γ mRNA expression and reduced viraemia [82].

In conclusion, there is a variety of inactivated vaccines and one MLV in the market with several alternative vaccine designs in clinical trials. Although most of them induce high serum titers of neutralizing antibodies, induction of mucosal as well as cellular immunity, which are considered key players in protection from and control of EHV-1 infections, is very poor.

4.2 The Major Histocompatibility Complex

Early in the 20th century CC Little and EE Tyzzer working with Japanese waltzing mice observed successful transplantation of tumors in F1 generations parented by mice of different strains but rejection by almost all mice in the F2 generation. They concluded that the compatibility of tissues is dependent on a number of different hereditable factors and that chances of an F2 offspring to possess all those factors are very low [83, 84]. Subsequently, PA Gorer found rejection reactions to be associated with four different antigens (antigen I, II, III and IV) and GD Snell located them to a distinct genetic area, which he called H (histocompatibility) genes, defining the H-2 haplotype in mice [85, 86]. Decades later, Zinkernagel and Doherty realized that T cell responses are restricted to major histocompatibility antigens, which became known in humans as the human leukocyte antigen (HLA) located on chromosome 6 [87]. Soon it became clear that the genetic locus coding for HLA or H-2 molecules comprises an abundance of genes and that this cluster is common to all jawed vertebrates [88], so it was renamed to major histocompatibility complex (MHC).

As it is defined today, the human MHC encodes several hundred genes and pseudogenes participating in the adaptive and innate immune response, and spans over 4 million base pairs [89]. It can be subdivided into MHC I, II and III regions, which accumulate immunomodula-

tory genes. The genes for classical as well as nonclassical MHC class I molecules are located in the MHC I region, for class II molecules in the MHC II region, and the MHC III region harbors genes, most of which are involved in the production of functional class I and II proteins or take part in inflammatory processes [89].

4.2.1 MHC class I genetics

Ubiquitously found on all nucleated cells, primary function of MHC class I molecules (MHC-I) is to convey the immune status of a cell by sampling the current proteome [1]. These molecules bind intracellular peptides with certain corresponding chemical specificities and transport them to the cell surface for T cell scrutiny. In context of an intracellular infection, for example, T cell clones with TCRs specific for a given complex of MHC-I and viral peptide bind, get activated and consequently destroy the target cell [1].

As a result, and perhaps not surprisingly, a strong selection pressure drives pathogen evolution towards escape variants that circumvent MHC-I and TCR specificities. To counter that, the classical MHC-Ia locus is polygenic and highly polymorphic. In humans, three different genetic loci encoding for classical MHC class Ia molecules are expressed. These loci, called Human Leukocyte Antigen (HLA)-A, -B or –C, are inherited as a unit (haplotype) by Mendelian rules. The A, B and C loci are amongst the most diverse genes in the human genome, with 3285, 4077 and 2801 alleles described to date, respectively, leading to a tremendous amount of different allele combinations on population levels (http://hla.alleles.org/nomenclature/stats.html, September 2015). This high allelic variation is due to point mutations and shared sequence motifs arisen by gene conversion events [90]. Most individuals are heterozygous doubling the number of MHC-I molecules to six.

In humans, three nonclassical MHC-Ib loci (HLA-E, -F, -G), characterized by lower allele numbers ranging from 18 to 51 (http://hla.alleles.org/nomenclature/stats.html, September 2015) are also expressed. The MHC-Ib molecules seem to have some peptide presenting properties, as a few viral peptides have been found to stabilize HLA-E molecules at 26 °C, although no evidence that they can interact with TCRs has been provided. Rather, they appear to play immunomodulatory roles by regulating T cells, NK cells, macrophages, dendritic cells and other immune cells [91].

The equine MHC class I counterpart (Equine Leukocyte Antigen: ELA) is located on chromosome 20 and comprises up to 30 classical and nonclassical MHC class I loci as well as pseudogenes spanning 4000 bp in eight exons [92, 93]. Serologically, 19 ELA-A haplotypes have been distinguished, to date, each with similar levels of polymorphism estimated using intra-MHC microsatellites and sequencing [94-96]. Using the same techniques, 27 microsatellite haplotypes that did not correspond with any serological haplotype were identified [94]. The number of expressed loci, however, may vary but up to seven different class I molecules can be detected in horses homozygous for the ELA-A3 haplotype [97, 98]. In 2010, Tallmadge and al. were able to assign almost 50 alleles to shared loci of ten different haplotypes (Figure 3), and to construct a phylogenetic tree based on sequence similarities [95].

Equine MHC Class I Loci

ELA Haplotypes	1	16*	2	17*	3	4	18*	5	6	7
A2		16*00101	2*00201					5*00101	6*00101	7*00101
A3	1*00101		2*00101		3*00101	4*00101		5*00101	6*00101	7*00201
A5		16*00201	2*00302	N*00201					6*00101	7*00101
A9	1*00201	16*00301	2*00102	N*00301				5*00101	6*00201	7*00101
A10	1*00301		2*00303	N*00401			N*00501	5*00101	6*00201	7*00101
A1			2*00202	N*00302			N*00601			
A4	1*00401		2*00301							
A6	1*00501	16*00401								
A7	1*00601	16*00501	2*00401							
W11	1*00701	16*00601	2*00304							
Total	7	6	9	4	1	1	2			

Figure 3: Shared classical (blue numbers) and non-classical (black numbers) MHC-I loci for 10 ELA-A haplotypes [95].

As racing horse breeds, Thoroughbreds and Standardbreds are economically important. Serological as well as molecular typing of these breeds indicates genetic diversity that is represented by only a limited number of MHC haplotypes (ELA-A2, -A3, -A5, -A9 and -A10) [94, 99]. The ELA haplotypes A3, A2 and A9, for example, are present in 25 %, 17 %, and 14 %, respectively, of Thoroughbred horses [99, 94]. The evolution of the domestic horse lineage is still controversially debated. The origin of wild horse *Equus ferus* domestication some 160 thousand years (kyr) ago can be located to the western Eurasian steppe followed by a demic

spread of herds to the west adjoined by an east-to-west decline in genetic diversity [100]. The high level of matrilineal diversity and low level of Y chromosome variability suggest a predominantly female introgression into domestic herds, which might reflect the strong patrilineal bottleneck in western central Eurasia around 125 kyr ago [101-103]. The discovery of the fossils of a Pleistocene horse confirmed that the only still existing wild horse, the Przewalski horse, falls into a different phylogenetic clade than modern domestic horses, and thusly is not a direct ancestor. [104]. However, 29 genetic loci have been identified that show significantly lower polymorphism in modern compared to Przewalski horses, probably due to domestication [104]. In contrast, evolution of dogs is the result of more than 500 individual domestication events of wolfs leading to a significantly higher MHC diversity [105]. In that light, it seems possible that the limited equine MHC haplotype diversity might be a consequence of millennia of breeding procedures aimed at optimizing particular characteristics of the horse, MHC diversity not being one of them.

4.2.2 Structure of MHC class I molecules

The functional MHC class I molecule is a type I transmembrane protein consisting of an α chain with three distinct domains (α1, α2 and α3) and a smaller β chain, the soluble β2-microglobulin. The α1 and α2 domains together form a binding cleft or groove with each domain contributing an α helix forming the walls and a β sheet on the bottom [1]. The α3 domain spans the cellular membrane with a short C-terminal domain in the cytosol, and provides an interaction site for the T lymphocyte co-receptor CD8 [1, 106]. The binding groove accommodates peptide ligands of, typically, eight to ten amino acids in length, although longer and rarely shorter peptides have been observed. As both ends of the class I groove are closed, bound peptides generally fit their N- and C-termini completely into the groove, although in some cases, as has been shown for a calreticulin-derived decamer, peptides may extend beyond the N- or C-terminus. Peptides bind in an extended conformation, but may bulge out in the middle, exposing certain residues for T cell scrutiny [107, 108].

The highly polymorphic genes of the MHC-I loci give rise to very diverse proteins where most of the polymorphic positions can be found lining the peptide binding cleft [1]. In general, the binding cleft manifest six pockets (A-F), each engaging certain positions in the peptide and mediating specific interactions [109]. The residues forming the pockets confer a high degree of chemical specificity, with high affinity ligands generally possessing corresponding

nature (e.g., hydrophobic pockets engage hydrophobic residues). Of greatest importance are the B and F pockets, which bind the residues in position 2 (P2) and the C-terminus, respectively, of the peptide. In addition to the B and F primary anchor pockets, an important role is played by pocket A, which typically binds the positively charged peptide's N-terminus (P1) through conserved tyrosine residues [110, 111], although A pocket engagement may also be driven by other specificities. Other secondary anchor positions along the peptide, such as P3 and P7, also contribute to specificity, but to a much lesser extent (Figure 4). The single contributions of any position to overall binding define a binding motif, which can be summarized in terms of primary and secondary anchor residues.
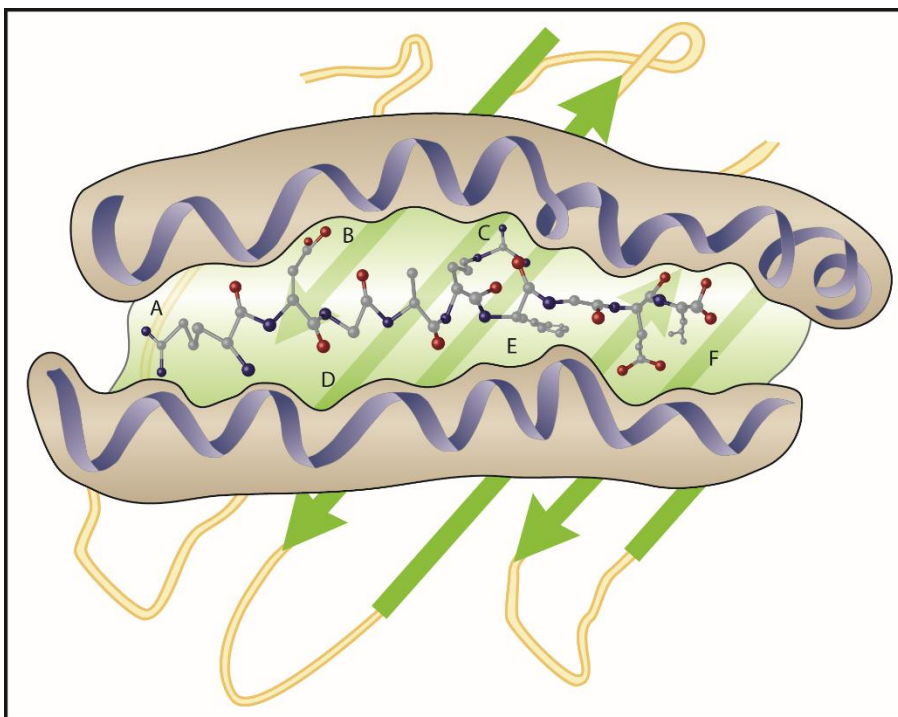
Figure 4: Schematic representation of the binding groove of an MHC class I molecule with bound peptide. Letters A – F represent binding pockets.

The majority of HLA class I alleles have primary anchors in the second and last position but other combinations have been described. Alleles with similar specificities can be organized into about twelve different HLA supertypes, which describe sets of alleles sharing common structural motifs in the respective pockets of the binding groove [112], and which thereby share correspondingly similar chemical specificity and somewhat overlapping peptide repertoires. Hydrophobic and aromatic F pocket specificities are common to most of the HLA alleles with exceptions preferring basic amino acids [112]. Specificities in the B pocket are more varied. For example, some supertypes are associated with acidic residues, while others demonstrate specificity for aliphatic/hydrophobic amino acids. The polymorphic residues lin-

ing the wall of the MHC groove have an intimate impact on TCR recognition. As shown for two alleles of the ELA-A1 haplotype, Eqca-N*00601 and Eqca-N*00602, which differ by only one amino acid in the α2 domain, CTL activity was abolished by a single substitution of E to A causing the EIAV peptide Gag-GW12 to bulge differently in the middle and consequently abolish TCR binding [113, 114].

MHC-I binding specificities can be quantified as a function of the affinities associated with the presence of specific amino acids in a respective position. These types of analyses can be undertaken by assessing the frequency of various residues in specific positions of peptides eluted from MHC molecules, or by screening larger panels of peptides for binding to purified MHC molecules. For humans, mice, primates and some other species, an impressive array of MHC class I specificities has been described, whereas for most species, including equids, data is very limited [115].

4.2.3 Antigen processing for the MHC class I pathway

From its origin in a protein to loading onto an MHC-I molecule, a peptide ligand undergoes a variety of processing steps that aim at generating a peptide pool with preferred specificities for most MHC-I molecules. Many peptides in this pool have their origin in once functional proteins that are recycled as part of normal protein turnover. However, cells infected with influenza virus are recognized by T cells within 1.5 hours pi, rather than after 10 hours when degradation of stable proteins takes place [116]. Due to defective transcription or translation, alternative reading frame usage or missfolding, 30-70 % of all proteins are defective ribosomal products (DRiPs), which are immediately degraded providing real-time surveillance of proteins that are being produced [117].

Proteins for degradation are often tagged with polyubiquitin molecules by E1, E2 and E3 proteins to direct them to a multicatalytic protein complex, the 26S proteasome [118], which is the major source of peptides in a cell. It is present in the cytoplasm as well as the nucleus of all eukaryotic nucleated cells [119]. Structurally, it resembles a barrel with two 19S caps engulfing the 20S catalytic core [120]. The actual proteolytic activity falls to two β rings of the 20S core, each of them containing three active sites with distinct but partially overlapping cleaving specificities (chymotrypsin-, trypsin-, and post-glutamyl peptide hydrolase-like) [121]. Human proteasomes have been found to cleave a substrate preferably after hydrophob-

ic, basic and aromatic residues, generating preferred C-termini for the peptide transporter TAP and MHC-I molecules [122].

Upon IFN-γ stimulation, subunits β1, β2 and β5 of the β rings are replaced by the alternative subunits LMP1 and 7, which are encoded in the MHC, as well as MECL-1, forming the immunoproteasome [122]. This modification alters the catalytic activity towards an increased production of N-terminally extended peptides to broaden downstream trimming by peptidases, which is thought to increase the pool of putative MHC-I ligands [124]. However, even more importantly, it seems that the proteolytic activity of the immunoproteasome is increased in an inflammatory background. As part of the innate immune reaction, IFN-γ stimulated activity of iNOS (inducible nitric oxide synthase) creates free nitric oxide radicals that aim at generally promoting inflammation by vasodilation, and damaging the DNA and proteins of the pathogen [125-127]. The unselective oxidation of DNA and proteins also affects the cell leading to an increased number of DRiPs, which have to be degraded in order to prevent protein aggregation. The increased activity of the immunoproteasome could account for those high protein contents and provide a way to quickly generate MHC-I ligands for detection by T cells. In a subset of cortical thymic epithelial cells (cTEC), the immunoproteasome is constitutively expressed. However, predominantly present is the thymoproteasome that like the immunoproteasome harbors LMP-1 and -2, but also another alternative subunit β5i, whose structural differences to the constitutive β5 subunit suggest a weaker chymotrypsin-like activity [128]. This subunit can also directly interact with the cytoplasmic domains of TAP1 and TAP2 suggesting a direct transfer of peptides from the proteasome through TAP into the lumen of the ER [129].

Many herpes viruses have adopted evasion strategies that target one or several steps in the antigen processing machinery to conceal the infection from the adaptive immune system, which is necessary to establish long-lived, latent infections. Both Epstein-Barr virus (EBV) as well as Kaposi sarcoma associated herpesvirus (KSHV) encode at least one protein that is present in latently infected cells, and yet no CTL response against epitopes derived from those proteins can be detected [130]. The EBV nuclear antigen 1 (EBNA1) and KSHV latency-associated nuclear antigen 1 (LANA1), respectively, may escape proteasomal degradation due to long repeats of two or three amino acids, even though presentation of EBNA1 peptides can be seen in some cell types [131-133]. Interestingly, EBNA1 and LANA1 share no sequence homology, which is one reason why the exact mechanism of inhibition is yet unknown [133].

The proteasome is the major but not exclusive source of peptides as inhibition experiments have shown [134]. Furthermore, even if cytosolic proteasomes interact with TAP, nuclear

proteasomes, however, do not have direct contact with TAP as it is excluded from the nuclear face of the membrane. Nuclear peptides, therefore, diffuse freely out into the cytosol where they encounter cytosolic peptidases like IFN-γ-inducible leucine amino peptidase or tripeptidyl peptidase II (TPPII) [135]. Since degradation is a very rapid process many peptides will be lost for antigen presentation. Thus, cytosolic peptidases contribute to the peptide pool by trimming longer peptides, albeit their activity is not essential for the generation of most antigenic peptides [136-138]. To shuttle peptides to the place of MHC-I loading, the transporter associated with peptide processing (TAP) located in the membrane of the ER translocates cytosolic peptides of 8 – 16 amino acids length into the ER lumen [139]. The TAP complex is a heterodimer consisting of two subunits, TAP1 and TAP2 that belong to the family of ABC transporters. Both are encoded in the MHC class II region and share highly homologues regions constituting the cytosolic ATP-binding and the N-terminal transmembrane domain [140, 141]. Both TAP subunits are polymorphic, which can influence substrate specificity as has been shown for rat TAP and a human splice variant TAP2iso [142-144]. Both preferably transport peptides with a hydrophobic and basic C-terminus [145]. Peptide binding and subsequent translocation leads to a conformational change and stimulation of ATP hydrolysis [146]. The expression of TAP1 and 2 genes is greatly increased upon IFN-γ stimulation as is the activity of the immunoproteasome and synthesis of MHC-I molecules upregulating antigen presentation in inflammation [147].

HSV-1 ICP47 inhibits peptide binding to TAP by competing for its binding site with an affinity 10-1000 times higher than for most peptides [148, 149]. The US6 gene product of human cytomegalovirus (HCMV) is also capable of inhibiting TAP transport by associating with its nucleotide-binding domain and thereby preventing ATP hydrolysis and conformational rearrangement, which is imperative for peptide translocation [150]. The UL49.5 protein of bovine herpesvirus-1 (BHV-1) targets TAP for proteasomal degradation whereas its EHV-1 and pseudorabies virus (PrV) homologues inhibit ATP binding [151, 152]. EBV BNFL2a protein was shown to block peptide and ATP binding to TAP [153].

Inside the ER, peptides longer than the preferred length for MHC-I loading can be trimmed by ER-resident amino peptidases (ERAP). Two peptidases in humans and one in mice have been described, which have distinct cleavage preferences for the generation of certain N-termini. IFN-γ inducible ERAP1 was the first to be identified and found to cleave all peptide bonds except those upstream of proline residues, thereby generating ligands with P at the putative anchor position 2 [154, 155]. ERAP1 acts in concert with ERAP2, which removes N-terminal

basic residues [156]. Murine ERAAP on the other hand trims all peptides except those with proline at the second position [157].
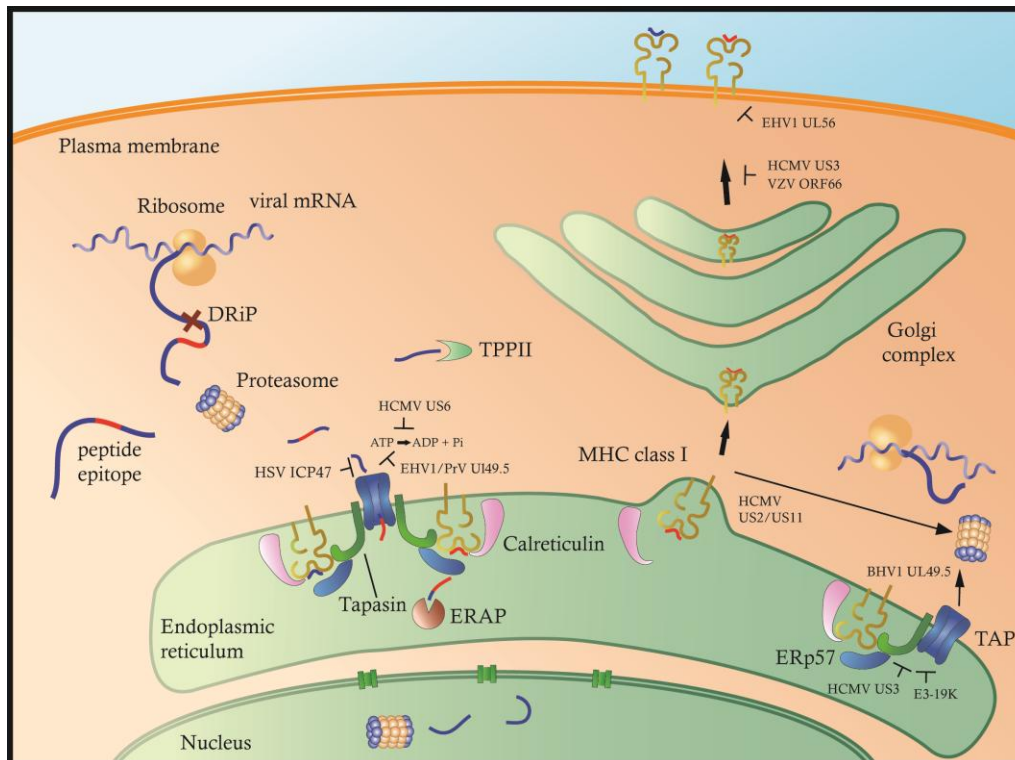
Figure 5: Antigen processing pathway for MHC class I molecules. For details see text. Briefly, defective ribosomal products are degraded by the proteasome, shuttled into the ER by TAP, bound to MHC-I with the help of the PLC and transported to the surface through the secretory pathway. Viral proteins interfere with this pathway at multiple sites.

The luminal domain of TAP interacts with the MHC-I chaperone tapasin, which with the help of two additional chaperons, calreticulin and ERp57, stabilizes the partially folded MHC-I-β2-microglobulin complex [158, 159]. This structure is commonly referred to as peptide loading complex (PLC). ERp57 is a thiol oxidoreductase ensuring correct formation of disulfide bonds in the MHC-I heavy chains and acts in concert with calreticulin [160]. The glycoprotein tapasin not only bridges MHC-I to TAP but it also stabilizes class I molecules in a peptide receptive state, enhances expression of TAP and prevents premature release of MHC-I [161-163]. It is therefore crucial for peptide loading, and tapasin-deficient cells fail to assemble the PLC, which abrogates antigen presentation for most HLA alleles [164]. That makes it a target for viral evasion strategies.

HCMV encodes the protein pUS3, which inhibits tapasin by associating with it whereas the adenoviral E3-19K binds to TAP preventing bridging with tapasin [166, 167].

The last step in the processing pathway is then loading of the peptide onto MHC-I molecules according to its binding preferences.

Some viruses target MHC-I molecules directly. US2 and US11 proteins of HCMV as well as the murine CMV E3 ligase mK3 physically bind to HLA-A2 molecules and redirect them to the ubiquitin-proteasome pathway or lysosomes for degradation [168, 169]. US3 and ORF66 products of HCMV and varicella zoster virus (VZV), respectively, both retain MHC-I in the Golgi or ER preventing it from reaching its destination on the surface [170, 171]. EHV-1 downregulates MHC-I by dynamin-dependent endocytosis through the action of the gene product of ORF1, an UL56 homologue [61]. Interestingly, MHC-I molecules of certain loci and alleles seem to be more affected than others, suggesting a selective downregulation to avoid complete depletion of MHC-I from the surface, as this would trigger NK cells to destroy the cell [1, 172].

The antigen processing pathway involves a number of steps with certain substrate specificities that aim at optimizing a putative MHC class I ligand pool. The importance of every step is highlighted by the abundance of different viral evasive strategies, which although targeting the same components show a remarkable diversity of actions.

4.3. Project introduction

4.3.1 Equine MHC class I

As outlined before, infections of equids with the equine herpesvirus type 1 (EHV-1) are still a major threat to animal health and a financial burden for the horse industry. Previous studies established a correlation between precursor CD8[+] cytotoxic T lymphocytes frequencies in peripheral blood and clinical protection from challenge infection [64, 65]. Classical vaccination strategies with attenuated modified-live virus or inactivated vaccines, however, regularly fail to induce protective CTL frequencies [71]. Since CTL responses are restricted to MHC class I molecules, detailed knowledge of MHC specificities, which determine potential CTL epitopes, is crucial in order to design a new generation of EHV-1 vaccines that would stimulate cellular immunity. For alleles of the MHC class I haplotypes A3 and A2, CTL epitopes could be mapped to the EHV-1 ICP4 protein, but distinct peptides remained to be defined [77, 78].

This project focuses on elucidating the binding motifs of four equine classical MHC-I alleles (Eqca-1*00101, -N*00101, -1*00201 and -16*00101) that are specific for the ELA haplotypes A3, A2 and A9, respectively (Aim 1), and identification of their EHV-1-derived CTL epitopes (Aim 2).

Table 1: ELA class I alleles used in this study

ELA haplotype	Locus	Systematic nom.	Alternative nom.	Allele
A3	1	Eqca-1*00101	ELA-A3.1	B2
A2	not assigned	Eqca-N*00101	ELA-A2.x	8-9
A2	16	Eqca-16*00101	ELA-A2.16	1-29
A9	1	Eqca-1*00201	ELA-A9.1	-

Throughout this thesis, the systematic nomenclature will be used, albeit all three are prominent in the literature. The alleles were chosen because they belong to common haplotypes in common horse breeds and because they had already been expressed in different mouse cell lines, RMA-S and P815.

RMA-S cells produce 'empty', temperature-sensitive MHC-I molecules, and this instability can be overcome by binding of externally supplied peptides [173]. RMA-S clones expressing the desired ELA alleles were used to screen a peptide library derived from the EHV-1 ICP4 protein for MHC-I binding peptides since previous studied had identified ICP4 as a source of

EHV-1-specific CTL epitopes [77, 78]. Alignment of the peptide sequences might suggest a putative binding motif and possible CTL epitope candidates. In a second, more refined approach, equine MHC-I molecules were purified from clonal P815 cells, their endogenous peptide ligand repertoires eluted and sequenced, and quantitative binding motifs identified by *in vitro* binding assays using positional scanning combinatorial libraries, which allow prediction of peptide binders from any equine pathogen. Additionally, the binding-peptide repertoires of Eqca-1*00101 and Eqca-N*00101 were determined.

Finally, binding motifs were used to predict EHV-1-derived peptide binders for Eqca-1*00101 and Eqca-N*00101. Analyzed in *in vitro* binding studies, high affinity binders were tested for T cell reactivity in peripheral blood of EHV-1 primed and naïve horses in *ex vivo* ELISpot assays (Aim 2).

The results for Eqca-1*00101 were published in September 2015.

4.3.2 T cell frequencies after *in vitro* stimulation

As previously established, EHV-1-specific T lymphocytes play a critical role in virological and clinical protection from EHV-1 infection [64, 65]. Preliminary data point to insufficient induction of virus-specific T cell frequencies by commercial vaccines. For many years, the gold standard for enumeration of CTL frequencies in peripheral blood has been the cytotoxic T cell assay with limiting dilution analysis (LDA). However, detection of virus-specific precursor CTLs relies on the ability of individual T cells to undergo clonal proliferation and differentiation, which might underestimate frequencies of responsive effector CTLs [174]. Furthermore, detection is restricted to cells with cytolytic capacity. Additionally, LDAs are very laborious and time-consuming to perform and strongly dependent on culture conditions [175]. Lately, new approaches to determine CTL frequencies have been established. Peptide-specific MHC tetramers allow quantification of T cells harboring a TCR capable of binding that particular peptide-MHC complex. However, they do not provide any information about the functional phenotype of these cells [176]. ELISpot assays are well suited to detect IFN-γ secreting CTLs after stimulation with their cognate peptide antigen [177, 178]. IFN-γ is generally considered to be a surrogate marker for T cell activity, which are the only source of that cytokine besides NK cells [179-181]. ELISpot analyses have been extensively used to characterize T cell epitopes for various MHC class I and II alleles for many pathogens in a few species [182-184]. Diagnostic ELISpot-based tools for infection with *Mycobacterium tuberculosis* or

Borrelia burgdorferi are commercially available [185, 186]. Both rely on stimulation of PBMCs with a pathogen-derived peptide antigen. However, peptide stimulation either in LDAs or ELISpot as well as peptide-tetramer approaches limit T cell detection to clones specific for only one antigen and ignore the variety of pathogen-derived antigens. Since CTL-mediated immune responses *in vivo* are likely to be directed against several epitopes restricted by multiple MHC class I alleles, development of robust detection systems that take all pathogen-specific CTL populations into account would be highly desirable as it would enable reproducible, multi-centered trials for determination of protective CTL frequencies against various infectious diseases. Furthermore, $CD4^+$ T lymphocytes also participate in cell-mediated immunity, not only by orchestrating the immune response but also by secreting IFN-γ. In the recent past, antiviral activity of the adaptive cellular immune system is not only attributed to direct cytotoxic effects, but also to noncytopathic actions of antiviral cytokines, e.g. IFN-γ and TNF-α, released by $CD8^+$ and $CD4^+$ T cells [67, 187]. ELISpot analyses can easily detect and enumerate cytokine-secreting cells, are easy to perform, only need a limited number of input cells and almost no specialized equipment.

In theory, re-stimulation of EHV-1-primed PBMCs with replication-competent virus should result in presentation of viral antigens on infected cells and professional APCs, and subsequent MHC-I- and -II-restricted activation of $CD8^+$ and $CD4^+$ T cells, respectively. Since the immune response against EHV-1 infections is associated with elevated INF-γ concentrations in peripheral blood coinciding with a decline in cell-associated viraemia [67, 82, 188], regardless of the source of the cytokine ($CD4^+$ or $CD8^+$ T lymphocytes), enumeration of IFN-γ-producing cells could provide a useful correlate of protection from viraemia and consequently abortion and EHM.

The aim of this project is to evaluate the potential of a modified ELISpot assay to quantify T cell responses against EHV-1 of primed and naïve horses.

5. Materials and Methods

5.1. Materials

5.1.1 Chemicals, consumables and equipment

5.1.1.1 Chemicals

Name	**Type/Cat.No.**	**Company**
Acetic acid	20103.295	VWR, Radnor
Agarose – Standard Roti® grade	[A160, 2500]	Applichem, Darmstadt
Albumin Fraction V	[A1391,0250]	Applichem, Darmstadt
Aminoethylcarbazole (AEC)	[A6926]	Sigma-Aldrich, St. Louis
Aprotinin	[A6106]	Sigma-Aldrich, St. Louis
β-mercaptoethanol	[28625]	Serva, Heidelberg
β2-microglobulin		Scripps Laboratories,
Chloramine T	[857319]	Sigma-Aldrich, St. Louis
cholamidopropyl dimethylammonio propanesulfonate (CHAPS)	[226947]	Sigma-Aldrich, St. Louis
Diethylamine	[471216]	Sigma-Aldrich, St. Louis
Dimethyl formamide	[UN2205]	Carl Roth, Karlsruhe
Dimethyl sulfoxide (DMSO)	[1.02952.2500]	Merck Millipore, Darmstadt
Ethanol, absolute	[34923]	Sigma-Aldrich, St. Louis
Ethylendiaminetetraacetic acid (EDTA)	[A2937,1000]	Applichem, Darmstadt

FACS rinse	[340346]	BD, San Jose
FACS clean	[340345]	BD, San Jose
FACS sheath fluid	[B51503]	Beckman Coulter, Krefeld
FACS FlowClean	[A64669]	Beckman Coulter, Krefeld
G-10 beads	[GE 17-0010-01]	GE Healthcare, Little Chalfont
Leupeptin	[L2884]	Sigma-Aldrich, St. Louis
Metabisulfite	[S9000]	Sigma-Aldrich, St. Louis
Methanol	[UN1230]	VWR, Radnor
Methyl cellulose	[M0262]	Sigma-Aldrich, St. Louis
Microscint	[6013621]	PerkinElmer, Waltham
NP40	[18896]	Sigma-Aldrich, St. Louis
Octylglucoside	[29836-26-8]	Anatrace, Maumee
Pepstatin A	[P5318]	Sigma-Aldrich, St. Louis
Phosphatase inhibitor cocktail II	[P5726]	Sigma-Aldrich, St. Louis
Phosphatase inhibitor cocktail III	[P0044]	Sigma-Aldrich, St. Louis
phenylmethylsulfonyl fluoride (PMSF)	[10837091001]	Sigma-Aldrich, St. Louis
Protein-A sepharose	[P3391]	Sigma-Aldrich, St. Louis
Sepharose	[4B200]	Sigma-Aldrich, St. Louis
Sodium acetate (NaAc)	[A4279,0100]	Applichem, Darmstadt
Sodium azide (NaAz)	[UN1687]	Applichem, Darmstadt
Sodium chloride (NaCl)	A3597,1000]	Applichem, Darmstadt
Sodium iodine 125 (NaI^{125})	[NEZ033002MC]	PerkinElmer, Waltham

Tris(hydroxymethyl)aminomethane [443866G]   VWR, Radnor
(Tris)

5.1.1.2 Consumables

Name	**Feature/Cat.No.**	**Company**
Acrodisc syringe filter	[PN4187]	Life Science Laboratories
Amicon Ultrafiltration unit	[UFC900308]	Merck Millipore, Darmstadt
Anti-static wipes	[PP9226]	Desco
Cell culture flasks	25 mL, 75 mL	Sarstedt, Nümbrecht
Cell culture plate, 10 cm	[83.1802]	Sarstedt, Nümbrecht
Conical test tubes 17x120		Sarstedt, Nümbrecht
Conical test tubes 30x115		Sarstedt, Nümbrecht
Costar seals	[3080]	PerkinElmer, Waltham
Cryotubes		Nunc, Kamstupvej
Econo-Column, 10 cm	[7371512]	Biorad, Munich
ELISpot PVDF plates	[MSIPS410]	Merck Millipore, Darmstadt
FACS tubes	[55.1579]	Sarstedt, Nümbrecht
Greiner 96-Well U-shape	[650201]	Greiner Bio-One
Greiner 96-Well white high binding	[655074]	Greiner Bio-One
Microtubes, low protein-binding	[727.706.600]	Sarstedt, Nümbrecht
Pipettes	5, 10, 25 ml	Sarstedt, Nümbrecht
Pipettetips	P1000, 200, 100, 10	VWR, West Chester
Roller bottles	Cellmaster	Greiner Bio-One

Spuncolumns	[SP-25]	ABT, USA
SuperFrost® Plus	[J1800AMNZ]	Menzel Glaser, Braunschweig
TopSeal	[6050195]	PerkinElmer, Waltham
U-bottom 96-well plates	[92697]	TPP, Trasadingen

5.1.1.3 Equipment

Name	**Feature/Cat.No.**	**Company**
Casy cell counter	Casy TT 150	Roche, Berlin
Cell incubators	Excella ECO-1	New Brunswick Scientific
Labofuge 400R	Rotor 8177	Heraeus, Hanau
Centrifuge 5424R	Rotor 5424	Eppendorf, Hamburg
CytoFlex	flow cytometer	BeckmanCoulter, Krefeld
Bioreader-6000 FZβ	ELISPOT	Biosys
FACSCalibur	flow cytometer	BD, San Jose
Freezer -20 °C		Liebherr, Bulle
Freezer -80 °C		GFL, Burgwedel
Gel electrophoresis chamber		VWR International, West Chester
Ice machine	AF100	Scotsman, Vernon Hills
INTEGRA Pipetboy		IBS Integrated Bioscience, Fernwald
Microscope AE20	AE20	Motic, Wetzlar
Nanodrop 1000		Peqlab, Erlangen
Newbauer counting chamber		Assitant, Sondheim/Rhön

Nitrogen tank	ARPEGE70	Air liquide, Düsseldorf
Oven		Mennert, Schwabach
pH-meter	RHBKT/C WTW	Inolab, Weilheim
Refrigerator		Erwin Bonn, Duisburg
Rollerbottle incubator	Incudrive D-1	Schuettbiotec, Göttingen
Sterile laminar flow chambers		Bleymehl, Inden
Thermocycler	Professional Trio	Biometra, Jena
Thermomixer	comfort	Eppendorf, Hamburg
TopCount microscintillator		Packard Instrument Co.,
Ultracentrifuge	L7-65	Beckman, Krefeld
UV Spectrophotometer	SmartSpec Plus	Biorad, Munich
UV transilluminator	Bio-Vision-3026	Peqlab, Erlangen
Vortex	Genie 2™	Bender&Hobein AG, Zürich
Waterbath	TW2	Julabo, Seebach

5.1.1.4 Software

FlowJo	V.7.6.5	TreeStar, Ashland
Graphpad Prism 5	Version 5	Graphpad Software inc, La Jolla
ND1000	V.3.0.7	Peqlab, Erlangen
Vector NTI 9	Version 9	Invitrogen Life Technologies, Grand Island

5.1.2 Enzymes and markers

Name	**Cat.No.**	**Company**
Generuler TM 1kb Plus ladder	[Cat. No. SM0311]	Fermentas, Mannheim
Taq DNA-Polymerase	[Cat. No.01-1020]	PeqLab, Erlangen

5.1.3 Antibodies

Name	**Cat.No**	**Company**
Mouse MKD6 IgG	[523-56-3-4]	
Mouse anti-bovine IFNγ IgG	[3115-3-1000]	Mabtech, Nacka Strand
Mouse anti-bovine IFNγ-biotin IgG	[3115-6-500]	Mabtech, Nacka Strand
Mouse anti-equine MHC-I IgG (CZ3)		
Alexa goat anti-mouse IgG (H+L) 488	[A11001]	Thermo Fisher, Waltham
Alexa goat anti-mouse IgG (H+L) 647	[A-21236]	Thermo Fisher, Waltham

5.1.4 Cells

Name	**Features**	**Reference**
RMA-S	*Mus musculus* lymphocytoma TAP2 deficiency	[189]
P815	*Mus musculus* mastocytoma cell line DBA/2 strain	ATCC TIB-64

| RK13 | Rabbit epithelia kidney cell line, | ATCC CCL-37 |
| Equine PBMC | mononuclear cells | primary cells |

5.1.5 Virus

Name	**Features**	**Reference**
EHV-1 strain RacL11	ΔORF1/2, Δgp2, ΔIR6, GFP$^+$	[190]

5.1.6 Kits for molecular biology

Name	**Cat.No.**	**Company**
RTP® DNA/RNA Virus Mini Kit	[10400100300]	Statec
GF-1 Nucleic acid extraction Kit	[GF-GC-100]	Vivantis
Vectastain Kit	[PK-6100]	Biozol, Eching

5.1.7 Buffers, media and antibiotics

5.1.7.1 Buffers

1x Phosphate saline buffer	**1x Tris-Acetate-EDTA buffer**	**1% Agarose Gel**
2 mM KH_2PO_4	40 mM Tris	100 mM agarose
10 mM Na_2HPO_4	1 mM Na_2EDTA x 2 H2O	1x TAE buffer
137 mM NaCl	20 mM Acetic acid 99 %,	1.5 µL ethidium bromide
2.7 mM KCl, pH 7.3	pH 8.0	

1x PBST

1x PBS

0.05 % Tween20

1x PBSN

1x PBS

0.05 % NP40

Binding buffer (BB)

1.5 M Glycin

3 M NaCl, pH 9

Elution buffer (EB)

0,2 M Glycin, pH 2,5

Neutralization buffer

1 M Tris/HCl, pH 9

Lysis buffer (MHC)

20 mM Tris

150 mM NaCl

1 % NP40

200 mM PMSF

Wash buffer

10 mM Tris/HCl, pH 8

1 % NP40

Elution buffer

150 mM NaCl, pH 11.5

50 mM diethyl amine

0.02 % sodium azide

1 % octylglucoside

Tris/HCl buffer pH 7.4

1 M Tris/HCl, pH 7.4

HEPES buffer pH 7.4

1 M HEPES/HCl, pH 7.4

Tris saline buffer pH 8.0

100 mM Tris/HCL, pH8

Lysis buffer (peptide elution)

20 mM Tris/HCl, pH 8

150 mM NaCl

10 mg/mL CHAPS

High salt buffer

20 mM Tris/HCl, pH 8

1 M NaCl

Low salt buffer

20 mM Tris/HCl, pH 8

150 mM NaCl

No salt buffer

20 mM Tris/HCl, pH 8

AEC buffer

3.5 mM acetic acid

8.5 mM sodium acetate

5.1.7.2 Cell culture supplements

RPMI 1640 (bicarbonate buffered)	[P04-18500]	PAN, Aidenbach
RPMI 1640 (unbuffered)		PAN, Aidenbach
Minimum essential medium	[P04-09500]	PAN, Aidenbach
Fetal calf serum (FCS)	[P30-3306]	PAN, Aidenbach
L-alanyl-L-glutamine	[K 0302]	Biochrom AG, Berlin
Non-essential amino acids (NEAA)	[K 0293]	Biochrom AG, Berlin
Sodium pyruvate	[L 0473]	Biochrom AG, Berlin
Ascorbic acid	[3525]	Carl Roth, Karlsruhe
Bovine insulin	[I1882]	Sigma-Aldrich, St. Louis

5.1.7.3 Media

RMA-S Medium	**P815 Medium**	**PBMC Medium**
RPMI 1640	RPMI 1640 (unbuffered)	RPMI 1640
10 % FCS	25 mM HEPES, pH 7.4	10 % FCS
2 mM glutamine	10 % FCS	2 mM glutamine
1x NEAA	2 mM glutamine	100 U/mL penicillin
100 μM sodium pyruvate	0.5 μg/mL ascorbic acid	65 ng/mL streptomycin
10 μM β-mercaptoethanol	4 μg/mL bovine insulin	
100 U/mL penicillin	10 μM β-mercaptoethanol	

100 µg/mL streptomycin   100 U/mL penicillin

                                      100 µg/mL streptomycin

RK13 MEM

MEM

10 % FCS

2 mM glutamine

100 U/mL penicillin

100 µg/mL streptomycin

5.1.7.4 Antibiotics

Name	Working concentration	Company
Penicillin	100 U/mL	Carl Roth, Karlsruhe
Streptomycin	100 µg/mL	Applichem, Darmstadt

5.2 Methods

5.2.1 Cell culture methods

5.2.1.1 Clones of RMA-S and P815 cell lines stably expressing various ELA class I loci

RMA-S cells are derived from Rauscher leukemia virus-induced mouse T-cell lymphomas of C57BL/6 origin. The resulting cell line (RBL-5) was modified such that the TAP2 gene was removed, which led to dysfunctional peptide loading of MHC-I molecules [191].

