Immunogenetics (2011) 63:275–290
DOI 10.1007/s00251-010-0502-8

ORIGINAL PAPER

Functional analysis of frequently expressed Chinese rhesus macaque MHC class I molecules Mamu-A1*02601 and Mamu-B*08301 reveals HLA-A2 and HLA-A3 supertypic specificities

Scott Southwood · Christopher Solomon · Ilka Hoof · Richard Rudersdorf · John Sidney · Bjoern Peters · Angela Wahl · Oriana Hawkins · William Hildebrand · Bianca R. Mothé · Alessandro Sette

Received: 28 September 2010 / Accepted: 7 December 2010 / Published online: 28 January 2011

Abstract The Simian immunodeficiency virus (SIV)-infected Indian rhesus macaque (*Macaca mulatta*) is the most established model of HIV infection and AIDS-related research, despite the potential that macaques of Chinese origin is a more relevant model. Ongoing efforts to further characterize the Chinese rhesus macaques' major histocompatibility complex (MHC) for composition and function should facilitate greater utilization of the species. Previous studies have demonstrated that Chinese-origin *M. mulatta* (Mamu) class I alleles are more polymorphic than their Indian counterparts, perhaps inferring a model more representative of human MHC, human leukocyte antigen (HLA). Furthermore, the Chinese rhesus macaque class I allele Mamu-*A1*02201*, the most frequent allele thus far identified, has recently been characterized and shown to be an HLA-B7 supertype analog, the most frequent supertype in human populations. In this study, we have characterized two additional alleles expressed with high frequency in Chinese rhesus macaques, Mamu-*A1*02601* and Mamu-*B*08301*. Upon the development of MHC–peptide-binding assays and definition of their associated motifs, we reveal that these Mamu alleles share peptide-binding characteristics with the HLA-A2 and HLA-A3 supertypes, respectively, the next most frequent human supertypes after HLA-B7. These data suggest that Chinese rhesus macaques may indeed be a more representative model of HLA gene diversity and function as compared to the species of Indian origin and therefore a better model for investigating human immune responses.

Keywords HLA supertype · MHC · Peptide-binding motif · Rhesus macaque

Electronic supplementary material The online version of this article (doi:10.1007/s00251-010-0502-8) contains supplementary material, which is available to authorized users.

S. Southwood · C. Solomon · J. Sidney · B. Peters · B. R. Mothé · A. Sette (✉)
Department of Vaccine Discovery,
La Jolla Institute for Allergy and Immunology,
La Jolla, CA 92037, USA
e-mail: alex@liai.org

C. Solomon · B. R. Mothé
Department of Biological Sciences,
California State University—San Marcos,
San Marcos, CA 92096, USA

I. Hoof
Department of Systems Biology, Center for Biological Sequence Analysis, Technical University of Denmark,
2800 Lyngby, Denmark

R. Rudersdorf
Department of Pathology and Laboratory Medicine,
University of Wisconsin—Madison,
Madison, WI 53706, USA

A. Wahl · O. Hawkins · W. Hildebrand
Department of Microbiology and Immunology,
The University of Oklahoma,
Oklahoma City, OK 73104, USA

Introduction

Non-human primates, notably rhesus macaques, are widely used as animal models in biomedical and immunological research. In particular, both the Indian-origin and Chinese-origin rhesus macaque models have contributed invaluable information to the study of disease pathogenesis and novel vaccine evaluation (Desrosiers 1990; Gardner and Luciw

2008; Haigwood 2009; ILAR 2003; Kindt et al. 1992; Persidsky and Fox 2007). However, these two distinct geographic populations have displayed decidedly different outcomes to infection despite their apparent physical similarity (Joag et al. 1994; Otting et al. 2005, 2007), perhaps attributive to genetic profile variations that affect immune responses. Specifically, upon Simian immunodeficiency virus (SIV) infection, Chinese rhesus macaques have a more delayed and prolonged progression to AIDS as compared to the animal populations of Indian origin (Ling et al. 2002).

A number of studies have shown that several macaque major histocompatibility complex (MHC; Mamu) class I alleles, including Mamu-*A*01* (Allen et al. 1998), Mamu-*B*17* (Mothe et al. 2002), and Mamu-*B*01* (Loffredo et al. 2005) among others (Loffredo et al. 2004; Sette et al. 2005), are expressed with high frequency in Indian rhesus macaque populations. Interbreeding of the Indian-origin animals in the USA since 1978, when the exportation of these animals from India was discontinued (Southwick and Siddiqi 1988), has likely played a significant role in this observation. The peptide-binding specificities for several of these and other Indian rhesus MHC allelic forms have been extensively characterized, leading to the identification of specific alleles which influence disease progression (Mothe et al. 2003; O'Connor et al. 2003; Yant et al. 2006; Loffredo et al. 2007) as well as the discovery of viral evasion from cytotoxic T lymphocyte (CTL) responses (Evans et al. 1999; Allen et al. 2000) in the SIV arena. Indeed, Indian rhesus macaques are the model most utilized in HIV- and AIDS-related research studies (Persidsky and Fox 2007; Patterson and Carrion 2005; Gardner and Luciw 2008; Watkins et al. 2008). However, the increased demand for these animals and, more importantly, the rapid progression to disease displayed after SIV infection of the Indian-origin populations (Ling et al. 2002) have underscored the advantages for developing alternative animal models.

Because of their relative accessibility, Chinese rhesus macaques are becoming more widely employed as non-human primate models in infectious disease research. They are utilized for the evaluation of vaccines and the study of immune responses in pathogen systems ranging from Marburg virus, Ebola virus, and influenza virus to the more well-studied SIV (Geisbert et al. 2007; Larsen et al. 2007; Carroll et al. 2008; Degenhardt et al. 2009; Ling et al. 2007, 2002). These animals, however, have not been characterized at the MHC loci to the same extent as their Indian counterparts. Studies to address this disparity have revealed a surprisingly high degree of MHC polymorphism (Otting et al. 2005, 2007, 2008; Karl et al. 2008; Ma et al. 2009; Wiseman et al. 2009; Ouyang et al. 2008). However, it is largely non-overlapping with Indian-origin macaques (Solomon et al. 2010). This polymorphism may be due to the diverse geographic origins from which the animals have been derived, comparable to human population distribution, suggesting that Chinese rhesus macaques may represent human leukocyte antigen (HLA) diversity more effectively than those of Indian origin.

HLA polymorphism and its function to bind a diverse array of antigenic peptides for CTL scrutiny have been well documented, as has the existence of HLA supertypes, groups of MHC molecules which share similar peptide-binding specificities (Bjorkman and Parham 1990; Maryanski et al. 1986; Parham et al. 1995; Sette and Sidney 1999; Sidney et al. 1995, a, b; Townsend et al. 2006). Previous studies have demonstrated CTL repertoire overlaps between humans and chimpanzees (Bertoni et al. 1998), as well as humans and Indian rhesus macaques (Loffredo et al. 2009), suggesting that HLA binding supertypes may extend to non-human primates. Recently, the peptide-binding specificity associated with the most frequent Chinese-origin allele, Mamu-*A1*02201*, has been characterized and shown to be an HLA-B7 supertype analog (Solomon et al. 2010), though it is the only one to date. In order for Chinese-origin macaques to become more valuable as animal models of infectious disease research, continued investigation and characterization of their MHC profile is necessary so cellular immunity and immune correlates of protection can be more accurately assessed.

We previously reported that 11 alleles were found in multiple Chinese rhesus animals in different cohorts/colonies, each with a cumulative frequency greater than or equal to 5.6%. Combined, these alleles provide coverage for approximately 68% of all Chinese rhesus animals studied thus far and therefore represent a logical target for further analysis and characterization (Solomon et al. 2010). Accordingly, herein we sought to characterize Mamu-*A1*02601* (6.7%) and Mamu-*B*08301* (5.8%), two of the most frequently expressed Chinese-origin class I alleles. We report the specific peptide-binding motifs associated with these allelic forms and utilize their respective motifs to map SIV-derived Mamu-A1*02601 and Mamu-B*08301 binding peptides.