P815 is a mouse mastocytoma-derived cell line of DBA/2 origin that has been used extensively to characterize human and murine MHC class I alleles [192, 193]. The cells can easily be transfected and are routinely used as targets in cytotoxicity assays [194].

RMA-S and P815 cell clones stably expressing equine MHC-I loci Eqca-1*00101, Eqca-N*00101, Eqca-16*00101 and Eqca-1*00201 were generated by random integration and following clonal selection by limited dilution procedures. Integration and expression of the respective genes was confirmed by sequencing, qPCR and immunoblotting.

This work was done by Rebecca M Harman in the laboratory of Prof. Dr. Douglas Antczak at the Baker Institute at Cornell University, Ithaca, USA.

5.2.1.2 Culturing of RMA-S and P815 cells

RMA-S clones were cultured in standard RPMI 1640 medium supplemented with 2 mM glutamine, 10 % FCS, 10 µM β-mercaptoethanol and 1 % penicillin/streptomycin (P/S) in a 37 °C incubator with 5 % CO_2 atmosphere. P815 clones were grown in RPMI 1640 medium buffered with 25 mM HEPES buffer, pH 7.4, and supplemented with 2 mM glutamine, 10 % FCS, 4 µg/mL bovine Insulin, 0.5 µg/mL ascorbic acid, 10 µM β-mercaptoethanol and 1% P/S at 37 °C. Cell densities were kept between 1×10^5 and 1×10^6 cells/mL. P815 cells were grown to large volumes by use of a roller bottle incubator. 400 mL of cell suspension were grown in two standard T175 tissue culture flasks, united and transferred to a 2 L roller bottle. The volume was expanded to reach 1 L before another roller bottle was introduced.

5.2.1.3 Culturing of antibody-producing hybridoma cells

Hybridoma cells derived from Balb/C mice producing the equine MHC-I-specific antibody CZ3 were cultured in RPMI 1640 medium buffered with 20 mM HEPES, pH 7.4, supplemented with 20 mM glutamine, 10 % FCS and 1 % P/S at 37 °C. Cell densities were kept between 1×10^5 and 1×10^6 cells/mL and the volume was expanded to roller bottles as described before. Cells were incubated to reach a density of 1×10^7 cells/mL in a final volume of 1 L.

5.2.1.4 Culturing of RK13 cells

RK13 cells were cultured in standard MEM medium supplemented with 10 % FCS, 2 mM glutamine and 1 % P/S at 37 °C in 5 % CO_2 atmosphere. Cells were grown to confluency, medium was removed, and cells were washed with PBS. Subsequently they were incubated with 0.25 % trypsin for several minutes at 37 °C to detach cells from the culture well surface.

For virus propagation, 5 X 10^6 cells were seeded in 10 cm tissue culture plates and incubated for 24 h to gain confluency.

5.2.1.5 Isolation of peripheral blood mononuclear cells (PBMC) from horses

Horses sampled for this study were maintained in AAALAC approved facilities at the Equine Genetics Center at Cornell University and the Veterinary Medical Center at Michigan State University. Animal care was performed in accordance with the guidelines set forth by the Institutional Animal Care and Use Committees. The horses at Cornell had been vaccinated at regular intervals with killed and twice in the months before sampling with modified live vaccines against EHV-1 and equine influenza virus. The horses at Michigan State University had been infected experimentally with virulent EHV-1 on at least three occasions. Peripheral blood mononuclear cells (PBMC) were isolated from heparinized samples of venous jugular blood using methods previously described [195]. 500 mL blood or leukocyte-enriched plasma was pelleted by centrifugation at 350 x g for 10 min at RT in 50 mL conical tubes. Supernatant was removed and the cell pellet was resuspended in 35 mL of RPMI 1640 medium supplemented with 10 % FCS. 15 mL of Ficoll (1.077 g/mL) were overlaid with the suspension and tubes were centrifuged at 350 x g for 15 min at RT. The interface was collected and washed three times with RPMI medium. Cells were frozen in cold freezing media containing FCS with 10 % DMSO at a density of 30 X 10^6/mL in a freezing container for 24 h and subsequently maintained in liquid nitrogen.

5.2.1.6 Recovery of cryopreserved PBMC

The required number of cryotubes containing frozen PBMCs were recovered from liquid nitrogen and thawed in a 37 °C water bath for 45 sec. Cells were transferred to 15 mL conical tubes with 10 mL of preheated RPMI 1640 medium with 10 % FCS and 50 U/mL Benzonase to avoid cell clumping. Cells were centrifuged for 10 min at 350 x g at RT, washed once with medium and resuspended in 10 mL medium without Benzonase. Cell numbers and viabilities were determined with a Casy cell counter and the volume was adjusted according to the desired cell concentrations. Excess cells were refrozen as previously described.

5.2.1.7 Sample preparation for sequencing

To confirm the identity of all P815 and RMA-S clones used in this study, the respective MHC-I alleles were sequenced. Roughly 1×10^4 cells were collected and genomic DNA was isolated with the RTP® DNA/RNA Virus Mini Kit. MHC coding sequences were amplified via PCR with primers CTCACTATAGGGAGACCCAA (primer 1) and TAGAAGGCA-CAGTCGAGG (primer 2) synthesized by LGC Genomics, Berlin, following manufactures instructions for the NEB Taq polymerase. Following amplification conditions were used: initial denaturation at 95 °C for 2 min followed by 30 cycles of denaturation at 95 °C for 20 s, primer annealing at 52 °C for 20 s and elongation at 65 °C with a final extension step for 5 min at 65 °C. PCR products were analyzed on a 1 % agarose gel stained with ethidium bromide and subsequently purified with the GF-1 Nucleic acid extraction Kit. 400 ng of the PCR product and 1.6 µg/mL primer 1 were sent for sequencing to LGC Genomics (Berlin). Sequencing results were analyzed with Vector NTI 9 software.

5.2.1.8 Virus propagation

Confluent RK13 cells in a 10 cm tissue culture plate were infected with EHV-1 strain RacL11, and the infection was allowed to spread to completion. Subsequently, cells and supernatant were harvested, subjected to two freeze-thaw cycles and centrifuged for 10 min at 300 x g. The supernatant was collected and titrated on confluent RK13 cells in six serial dilutions with an overlay of semi-fluid methyl cellulose. After 48 hours, plaque numbers were assessed using a standard light microscope and titers calculated accordingly.

5.2.2 Biochemical Methods

5.2.2.1 Purification of CZ3 antibody from cell culture supernatant

CZ3 antibody-producing hybridoma cells were grown as described before. 1 L of cell suspension was centrifuged at 300 x g for 15 min. 2 g of protein A-conjugated sepharose beads were hydrated in 30 mL PBS, poured into a 10 cm Econo-Column and equilibrated with 15 mL of binding buffer (BB) containing 1.5 M glycine and 3 M sodium chloride (NaCl), pH 9. Cell culture supernatant was diluted with 1 L BB, filtrated through a 0.2 µm pore size filter and 100 mL were applied to the column, which was subsequently washed with 20 mL of BB. Bound antibody was eluted with 20 mL of elution buffer (EB) containing 0.2 M Glycine, pH

2.5, and fractions of 2 mL were individually collected into tubes containing 60 μL of 1 M Tris/HCl, pH 9, to neutralize pH. The column was washed with 10 mL of EB and equilibrated with 20 mL of BB. This procedure was repeated until all cell culture supernatant had been processed. Fractions with OD_{280} of 1 or higher as determined by UV spectroscopy were pooled and concentrated in an Ultra-15 3K concentrator by centrifugation at 5000 rpm for 2 h.

5.2.2.2 Affinity purification of ELA class I molecules from cell cultures

P815 clones expressing the desired ELA class I allele [97] were used to purify MHC class I molecules by immunoaffinity columns as described previously [196]. MHC class I expression was confirmed by flow-cytometry analysis. 1.5×10^{13} cells were harvested by centrifugation for 10 min at 400 x g, washed once with PBS and stored at -80 °C until usage. When needed, cell pellet (10-12 mL) was resuspended in 120 mL lysis buffer containing 20 mM Tris, 150 mM NaCl, 1% NP40 and 200 mM phenylmethylsulfonyl fluoride (PMSF) by stirring for 30 min at 4 °C. The lysate was ultracentrifuged for 20 min at 15000 rpm at 4 °C and subsequently passed through a 0.8 μm filter. Four 10 cm purification Econo-Columns in total were prepared. For two, 5 g of sepharose beads were hydrated in 20 mL of washing buffer containing 10 mM Tris, pH 8, and 1 % NP40 and equally distributed to both columns. For a third column, the procedure was repeated but with protein A-conjugated sepharose beads. All three columns were stripped with 30 mL of elution buffer containing 150 mM NaCl, pH 11.5, 50 mM diethylamine, 0.02 % azide (w/v) and 1 % octyl glucoside (w/v), neutralized with 200 mL of washing buffer and equilibrated with 30 ml of lysis buffer. A fourth column was prepared by also hydrating protein A-conjugated sepharose beads in borate buffer, pH 8.2, containing 0.6 % boric acid, 0.95 % borax and 0.2 % NaCl. Beads were transferred to the column, excess buffer was drained, and beads were washed twice with 10 mL of borate buffer. 40 mL of a borate buffer solution containing 0.75 mg/mL CZ3 antibody was prepared and incubated with the beads on a rotator for 1 h at 4° C. The flow-through was collected and the OD_{280} was measured before and after coupling to determine the amount of antibody that had bound to the column. The column was washed with 50 mL of borate buffer, 20 mL of 200 mM triethanolamine and incubated with 40 mL of 20 mM dimethyl pimelimidate in triethanolamine on a rotator for 45 min to cross-link the antibody to protein A beads. Subsequently, beads were washed twice with 10 mL of 20 mM ethanolamine, 200 mL of borate buffer and 100 mL of PBS with 0.05 % NP40. Right before use, the column was washed with 30 mL of elution buffer, 2 M glycine solution pH 2.5 until a pH of 7-8 was reached, and finally with washing buf-

fer. The first three columns were assembled in sequence, starting with both plain sepharose beads columns and followed by the protein A-beads column. The lysate was precleared by repeated runs over the column assembly with stripping steps using 30 mL of elution buffer, 20 mL of washing buffer and 30 mL of lysis buffer in between. Subsequently, the lysate was applied twice to the specific antibody column. After each run, the column was washed with 200 mL of washing buffer and 50 mL of PBS with 0.4 % octylglycoside. Bound MHC-I was finally eluted in 50 mL of elution buffer and collected in a 50 mL conical tube containing 3 mL of 2 M Tris, pH 6.8. The volume was reduced to 800 µL and possibly co-purified murine β2-microglobulin was size-excluded by several centrifugation steps in Centriprep Ultracel YM-30 tubes at 2000 rpm. Protein purity, concentration, and the effectiveness of depletion steps were monitored by SDS-PAGE and BCA assay. Purification of Eqca-1*00101, Eqca-N*00101 and Eqca-16*00101 were done by Carrie Moore, Erin Moore, and Mikaela Lindvall.

5.2.2.3 Acidic elution of endogenous MHC class I ligands

P815 cells were grown to a volume of 10 L with a density of 1×10^6 cells/mL. MHC class I molecules were purified from the cell surface as follows. 1×10^{10} cells were pelleted by centrifugation at 300 x g for 10 min and washed once with PBS. The pellet was lysed in 35 mL lysis buffer containing 20 mM Tris/HCl, pH 8.0, 150 mM NaCl, 10 mg/mL CHAPS, the protease inhibitors aprotinin (5 µg/mL), leupeptin (1 µg/mL), pepstatin A (10 µg/mL) and PMSF (1 mM) as well as phosphatase inhibitor cocktail I and II by stirring at 4 °C for at least one hour. The lysate was ultracentrifuged for 1 h at 10,000 x g and 4° C to remove cellular debris and subsequently filtered through 0.2 µm pore size filters to remove remaining lipids. In parallel, affinity columns were prepared. 1 g of protein A-sepharose beads were hydrated in 15 mL TBS and rotated for 1 h at 4° C. The suspension was poured into a 10 cm Econo-Column. This was done in duplicates. The equine MHC class I-specific antibody CZ3 and the murine MHC-II-specific antibody NKD6 were each diluted in TBS to a concentration of 0.5 mg/mL. The columns were equilibrated with 20 mL of cold lysis buffer (without inhibitors) and 30 mL of antibody solution were applied to the respective columns and the flow through (FT) was collected. The OD_{280} of the antibody solution and the FT was determined by UV-spectroscopy to control for antibody binding to the columns. Filtered lysate was passed over the anti-murine MHC-II antibody-coupled protein A column to control for nonspecific binding on top of the CZ3-coated column. The columns were washed in parallel with 8 mL of lysis buffer (without inhibitors), 80 mL of low salt buffer (20 mM Tris/HCL, pH 8, with

150 mM NaCl), 80 mL of high salt buffer (20 mM Tris/HCl, pH 8.0, with 1 M NaCl) and80 mL of no salt buffer (20 mM Tris/HCl, pH 8.0). Antibody-MHC-peptide complex were eluted in 15 mL conical tubes with 12 mL of 0.2 N acetic acid. 1.06 mL of glacial acetic acid was added to reach a final concentration of 10 % acetic acid. Antibody and MHC molecules were size-excluded by centrifugation through an Ultrafree-CL filter and the volume of the peptide-containing FT was reduced to 250 µL by vacuum centrifugation. Peptide elution for Eqca-1*00101 was done by Carrie Moore. Peptides were sent for HPLC purification with subsequent sequencing by mass spectrometry to University of Virginia, department of chemistry. HPLC purification and sequencing were done by Amanda Wriston and Paisley Myers in Donald Hunt's laboratory.

5.2.2.4 Peptide synthesis and handling

Peptide libraries were designed by Dr. John Sidney. Peptides used for screening studies in *in vitro* binding assays were synthesized as crude, for radiolabeled ligands as purified (> 95 %) material on a 1-mg scale by A&A (SyntheticBiomolecules; San Diego). Peptide 15mers used in RMA-S MHC stabilization assays were synthesized by ChinaPeptides Co., Ltd. as crude material on a 1-mg scale.

Peptide stocks (1 mg) used for screening were dissolved in 25 µL DMSO to yield a sock concentration of 40 mg/mL. For binding assays, stocks were diluted 1:1 with DMSO.

Peptides used for RMA-S MHC stabilization were dissolved in 660 µL DMSO to yield a concentration of 15 mg/mL.

5.2.2.5 Radiolabeling of peptides with iodine[125]

Peptides were labeled following the microlabeling method [196]. Spun columns containing 0.8 mL of hydrated G10 beads and collection vessels containing 10 µL of 100 % EtOH, 5 µL of 10 % NaN_3 and 25 µL of 0.82 % NP-40 in PBS were prepared. 1.5 µL of NaI[125] were diluted with 6 µL of 0.05% NP40 in PBS. 100 ng of tyrosinated peptides were labeled with 5 µL of diluted I[125] in the presence of 30 µg/mL chloramine T as an oxidizing catalyst. In order to prevent the peptide from being destroyed by gamma-radiation, reaction was quenched by the addition of 20 µg/mL sodium metabisulfate after 1 min. 25 µL of PBS with 0.05 % tween20 were added and peptides were transferred to spun columns and centrifuged four times for 2 min each at 82 x g with 50, 150 and 75 µl of PBS added after the first, second and

third run, respectively, to separate labeled from non-labeled peptides into collection vessels by size-exclusion gel chromatography. To determine the efficiency of the labeling procedure, 10 µL of peptides diluted 1:10 with PBS were mixed with 100 µL of semi-fluidmicroscint in a sealed U-bottom assay plate by vigorously vortexing for 2 min. Samples were prepared in triplicates and specific activity was determined with a TopCount microscintillation counter (Packard Instrument Co., Meriden, CT).

5.2.2.6 Measurement of MHC-I-peptide interactions by monoclonal antibody capture

The *in vitro* binding assay using purified MHC-I molecules and iodine[125]-labeled high affinity (standard) peptides provides a useful tool for the quantification of binding affinities of a (inhibitory) peptide to MHC-I molecules. Inhibitory peptides at a given concentration compete with the standard peptide for MHC-I binding, which is detected by a decreased radioactive output (Figure 6). The concentration of inhibitory peptide yielding 50% competition is referred to as inhibitory concentration 50 (IC_{50}), and within the conditions utilized here with [label] < [MHC] and $IC_{50} \geq$ [MHC], reflects a good approximation of the real dissociation constant [196]. A peptide library of inhibitory peptides was tested for competition of binding to purified MHC-I molecules with an I[125]-labeled standard peptide. MHC class I molecules were produced in cell culture and purified as described above. To establish a binding assay MHC-I molecules in a range of 0.3 nM – 1 µM were incubated with 1 µM human β2-microglobulin (β2m) (Scripps laboratories) and various labeled standard peptide candidates in an amount that corresponded to 8500 cpm in the presence of protease inhibitors (2 mg/mL EDTA, 120 µg/mL pepstatin A, 620 µg/mL phenanthroline, 480 µg/mL PMSF and 120 µg/mL tosyl lyschloromethyl ketone hydrochloride) for 48 h. Flat-bottom 96-well plates were coated with 30 µg/mL CZ3 anti-equine MHC-I antibody in 100 mMTris buffer for 24 h and then blocked with 0.3 % tween20 in PBS for 2 h. They were washed twice with 0.05 % tween20 in PBS (PBST) and peptide-MHC class I complexes were transferred to the antibody-coated plates for capturing. After 3 h, plates were washed twice with PBST and overlaid with 150 µL of semi-fluid microscint to distribute radiation equally inside the well. MHC concentrations and peptide candidates producing high counts with a low signal-to-noise ratio were chosen for further experiments.

For binding assays, MHC-I standard peptide mix was prepared as described. In U-bottom 96-well assay plates, MHC-I mix was subsequently incubated with an inhibitory peptide library titrated in a range of 33 µM to 0.33 nM for 48 h to allow the MHC-I-β2m-peptide complex to

form. The unlabeled version of radiolabeled probe was included in each analysis as a positive control. All assays were done in triplicates. Radioactivity was measured with a TopCount microscintillation counter
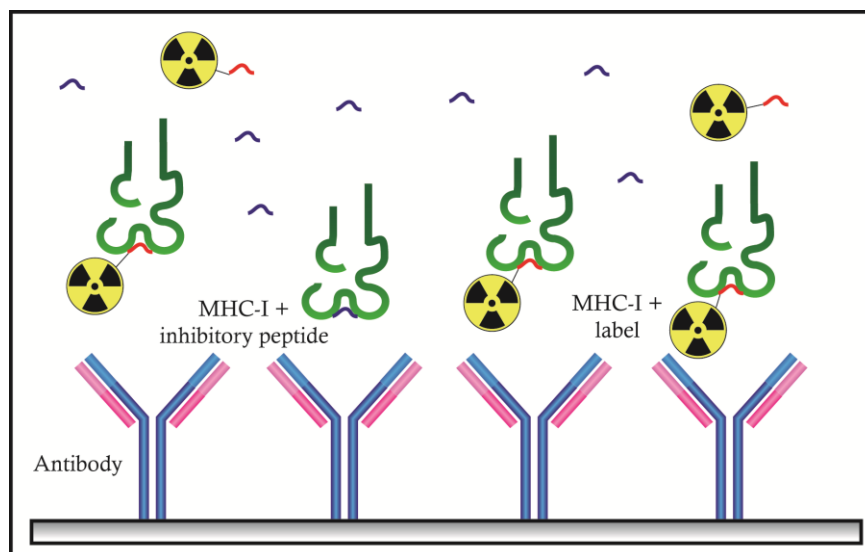
Figure 6: In vitro binding assay. Immobilized antibody captures MHC-I molecules in complex with either radioactively labeled (standard peptide) or unlabeled peptides (inhibitory peptides). Inhibitory peptide concentration yielding 50 % inhibition is referred to as inhibitory concentration 50 (IC_{50}).

5.2.2.7 Positional scanning combinatorial libraries (PSCL)

A PSCL is a peptide library of 9mers consisting of 180 pools, each one with one fixed amino acid residue at a single position and each of the 20 naturally occurring residues represented at each position along the 9mer backbone (Figure 7). This allows evaluating the individual contribution of each amino acid in each position to overall binding. A library completely randomized at each position was included. PSCL were tested in a competition assay as described above and results were analyzed as follows. IC_{50} nM values for each library were standardized as a ratio to the geometric mean IC_{50} nM value of the entire set of 180 libraries thereby evaluating the contribution of each position in the peptide to binding. Individual mixtures were normalized at each position so that the value of a specific residue associated with optimal binding corresponded to 1, yielding an average relative binding (ARB) affinity for each position. The ratio of individual ARBs to the ARB of the entire library is denoted as specificity factor (SF) and describes the influence of all 20 amino acids at a specified position to binding in respect to the average affinity of the library. Highly specific positions will be associated with high SF values and primary anchor positions were defined as those with an SF > 2.4, identifying positions where the majority of residues is associated with decreased binding. Positions with standard deviation (SD) > 2.4 were defined as secondary anchor positions.
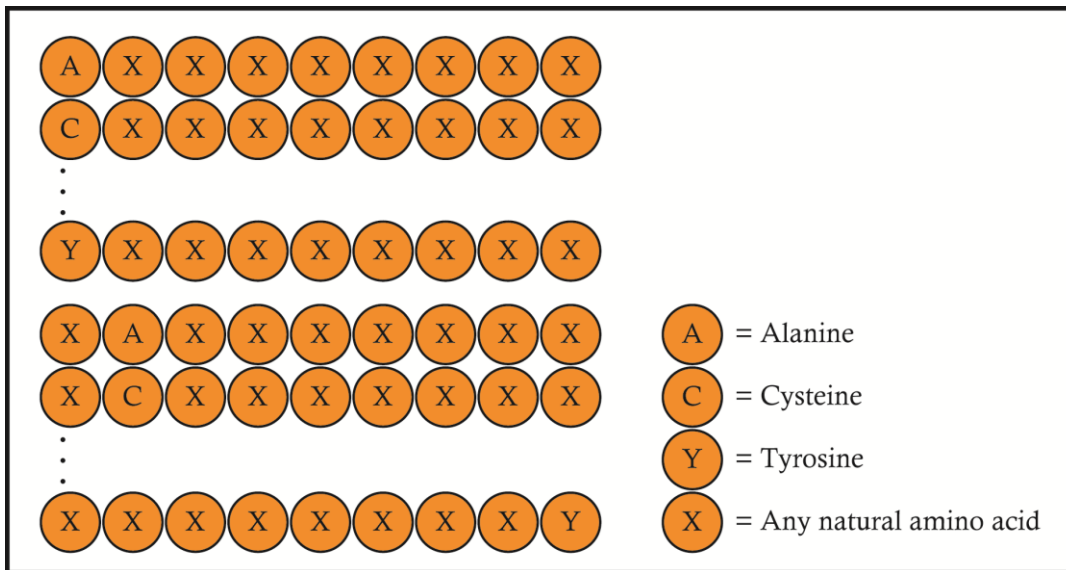
Figure 7: Positional scanning combinatorial libraries (PSCL). In each 9mer of 180 peptide libraries, one amino acid in one position remains fixed while any other position is represented by any of the 20 natural amino acids

5.2.3 Cell assays

5.2.3.1 RMA-S MHC stabilization assay

RMA-S cells exhibit dysfunctional peptide loading due to a deficiency in the TAP2 gene. Initially derived from the Rauscher virus-induced H2b lymphoma RBL-5 by exposure to the mutagen ethyl methane sulfonate and subsequent selection with antisera against MHC-I molecules and complement factors, RMA-S cells synthesize both class I α and β chains but they do not associate and remain mostly intracellular [189, 197]. However, at low temperature around 26 °C, increased expression of MHC-I molecules at the cell surface can be observed, and they seem to be folded correctly and associated with β2m [173]. Nevertheless, despite increased MHC-I density at the surface, influenza-infected RMA-S cells are not recognized by specific CTL clones as opposed to the control cell line RMA, which is thought to be the result of 'empty' MHC molecules being transported to the surface at 26 °C [173]. This thermolability also can be overcome by binding of externally supplied peptides, leading to accumulation of MHC-I-peptide complexes at the cell surface [198]. Stabilized MHC-I molecules can be detected with specific antibodies in flow cytometric analyses (Figure 8) [191].

One day prior to feeding experiments, 1.5×10^6 RMA-S cells expressing the desired ELA class I allele were seeded in 50 mL conical tubes in 15 mL medium. Tubes were incubated with an open lid at 37° C and 5 % CO_2 for two hours to allow for gas exchange before the lid

was closed to retain the atmosphere, and the tubes were moved to 24° C for 20 h. The next day, cells were pelleted by centrifugation for 2 min at 300 x g, pooled and resuspended in 2.25 mL of medium to reach a density of 2 X 10^6 cells/mL. 100 µL of cell suspension were seeded in a well of an U-bottom 96-well plate.
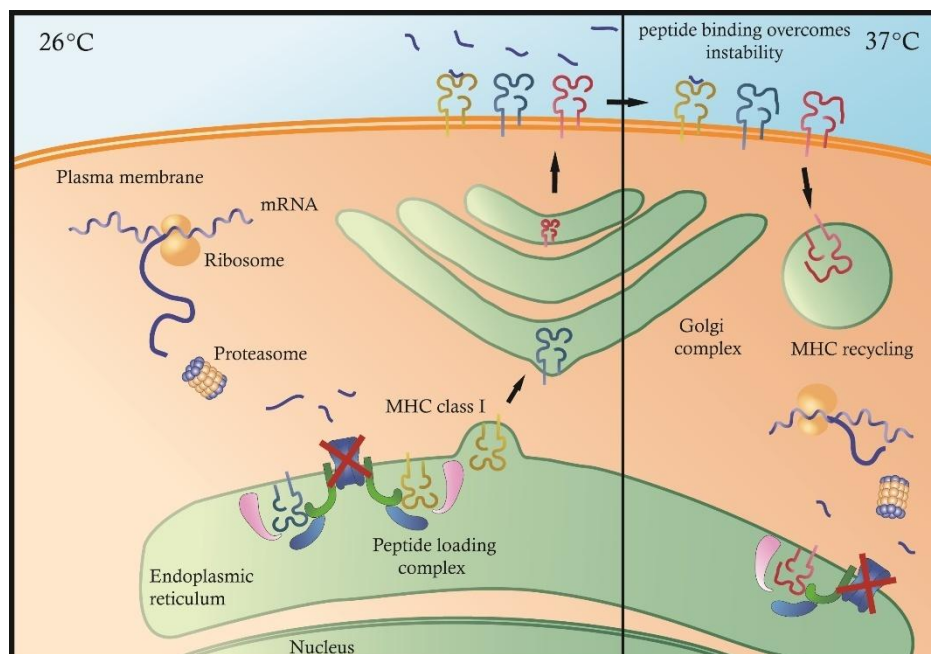
Figure 8: RMA-S MHC stabilization assay. TAP-deficiency depletes peptide pool in the ER. However, decreased temperature (left side) stabilizes MHC-I molecules in an open conformation. Binding of extracellular peptides to MHC-I molecules overcomes temperature-sensitivity and retains class I molecules on the cell surface (right side).

Peptides were thawed from a 10 mM (15 mg/mL) DMSO stock, diluted 1:5 with sterile PBS and 20 µM (30 µg/mL) were added to the cells, which were incubated for 30 min at RT and subsequently at 37° C for five hours. Experiments were done in duplicates and repeated four times. Of those, three times were carried out in a blind fashion.

5.2.3.2 Flow cytometry analysis

For RMA-S MHC stabilization assays, 2 X 10^5 cells were seeded in an U-bottom 96-well plate in 100 µL, centrifuged at 1200 rpm a 4° C for 2 min. They were washed three times with cold PBS and incubated for 20 min with the anti-equine MHC-I antibody CZ3 in a 1:100 dilution in PBS + 1 % FCS. After washing again three times, Alexa488-conjugated anti-mouse IgG (A11001, Life technologies) (1:1000) was added and incubated for another 20 min at 4° C in the dark. After a final washing step, cells were resuspended in 300 µL PBS + 1 % FCS and transferred to FACS tubes.

For detection of MHC-I on the surface and GFP as an infection marker in equine PBMC, 2×10^5 PBMC were infected with EHV-1 strain RacL11 at a multiplicity of infection (MOI) of 1. After 3 h, cells were washed three times with RPMI medium and resuspended in 100 μL. After incubation for 24 h at 37 °C, cells were washed three times with cold PBS, resuspended in 100 μL PBS + 0.5 % BSA and incubated with MHC-I-specific antibody CZ3 or an isotype control antibody in a 1:100 dilution. After 20 min incubation, cells were washed twice with cold PBS and in the presence of 0.5 % BSA incubated with Alexa647-conjugated anti-mouse IgG for 20 min at 4 °C in the dark. Finally, cells were washed again three times in cold PBS, resuspended in 300 μL PBS + 0.5 % BSA and transferred to FACS tubes.

5.2.3.3 *Ex Vivo* primary ELISpot assay

Ex vivo ELISpot assays are used to detect cytokine secretion by T cells upon stimulation. Target cells present in PBMCs are pulsed with peptides directed against an MHC-I molecule with known binding specificities to be presented to CD8$^+$ cytotoxic T lymphocytes. Upon TCR recognition of the MHC-peptide complex, the CTL clone is activated to secrete IFN-γ, which is detected by immobilized specific antibodies and a subsequent color reaction [199]. 96-well ELISpot plates containing a PVDF membrane on the bottom of the wells were activated with methanol, washed three times with ddH$_2$O, coated with 50 μL of anti-bovine IFN-γ antibody (5 μg/mL) cross-reactive with horse IFN-γ at 4° C over night. Subsequently, PBMCs from donor horses with the haplotype of interest were thawed from liquid nitrogen stocks and 2×10^5 cells were seeded in coated ELISpot plates. To each well were added, 10 μg/mL peptide, DMSO in comparable concentrations as a negative control, or 5 μg/mL concanavalin A as a positive control, respectively. Plates were incubated for 20 h at 37° C and 5 % CO$_2$ before they were washed six times with PBS+0.05 % tween20 (PBST). Secreted and captured IFN-γ was detected with 0.25 μg/mL of a biotinylated secondary anti-INF-γ antibody for 2 h at 37° C. Thirty minutes prior to use, avidin peroxidase complex (APC) (Vectastain ABC Kit) was prepared by diluting 20 μL solution A with 20 μL solution B in 5 mL PBST. Plates were washed again six times with PBST and incubated for 1 h at RT with 100 μL of APC. Following six times washing with ddH$_2$O. IFN-γ spots were visualized by the addition of 100 μL of AEC solution for 10 min. AEC solution was prepared by dissolving an AEC tablet in 2.5 mL dimethyl formamide, diluting with 47.5 mL of AEC buffer containing 3.5 mM acetic acid and 8.5 mM NaAc, and adding 25 μL of hydrogen peroxide. After a final washing step with deionized water, plates were dried and spots were counted in an ELISpot reader. Responses

against the negative control (DMSO) were subtracted from experimental values and statistical significance was assessed using a one-tailed Student's-test. All samples were done in triplicates. A sample was considered positive if net spot-forming cells (SFCs) per one million cells exceeded 20, the stimulation index (SI) was greater than 2.0 and probability ($p \leq 0.05$) was reached.

5.2.3.4 Determining frequencies of IFN-γ-secreting T lymphocytes by the use of ELISpot

PBMCs from EHV-1 infected or naïve horses were isolated and stored as previously described. 2×10^5 cells were infected with EHV-1 strain RacL11 at an MOI of 1 or mock-treated with medium. After 3 h of incubation at 37 °C, cells were washed three times with RPMI medium. Subsequently, EHV-1- as well as mock-infected or with 5 µg/mL concanavalin A stimulated PBMCs were transferred to antibody-coated ELISpot plates. After 20 h incubation at 37 °C, ELISpot plates were developed and count numbers were determined as described before.

5.2.3.5 Statistical analysis

Statistical analyses for RMA-S MHC stabilization assays were performed using GraphPad Prism 5 software. Fluorescent shift data were tested for normality of distribution and analyzed for significance using one-way ANOVA.

6. Results

6.1 Equine MHC class I

6.1.1 Stabilization of equine MHC-I molecules by externally supplied peptides

RMA-S MHC stabilization assays have been previously used to identify peptide binders of MHC-I molecules [200]. For all four alleles, respective RMA-S clones were available and their idendity was confirmed by sequencing. As at least one EHV-1 ICP4-derived CTL epitopes had previously been found to be presented by Eqca-1*00101, a peptide library consisting of 295 peptides of 15 aa in length each overlapping with the previous one by ten residues, and spanning the entire deduced amino acid sequence of ICP4 was used in MHC-I stabilization assays. Cells maintained at room temperature (RT) served as positive control; subsequent incubation at 37 °C (RT/37) for five hours yielded the negative control. A peptide was considered a stabilizer if it induced a fluorescent shift of 20 % in regard to the negative control (RT/37). Putative stabilizers were tested in five independent experiments and considered positive if they showed a statistical significant difference ($p < 0.05$) to the negative control. Peptide p250 was consistently negative for all alleles, and therefore chosen as an additional negative control. Experiments were done in duplicates and repeated four times. Of those, three times were carried out in a blind fashion. Figure 9 exemplary shows a histogram for the negative and positive control as well as one non-stabilizing (p250) and one stabilizing (p244) peptide for the Eqca-1*00101 allele.
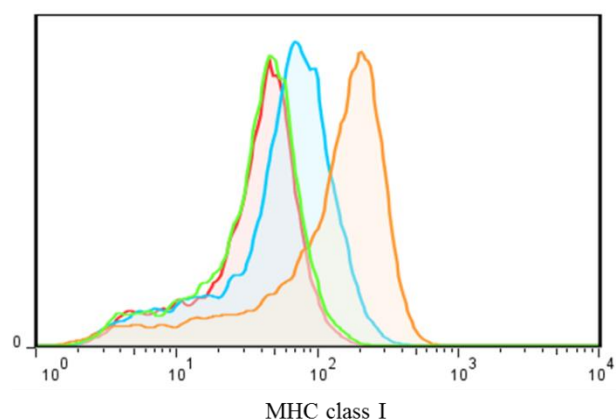
Figure 9: FACS histogram for negative control (red), p250 (green), p244 (blue) and positive control (orange) for Eqca-1*00101. Peptide p244 is a stabilizer.

Eqca-1*00101

Five peptides derived from the ICP4 protein (IE peptides) of EHV-1 were found to stabilize equine MHC-I molecules of the Eqca-1*00101 allele (Figure 10, A) in the RMA-S MHC stabilization assay.
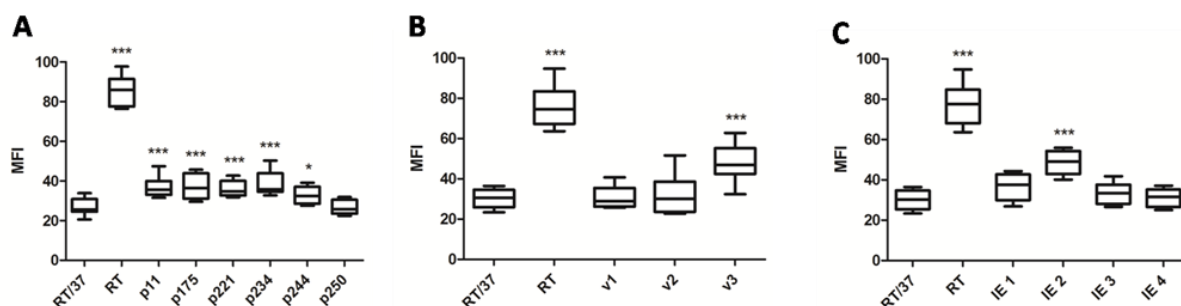
Figure 10: Peptide stabilizers for Eqca-1*00101. A) Five ICP4-derived peptides stabilize; p250 serves as negative control. B) PolyA variants confirm putative anchor positions and C) four IE peptides selected by the presence of the preliminary binding motif

All peptides have an acidic residue, predominantly aspartic acid (D), near the N-terminus and a small hydrophobic one (isoleucine (I), leucine (L), alanine (A), phenylalanine (F), and valine (V)) seven to ten positions downstream (Table 2). Those putative primary anchors were incorporated into polyalanine (polyA) variants of nine amino acids in length and consecutively replaced by alanine to confirm their specificity. The presence of D in P2 and I as a surrogate for small hydrophobic residues in P9 was sufficient for stabilization (v3) (Figure 10, B).

Table 2: Peptide sequences for IE 15mers, polyA variants and putative IE epitopes for Eqca-1*00101

peptides	No.	Stab.	1	2	3	4	5	6	7	8	9	10	11	12					
IE-derived 15mers	p11	√			S	D	F	A	P	D	L	Y	D	F	I	E	S	N	D
	p175	√		A	E	G	S	L	Q	T	L	P	P	L	W	P	T	V	
	p221	√	P	P	A	R	D	G	A	R	F	G	E	L	A	A	S		
	p234	√		W	S	D	L	K	G	G	L	S	A	L	L	A	A	L	
	p244	√			G	D	L	A	F	T	G	C	V	E	Y	L	C	L	R
polyA variants	v1	x	A	A	A	A	A	A	A	A	I								
	v2	x	A	D	A	A	A	A	A	A	A								
	v3	√	A	D	A	A	A	A	A	A	I								
putative IE epitopes	IE1	x	G	D	S	D	P	T	H	R	L								
	IE2	√	R	D	G	A	R	F	G	E	L								
	IE3	x	Q	D	E	D	L	L	G	G	L								
	IE4	x	I	D	G	A	G	D	V	E	L								

Based on these results, four peptides carrying the putative primary anchor residues from the ICP4 protein were tested for stabilization and only one (IE2) was found to stabilize Eqca-1*00101 (Figure 10, C). IE2 is also present as nested peptide in p221. In conclusion, four pep-

tides derived from the ICP4 protein were shown to stabilize Eqca-1*00101. They share specificities for acidic residues, predominantly aspartic acid, at P2, and leucine at P9.

Eqca-N*00101

Eqca-N*00101 was found to be stabilized by 14 IE peptides (Figure 11, A).
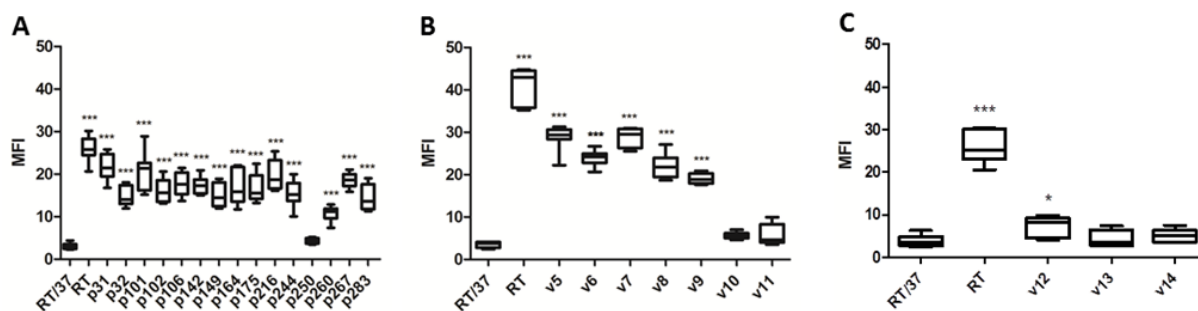
Figure 11: Peptide stabilizers for Eqca-N*00101. A) 14 IE peptides are stabilizers; p250 serves as additional negative control. B) Peptide variants based on stabilizer p31. C) PolyA variants

The majority (n = 13; 92.8 %) of these peptides had proline (P) or A near or at the N-terminus and a small hydrophobic residue seven to ten positions downstream. Eight peptides show a preference for proline (57.1 %), four peptides are more likely to use alanine (28.5 %) either because there is no proline present at all (p142 and p244) or only in a position that would restrict peptide lengths to less than nine residues (p267 and p175). Only p106 does not conform to these criteria (Table 3).

Table 3: Peptide sequences for IE 15mers, p31 variants and polyA variants for Eqca-N*00101

peptides	No.	Stab.									1	2	3	4	5	6	7	8	9	10	11	12				
IE-derived 15mers	p31	√					A	P	S	P	A	P	G	L	A	A	M	L	K	M	V					
	p216	√						F	P	E	A	W	R	P	A	L	T	F	D	P	Q	A				
	p101	√			Y	A	E	A	Q	A	P	V	P	V	F	V	P	E	M							
	p142	√						L	G	D	A	C	V	L	A	C	Q	A	V	F	E	A				
	p267	√								R	A	G	E	R	T	R	V	P	L	A	P	R	E	Y		
	p106	√	V	R	M	V	F	E	S	R	E	A	M	S	W	L	Q									
	p175	√						A	E	G	S	L	Q	T	L	P	P	L	W	P	T	V				
	p164	√					R	A	R	A	S	A	W	T	V	T	Q	A	V	F	S					
	p102	√								A	P	V	P	V	F	V	P	E	M	G	D	S	T	K		
	p32	√										P	G	L	A	A	M	L	K	M	V	H	S	S	V	A
	p244	√					G	D	L	A	F	T	G	C	V	E	Y	L	C	L	R					
	p283	√			S	E	L	L	I	P	E	L	L	S	E	F	C	R	V							
	p149	√						S	P	A	C	P	P	E	A	L	C	S	N	P	A	G				
	p260	√										P	G	L	F	A	R	A	E	A	A	F	A	R	L	Y
p31 variants	v5	√					A	P	S	P	A	P	G	L	A	A	M	L	K	M	V					
	v6	√						P	S	P	A	P	G	L	A	A	M	L	K	M	V					
	v7	√							S	P	A	P	G	L	A	A	M									
	v8	√							A	P	A	P	G	L	A	A	M									
	v9	√							S	A	A	P	G	L	A	A	M									
	v10	x								P	A	P	G	L	A	A	M	L								
	v11	x									A	P	G	L	A	A	M	L	I							

polyA variants	v12	√		A	P	A	A	A	A	A	A	M
	v13	x		A	P	A	A	A	A	A	A	L
	v14	x		A	P	A	A	A	A	A	A	A

To further elucidate these results, the best stabilizer (p31 = v5) was successively trimmed down to a putative minimal binding epitope (v6 and v7) and possible N-terminal anchors were replaced by alanine (v8 and v9). V10 represents a peptide with alanine at P2 and proline at P1 and P3. In v11, the second proline was assumed to be the anchor at P2 and a common C-terminus was added. Interestingly, substitution of proline at P2 with alanine (v9) did not abolish stabilization. Alanine, however, has been found to be sometimes a sufficient surrogate for proline, probably because they share similar spatial configurations, which could reflect the presence of P and A in the stabilizing peptides. The last peptides did not stabilize even in the presence of alanine or proline at P2 possibly due to the lack of a correct C-terminus or other deleterious effects (Figure 11, B). PolyA variants confirmed the need of methionine at the C-terminus albeit stabilization was rather weak (Figure 11, C).

In total, Eqca-N*00101 was found to be stabilized by 14 IE peptides. Anchor positions can probably be found at P2 and the C-terminus with specificities for proline or alanine and possibly methionine, respectively.

Eqca-16*00101

For Eqca-16*00101 no stabilizers could be identified.

Eqca-1*00201

For Eqca-1*00201, ten IE peptides were found to be stabilizers (Figure 12, A).
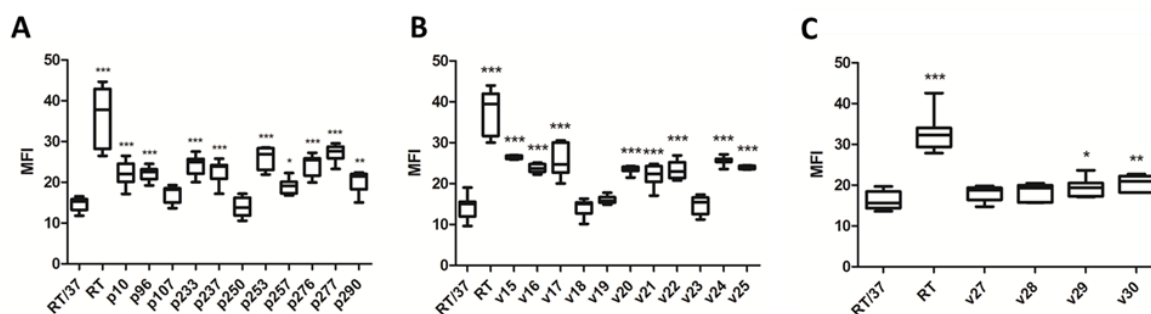
Figure 12: Peptide stabilizers for Eqca-1*00201. A) 9 IE peptides were found to stabilize Eqca-1*00201. B) Peptide variants based on stabilizer p277. C) PolyA variants

All peptides had arginine near the N-terminus preceded by polar or positively charged residues in most cases. Only three exceptions could be observed (p253, p96 and p290). Near the C-terminus, hydrophobic residues are predominant (Table 4). This could point to primary anchors in P3 and at the C-terminus, and maybe a strong secondary anchor in P2. To further assess these observations, variants based on p277 were tested.

Table 4: Peptide sequences for IE 15mers, p277 variants and polyA variants for Eqca-1*00201

peptides	No.	Stab.									1	2	3	4	5	6	7	8	9	10	11	12				
IE-derived 15mers	p277	√						H	R	A	A	N	R	W	G	L	G	A	W	L	R	P				
	p253	√								I	Y	M	R	A	A	L	T	P	R	V	A	C	A	V		
	p233	√										R	R	E	P	L	W	S	D	L	K	G	G	L	S	A
	p276	√	G	A	A	Q	S	H	R	A	A	N	R	W	G	L	G										
	p10	√							M	A	S	Q	R	S	D	F	A	P	D	L	Y	D	F			
	p237	√									G	N	R	I	L	T	K	R	S	H	A	W	A	G	N	
	p96	√										V	R	Y	G	G	T	G	D	S	R	D	G	L	W	
	p290	√						R	R	R	A	P	R	V	D	W	E	P	G	F	G	S				
	p257	√							R	H	L	S	R	A	V	L	T	S	S	T	L	F	G			
	p107	√									E	S	R	E	A	M	S	W	L	Q	N	S	K	L	S	
p277 variants	v15	√						H	R	A	A	N	R	W	G	L	G	A	W	L	R	P				
	v16	√									A	N	R	W	G	L	G	A	W							
	v17	√									D	N	R	W	G	L	G	A	W							
	v18	x									A	A	R	W	G	L	G	A	W							
	v19	x									A	N	A	W	G	L	G	A	W							
	v20	√									A	N	R	A	G	L	G	A	W							
	v21	√									A	N	R	W	A	L	G	A	W							
	v22	√									A	N	R	W	G	A	G	A	W							
	v23	x									A	N	R	W	G	L	A	A	W							
	v24	√									A	N	R	W	G	L	G	D	W							
	v25	√									A	N	R	W	G	L	G	A	A							
polyA variants	v27	x									A	N	A	A	A	A	G	A	A							
	v28	x									A	A	R	A	A	A	G	A	A							
	v29	√									A	N	R	A	A	A	A	A	A							
	v30	√									A	N	R	A	A	A	G	A	A							

A minimal binding epitope was assumed positioning the arginine residues in P3 and asparagine in this case in P2. Consecutively, all residues were replaced by alanine or aspartic acid where alanine was already present. If P2, P3 or P7 were exchanged, stabilization was abolished strongly suggesting them to be specific anchor positions (Figure 12, C). Whether all three positions can be considered primary anchors remains to be elucidated in more detailed approaches. Interestingly, polyA variants carrying alanine at P7 did stabilize (v29), albeit very poorly, which is in contrast to previous findings (v23). In v23, stabilization is possibly rescued by effects of downstream residues.

In total, Eqca-1*00201 was stabilized by ten ICP4 peptides, and probably uses positions 2, 3 and 9 as anchors with preferred specificities for N, R and hydrophobic/aromatic residues, respectively.

6.1.2 Elution of natural equine MHC class I ligands

Elution and identification of natural peptide ligands of four equine MHC class I alleles from stably transfected P815 cells by affinity chromatography and mass spectrometry was performed to derive preliminary motifs characterizing their individual binding specificities and to obtain high affinity binders for subsequent assay development. To control for potential contaminations with co-purified P815-endogenous H-2 class I MHC, equine MHC-I-specific monoclonal as well as isotype control antibodies were used for purification.

Eqca-1*00101

For the Eqca-1*00101 allele, 24 different peptides were identified (Table 5). Most of them (n = 17; 71 %) had a length of nine aa, which is in agreement with the canonical ligand size for most human, non-human primate and murine MHC-I alleles [115]. Fewer peptides were found to have ten or eleven residues (n = 4; 16.7 % and n= 3; 12.5 %, respectively) and none with less than nine (Figure 13). The majority of peptides (n = 20; 83.3 %) had an aspartic acid at P2 and isoleucine at the C-terminus (n = 19; 79.2 %). The only other residues found at position 2 were glutamine and glutamic acid, which share conserved or semi-conserved chemical specificities with D. At the C-terminus, isoleucine-similar hydrophobic and/or aliphatic residues (L, V and F) were observed as well as one peptide with N. Around half of the peptides (n = 13, 54.2 %) carried a positively charged aa (arginine (R) or lysine (K)) at position 1 and even 75 % (n = 18) had a charged or polar (N, Q and S) residue at position C-1.

Taken together, these data point towards position 2 and the C-terminus being a primary anchor position with favored acidic and hydrophobic/aliphatic specificities, respectively, and a predominant ligand size of 9 aa.

Table 5: Sequences, length and affinities (*expressed in IC$_{50}$nM) for peptides eluted from Eqca-1*00101

Sequence	Len	Eqca-1*00101 binding*	Sequence	Len	Eqca-1*00101 binding*
SDYVKVSNI	9	0.66	LDKAQLLEI	9	59
SDIYKIREI	9	1.4	KDTKQILRI	9	77
GDYNGHVGL	9	1.5	LEQEELDKN	9	90
RDLERGVEI	9	2.2	SQSTVKVLI	9	650
RDLPSHTKI	9	4.5	ADDPSLSLI	9	1073
ADLKKAQEI	9	7.7	KDYSKKVTMI	10	16
RDFKGKILI	9	9.1	RQADFEAHNI	10	67
GDMKNKVVI	9	16	ADFDGVLYHI	10	92
KDMDIITKF	9	16	RQMSGAQIKI	10	357
KDAGLVKSL	9	20	KDVPNSQLRHI	11	45
RDIPIVHRV	9	37	RDIETFYNTSI	11	0.14
VDIAAFNKI	9	48	KDVEGSTSPQI	11	43

Eqca-N*00101

For Eqca-N*00101, 28 peptides were recovered with a predominant ligand size of nine residues (21/28; 75%). Interestingly, there seem to be present two different sets of peptides with different binding motifs. The first set comprises peptides 1 – 20 (Table 6, black). Of those, four peptides had eleven, and two peptides twelve or 13 residues, respectively. No 10mers were found (Figure 13). The majority of these peptides (13/20; 65 %) carried proline in position 2. The residual ligands had in low frequencies alanine, valine, isoleucine or leucine in that position, which all share hydrophobic chemical properties. These data suggest position 2 to be a main anchor position with preferred specificities for proline. A comparably narrow specificity can be observed for the C-terminal position suggesting a second anchor where predominantly aromatic residues (14/20; 70 %) or to a lesser extend small hydrophobic ones can be found, although two examples of polar amino acids are present. Also notable are positions 1 and C-1 where with 95 % (19/20) polar or with 30 % (6/20) positively charged residues are predominant, respectively.

The second set (Table 6, grey) resembles specificities identified for Eqca-1*00101 with D and Q in P2. In cross-contamination tests, it was confirmed that the first set of peptides primarily binds Eqca-N*00101, and the second set Eqca-1*00101. Therefore, it is likely that the second peptide set is the result of a contamination. Both alleles (Eqca-1*00101 and Eqca-N*00101) were purified using the same columns, which even after extensive washing could be a source of the contamination. Nevertheless, cross-contamination tests as well as binding motifs identified later on strongly point to contamination of Eqca-N*00101-binding peptides with Eqca-1*00101 peptides. Therefore, the first set of eluted peptides were used to establish a binding assay for Eqca-N*00101.

In summary, position 2 and the C-terminus seem to be primary anchor positions with preferences for proline and small aromatic residues.

Table 6: Sequences, length and affinities (*expressed in IC_{50} nM) for peptides eluted from Eqca-N*00101

Sequence	Len	Eqca-N*00101 binding*	Eqca-1*00101 binding*	Sequence	Len	Eqca-N*00101 binding*	Eqca-1*00101 binding*
DAANLRETF	9	0,96	-	DIVMTQSHKFM	11	9,95	63.291
DAKEKMREF	9	6,87	-	DPHNSRRGFFF	11	10,82	63.291
DPFGRPTSF	9	7,40	-	TPAPAGGYQEF	11	1447,60	63.293
LPYNHQHEY	9	8,41	1.663	YLKKTISRSPG	11	7012,62	63.291
NPWDNKAVY	9	8,97	15.060	YYLKKTISRSPG	12	3196,93	32.258
NVKVDPEIQ	9	14,91	27.027	NYYLKKTISRSPG	13	5307,86	53.763
SPRIEITSY	9	16,73	9.633	RDFKGKILI	9	64.102	4
SPYHQPTSY	9	18,18	-	KDIKESHLI	9	8928,00	4
TPAEIREEF	9	26,94	-	KQMEQISQF	9	1497,00	27
TPAHHYNDF	9	72,65	-	KQAEIVKRL	9	642,50	68
TPIEGMLSH	9	88,48	-	RDLERGVEI	9	-	175
TPIQSKEAY	9	102,27	-	KQLEDGRTL	9	34.246	1.766
TPLDRVTEF	9	1616,55	17.006	AQFSQEIGL	9	8143,00	10.917
TPYEGQRSY	9	16181,23	-	RQYDSRGRFGEL	12	58139	754

Eqca-16*00101

From Eqca-16*00101, 40 peptides were recovered, 90 % (n = 36) of them being nine aa in length and 10 % (n = 4) ten aa (Table 7). A putative binding motif is not as obvious, though. Position 2 favors small hydrophobic/aliphatic residues like valine (n = 22; 55%) or T, L, I, M and A (n = 14; 35 %), and with L (n = 19; 47.5 %), M, T, I and V similar specificities are present at the C-terminus. Positions 1, 4 and 8 seem to favor acidic and related residues (n = 19; 47.5 %; n = 24; 60 %; or n = 16; 40 %), but 12.5 %, 22.5 % and 27.5 %, respectively, of the peptides carry basic residues, which are oppositely charged, in these positions.

In total, Eqca-16*00101 seems to use classical anchors in P2 and at the C-terminus with predominantly hydrophobic residues at both positions.

Table 7: Sequences, length and affinities (*expressed in IC$_{50}$nM) for peptides eluted from Eqca-16*00101

Sequence	Len	Eqca-16*00101 binding*	Sequence	Len	Eqca-16*00101 binding*
GTIDEIQKL	9	0,28	GVIKVFNDM	9	18,25
DVKEQIYKL	9	0,99	EVEESINRL	9	26,69
EIQDKLSEV	9	1,09	KVIDGLETL	9	28,15
DVANKIGII	9	1,64	AVHEQLAAL	9	33,77
ETQKSIYYI	9	1,64	AVKDTILQL	9	42,61
EIPDDLKQL	9	2,50	ETTGQLTKI	9	44,77
EVAEHVQYM	9	2,96	EMHNKIYSI	9	49,59
EIIKDIHNL	9	3,10	EVQNEIDRL	9	52,29
SVLGKVNEI	9	3,24	EIKDFLLTA	9	86,34
GVYGDVQRV	9	6,75	EVVQKLTQM	9	203,05
KIYEGQVEV	9	7,37	EVVRQLQEI	9	242,39
AVIGDVIRV	9	8,32	EVLEVMQHM	9	277,44
EAVKIINEL	9	8,64	RVSDSVGFL	9	320,77
AIVDKVPSV	9	12,53	HVLPELNKM	9	803,73
TTIEAIHGL	9	12,88	SVVRDIQNT	9	2551,60
SSLEKSYEL	9	13,10	DVIAQGVGKL	9	0,20
SVLKGLEKM	9	13,43	EVTDHNLHEL	10	0,60
TIMPKDIQL	9	13,98	SQYNGNLQDL	10	1,74
IQKESTLHL	9	14,13	RLYPEGLAQL	10	4,10
KVPDDIYKT	9	15,89	EQIRNQLTAM	10	8,91

Eqca-1*00201

A total of 40 peptides were identified for the Eqca-1*00201 allele with 62.5 % (n = 25) 9mers, 10 % (n = 4) 10mers and 27.5 % (n = 11) 11mers (Table 8). Interestingly, three positions seem to be important for specificity. Polar residues (N, S, Y, H and T), predominantly N, (n = 30; 75 %) are preferred at P1; at P2 all but one peptide (n = 39; 97.5 %) carry an arginine, and small hydrophobic residues (L, I, V and F) make up the C-terminus. Less striking but also notable is position C-2 where predominantly polar or charged residues (T, S, R, L, D and E) can be found. Taken together, two or three anchor positions can be found at P1, P2 and

the C-terminus with specificities for asparagine, arginine and hydrophobic/aromatic residues, respectively.

Table 8: Sequences, length and affinities (*expressed in $IC_{50}nM$) for peptides eluted from Eqca-1*00201

Sequence	Len	Eqca-1*00201 binding*	Sequence	Len	Eqca-1*00201 binding*
NRMPQPLSM	9	1,31	SRDPTVTMV	9	272,11
MRVNDSTML	9	3,73	GYKAGMTHI	9	272,51
NRMADAIIL	9	3,83	NRQVQHILA	9	356,38
NRTPSTVTL	9	6,86	NRDVPNETL	9	437,68
NRAPNTAEL	9	7,40	NRGEVVTTF	9	950,28
NRDRSWTEL	9	7,62	NRVAEFHTEL	10	1,26
MRPLDAVEL	9	9,96	NRQPYAVSEL	10	3,04
NRLVQTAEL	9	13,18	NRDAAARLTI	10	10,41
NRFSSYTRL	9	15,23	NRDKYIKDDF	10	12,58
SRPLSAVEL	9	15,66	NRDKHSAVAYF	11	0,83
YRRKDGVFL	9	15,84	NRLENDGATAL	11	1,61
MRDSVLTHL	9	15,98	NRVEVGKDQEF	11	5,08
NRIANPDQL	9	19,80	NRDKNSIYSGL	11	5,16
NRTPAPVSV	9	23,36	NRDLQADNQRL	11	9,46
HRIPTNDEL	9	36,52	NRVDPNGSRYL	11	10,19
NRVTALTQM	9	49,20	NRMQAGEIGEM	11	10,35
NRILTTDGL	9	53,26	MRANNPEQNRL	11	10,54
NRRELTDKL	9	96,04	TRKPMTTKDLL	11	31,21
NRMITQTSI	9	107,06	NRIQHPSNVLH	11	67,31
NRDVAKSVL	9	144,55	NRVSQQNKINL	11	151,41
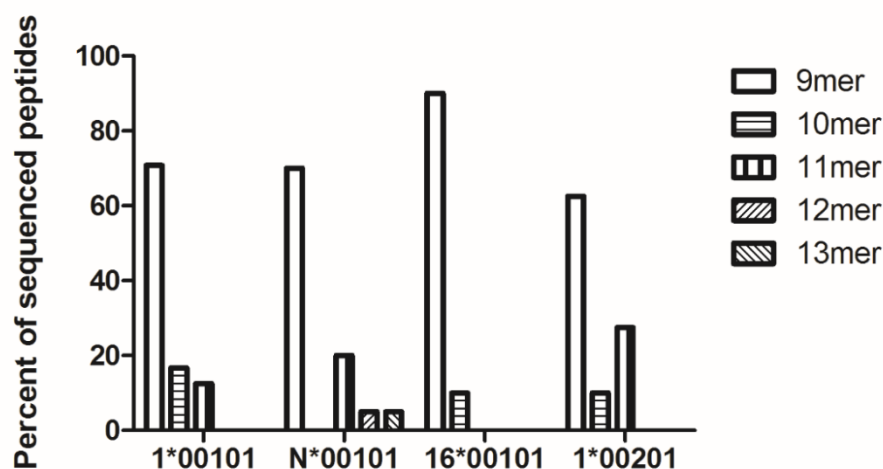
Figure 13: Length distribution of eluted peptides. For all alleles, the majority of all peptides had a length of nine residues. Limited numbers of ligands with more than nine residues and none with less were found.

In total, peptide sets for all alleles conform to the conical ligand size with predominantly 9mers and to a lesser extend longer peptides (Figure 13). It should be noted, however, that the number of eluted and sequenced peptides is not necessarily indicative of the respective repertoire size. Only limiting sequencing was performed for the purpose of defining putative binding motifs and identifying candidate peptides for assay development.

6.1.3 Development of high throughput equine MHC-I binding assays

To identify quantitative binding motifs describing detailed specificities of Eqca-1*00101, Eqca-N*00101, Eqca-1*00201 and Eqca-16*00101, MHC-I molecules were purified from respective P815 clones by affinity chromatography. From the sets of eluted peptides, candidates from all alleles were selected carrying a tyrosine residue necessary for I^{125} radiolabeling and tested against serial dilutions of the respective MHC preparations. MHC concentrations and standard peptides associated with the highest counts and signal-to-noise ratio were chosen for subsequent binding assays (Table 9). Selected standard peptides were tested for binding to MHC-I molecules against a panel of eluted and hence highly affine peptides (Tables 5, 6, 7 and 8).

Table 9: Assay specifications of *in vitro* binding assays for all ELA alleles

Allele	Standard Peptide	*Mus musculus* provenance	Affinity	Signal-to-noise ratio	MHC concentration
Eqca-1*00101	SDIYKIREI	mCG1059215-23	1.4 nM	10	0.3 nM
Eqca-N*00101	SPRIEITSY	nuclear factor of activated T-cells	16.73 nM	20	0.1 μM
Eqca-1*00201	NRQPYAVSEL	ribosomal protein L4	3.05 nM	15	6 nM
Eqca-16*00101	EVAEHVQYM	Geminin	2.96 nM	12	0.1 μM

For Eqca-1*00101, sixteen of the eluted peptides (66.7 %) bound with high affinities of 50 nM or better and six (25 %) with moderate IC_{50} of 500 nM (Table 5). These cut-offs have previously been found to correspond with the majority of MHC-I epitopes [201].

Eqca-N*00101 bound 55 % (n = 11/20) of the eluted peptides with high affinities and 30 % (n = 3/20) with moderate affinities (Table 6).

Twenty-nine peptides (72.5 %) bound Eqca-1*00201 with high, and 10 peptides (25 %) with moderate affinities (Table 7).

Eqca-16*00101 was found to bind 80 % (n = 32/40) of its eluted peptides with high affinities, and 15 % (n = 6/20) with moderate affinities (Table 8).

6.1.4 Determination of quantitative peptide-binding motifs for all ELA alleles

Positional scanning combinatorial libraries (PSCL) tested for binding to MHC-I molecules *in vitro* allow evaluation of the contributions of any of the 20 natural amino acids in any position along the peptide backbone to overall binding. For all four ELA alleles, PSCL were tested and average relative binding (ARB) matrices were calculated as described in Materials and Methods. The specificity factor (SF) represents the average affinity of all residues in one position in respect to the average affinity of the entire library thereby identifying positions associated with specific binding. Primary anchor positions are associated with an SF > 2.4 and a standard deviation (SD) > 2.4, secondary anchors with an SF < 2.4 and an SD > 2.4. The number of residues in each position associated with an ARB 10-fold lower than optimal (0.01) is shown on the bottom.

Eqca-1*00101

Table 10: PSCL-derived ARB matrix for Eqca-1*00101

Residue	Position								
	1	2	3	4	5	6	7	8	9
A	0,660	0,021	0,232	0,249	1,000	0,713	0,327	0,111	0,006
C	0,145	0,015	0,114	0,121	0,380	0,384	0,137	0,212	0,003
D	0,073	**1,000**	0,072	0,446	0,156	0,453	0,081	0,288	0,001
E	0,103	0,025	0,082	0,312	0,240	0,616	0,076	0,175	0,001
F	0,328	0,008	0,592	0,295	0,443	0,938	0,084	0,315	0,109
G	0,200	0,020	0,067	0,243	0,521	0,594	0,198	0,288	0,003
H	0,333	0,030	0,112	0,565	0,452	0,511	0,083	0,084	0,002
I	0,309	0,031	0,991	0,244	0,591	0,655	0,241	0,127	**1,000**
K	1,000	0,011	0,212	0,168	0,278	0,270	0,057	0,121	0,016
L	0,123	0,008	1,000	0,289	0,947	0,759	0,224	0,215	0,196
M	0,699	0,037	0,810	0,538	0,437	0,551	0,308	0,196	0,031
N	0,068	0,017	0,129	0,511	0,357	0,412	0,285	0,234	0,002
P	0,092	0,011	0,068	1,000	0,466	0,248	0,615	0,135	0,001
Q	0,150	0,053	0,223	0,441	0,328	0,262	0,480	0,212	0,001
R	0,937	0,016	0,097	0,587	0,354	0,343	0,194	0,175	0,001
S	0,403	0,018	0,051	0,899	0,808	0,521	1,000	0,325	0,001
T	0,251	0,034	0,053	0,495	0,663	0,758	0,567	0,328	0,010
V	0,158	0,014	0,062	0,300	0,192	1,000	0,547	0,254	0,508
W	0,038	0,011	0,054	0,525	0,752	0,181	0,158	0,248	0,038
Y	0,212	0,011	0,092	0,519	0,259	0,225	0,045	1,000	0,002
German	0,22	0,02	0,15	0,39	0,43	0,47	0,20	0,22	0,01
SD	2,5	2,8	2,7	1,7	1,7	1,6	2,4	1,7	9,2
SF	0,64	6,35	0,95	0,36	0,32	0,30	0,68	0,64	16,98
10-fold	4	19	10	0	0	0	6	1	16

Positions 2 and 9 were found to be associated with the highest SF of 6.35 and 16.98, respectively (Table 10). Based on the previously defined criteria and in good agreement with the RMA-S stabilization and the elution data, P2 and P9 can be regarded as primary anchor residues with strong preferences for aspartic acid in P2 and isoleucine, valine or leucine (ARB = 0.508, 0.196) at the C-terminus. Glutamine, which shares some chemical properties with D, is

tolerated at P2, albeit with a 20-fold lower ARB (0.053). Secondary anchor positions with SD values > 2.4 can be found in positions 1, 3 and 7. Position 1 preferred basic residues lysine or arginine (ARB = 0.937), P3 small hydrophobic residues leucine, isoleucine or methionine (ARB = 0.991, 0.810), and at P7 small residues like serine, threonine and valine (ARB = 0.567, 0.547) can predominantly be found (Table 10). This motif conforms to the specificities known for the HLA supertype B44 [112].

Eqca-N*00101

Table 11: PSCL-derived ARB matrix for Eqca-N*00101

Residue	Position								
	1	2	3	4	5	6	7	8	9
A	0,085	0,504	0,629	0,310	0,239	0,256	0,296	0,292	0,029
C	0,117	0,024	0,078	0,118	0,087	0,219	0,160	0,030	0,006
D	1,000	0,024	0,091	1,000	0,100	0,087	0,129	0,420	0,002
E	0,114	0,025	0,074	0,253	0,220	0,109	0,158	1,000	0,001
F	0,406	0,035	0,950	0,153	0,408	0,164	0,228	0,149	**1,000**
G	0,070	0,033	0,090	0,186	0,283	0,181	0,111	0,173	0,009
H	0,207	0,024	0,244	0,063	0,198	0,107	0,314	0,043	0,031
I	0,066	0,059	0,683	0,096	0,129	0,165	0,294	0,047	0,017
K	0,027	0,037	0,072	0,062	0,188	0,243	0,098	0,091	0,002
L	0,092	0,120	0,803	0,113	1,000	0,208	0,222	0,124	0,030
M	0,118	0,129	0,988	0,136	0,272	0,228	0,274	0,076	0,006
N	0,431	0,026	0,724	0,131	0,196	0,316	0,421	0,051	0,001
P	0,182	**1,000**	1,000	0,207	0,502	0,409	0,980	0,773	0,001
Q	0,086	0,024	0,134	0,055	0,122	0,158	0,297	0,034	0,001
R	0,086	0,026	0,209	0,294	0,211	1,000	0,151	0,084	0,001
S	0,350	0,031	0,048	0,330	0,317	0,182	0,132	0,146	0,003
T	0,526	0,024	0,060	0,206	0,116	0,140	1,000	0,120	0,032
V	0,095	0,283	0,075	0,214	0,083	0,090	0,166	0,161	0,107
W	0,059	0,158	0,521	0,227	0,254	0,061	0,214	0,030	0,575
Y	0,065	0,048	0,238	0,166	0,319	0,987	0,093	0,059	0,220
Geomean	0,14	0,06	0,23	0,17	0,21	0,20	0,22	0,11	0,01
SD	2,5	3,1	3,0	2,0	1,9	2,0	1,9	2,8	8,7
SF	0,85	2,00	0,51	0,68	0,54	0,58	0,51	1,03	10,17
10-fold	10	14	8	4	2	3	2	10	16

With an SF of 10.17, position 9 conforms to the criteria of primary anchor positions with strong preference for phenylalanine. Position 2 is associated with a slightly decreased SF, which reflects not only a preference for proline but also alanine (ARB = 0.504), and several other residues like V, M, L and W may be tolerated. Tryptophan (ARB = 0.575) as well as tyrosine (ARB = 0.22) can be found at the C-terminus (Table 11). With SD values > 2, positions 1, 3 and 8 can be considered secondary anchors with aspartic acid or its amino derivative as well as polar residues (T, S) in P1 (ARB = 0.526 and 0.35), proline, methionine (ARB = 0.988), phenylalanine (ARB = 0.95) or hydrophobic amino acids (L, I, A) (ARB = 0.803, 0.683, 0.629) in P3, and negatively charged residues as well as proline (ARB = 0.773) in P8 (Table 11). These results are in good agreement with the RMA-S stabilization and elution data. This binding motif falls into the HLA supertype B07 [112].

Eqca-16*00101

Table 12: PSCL-derived ARB matrix for Eqca-16*00101

Residue	Position								
	1	2	3	4	5	6	7	8	9
A	0,048	0,105	0,132	0,354	1,000	0,102	0,470	1,000	0,022
C	0,050	0,010	0,065	0,293	0,577	0,218	0,165	0,828	0,011
D	0,102	0,010	0,055	0,952	0,502	0,126	0,145	0,874	0,001
E	1,000	0,010	0,032	0,649	0,413	0,049	0,137	0,299	0,002
F	0,052	0,010	0,125	0,428	0,676	0,171	0,590	0,509	**1,000**
G	0,042	0,010	0,037	0,261	0,869	0,117	0,517	0,185	0,013
H	0,128	0,010	0,053	0,221	0,552	0,063	0,354	0,113	0,004
I	0,048	0,115	1,000	0,284	0,275	1,000	0,158	0,074	0,260
K	0,034	0,010	0,104	0,214	0,572	0,075	0,051	0,100	0,005
L	0,046	0,329	0,759	0,274	0,328	0,799	0,214	0,138	0,257
M	0,204	0,097	0,891	0,333	0,628	0,327	0,195	0,100	0,019
N	0,100	0,010	0,040	0,471	0,818	0,117	0,221	0,128	0,002
P	0,006	0,010	0,058	0,740	0,215	0,058	1,000	0,398	0,002
Q	0,362	0,376	0,030	0,472	0,256	0,075	0,176	0,155	0,001
R	0,074	0,019	0,022	0,295	0,483	0,139	0,184	0,239	0,001
S	0,672	0,038	0,034	1,000	0,503	0,109	0,114	0,270	0,002
T	0,098	0,089	0,047	0,818	0,372	0,136	0,067	0,302	0,022
V	0,058	**1,000**	0,163	0,218	0,267	0,133	0,118	0,134	0,335
W	0,007	0,012	0,054	0,319	0,432	0,046	0,193	0,128	0,011
Y	0,030	0,010	0,071	0,373	0,355	0,167	0,228	0,114	0,004
Geomean	0,07	0,03	0,09	0,40	0,46	0,13	0,20	0,22	0,01
SD	3,5	4,6	3,1	1,6	1,5	2,3	2,1	2,2	8,1
SF	1,52	3,52	1,29	0,28	0,24	0,82	0,54	0,50	10,03
10-fold	14	15	13	0	0	6	2	2	16

The quantitative binding motif for Eqca-16*00101 involves primary anchors in position 2 (SD = 3.52) and at the C-terminus (SD = 10.03) with preferred residues being valine, and to a lesser extent leucine and, considering its differing chemical properties, surprisingly glutamine (ARB = 0.329 and 0.376) or phenylalanine and valine (ARB 0= 0.335), respectively. With SD > 2.4 but SF < 2.4, positions 1 and 3 qualify as secondary anchor positions with preferences for negatively charged residues and small hydrophobic residues (I, M, L), respectively (Table 12). However, the positively charged histidine in P1 is tolerated albeit with an ARB eightfold less (ARB = 0.128). This data reflects the elution results well and resembles specificities known for HLA-A02 [112].

Eqca-1*00201

With SF values of 7.48 and 3.51, Eqca-1*00201 uses position 2 and the C-terminus as primary anchors. P2 only tolerates arginine, and the C-terminus prefers phenylalanine and hydrophobic residues leucine and valine (ARB = 0.835, 0.643). Secondary anchors can be found in P1 and P3 with preferences for asparagine, methionine and histidine (ARB = 0.886, 0.22) or small hydrophobic residues, respectively. P1 is associated with an SF only slightly below the

threshold of 2.4, which qualifies as a very strong secondary anchor. In contrast, P3 tolerates a number of residues with hydrophobic properties, making it a weak secondary anchor (Table 13).

Table 13: PSCL-derived ARB matrix for Eqca-1*00201

Residue	Position								
	1	2	3	4	5	6	7	8	9
A	0,055	0,018	0,172	0,348	0,450	0,391	0,554	0,570	0,083
C	0,154	0,026	0,265	0,298	0,304	0,230	0,418	0,331	0,077
D	0,029	0,018	0,339	0,369	0,368	0,166	0,335	0,696	0,036
E	0,021	0,020	0,299	0,361	0,230	0,367	0,229	0,838	0,024
F	0,117	0,018	0,236	0,344	1,000	0,459	0,257	0,540	**1,000**
G	0,029	0,018	0,066	0,330	0,561	0,519	0,174	0,602	0,033
H	0,220	0,018	0,061	0,227	0,380	0,228	0,250	0,218	0,016
I	0,160	0,018	0,443	0,395	0,586	0,467	0,701	0,420	0,294
K	0,033	0,018	0,090	0,321	0,191	0,185	0,142	0,553	0,012
L	0,086	0,018	0,450	0,416	0,545	0,331	1,000	0,649	0,835
M	0,886	0,021	1,000	0,582	0,582	0,209	0,308	1,000	0,041
N	1,000	0,018	0,167	0,230	0,537	0,599	0,538	0,568	0,013
P	0,035	0,018	0,243	1,000	0,742	0,278	0,647	0,524	0,019
Q	0,043	0,018	0,175	0,817	0,341	0,190	0,352	0,373	0,019
R	0,068	**1,000**	0,266	0,567	0,214	0,324	0,232	0,260	0,016
S	0,091	0,036	0,061	0,421	0,515	0,386	0,272	0,465	0,019
T	0,024	0,018	0,061	0,383	0,418	0,227	0,320	0,509	0,026
V	0,018	0,018	0,135	0,608	0,411	0,262	0,535	0,297	0,643
W	0,054	0,018	0,128	0,298	0,419	0,285	0,165	0,366	0,096
Y	0,079	0,018	0,281	0,390	0,715	1,000	0,359	0,521	0,012
Geomean	0,07	0,02	0,19	0,40	0,44	0,32	0,34	0,48	0,05
SD	3,1	2,5	2,1	1,5	1,5	1,6	1,7	1,5	4,3
SF	2,35	7,48	0,93	0,43	0,40	0,55	0,51	0,36	3,51
10-fold	14	19	5	0	0	0	0	0	16

This pattern reflects the elution as well as the RMA-S stabilization data and has been described in the HLA supertype B27 [112].