Materials and methods

Creation of stable Mamu-A1*02601, Mamu-B*08301 transfectant cell lines

Stable MHC class I transfectants were produced in the MHC class I deficient EBV-transformed B-lymphoblastoid cell line 721.221. An expression construct was created for Mamu-*A1*02601* and Mamu-*B*08301* by sub-cloning a full-length allele transcript into separate pcDNA 3.1 vectors (Invitrogen). These constructs were then used to transfect

MHC class I-null 721.221 cells using an Amaxa Nucleofector II transfection machine (Lonza AG, Walkersville, MD, USA).

To produce secreted Mamu-*A1*02601* molecules in the context of endogenous ligand identification, α-chain cDNAs of Mamu-*A1*02601* were modified at the 3′ end by PCR mutagenesis to delete codons 5–7 encoding the transmembrane and cytoplasmic domains and to add a 30-bp tail encoding the ten amino acid rat very low density lipoprotein receptor (VLDLr), SVVSTDDDLA, for purification purposes (Hickman et al. 2000). sMHC-VLDLr were cloned into the mammalian expression vector pcDNA3.1 (Invitrogen); 721.221 cells were transfected with sMHC Mamu-A*26TVLDLr by electroporation. After 48 h incubation, cells were plated in 96-well plates (Falcon) in RPMI 1640 containing the antibiotic Geneticin. Transfectants were tested for production of sMHC molecules by a VLDLr-specific ELISA (Hawkins et al. 2008).

Mamu-A1*02601 endogenous ligand determination

Approximately 25 mg of Mamu-A*26TVLDLr molecules from the 721.221 cell line were purified over an affinity column composed of anti-VLDLr antibody (ATCC clone CRL-2197) coupled to CNBr activated Sepharose 4B (GE Healthcare, Piscataway, NJ, USA). sMHC molecules were then eluted in 0.2 N acetic acid, brought up to 10% acetic acid, and heated to 76°C for 10 min. Peptides were separated from heavy and light chains by ultra-filtration in a stirred cell with a 3-kDa molecular weight cutoff cellulose membrane (Millipore, Bedford, MA, USA). The peptide batch was flash frozen and lyophilized. The peptides were then reconstituted in 10% acetic acid.

Following isolation, 10% of the peptide pool was subjected to 14 rounds of N-terminal sequencing by Edman degradation. A motif was generated by calculating the fold increase of each amino acid over the prior round. A hierarchy was then determined based on the amino acid composition at each position (Falk et al. 1991).

Peptides were reverse-phase HPLC fractionated using a Jupiter Proteo C12 column (Phenomenex, Torrance, CA, USA) on a Paradigm MG4 system (Michrom Bioresources, Auburn, CA, USA). A standard CH_3CN gradient was employed to generate approximately 40 peptide-containing fractions. UV absorption was monitored at 215 nm. Peptide fractions were concentrated to dryness and reconstituted in 20 μl of nanospray buffer composed of 50% methanol, 50% H_2O, and 0.5% acetic acid. Nano-electrospray capillaries (Proxeon, Denmark) were loaded with 1 μl of each peptide fraction and infused at 1,100 V on a Q-Star Elite quadrupole mass spectrometer with a time of flight detector (Applied Biosystems, Foster City, CA, USA). Ion maps were generated for each fraction in a mass range of 300–1,200 amu. Using independent data acquisition for selection, ions (putative peptides) were fragmented by tandem mass spectrometry (MS/MS). An amino acid sequence was assigned using the publicly available, web-based MASCOT (Matrix Science Ltd., London, UK) and/or de novo sequencing.

Positional scanning combinatorial library and peptide synthesis

Positional scanning combinatorial libraries (PSCL) were synthesized as previously described (Pinilla et al. 1999). In the PSCL, each pool in the library contains randomized 9-mer peptides with one fixed residue at a single position. With each of the 20 naturally occurring residues represented at each position along a 9-mer backbone, the entire library consisted of 180 peptide mixtures.

Peptides utilized in screening studies were purchased as crude or purified material from Mimotopes (Minneapolis, MN, USA/Clayton, Victoria, Australia), Pepscan Systems B. V. (Lelystad, The Netherlands), A&A Labs (San Diego, CA, USA), Genescript Corporation (Piscataway, NJ, USA), or the Biotechnology Center at the University of Wisconsin—Madison (Madison, WI, USA). Peptides synthesized for use as radiolabeled ligands were synthesized by A&A Labs and purified to >95% homogeneity by reverse-phase HPLC. Peptide purity was determined with analytical reverse-phase HPLC and amino acid analysis, sequencing, and/or mass spectrometry. Peptides were radiolabeled utilizing the chloramine T method (Sidney et al. 2001b). Lyophilized peptides were re-suspended at 4–20 mg/ml in 100% DMSO, then diluted to required concentrations in PBS+0.05% (*v*/*v*) nonidet P40 (Fluka Biochemika, Buchs, Switzerland).

SIV peptide sequences were derived from the SIVmac239 sequence, GenBank accession M33262 (Kestler et al. 1990).

MHC purification for peptide-binding assays

HLA and Mamu class I MHC purification was performed by affinity chromatography using the W6/32 and/or B123.2 class I antibodies, as previously described (Sidney et al. 2001b, 2005; Loffredo et al. 2009). Protein purity, concentration, and depletion efficiency steps were monitored by SDS-PAGE.

Quantitative assays for peptide binding to detergent solubilized MHC class I molecules were based on the inhibition of binding of a high-affinity radiolabeled standard probe peptide and performed as detailed in prior studies (Loffredo et al. 2004; Schneidewind et al. 2008; Sidney et al. 2001b, 2005). Peptides were tested at six different concentrations covering a 100,000-fold dose range in three or more independent assays. Control wells to measure non-specific (background) binding were also

included. In each experiment, a titration of the unlabeled version of the radiolabeled probe was also tested as a positive control for inhibition.

The radiolabeled peptide utilized for the Mamu-A1*02601 assay was 3317.02 (sequence YLPTQQDVL), representing a sequence identified by Edman degradation and mass spectrometry analysis (described above). For Mamu-B*08301 assays, peptide 3317.04 (KSINKVYGK, an R9->K analog of Vaccinia B13R, an HLA-A3 supertype degenerate binder) was used. For each peptide, the concentration of peptide yielding 50% inhibition of the binding of the radiolabeled probe peptide (IC_{50}) was calculated. Under the conditions used, where [radiolabeled probe] < [MHC] and $IC_{50} \geq$ [MHC], the measured IC_{50} values are reasonable approximations of the true K_d values (Gulukota et al. 1997; Sette et al. 1994a; Cheng and Prusoff 1973).

Bioinformatic analysis

We performed analysis of the PSCL data as described previously (Sidney et al. 2008). Briefly, IC_{50} (nanomolars) values for each residue/position mixture were standardized as a ratio to the geometric mean IC_{50} (nanomolars) value of the entire set of 180 mixtures and further normalized to the average of libraries tested at each position. To identify predicted binders, all possible 9-mer peptides in SIVmac239 sequences were scored using the matrix values derived from the PSCL analyses of Mamu-A1*02601 and Mamu-B*08301. The final score for each peptide represents the product of the corresponding matrix values for each peptide residue–position pair. Peptides scoring among the top 3.0% (n=100) were selected for binding analysis.

Phylogenetic analysis

We assembled representative MHC class I sequences from humans, Chinese, and Indian rhesus macaques. Sequences were normalized to 1,068 nucleotides in length and aligned using the ClustalX program (Thompson et al. 2002). A phylogenetic tree was built with the neighbor-joining method (Saitou and Nei 1987) using the Tamura three-parameter distance model (Tamura 1992). One thousand bootstrap samples were analyzed to ensure reliable clustering.