In summary, the Eqca-1*00101 binding motif has primary anchors in position 2 with specificities for D, and at the C-terminus for hydrophobic/aliphatic (I, L, V and F) residues. Secondary anchors are present in positions 1, 3 and 7. The Eqca-N*00101 binding motif is for P in position 2, and aromatic-hydrophobic (F, W, Y) residues at the C-terminus with secondary anchors in positions 1, 3 and 8. Eqca-16*00101's primary binding motif involves V, Q and L in position 2 as well as F, W and L at the C-terminus. Secondary anchors can be found in positions 1, 3 and 6. The binding motif for Eqca-1*00201 is characterized also by primary anchors in position 2 where only R is tolerated and at the C-terminus, which prefers F, L and V. Secondary anchors are in positions 1 and 3. All alleles show specificities known from common HLA alleles.

Figure 14 shows a summary cartoon depicting the ELA allele motifs.
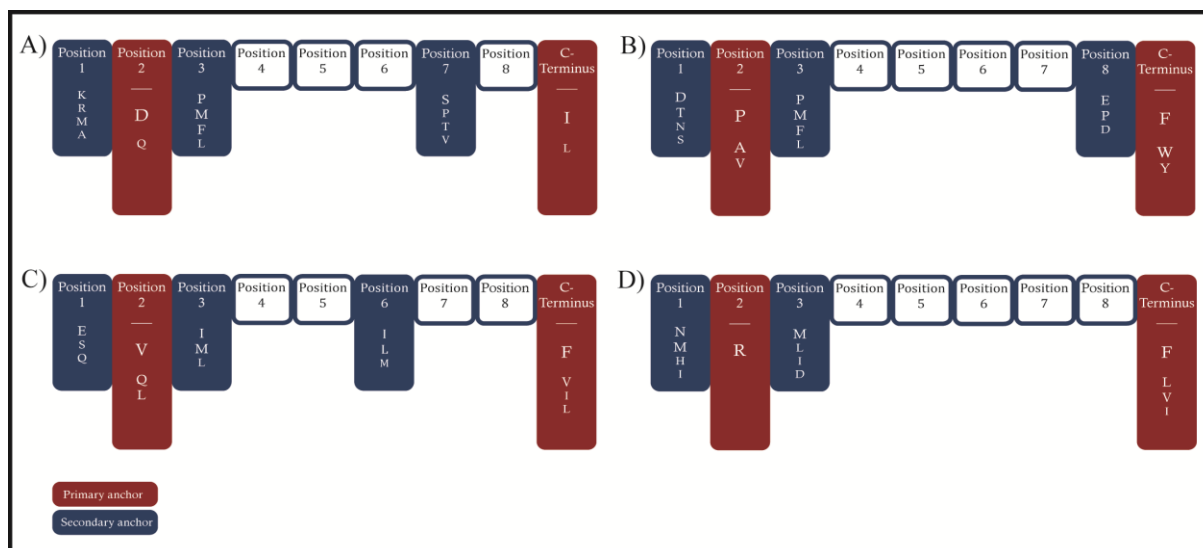
Figure 14: Summary of the binding motifs for Eqca-1*00101 (A), Eqca-N*00101 (B), Eqca-16*00101 (C) and Eqca-1*00201 (D). As determined using PSCL binding analyses, the motives are summarized to show the primary (red fill) and secondary (blue fill) anchor positions. The most preferred residues at the primary anchor positions are indicated by enlarged font. Tolerated anchor residues are also shown

6.1.5 Identification of equine herpes virus type-1-derived binding peptides

Quantitative binding motifs based on PSCL matrices have proven to be very efficient in identifying MHC class I binding peptides [202-204]. All 9mer and some 10mer peptides in the EHV1 proteome (strain Ab4) were scored according to the PSCL-based matrices for Eqca-1*00101 and Eqca-N*00101. The score for each peptide is the product of the values for each peptide residue-position pair. The top 0.5 % scoring 9mer and 0.1 % scoring 10mer peptides as well as a control set representative for the 0.5-100 % scoring range for Eqca-1*00101; and the top 0.5 % scoring 9mers (including the top 2 % scoring 9- and 10mers from ICP4) and a control set for Eqca-N*00101 were selected, synthesized and tested in *in vitro* binding assays. Significant binding was defined using a binding affinity threshold of $IC_{50} \leq 500$ nM. This threshold has previously been associated with T cell epitopes in humans, non-human primates and mice [201, 205].

Eqca-1*00101

For Eqca-1*00101, 194 peptides representing the top 0.5 % scoring 9mers, 38 peptides representing the top 0.1 % scoring 10mers, and 319 9mers representative of the 0.5 – 100 % scoring range as well as 84 10mers representative of the 0.1 – 20 % scoring range were selected and tested for *in vitro* binding. Of the predicted binding 9mers, 55 % in the 0.1 % range

actually bound, 20 % in the $0.1 - 5$ % range, 10 % in the $0.5 - 2$ % range, and 5 % in the $2 - 10$ % range. None were found to bind in the $10 - 100$ % range (Figure 15). In total, 59 of the 513 (11.5 %) 9mer peptides bound Eqca-1*00101 with affinities of 500 nM or better, including 20 that bound with $IC_{50} \leq 50$ nM. This set represents 35 different EHV-1 ORFs. Thirteen (34.2 %) 10mers of the top 0.1 %, 41.4 % of the peptides (n = 12/29) in the $0.1 - 1$ percentile range but only 10 % (n = 2/20) in the $1 - 5$ percentile range bound with $IC_{50} \leq 500$ nM. No peptides included in the $5 - 20$ % range were found to bind Eqca-1*00101.

Eqca-N*00101

For Eqca-N*00101, 214 peptides derived from the entire proteome of EHV-1 were tested for binding in *in vitro* assays. Of 38 peptides representing the top 0.1 % scoring 9mers, 23.7 % (n = 9/38) bound, and of 78 peptides representative of the $0.1 - 0.5$ percentile range, 15.2 % (n = 14/92) were found to bind Eqca-N*00101. In the percentile range of $0.5 - 2$ % scoring peptides, 7.7 % (n = 7/91) of the peptides bound whereas only 4.2 % (n = 1/24) of the peptides in the $2 - 10$ % range and none in the $10 - 100$ % range bound (Figure 15). In total, 26 of 214 peptides bound Eqca-N*00101 with affinities of 500 nM or better, including 7 peptides with $IC_{50} < 50$ nM. This set represents 20 different EHV-1 ORFs.
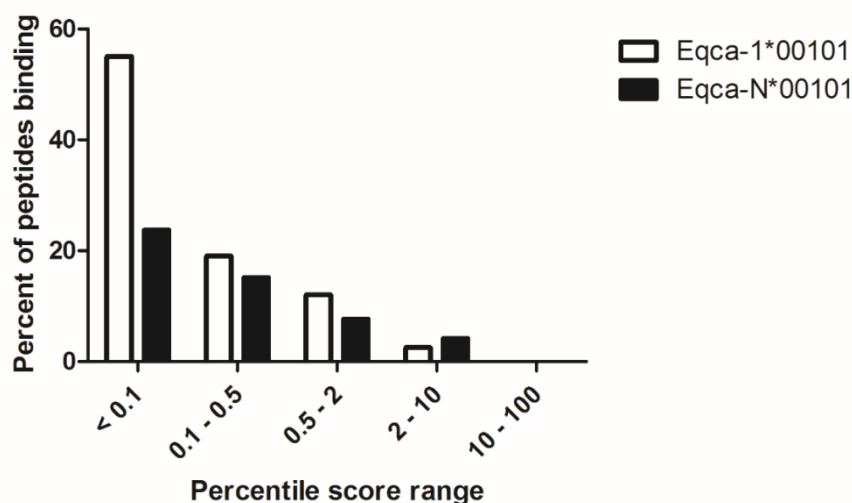
Figure 15: Percent of predicted peptides binding to Eqca-1*00101 or -N*00101 in the respective percentile score ranges

Both algorithms used to predict binding scores worked sufficiently well, and identified several high affinity binders for their respective ELA alleles. For Eqca-1*00101, a larger set of binding data was available, which is reflected by a higher prediction accuracy. The algorithm for

Eqca-1*00101 correctly predicted 55 % of the peptides in the 0.1 % scoring range to bind whereas for Eqca-N*00101 only 23 % actually bound.

Table 14: EHV-1-derived peptides for Eqca-1*00101 (upper table) and Eqca-N*00101 (lower table) and their binding affinities. Depicted are the top ten binders for both alleles. A list of all tested peptides can be found in the supplemental material.

Peptide	Len	Protein	GI no.	Locus	Position	IC50 nM
RDFAVGGGLL	10	nuclear egress membrane protein	50313267	YP_053071.1	7	1.1
LDLSVHPSL	9	envelope glycoprotein C	50313257	YP_053061.1	222	2.5
SDILWQASLI	10	envelope glycoprotein D	126801894	YP_053116.2	151	4.7
KSMRHIGDI	9	capsid portal protein	50313297	YP_053100.1	388	5.1
ADANHTMSI	9	single-stranded DNA-binding protein	50313272	YP_053076.1	655	6.1
SDILWQASL	9	envelope glycoprotein D	126801894	YP_053116.2	151	6.9
VDLRIQADV	9	helicase-primase primase subunit	50313248	YP_053052.1	319	7.9
RDGARFGEL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1059	8.1
RDIPQDGML	9	tegument protein UL21	50313282	YP_053086.1	140	8.8
LDMLNGGFI	9	single-stranded DNA-binding protein	50313272	YP_053076.1	723	10
NPIFYRLAY	9	deoxyribonuclease	50313291	YP_053046.1	101	2,9
TVWYGRVAF	9	membrane protein V1	50313243	YP_053047.1	55	8,4
TPAAAAQPA	9	transcriptional regulator ICP4	50313321	YP_053124.1	95	14,6
SAWNGPMAF	9	ubiquitin E3 ligase ICP0	50313304	YP_053107.1	100	22,8
NPFIGKRPF	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	17	25,7
VPWLNVIPL	9	envelope glycoprotein J	50313312	YP_053115.1	621	30,4
GPFRNQNEF	9	envelope glycoprotein K	50313247	YP_053051.1	148	45,7
DPITFQCLF	9	DNA polymerase processivity subunit	50313259	YP_053063.1	212	53,7
YAASPMPVF	9	virion protein US10	50313319	YP_053122.1	95	57,5
SAFASYSSF	9	envelope glycoprotein H	50313281	YP_053085.1	796	77,7

6.1.6 Quantifying the binding repertoires of Eqca-1*00101 and Eqca-N*00101

The binding rates associated with various prediction percentile intervals were quantified to estimate the total number of binders in the entire set of 39,088 EHV-1 9mers. It was found that Eqca-1*00101 should bind around 77 and Eqca-N*00101 around 211 9mer peptides, which corresponds to a repertoire size of 0.2 % and 0.5 %, respectively. Previous studies evaluated the binding repertoire of 27 well-characterized HLA class I alleles [206, 207], and found an average binding repertoire of 3 %, ranging from 0.07 − 10.4 %. Therefore, the pre-

dicted repertoires of Eqca-N*00101 and Eqca-1*00101 are at the lower limit of what is known for most MHC class I alleles.

6.1.7 Screening for *ex vivo* T cell reactivity in PBMCs from vaccinated or infected horses

To identify immunogenic CTL epitopes, PBMCs from ELA serotyped and multiply EHV-1-infected or with both inactivated and modified-life vaccines multiply immunized horses (Table 15) were used in IFN-γ enzyme-linked immunospot (ELISpot) assays to test previously identified binding peptides of the respective ELA alleles for T cell reactivity.

Table 15: Blood donors for ELISpot analyses. n/a: not available. n/d: not determined. mult. inf: multiply infected

Name	ser. Haplotype	Acc	YOB	Sex	Age	MLV	SN Titer
Crowd Pleaser	ELA-A3	2885	1988	F	28	Yes	512
Twilight	ELA-A3	3729	2004	F	12	Yes	768
Bravo	ELA-A3	3474	1999	M	17	Yes	192
pos. Ctrl.	ELA-A3/-A6	G0	n/a	F	n/a	mult. inf.	16
Leaper	ELA-A3/-A1	G1	n/a	F	n/a	mult. inf.	n/d
Yeager	ELA-A3/-W11	G2	n/a	F	n/a	mult. inf.	n/d
Biter	ELA-A3/-A6	G3	n/a	F	n/a	mult. inf.	n/d
Fiona	ELA-A3	4231	2011	F	5	No	6
Gypsy	ELA-A2	3919	2007	F	9	Yes	2048
Brown Bear	ELA-A2	3475	1999	M	17	Yes	768
Pearl	ELA-A2	3908	2007	F	9	Yes	512
Farah	ELA-A2	4067	2009	F	7	Yes	384
Maya	ELA-A2	4753	2014	F	0.42	No	12

Eqca-1*00101

For Eqca-1*00101, ELA-A3$^+$ PBMCs from one EHV-1 naïve control horse, three repeatedly vaccinated (inactivated and live-attenuated) horses and initially one multiply infected horse (G0) (Table 15) were tested for T cell reactivity against 87 binding peptides. The proteins of provenance of these peptides included the immediate early and several early as well as late proteins, and represented a total of 40 different proteins. Interestingly, only one peptide with the sequence RDGARFGEL (ICP4$_{1059-1067}$) elicited responses only in the multiply infected

horse (G0). No response was noted from any of the vaccinated horses to that or any other of the 86 Eqca-1*00101 binding 9- and 10mer peptides.
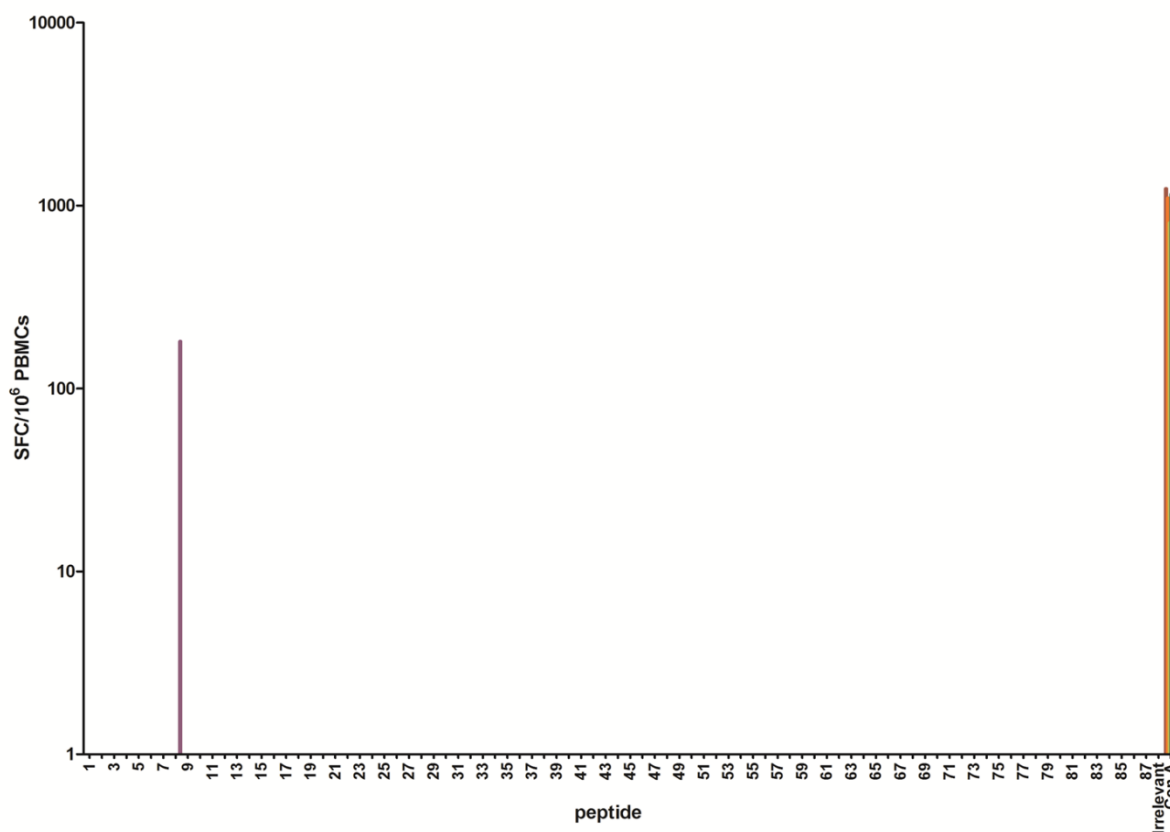
Figure 16: ELISpot interferon-γ T cell responses to Eqca-1*00101 binding-peptides. Binding-peptides (IC_{50}< 500 nM) were tested for T cell responses in five Eqca-1*00101$^+$ horses. Responses ($SFC/10^6$ PBMC) were detected to only the ICP4$_{1059-1067}$ peptide RDGARFGEL (peptide 8), and in only the multiply infected horse. Peptides are sorted left to right from highest to lowest, corresponding to the order in supplemental data.

Eqca-N*00101

All 189 of the EHV-1-derived peptides for Eqca-N*00101 arranged in 19 pools were also tested in ELISpot assays on peripheral blood cells of five ELA-A2 haplotyped horses, which had been repeatedly vaccinated with both inactivated and MLV vaccines. Included were cells from one EHV-1 naïve control horse (Table 15). As opposed to the study with Eqca-1*00101, no multiply infected donor horse was available.

Despite the functionality of the assay, no CTL response after stimulation with any of the peptide pools could be detected in *ex vivo* ELISpot analyses.
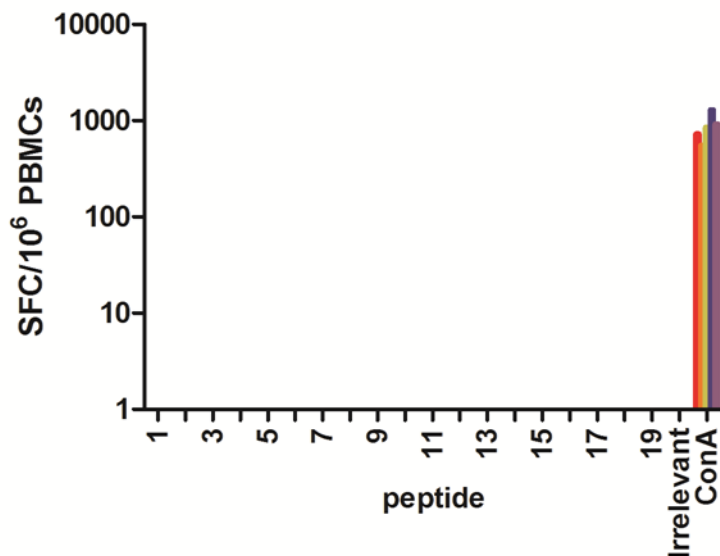
Figure 17: ELISpot interferon-γ T cell responses to Eqca-N*00101 binding-peptides. Binding-peptides ($IC_{50} < 500$ nM) were tested for T cell responses in five Eqca-N*00101[+] horses. No T cell response could be observed.

6.2 Detection of IFN-γ-secreting cells upon virus stimulus

As outlined before, control of EHV-1 infections of horses is mainly attributed to the action of virus-specific T cells. So far, only virulent field strain infections seem to induce protective T cell frequencies after multiple infections and/or reactivations. But robust detection systems for quantification of a cellular immune response are absent. IFN-γ is generally considered a surrogate marker for T cell activity [179]. Thusly, in order to correlate frequencies of virus-specific, IFN-γ-secreting T lymphocytes with immunity to acute and recurrent EHV-1 infections, classical ELISpot approaches were modified to detect responses of EHV-1-specific T lymphocyte populations upon virus stimulus.

6.2.1 RacL11-infected PBMCs co-express MHC-I

As previously reported, equine PBMCs are susceptible to infections with EHV-1. PBMCs were infected with the laboratory strain RacL11, which codes for the green fluorescent protein (GFP) as infection marker at an MOI of 1. This strain was chosen owing to its diminished capacity to downregulate MHC-I surface presentation [61, 190]. After 24 h, cells were incu-

bated with MHC-I-specific antibody and subjected to flow-cytometric analyses. GFP and MHC-I double positive lymphocytes are indicative of MHC-I-restricted peptide presentation.
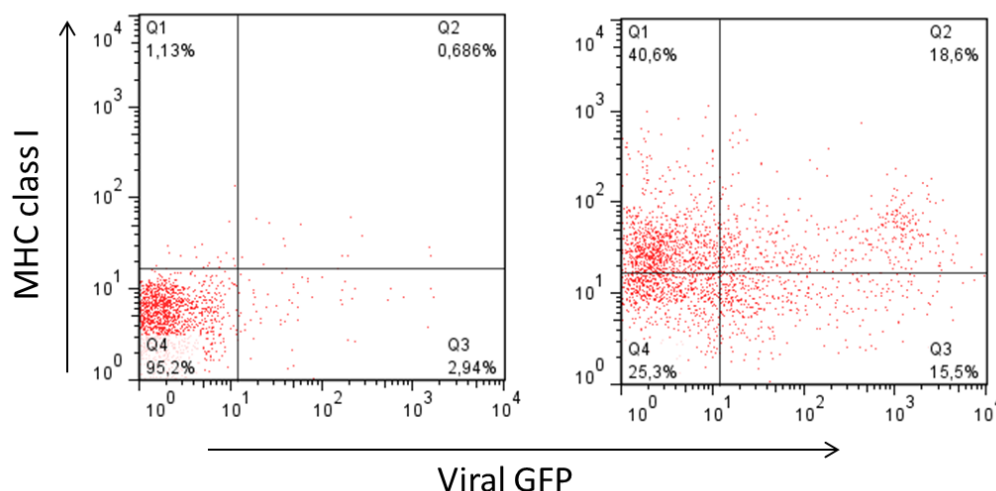
Figure 18: Equine PBMCs infected with EHV-1 strain RacL11 expressing the gene for GFP, and immunostained for MHC-I. Mock-infected and isotype-controlled cells (left) only show background fluorescence. Upon infection, around 30 % of the lymphocyte subpopulation expresses virally encoded GFP, 57 % MHC-I and around 18 % both (right)

Around 18 % of the PBMCs are infected as evident by the production of GFP, and express detectable levels of MHC-I on the cell surface (Figure 18). Although that is no direct proof of MHC-I-restricted presentation of viral antigens, ongoing viral protein production and presence of MHC-I are likely to result in epitope presentation.

6.2.2 PBMCs from a horse with unknown infection history respond to RacL11 infection

In a first pilot experiment, 2×10^5 equine PBMCs from a donor horse with unknown infection history were infected with EHV-1 strain RacL11 at an MOI of 1. The cells responded to the EHV-1 stimulus by secreting IFN-γ. Mock-infected control cells only showed background signals.

Figure 19: ELISpot analysis of mitogen-treated, mock- or RacL11-infected PBMCs. ConA-treated (60 SFC) and RacL11-infected (32 SFC) cells show significantly increased IFN-γ signals. Mock-infected cells are negative (3 SFC).

Although no data on the infection history of the donor horse was available, it seems likely that a horse not housed under quarantine conditions would be naturally infected with virulent EHV-1.

6.2.3 PBMCs from EHV-1 primed but not naïve horses respond to RacL11 stimulus

PBMCs from the ELA-A3-typed horses that had been used in the Eqca-1*00101 epitope experiments, namely G0 and 4231, which were infected with virulent EHV-1 strains on several occasions (primed) or did not come into contact (naïve) with EHV-1 before, respectively, were utilized to detect EHV-1-specific T lymphocyte activity. Two hundred thousand cells were treated with mitogen, mock-infected or infected with RacL11 at an MOI of 1. The previously mapped epitope RDGARFGEL (peptide) and a respective DMSO control were included as additional controls.
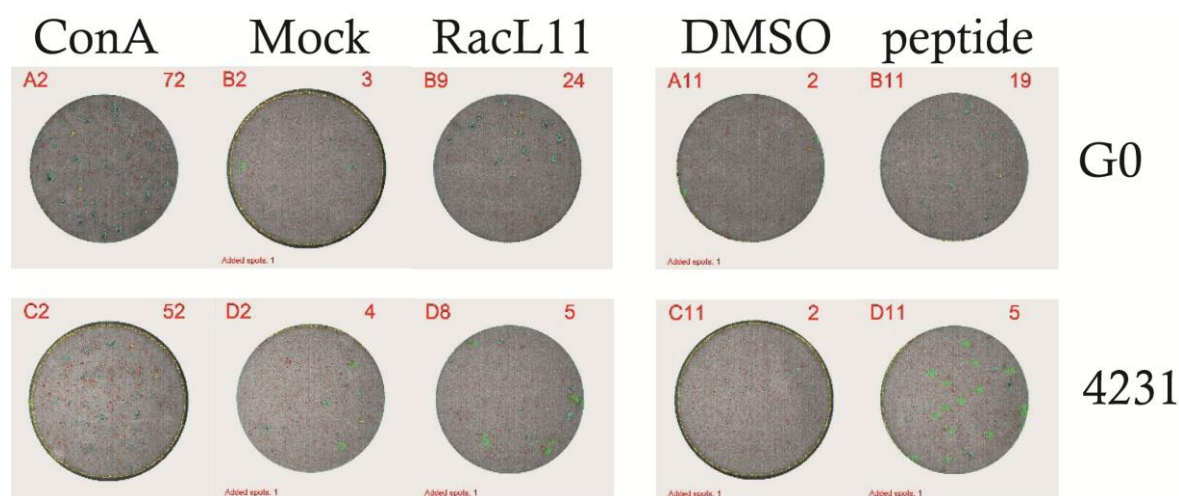
Figure 20: ELISpot analysis of PBMCs from EHV-1 primed (G0, upper panel) and naive (4231, lower panel) horses. ConA-treated cells are positive, mock-infected are negative for both horses (72 and 52 SFC, or 3 and 4 SFC, respectively). RacL11-stimulated cells are only positive for the EHV-1-primed horse (24 SFC as opposed to 5 SFC). The previously identified peptide RDGARFGEL (restricted by Eqca-1*00101) also only elicits a response in the primed PBMCs (19 SFC as opposed to 5 SFC).
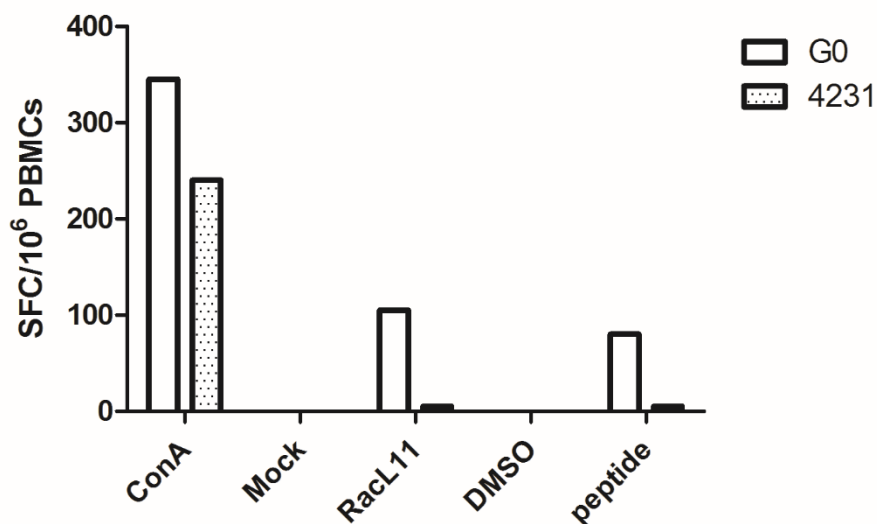
Figure 21: SFC per 10^6 PBMCs for an EHV-1 experienced (G0) and naïve (4231) horse. G0 responds with 120 and 95 SFC/10^6 PBMCs to RacL11 or peptide stimulation, respectively. 4231 remained negative.

As expected, PBMCs only from the EHV-1–primed horse (G0) were reactive upon stimulus with EHV-1 and peptide with 120 and 95 SFC/10^6 cells, respectively. PBMCs from the naïve horse only reacted upon ConA treatment but remained otherwise negative.

To further confirm these results, cells from three additional, multiply infected donor horses (G1-G3) were tested. Also included were negative control cells from horse 4231.
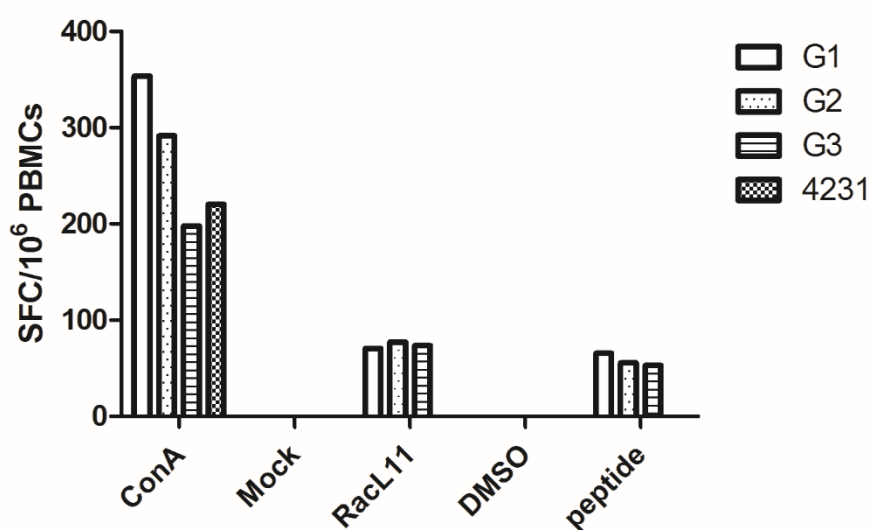
Figure 22: SFC per 10^6 PBMCs for three multiply infected (G1-G3) horses and one EHV-1 naïve horse (4231). As previously seen only EHV-1 experienced horses were responsive to RacL11 and peptide stimulation.

Multiply primed horses G1 – G3 reacted upon stimulation with RacL11 with detectable IFN-γ synthesis. Cells from the naïve control horse remained negative. Additionally, horses G1 - G3 are all Eqca-1*00101$^+$ (Table 15), and confirmed immunogenicity of the RDGARFGEL epitope.

6.2.4 PBMCs from vaccinated horses did not respond to RacL11 stimulus

Next, PBMCs from repeatedly vaccinated horses and another EHV-1 naïve control horse were used.
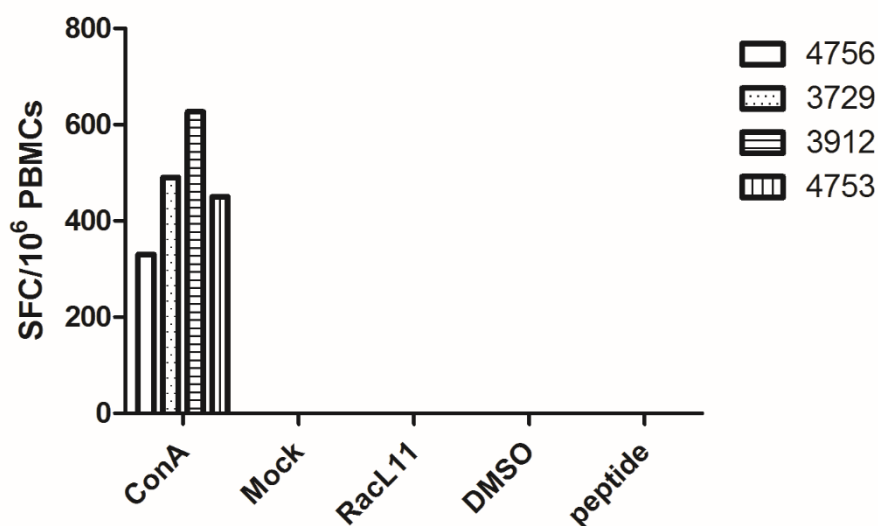
Figure 23: SFC per 10^6 PBMCs for three vaccinated horses and one naïve horse (4753). Cells from all horses were negative upon virus as well as peptide stimulation.

PBMCs from all vaccinated horses and the control horse were unresponsive upon virus and peptide stimulation.

7. Discussion

Cellular immunity against viral infections is of great importance for most viral diseases, albeit some infections can be controlled or prevented by the action of neutralizing and non-neutralizing antibodies alone, as observed for challenge infections after immunization with inactivated equine influenza virus [209]. In general, virus-specific antibodies are thought to play an important role in prevention of cell-to-cell spread and resistance to re-infection. However, the parasitic nature of viral pathogens most of the time deprives circulating antibodies of their targets. Thusly, CD8$^+$ T lymphocytes are key players in the adaptive immune response to intracellular infections. In EHV-1 infections in particular, CTL responses have been found to be superior to antibody-mediated immunity, probably due to the strict cell-association of the virus during viraemic phases.

MHC class I restriction of CTL epitopes is an integral part in the recognition process of infected cells. MHC-I allele-specific peptide presentation determines the T cell receptor repertoire of an individual. Therefore, determining binding motifs for MHC-I alleles can help design new MHC haplotype-specific peptide vaccines that directly stimulate T cell-mediated immunity.

In the present study, detailed binding motifs for the four common horse MHC class I alleles Eqca-1*00101, Eqca-N*00101, Eqca-16*00101 and Eqca-1*00201 were determined. Based on those results, in *in vitro* binding assays high affinity binders derived from EHV-1 were identified for Eqca-1*00101 and Eqca-N*00101.

At first, RMA-S stabilization assays for all alleles were established in our laboratory. As previous studies mapped CTL epitopes restricted by Eqca-1*00101 to the EHV-1 ICP4 protein [77, 78], a peptide library consisting of 295 15mers spanning the entire ICP4 protein was screened for MHC-I stabilizers on RMA-S cells expressing the desired ELA allele. For Eqca-1*00101, Eqca-N*00101 and Eqca-1*00201, five, fourteen and nine peptides, respectively, were identified as stabilizers (Figures 10, 11, and 12). Differences in the mean fluorescence intensity (MFI) among the alleles can be attributed to different expression levels of the respective MHC-I alleles, which could also be confirmed by immunoblotting (not shown). This might be due to the locus of integration in the genome of the respective RMA-S cell clones,

which could influence gene expression, even though all alleles are under the control of the same CMV immediate early promoter.

Since 15mers cannot bind to MHC-I molecules directly, it is likely that extracellular peptidases trim peptides to various length leaving some with the right configuration to bind [210]. That might be one reason, why peptide stabilization never reaches the presentation level of the positive control (cells only incubated at 26 °C) (Figure 9). Nevertheless, a peptide was considered a stabilizer if it caused a fluorescence shift of 20 % in respect to the negative control. This arbitrary threshold proved to be of statistical significance, and identified several high affinity binders as can be seen by comparison with the *in vitro* binding data for Eqca-1*00101 and Eqca-N*00101. Four of the five stabilizers contain the top four minimal binding peptides from ICP4 identified in the *in vitro* binding analyses within their sequence for Eqca-1*00101, and three stabilizers contain peptides found among the top five ICP4 peptides for Eqca-N*00101 (Table 16).

Table 166: ICP4-derived MHC stabilizers and binding peptides for
Eqca-1*00101 (upper panel) and Eqca-N*00101 (lower panel)

15mer stabilizer	Binding peptide	IC_{50} in nM
PPA**RDGARFGEL**AAS	RDGARFGEL	8.1
W**SDLKGGLSAL**LAAl	SDLKGGLSAL	18
SDFAPDLYDFIESND	SDFAPDLYDF	128
GDLAFTGCVEYLCLR	GDLAFTGCV	202
FP**EAWRPALTF**DPQA	EAWRPALTF	127
GD**LAFTGCVEY**LCLR	LAFTGCVEY	130
RARA**SAWTVTQAVF**S	SAWTVTQAVF	248

Interestingly, Eqca-N*00101-stabilizing ICP4 peptides did not bind their MHC molecule with high but only moderate affinities ($IC_{50} > 50$ nM), and the only high affinity ICP4 peptide (TPAAAAQPA: 14.6 nM) did not stabilize. This might be due to the quadruple alanine motif or terminal residues that could prevent extracellular peptidases to trim the 15mer to an appropriate size as has been observed for mitochondrial peptidases [211].

For Eqca-16*00101 no stabilizers were identified. However, fluorescence intensity of the positive control most of the time did not even double in regard to the negative control suggesting low expression levels of this allele. Therefore, the abundance of peptide-stabilized MHC-I on the cell surface might have been below detection limits.

In general, RMA-S MHC stabilization assays identified several peptide stabilizers for three of the four alleles and assumptions on the position and specificities of primary anchors could be

drawn. These premature motifs were confirmed later on. The assay provides a quick way to confirm high affinity MHC binders, or to derive a very preliminary motif. For detailed insights into binding specificities and large scale screening endeavors, however, it is not well suited.

Therefore, we used a more sophisticated approach. A preliminary binding motif was derived for all four alleles by eluting endogenously bound peptides from affinity purified MHC-I. Between 20 and 40 peptides were recovered for each allele and all sets had a predominant ligand size of nine amino acids. Only few peptides with ten, eleven or even thirteen residues were observed, which conforms to the canonical ligand size for most human, non-human primate and mouse alleles known [115] (Figure 13). The elution reactions as well as peptide sequencing methods are not quantitative, so the limited number of sequenced ligands is not necessarily indicative of the real binding repertoire, which was evaluated in *in vitro* binding studies. To establish high-throughput binding assays, tyrosinated endogenous peptides were radiolabeled with iodine[125] and their MHC binding capacities assessed. The assays in turn were used to test a panel of positional scanning combinatorial libraries, resulting in detailed quantitative binding motifs for all four alleles (Figure 14). Briefly summarized, Eqca-1*00101 prefers aspartic acid in P2 and leucine or isoleucine at the C-terminus with secondary anchors at P1, P2 and P7. For Eqca-N*00101, proline in P2 and phenylalanine at the C-terminus could be identified as preferred primary anchors and P1, P3 and P8 as secondary anchor positions. Eqca-16*00101 favors valine in P2 and F at the C-terminus, and has secondary anchor positions in P1, P3 and P6. Eqca-1*00201 has primary specificity for arginine in P2 and F at the C-terminus, and secondary anchors in P1 and P3. These allele motifs can be classified under known HLA supertypes (Eqca-1*00101: HLA-B44; -N*00101: HLA-B07; -16*00101: HLA-A02; -1*00201: HLA-B27) underscoring the general validity of this concept and strengthening the hypothesis of convergent evolution of mammalian MHC [212].

With the motifs for Eqca-1*00101 and Eqca-N*00101, a complete scan of the EHV-1 proteome was undertaken. In total, 87 and 26 EHV-1 peptides were identified that have the capacity to bind Eqca-1*00101 or Eqca-N*00101, respectively. Eqca-1*00101 peptides originated from 40 different viral proteins, including 1, 12 and 27 from immediate early, early and late proteins. Eqca-N*00101 bound peptides from 19 different EHV-1 proteins, including 1, 7 and 11 from immediate early, early and late proteins.

Finally, ELISpot assays to test Eqca-1*00101 and Eqca-N*00101 binding peptides for immunogenicity revealed a narrow CTL epitope reservoir. In fact, only one epitope with the se-

quence RDGARFGEL restricted by Eqca-1*00101 elicited a detectable IFN-γ response, whereas PBMCs from all Eqca-N*00101 expressing A2 horses were unresponsive upon peptide stimulus. However, from 13 donor horses only four had been exposed to virulent EHV-1 strains on multiple occasions in the past years. All other horses were either EHV-1 naïve or had been vaccinated with inactivated vaccines and MLV in the months prior to sampling (Table 15). Only multiply infected horses showed detectable frequencies of MHC-I-restricted and IFN-γ-producing T cells upon peptide stimulation (Figures 16 and 22).

As outlined in the introduction, protection from disease and cell-associated viraemia can only be achieved in multiply infected horses, whereas vaccinated animals only show reduced symptoms. On the other hand, protection from abortion is clearly associated with elevated virus-specific CTL frequencies in multiply infected and vaccinated horses [65]. However, assays used in this study could only detect cytolytic activity of precursor CTLs and ignored the noncytopathic effects of IFN-γ secreted by CD8$^+$ and CD4$^+$ T cells. IFN-γ is likely to play an important role in acute EHV-1 infections as high protein and mRNA levels can be detected after challenge in multiply primed and protected horses. In young EHV-1 inexperienced foals, however, IFN-γ levels are significantly lower, since these foals had not been primed repeatedly [67, 68, 213]. In this light, it seems possible that a broader T cells response than cytolytic effector functions are necessary for protection, and that this broad response is not well stimulated by commercial vaccines or rapidly wanes after vaccination. Our findings also point in that direction, and it would explain why no CTL epitopes could be found for Eqca-N*00101 as no multiply infected donor horses expressing the allele were available.

MHC class I peptide-binding and T cell epitope repertoires should be regarded as fairly independent. The peptide repertoire determined in *in vitro* studies does not depend on the source of the proteins of provenance or their processing, but only on the binding motif. Hence, the repertoire is not expected to change significantly when different pathogens are used to derive-binding peptides. Differences in the amino acid frequencies in proteins of evolutionary distant species could, in theory, influence the peptide repertoire. But these differences are mainly thought to be insignificant [206].

T cell epitope repertoires, however, are not only determined by their ability to bind to MHC-I molecules in *in vitro* settings but rely on several other factors. Abundance of the protein of provenance, substrate specificities of the antigen processing pathway or MHC-peptide stability significantly influence availability of a specific MHC-I-peptide complex at the cell surface [214]. Finally, as TCR repertoires are the result of negative selection in the thymus, TCRs

cross-reactive with self- and similar non-self peptide-MHC complexes are depleted by thymic intolerance [215].

The peptide binding repertoires for both alleles Eqca-1*00101 and Eqca-N*00101 with 0.2 % and 0.5 %, respectively, of all possible 9mer peptides are rather limited, compared to about 1.5 – 3 % on average for human, mouse and macaque class I alleles [206].

When compared to HLA peptide-binding motifs, Eqca-1*00101 ligands are very similar to peptides restricted by HLA alleles in the B44 supertype [112, 216] and the non-human primate alleles Mamu-B*01, Mamu-A*11, and Patr-A*24. In the majority of these cases, however, the predominant specificity is for glutamic acid in position 2, with D only rarely tolerated. The notable exception is Mamu-B*01, which shares a strong preference for D in P2 as well as for I at the C-terminus with Eqca-1*00101. Mamu-B*01 also has a very limited peptide binding repertoire and only few T cell epitopes have been identified to date that are Mamu-B*01-restricted [115].

Eqca-N*00101 also displays a rather limited peptide binding repertoire. However, ligands with P at P2 are commonly produced in the human antigen processing pathway, and alleles associated with the B07 supertype are also very common [154, 155]. Whether that is true for equine antigen processing is not known. However, for epitopes from CMV it could be shown that B07-restricted peptides are recognized more efficiently than those restricted by HLA-A02 pointing to a possibly general immunodominance associated with B07 alleles, which could neutralize the constraints of a limited binding repertoire [217]. Further investigation of additional equine MHC-I alleles and the antigen processing specificities are required to draw more detailed conclusions.

Another hypothesis to explain the identification of limited binding repertoires pertains the various sources of β2-microglobulin (β2m) present in different contexts. The ELA molecules analyzed with the elution studies are largely heterodimers associated with murine β2m or, to a minor extent, with bovine β2m. Additionally, the binding assays are performed with human β2m, and the entire peptide library was tested in ELISpot assays where equine β2m is predominant. While the affinity of β2m may affect peptide binding, especially weak binders, in general it is also accepted that the specificity of peptide selection would not change much. Indeed, given the concordance between the elution (largely mouse β2m) and *in vitro* assays (human β2m) in terms of specificity, and that the binding and T cell (equine β2m) repertoires were addressed using a large set of peptides covering a broad range of predicted affinities, it seems unlikely that the source of β2m has a significant influence.

While no specific mechanism implicating D as CTL epitope repertoire limiting factor is apparent, it is tempting to speculate that the presence of D near the N-terminus of potential ligands somehow influences the efficiency of ligand generation by processing. Whereas nothing specific is known for horse TAP, rat TAP display a remarkably reduced affinity to peptides carrying an aspartic acid or related residues in position 2 or 3 [218]. Also, human ERAP have preferences for N-terminal hydrophobic residues, which also might influence peptide repertoire for the peptide loading complex and hence CTL epitope repertoire [154, 219]. In addition, EHV-1 inhibits TAP function by the action of the early protein pUL49.5, which could prevent any other peptide than the one derived from the immediate early protein to be shuttled into the ER and presented on Eqca-1*00101 [60]. Certain strains of EHV-1 also encode an additional early protein, pUL56, which downregulates MHC-I surface expression by dynamin-dependent endocytosis [61]. Both proteins affect antigen presentation and therefore could limit the antigen repertoire presented to CTLs.

In the sequences of 11 EHV-1 strains deposited in GenBank, no DNA sequence variation within the ICP4 epitope could be found to implicate that escape variants have arisen. Several interaction sites have been reported for the EHV-1 sole IE (ICP4) protein. It physically binds to the viral proteins ICP0, ICP22, and ICP27, as well as several cellular transcription factors [20-23]. Although none of these interaction sites include the sequence of the identified epitope, a mutation leading to an escape variant might abolish binding or functional folding of the ICP4 protein, which in turn might interrupt the viral replication cycle. However, the EHV-1 closely related virus EHV-4 encodes for a substitution of aspartic acid with glutamic acid in the identified CTL epitope (REGARFGEL). The ICP4 proteins of EHV-1 and EHV-4 share a homology of 89 % on the amino acid level, which suggests that mutations in the epitope can be tolerated provided no complementary mutations arose. On the other hand, it could imply that the Eqca-1*00101 allele, while common (allele frequency estimated at 13.4 %), does not contribute substantially to the control of EHV-1 field infections, and therefore does not exert a strong selection pressure. That actually might not be necessary as the infection could be controlled by CTL activity directed against other class I loci, or the Eqca-1*00101 allele confers protection against other important equine infections.

It is tempting to speculate on a potential immunodominance of the ICP4 epitope. However, that is only true for the Eqca-1*00101 allele so far. Whether the majority of the combined CTL responses of all alleles expressed in all haplotypes is directed against this epitope remains to be investigated. However, the fact that stimulation of IFN-γ^+ T cells adjoined by a reduction of cell-associated viraemia could be achieved in Eqca-1*00101$^+$ horses after vacci-

nation with a NYVAC-IE construct suggests [82, 213], paradoxically, that a narrow repertoire might be beneficial with respect to control of, or protection against, EHV-1 infection, perhaps allowing a more focused response, or avoiding immune diversion or decoy responses, which have been described in the case of other herpesviruses [220]. A narrow epitope repertoire might also be beneficial in context of immunity towards other infections. In patients seropositive for HCMV, for example, around 10 % of the memory $CD4^+$ and $CD8^+$ T cell pool is HCMV-specific, significantly reducing the T cell repertoire, which is thought to contribute to immune senescence [221-223]. A limited but effective T cell response could ensure a sufficiently diverse T cell repertoire against various infectious diseases.

In summary, we determined quantitative peptide-binding motifs for four common equine MHC class I alleles to predict potential CTL epitope candidates from various pathogens of the horse. All alleles have binding motifs similar to human MHC-I alleles organized in HLA supertypes. Based on the motifs, we predicted and identified high affinity binders derived from EHV-1 for two of these alleles in *in vitro* binding studies, and found one peptide (RDGARF-GEL) from ICP4 to elicit IFN-γ secretion in PBMCs from multiply infected and Eqca-$1*00101^+$ horses in ELISpot assays. Cells from donor horses vaccinated with both inactivated and modified-live vaccines (MLV) did not respond to any of the tested high affinity binding peptides, which ties into the hypothesis that MLV-based vaccination does not sufficiently induce T cell-mediated immunity.

To further elucidate this working hypothesis, ELISpot assays were adapted to detect IFN-γ production in equine PBMCs after stimulation with replication-competent virus. PBMCs from a horse with unknown infection history responded with detectable IFN-γ secretion to virus stimulation but not to mock treatment (Figure 19). Since seroprevalence of EVH-1 in horse populations ranges between 30 – 80 %, it seems likely that this horse also came into contact with virulent EHV-1 strains in its life [31]. After these promising results it could be shown that only cells from multiply infected donor horses (G0 – G4) but not from EHV-1 naïve horses (4231, 4753) were responsive (Figures 21 and 22) excluding a potential NK cell involvement in the IFN-γ response. Cells from vaccinated horses were not responsive (Figure 23) further strengthening the working hypothesis.

The reasons for the limited efficacy in T cell induction of the MLV are not yet known but speculated upon. As shown for herpes viruses HSV-1, VZVand HCMV, $CD8^+$ and $CD4^+$ T cells play an important role in immunity, but their frequencies detectable after short-term antigenic stimulation differ significantly. While no IFN-γ or TNF-α production could be de-

tected in HSV and VZV primed PBMCs by intracellular staining, HCMV-specific T cell activity was observed. The authors attribute these differences in part to the differing sites of latency arguing that higher memory T cell frequencies are required to control reactivation from cells of the monocyte/macrophage lineage as site of HCMV latency than from cells of sensory ganglia for HSV and VZV [224]. Since EHV-1 like HCMV also latently infects lymphocytes and monocytes, a T cell boost after reactivation of virulent strains could be a consequence. For the vaccine strain RacH, however, no latent infections have been described, so that immunity might wane more quickly without frequent reactivations.

In addition, EHV-1 has some MHC-I-downregulating capacities. In the vaccine strain, however, one of the proteins responsible encoded by UL56 (ORF1) is deleted resulting in higher MHC-I density on the cell surface [225]. Intuitively, that would suggest enhanced T cell priming and, hence, better stimulation of T cell mediated immunity. These assumptions are contrary to actual observations. However, MHC-I downregulation seems to be locus and allele specific [172], which would alter the MHC-I composition on the surface of an infected cell. As a consequence, priming after field-strain infections could result in a different TCR repertoire than after vaccination with RacH-based MLV. Epitopes evoking a T cell response in infections with the vaccine strain but presented by downregulation-sensitive MHC-I could be underrepresented under field conditions. A TCR repertoire primed by vaccination would be less efficacious or even completely useless. On the other hand, priming of T cells with immunodominant epitopes identified in multiply infected horses (like RDGARFGEL) and presented on downregulation-insensitive MHC-I could be less efficient after vaccination as the presenting MHC-I is present in lower relative frequencies. Although unlikely, these effects could contribute to the low CTL induction efficiency in horses after vaccination with MLV at least for the haplotype A3.

Another explanation, of course, could be reduced replication and antigen production of the vaccine strain leading to reduced availablitiy of epitope-containing proteins.

Interestingly, T cell frequencies after whole virus or peptide stimulation were comparable (Figures 21 and 22) suggesting a strong immunodominance of the RDGARFGEL epitope even in ELA-A heterozygous horses. However, peptide stimulation involves continously high concentration (10 µg/mL) of peptide in the cell medium probably resulting in MHC-I saturation on more or less any cell present, which might facilitate activation and detection of clonal T cells [226]. *In vitro* infections done in this assay, however, used an MOI of 1, which results in infection of only 63 % of the cell population provided all cells are equally susceptible. Considering the heterogeneity of PBMCs and the cellular tropism of EHV-1, an even lower

infection rate can be assumed restricting the number of MHC-peptide presenting cells and thusly availability for T cell recognition. Epitope-specific responses might be overestimated in peptide stimulation. Furthermore, infected cells present various epitopes on several different MHC-I possibly activating different T cell clones in close proximity resulting in one larger spot rather than in several distinct spots on the membrane, which could underestimate T cell frequencies [227]. Additionally, the contribution of CD4$^+$ T cells to the IFN-γ response in this assay is not known. Further characterization of the T cell response after *in vitro* infection with this and alternative approaches is clearly needed if a reliable correlate of IFN-γ$^+$ T cell frequencies with protection wants to be established.

However, the clear absence of responsive cells in all vaccinated horses indicates a diagnostic and scientific potential of the modified ELISpot.

These preliminary results are encouraging but need further evaluation. In future experiments, it should be possible to correlate protection from viraemia and consequently abortion and EHM with a certain frequency of virus-specific, IFN-γ-producing T cells with this assay, and to evaluate efficacy of T cell induction by vaccines.

In conclusion, several experimental findings and the current model of EHV-1 pathogenesis suggest a predominantly T cell-mediated immunity against EHV-1 and an insufficient induction of this response by commercial vaccines. Our results strongly support this hypothesis. However, robust assays to correlate T cell activity with protection are absent. The modified ELISpot assay presented in this thesis could close that gap and prove to become a valuable tool in research and diagnostics.

8. Outlook

The immune system plays a central role in the defense against infections. A potent activation of the adaptive immune response is a necessity not only for clearing the acute infection but also in achieving immunity against re-infection. Defenses against viruses, and in particular herpes viruses, rely on MHC-restricted cytopathic and noncytopathic actions of T lymphocytes [228]. Attenuated virus strains have proven to be valuable tools for inducing protective B and T cell-mediated immunity, and passaging virulent viruses in animals or on cell cultures to derive attenuated strains has been the gold standard in the development of live vaccines for decades [3]. However, as seen in the case of EHV-1, these classical approaches can fail to induce protective immunity in genetically diverse populations. Vaccine design of the future will have to take into account the genetic and immunological background of vaccinees.

With increasing knowledge of MHC specificities and T cell epitopes, a new generation of peptide vaccines that cover a broad range of MHC haplotypes is conceivable. To better understand MHC motifs and repertoire sizes in horses, characterization of common MHC-I alleles from various haplotypes has to be continued. It is also notable that equine immune responses to EHV-1 are particularly short-lived, but there is very little information on whether there are differences in immune responses that would correlate with the ELA haplotype. Methods to approach this question have been introduced and established for equine MHC class I molecules in this work, but also MHC-II specificities should be addressed in the near future.

The current problem in peptide-based vaccination approaches of efficiently targeting peptides to MHC molecules for presentation could be tackled in different ways. Artificial sugar beads containing synthetic proteins combining several immunodominant epitopes could be routed through the endosomal pathway for cross-presentation on MHC-I [229, 230]. However, the class I antigen processing pathway would be ignored in this approach. Alternatively, virus-like particles as antigen carriers have been improved in the recent past. Particles carrying fusion proteins with the translocation motif (TLM) of hepatitis B virus (HBV) at the surface are cell-permeable and deliver their cargo directly into the cytoplasm of a DC for intracellular processing [231-233]. To date, there are no peptide vaccines commercially available but increasing understandings of MHC specificitites and antigen delivery systems will help efficiently target cell-mediated immunity.

For detection of virus-specific T cell frequencies after EHV-1 infections in the horse, the modified ELISpot assay needs to be evaluated in further experiments involving more multiply infected, vaccinated, and control animals. Classical approaches to enumerate virus-specific T cell frequencies (e.g. Cr^{51}-release assays) could confirm the results yielded by ELISpot. If consistency and reproducibility are given, subsequent challenge studies can hopefully help define protective T cell frequencies against EHV-1 infections. Depletion experiments could shed more light on the contributions of $CD8^+$ and $CD4^+$ T cells to the IFN-γ response.

9. Summary

Infections of horses with the equine herpes virus type 1 (EHV-1) can result in rhinopneumonitis, abortion in pregnant mares, and sometimes fatal neurological disorder. The virus primarily replicates in the epithelium of the upper respiratory tract, and subsequently spreads via leukocyte-associated viraemia to secondary sites of infection (endothelia of blood and lymphatic vessels). Thusly, the infection is thought to be mainly controlled by the action of virus-specific cytotoxic T lymphocytes (CTL) as precursor CTL frequencies clearly correlate with protection. High neutralizing antibody titers can be induced by several commercial vaccines, but complete protection from infection and disease is only achieved in horses multiply infected with virulent strains.

CTL recognize an infected cell by binding with their polymorphic T cell receptors (TCR) to a complex of pathogen-derived peptide and surface major histocompatibility complex class I molecules (MHC-I). MHC-I molecules bind peptides of intracellular origin according to their binding motifs mediated by polymorphic residues in the peptide-binding groove. These binding motifs can be summarized in terms of primary and secondary anchors in the peptide, and used to predict putative CTL epitopes from a pathogen.

In this study, we identified the MHC-I binding motifs for four common equine MHC-I alleles, namely Eqca-1*00101, Eqca-N*00101, Eqca-16*00101 and Eqca-1*00201 by eluting and sequencing endogenous peptide binders, and by *in vitro* binding approaches using positional scanning combinatorial libraries (PSCL). We found these binding motifs to be similar to those of known human MHC-I alleles organized in HLA supertypes. The peptide binding repertoires of Eqca-1*00101 and Eqca-N*00101 with 0.2 % and 0.5 %, respectively, are rather limited compared to 3 % on average for most human and mouse class I alleles.

Based on the motifs of the first two alleles, we predicted and identified several high affinity peptides derived from EHV-1, and found peptide RDGARFGEL restricted by Eqca-1*00101 to be immunogenic in *ex vivo* ELISpot assays.

In addition, ELISpot assays were adapted to detect virus-specific T cell activity in peripheral blood mononuclear cells (PBMC) after stimulation with replication-competent virus. This method was used to evaluate induction of T cell activity after multiple infections with virulent EHV-1 strains or vaccination with attenuated virus. Detectable T cell frequencies were found

only in multiply infected but not in vaccinated horses. The modified ELISpot assay can hopefully provide a reliable tool to correlate virus-specific T cell frequencies with protective immunity against EHV-1-associated diseases of the horse.

10. Zusammenfassung

Infektionen des Pferds mit dem equinen Herpesvirus Typ 1 (EHV-1) führen zu Rhinopneu-monitis, Aborten in trächtigen Stuten und neurologischen Erkrankungen. Das Virus repliziert primär in Epithelzellen der oberen Atemwege, bevor es via Leukozyten-assoziierter Virämie zu sekundären Infektionsorten (Endothel von Blut- und Lymphgefäßen) transportiert wird. Frequenzen virusspezifischer zytotoxischer T Lymphozyten (CTL) korrelieren eindeutig mit einem Immunschutz, weswegen davon ausgegangen wird, dass die Infektion größtenteils durch CTL-Aktivität kontrolliert wird. Kommerzielle Impfstoffe können hohe virusneutrali-sierende Antikörpertiter induzieren, ein kompletter Schutz vor Infektion und Erkrankung kann allerdings nur in Pferden beobachtet werden, die mehrmals mit virulenten Stämmen infiziert wurden.

CTL erkennen infizierte Zellen, indem sie mit ihrem polymorphen T-Zellrezeptor (TCR) ei-nen Komplex aus Peptiden pathogenen Ursprungs und Haupthistokompatibilitätskomplex Klasse I-Molekülen (MHC-I) auf der Oberfläche der Zelle binden. MHC-I-Moleküle assoziie-ren mit Peptiden intrazellulären Ursprungs entsprechend ihres Bindemotivs, das durch poly-morphe Aminosäurereste in der Bindespalte für Peptide bestimmt wird. Diese Bindemotive können als primäre und sekundäre Ankerreste im Peptid zusammengefasst und zur Vorhersa-ge vermeintlicher CTL-Epitope genutzt werden.

In dieser Arbeit wurden die Bindemotive für vier häufige, equine MHC-I-Allele namens Eq-ca-1*00101, Eqca-N*00101, Eqca-16*00101 und Eqca-1*00201 durch Eluierung und Se-quenzierung endogener Peptidliganden sowie *in vitro* Bindungsstudien mit Hilfe von positio-nal scanning combinatorial libraries (PSCL) identifiziert. Die Bindemotive ähneln denen be-kannter humaner MHC-I-Allele, die entsprechend ihrer Spezifitäten in Supertypen zusam-mengefasst werden können. Die Peptidrepertoires von Eqca-1*00101 und Eqca-N*00101 sind mit 0,2 % bzw. 0,6 % eher limitiert verglichen mit durchschnittlich 3 % für die meisten hu-manen und murinen Klasse I-Allele.

Basierend auf den Motiven für die ersten zwei Allele wurden mehrere, hoch affine Bindepep-tide aus dem EHV-1-Proteom vorhergesagt und identifiziert. Das Peptid RDGARFGEL wird von Eqca-1*00101 präsentiert und ist in *ex vivo* ELISpot Assays immunogen.

Zusätzlich wurde der ELISpot Assay adaptiert, um virusspezifische T-Zellaktivität in mono-nuklearen Zellen aus peripherem Blut (PBMC) nach Stimulierung mit replikationskompeten-tem Virus zu detektieren. Diese Methode wurde angewandt, um die Induzierung der T-Zellantwort nach multiplen Infektionen mit virulenten EHV-1 Stämmen oder Immunisierung mit attenuiertem Virus zu evaluieren. T-Zellfrequenzen waren nur in infizierten aber nicht in geimpften Tieren nachweisbar. Der modifizierte ELISpot Assay wird hoffentlich ein verläss-liches System darstellen, mit dem virusspezifische T-Zellfrequenzen mit protektiver Immuni-tät gegen EHV-1-assoziierte Erkrankungen des Pferdes korreliert werden können.

11. References

1.   Murphy, K., et al., *Janeway's immunobiology*. 8th ed2012, New York: Garland Science. xix, 868 p.
2.   Davison, A.J., et al., *The order Herpesvirales.* Arch Virol, 2009. **154**(1): p. 171-7.
3.   Fields, B.N., et al., *Fields virology*, 2002, Lippincott Williams & Wilkins,: Philadelphia, PA. p. 1 computer laser optical disc.
4.   Pellett, P.E., *Trunkloads of viruses.* J Virol, 2014. **88**(23): p. 13520-2.
5.   Boehmer, P.E., et al., *Association of origin binding protein and single strand DNA-binding protein, ICP8, during herpes simplex virus type 1 DNA replication in vivo.* J Biol Chem, 1994. **269**(46): p. 29329-34.
6.   Browning, G.F. and M.J. Studdert, *Genomic heterogeneity of equine betaherpesviruses.* J Gen Virol, 1987. **68 (Pt 5)**: p. 1441-7.
7.   Browning, G.F., N. Ficorilli, and M.J. Studdert, *Asinine herpesvirus genomes: comparison with those of the equine herpesviruses.* Arch Virol, 1988. **101**(3-4): p. 183-90.
8.   Maramorsch, K., F.A. Murphy, and A.J. Shatkin, *Advances in virus research. Volume 55*, in *Advances in virus research v 55*2000, Academic Press,: New York. p. 1 online resource (x, 638 p.).
9.   Genbank.
10.   Telford, E.A., et al., *The DNA sequence of equine herpesvirus-1.* Virology, 1992. **189**(1): p. 304-16.
11.   Telford, E.A., et al., *The DNA sequence of equine herpesvirus-4.* J Gen Virol, 1998. **79 (Pt 5)**: p. 1197-203.
12.   Azab, W., et al., *Glycoprotein C of equine herpesvirus 4 plays a role in viral binding to cell surface heparan sulfate.* Virus Res, 2010. **151**(1): p. 1-9.
13.   Osterrieder, N., *Construction and characterization of an equine herpesvirus 1 glycoprotein C negative mutant.* Virus Res, 1999. **59**(2): p. 165-77.
14.   Spear, P.G., *Herpes simplex virus: receptors and ligands for cell entry.* Cell Microbiol, 2004. **6**(5): p. 401-10.
15.   Kurtz, B.M., et al., *Equus caballus major histocompatibility complex class I is an entry receptor for equine herpesvirus type 1.* J Virol, 2010. **84**(18): p. 9027-34.
16.   Sasaki, M., et al., *Equine major histocompatibility complex class I molecules act as entry receptors that bind to equine herpesvirus-1 glycoprotein D.* Genes Cells, 2011. **16**(4): p. 343-57.
17.   Sasaki, M., et al., *Single amino acid residue in the A2 domain of major histocompatibility complex class I is involved in the efficiency of equine herpesvirus-1 entry.* J Biol Chem, 2011. **286**(45): p. 39370-8.
18.   Azab, W., M.J. Lehmann, and N. Osterrieder, *Glycoprotein H and alpha4beta1 integrins determine the entry pathway of alphaherpesviruses.* J Virol, 2013. **87**(10): p. 5937-48.
19.   Sodeik, B., M.W. Ebersold, and A. Helenius, *Microtubule-mediated transport of incoming herpes simplex virus 1 capsids to the nucleus.* J Cell Biol, 1997. **136**(5): p. 1007-21.
20.   Albrecht, R.A., et al., *Direct interaction of TFIIB and the IE protein of equine herpesvirus 1 is required for maximal trans-activation function.* Virology, 2003. **316**(2): p. 302-12.
21.   Derbigny, W.A., et al., *The EICP22 protein of equine herpesvirus 1 physically interacts with the immediate-early protein and with itself to form dimers and higher-order complexes.* J Virol, 2000. **74**(3): p. 1425-35.
22.   Kim, S.K., et al., *Interaction of the equine herpesvirus 1 EICP0 protein with the immediate-early (IE) protein, TFIIB, and TBP may mediate the antagonism between the IE and EICP0 proteins.* J Virol, 2003. **77**(4): p. 2675-85.

23.   Zhao, Y., et al., *Regulatory function of the equine herpesvirus 1 ICP27 gene product.* J Virol, 1995. **69**(5): p. 2786-93.

24.   Taddeo, B. and B. Roizman, *The virion host shutoff protein (UL41) of herpes simplex virus 1 is an endoribonuclease with a substrate specificity similar to that of RNase A.* J Virol, 2006. **80**(18): p. 9341-5.

25.   Boehmer, P.E. and I.R. Lehman, *Herpes simplex virus DNA replication.* Annu Rev Biochem, 1997. **66**: p. 347-84.

26.   Boehmer, P.E. and A.V. Nimonkar, *Herpes virus replication.* IUBMB Life, 2003. **55**(1): p. 13-22.

27.   Morissette, G. and L. Flamand, *Herpesviruses and chromosomal integration.* J Virol, 2010. **84**(23): p. 12100-9.

28.   Smith, D.J., et al., *In vitro reactivation of latent equid herpesvirus-1 from CD5+/CD8+ leukocytes indirectly by IL-2 or chorionic gonadotrophin.* J Gen Virol, 1998. **79 (Pt 12)**: p. 2997-3004.

29.   Slater, J.D., et al., *The trigeminal ganglion is a location for equine herpesvirus 1 latency and reactivation in the horse.* J Gen Virol, 1994. **75 (Pt 8)**: p. 2007-16.

30.   Allen, G.P. and J.T. Bryans, *Molecular epizootiology, pathogenesis, and prophylaxis of equine herpesvirus-1 infections.* Prog Vet Microbiol Immunol, 1986. **2**: p. 78-144.

31.   Patel, J.R. and J. Heldens, *Equine herpesviruses 1 (EHV-1) and 4 (EHV-4)--epidemiology, disease and immunoprophylaxis: a brief review.* Vet J, 2005. **170**(1): p. 14-23.

32.   Vandekerckhove, A.P., et al., *Equine alphaherpesviruses (EHV-1 and EHV-4) differ in their efficiency to infect mononuclear cells during early steps of infection in nasal mucosal explants.* Vet Microbiol, 2011. **152**(1-2): p. 21-8.

33.   Baghi, H.B. and H.J. Nauwynck, *Impact of equine herpesvirus type 1 (EHV-1) infection on the migration of monocytic cells through equine nasal mucosa.* Comparative immunology, microbiology and infectious diseases, 2014. **37**(5): p. 321-329.

34.   Edington, N., B. Smyth, and L. Griffiths, *The role of endothelial cell infection in the endometrium, placenta and foetus of equid herpesvirus 1 (EHV-1) abortions.* J Comp Pathol, 1991. **104**(4): p. 379-87.

35.   Edington, N., C.G. Bridges, and J.R. Patel, *Endothelial cell infection and thrombosis in paralysis caused by equid herpesvirus-1: equine stroke.* Arch Virol, 1986. **90**(1-2): p. 111-24.

36.   Smith, K.C., et al., *Abortion of virologically negative foetuses following experimental challenge of pregnant pony mares with equid herpesvirus 1.* Equine Vet J, 1992. **24**(4): p. 256-9.

37.   Smith, K.C., et al., *An immunohistological study of the uterus of mares following experimental infection by equid herpesvirus 1.* Equine Vet J, 1993. **25**(1): p. 36-40.

38.   Patel, J.R., N. Edington, and J.A. Mumford, *Variation in cellular tropism between isolates of equine herpesvirus-1 in foals.* Arch Virol, 1982. **74**(1): p. 41-51.

39.   Kydd, J.H., et al., *Distribution of equid herpesvirus-1 (EHV-1) in respiratory tract associated lymphoid tissue: implications for cellular immunity.* Equine Vet J, 1994. **26**(6): p. 470-3.

40.   Kydd, J.H., et al., *Distribution of equid herpesvirus-1 (EHV-1) in the respiratory tract of ponies: implications for vaccination strategies.* Equine Vet J, 1994. **26**(6): p. 466-9.

41.   Scott, J.C., S.K. Dutta, and A.C. Myrup, *In vivo harboring of equine herpesvirus-1 in leukocyte populations and subpopulations and their quantitation from experimentally infected ponies.* Am J Vet Res, 1983. **44**(7): p. 1344-8.

42.   Allen, G.P. and C.C. Breathnach, *Quantification by real-time PCR of the magnitude and duration of leucocyte-associated viraemia in horses infected with neuropathogenic vs. non-neuropathogenic strains of EHV-1.* Equine Vet J, 2006. **38**(3): p. 252-7.

43.   Smith, D., A. Hamblin, and N. Edington, *Equid herpesvirus 1 infection of endothelial cells requires activation of putative adhesion molecules: an in vitro model.* Clin Exp Immunol, 2002. **129**(2): p. 281-7.

44.   Goehring, L.S., et al., *Infection of central nervous system endothelial cells by cell-associated EHV-1.* Vet Microbiol, 2011. **148**(2-4): p. 389-95.

45.   Smith, K.C., *Herpesviral abortion in domestic animals.* Vet J, 1997. **153**(3): p. 253-68.

46.   Smith, K. C., and K. Borchers. *"A study of the pathogenesis of equid herpesvirus-1 (EHV-1) abortion by DNA in-situ hybridization."* Journal of comparative pathology 125.4 (2001): 304-310.

47.   Perkins, G.A., et al., *Investigation of the prevalence of neurologic equine herpes virus type 1 (EHV-1) in a 23-year retrospective analysis (1984-2007).* Vet Microbiol, 2009. **139**(3-4): p. 375-8.

48.   Nugent, J., et al., *Analysis of equid herpesvirus 1 strain variation reveals a point mutation of the DNA polymerase strongly associated with neuropathogenic versus nonneuropathogenic disease outbreaks.* J Virol, 2006. **80**(8): p. 4047-60.

49.   Goodman, L.B., et al., *A point mutation in a herpesvirus polymerase determines neuropathogenicity.* PLoS Pathog, 2007. **3**(11): p. e160.

50.   Van de Walle, G.R., et al., *A single-nucleotide polymorphism in a herpesvirus DNA polymerase is sufficient to cause lethal neurological disease.* J Infect Dis, 2009. **200**(1): p. 20-5.

51.   van der Meulen, K.M., et al., *Immune escape of equine herpesvirus 1 and other herpesviruses of veterinary importance.* Vet Immunol Immunopathol, 2006. **111**(1-2): p. 31-40.

52.   van der Meulen, K.M., H.J. Nauwynck, and M.B. Pensaert, *Absence of viral antigens on the surface of equine herpesvirus-1-infected peripheral blood mononuclear cells: a strategy to avoid complement-mediated lysis.* J Gen Virol, 2003. **84**(Pt 1): p. 93-7.

53.   Huemer, H.P., et al., *gp13 (EHV-gC): a complement receptor induced by equine herpesviruses.* Virus Res, 1995. **37**(2): p. 113-26.

54.   Bryant, N.A., et al., *Glycoprotein G isoforms from some alphaherpesviruses function as broad-spectrum chemokine binding proteins.* EMBO J, 2003. **22**(4): p. 833-46.

55.   Sheoran, A.S., et al., *Serum and mucosal antibody isotype responses to M-like protein (SeM) of Streptococcus equi in convalescent and vaccinated horses.* Vet Immunol Immunopathol, 1997. **59**(3-4): p. 239-51.

56.   Hannant, D., R. Easeman, and J. Mumford. *Equine mucosal immune system: intranasal vaccination with inactivated equine influenza virus protects from infection*. in *Proc. 8th int. Conf. equine inf. Dis*. 1999.

57.   Breathnach, C.C., et al., *The mucosal humoral immune response of the horse to infective challenge and vaccination with equine herpesvirus-1 antigens.* Equine Vet J, 2001. **33**(7): p. 651-7.

58.   Hannant, D., et al., *Responses of ponies to equid herpesvirus-1 ISCOM vaccination and challenge with virus of the homologous strain.* Res Vet Sci, 1993. **54**(3): p. 299-305.

59.   Kydd, J.H., H.G. Townsend, and D. Hannant, *The equine immune response to equine herpesvirus-1: the virus and its vaccines.* Vet Immunol Immunopathol, 2006. **111**(1-2): p. 15-30.

60.   Koppers-Lalic, D., et al., *Varicellovirus UL 49.5 proteins differentially affect the function of the transporter associated with antigen processing, TAP.* PLoS Pathog, 2008. **4**(5): p. e1000080.

61.   Huang, T., et al., *Major histocompatibility complex class I downregulation induced by equine herpesvirus type 1 pUL56 is through dynamin-dependent endocytosis.* J Virol, 2014. **88**(21): p. 12802-15.

62.   Heldens, J.G., et al., *Clinical and virological evaluation of the efficacy of an inactivated EHV1 and EHV4 whole virus vaccine (Duvaxyn EHV1,4). Vaccination/challenge experiments in foals and pregnant mares.* Vaccine, 2001. **19**(30): p. 4307-17.

63.   Lunn, D.P., et al., *Equine herpesvirus-1 consensus statement.* J Vet Intern Med, 2009. **23**(3): p. 450-61.

64.   O'Neill, T., et al., *Determination of equid herpesvirus 1-specific, CD8+, cytotoxic T lymphocyte precursor frequencies in ponies.* Vet Immunol Immunopathol, 1999. **70**(1-2): p. 43-54.

65.   Kydd, J.H., E. Wattrang, and D. Hannant, *Pre-infection frequencies of equine herpesvirus-1 specific, cytotoxic T lymphocytes correlate with protection against abortion following*

experimental infection of pregnant mares. Vet Immunol Immunopathol, 2003. **96**(3-4): p. 207-17.

66. Tewari, D., et al., *Characterization of immune responses to baculovirus-expressed equine herpesvirus type 1 glycoproteins D and H in a murine model.* J Gen Virol, 1994. **75 (Pt 7)**: p. 1735-41.

67. Paillot, R., et al., *Equine interferon gamma synthesis in lymphocytes after in vivo infection and in vitro stimulation with EHV-1.* Vaccine, 2005. **23**(36): p. 4541-51.

68. Paillot, R., et al., *Frequency and phenotype of EHV-1 specific, IFN-gamma synthesising lymphocytes in ponies: the effects of age, pregnancy and infection.* Dev Comp Immunol, 2007. **31**(2): p. 202-14.

69. Osterrieder, N., et al., *A touchdown PCR for the differentiation of equine herpesvirus type 1 (EHV-1) field strains from the modified live vaccine strain RacH.* J Virol Methods, 1994. **50**(1): p. 129-136.

70. Soboll Hussey, G., et al., *Evaluation of immune responses following infection of ponies with an EHV-1 ORF1/2 deletion mutant.* Vet Res, 2011. **42**: p. 23.

71. Goodman, L.B., et al., *Comparison of the efficacy of inactivated combination and modified-live virus vaccines against challenge infection with neuropathogenic equine herpesvirus type 1 (EHV-1).* Vaccine, 2006. **24**(17): p. 3636-45.

72. Goehring, L.S., et al., *Control of EHV-1 viremia and nasal shedding by commercial vaccines.* Vaccine, 2010. **28**(32): p. 5203-11.

73. Bresgen, C., et al., *Serological responses and clinical outcome after vaccination of mares and foals with equine herpesvirus type 1 and 4 (EHV-1 and EHV-4) vaccines.* Vet Microbiol, 2012. **160**(1-2): p. 9-16.

74. McGuire, T.C., et al., *Equine infectious anaemia virus proteins with epitopes most frequently recognized by cytotoxic T lymphocytes from infected horses.* J Gen Virol, 2000. **81**(Pt 11): p. 2735-9.

75. Ridgely, S.L. and T.C. McGuire, *Lipopeptide stimulation of MHC class I-restricted memory cytotoxic T lymphocytes from equine infectious anemia virus-infected horses.* Vaccine, 2002. **20**(13-14): p. 1809-19.

76. Zhang, W., S.M. Lonning, and T.C. McGuire, *Gag protein epitopes recognized by ELA-A-restricted cytotoxic T lymphocytes from horses with long-term equine infectious anemia virus infection.* J Virol, 1998. **72**(12): p. 9612-20.

77. Soboll, G., et al., *Identification of equine herpesvirus-1 antigens recognized by cytotoxic T lymphocytes.* J Gen Virol, 2003. **84**(Pt 10): p. 2625-34.

78. Kydd, J.H., et al., *A molecular approach to the identification of cytotoxic T-lymphocyte epitopes within equine herpesvirus 1.* J Gen Virol, 2006. **87**(Pt 9): p. 2507-15.

79. Kydd, J.H., et al., *Immediate-early protein of equid herpesvirus type 1 as a target for cytotoxic T-lymphocytes in the Thoroughbred horse.* J Gen Virol, 2014. **95**(Pt 8): p. 1783-9.

80. Soboll, G., et al., *Antibody and cellular immune responses following DNA vaccination and EHV-1 infection of ponies.* Vet Immunol Immunopathol, 2006. **111**(1-2): p. 81-95.

81. Paillot, R., et al., *Antibody and IFN-gamma responses induced by a recombinant canarypox vaccine and challenge infection with equine influenza virus.* Vet Immunol Immunopathol, 2006. **112**(3-4): p. 225-33.

82. Soboll, G., et al., *Vaccination of ponies with the IE gene of EHV-1 in a recombinant modified live vaccinia vector protects against clinical and virological disease.* Vet Immunol Immunopathol, 2010. **135**(1-2): p. 108-17.

83. Little, C.C. and E.E. Tyzzer, *Further experimental studies on the inheritance of susceptibility to a Transplantable tumor, Carcinoma (J. W. A.) of the Japanese waltzing Mouse.* J Med Res, 1916. **33**(3): p. 393-453.

84. Little, C.C., *A Possible Mendelian Explanation for a Type of Inheritance Apparently Non-Mendelian in Nature.* Science, 1914. **40**(1042): p. 904-6.

85. Gorer, P.A., *The genetic and antigenic basis of tumour transplantation.* The Journal of Pathology and Bacteriology, 1937. **44**(3): p. 691-697.

86. Gorer, P., *The antigenic basis of tumour transplantation.* The Journal of Pathology and Bacteriology, 1938. **47**(2): p. 231-252.

87. Zinkernagel, R.M. and P.C. Doherty, *H-2 compatability requirement for T-cell-mediated lysis of target cells infected with lymphocytic choriomeningitis virus. Different cytotoxic T-cell specificities are associated with structures coded for in H-2K or H-2D.* J Exp Med, 1975. **141**(6): p. 1427-1436.