Results

Determination of natural Mamu-A1*02601 ligands

Previous studies have demonstrated that the elution and characterization of naturally bound ligands is an effective method for determining the peptide-binding specificity of class I MHC molecules (Hickman-Miller et al. 2005; Kubo et al. 1994). Accordingly, an initial evaluation of the peptide-binding specificity of soluble Mamu-A1*02601 was determined by sequencing 30 different endogenously loaded Mamu-A1*02601 peptide ligands representing 14 different peptide fractions (Table 1). The majority of ligands (80%; 24 out of 30) were nine residues in length, with only four and two ligands identified of ten and 11 residues, respectively. This observation suggests that 9-mers represent the preferred size for peptide ligands bound to Mamu-A1*02601.

Next, peptides were aligned and at each position, the frequency of each residue was tabulated (Supplemental Table 1, online resource). At position 2, leucine (L) was found in 15 of the 30 ligands, and other aliphatic residues valine (V) and isoleucine (I) were present in ten and three ligands, respectively. The amide residue glutamine (Q) was found in 15 peptides at position 6, and related residues

Table 1 Endogenous Mamu-A1*02601 ligands identified by MS/MS sequencing analysis

Fraction	Ion	Length	Peptide sequence
53	457.7	9	ALGPTNKTL
53	508.7	9	RLPTSQSDL
53	536.2	9	LVADEQQRL
55	520.7	9	SVTLHQDQL
55	538.7	9	YVPEAQTRL
57	507.2	9	GLVVKQEQL
57	552.2	9	WVGESERQL
59	514.7	9	VVSVPQRIM
59	534.3	9	RVGERQPLL
61	508.2	9	VLIKTAEEL
61	515.7	9	SLEIGENRL
61	538.7	9	YLPTQQDVL
61	568.3	10	YAAISQVDRL
63	526.7	9	RVHIGQVIM
63	612.3	10	SLPNFKQNEF
65	541.2	9	LINITEHEL
65	554.2	9	FLIRESETL
67	488.7	9	SLPHPEALL
67	552.7	9	AVRIEQEFL
69	478.7	9	ASLLIQQAL
69	535.8	9	IVTILQDRL
71	485.7	9	GLDIPEVDL
71	490.2	9	HINIGALEL
71	505.2	9	HVSLVQLTL
71	590.8	9	FLRFQEEVL
73	541.8	10	ALPDIKVLTL
73	649.8	11	RLPTDTLQELL
77	564.3	10	AISILQQLEI
77	592.3	11	ALPSDLLENVL
79	516.7	9	ALPEIFTEL

glutamic acid (E) or asparagine (N) were found in seven ligands. At the C terminus, L was dominant (26 out of 30 peptides). The related aliphatic residues methionine (M) and I were also present at the C terminus. This information presents a preliminary motif with positions 2 and 6 and the C terminus, tentatively assigned as the main anchor positions based on ≥50% residue frequency.

Establishment of peptide-binding assays for Mamu-A1*02601 and Mamu-B*08301

To functionally characterize Mamu-A1*02601 further, MHC class I molecules expressed in single cell transfectants of 721.221 lines were purified by affinity chromatography. To establish a Mamu-A1*02601 binding assay, we took advantage of the natural ligand sequence information previously described. Accordingly, Mamu-A1*02601 natural ligands containing tyrosine (Y), an amino acid to which gamma-radioactive isotopes of iodine can attach, were synthesized and their capacity to bind purified MHC investigated. We found that the ligand YLPTQQDVL (peptide ID 3317.02) was associated with prominent binding to Mamu-A1*02601 (Fig. 1a). Significant binding was observed with as little as 700 pM of purified Mamu-A1*02601. The binding of YLPTQQDVL was also selective, since no significant binding was observed to purified Mamu-B*08301 molecules.

Endogenous ligands or defined epitopes for the Mamu-B*08301 molecule have not been reported. However, previous studies demonstrated that HLA supertype ligands also bind to MHC molecules expressed in other species, such as chimpanzees (*Pan troglodytes*), mice (*Mus musculus*; McKinney et al. 2000; Sette et al. 2005; Sidney et al. 2006), and, recently, Chinese rhesus macaques (Solomon et al. 2010), where the most frequent allele Mamu-*A1*02201* was shown to be an analog of the HLA-B7 supertype. Hierarchical clustering analysis performed by our group (data not shown) predicted that Mamu-B*08301 might be associated with an HLA-A3 supertypic peptide-binding specificity. Experiments examining direct binding for a panel of HLA-A3-supertype peptides to Mamu-B*08301 were performed. An analog of a *Vaccinia*-derived HLA-A3 supertype ligand (KSINKVYGK, peptide 3317.04) bound Mamu-B*08301 with significant counts and showed specificity, displaying negligible binding to Mamu-A1*02601 (Fig. 1b).

Furthermore, the binding for both alleles was specific at the level of inhibition by unlabeled ligands. Mamu-A1*02601 binding to the radiolabeled YLPTQQDVL ligand could be inhibited by an excess of unlabeled 3317.02 peptide with an IC_{50} value of approximately 0.2 nM. Mamu-B*08301 binding to the radiolabeled KSINKVYGK ligand could be inhibited by an excess of unlabeled 3317.04 with an IC_{50} value of approximately

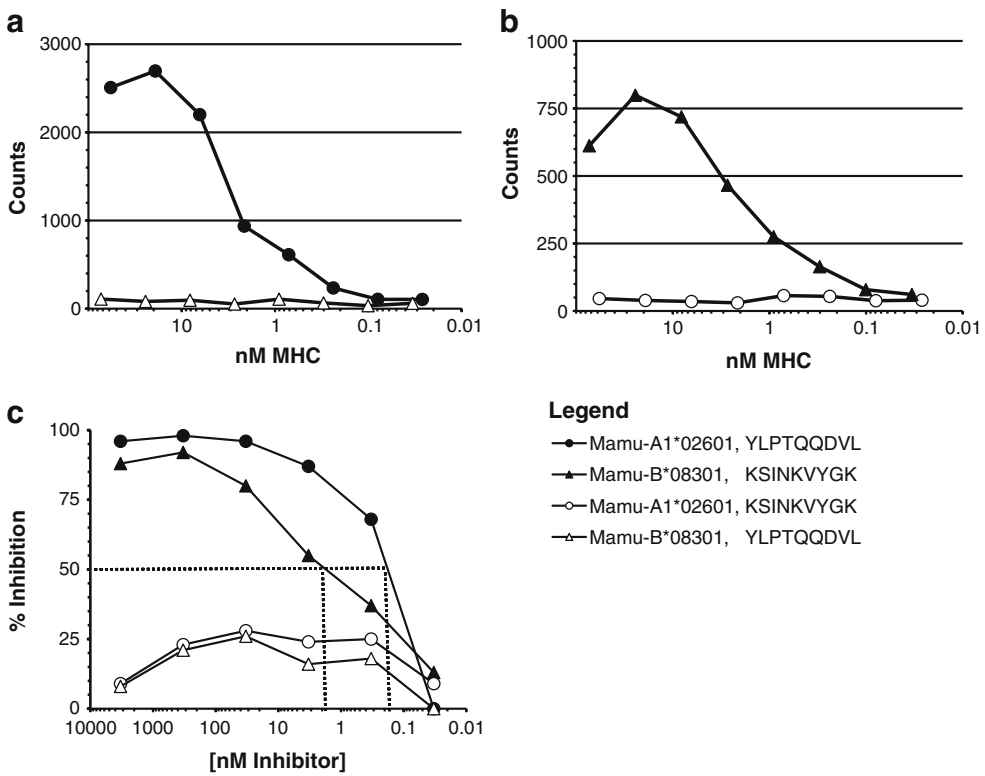

Fig. 1 The development of the Mamu-A1*02601 and Mamu-B*08301 MHC–peptide-binding assays. The endogenous Mamu-A1*02601 ligand YLPTQQDVL (peptide 3317.02) and HLA-A3 supertype ligand *Vaccinia* B13R analog KSINKVYGK (peptide 3317.04) were utilized as radiolabeled probes in direct binding dose titration experiments to ascertain binding potential to purified Mamu MHC molecules. **a** The radiolabeled ligand YLPTQQDVL binds Mamu-A1*02601 (*solid circle*), but not B*08301 (*open triangle*). **b** The radiolabeled ligand KSINKVYGK binds Mamu-B*08301 (*solid triangle*), but not A1*02601 (*open circle*). **c** Inhibition of radiolabeled ligand binding to Mamu-A1*02601 and Mamu-B*08301 by excess unlabeled 3317.02 and 3317.04 ligands in dose titration binding experiments demonstrated specificity of binding

2.0 nM (Fig. 1c). In conclusion, these results demonstrate that the binding assays are specific for the purified MHC molecule and thus enable detailed investigation of the peptide-binding specificity for both Mamu-A1*02601 and Mamu-B*08301.