88. Danchin, E., et al., *The major histocompatibility complex origin.* Immunol Rev, 2004. **198**(1): p. 216-232.

89. The, M., *Complete sequence and gene map of a human major histocompatibility complex.* Nature, 1999. **401**(6756): p. 921-923.

90. Parham, P., et al., *Nature of polymorphism in HLA-A, -B, and -C molecules.* Proc Natl Acad Sci U S A, 1988. **85**(11): p. 4005-9.

91. Braud, V.M., D.S. Allan, and A.J. McMichael, *Functions of nonclassical MHC and non-MHC-encoded class I molecules.* Curr Opin Immunol, 1999. **11**(1): p. 100-8.

92. Carpenter, S., et al., *Molecular and functional characterization of genes encoding horse MHC class I antigens.* Immunogenetics, 2001. **53**(9): p. 802-9.

93. Vaiman, M., P. Chardon, and D. Cohen, *DNA polymorphism in the major histocompatibility complex of man and various farm animals.* Anim Genet, 1986. **17**(2): p. 113-33.

94. Tseng, C.T., et al., *Identification of equine major histocompatibility complex haplotypes using polymorphic microsatellites.* Anim Genet, 2010. **41 Suppl 2**: p. 150-3.

95. Tallmadge, R.L., et al., *Analysis of MHC class I genes across horse MHC haplotypes.* Immunogenetics, 2010. **62**(3): p. 159-72.

96. Kydd, J., et al., *Report of the first international workshop on equine leucocyte antigens, Cambridge, UK, July 1991.* Vet Immunol Immunopathol, 1994. **42**(1): p. 3-60.

97. Azab, W., et al., *Equid herpesvirus type 4 uses a restricted set of equine major histocompatibility complex class I proteins as entry receptors.* J Gen Virol, 2014. **95**(Pt 7): p. 1554-63.

98. Tallmadge, R.L., T.L. Lear, and D.F. Antczak, *Genomic characterization of MHC class I genes of the horse.* Immunogenetics, 2005. **57**(10): p. 763-74.

99. Antczak, D.F., et al., *Joint report of the Third International Workshop on Lymphocyte Alloantigens of the Horse, Kennett Square, Pennsylvania, 25-27 April 1984.* Anim Genet, 1986. **17**(4): p. 363-73.

100. Warmuth, V., et al., *Reconstructing the origin and spread of horse domestication in the Eurasian steppe.* Proc Natl Acad Sci U S A, 2012. **109**(21): p. 8202-6.

101. Lindgren, G., et al., *Limited number of patrilines in horse domestication.* Nat Genet, 2004. **36**(4): p. 335-6.

102. Cieslak, M., et al., *Origin and history of mitochondrial DNA lineages in domestic horses.* PLoS One, 2010. **5**(12): p. e15311.

103. Ling, Y., et al., *Identification of Y chromosome genetic variations in Chinese indigenous horse breeds.* J Hered, 2010. **101**(5): p. 639-43.

104. Orlando, L., et al., *Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse.* Nature, 2013. **499**(7456): p. 74-78.

105. Niskanen, A.K., et al., *MHC variability supports dog domestication from a large number of wolves: high diversity in Asia.* Heredity (Edinb), 2013. **110**(1): p. 80-5.

106. Sun, J., D.J. Leahy, and P.B. Kavathas, *Interaction between CD8 and major histocompatibility complex (MHC) class I mediated by multiple contact surfaces that include the alpha 2 and alpha 3 domains of MHC class I.* J Exp Med, 1995. **182**(5): p. 1275-1280.

107. Guo, H.-C., et al., *Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle.* 1992.

108. Collins, E.J., D.N. Garboczi, and D.C. Wiley, *Three-dimensional structure of a peptide extending from one end of a class I MHC binding site.* Nature, 1994. **371**(6498): p. 626-9.

109. Engelhard, V.H., *Structure of peptides associated with class I and class II MHC molecules.* Annu Rev Immunol, 1994. **12**: p. 181-207.

110. Saper, M.A., P.J. Bjorkman, and D.C. Wiley, *Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution.* J Mol Biol, 1991. **219**(2): p. 277-319.

111. Latron, F., et al., *A critical role for conserved residues in the cleft of HLA-A2 in presentation of a nonapeptide to T cells.* Science, 1992. **257**(5072): p. 964-967.

112. Sidney, J., et al., *HLA class I supertypes: a revised and updated classification.* BMC Immunol, 2008. **9**: p. 1.

113. Mealey, R.H., et al., *A single amino acid difference within the alpha-2 domain of two naturally occurring equine MHC class I molecules alters the recognition of Gag and Rev epitopes by equine infectious anemia virus-specific CTL.* J Immunol, 2006. **177**(10): p. 7377-90.

114. Yao, S., et al., *Structural Illumination of Equine MHC Class I Molecules Highlights Unconventional Epitope Presentation Manner That Is Evolved in Equine Leukocyte Antigen Alleles.* J Immunol, 2016. **196**(4): p. 1943-54.

115. Rammensee, H., et al., *SYFPEITHI: database for MHC ligands and peptide motifs.* Immunogenetics, 1999. **50**(3-4): p. 213-9.

116. Khan, S., et al., *Cutting edge: neosynthesis is required for the presentation of a T cell epitope from a long-lived viral protein.* The Journal of Immunology, 2001. **167**(9): p. 4801-4804.

117. Neefjes, J., et al., *Towards a systems understanding of MHC class I and MHC class II antigen presentation.* Nat Rev Immunol, 2011. **11**(12): p. 823-36.

118. Wilkinson, K.D., *Ubiquitin-dependent signaling: the role of ubiquitination in the response of cells to their environment.* J Nutr, 1999. **129**(11): p. 1933-6.

119. Peters, J.M., W.W. Franke, and J.A. Kleinschmidt, *Distinct 19 S and 20 S subcomplexes of the 26 S proteasome and their distribution in the nucleus and the cytoplasm.* J Biol Chem, 1994. **269**(10): p. 7709-18.

120. Zwickl, P., D. Voges, and W. Baumeister, *The proteasome: a macromolecular assembly designed for controlled proteolysis.* Philos Trans R Soc Lond B Biol Sci, 1999. **354**(1389): p. 1501-11.

121. Adams, J., *The proteasome: structure, function, and role in the cell.* Cancer Treat Rev, 2003. **29 Suppl 1**: p. 3-9.

122. Kloetzel, P.M., *The proteasome and MHC class I antigen processing.* Biochim Biophys Acta, 2004. **1695**(1-3): p. 225-33.

123. Sesma, L., et al., *Species-specific differences in proteasomal processing and tapasin-mediated loading influence peptide presentation by HLA-B27 in murine cells.* J Biol Chem, 2003. **278**(47): p. 46461-72.

124. Cascio, P., et al., *26S proteasomes and immunoproteasomes produce mainly N-extended versions of an antigenic peptide.* EMBO J, 2001. **20**(10): p. 2357-66.

125. Wink, D.A., et al., *DNA deaminating ability and genotoxicity of nitric oxide and its progenitors.* Science, 1991. **254**(5034): p. 1001-3.

126. Green, S.J., et al., *Leishmania major amastigotes initiate the L-arginine-dependent killing mechanism in IFN-gamma-stimulated macrophages by induction of tumor necrosis factor-alpha.* J Immunol, 1990. **145**(12): p. 4290-7.

127. Frostell, C.G., et al., *Inhaled nitric oxide selectively reverses human hypoxic pulmonary vasoconstriction without causing systemic vasodilation.* Anesthesiology, 1993. **78**(3): p. 427-35.

128. Murata, S., et al., *Regulation of CD8+ T cell development by thymus-specific proteasomes.* Science, 2007. **316**(5829): p. 1349-53.

129. Begley, G.S., et al., *Cytoplasmic domains of the transporter associated with antigen processing and P-glycoprotein interact with subunits of the proteasome.* Mol Immunol, 2005. **42**(1): p. 137-41.

130. Murray, R.J., et al., *Identification of target antigens for the human cytotoxic T cell response to Epstein-Barr virus (EBV): implications for the immune control of EBV-positive malignancies.* J Exp Med, 1992. **176**(1): p. 157-68.

131. Levitskaya, J., et al., *Inhibition of antigen processing by the internal repeat region of the Epstein-Barr virus nuclear antigen-1.* Nature, 1995. **375**(6533): p. 685-8.

132. Kwun, H.J., et al., *Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 mimics Epstein-Barr virus EBNA1 immune evasion through central repeat domain effects on protein processing.* J Virol, 2007. **81**(15): p. 8225-35.

133. Tellam, J., et al., *Endogenous presentation of CD8+ T cell epitopes from Epstein-Barr virus-encoded nuclear antigen 1.* J Exp Med, 2004. **199**(10): p. 1421-31.

134. Vinitsky, A., et al., *The generation of MHC class I-associated peptides is only partially inhibited by proteasome inhibitors: involvement of nonproteasomal cytosolic proteases in antigen processing?* J Immunol, 1997. **159**(2): p. 554-64.

135. Reits, E., et al., *Peptide diffusion, protection, and degradation in nuclear and cytoplasmic compartments before antigen presentation by MHC class I.* Immunity, 2003. **18**(1): p. 97-108.

136. Beninga, J., K.L. Rock, and A.L. Goldberg, *Interferon-gamma can stimulate post-proteasomal trimming of the N terminus of an antigenic peptide by inducing leucine aminopeptidase.* J Biol Chem, 1998. **273**(30): p. 18734-42.

137. York, I.A., et al., *Tripeptidyl peptidase II is the major peptidase needed to trim long antigenic precursors, but is not required for most MHC class I antigen presentation.* J Immunol, 2006. **177**(3): p. 1434-43.

138. Glas, R., et al., *A proteolytic system that compensates for loss of proteasome function.* Nature, 1998. **392**(6676): p. 618-22.

139. Parcej, D. and R. Tampe, *ABC proteins in antigen translocation and viral inhibition.* Nat Chem Biol, 2010. **6**(8): p. 572-80.

140. Russ, G., et al., *Assembly, intracellular localization, and nucleotide binding properties of the human peptide transporters TAP1 and TAP2 expressed by recombinant vaccinia viruses.* J Biol Chem, 1995. **270**(36): p. 21312-8.

141. Trowsdale, J., et al., *Sequences encoded in the class II region of the MHC related to the 'ABC' superfamily of transporters.* Nature, 1990. **348**(6303): p. 741-4.

142. Powis, S.J., et al., *Effect of polymorphism of an MHC-linked transporter on the peptides assembled in a class I molecule.* Nature, 1992. **357**(6375): p. 211-5.

143. Neumann, L. and R. Tampe, *Kinetic analysis of peptide binding to the TAP transport complex: evidence for structural rearrangements induced by substrate binding.* J Mol Biol, 1999. **294**(5): p. 1203-13.

144. Yan, G., L. Shi, and D. Faustman, *Novel splicing of the human MHC-encoded peptide transporter confers unique properties.* J Immunol, 1999. **162**(2): p. 852-9.

145. Heemels, M.T., et al., *Peptide translocation by variants of the transporter associated with antigen processing.* Science, 1993. **262**(5142): p. 2059-63.

146. Gorbulev, S., R. Abele, and R. Tampe, *Allosteric crosstalk between peptide-binding, transport, and ATP hydrolysis of the ABC transporter TAP.* Proc Natl Acad Sci U S A, 2001. **98**(7): p. 3732-7.

147. Ma, W., et al., *Interferon-gamma rapidly increases peptide transporter (TAP) subunit expression and peptide transport capacity in endothelial cells.* J Biol Chem, 1997. **272**(26): p. 16585-90.

148. Tomazin, R., et al., *Stable binding of the herpes simplex virus ICP47 protein to the peptide binding site of TAP.* EMBO J, 1996. **15**(13): p. 3256-66.

149. Ahn, K., et al., *Molecular mechanism and species specificity of TAP inhibition by herpes simplex virus ICP47.* EMBO J, 1996. **15**(13): p. 3247-55.

150. Hewitt, E.W., S.S. Gupta, and P.J. Lehner, *The human cytomegalovirus gene product US6 inhibits ATP binding by TAP.* EMBO J, 2001. **20**(3): p. 387-396.

151. Koppers-Lalic, D., et al., *Varicellovirus UL49. 5 proteins differentially affect the function of the transporter associated with antigen processing, TAP.* PLoS Pathog, 2008. **4**(5): p. e1000080.

152. Verweij, M.C., et al., *The varicellovirus UL49.5 protein blocks the transporter associated with antigen processing (TAP) by inhibiting essential conformational transitions in the 6+6 transmembrane TAP core complex.* J Immunol, 2008. **181**(7): p. 4894-907.

153. Horst, D., et al., *Specific targeting of the EBV lytic phase protein BNLF2a to the transporter associated with antigen processing results in impairment of HLA class I-restricted antigen presentation.* The Journal of Immunology, 2009. **182**(4): p. 2313-2324.

154. Serwold, T., S. Gaw, and N. Shastri, *ER aminopeptidases generate a unique pool of peptides for MHC class I molecules.* Nat Immunol, 2001. **2**(7): p. 644-51.

155. Chang, S., T. Saric, and A. Goldberg. *An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides*. in *Second Charite Zeuthener See Workshop*. 2003.

156. Saveanu, L., et al., *Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum.* Nat Immunol, 2005. **6**(7): p. 689-697.

157. Hammer, G.E., et al., *The aminopeptidase ERAAP shapes the peptide repertoire displayed by major histocompatibility complex class I molecules.* Nat Immunol, 2006. **7**(1): p. 103-112.

158. Wearsch, P.A. and P. Cresswell, *The quality control of MHC class I peptide loading.* Current opinion in cell biology, 2008. **20**(6): p. 624-631.

159. Park, B., et al., *Redox regulation facilitates optimal peptide selection by MHC class I during antigen processing.* Cell, 2006. **127**(2): p. 369-382.

160. Hughes, E.A. and P. Cresswell, *The thiol oxidoreductase ERp57 is a component of the MHC class I peptide-loading complex.* Current biology, 1998. **8**(12): p. 709-713.

161. !!! INVALID CITATION !!!

162. Bangia, N., et al., *The N-terminal region of tapasin is required to stabilize the MHC class I loading complex.* Eur J Immunol, 1999. **29**(6): p. 1858-1870.

163. Barnden, M.J., et al., *Tapasin-mediated retention and optimization of peptide ligands during the assembly of class I molecules.* The Journal of Immunology, 2000. **165**(1): p. 322-330.

164. Peh, C.A., et al., *HLA-B27–restricted antigen presentation in the absence of tapasin reveals polymorphism in mechanisms of HLA class i peptide loading.* Immunity, 1998. **8**(5): p. 531-542.

165. Howarth, M., et al., *Tapasin enhances MHC class I peptide presentation according to peptide half-life.* Proc Natl Acad Sci U S A, 2004. **101**(32): p. 11737-11742.

166. Lee, S., et al., *Structural and functional dissection of human cytomegalovirus US3 in binding major histocompatibility complex class I molecules.* J Virol, 2000. **74**(23): p. 11262-11269.

167. Bennett, E.M., et al., *Cutting Edge: Adenovirus E19 Has Two Mechanisms for Affecting Class I MHC Expression1.* The Journal of Immunology, 1999. **162**(9): p. 5049-5052.

168. Gewurz, B.E., et al., *Antigen presentation subverted: Structure of the human cytomegalovirus protein US2 bound to the class I molecule HLA-A2.* Proceedings of the National Academy of Sciences, 2001. **98**(12): p. 6794-6799.

169. Boname, J.M. and P.G. Stevenson, *MHC class I ubiquitination by a viral PHD/LAP finger protein.* Immunity, 2001. **15**(4): p. 627-636.

170. Jones, T.R., et al., *Human cytomegalovirus US3 impairs transport and maturation of major histocompatibility complex class I heavy chains.* Proceedings of the National Academy of Sciences, 1996. **93**(21): p. 11327-11333.

171. Eisfeld, A.J., et al., *Downregulation of class I major histocompatibility complex surface expression by varicella-zoster virus involves open reading frame 66 protein kinase-dependent and-independent mechanisms.* J Virol, 2007. **81**(17): p. 9034-9049.

172. Rappocciolo, G., J. Birch, and S.A. Ellis, *Down-regulation of MHC class I expression by equine herpesvirus-1.* J Gen Virol, 2003. **84**(Pt 2): p. 293-300.

173. Ljunggren, H.-G., et al., *Empty MHC class I molecules come out in the cold.* 1990.

174. Moss, P., et al., *Persistent high frequency of human immunodeficiency virus-specific cytotoxic T cells in peripheral blood of infected donors.* Proceedings of the National Academy of Sciences, 1995. **92**(13): p. 5773-5777.

175. Beverley, P., *Lymphocytes: A Practical Approach*, 1988, Elsevier.

176. Altman, J.D., et al., *Phenotypic analysis of antigen-specific T lymphocytes.* Science, 1996. **274**(5284): p. 94-96.

177. Schmittel, A., U. Keilholz, and C. Scheibenbogen, *Evaluation of the interferon-gamma ELISPOT-assay for quantification of peptide specific T lymphocytes from peripheral blood.* J Immunol Methods, 1997. **210**(2): p. 167-74.

178. Sette, A., et al., *Characterization of the peptide-binding specificity of Mamu-A* 11 results in the identification of SIV-derived epitopes and interspecies cross-reactivity.* Immunogenetics, 2005. **57**(1-2): p. 53-68.

179. Ghanekar, S.A., et al., *Gamma interferon expression in CD8(+) T cells is a marker for circulating cytotoxic T lymphocytes that recognize an HLA A2-restricted epitope of human cytomegalovirus phosphoprotein pp65.* Clin Diagn Lab Immunol, 2001. **8**(3): p. 628-31.

180. Scheibenbogen, C., et al., *Quantitation of antigen-reactive T cells in peripheral blood by IFNγ-ELISPOT assay and chromium-release assay: a four-centre comparative trial.* J Immunol Methods, 2000. **244**(1): p. 81-89.

181. Hines, S.A., et al., *Clearance of virulent but not avirulent Rhodococcus equi from the lungs of adult horses is associated with intracytoplasmic gamma interferon production by CD4+ and CD8+ T lymphocytes.* Clin Diagn Lab Immunol, 2003. **10**(2): p. 208-215.

182. Komori, H., et al., *Identification of HLA-A2-or HLA-A24-restricted CTL epitopes possibly useful for glypican-3-specific immunotherapy of hepatocellular carcinoma.* Clinical Cancer Research, 2006. **12**(9): p. 2689-2697.

183. Wang, M., et al., *CTL epitopes for influenza A including the H5N1 bird flu; genome-, pathogen-, and HLA-wide screening.* Vaccine, 2007. **25**(15): p. 2823-2831.

184. Altfeld, M.A., et al., *Identification of dominant optimal HLA-B60-and HLA-B61-restricted cytotoxic T-lymphocyte (CTL) epitopes: rapid characterization of CTL responses by enzyme-linked immunospot assay.* J Virol, 2000. **74**(18): p. 8541-8549.

185. Shrestha, R., et al., *In-vitro assessment of cell-mediated immunity by demonstrating effector-t cells for diagnosis of tuberculosis in Nepalese subjects.* Nepal Med Coll J, 2011. **13**(4): p. 275-278.

186. Jin, C., et al., *An enhanced ELISPOT assay for sensitive detection of antigen-specific T cell responses to Borrelia burgdorferi.* Cells, 2013. **2**(3): p. 607-620.

187. Guidotti, L.G. and F.V. Chisari, *Noncytolytic control of viral infections by the innate and adaptive immune response.* Annu Rev Immunol, 2001. **19**: p. 65-91.

188. Breathnach, C., et al., *Equine herpesvirus-1 infection induces IFN-γ production by equine T lymphocyte subsets.* Vet Immunol Immunopathol, 2005. **103**(3): p. 207-215.

189. Kärre, K., et al., *Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy.* Nature, 1985. **319**(6055): p. 675-678.

190. Rudolph, J., D. O'CALLAGHAN, and N. Osterrieder, *Cloning of the Genomes of Equine Herpesvirus Type 1 (EHV-1) Strains KyA and RacL11 as Bacterial Artificial Chromosomes (BAC).* Journal of Veterinary Medicine, Series B, 2002. **49**(1): p. 31-36.

191. Ljunggren, H.G., et al., *Empty MHC class I molecules come out in the cold.* Nature, 1990. **346**(6283): p. 476-80.

192. Falk, K., et al., *Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules.* Nature, 1991. **351**(6324): p. 290-6.

193. Lazetic, S., et al., *Human natural killer cell receptors involved in MHC class I recognition are disulfide-linked heterodimers of CD94 and NKG2 subunits.* J Immunol, 1996. **157**(11): p. 4741-5.

194.   Fu, T.M., et al., *Dose dependence of CTL precursor frequency induced by a DNA vaccine and correlation with protective immunity against influenza virus challenge.* J Immunol, 1999. **162**(7): p. 4163-70.

195.   Antczak, D.F., et al., *Lymphocyte alloantigens of the horse. I. Serologic and genetic studies.* Tissue Antigens, 1982. **20**(3): p. 172-87.

196.   Sidney, J., et al., *Measurement of MHC/peptide interactions by gel filtration or monoclonal antibody capture.* Curr Protoc Immunol, 2013. **Chapter 18**: p. Unit 18 3.

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

198.   Rock, K.L., C. Gramm, and B. Benacerraf, *Low temperature and peptides favor the formation of class I heterodimers on RMA-S cells at the cell surface.* Proceedings of the National Academy of Sciences, 1991. **88**(10): p. 4200-4204.

199.   Tangri, S., et al., *Structural features of peptide analogs of human histocompatibility leukocyte antigen class I epitopes that are more potent and immunogenic than wild-type peptide.* J Exp Med, 2001. **194**(6): p. 833-46.

200.   Ross, P., et al., *A cell-based MHC stabilization assay for the detection of peptide binding to the canine classical class I molecule, DLA-88.* Vet Immunol Immunopathol, 2012. **150**(3-4): p. 206-12.

201.   Sette, A., et al., *The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes.* J Immunol, 1994. **153**(12): p. 5586-92.

202.   Udaka, K., et al., *An automated prediction of MHC class I-binding peptides based on positional scanning with peptide libraries.* Immunogenetics, 2000. **51**(10): p. 816-28.

203.   Udaka, K., et al., *Decrypting the structure of major histocompatibility complex class I-restricted cytotoxic T lymphocyte epitopes with complex peptide libraries.* J Exp Med, 1995. **181**(6): p. 2097-2108.

204.   Stryhn, A., et al., *Peptide binding specificity of major histocompatibility complex class I resolved into an array of apparently independent subspecificities: quantitation by peptide libraries and improved prediction of binding.* Eur J Immunol, 1996. **26**(8): p. 1911-1918.

205.   Kotturi, M.F., et al., *The CD8+ T-cell response to lymphocytic choriomeningitis virus involves the L antigen: uncovering new tricks for an old virus.* J Virol, 2007. **81**(10): p. 4928-4940.

206.   Paul, S., et al., *HLA class I alleles are associated with peptide-binding repertoires of different size, affinity, and immunogenicity.* The Journal of Immunology, 2013. **191**(12): p. 5831-5839.

207.   Assarsson, E., et al., *A quantitative analysis of the variables affecting the repertoire of T cell specificities recognized after vaccinia virus infection.* The Journal of Immunology, 2007. **178**(12): p. 7890-7901.

208.   Loffredo, J.T., et al., *The high frequency Indian rhesus macaque MHC class I molecule, Mamu-B* 01, does not appear to be involved in CD8+ T lymphocyte responses to SIVmac239.* The Journal of Immunology, 2005. **175**(9): p. 5986-5997.

209.   Newton, J., et al., *Immunity to equine influenza: relationship of vaccine-induced antibody in young Thoroughbred racehorses to protection against field infection with influenza A/equine-2 viruses (H3N8).* Equine Vet J, 2000. **32**(1): p. 65-74.

210.   Amoscato, A.A., D.A. Prenovitz, and M.T. Lotze, *Rapid extracellular degradation of synthetic class I peptides by human dendritic cells.* J Immunol, 1998. **161**(8): p. 4023-32.

211.   Song, M.-C., T. Ogishima, and A. Ito, *Importance of residues carboxyl terminal relative to the cleavage site in substrates of mitochondrial processing peptidase for their specific recognition and cleavage.* J Biochem, 1998. **124**(5): p. 1045-1049.

212.   Yeager, M. and A.L. Hughes, *Evolution of the mammalian MHC: natural selection, recombination, and convergent evolution.* Immunol Rev, 1999. **167**(1): p. 45-58.

213.   Paillot, R., et al., *Characterisation of CTL and IFN-gamma synthesis in ponies following vaccination with a NYVAC-based construct coding for EHV-1 immediate early gene, followed by challenge infection.* Vaccine, 2006. **24**(10): p. 1490-500.

214. Dutoit, V., et al., *Functional avidity of tumor antigen-specific CTL recognition directly correlates with the stability of MHC/peptide multimer binding to TCR.* The Journal of Immunology, 2002. **168**(3): p. 1167-1171.

215. Deng, Y., et al., *MHC affinity, peptide liberation, T cell repertoire, and immunodominance all contribute to the paucity of MHC class I-restricted peptides recognized by antiviral CTL.* The Journal of Immunology, 1997. **158**(4): p. 1507-1515.

216. Lund, O., et al., *Definition of supertypes for HLA molecules using clustering of specificity matrices.* Immunogenetics, 2004. **55**(12): p. 797-810.

217. Lacey, S.F., et al., *Relative dominance of HLA-B* 07 restricted CD8+ T-lymphocyte immune responses to human cytomegalovirus pp65 in persons sharing HLA-A* 02 and HLA-B* 07 alleles.* Hum Immunol, 2003. **64**(4): p. 440-452.

218. Gubler, B., et al., *Substrate selection by transporters associated with antigen processing occurs during peptide binding to TAP.* Mol Immunol, 1998. **35**(8): p. 427-33.

219. Blanchard, N. and N. Shastri, *Coping with loss of perfection in the MHC class I peptide repertoire.* Curr Opin Immunol, 2008. **20**(1): p. 82-88.

220. Gillet, L., et al., *The murine gammaherpesvirus-68 gp150 acts as an immunogenic decoy to limit virion neutralization.* PLoS One, 2007. **2**(8): p. e705.

221. Sylwester, A.W., et al., *Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects.* J Exp Med, 2005. **202**(5): p. 673-685.

222. Khan, N., et al., *Cytomegalovirus seropositivity drives the CD8 T cell repertoire toward greater clonality in healthy elderly individuals.* The Journal of Immunology, 2002. **169**(4): p. 1984-1992.

223. Pourgheysari, B., et al., *The cytomegalovirus-specific CD4+ T-cell response expands with age and markedly alters the CD4+ T-cell repertoire.* J Virol, 2007. **81**(14): p. 7759-7765.

224. Asanuma, H., et al., *Frequencies of memory T cells specific for varicella-zoster virus, herpes simplex virus, and cytomegalovirus by intracellular detection of cytokine expression.* Journal of Infectious Diseases, 2000. **181**(3): p. 859-866.

225. Hübert, P., et al., *Alterations in the Equine Herpesvirus Type-1 (EHV-1) Strain RacH During Attenuation.* Journal of Veterinary Medicine, Series B, 1996. **43**(1-10): p. 1-14.

226. Ishioka, G.Y., et al., *Failure to demonstrate long-lived MHC saturation both in vitro and in vivo. Implications for therapeutic potential of MHC-blocking peptides.* The Journal of Immunology, 1994. **152**(9): p. 4310-4319.

227. Cella, M., et al., *Ligation of CD40 on dendritic cells triggers production of high levels of interleukin-12 and enhances T cell stimulatory capacity: TT help via APC activation.* J Exp Med, 1996. **184**(2): p. 747-752.

228. Guidotti, L.G. and F.V. Chisari, *Noncytolytic control of viral infections by the innate and adaptive immune response.* Annual review of immunology, 2001. **19**(1): p. 65-91.

229. Nair-Gupta, P., et al., *TLR signals induce phagosomal MHC-I delivery from the endosomal recycling compartment to allow cross-presentation.* Cell, 2014. **158**(3): p. 506-521.

230. Blander, J.M., *Coupling Toll-like receptor signaling with phagocytosis: potentiation of antigen presentation.* Trends in immunology, 2007. **28**(1): p. 19-25.

231. Oess, S. and E. Hildt, *Novel cell permeable motif derived from the PreS2-domain of hepatitis-B virus surface antigens.* Gene Ther, 2000. **7**(9).

232. Bleifuss, E., et al., *The translocation motif of hepatitis B virus improves protein vaccination.* Cellular and molecular life sciences, 2006. **63**(5): p. 627-635.

233. Koho, T., *Virus-Like Particles for Vaccine Development.* 2015.

12. Publications

The common equine class I molecule Eqca-1*00101 (ELA-A3.1) is characterized by narrow peptide binding and T cell epitope repertoires (2015). Tobias Bergmann, Carrie Moore, John Sidney, Donald Miller, Rebecca Tallmadge, Rebecca M. Harman, Carla Oseroff, Amanda Wriston, Jeffrey Shabanowitz, Donald F. Hunt, Nikolaus Osterrieder, Bjoern Peters, Douglas F. Antczak, Alessandro Sette. Immunogenetics DOI 10.1007/s00251-015-0872-z

Peptide binding motifs of two common equine class I MHC molecules. Tobias Bergmann; Mikaela Lindvall; Erin Moore; Eugene Moore; John Sidney; Donald Miller; Rebecca L. Tallmadge; Paisley T. Myers; Stacy A. Malaker; Jeffrey Shabanowitz; Nikolaus Osterrieder; Bjoern Peters; Donald F. Hunt; Douglas F. Antczak; Alessandro Sette. Under review

13. Acknowledgements

14. Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Ich versichere, dass diese Arbeit in dieser oder anderer Form noch keiner Prüfungsbehörde vorgelegt wurde. Der Inhalt der Promotionsordung des Fachbereichs Biologie, Chemie, Pharmazie der Freien Universität Berlin vom 10. Januar und 11. Juli 2007 ist mir bekannt.

15. Curriculum vitae

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

16. Supplemental data

Supplemental Table 1: EHV-1-derived peptides tested for binding to Eqca-1*00101

Peptide	Len	Protein	GI no.	Locus	Position	IC50 nM
RDFAVGGGLL	10	nuclear egress membrane protein	50313267	YP_053071.1	7	1.1
LDLSVHPSL	9	envelope glycoprotein C	50313257	YP_053061.1	222	2.5
SDILWQASLI	10	envelope glycoprotein D	126801894	YP_053116.2	151	4.7
KSMRHIGDI	9	capsid portal protein	50313297	YP_053100.1	388	5.1
ADANHTMSI	9	single-stranded DNA-binding protein	50313272	YP_053076.1	655	6.1
SDILWQASL	9	envelope glycoprotein D	126801894	YP_053116.2	151	6.9
VDLRIQADV	9	helicase-primase primase subunit	50313248	YP_053052.1	319	7.9
RDGARFGEL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1059	8.1
RDIPQDGML	9	tegument protein UL21	50313282	YP_053086.1	140	8.8
LDMLNGGFI	9	single-stranded DNA-binding protein	50313272	YP_053076.1	723	10
VDMNCTTSV	9	envelope glycoprotein C	50313257	YP_053061.1	89	11
VDISNVVSPL	10	nuclear protein UL55	50313245	YP_053049.1	39	13
ADIRAILNM	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	307	14
RDIALEVEI	9	helicase-primase subunit	126801893	YP_053098.2	198	14
ADMERLSRL	9	capsid portal protein	50313297	YP_053100.1	557	15
RDLSKTMSL	9	deoxyribonuclease	50313291	YP_053046.1	149	17
SDLKGGLSAL	10	transcriptional regulator ICP5	50313321	YP_053124.1	1122	18
TDLPISDDL	9	ribonucleotide reductase subunit 2	50313261	YP_053065.1	45	22
TDINRSLLL	9	large tegument protein	50313265	YP_053069.1	3403	22
RDVDGTPTL	9	virion protein US2	50313309	YP_053112.1	286	25
LDLTHLTSF	9	helicase-primase primase subunit	50313248	YP_053052.1	286	27
KDVPIGIGI	9	envelope protein UL43	50313258	YP_053062.1	180	31
KDILDHLDFI	10	ribonucleotide reductase subunit 1	50313262	YP_053066.1	105	38
IDMKLIATL	9	thymidine kinase	50313280	YP_053084.1	203	47
FDTNLAVDL	9	DNA packaging tegument protein UL25	50313278	YP_053082.1	151	48
KDMSMAPLT	9	tegument protein UL37	50313264	YP_053068.1	222	50
RDLEHHSTA	9	tegument protein UL11	50313289	YP_053093.1	230	58
RDQHGFMSV	9	helicase-primase helicase subunit	50313298	YP_053101.1	766	62
RDLSVSGYM	9	nuclear egress lamina protein	50313270	YP_053074.1	225	65
SDIYDSTSF	9	nuclear egress lamina protein	50313270	YP_053074.1	6	67
SQIEKIVSDI	10	envelope glycoprotein H	50313281	YP_053085.1	433	70
KDMLPVDVI	9	small capsid protein	50313266	YP_053070.1	28	72
SDMSTLHQSV	10	helicase-primase subunit	126801893	YP_053098.2	616	75
SDINVEAEL	9	envelope glycoprotein H	50313281	YP_053085.1	440	77
RDFAVGGGL	9	nuclear egress membrane protein	50313267	YP_053071.1	7	83
RNIDKGPVI	9	tegument protein VP13/14	50313255	YP_053059.1	235	87
SDMLKVPQL	9	large tegument protein	50313265	YP_053069.1	1384	87
RDLIPSSDV	9	DNA packaging protein UL33	126801892	YP_053072.2	36	91
SDLAVLPVI	9	single-stranded DNA-binding protein	50313272	YP_053076.1	41	101

RDLVGEMEL	9	deoxyribonuclease	50313291	YP_053046.1	111	105
SDILINTLV	9	large tegument protein	50313265	YP_053069.1	2048	106
GDMARVSMEV	10	single-stranded DNA-binding protein	50313272	YP_053076.1	764	110
KDASLFLSL	9	envelope glycoprotein H	50313281	YP_053085.1	529	116
ADLMGAVSSL	10	capsid matruation protease	50313276	YP_053080.1	488	121
SDFAPDLYDF	10	transcriptional regulator ICP6	50313321	YP_053124.1	6	128
IDQSLMPSI	9	DNA packaging terminase subunit 2	50313273	YP_053077.1	589	129
TDAKVLRNI	9	tegument protein VP13/14	50313255	YP_053059.1	229	136
SDFVQVRLI	9	tegument protein UL22	50313282	YP_053086.1	246	137
VDMNLRLIV	9	envelope glycoprotein H	50313281	YP_053085.1	394	153
ADMERFRVL	9	large tegument protein	50313265	YP_053069.1	858	155
FDMPNVMPTV	10	major capsid protein	50313284	YP_053088.1	561	155
RDYTTTISTL	10	tegument protein UL37	50313264	YP_053068.1	452	161
GDMSMCALI	9	thymidine kinase	50313280	YP_053084.1	156	163
YDLQAVKDFI	10	helicase-primase primase subunit	50313248	YP_053052.1	263	164
QDVISRVDAI	10	ribonucleotide reductase subunit 1	50313262	YP_053066.1	16	181
RDLQLTRTT	9	helicase-primase primase subunit	50313248	YP_053052.1	607	189
HDVGLTIDL	9	tegument protein UL15	50313290	YP_053094.1	444	190
GDLAFTGCV	9	transcriptional regulator ICP4	50313321	YP_053124.1	1171	202
KDLRVTLAKF	10	nuclear protein UL3	50313301	YP_053104.1	96	204
RDLVERARLL	10	transcriptional regulator ICP8	50313321	YP_053124.1	722	204
SDLLAALDKI	10	large tegument protein	50313265	YP_053069.1	1243	207
TDLILMGCDI	10	tegument host shutoff protein	50313260	YP_053064.1	221	217
MDFKYSDTVI	10	tegument protein UL23	50313282	YP_053086.1	1	217
RDIVYYWEVL	10	helicase-primase helicase subunit	50313298	YP_053101.1	188	219
ADLNAYGYRL	10	tegument protein UL16	50313290	YP_053094.1	302	255
ADYNRFQSL	9	DNA replication origin-binding helicase	50313294	YP_053097.1	713	263
SDREFITYI	9	helicase-primase primase subunit	50313248	YP_053052.1	347	265
RQLSTTEVV	9	tegument protein UL51	50313249	YP_053053.1	212	273
RDLVERARL	9	transcriptional regulator ICP4	50313321	YP_053124.1	722	273
RDLAGKLYV	9	capsid portal protein	50313297	YP_053100.1	231	289
KDMRAQAEGL	10	transcriptional regulator ICP4	50313321	YP_053124.1	1312	293
SDDMLVALV	9	DNA packaging terminase subunit 10	50313286	YP_053090.1	705	312
RRAAWMHQI	9	transcriptional regulator ICP4	50313321	YP_053124.1	1075	319
HDYFRVNASV	10	envelope glycoprotein I	50313314	YP_053117.1	115	357
RDCPHFLEA	9	uracil-DNA glycosylase	50313302	YP_053105.1	287	377
HDFSVWFEDL	10	DNA packaging protein UL33	126801892	YP_053072.2	59	397
RDEHLAGSI	9	DNA packaging terminase subunit 19	50313287	YP_053091.1	425	409
SLFGQFREAL	10	large tegument protein	50313265	YP_053069.1	1007	412
CDHQRLSDV	9	multifunctional expression regulator	50313246	YP_053050.1	373	419
IDLWNVMYTL	10	tegument host shutoff protein	50313260	YP_053064.1	33	420
KILSVARGI	9	nuclear protein UL55	50313245	YP_053049.1	180	428
RDVKTENIFI	10	serine/threonine protein kinase US3	50313310	YP_053113.1	206	435
RDLLEHVQAL	10	major capsid protein	50313284	YP_053088.1	691	440
VDLLATVFV	9	DNA packaging tegument protein UL26	50313278	YP_053082.1	157	451
SDGRTHPVV	9	nuclear protein UL3	50313301	YP_053104.1	176	456
CDLSFDGEL	9	nuclear egress lamina protein	50313270	YP_053074.1	300	475