Definition of Mamu-A1*02601 peptide-binding motif

Building upon our mass spectrometric analysis of Mamu-A1*02601 eluted ligands, we next tested the capacity of 9-mer PSCL to bind purified Mamu-A1*02601 MHC to further define the peptide-binding motif (Fig. 2). The binding capacity relative to the geometric mean for the entire PSCL (IC_{50}=2,438 nM) was determined for each positional/residue mixture and further normalized to the average of libraries tested at each position, as previously described (Sidney et al. 2008). For this analysis, we defined preferred residues as those that displayed 20-fold higher binding capacity than the position average and defined tolerated residues as those that displayed 4-fold higher binding. We defined a main binding anchor as a position in which at least half of the residues (≥10) are associated with binding capacity either 4-fold higher (in blue in Fig. 2) or lower (orange) than the position average. Furthermore, secondary anchors are defined as those positions associated with at least five residues (25%), but less than ten, employing the same binding criteria.

Accordingly, position 2 and the C terminus are the main anchor positions. At position 2, L, M, and V were overwhelmingly preferred above all other residues, while I and N were well tolerated. At the C terminus, residues L, I, and M were preferred and alanine (A), phenylalanine (F), glycine (G), and V were tolerated. Using the aforementioned criteria, position 1 is a secondary anchor with preferences for F and Y. Position 6, although failing to meet our secondary anchor criteria by one amino acid, has a very strong preference for Q, as displayed by a PSCL ratio of 41. These data are in overall agreement with the data observed with the natural ligands and were used to derive a Mamu-A1*02601 motif (Fig. 3). The observation that M was not identified in the natural ligand analysis, an exercise not intended to be comprehensive, but shown to be a preferred p2 anchor residue by PSCL, could in part be attributed to the residue's low observed frequency in nature. The fact that both L and M contribute comparable binding potential at an anchor position stems from their shared chemical similarity as hydrophobic, aliphatic residues.

Identification of Mamu-A1*02601 binding SIV-derived peptides

Next, we determined if the PSCL-based motifs could be used to identify Mamu-A1*02601 binding peptides. With Chinese rhesus macaques becoming an increasingly utilized animal model in SIV pathogenesis studies, we targeted the SIV proteome for these experiments (Joag et al. 1994; Ling et al. 2002; Trichel et al. 2002; Burdo et al. 2005; Monceaux et al. 2007; Degenhardt et al. 2009). The PSCL

Fig. 2 The relative influence of amino acids at each position on the binding to Mamu-A1*02601. The 9-mer positional scanning combinatorial library was tested for binding to Mamu-A1*02601. Values represent the ratio of the IC_{50} (nanomolars) of each pool at a given position relative to the geometric mean for the entire library (2,438 nM) and further normalized to the average of libraries tested at each position. Ratios of 0.25 or less are *highlighted in orange*; >4 are in *blue*

	Position								
Residue	1	2	3	4	5	6	7	8	9
A	3.6	1.1	2.2	1.0	1.5	8.1	1.1	6.0	8.0
C	3.3	0.49	0.94	0.33	1.3	0.30	2.0	0.56	0.33
D	0.051	0.13	0.48	0.28	0.19	1.2	1.3	0.82	0.23
E	0.17	0.066	0.077	0.081	0.34	0.32	0.022	0.13	0.17
F	6.9	1.5	0.98	1.8	1.8	0.44	2.6	1.4	7.5
G	2.6	0.23	2.1	0.41	1.2	0.34	1.4	0.77	4.6
H	1.2	0.17	0.26	0.96	0.42	1.9	0.46	0.58	0.16
I	0.75	17	1.7	3.4	1.9	0.53	2.9	5.9	31
K	2.8	0.19	1.3	0.39	0.47	0.17	0.22	0.32	0.17
L	1.5	25	0.77	3.7	0.72	0.45	1.4	1.2	110
M	3.2	26	1.0	2.2	1.2	3.1	2.5	0.61	29
N	0.63	4.5	1.7	0.90	0.57	12	0.85	0.30	0.26
P	0.090	0.41	0.85	1.5	1.9	0.37	1.3	1.3	0.098
Q	0.43	2.0	1.1	1.1	1.3	41	2.2	2.2	0.20
R	2.8	0.28	0.83	1.2	1.2	1.1	0.31	1.1	0.098
S	0.39	0.65	2.0	0.91	0.99	2.2	0.97	1.0	0.17
T	0.22	0.89	2.5	1.1	0.78	1.3	1.8	2.4	0.79
V	0.81	20	2.0	3.2	2.4	0.47	1.2	3.8	5.3
W	2.1	0.32	1.3	2.1	2.6	0.33	0.71	0.34	0.31
Y	4.4	0.46	0.91	1.5	1.3	0.37	3.7	1.1	1.0

Residues with 4-fold ratio increase	FY	**MLV** IN				**Q** NA		AI	**LIM** AFVG
			Position						
	1	2	3	4	5	6	7	8	9
Residues with 4-fold ratio decrease	DPTE	EDHKG	E	E	D	K	EK	E	RPHKESQD
Number of residues affecting binding per position	6	10	1	1	1	4	2	3	15
	2^o anchor	1^o anchor							1^o anchor

Fig. 3 Map of the Mamu-A1*02601 motif. Pictorial summary representation of the Mamu-A1*02601 PSCL matrix, indicating residues contributing positive or negative binding potential by position. The total number of residues affecting binding by position is summed, with a double-digit score indicating primary anchor position. Preferred residues at anchor positions, defined as residues showing a 20-fold increase in binding capacity versus the position average, are *highlighted in larger font*. Positions 2 and the C terminus are identified as main anchor positions

matrix was used to score all 9-mer peptides in the SIVmac239 proteome. Peptides scoring in the upper 3.0% range (n=100) were tested for binding capacity (Supplemental Table 2, online resource). Additionally, a set of 35 control peptides, ranging from 3.1% to 70% in rank (five peptides randomly selected within each 10% demarcation), was also tested.

Previous SIV epitope and in vivo T cell recognition studies have established 500 nM as an appropriate peptide-binding threshold of immunogenicity (Sette et al. 1994a, b; van der Most et al. 1998; Vitiello et al. 1996; Allen et al. 2001; Loffredo et al. 2004, 2005; Mothe et al. 2002; Sette et al. 2005). In total, 45 of the 100 "top 3% score" peptides bound Mamu-A1*02601 with an affinity of 500 nM or less (Table 2). Of those, 21 were high-affinity binders, possessing an $IC_{50} \leq 50$ nM. By contrast, none of the 35 lower-scoring control peptides resulted in similar affinities. Furthermore, 25 of the 45 binders with affinity ≤500 nM (56%) scored in the top 1% of predicted peptides, as well as 76% (16 of 21) of high-affinity binders.

Table 2 The efficiency of PSCL matrices in predicting Mamu-A1*02601 binders

Peptides tested[a]		Mamu-A1*02601 binding affinity (IC_{50}, nM)		
Prediction rank	n	<50	50–500	>500
Top 1.0%	33	16	9	l8
1.1–2.0%	33	3	12	18
2.1–3.0%	34	2	3	29
3.1–70%[b]	35	0	0	35
Total	135	21	24	90

[a] All possible 9-mer peptides in the SIVmac239 proteome were scored using the matrix values derived from the Mamu-A1*02601 PSCL analysis. The final score for each peptide represents the product of the corresponding matrix values for each peptide residue–position pair. Peptides scoring among the top 3.0% (n=100) were selected for characterization

[b] A set of 35 control peptides scoring in the 3.1–70% range, five peptides randomly selected within each 10% demarcation, was also examined

Mamu-A1*02601 and the HLA-A2 MHC molecules share overlapping binding repertoires

The results described above define a Mamu-A1*02601 motif that appeared to match the motif associated with HLA-A2 supertype alleles, which is characterized by a preference for aliphatic and hydrophobic residues at position 2 and at the C terminus (Sidney et al. 2001a). Based on this observation, we predicted that cross-reactivity between these human and macaque molecules may exist (Dzuris et al. 2000). To investigate this hypothesis, we tested the 45 top SIV-derived binders ($IC_{50} \leq 500$ nM) described above for binding to HLA-A2 supertype molecules (HLA-A*0201, HLA-A*0202, HLA-A*0203, HLA-A*0206, and HLA-A*6802). Twenty-seven of these 45 Mamu-A1*02601 binders (60%) bound at least one molecule of the HLA-A2 superfamily with an $IC_{50} \leq$ 500 nM (Fig. 4), with 18 (40%) being promiscuous binders, binding three or more HLA-A2 alleles. One peptide, FLIRQLIRL (Env.771), bound all five HLA-A2 supertype molecules examined.

HLA-A*0202 displayed the highest degree of cross-reactivity with Mamu-A1*02601. Of the 27 Mamu-A1*02601 binders that bound an A2-like allele, 23 (85%) bound HLA-A*0202. Additionally, 17 of the 18 promiscuous binders (94%) bound HLA-A*0202. HLA-A*0201 and HLA-A*0203 were also highly cross-reactive, as both

Fig. 4 Mamu-A1*02601 binders tested against HLA-A2 supertype alleles. SIVmac239-derived, Mamu-A1*02601 predicted peptides which had an $IC_{50} \leq 500\,nM (n = 45)$ were tested for binding capacity to HLA-A2 supertype molecules. Affinities *highlighted in blue* possess an IC_{50}<50 nM. Those affinities in the 50–500 nM range are *highlighted in green*. *Dashes* indicate binding affinity >20,000 nM

Sequence	Source	IC_{50} nM to purified molecules					
		Mamu-A1*02601	HLA-A*0201	HLA-A*0202	HLA-A*0203	HLA-A*0206	HLA-A*6802
LIRILQRAL	Vpr.61	0.00091	3748	362	492	1225	3373
YLCLIQKAL	Vpx.71	0.042	57	29	167	4345	--
MLTACQGVG	Gag.347	0.77	10,267	985	927	--	13,395
FLIRQLIRL	Env.771	0.83	1.2	0.64	1.1	2.2	489
IIVDSQYVM	Pol.702	1.0	832	215	5189	1136	6002
IVQQQQQLL	Env.561	1.0	4578	977	8337	9921	9547
GIVQQQQQL	Env.560	1.7	5818	871	8547	--	--
TVPWPNASL	Env.620	2.7	821	222	1878	6135	102
LVQYMDDIL	Pol.388	9.7	253	72	173	7831	12,322
FLATAGSAM	Env.535	12	161	8.5	7.7	1467	5558
RVVLQSKEL	Pol.408	13	6239	1457	2475	1300	2143
AMGAASLTL	Env.542	15	174	59	56	7521	--
WIKYIQYGV	Env.687	16	230	20	1.7	944	1.1
KNWMTQTLL	Gag.315	17	1979	4494	11,287	7753	11,442
KIGKEAIVI	Pol.581	21	16,219	1802	2651	--	--
SPAIFQYTM	Pol.364	26	--	3189	--	--	1087
LLTALGMSL	Pol.198	28	19	7.9	37	38	10,106
CVGDHQAAM	Gag.192	38	--	2691	4596	--	--
RMYNPTNIL	Gag.276	41	418	78	101	6560	381
GPKLKQWPL	Pol.226	42	--	4575	12,855	--	2009
CLIQKALFM	Vpx.73	45	788	1776	2353	7103	--
HVVWAANEL	Gag.33	72	10,997	1510	11,074	1031	19
RIVARQIVD	Pol.797	94	5910	4718	6544	10,467	2714
RILQRALFM	Vpr.63	104	--	6027	11,322	--	9890
FLGFLATAG	Env.532	106	309	200	23	3095	--
GVRLLAHVI	Pol.571	107	--	10,038	4844	--	--
NQLLIAILL	Env.6	117	5600	4955	4479	185	7638
YRRWIQLGL	Gag.263	119	14,959	5165	2224	--	--
KILSVLAPL	Gag.59	123	3.0	23	47	2.2	3183
LLSRVYQIL	Env.790	135	35	2.5	28	4578	4921
WMYRQQNPI	Gag.249	138	169	24	72	893	996
WLSTYAVRI	Vif.81	170	108	7.0	108	1097	14,675
VYQILQPIL	Env.794	184	9223	128	410	9321	--
RLIRGKMTL	Pol.489	217	151	72	83	908	--
VIYIVQMLA	Env.707	243	1438	2360	1176	293	4533
YNPQSQGVV	Pol.910	251	--	3863	2176	--	--
KIIIVAVHV	Pol.838	340	3883	3655	1542	1594	7353
GNYPVQQIG	Gag.133	343	4992	329	2251	13,973	14,992
IVIYIVQML	Env.706	354	382	644	1010	258	219
GFAAPQFSL	Pol.106	372	11,843	548	5309	--	15,237
MAQVHQGLM	Gag.449	431	18,523	4361	3975	--	7121
KLSPLCITM	Env.103	440	77	85	1158	85	--
QPAPQQGQL	Gag.220	452	4669	497	979	--	--
ILQRLSATL	Env.801	475	1.4	0.22	6.1	1381	6833
GLVFHSQPI	Env.327	491	234	9.7	34	4758	5130

alleles bound 18 Mamu-A1*02601 binders (67%); 16 of these cross-reactive peptides were shared between the two alleles.

Interestingly, the amino acid present at position 6 greatly influences cross-reactivity. More specifically, while Q at position 6 is strongly preferred for Mamu-A1*02601 binding, peptides possessing Q at position 6 are less likely to cross-react with HLA-A2 supertype molecules. Indeed, only 14% (three out of 21) of Mamu-A1*02601 binders with Q at position 6 were promiscuous HLA-A2 supertype

binders, while 63% (15 out of 24) of Mamu-A1*02601 binders that did not have Q at p6 bound three or more HLA-A2 supertype molecules.

Definition of Mamu-B*08301 peptide-binding motif

As was performed for Mamu-A1*02601, we tested the capacity of 9-mer PSCLs to bind Mamu-B*08301 molecules (Fig. 5). The binding affinity relative to the geometric mean for the entire PSCL (518 nM) was determined. Using the criteria outlined above, the C terminus was identified as the dominant anchor position, with the basic positively charged residues lysine (K) and arginine (R) being overwhelmingly preferred and residues F, L, and Y being tolerated. Unlike Mamu-A1*02601, a second main anchor position was not identified for Mamu-B*08301. Instead, position 2 was defined as a secondary anchor. In this position, the most preferred residues were N and serine (S), while threonine (T) was tolerated. These data were subsequently used to derive a Mamu-B*08301 motif (Fig. 6).

Identification of Mamu-B*08301 binding SIV-derived peptides

Repeating the exercise performed for Mamu-A1*02601, all 9-mer peptides in the SIVmac239 proteome were scored using the Mamu-B*08301 PSCL matrix. Peptides scoring in the upper 3.0% range (n=100) were tested for binding capacity (Supplemental Table 3, online resource), and a set of control peptides was included.

In total, 64 of the 100 "top 3% score" peptides bound Mamu-B*08301 with an affinity of 500 nM or less, and of those, 32 were high-affinity binders (Table 3). Furthermore, 30 of the 64 binders with affinity ≤500 nM (47%) scored in the top 1% of predicted peptides, as well as 72% (23 of 32) of high-affinity binders. By contrast, none of the 35 control peptides had an IC_{50}<1,900 nM.