FDLRRKGAI	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	424	496
ADLRADVVTL	10	tegument protein UL37	50313264	YP_053068.1	803	502
REMIFKMHTL	10	tegument protein UL12	50313289	YP_053093.1	20	512
RDTPEASPV	9	envelope glycoprotein E	50313315	YP_053118.1	518	513
RDFKIREGLA	10	single-stranded DNA-binding protein	50313272	YP_053076.1	646	524
RDNMLEDLEI	10	tegument protein VP13/14	50313254	YP_053058.1	776	526
YDIDSVVNV	9	envelope glycoprotein B	50313274	YP_053078.1	792	570
KSNSTFVGI	9	envelope glycoprotein D	126801894	YP_053116.2	344	571
LDADQTAYI	9	DNA packaging tegument protein UL27	50313278	YP_053082.1	345	630
RDLVVVDSV	9	myristylated tegument protein CIRC	50313244	YP_053048.1	122	636
KDILDHLDF	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	105	647
PDIASTSPL	9	ubiquitin E3 ligase ICP0	50313304	YP_053107.1	486	648
KDQRPKSAL	9	transcriptional regulator ICP12	50313321	YP_053124.1	336	649
ADQLLGLDVI	10	DNA packaging terminase subunit 3	50313273	YP_053077.1	372	658
VDHVKGENI	9	DNA packaging terminase subunit 11	50313286	YP_053090.1	315	668
ADLQNFTMTL	10	envelope glycoprotein H	50313281	YP_053085.1	206	671
YDMPFVQEGL	10	tegument protein UL37	50313264	YP_053068.1	192	703
RGYHKRRRF	9	capsid portal protein	50313297	YP_053100.1	709	704
TDVEAVPKI	9	envelope protein UL43	50313258	YP_053062.1	233	707
PDPEDVKVV	9	transcriptional regulator ICP13	50313321	YP_053124.1	1084	749
MDATANAQL	9	DNA replication origin-binding helicase	50313294	YP_053097.1	236	752
ADLRADVVT	9	tegument protein UL37	50313264	YP_053068.1	803	756
KQYNALVRM	9	transcriptional regulator ICP14	50313321	YP_053124.1	475	766
RYLDGESGV	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	346	768
MDYDRVNIF	9	capsid portal protein	50313297	YP_053100.1	264	768
HDILQLRFGL	10	large tegument protein	50313265	YP_053069.1	1906	799
EDIETITAI	9	single-stranded DNA-binding protein	50313272	YP_053076.1	866	839
IDFALTADSI	10	large tegument protein	50313265	YP_053069.1	55	849
GDIVYASPF	9	envelope glycoprotein B	50313274	YP_053078.1	317	861
KDIVVTRVW	9	envelope glycoprotein B	50313274	YP_053078.1	190	869
MDQRPMVVI	9	single-stranded DNA-binding protein	50313272	YP_053076.1	982	875
KDMRAQAEG	9	transcriptional regulator ICP15	50313321	YP_053124.1	1312	887
IDNSAFVPI	9	tegument protein UL37	50313264	YP_053068.1	267	903
LDLPAFKGF	9	tegument protein UL7	50313296	YP_053099.1	100	910
ADIWSAGIV	9	serine/threonine protein kinase US4	50313310	YP_053113.1	262	916
ADQLLGLDV	9	DNA packaging terminase subunit 4	50313273	YP_053077.1	372	927
SDNPNLNAV	9	tegument protein UL51	50313249	YP_053053.1	93	937
RDFSFIDTGV	10	helicase-primase primase subunit	50313248	YP_053052.1	843	941
RDAVMRGENL	10	multifunctional expression regulator	50313246	YP_053050.1	299	988
RDTVISEHI	9	helicase-primase helicase subunit	50313298	YP_053101.1	860	990
LHLYTVPTI	9	tegument protein UL24	50313282	YP_053086.1	148	1005
RDPIISSSV	9	multifunctional expression regulator	50313246	YP_053050.1	335	1008
RGFAICTYV	9	helicase-primase subunit	126801893	YP_053098.2	493	1009
KQFHMLVTF	9	tegument protein UL6	50313288	YP_053092.1	99	1016
RSLISALYV	9	deoxyribonuclease	50313291	YP_053046.1	233	1019
YHIYMRAAL	9	transcriptional regulator ICP16	50313321	YP_053124.1	1214	1026
VDLPGGDEV	9	myristylated tegument protein CIRC	50313244	YP_053048.1	34	1050

RERHLSRAVL	10	transcriptional regulator ICP17	50313321	YP_053124.1	1234	1054
RKYDSVQQP	9	helicase-primase subunit	126801893	YP_053098.2	215	1056
RELPFRPSV	9	protein V32	50313275	YP_053079.1	103	1107
TLVALRVDI	9	DNA packaging terminase subunit 20	50313287	YP_053091.1	130	1111
PDMTPGASV	9	tegument protein VP13/14	50313254	YP_053058.1	810	1132
VDASASTGL	9	capsid matruation protease	50313276	YP_053080.1	620	1144
VDQMFVLLEI	10	tegument protein UL37	50313264	YP_053068.1	106	1158
YDHLLAVSI	9	helicase-primase primase subunit	50313248	YP_053052.1	369	1184
GLSALLAAL	9	transcriptional regulator ICP18	50313321	YP_053124.1	1127	1193
RDMLLKTRDY	10	DNA polymerase catalytic subunit	50313271	YP_053075.1	843	1208
KIITPHPSV	9	envelope glycoprotein E	50313315	YP_053118.1	384	1208
PDIEHLRSL	9	ribonucleotide reductase subunit 2	50313261	YP_053065.1	28	1223
AGLVLVGVI	9	envelope glycoprotein D	126801894	YP_053116.2	359	1230
DDLLNVVTYI	10	DNA packaging terminase subunit 12	50313286	YP_053090.1	405	1239
VDLYTEIAL	9	transactivating tegument protein VP16	126801890	YP_053057.2	102	1244
KQLQHYRETI	10	tegument protein VP13/14	50313254	YP_053058.1	832	1257
SDSESASSL	9	thymidine kinase	50313280	YP_053084.1	331	1264
RLVKLIVNM	9	capsid portal protein	50313297	YP_053100.1	379	1386
RDGLWDDPEI	10	transcriptional regulator ICP19	50313321	YP_053124.1	440	1389
RDETALSQAL	10	tegument protein VP13/14	50313254	YP_053058.1	668	1417
RDQPQHRAV	9	DNA packaging tegument protein UL28	50313278	YP_053082.1	485	1430
ADLVLVGDRL	10	major capsid protein	50313284	YP_053088.1	364	1440
RDWLAMRKAV	10	DNA polymerase catalytic subunit	50313271	YP_053075.1	780	1446
IDVDMNLRLI	10	envelope glycoprotein H	50313281	YP_053085.1	392	1457
LDAPPTQII	9	tegument protein UL37	50313264	YP_053068.1	579	1487
SEFCRVALL	9	transcriptional regulator ICP20	50313321	YP_053124.1	1375	1490
SDYMLLQEA	9	envelope glycoprotein H	50313281	YP_053085.1	383	1548
RDLGLAVERL	10	tegument protein UL37	50313264	YP_053068.1	935	1568
PDIASTSPLI	10	ubiquitin E3 ligase ICP0	50313304	YP_053107.1	486	1623
VDYAAIPWV	9	major capsid protein	50313284	YP_053088.1	761	1645
IDLWNVMYT	9	tegument host shutoff protein	50313260	YP_053064.1	33	1650
PPARDGARF	9	transcriptional regulator ICP21	50313321	YP_053124.1	1056	1655
VDQPTHTYM	9	major capsid protein	50313284	YP_053088.1	110	1667
RDIVYYWEV	9	helicase-primase helicase subunit	50313298	YP_053101.1	188	1681
KMIAAALGV	9	capsid triplex subunit 2	50313285	YP_053089.1	82	1692
RDTTSTQVL	9	tegument protein UL37	50313264	YP_053068.1	546	1761
LDADISRLI	9	tegument protein UL7	50313288	YP_053092.1	335	1771
AQLLLTEDI	9	deoxyuridine triphosphatase	50313250	YP_053054.1	260	1783
ADQLLRVLL	9	major capsid protein	50313284	YP_053088.1	179	1833
TDLHEPQTL	9	tegument host shutoff protein	50313260	YP_053064.1	267	1883
LDLKQIRVL	9	large tegument protein	50313265	YP_053069.1	3411	1946
ARLQALSTL	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	672	1969
AMLLTVADV	9	DNA packaging protein UL33	50313269	YP_053073.1	232	2001
GSYRFTIRSI	10	envelope glycoprotein B	50313274	YP_053078.1	407	2007
KDVILSLREA	10	large tegument protein	50313265	YP_053069.1	1830	2018
ADLYRFSSSL	10	virion protein US2	50313309	YP_053112.1	54	2039
RLLYLHLYV	9	transactivating tegument protein VP16	126801890	YP_053057.2	231	2045

KRMGVKVLL	9	helicase-primase subunit	126801893	YP_053098.2	248	2076
RDLRVKNRVL	10	single-stranded DNA-binding protein	50313272	YP_053076.1	775	2093
YDLQAVKDF	9	helicase-primase primase subunit	50313248	YP_053052.1	263	2113
GSLVLLSLI	9	capsid triplex subunit 2	50313285	YP_053089.1	211	2146
FDEGMLSAI	9	transactivating tegument protein VP16	126801890	YP_053057.2	34	2173
KRYFKFGKEY	10	envelope glycoprotein B	50313274	YP_053078.1	713	2262
QDEDLLGGL	9	transcriptional regulator ICP22	50313321	YP_053124.1	1099	2263
RDEVSFARI	9	envelope glycoprotein C	50313257	YP_053061.1	307	2327
RTLPPGSGEI	10	nuclear protein UL24	50313279	YP_053083.1	133	2339
EDVAVLRLI	9	tegument protein UL8	50313288	YP_053092.1	38	2364
FDREAPSSL	9	deoxyuridine triphosphatase	50313250	YP_053054.1	307	2661
ADLGLAGTV	9	serine/threonine protein kinase US5	50313310	YP_053113.1	237	2680
TDVSYYMYRI	10	envelope glycoprotein H	50313281	YP_053085.1	356	2710
RDLNAIEAAV	10	large tegument protein	50313265	YP_053069.1	661	2751
AALALVVLI	9	envelope glycoprotein C	50313257	YP_053061.1	440	2752
KDFIRTYNV	9	helicase-primase primase subunit	50313248	YP_053052.1	269	2779
SDFVQVRLIP	10	tegument protein UL25	50313282	YP_053086.1	246	2791
RDLRVKNRV	9	single-stranded DNA-binding protein	50313272	YP_053076.1	775	2939
SAMEKWNEDM	10	transactivating tegument protein VP16	126801890	YP_053057.2	85	3000
SLFVQIECI	9	tegument protein UL7	50313296	YP_053099.1	165	3026
KGGLFKVWDI	10	DNA polymerase catalytic subunit	50313271	YP_053075.1	490	3047
RDVWRRRPI	9	DNA packaging terminase subunit 21	50313287	YP_053091.1	386	3085
GIITTLSLI	9	helicase-primase primase subunit	50313248	YP_053052.1	676	3364
RALFLLSLV	9	tegument protein UL37	50313264	YP_053068.1	231	3377
EDRPVLTSI	9	envelope glycoprotein C	50313257	YP_053061.1	424	3414
MDKTSGARF	9	large tegument protein	50313265	YP_053069.1	741	3522
SQLRAIQTV	9	capsid matruation protease	50313276	YP_053080.1	502	3557
SDVASYKSI	9	large tegument protein	50313265	YP_053069.1	541	3688
RDETVRCDNI	10	ribonucleotide reductase subunit 1	50313262	YP_053066.1	321	3866
SDPRVKIFMA	10	tegument protein UL9	50313288	YP_053092.1	48	3936
TQAVFSIPSL	10	transcriptional regulator ICP23	50313321	YP_053124.1	780	4153
SDYMLLQEAI	10	envelope glycoprotein H	50313281	YP_053085.1	383	4162
RDFKIREGL	9	single-stranded DNA-binding protein	50313272	YP_053076.1	646	4236
SKFAERHKNL	10	large tegument protein	50313265	YP_053069.1	1707	4568
FDPQALATI	9	transcriptional regulator ICP24	50313321	YP_053124.1	1041	4610
TDINDPRSV	9	large tegument protein	50313265	YP_053069.1	3328	4641
IQISSAGVI	9	capsid triplex subunit 1	50313263	YP_053067.1	80	4740
KTFMTTAVV	9	DNA packaging tegument protein UL29	50313278	YP_053082.1	208	4962
SDAHPATPL	9	transactivating tegument protein VP16	126801890	YP_053057.2	355	5039
HVLDLLSVI	9	envelope glycoprotein H	50313281	YP_053085.1	597	5051
AWTVTQAVF	9	transcriptional regulator ICP4	50313321	YP_053124.1	776	5081
HDCESPSPV	9	helicase-primase subunit	126801893	YP_053098.2	684	5224
RTINGAADV	9	single-stranded DNA-binding protein	50313272	YP_053076.1	942	5311
ADQRRGVHV	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	67	5360
REMIKYMSM	9	envelope glycoprotein B	50313274	YP_053078.1	920	5487
RVGVRKREV	9	large tegument protein	50313265	YP_053069.1	918	5579
TDIDYDIDEL	10	tegument protein VP13/14	50313254	YP_053058.1	334	5844

RELGDACVL	9	transcriptional regulator ICP4	50313321	YP_053124.1	659	5847
ALLPATLTI	9	tegument protein UL51	50313249	YP_053053.1	45	5858
RFVTEVREV	9	tegument protein UL7	50313296	YP_053099.1	46	5899
SEFKKMMDL	9	membrane protein US8A	50313316	YP_053119.1	68	6338
RDIGSHGRN	9	nuclear protein UL55	50313245	YP_053049.1	16	6386
HHLQYIINTV	10	tegument protein VP13/14	50313255	YP_053059.1	323	6386
SDNETASMFI	10	membrane protein US9	50313317	YP_053120.1	165	6386
QAPRALGPM	9	transcriptional regulator ICP4	50313321	YP_053124.1	976	6386
IDILLPSDL	9	capsid triplex subunit 2	50313285	YP_053089.1	8	6410
RDFWHAACS	9	envelope glycoprotein N	50313251	YP_053055.1	44	6443
ADDPRTTEV	9	DNA packaging terminase subunit 22	50313287	YP_053091.1	139	6532
MILMHATYFI	10	DNA packaging terminase subunit 13	50313286	YP_053090.1	208	6787
HIFTGRVTL	9	envelope glycoprotein G	50313311	YP_053114.1	162	7087
RDEAGRPDSL	10	transcriptional regulator ICP4	50313321	YP_053124.1	584	7124
QDFKEKLLF	9	envelope glycoprotein L	155041723	YP_053106.2	118	7144
PEDVKVVVL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1086	7338
REFVMTAEL	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	1030	7430
SDKRLNPIF	9	deoxyribonuclease	50313291	YP_053046.1	96	7495
GDAHAEAAV	9	protein V32	50313275	YP_053079.1	12	7555
SDAGALESV	9	DNA replication origin-binding helicase	50313294	YP_053097.1	656	7563
TDYNTVLVNY	10	tegument host shutoff protein	50313260	YP_053064.1	477	7721
RDGTLLRTL	9	DNA replication origin-binding helicase	50313294	YP_053097.1	272	7762
YDLGDAVGV	9	large tegument protein	50313265	YP_053069.1	433	7878
ADFINNSGL	9	tegument protein UL17	50313290	YP_053094.1	427	8022
RDYVHSRWA	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	850	8047
HEFNNYALM	9	DNA polymerase processivity subunit	50313259	YP_053063.1	175	8195
RFFERWHMLM	10	major capsid protein	50313284	YP_053088.1	504	8240
RELEAVIKEI	10	large tegument protein	50313265	YP_053069.1	834	8249
YYWEVLVDI	9	helicase-primase helicase subunit	50313298	YP_053101.1	192	8250
KDWQTHLDS	9	capsid portal protein	50313297	YP_053100.1	114	8327
KFFAQVRIQL	10	helicase-primase primase subunit	50313248	YP_053052.1	399	8371
HDFLETILM	9	nuclear protein UL4	50313299	YP_053102.1	192	8463
SDTVIHNGV	9	tegument protein UL26	50313282	YP_053086.1	6	8489
KQYNALVRMV	10	transcriptional regulator ICP4	50313321	YP_053124.1	475	8712
RHLPHALSV	9	DNA packaging terminase subunit 5	50313273	YP_053077.1	244	8737
SVFIKFVGL	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	890	8895
VDPARLREI	9	virion protein US2	50313309	YP_053112.1	222	8932
TDCDSDSSV	9	myristylated tegument protein CIRC	50313244	YP_053048.1	244	9146
RDLLNINVR	9	large tegument protein	50313265	YP_053069.1	3278	9209
ADENTLTYAL	10	major capsid protein	50313284	YP_053088.1	1015	9242
ADVSMDGFV	9	large tegument protein	50313265	YP_053069.1	1661	9287
KHYHFDLRRK	10	ribonucleotide reductase subunit 1	50313262	YP_053066.1	420	9444
HDEMWEVVI	9	tegument protein VP13/14	50313255	YP_053059.1	106	9564
RVYRMFCDM	9	virion protein US10	50313319	YP_053122.1	83	9607
PELLSEFCRV	10	transcriptional regulator ICP4	50313321	YP_053124.1	1371	9626
CDAYSFTLI	9	envelope glycoprotein G	50313311	YP_053114.1	118	9716
SQRSDFAPDL	10	transcriptional regulator ICP4	50313321	YP_053124.1	3	9842

FGPGLFARA	9	transcriptional regulator ICP4	50313321	YP_053124.1	1249	9952
RPLQAMMQI	9	nuclear egress membrane protein	50313267	YP_053071.1	141	9994
HDCCETAQI	9	DNA packaging terminase subunit 14	50313286	YP_053090.1	68	9994
PDYKLMCFDI	10	DNA polymerase catalytic subunit	50313271	YP_053075.1	355	10008
VDGVWMRQI	9	envelope glycoprotein M	50313293	YP_053096.1	79	10152
MAIAILSVV	9	envelope glycoprotein D	126801894	YP_053116.2	16	10393
QMYINRNEI	9	helicase-primase primase subunit	50313248	YP_053052.1	631	10461
TDYYFSHLL	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	1136	10751
VDFWRARDNM	10	tegument protein VP13/14	50313254	YP_053058.1	770	10824
HFVGKLVPI	9	tegument protein VP13/14	50313255	YP_053059.1	361	10850
RRFLHPDFI	9	tegument protein UL13	50313289	YP_053093.1	95	10865
ADQLATALL	9	helicase-primase primase subunit	50313248	YP_053052.1	75	11164
HDAENFLKQL	10	DNA packaging terminase subunit 6	50313273	YP_053077.1	147	11198
KQLRGPGGI	9	DNA replication origin-binding helicase	50313294	YP_053097.1	564	11226
AAILAFVFV	9	nuclear egress membrane protein	50313267	YP_053071.1	254	11270
RRQVTIVRI	9	thymidine kinase	50313280	YP_053084.1	20	11278
FYLIAAPYV	9	single-stranded DNA-binding protein	50313272	YP_053076.1	411	11345
MILMHATYF	9	DNA packaging terminase subunit 15	50313286	YP_053090.1	208	11402
GKYFYCNQI	9	multifunctional expression regulator	50313246	YP_053050.1	461	11532
TMFVGRYSI	9	helicase-primase helicase subunit	50313298	YP_053101.1	670	11601
RTFNAEMGV	9	thymidine kinase	50313280	YP_053084.1	344	11604
APRVRSISI	9	transcriptional regulator ICP4	50313321	YP_053124.1	176	11618
AVEPSELLI	9	transcriptional regulator ICP4	50313321	YP_053124.1	1362	11660
DELRLAVRAV	10	transcriptional regulator ICP4	50313321	YP_053124.1	743	11667
IDYDIDELI	9	tegument protein VP13/14	50313254	YP_053058.1	336	11804
TRYYLFQAI	9	helicase-primase primase subunit	50313248	YP_053052.1	246	11837
QKSFILQSL	9	transcriptional regulator ICP4	50313321	YP_053124.1	565	11868
AALSLLSLL	9	envelope glycoprotein G	50313311	YP_053114.1	6	11971
LDPDSIPMI	9	DNA polymerase processivity subunit	50313259	YP_053063.1	308	12146
RHFFLHVCF	9	DNA packaging protein UL33	50313269	YP_053073.1	179	12250
VDATTDTMV	9	tegument protein UL51	50313249	YP_053053.1	130	12266
SDYGYANEL	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	26	12459
AKFVYLIFVY	10	capsid triplex subunit 1	50313263	YP_053067.1	305	12471
PEPDDELEI	9	transcriptional regulator ICP4	50313321	YP_053124.1	1431	12572
DDFTLAAPI	9	large tegument protein	50313265	YP_053069.1	1331	12633
TVTQAVFSI	9	transcriptional regulator ICP4	50313321	YP_053124.1	778	13177
AGERTRVPL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1287	13356
SPAPGLAAM	9	transcriptional regulator ICP4	50313321	YP_053124.1	108	13381
ALPHAAASV	9	transcriptional regulator ICP4	50313321	YP_053124.1	547	13556
DWPQDGPAI	9	transcriptional regulator ICP4	50313321	YP_053124.1	1203	13831
ADFGAAQFPV	10	serine/threonine protein kinase US6	50313310	YP_053113.1	224	13853
GDFAAVVRAV	10	virion protein US11	50313319	YP_053122.1	53	14127
SRGTFVMREN	10	tegument protein VP22	50313252	YP_053056.1	36	14648
HCYWLYNRF	9	helicase-primase subunit	126801893	YP_053098.2	571	14972
RFFETRLGI	9	virion protein US12	50313319	YP_053122.1	162	15003
MDPAWRRRI	9	membrane rotein V1	50313243	YP_053047.1	1	15278
RIAATVTTI	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	833	15321

NDLILALAI	9	helicase-primase primase subunit	50313248	YP_053052.1	155	15369
SLFKFFKEM	9	envelope protein UL43	50313258	YP_053062.1	250	15534
HKIDLPYKEI	10	DNA polymerase catalytic subunit	50313271	YP_053075.1	546	15545
FDPDLFSPV	9	deoxyribonuclease	50313291	YP_053046.1	326	15566
RDYIAHNVN	9	helicase-primase primase subunit	50313248	YP_053052.1	411	15619
FAPDLYDFI	9	transcriptional regulator ICP4	50313321	YP_053124.1	8	15885
DDAPLSTPAV	10	transcriptional regulator ICP4	50313321	YP_053124.1	56	16096
ADAYLRRVV	9	major capsid protein	50313284	YP_053088.1	1143	16199
SDVSEFKKMM	10	membrane protein US8A	50313316	YP_053119.1	65	16407
AVLTSSTLF	9	transcriptional regulator ICP4	50313321	YP_053124.1	1241	16495
FTLPAYTFV	9	helicase-primase helicase subunit	50313298	YP_053101.1	428	16499
KVYLHRDNVI	10	nuclear protein UL3	50313301	YP_053104.1	196	16538
YDLLYFVCL	9	DNA packaging tegument protein UL30	50313278	YP_053082.1	569	17276
SVANPLPHI	9	transcriptional regulator ICP4	50313321	YP_053124.1	527	17366
SRLDIASTV	9	large tegument protein	50313265	YP_053069.1	2168	17427
KLIVNMKSM	9	capsid portal protein	50313297	YP_053100.1	382	17693
HYYIEQESI	9	ribonucleotide reductase subunit 2	50313261	YP_053065.1	99	17937
GDSDPTHRL	9	transcriptional regulator ICP4	50313321	YP_053124.1	732	18674
RDWLAMRKA	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	780	18727
SDNETASMF	9	membrane protein US9	50313317	YP_053120.1	165	18890
RFERFHVYDI	10	DNA polymerase catalytic subunit	50313271	YP_053075.1	122	18962
QDLLALNSGV	10	tegument protein UL51	50313249	YP_053053.1	173	19068
RKFVKHVVSM	10	tegument host shutoff protein	50313260	YP_053064.1	412	19247
RIVAKKINWI	10	tegument protein UL37	50313264	YP_053068.1	202	19343
SSYDEVRNM	9	tegument protein UL27	50313282	YP_053086.1	443	19751
AQAPVPVFV	9	transcriptional regulator ICP4	50313321	YP_053124.1	459	19826
RDLLEHVQA	9	major capsid protein	50313284	YP_053088.1	691	20172
FDKFGRVLV	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	1008	20620
RCLYAVSEL	9	DNA packaging terminase subunit 7	50313273	YP_053077.1	305	20847
SDLYVYVDP	9	DNA packaging terminase subunit 16	50313286	YP_053090.1	504	20968
MDPPSPSAV	9	transactivating tegument protein VP16	126801890	YP_053057.2	389	21216
SYYMYRIVAR	10	envelope glycoprotein H	50313281	YP_053085.1	359	21371
IDFTLCIIQM	10	tegument protein UL7	50313296	YP_053099.1	143	21570
KQLNDGTFV	9	virion protein V67	50313318	YP_053121.1	132	21787
SVLRDFSFI	9	helicase-primase primase subunit	50313248	YP_053052.1	840	22433
KHFFERCTL	9	DNA replication origin-binding helicase	50313294	YP_053097.1	502	22641
SDVWLFSEP	9	deoxyribonuclease	50313291	YP_053046.1	412	22642
HDFLKFKGLS	10	myristylated tegument protein CIRC	50313244	YP_053048.1	111	23054
LQDEDLLGGL	10	transcriptional regulator ICP4	50313321	YP_053124.1	1098	23173
RKVARNREAI	10	helicase-primase primase subunit	50313248	YP_053052.1	931	23256
HDFSVWFED	9	DNA packaging protein UL33	126801892	YP_053072.2	59	23385
CEYHVRVHL	9	DNA packaging terminase subunit 8	50313273	YP_053077.1	124	23523
LDASSDVWL	9	deoxyribonuclease	50313291	YP_053046.1	408	24783
SPFDTFVDNI	10	helicase-primase helicase subunit	50313298	YP_053101.1	690	25214
RDSEHIYET	9	tegument protein VP22	50313252	YP_053056.1	83	25445
IPIPSERYV	9	virion protein V67	50313318	YP_053121.1	204	26388
GILRSEVRI	9	envelope glycoprotein J	50313312	YP_053115.1	718	26457

QDMAFIIVRS	10	ribonucleotide reductase subunit 1	50313262	YP_053066.1	394	26633
HDQASFIGL	9	single-stranded DNA-binding protein	50313272	YP_053076.1	615	26633
CDIILDIMPL	10	tegument host shutoff protein	50313260	YP_053064.1	228	27040
VDYTQKNHS	9	envelope glycoprotein M	50313293	YP_053096.1	69	27360
ADEDLYSDI	9	serine/threonine protein kinase US7	50313310	YP_053113.1	45	28012
RRFFERWHML	10	major capsid protein	50313284	YP_053088.1	503	28601
RDEGQQTDM	9	nuclear egress lamina protein	50313270	YP_053074.1	277	29740
YDLQIARAV	9	large tegument protein	50313265	YP_053069.1	3157	29966
RDCGQREIV	9	helicase-primase subunit	126801893	YP_053098.2	671	30257
ADVLQAVGV	9	large tegument protein	50313265	YP_053069.1	1418	30359
SDVSEFKKM	9	membrane protein US8A	50313316	YP_053119.1	65	31558
KGGLSALLA	9	transcriptional regulator ICP4	50313321	YP_053124.1	1125	31587
CCRATQTNA	9	transcriptional regulator ICP4	50313321	YP_053124.1	872	32359
AVLGAAVYV	9	envelope glycoprotein M	50313293	YP_053096.1	178	32444
AALPPASRL	9	capsid matruation protease	50313276	YP_053080.1	41	32907
AMIGLIVYI	9	envelope glycoprotein G	50313311	YP_053114.1	373	33107
AQAGASASF	9	transcriptional regulator ICP4	50313321	YP_053124.1	939	33148
FESREAMSWL	10	transcriptional regulator ICP4	50313321	YP_053124.1	485	33498
KIFMAVSVL	9	tegument protein UL10	50313288	YP_053092.1	53	34026
PDAEAEPLVL	10	transcriptional regulator ICP4	50313321	YP_053124.1	1385	34250
TQKSFILQSL	10	transcriptional regulator ICP4	50313321	YP_053124.1	564	34300
GGEPFSGSA	9	transcriptional regulator ICP4	50313321	YP_053124.1	816	34862
RRLRTRARI	9	thymidine kinase	50313280	YP_053084.1	191	34879
AEAEPLVLPI	10	transcriptional regulator ICP4	50313321	YP_053124.1	1387	34879
GDAMAAGNA	9	transcriptional regulator ICP4	50313321	YP_053124.1	536	34966
CDWGYRLIAV	10	nuclear protein UL4	50313299	YP_053102.1	36	35513
KNLVSEHEI	9	tegument host shutoff protein	50313260	YP_053064.1	453	35633
KTWFLVPLI	9	DNA packaging terminase subunit 17	50313286	YP_053090.1	265	35702
PSLVGGMLG	9	transcriptional regulator ICP4	50313321	YP_053124.1	787	35773
FPEAWRPAL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1031	36037
RYMQAHSRV	9	tegument protein UL14	50313289	YP_053093.1	59	36096
EDMAADVPV	9	major capsid protein	50313284	YP_053088.1	295	36934
ITGSVANPL	9	transcriptional regulator ICP4	50313321	YP_053124.1	524	37987
KVYLHRDNV	9	nuclear protein UL3	50313301	YP_053104.1	196	38243
LDLSAVPTRV	10	large tegument protein	50313265	YP_053069.1	700	39443
STASSTTSI	9	envelope glycoprotein J	50313312	YP_053115.1	92	39925
VDYTQKNHSV	10	envelope glycoprotein M	50313293	YP_053096.1	69	40419
RKFAAINII	9	membrane rotein V1	50313243	YP_053047.1	175	41474
TDAYAESPAV	10	membrane rotein V1	50313243	YP_053047.1	116	42279
ADRDASANL	9	large tegument protein	50313265	YP_053069.1	1532	42526
YDRTQKSFIL	10	transcriptional regulator ICP4	50313321	YP_053124.1	561	43103
VDFWRARDN	9	tegument protein VP13/14	50313254	YP_053058.1	770	43326
APRRRAPRV	9	transcriptional regulator ICP4	50313321	YP_053124.1	1399	43384
FDIPPPPPL	9	myristylated tegument protein	50313292	YP_053095.1	38	43901
AQALRRAPI	9	transcriptional regulator ICP4	50313321	YP_053124.1	1459	44141
GDAMAAGNAL	10	transcriptional regulator ICP4	50313321	YP_053124.1	536	44916
GDLRQRRVL	9	envelope glycoprotein B	50313274	YP_053078.1	21	45029

114

IDGAGDVELV	10	transcriptional regulator ICP4	50313321	YP_053124.1	1439	45828
RVIFDGQQI	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	606	45895
NDLISRDEAI	10	ribonucleotide reductase subunit 2	50313261	YP_053065.1	198	46532
EDWPQDGPAI	10	transcriptional regulator ICP4	50313321	YP_053124.1	1202	46953
LNAQGVLLL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1160	47514
YDFIESNDF	9	transcriptional regulator ICP4	50313321	YP_053124.1	13	47673
DSRDGLWDD	9	transcriptional regulator ICP4	50313321	YP_053124.1	438	48314
SPAPRVRSI	9	transcriptional regulator ICP4	50313321	YP_053124.1	174	48580
RERHLSRAV	9	transcriptional regulator ICP4	50313321	YP_053124.1	1234	49529
LDAWGVQQV	9	envelope glycoprotein E	50313315	YP_053118.1	19	50889
RINYRRREL	9	transactivating tegument protein VP16	126801890	YP_053057.2	301	51031
RDRRTEQRL	9	capsid triplex subunit 1	50313263	YP_053067.1	214	51111
APLPPGAFI	9	capsid triplex subunit 2	50313285	YP_053089.1	91	51717
RRYARGKEC	9	envelope glycoprotein M	50313293	YP_053096.1	128	52703
KFFAQVRIQ	9	helicase-primase primase subunit	50313248	YP_053052.1	399	53024
RHFRQRATV	9	DNA packaging terminase subunit 18	50313286	YP_053090.1	248	53838
RIVAKKINW	9	tegument protein UL37	50313264	YP_053068.1	202	54882
HDFLKFKGL	9	myristylated tegument protein CIRC	50313244	YP_053048.1	111	55550
RRYARKRRH	9	capsid matruation protease	50313276	YP_053080.1	444	55679
GELAASGPL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1065	56183
GLYRRVIEI	9	envelope glycoprotein D	126801894	YP_053116.2	183	56467
EDLLAEIDAI	10	tegument protein UL37	50313264	YP_053068.1	1006	56574
AALPSEVYL	9	membrane rotein V1	50313243	YP_053047.1	139	56832
VEHPGPSPGV	10	transcriptional regulator ICP4	50313321	YP_053124.1	1449	57166
REANLAKTL	9	transcriptional regulator ICP4	50313321	YP_053124.1	306	57306
AISQYHIYM	9	transcriptional regulator ICP4	50313321	YP_053124.1	1210	58062
HQIPDPEDV	9	transcriptional regulator ICP4	50313321	YP_053124.1	1081	58621
PEAPAPSVRV	10	transcriptional regulator ICP4	50313321	YP_053124.1	613	58926
DDKEIMSLI	9	DNA packaging terminase subunit 23	50313287	YP_053091.1	242	59010
RKSIHVTSI	9	deoxyuridine triphosphatase	50313250	YP_053054.1	144	59528
QDGPAISQY	9	transcriptional regulator ICP4	50313321	YP_053124.1	1206	59591
KLFFVHAHI	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	762	59942
ACVLACQAV	9	transcriptional regulator ICP4	50313321	YP_053124.1	664	60523
FEYDQRTLI	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	707	60881
WDDPEIVLA	9	transcriptional regulator ICP4	50313321	YP_053124.1	444	61443
EGARARASA	9	transcriptional regulator ICP4	50313321	YP_053124.1	768	61770
LTKSLLSDI	9	large tegument protein	50313265	YP_053069.1	2042	62520
AELIDNQLF	9	transcriptional regulator ICP4	50313321	YP_053124.1	1023	63215
LEGVREANL	9	transcriptional regulator ICP4	50313321	YP_053124.1	302	64550
SDLYTFLTK	9	serine/threonine protein kinase US8	50313310	YP_053113.1	166	64800
RLYVVCGEL	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	67	65841
TQPAAPDLL	9	transcriptional regulator ICP4	50313321	YP_053124.1	38	66157
YFWAQRRKF	9	transactivating tegument protein VP16	126801890	YP_053057.2	265	66296
GYFLSKRDI	9	helicase-primase subunit	126801893	YP_053098.2	192	67298
QQHQPQKAR	9	transcriptional regulator ICP4	50313321	YP_053124.1	885	68295
SRFAYVKFI	9	myristylated tegument protein CIRC	50313244	YP_053048.1	64	-
PDLTNFKCI	9	nuclear protein UL55	50313245	YP_053049.1	107	-

RDAVMRGEN	9	multifunctional expression regulator	50313246	YP_053050.1	299	-
FVVAVTIVEI	10	envelopeenvelope glycoprotein K	50313247	YP_053051.1	259	-
FDYVCRENW	9	envelopeenvelope glycoprotein K	50313247	YP_053051.1	79	-
KELHDKVRI	9	envelopeenvelope glycoprotein K	50313247	YP_053051.1	100	-
SRTFKRDPI	9	envelopeenvelope glycoprotein K	50313247	YP_053051.1	198	-
LHYEQRIQI	9	envelopeenvelope glycoprotein K	50313247	YP_053051.1	326	-
LDVDFGIKRR	10	helicase-primase primase subunit	50313248	YP_053052.1	651	-
VDFGIKRRV	9	helicase-primase primase subunit	50313248	YP_053052.1	653	-
GDPGVKQRI	9	envelope glycoprotein N	50313251	YP_053055.1	27	-
GHVDLYTEI	9	transactivating tegument protein VP16	126801890	YP_053057.2	100	-
SLMHKFKQV	9	transactivating tegument protein VP16	126801890	YP_053057.2	209	-
HKFKQVVRD	9	transactivating tegument protein VP16	126801890	YP_053057.2	212	-
SEYHDHGEM	9	tegument protein VP13/14	50313254	YP_053058.1	169	-
SDVVTKQNA	9	tegument protein VP13/14	50313254	YP_053058.1	447	-
SSVGKRGLI	9	tegument protein VP13/14	50313254	YP_053058.1	848	-
KDVTENIVGQ	10	tegument protein VP13/14	50313255	YP_053059.1	42	-
HDGFIRRTM	9	tegument protein VP13/14	50313255	YP_053059.1	720	-
VDFTKYVTN	9	envelope glycoprotein C	50313257	YP_053061.1	261	-
RRGGQLGVI	9	envelope glycoprotein C	50313257	YP_053061.1	150	-
RWYKNAREV	9	envelope glycoprotein C	50313257	YP_053061.1	253	-
SKVVFEIENY	10	DNA polymerase processivity subunit	50313259	YP_053063.1	125	-
HDYDSSTED	9	tegument host shutoff protein	50313260	YP_053064.1	323	-
RKFVKHVVS	9	tegument host shutoff protein	50313260	YP_053064.1	412	-
KFVKHVVSM	9	tegument host shutoff protein	50313260	YP_053064.1	413	-
RIYELFSEA	9	ribonucleotide reductase subunit 2	50313261	YP_053065.1	230	-
KYWLNSTEI	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	89	-
THVFPHKHM	9	capsid triplex subunit 1	50313263	YP_053067.1	235	-
HVFPHKHMT	9	capsid triplex subunit 1	50313263	YP_053067.1	236	-
KRMESIEEF	9	tegument protein UL37	50313264	YP_053068.1	171	-
REYLKKEDL	9	tegument protein UL37	50313264	YP_053068.1	1000	-
REYLKKEDLL	10	tegument protein UL37	50313264	YP_053068.1	1000	-
GDGFQSTQI	9	large tegument protein	50313265	YP_053069.1	124	-
KLVARVEEL	9	large tegument protein	50313265	YP_053069.1	581	-
HDKFDAIEI	9	large tegument protein	50313265	YP_053069.1	674	-
SDTTNMDEL	9	large tegument protein	50313265	YP_053069.1	688	-
KVIKKVESM	9	large tegument protein	50313265	YP_053069.1	711	-
SLFGQFREA	9	large tegument protein	50313265	YP_053069.1	1007	-
REFAKRAEQ	9	large tegument protein	50313265	YP_053069.1	1278	-
SKFAERHKN	9	large tegument protein	50313265	YP_053069.1	1707	-
GLYDEYQKI	9	large tegument protein	50313265	YP_053069.1	1783	-
HDILQLRFG	9	large tegument protein	50313265	YP_053069.1	1906	-
RALEEVTYV	9	large tegument protein	50313265	YP_053069.1	1433	-
RTLQASMGV	9	large tegument protein	50313265	YP_053069.1	1676	-
WDQHLVPAV	9	large tegument protein	50313265	YP_053069.1	3302	-
LDYTKYTED	9	small capsid protein	50313266	YP_053070.1	46	-
MDSYNYRDF	9	nuclear egress membrane protein	50313267	YP_053071.1	1	-
RRVLTLVDI	9	DNA packaging protein UL33	50313269	YP_053073.1	169	-