Mamu-B*08301 and the HLA-A3 MHC molecules share overlapping binding repertoires

The results presented for Mamu-B*08301 depict a motif that closely resembles the one associated with HLA-A3 supertype alleles, which is characterized by a preference for basic residues at the C terminus and small and aliphatic residues at position 2 (Sidney et al. 1996a). Revisiting our hypothesis of the potential for cross-reactivity between human and macaque alleles, we tested the 64 top SIV-derived binders (IC_{50}≤500 nM) for binding to HLA-A3 supertype molecules (HLA-A*0301, HLA-A*1101, HLA-A*3101, HLA-A*3301, and HLA-A*6801). Fifty-two of the 64 Mamu-B*08301 binders (81%) bound at least one molecule of the HLA-A3 superfamily with an IC_{50}≤ 500 nM (Fig. 7).

HLA-A*3101 displayed the highest degree of cross-reactivity with Mamu-B*08301. Of the 52 Mamu-B*08301 binders that bound an A3-like allele, 39 (75%) bound HLA-

Fig. 5 The relative influence of amino acids at each position on the binding to Mamu-B*08301. The 9-mer positional scanning combinatorial library was tested for binding to Mamu-B*08301. Values represent the ratio of the IC_{50} (nanomolars) of each pool at a given position relative to the geometric mean for the entire library (518 nM) and further normalized to the average of libraries tested at each position. Ratios of 0.25 or less are *highlighted in orange*; >4 are in *blue*

	Position								
Residue	1	2	3	4	5	6	7	8	9
A	2.4	2.4	1.1	0.71	1.5	0.77	0.61	2.2	2.3
C	3.1	0.48	1.0	0.73	2.0	0.98	1.8	0.62	0.28
D	0.63	0.50	1.7	1.2	0.98	0.67	2.0	0.20	0.24
E	0.14	0.15	0.42	0.23	0.95	1.2	0.77	0.21	0.18
F	2.0	0.97	8.5	1.6	4.4	6.1	1.6	7.8	8.6
G	3.2	2.0	0.65	1.5	0.78	1.0	1.1	1.4	0.28
H	0.57	0.18	0.71	0.43	0.63	1.7	0.70	1.2	0.82
I	0.50	0.68	1.6	0.22	1.3	1.4	3.1	0.64	0.83
K	0.96	0.79	0.38	0.77	0.21	0.32	0.18	0.89	154
L	0.45	0.42	1.5	0.82	0.55	0.71	2.1	1.6	5.1
M	4.5	0.79	5.5	1.4	0.89	1.0	1.7	0.71	2.5
N	1.2	70	1.6	2.5	0.29	1.8	0.58	0.82	0.28
P	0.084	1.1	0.14	1.1	0.62	0.90	1.0	1.4	0.058
Q	2.8	0.92	1.1	0.91	3.1	1.0	0.62	1.7	0.062
R	2.1	0.21	0.19	0.64	0.26	0.48	0.17	1.2	126
S	2.0	22	1.2	5.8	1.3	0.76	1.3	0.66	0.33
T	0.70	6.8	0.74	0.67	0.56	0.56	0.98	0.95	0.16
V	0.31	1.3	0.86	0.93	0.77	1.0	0.74	0.68	1.3
W	1.4	0.19	0.72	1.8	5.7	1.5	1.7	0.57	0.31
Y	1.2	0.34	2.6	3.0	2.6	0.90	2.2	3.3	4.2

Residues with 4-fold ratio increase	M	**NS** T	FM	S	WF	F		F	**KR** FLY
			Position						
	1	2	3	4	5	6	7	8	9
Residues with 4-fold ratio decrease	PE	EHWR	PR	IE	K		KR	DE	PQDET
Number of residues affecting binding per position	3	7	4	3	3	1	2	3	10
		2° anchor							1° anchor

Fig. 6 Map of the Mamu-B*08301 motif. Pictorial summary representation of the Mamu-B*08301 PSCL matrix, indicating residues contributing positive or negative binding potential by position. The total number of residues affecting binding by position is summed, with a double-digit score indicating primary anchor position. Preferred residues at anchor positions, defined as residues showing a 20-fold increase in binding capacity versus the position average, are *highlighted in larger font*. The C terminus position is identified as the sole main anchor position

A*3101. HLA-A*0301 and HLA-A*6801 also displayed good cross-reactivity, as both alleles bound at least 26 Mamu-B*08301 binders. Additionally, 41% of Mamu-B*08301 binders (26 out of 64) were promiscuous HLA-A3 supertype binders, a percentage that is comparable to that seen between Mamu-A1*02601 and HLA-A2 (40%).

Evolutionary origin of Chinese rhesus macaque HLA-A2 and HLA-A3 functional analogy

The two common Chinese rhesus class I alleles described above are associated with HLA-A2 and HLA-A3 supertype specificities, which have not been detected in Indian rhesus macaques, although functional analysis of many of the most common alleles expressed in Indian-origin animals has been performed. To investigate the evolutionary origin of the functional analogy between the HLA and Chinese rhesus macaque MHC class I alleles in this study, we built a phylogenetic tree using representative MHC allele sequences.

For humans, we included one *HLA* allele (*HLA-A*01010101*, *HLA-A*02010101*, *HLA-A*03010101*, *HLA-A*24020101*, *HLA-B*070201*, *HLA-B*080101*, *HLA-B*270202*, *HLA-B*44020101*, *HLA-B*580101*, and *HLA-B*15010101*) from each of the known supertypes (A1, A2, A3, A24, B7, B8, B27, B44, B58, and B62 respectively) with an additional four HLA-B7 supertype alleles (*HLA-B*35010101*, *HLA-B*510101*, *HLA-B*530101*, and *HLA-B*5401*) and the four additional HLA-A2 supertype alleles (*HLA-A*0202*, *HLA-A*020301*, *HLA-A*020601*, and *HLA-A*68020101*) and HLA-A3 supertype alleles (*HLA-A*110101*, *HLA-A*310102*, *HLA-A*330101*, and *HLA-A*680101*) that were used in this study. For Indian rhesus macaques, we selected 14 sequences represented among the most common specificities (Boyson et al. 1996; Kaizu et al. 2007; Knapp et al. 1997; Loffredo et al. 2007; Voss and Letvin 1996). Similarly, the 14 Chinese rhesus macaque sequences included in the analysis were selected from the most frequent alleles (Solomon et al. 2010).

Table 3 The efficiency of PSCL matrices in predicting Mamu-B*08301 binders

Peptides tested[a]		Mamu-B*08301 binding affinity (IC_{50}, nM)		
Prediction rank	*n*	<50	50–500	>500
Top 1.0%	33	23	7	3
1.1–2.0%	33	4	13	16
2.1–3.0%	34	5	12	17
3.1–70%[b]	35	0	0	35
Total	135	32	32	71

[a] All possible 9-mer peptides in the SIVmac239 proteome were scored using the matrix values derived from the Mamu-B*08301 PSCL analysis. The final score for each peptide represents the product of the corresponding matrix values for each peptide residue–position pair. Peptides scoring among the top 3.0% (n=100) were selected for characterization

[b] A set of 35 control peptides scoring in the 3.1–70% range, five peptides randomly selected within each 10% demarcation, was also examined

Fig. 7 Mamu-B*08301 binders tested against HLA-A3 supertype alleles. SIVmac239-derived, Mamu-B*08301 predicted peptides which had an $IC_{50} \leq 500\,\text{nM} (n = 64)$ were tested for binding capacity to HLA-A3 supertype molecules. Affinities *highlighted in blue* possess an IC_{50}<50 nM. Those affinities in the 50–500-nM range are *highlighted in green*. *Dashes* indicate binding affinity >20,000 nM

Sequence	Source	IC_{50} nM to purified molecules					
		Mamu-B*08301	HLA-A*0301	HLA-A*1101	HLA-A*3101	HLA-A*3301	HLA-A*6801
TNYSGFMPK	Env.246	0.10	37	16	111	1.4	9.0
MSLNFPIAK	Pol.204	0.15	2.3	2.7	41	114	21
ANFASQEVK	Pol.886	0.18	1287	158	2467	10,380	188
QSYVDRFYK	Gag.295	0.23	15	5.0	11	77	8.2
YNTPTFAIK	Pol.264	0.38	442	1808	96	336	45
SNVKELVFK	Pol.784	0.41	2185	976	390	698	552
KNSKFKNFR	Pol.983	0.46	18,379	--	26	872	2904
ASNKPISNR	Tat.96	0.51	2.2	11	0.99	17	65
RNYVPCHIR	Env.427	0.58	2131	4257	1.3	52	586
ANTSSASNK	Tat.91	0.93	854	1306	2746	1959	2363
MSPSYVKYR	Vpx.62	1.4	139	194	29	58	2.5
GNNTGNESR	Env.197	1.8	287	253	76	473	1024
LNAWGCAFR	Env.606	1.8	369	422	45	78	15
SNCRTLLSR	Env.785	1.9	938	--	24	35	321
SSLSCEGQK	Nef.51	2.0	242	127	1091	398	614
RQTALFLLK	Pol.862	2.2	3.4	16	12	7637	454
KTGMLEMWK	Pol.19	2.2	6.3	11	150	12,587	1365
RNSVLSGKK	Gag.4	2.2	522	2911	58	1573	1851
SGFMPKCSK	Env.249	2.4	250	482	408	4538	124
RNTANQKPK	Env.414	3.1	3210	13,062	2567	4354	1857
RSRPSGDLR	Nef.9	6.1	316	8864	2.1	584	722
YNLTMKCRR	Env.305	8.3	3872	--	142	66	802
SQDNQWSYK	Pol.540	8.3	144	27	87	243	542
NTPTFAIKK	Pol.265	8.9	113	175	2938	2488	2.5
GANFPGLAK	Vif.200	16	22	14	448	2966	373
CNKQSKEGK	Pol.653	19	17,118	--	151	4080	2406
FKNFRVYYR	Pol.987	21	9899	14,370	6.0	6.3	121
AIDMSHFIK	Nef.116	21	89	15	3277	19,014	841
RTYIYWHGR	Env.285	24	4.0	27	0.89	66	46
LVFKFGLPR	Pol.789	36	18	57	1.0	15	2.1
FNGTRAENR	Env.277	39	--	--	848	328	1860
QNSRGDKQR	Vif.184	42	--	13,654	565	14,746	1612
ACYNTCYCK	Tat.49	53	225	2656	1437	8729	1756
VNNAEPGKR	Pol.343	56	1355	--	203	17,345	2822
SAEVAELYR	Env.483	64	12,730	781	1038	1765	4.3
KNTHTNGVR	Pol.565	65	1633	17,663	729	7140	7162
QRALFMHFR	Vpr.66	81	2373	11,314	19	207	24
YRYLCLIQK	Vpx.69	82	1355	2049	940	786	1828
KVCYVPHFK	Vif.38	102	11	205	3.0	2126	128
RGDKQRGGK	Vif.187	108	115	17,868	3573	--	--
MAVHCMNFK	Pol.945	111	59	194	13	1876	141
MGYELWPTK	Pol.438	124	1541	823	2842	874	7999
QKALFMHCK	Vpx.76	125	308	1194	799	3183	1454
KALFMHCKK	Vpx.77	125	70	698	68	10,686	3861
PSLQYLALK	Vif.146	138	1165	1425	2210	6485	1589
GNSSWPWQI	Env.758	139	2831	3214	443	6931	8674
APQFSLWRR	Pol.109	164	3448	1108	48	135	303
QILQPILQR	Env.796	168	683	1530	148	147	114
GFINTKEYK	Pol.161	175	936	2576	336	820	2634
HCQFCFLKK	Tat.62	184	68	1087	431	--	1848
ETDRWGLTK	Env.117	206	142	545	417	--	9697
CNDTNYSGF	Env.243	223	8638	8485	3587	7655	325
TTSTTASAK	Env.133	247	23	141	12,497	12,075	11
QAWCWFGGK	Env.341	295	2611	2276	1883	18,923	1179
GPCYGQMPR	Pol.29	303	1113	8290	401	5706	370
KGEGAVILK	Pol.1011	311	1206	1499	2366	--	5242
SGLSEEEVR	Nef.237	382	16,769	--	5232	14,259	3413
ANFPGLAKV	Vif.201	407	6314	--	1019	6336	--
RGPSCGSAK	Pol.71	410	313	1419	3544	--	761
CKFNMTGLK	Env.168	427	812	10,722	739	2258	5351
GKMDHVMAK	Gag.418	428	593	1202	44	279	275
QGMSPSYVK	Vpx.60	476	406	857	92	2126	8334
VTFMWTNCR	Env.388	482	761	4570	30	1383	1206
GIGGFINTK	Pol.158	484	344	179	5659	19,497	3201