KSVHLHYRLI	10	nuclear egress lamina protein	50313270	YP_053074.1	241	-
KLYAVFDDF	9	nuclear egress lamina protein	50313270	YP_053074.1	314	-
RFHVYDIVET	10	DNA polymerase catalytic subunit	50313271	YP_053075.1	125	-
HVYGVRHYF	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	172	-
RHYFYMAKA	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	177	-
GDVLGEHKI	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	540	-
HKPYSVRVI	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	876	-
KKYIGVING	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	940	-
ARITLTRVI	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	600	-
HDAENFLKQ	9	DNA packaging terminase subunit 9	50313273	YP_053077.1	147	-
HKFFELING	9	DNA packaging terminase subunit 10	50313273	YP_053077.1	599	-
RANVYYKDI	9	envelope glycoprotein B	50313274	YP_053078.1	184	-
RVPVSVEEI	9	envelope glycoprotein B	50313274	YP_053078.1	213	-
TKWKEVDEL	9	envelope glycoprotein B	50313274	YP_053078.1	392	-
QDGVHREQI	9	envelope glycoprotein B	50313274	YP_053078.1	533	-
TLWNEMVKI	9	envelope glycoprotein B	50313274	YP_053078.1	611	-
KRYFKFGKE	9	envelope glycoprotein B	50313274	YP_053078.1	713	-
RYFKFGKEY	9	envelope glycoprotein B	50313274	YP_053078.1	714	-
KEYVYYENY	9	envelope glycoprotein B	50313274	YP_053078.1	720	-
TAWFMFRNI	9	envelope protein UL20	50313283	YP_053087.1	119	-
TLFDFFKQI	9	major capsid protein	50313284	YP_053088.1	32	-
THTYMMKHI	9	major capsid protein	50313284	YP_053088.1	114	-
KHERNRKEL	9	major capsid protein	50313284	YP_053088.1	234	-
RRFFERWHM	9	major capsid protein	50313284	YP_053088.1	503	-
RFFERWHML	9	major capsid protein	50313284	YP_053088.1	504	-
MDRRSEAFKI	10	major capsid protein	50313284	YP_053088.1	1	-
GDAHRRQCV	9	DNA packaging terminase subunit 24	50313287	YP_053091.1	273	-
SYFVHFHAM	9	envelope glycoprotein M	50313293	YP_053096.1	318	-
RRYANKVRG	9	envelope glycoprotein M	50313293	YP_053096.1	383	-
RPGVAHDEI	9	DNA replication origin-binding helicase	50313294	YP_053097.1	594	-
SFFGGLQHI	9	helicase-primase subunit	126801893	YP_053098.2	459	-
SRFEHKNLK	9	helicase-primase helicase subunit	50313298	YP_053101.1	305	-
RLYSSHKEV	9	helicase-primase helicase subunit	50313298	YP_053101.1	398	-
AKYNSKADI	9	serine/threonine protein kinase US9	50313310	YP_053113.1	256	-
KDGQPEPVV	9	serine/threonine protein kinase US10	50313310	YP_053113.1	112	-
GIVEFTIVNM	10	envelope glycoprotein G	50313311	YP_053114.1	130	-
HRNYQRLEY	9	envelope glycoprotein G	50313311	YP_053114.1	401	-
HDYFRVNAS	9	envelope glycoprotein I	50313314	YP_053117.1	115	-
TDFVVHGYHS	10	envelope glycoprotein E	50313315	YP_053118.1	198	-
RFYKQLNDG	9	virion protein V67	50313318	YP_053121.1	129	-
AEAQAPVPVF	10	transcriptional regulator ICP4	50313321	YP_053124.1	457	-
GDSTKQYNAL	10	transcriptional regulator ICP4	50313321	YP_053124.1	471	-
QQPEAPAPSV	10	transcriptional regulator ICP4	50313321	YP_053124.1	611	-
PQAAASPAHL	10	transcriptional regulator ICP4	50313321	YP_053124.1	895	-
AEGPDRRGGF	10	transcriptional regulator ICP4	50313321	YP_053124.1	986	-
PDRRGGFRRV	10	transcriptional regulator ICP4	50313321	YP_053124.1	989	-
PEAWRPALTF	10	transcriptional regulator ICP4	50313321	YP_053124.1	1032	-

PDPEDVKVVV	10	transcriptional regulator ICP4	50313321	YP_053124.1	1084	-
LEPDAEAEPL	10	transcriptional regulator ICP4	50313321	YP_053124.1	1383	-
DEAGRPDSL	9	transcriptional regulator ICP4	50313321	YP_053124.1	585	-
SDDRAIHGR	9	transcriptional regulator ICP4	50313321	YP_053124.1	914	-
GDAADPVAL	9	transcriptional regulator ICP4	50313321	YP_053124.1	141	-
NDFGEDPLI	9	transcriptional regulator ICP4	50313321	YP_053124.1	19	-
YGSQNMFGV	9	transcriptional regulator ICP4	50313321	YP_053124.1	47	-
VDDAPLSTP	9	transcriptional regulator ICP4	50313321	YP_053124.1	55	-
PLSTPAVVI	9	transcriptional regulator ICP4	50313321	YP_053124.1	59	-
PGLAAMLKM	9	transcriptional regulator ICP4	50313321	YP_053124.1	111	-
SSSSSSSSM	9	transcriptional regulator ICP4	50313321	YP_053124.1	186	-
EDDQADGAG	9	transcriptional regulator ICP4	50313321	YP_053124.1	196	-
DDQADGAGA	9	transcriptional regulator ICP4	50313321	YP_053124.1	197	-
GDRPAAGAA	9	transcriptional regulator ICP4	50313321	YP_053124.1	251	-
APSRPGGGL	9	transcriptional regulator ICP4	50313321	YP_053124.1	283	-
SDPPPMGRV	9	transcriptional regulator ICP4	50313321	YP_053124.1	423	-
GDSRDGLWD	9	transcriptional regulator ICP4	50313321	YP_053124.1	437	-
DGLWDDPEI	9	transcriptional regulator ICP4	50313321	YP_053124.1	441	-
EAQAPVPVF	9	transcriptional regulator ICP4	50313321	YP_053124.1	458	-
STKQYNALV	9	transcriptional regulator ICP4	50313321	YP_053124.1	473	-
ESREAMSWL	9	transcriptional regulator ICP4	50313321	YP_053124.1	486	-
GQDQNLAQF	9	transcriptional regulator ICP4	50313321	YP_053124.1	501	-
HAPRGHGSF	9	transcriptional regulator ICP4	50313321	YP_053124.1	515	-
APRGHGSFI	9	transcriptional regulator ICP4	50313321	YP_053124.1	516	-
AAGNALWAL	9	transcriptional regulator ICP4	50313321	YP_053124.1	540	-
PHAAASVAM	9	transcriptional regulator ICP4	50313321	YP_053124.1	549	-
YDRTQKSFI	9	transcriptional regulator ICP4	50313321	YP_053124.1	561	-
GRPDSLAAV	9	transcriptional regulator ICP4	50313321	YP_053124.1	588	-
VPGLDPSEI	9	transcriptional regulator ICP4	50313321	YP_053124.1	686	-
GSLQTLPPL	9	transcriptional regulator ICP4	50313321	YP_053124.1	828	-
GSSSDDRAI	9	transcriptional regulator ICP4	50313321	YP_053124.1	911	-
EGPDRRGGF	9	transcriptional regulator ICP4	50313321	YP_053124.1	987	-
IDNQLFPEA	9	transcriptional regulator ICP4	50313321	YP_053124.1	1026	-
EAWRPALTF	9	transcriptional regulator ICP4	50313321	YP_053124.1	1033	-
TGAPDVSAL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1152	-
PDAEAEPLV	9	transcriptional regulator ICP4	50313321	YP_053124.1	1385	-
EAEPLVLPI	9	transcriptional regulator ICP4	50313321	YP_053124.1	1388	-
FGSRSTSVL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1413	-
IDGAGDVEL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1439	-
ALRRAPIKI	9	transcriptional regulator ICP4	50313321	YP_053124.1	1461	-
GGDWCNPYL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1478	-
PDLYDFIES	9	transcriptional regulator ICP4	50313321	YP_053124.1	10	-
PLIRAASAA	9	transcriptional regulator ICP4	50313321	YP_053124.1	25	-
GLLPPGARI	9	transcriptional regulator ICP4	50313321	YP_053124.1	290	-
DDPEIVLAA	9	transcriptional regulator ICP4	50313321	YP_053124.1	445	-
GGASAVPGL	9	transcriptional regulator ICP4	50313321	YP_053124.1	681	-
AAEGSLQTL	9	transcriptional regulator ICP4	50313321	YP_053124.1	825	-

DPEDVKVVV	9	transcriptional regulator ICP4	50313321	YP_053124.1	1085	-
APDVSALNA	9	transcriptional regulator ICP4	50313321	YP_053124.1	1154	-
PSELLIPEL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1365	-
PDSLAAVAG	9	transcriptional regulator ICP4	50313321	YP_053124.1	590	-
AASQQQPEA	9	transcriptional regulator ICP4	50313321	YP_053124.1	607	-
RVREAYTRV	9	transcriptional regulator ICP4	50313321	YP_053124.1	621	-
DELRLAVRA	9	transcriptional regulator ICP4	50313321	YP_053124.1	743	-
GEAVSLLAP	9	transcriptional regulator ICP4	50313321	YP_053124.1	795	-
GPDRRGGFR	9	transcriptional regulator ICP4	50313321	YP_053124.1	988	-
AGAVEPSEL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1360	-
DMNSSQATG	9	transcriptional regulator ICP4	50313321	YP_053124.1	380	-
RPDSLAAVA	9	transcriptional regulator ICP4	50313321	YP_053124.1	589	-
GAGATTATC	9	transcriptional regulator ICP4	50313321	YP_053124.1	864	-
RSPQAAASP	9	transcriptional regulator ICP4	50313321	YP_053124.1	893	-
SPPAPKRRV	9	transcriptional regulator ICP4	50313321	YP_053124.1	962	-
APRALGPMP	9	transcriptional regulator ICP4	50313321	YP_053124.1	977	-
DEDLLGGLP	9	transcriptional regulator ICP4	50313321	YP_053124.1	1100	-
RHLSRAVLT	9	transcriptional regulator ICP4	50313321	YP_053124.1	1236	-
YGGTGDSRD	9	transcriptional regulator ICP4	50313321	YP_053124.1	433	-
AAAAAPGTR	9	transcriptional regulator ICP4	50313321	YP_053124.1	641	-
ARSPQAAAS	9	transcriptional regulator ICP4	50313321	YP_053124.1	892	-
FGELAASGP	9	transcriptional regulator ICP4	50313321	YP_053124.1	1064	-
ISQYHIYMR	9	transcriptional regulator ICP4	50313321	YP_053124.1	1211	-
HPGPSPGVA	9	transcriptional regulator ICP4	50313321	YP_053124.1	1451	-
QALRRAPIK	9	transcriptional regulator ICP4	50313321	YP_053124.1	1460	-
APDLYDFIE	9	transcriptional regulator ICP4	50313321	YP_053124.1	9	-
QNMFGVDDA	9	transcriptional regulator ICP4	50313321	YP_053124.1	50	-
AQPASPAPS	9	transcriptional regulator ICP4	50313321	YP_053124.1	100	-
RSPHRLPKD	9	transcriptional regulator ICP4	50313321	YP_053124.1	329	-
GASKRKRAN	9	transcriptional regulator ICP4	50313321	YP_053124.1	346	-
VRMVFESRE	9	transcriptional regulator ICP4	50313321	YP_053124.1	481	-
ALSLYELRD	9	transcriptional regulator ICP4	50313321	YP_053124.1	715	-
VSTEDWPQD	9	transcriptional regulator ICP4	50313321	YP_053124.1	1199	-
GARILEYLE	9	transcriptional regulator ICP4	50313321	YP_053124.1	295	-
LAASGPLRR	9	transcriptional regulator ICP4	50313321	YP_053124.1	1067	-
VLYSPLQDE	9	transcriptional regulator ICP4	50313321	YP_053124.1	1093	-
AAFARLYPD	9	transcriptional regulator ICP4	50313321	YP_053124.1	1259	-
DTRAGERTR	9	transcriptional regulator ICP4	50313321	YP_053124.1	1284	-
VLPDYDGCK	9	transcriptional regulator ICP4	50313321	YP_053124.1	1304	-
VEPSELLIP	9	transcriptional regulator ICP4	50313321	YP_053124.1	1363	-
LLEPDAEAE	9	transcriptional regulator ICP4	50313321	YP_053124.1	1382	-
WEPGFGSRS	9	transcriptional regulator ICP4	50313321	YP_053124.1	1409	-
CLPEPDDEL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1429	-
AEEAPQTAV	9	transcriptional regulator ICP4	50313321	YP_053124.1	367	-
AEAQAPVPV	9	transcriptional regulator ICP4	50313321	YP_053124.1	457	-
CQAVFEALL	9	transcriptional regulator ICP4	50313321	YP_053124.1	669	-
PQTAVWDLL	9	transcriptional regulator ICP4	50313321	YP_053124.1	371	-

Peptide	Len	Protein	GI no.	Locus	Position	
SELLIPELL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1366	-
PDYDGCKDM	9	transcriptional regulator ICP4	50313321	YP_053124.1	1306	-
GDVELVVEH	9	transcriptional regulator ICP4	50313321	YP_053124.1	1443	-
LEIDGAGDV	9	transcriptional regulator ICP4	50313321	YP_053124.1	1437	

Supplemental Table2: EHV-1-derived peptides tested for binding to Eqca-N*00101

Peptide	Len	Protein	GI no.	Locus	Position	IC50 nM
NPIFYRLAY	9	deoxyribonuclease	50313291	YP_053046.1	101	2,9
TVWYGRVAF	9	membrane protein V1	50313243	YP_053047.1	55	8,4
TPAAAAQPA	9	transcriptional regulator ICP4	50313321	YP_053124.1	95	14,6
SAWNGPMAF	9	ubiquitin E3 ligase ICP0	50313304	YP_053107.1	100	22,8
NPFIGKRPF	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	17	25,7
VPWLNVIPL	9	envelope glycoprotein J	50313312	YP_053115.1	621	30,4
GPFRNQNEF	9	envelope glycoprotein K	50313247	YP_053051.1	148	45,7
DPITFQCLF	9	DNA polymerase processivity subunit	50313259	YP_053063.1	212	53,7
YAASPMPVF	9	virion protein US10	50313319	YP_053122.1	95	57,5
SAFASYSSF	9	envelope glycoprotein H	50313281	YP_053085.1	796	77,7
GPWFAKFDF	9	helicase-primase subunit	126801893	YP_053098.2	412	91,1
DAFEMLTPI	9	DNA polymerase processivity subunit	50313259	YP_053063.1	45	93,3
FPHKHMTVF	9	capsid triplex subunit 1	50313263	YP_053067.1	238	101,3
LAFTGCVEY	9	transcriptional regulator ICP4	50313321	YP_053124.1	1173	126,9
EAWRPALTF	9	transcriptional regulator ICP4	50313321	YP_053124.1	1033	130,0
DPIGFLCEH	9	envelope glycoprotein K	50313247	YP_053051.1	204	133,2
DALRSSDAF	9	DNA packaging terminase subunit 1	50313286	YP_053090.1	146	150,1
TPLGASVFW	9	envelope glycoprotein H	50313281	YP_053085.1	81	200,4
WPALGMSAF	9	helicase-primase subunit	126801893	YP_053098.2	351	207,4
FPGPSRFAY	9	myristylated tegument protein CIRC	50313244	YP_053048.1	60	221,9
SPLPLDLTW	9	nuclear protein UL55	50313245	YP_053049.1	46	238,2
SAWTVTQAVF	10	transcriptional regulator ICP4	50313321	YP_053124.1	775	248
EARSCRNAY	9	DNA packaging terminase subunit 1	50313286	YP_053090.1	130	261,4
TCPGSPQPEF	10	transcriptional regulator ICP4	50313321	YP_053124.1	155	271
SPLEAEPLF	9	transactivating tegument protein VP16	126801890	YP_053057.2	324	292,4
NATALASPF	9	envelope glycoprotein D	126801894	YP_053116.2	61	393,5
NSMAMRAEF	9	tegument serine/threonine protein kinase	50313290	YP_053094.1	380	396,8
PPHTAGVPY	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	468	406,5
VLWKFTTDF	9	deoxyuridine triphosphatase	50313250	YP_053054.1	299	496,1
CPGSPQPEF	9	transcriptional regulator ICP4	50313321	YP_053124.1	156	558,2
CPGSPQPEF	9	transcriptional regulator ICP4	50313321	YP_053124.1	156	558
DGYKLLLSY	9	envelope glycoprotein H	50313281	YP_053085.1	839	559,0
DAYAESPAV	9	membrane protein V1	50313243	YP_053047.1	117	575,4
SPAPSPAPG	9	transcriptional regulator ICP4	50313321	YP_053124.1	104	615
RAFFGSLVY	9	tegument protein VP13/14	50313254	YP_053058.1	540	620,7
DTALLRTPH	9	major capsid protein	50313284	YP_053088.1	1339	723,0

NSAVFRAFF	9	tegument protein VP13/14	50313254	YP_053058.1	535	907,1
SPAPSPAPGL	10	transcriptional regulator ICP4	50313321	YP_053124.1	104	945
QAPYKKSDF	9	envelope glycoprotein C	50313257	YP_053061.1	460	945,7
VAAPLSLAF	9	envelope protein UL43	50313258	YP_053062.1	344	1019,0
NPKSERPAF	9	multifunctional expression regulator	50313246	YP_053050.1	84	1100,4
DPDSAYQSW	9	tegument protein VP13/14	50313254	YP_053058.1	298	1132,1
TPPRPVEDF	9	helicase-primase helicase subunit	50313298	YP_053101.1	26	1221,8
SNYTYPTSF	9	DNA packaging tegument protein UL25	50313278	YP_053082.1	250	1230,4
CPRVANNAF	9	envelope glycoprotein I	50313314	YP_053117.1	95	1231,2
QVASFVARF	9	DNA packaging terminase subunit 1	50313286	YP_053090.1	156	1285,1
DTLSLMAAF	9	capsid triplex subunit 2	50313285	YP_053089.1	59	1665,8
TAGVPYFDF	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	471	1681,3
TRPDLTVVF	9	thymidine kinase	50313280	YP_053084.1	129	1820,7
GLFARAEAAF	10	transcriptional regulator ICP4	50313321	YP_053124.1	1252	1849
DIMPLFPPT	9	tegument host shutoff protein	50313260	YP_053064.1	233	1852,6
TPRVACAVRW	10	transcriptional regulator ICP4	50313321	YP_053124.1	1223	2037
SPAPGLAAML	10	transcriptional regulator ICP4	50313321	YP_053124.1	108	2104
FPNEVYLNF	9	helicase-primase helicase subunit	50313298	YP_053101.1	39	2132,4
TVAAALPPA	9	capsid maturation protease	50313276	YP_053080.1	38	2162,6
AGLFLTAVY	9	tegument protein UL37	50313264	YP_053068.1	711	2190,9
DPAAENIPT	9	tegument protein UL16	50313288	YP_053092.1	149	2351,9
NPPLEMIVV	9	envelope glycoprotein H	50313281	YP_053085.1	262	2385,8
DPPPMGRVRY	10	transcriptional regulator ICP4	50313321	YP_053124.1	424	2759
HPLAMNPSW	9	tegument protein VP13/14	50313254	YP_053058.1	312	2970,8
IAAASKLSF	9	tegument protein UL37	50313264	YP_053068.1	568	3617,1
EALRTRRVF	9	large tegument protein	50313265	YP_053069.1	415	3944,4
DPQALATIAA	10	transcriptional regulator ICP4	50313321	YP_053124.1	1042	3950
NEAWANVTF	9	envelope glycoprotein G	50313311	YP_053114.1	217	4145,3
LANTLRHVW	9	tegument protein UL37	50313264	YP_053068.1	908	4156,3
SAFRLRELG	9	transcriptional regulator ICP4	50313321	YP_053124.1	654	4185
DVDGHLWEF	9	virion protein US2	50313309	YP_053112.1	24	4259,5
TVLEFDSEY	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	431	4307,2
NPAGLETAAL	10	transcriptional regulator ICP4	50313321	YP_053124.1	707	4816
QPERPPVIF	9	envelope glycoprotein G	50313311	YP_053114.1	41	4956,4
SADPLRRAV	9	virion protein US10	50313319	YP_053122.1	105	5027,8
DVPPVPHFF	9	major capsid protein	50313284	YP_053088.1	986	5400,0
EAARRRFGF	9	single-stranded DNA-binding protein	50313272	YP_053076.1	114	5409,7
IAPSLKNAF	9	DNA polymerase processivity subunit	50313259	YP_053063.1	53	5592,6
CPLAIIQDV	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	10	5598,3
DAFKAKTQT	9	DNA polymerase processivity subunit	50313259	YP_053063.1	114	5667,3
DAAPPRGPA	9	virion protein US2	50313309	YP_053112.1	268	5758,5
DALLLLQEV	9	major capsid protein	50313284	YP_053088.1	889	5766,8
PPGARILEY	9	transcriptional regulator ICP4	50313321	YP_053124.1	293	5870
LPFDSYPGV	9	ribonucleotide reductase subunit 1	50313262	YP_053066.1	583	6088,1
DPALLPPLI	9	major capsid protein	50313284	YP_053088.1	728	6466,7
LPPGARILEY	10	transcriptional regulator ICP4	50313321	YP_053124.1	292	6528
DPSYSGAAA	9	serine/threonine protein kinase US3	50313310	YP_053113.1	37	6585,4

DSPPISPEV	9	tegument protein VP11/12	50313255	YP_053059.1	656	6889,2
DAPTPTMAH	9	tegument protein UL37	50313264	YP_053068.1	660	6892,2
GPLRRRAAW	9	transcriptional regulator ICP4	50313321	YP_053124.1	1071	6979,0
GPLRRRAAW	9	transcriptional regulator ICP4	50313321	YP_053124.1	1071	6979
DMAYPRDEA	9	transcriptional regulator ICP4	50313321	YP_053124.1	579	7272,1
DMAYPRDEA	9	transcriptional regulator ICP4	50313321	YP_053124.1	579	7272
LIPELLSEF	9	transcriptional regulator ICP4	50313321	YP_053124.1	1369	7423,6
LIPELLSEF	9	transcriptional regulator ICP4	50313321	YP_053124.1	1369	7424
AAMVPPLEW	9	helicase-primase helicase subunit	50313298	YP_053101.1	70	7564,1
RVRDLQRAF	9	ubiquitin E3 ligase ICP0	50313304	YP_053107.1	501	7592,0
GAAALPSEV	9	membrane protein V1	50313243	YP_053047.1	137	7653,9
DPPSTPEEY	9	serine/threonine protein kinase US3	50313310	YP_053113.1	284	7939,8
FVFGFSGTF	9	helicase-primase helicase subunit	50313298	YP_053101.1	514	8056,8
SPAHLSQEA	9	transcriptional regulator ICP4	50313321	YP_053124.1	900	8060
SGPLRRRAAW	10	transcriptional regulator ICP4	50313321	YP_053124.1	1070	8215
FPDTPTAPW	9	protein V32	50313275	YP_053079.1	138	8692,6
SPSDPRHPL	9	major capsid protein	50313284	YP_053088.1	857	8892,6
SPACPPEAL	9	transcriptional regulator ICP4	50313321	YP_053124.1	696	8902
TYDEFLTEF	9	tegument host shutoff protein	50313260	YP_053064.1	254	8908,3
NPLGLYTYL	9	large tegument protein	50313265	YP_053069.1	599	8929,6
AAASCAPGVY	10	transcriptional regulator ICP4	50313321	YP_053124.1	398	9097
DGRRIYTDF	9	envelope glycoprotein D	126801894	YP_053116.2	192	9804,9
TAPPSYSEA	9	membrane protein UL56	50313242	YP_053046.1	119	10028,7
NAPKPSPSF	9	ribonucleotide reductase subunit 2	50313261	YP_053065.1	284	10125,9
TAPVPINEV	9	DNA packaging tegument protein UL25	50313278	YP_053082.1	518	10496,1
FLLYLSRNF	9	myristylated tegument protein CIRC	50313244	YP_053048.1	84	10679,3
NTLDWLPGF	9	capsid triplex subunit 1	50313263	YP_053067.1	56	10688,7
NALFACPEH	9	major capsid protein	50313284	YP_053088.1	961	11467,8
QPASPAPSPA	10	transcriptional regulator ICP4	50313321	YP_053124.1	101	11685
RAASAAEEGF	10	transcriptional regulator ICP4	50313321	YP_053124.1	28	11840
DEFVYSRPF	9	single-stranded DNA-binding protein	50313272	YP_053076.1	220	11911,9
TPMLLNHVV	9	DNA replication origin-binding helicase	50313294	YP_053097.1	448	12148,7
SPGAPAPAPA	10	transcriptional regulator ICP4	50313321	YP_053124.1	270	12168
VSDSIREYF	9	large tegument protein	50313265	YP_053069.1	720	12631,6
EPPRTCPKF	9	envelope glycoprotein B	50313274	YP_053078.1	152	13366,0
SVAVYRPPT	9	nuclear protein UL4	50313299	YP_053102.1	138	13403,4
RVPLAPREY	9	transcriptional regulator ICP4	50313321	YP_053124.1	1292	13586
GPQGPRPIF	9	major capsid protein	50313284	YP_053088.1	1158	14521,5
DALVYAEEL	9	helicase-primase subunit	126801893	YP_053098.2	521	14939,8
TLPPLWPTV	9	transcriptional regulator ICP4	50313321	YP_053124.1	832	15222
DRYSLYLDY	9	nuclear egress lamina protein	50313270	YP_053074.1	68	15630,9
SANALCCVF	9	large tegument protein	50313265	YP_053069.1	102	15869,0
QAPRPRADF	9	tegument protein VP13/14	50313254	YP_053058.1	98	15948,9
SGLDMSEDF	9	DNA replication origin-binding helicase	50313294	YP_053097.1	698	17351,4
DPQALATIA	9	transcriptional regulator ICP4	50313321	YP_053124.1	1042	17389
LPSDLSPAD	9	capsid triplex subunit 2	50313285	YP_053089.1	12	18028,3
TPPPQEPPA	9	large tegument protein	50313265	YP_053069.1	3179	18356,2

122

VPLDWWTET	9	large tegument protein	50313265	YP_053069.1	3245	18394,7
DAMAAGNALW	10	transcriptional regulator ICP4	50313321	YP_053124.1	537	18497
LAASRYAEA	9	transcriptional regulator ICP4	50313321	YP_053124.1	451	18500
FWGYLRRVF	9	tegument protein VP13/14	50313254	YP_053058.1	81	18577,9
VPNTLNILF	9	major capsid protein	50313284	YP_053088.1	871	19190,2
PPNALASLW	9	large tegument protein	50313265	YP_053069.1	2139	19702,4
FATGFHPGF	9	major capsid protein	50313284	YP_053088.1	1040	20099,3
LPLAGFVGL	9	large tegument protein	50313265	YP_053069.1	956	20601,4
TALVLDCGV	9	nuclear protein UL4	50313299	YP_053102.1	68	20889,9
VRMVFESRE	9	transcriptional regulator ICP4	50313321	YP_053124.1	481	21071,7
SAAAPATRV	9	membrane protein US9	50313317	YP_053120.1	130	23184,0
AAFARLYPD	9	transcriptional regulator ICP4	50313321	YP_053124.1	1259	23227,9
MAAAEAVAA	9	membrane protein US9	50313317	YP_053120.1	45	23880,1
LRADLRADV	9	tegument protein UL37	50313264	YP_053068.1	801	24012,8
DPTLLPTET	9	envelope glycoprotein B	50313274	YP_053078.1	109	24214,4
GASKRKRAN	9	transcriptional regulator ICP4	50313321	YP_053124.1	346	24310,8
TAASTSAET	9	envelope glycoprotein J	50313312	YP_053115.1	138	25096,3
AAAALRTPV	9	major capsid protein	50313284	YP_053088.1	1111	25296,3
SFLARRTVF	9	single-stranded DNA-binding protein	50313272	YP_053076.1	677	25326,7
ALMHLYLSV	9	tegument protein UL51	50313249	YP_053053.1	119	26503,6
LARDTAPRF	9	DNA packaging protein UL32	50313269	YP_053073.1	390	26693,5
EAQAPVPVF	9	transcriptional regulator ICP4	50313321	YP_053124.1	458	26784
SAMLRNSMH	9	large tegument protein	50313265	YP_053069.1	2010	27233,5
PAPPKTTPV	9	large tegument protein	50313265	YP_053069.1	2483	27554,2
SLQTLPPLW	9	transcriptional regulator ICP4	50313321	YP_053124.1	829	28129,6
SLQTLPPLW	9	transcriptional regulator ICP4	50313321	YP_053124.1	829	28130
PAISQYHIY	9	transcriptional regulator ICP4	50313321	YP_053124.1	1209	29173
TLATLMLEV	9	tegument protein VP13/14	50313254	YP_053058.1	602	30085,3
SPPAPKRRV	9	transcriptional regulator ICP4	50313321	YP_053124.1	962	30648
AWTVTQAVF	9	transcriptional regulator ICP4	50313321	YP_053124.1	776	30711,9
DLHEPQTLA	9	tegument host shutoff protein	50313260	YP_053064.1	268	31473,0
CAAYCPPELV	10	transcriptional regulator ICP4	50313321	YP_053124.1	1013	31572
NAPLLDLSV	9	envelope glycoprotein C	50313257	YP_053061.1	218	32081,1
PPPLPKPVF	9	myristylated tegument protein	50313292	YP_053095.1	43	32422,4
DAPSGRSGF	9	envelope glycoprotein E	50313315	YP_053118.1	505	32568,8
VALLCRQFV	9	DNA packaging tegument protein UL17	50313287	YP_053091.1	168	33633,5
RLPRLETPL	9	helicase-primase helicase subunit	50313298	YP_053101.1	81	34401,0
DAMAAGNAL	9	transcriptional regulator ICP4	50313321	YP_053124.1	537	34492
NPAGLETAA	9	transcriptional regulator ICP4	50313321	YP_053124.1	707	34656
RVLSWARDY	9	tegument protein UL37	50313264	YP_053068.1	623	34888,7
SVATSHPGI	9	large tegument protein	50313265	YP_053069.1	252	37098,2
HGRSLRLPF	9	helicase-primase primase subunit	50313248	YP_053052.1	855	37394,4
VLVEGRAWT	9	large tegument protein	50313265	YP_053069.1	67	37876,9
SAFVPITEV	9	tegument protein UL37	50313264	YP_053068.1	270	38531,3
SRLRLEPPG	9	envelope glycoprotein H	50313281	YP_053085.1	284	38682,2
MLLGGRTAY	9	envelope glycoprotein K	50313247	YP_053051.1	1	39597,0
PEFATGYNI	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	454	39969,8

CAPPWRPDV	9	regulatory protein ICP22	50313320	YP_053123.1	87	40654,8
AVLTSSTLF	9	transcriptional regulator ICP4	50313321	YP_053124.1	1241	40681
RPPGGRRDV	9	virion protein US2	50313309	YP_053112.1	280	41981,0
VPPDNHPGF	9	envelope glycoprotein D	126801894	YP_053116.2	279	42242,9
PPQDTCPPT	9	tegument host shutoff protein	50313260	YP_053064.1	139	43504,6
DAPLSTPAVV	10	transcriptional regulator ICP4	50313321	YP_053124.1	57	43524
FQPVLRRRF	9	single-stranded DNA-binding protein	50313272	YP_053076.1	714	43557,1
TRAPRPSAF	9	transcriptional regulator ICP4	50313321	YP_053124.1	648	43796
RLPFLYPEV	9	virion protein V67	50313318	YP_053121.1	141	44816,9
IVDDARGPF	9	capsid maturation protease	50313276	YP_053080.1	67	45282,7
PAAASCAPGV	10	transcriptional regulator ICP4	50313321	YP_053124.1	397	45520
DLLYGSQNMF	10	transcriptional regulator ICP4	50313321	YP_053124.1	44	45873
AATAAEREL	9	large tegument protein	50313265	YP_053069.1	1176	46217,1
PALFPTTPL	9	tegument serine/threonine protein kinase	50313290	YP_053094.1	506	46310,9
AAPILNTLF	9	large tegument protein	50313265	YP_053069.1	1336	47077,1
TAPRFARFV	9	DNA packaging protein UL32	50313269	YP_053073.1	394	47259,7
FALTTSVLF	9	envelope glycoprotein M	50313293	YP_053096.1	40	47329,9
TVPDPMPTT	9	envelope glycoprotein I	50313314	YP_053117.1	171	48596,4
GAPAPAPASA	10	transcriptional regulator ICP4	50313321	YP_053124.1	272	48690
IPPRLYTDL	9	envelope glycoprotein H	50313281	YP_053085.1	605	49188,9
TAATAGTEV	9	virion protein V67	50313318	YP_053121.1	7	49211,1
SVNDHLPGV	9	envelope glycoprotein C	50313257	YP_053061.1	355	51741,3
LMHAAGTEA	9	DNA packaging protein UL32	50313269	YP_053073.1	506	52086,2
EPADLGLAG	9	serine/threonine protein kinase US3	50313310	YP_053113.1	235	55815,1
FPPRFMQPI	9	small capsid protein	50313266	YP_053070.1	108	55890,8
AAPPPRVPT	9	tegument protein VP22	50313252	YP_053056.1	132	55970,8
CPPELVAEL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1017	56440
KDAGLLLRR	9	DNA packaging protein UL32	50313269	YP_053073.1	596	57265,1
PAPSPAPGLA	10	transcriptional regulator ICP4	50313321	YP_053124.1	105	57559
EPLVLPITEA	10	transcriptional regulator ICP4	50313321	YP_053124.1	1390	57658
CAAYCPPEL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1013	57735
FPEAWRPALT	10	transcriptional regulator ICP4	50313321	YP_053124.1	1031	58175
RLMGGYPEV	9	tegument protein UL37	50313264	YP_053068.1	943	59695,0
LPFDLNQTV	9	nuclear egress lamina protein	50313270	YP_053074.1	103	60598,3
HPLKLRAAG	9	virion protein US2	50313309	YP_053112.1	181	60991,7
THDILQLRF	9	large tegument protein	50313265	YP_053069.1	1905	60993,8
SRAVPRADV	9	large tegument protein	50313265	YP_053069.1	1655	61868,3
DVFSLLYAV	9	DNA packaging terminase subunit 2	50313273	YP_053077.1	756	61876,5
SLPSKQMVF	9	tegument serine/threonine protein kinase	50313290	YP_053094.1	289	62243,2
DMAYPRDEAG	10	transcriptional regulator ICP4	50313321	YP_053124.1	579	62520
HPGPSPGVA	9	transcriptional regulator ICP4	50313321	YP_053124.1	1451	63418,3
PKFDLGRNF	9	envelope glycoprotein B	50313274	YP_053078.1	158	64165,9
DYLALACGF	9	single-stranded DNA-binding protein	50313272	YP_053076.1	453	64271,8
VALDSDTETC	10	transcriptional regulator ICP4	50313321	YP_053124.1	147	65394
DDIDELAEY	9	large tegument protein	50313265	YP_053069.1	3149	67257,7
NVPAAPNPT	9	helicase-primase primase subunit	50313248	YP_053052.1	276	-
DINDPRSVV	9	large tegument protein	50313265	YP_053069.1	3329	-

124

LAALDKIET	9	large tegument protein	50313265	YP_053069.1	1246	-
LPKRRRPPW	9	large tegument protein	50313265	YP_053069.1	319	-
RPNIGLKRT	9	small capsid protein	50313266	YP_053070.1	99	-
DLAGLQPSV	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	743	-
DVATAAAEV	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	983	-
DWLAMRKAV	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	781	-
KVLDPVSGF	9	DNA polymerase catalytic subunit	50313271	YP_053075.1	701	-
DMARVSMEV	9	single-stranded DNA-binding protein	50313272	YP_053076.1	765	-
PCSAPTSVV	9	single-stranded DNA-binding protein	50313272	YP_053076.1	606	-
DVIDGMLVG	9	DNA packaging terminase subunit 2	50313273	YP_053077.1	379	-
TMALEYAEV	9	DNA packaging terminase subunit 2	50313273	YP_053077.1	85	-
AVSSLQQEV	9	capsid maturation protease	50313276	YP_053080.1	493	-
HPRNLPPET	9	DNA packaging tegument protein UL17	50313287	YP_053091.1	448	-
DAEEPIYDV	9	envelope glycoprotein M	50313293	YP_053096.1	416	-
LVTASLPQT	9	envelope glycoprotein M	50313293	YP_053096.1	51	-
VAFLYYTTL	9	envelope glycoprotein M	50313293	YP_053096.1	103	-
EVPGLPPSF	9	tegument protein UL7	50313296	YP_053099.1	53	-
DMMTAATAG	9	virion protein V67	50313318	YP_053121.1	4	-
DVNRLAGDV	9	regulatory protein ICP22	50313320	YP_053123.1	94	-
AGAVEPSEL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1360	-
ALSLYELRD	9	transcriptional regulator ICP4	50313321	YP_053124.1	715	-
APDLYDFIE	9	transcriptional regulator ICP4	50313321	YP_053124.1	9	-
AQPASPAPS	9	transcriptional regulator ICP4	50313321	YP_053124.1	100	-
ARSPQAAAS	9	transcriptional regulator ICP4	50313321	YP_053124.1	892	-
CCRATQTNA	9	transcriptional regulator ICP4	50313321	YP_053124.1	872	-
DEAGRPDSL	9	transcriptional regulator ICP4	50313321	YP_053124.1	585	-
FGELAASGP	9	transcriptional regulator ICP4	50313321	YP_053124.1	1064	-
GEAVSLLAP	9	transcriptional regulator ICP4	50313321	YP_053124.1	795	-
GGEPFSGSA	9	transcriptional regulator ICP4	50313321	YP_053124.1	816	-
ISQYHIYMR	9	transcriptional regulator ICP4	50313321	YP_053124.1	1211	-
KGGLSALLA	9	transcriptional regulator ICP4	50313321	YP_053124.1	1125	-
QNMFGVDDA	9	transcriptional regulator ICP4	50313321	YP_053124.1	50	-
QQHQPQKAR	9	transcriptional regulator ICP4	50313321	YP_053124.1	885	-
RSPHRLPKD	9	transcriptional regulator ICP4	50313321	YP_053124.1	329	-
RSPQAAASP	9	transcriptional regulator ICP4	50313321	YP_053124.1	893	-
VLYSPLQDE	9	transcriptional regulator ICP4	50313321	YP_053124.1	1093	-
VSTEDWPQD	9	transcriptional regulator ICP4	50313321	YP_053124.1	1199	-
RPDSLAAVA	9	transcriptional regulator ICP4	50313321	YP_053124.1	589	-
TVPVFAPPL	9	membrane protein UL45	126801891	YP_053048.2	33	-
LLVTLRPAA	9	helicase-primase subunit	126801893	YP_053098.2	131	-
EAPRRAPRV	10	transcriptional regulator ICP4	50313321	YP_053124.1	1398	-
DAPLSTPAV	9	transcriptional regulator ICP4	50313321	YP_053124.1	57	-
FAPDLYDFI	9	transcriptional regulator ICP4	50313321	YP_053124.1	8	-
DDAPLSTPAV	10	transcriptional regulator ICP4	50313321	YP_053124.1	56	-
VALDSDTET	9	transcriptional regulator ICP4	50313321	YP_053124.1	147	-
PPPMGRVRY	9	transcriptional regulator ICP4	50313321	YP_053124.1	425	-
FPEAWRPAL	9	transcriptional regulator ICP4	50313321	YP_053124.1	1031	-

PAPSPAPGL	9	transcriptional regulator ICP4	50313321	YP_053124.1	105	-
EGPDRRGGF	9	transcriptional regulator ICP4	50313321	YP_053124.1	987	-
LGAWLRPVY	9	transcriptional regulator ICP4	50313321	YP_053124.1	1344	-
HAPRGHGSF	9	transcriptional regulator ICP4	50313321	YP_053124.1	515	-
PPARDGARF	9	transcriptional regulator ICP4	50313321	YP_053124.1	1056	-
CPPELVAELI	10	transcriptional regulator ICP4	50313321	YP_053124.1	1017	-
GPPARDGARF	10	transcriptional regulator ICP4	50313321	YP_053124.1	1055	-
DDAPLSTPA	9	transcriptional regulator ICP4	50313321	YP_053124.1	56	-
DGGDWCNPY	9	transcriptional regulator ICP4	50313321	YP_053124.1	1477	-
AANRWGLGAW	10	transcriptional regulator ICP4	50313321	YP_053124.1	1338	-
PLLTPSGDPW	10	transcriptional regulator ICP4	50313321	YP_053124.1	411	-
GTRAPRPSAF	10	transcriptional regulator ICP4	50313321	YP_053124.1	647	-
AFARLYPDSA	10	transcriptional regulator ICP4	50313321	YP_053124.1	1260	-
IDNQLFPEAW	10	transcriptional regulator ICP4	50313321	YP_053124.1	1026	-