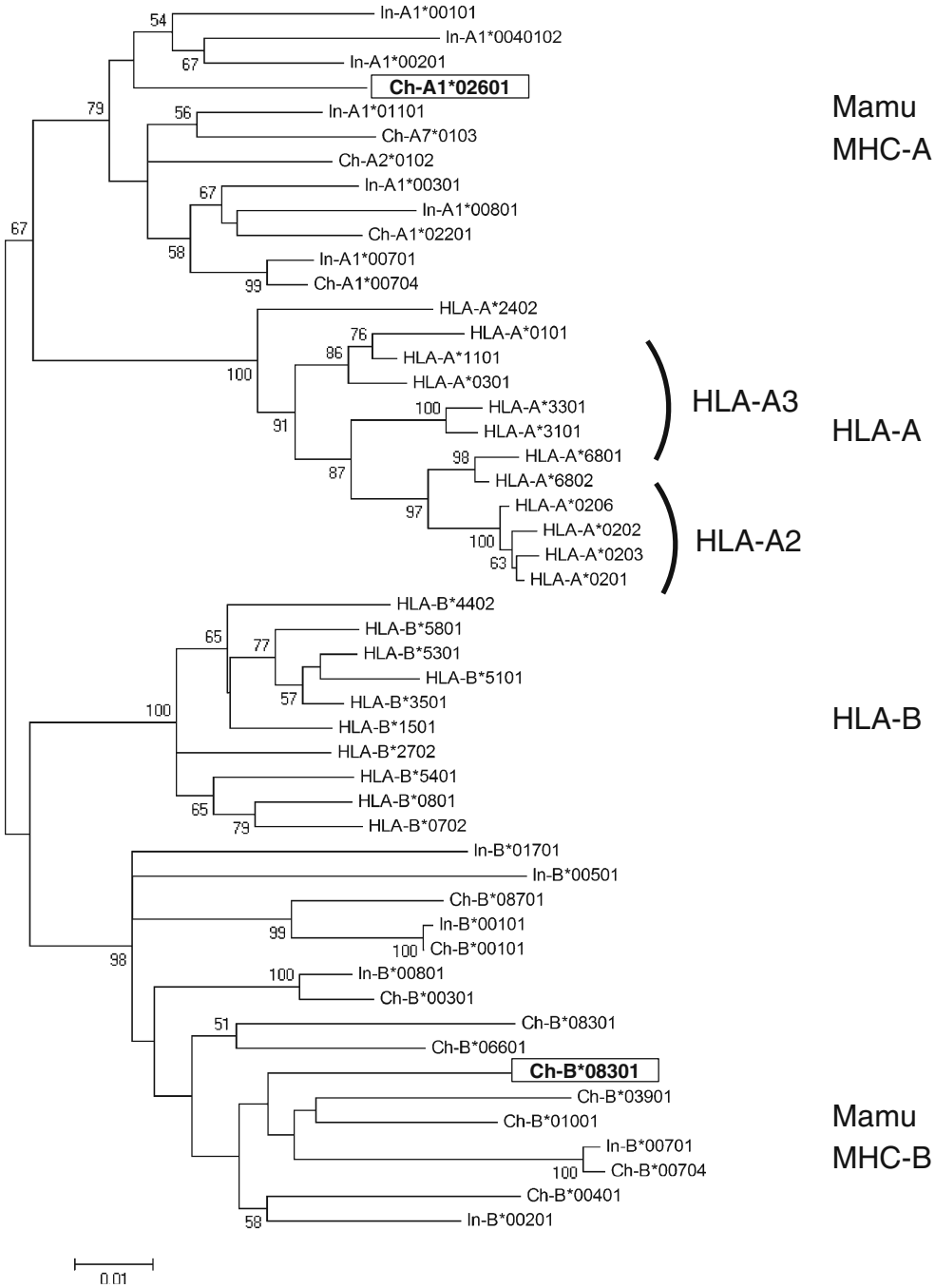

Fig. 8 Phylogenetic tree of HLA and Mamu MHC class I sequences. Phylogenetic analysis of 22 HLA, 14 Indian rhesus, and 14 Chinese rhesus MHC class I sequences is shown. Neighbor-joining tree created based on 1,068 aligned nucleotide sites. The percentage of bootstrap samples supporting the branch is shown (for values >50%). Mamu sequences derived from Indian animals are prepended with *In*. Sequences identified in Chinese animals (Solomon et al. 2010) are prepended with *Ch*. The Mamu allele *B*08301* demonstrating HLA-A3-like specificity appears in the Mamu-B group and not Mamu-A, indicating this motif specificity between humans and macaques likely did not arise through persistence of a common allele

Although Mamu-A1*02601 displays functional analogy to HLA-A2 supertype alleles and Mamu-B*08301 is associated with an HLA-A3 supertype peptide specificity, these *Mamu* alleles cluster separately from their respective HLA-A2 and HLA-A3 molecules (Fig. 8). Previous analysis of Mamu-A1*02201 showed a similar lack of sequence homology to its functional HLA cognate (Solomon et al. 2010). These results further support the hypothesis that these shared similarities in allele specificity have evolved independently in humans and macaques, most likely as a result of convergent evolution (Sette and Sidney 1999).

Discussion

Herein we report the peptide-binding motifs associated with the common Chinese-origin rhesus macaque class I molecules Mamu-A1*02601 and Mamu-B*08301. This is

the first description of a motif for Chinese-origin Mamu-B alleles and only the second motif for Mamu-A alleles, since until now Mamu-A1*02201 was the only Chinese-origin MHC molecule for which a peptide-binding motif had been described (Solomon et al. 2010). Thus, our study triples the number of alleles for which peptide-binding motifs are available.

The significance of our study is underscored by the fact that, while Mamu-*A1*02201* is the most frequently expressed class I allele in Chinese rhesus macaques, Mamu-*A1*02601* is the second most frequent allele and Mamu-*B*08301* is the most frequent Mamu-B allele, a distinction it shares with six other molecules (Solomon et al. 2010). These three alleles combined allow for potential coverage of approximately 20% of Chinese rhesus macaques used in biomedical research, irrespective of the geographical origin from China, thus significantly enhancing our knowledge of the functional MHC profile of these experimental animals.

Unexpectedly, we found that both of these molecules are associated with motifs that are overlapping with well-known HLA supermotifs. Specifically, Mamu-A1*02601 shares a high degree of cross-reactivity with the HLA-A2 supertype allele HLA-A*0202, as well as HLA-A*0201 and HLA-A*0203. Mamu-B*08301 is highly cross-reactive with HLA-A3 supertype alleles HLA-A*3101, HLA-A*0301, and HLA-A*6801. These results are even more remarkable in the context of the recently described HLA-B7 supertype specificity similarity of Mamu-A*02201 (Solomon et al. 2010). HLA-B7, HLA-A3, and HLA-A2 are the three most abundant supertypes in the human population. Separately, each of these three supertypes has a phenotypic frequency greater than 42% averaged across various ethnic groups. When these three supertypes are combined, over 86% of the human population is covered (Sette and Sidney 1999). Thus, each of the three common Chinese Mamu class I alleles thus far investigated are associated with a motif corresponding to one of the three most common HLA supertypes expressed in humans. It is possible that cross-reactivity with additional HLA or Mamu alleles will be identified, thus increasing the relevance and utility of the assays characterized herein for studies directed toward epitope discovery and characterization of the immune response to specific pathogens.

These observed similarities of *Mamu* with *HLA* could be explained by either common ancestry or convergent evolution (Sette et al. 2003). Previous studies have detected functional similarities between *Mamu* and *HLA* class I molecules, such as Mamu-B*08 with HLA-B*27 molecules. Notably, other similarities like Mamu-A*11 with HLA-B*44 and Mamu-A*22 with HLA-B*07 (Dzuris et al. 2000; Sette et al. 2005; Loffredo et al. 2009; Solomon et al. 2010) spanned across loci, arguing against the common ancestry hypothesis. In this study, we show that a *Mamu* molecule encoded by the B locus shares overlapping binding characteristics with HLA-A alleles, also arguing against common ancestry. This point was further and more formally demonstrated by phylogenetic analysis. Indeed, prior to this study, *Mamu* class I alleles, regardless of loci, have been purported to be functional HLA-B analogs only (Hickman-Miller et al. 2005), with rare associations linked to HLA-A thus far reported, even at the level of binding specificity. Our results present evidence of two *Mamu* class I molecules that share overlapping binding characteristics with HLA-A alleles.

The class I loci for both Indian- and Chinese-origin rhesus macaques are highly polymorphic, though Chinese alleles appear to be more polymorphic given the allele frequencies that have emerged in this population (Solomon et al. 2010). Additionally, the allelic variants are largely non-overlapping, possibly because of the geographical distance between India and China resulting in low probability of genetic exchange as demonstrated by mitochondrial DNA analysis of captive rhesus macaques (Kanthaswamy and Smith 2004; Satkoski et al. 2008). In this respect, the absence of the HLA-A2, HLA-A3, and HLA-B7 specificities from the seven common Indian rhesus macaque MHC class I alleles characterized to date (Loffredo et al. 2007; Knapp et al. 1997; Kaizu et al. 2007) is intriguing. This occurrence is likely the result of a combination of factors, including but not limited to a probable founder effect following the 1978 moratorium on importation of Indian-origin animals and the consequent breeding in the USA of captive populations.

Regardless of the evolutionary ramifications of MHC polymorphism and function, our findings have important practical implications because of the role of Chinese-origin rhesus macaques in biomedical research. The presence and identification of key HLA-like specificities in macaque populations of Chinese origin provides the scientific community valuable tools to evaluate disease pathogenesis and vaccine concepts in a setting more reflective of the global community, with broader human population coverage implications.

Acknowledgments We would like to thank A. Steen, S. Ngo, and C. Moore for MHC purification and binding assay assistance. This research is supported by the National Institutes of Health (grants R01 AI070902-01A2 to Alessandro Sette and Bianca R. Mothé and R15 AI064175-01 to Bianca R. Mothé).

References

Allen TM, Sidney J, del Guercio MF, Glickman RL, Lensmeyer GL, Wiebe DA, DeMars R, Pauza CD, Johnson RP, Sette A, Watkins DI (1998) Characterization of the peptide binding motif of a rhesus MHC class I molecule (Mamu-A*01) that binds an immunodominant CTL epitope from simian immunodeficiency virus. J Immunol 160(12):6062–6071

Allen TM, O'Connor DH, Jing P, Dzuris JL, Mothe BR, Vogel TU, Dunphy E, Liebl ME, Emerson C, Wilson N, Kunstman KJ, Wang X, Allison DB, Hughes AL, Desrosiers RC, Altman JD, Wolinsky SM, Sette A, Watkins DI (2000) Tat-specific cytotoxic T lymphocytes select for SIV escape variants during resolution of primary viraemia. Nature 407 (6802):386–390

Allen TM, Mothe BR, Sidney J, Jing P, Dzuris JL, Liebl ME, Vogel TU, O'Connor DH, Wang X, Wussow MC, Thomson JA, Altman JD, Watkins DI, Sette A (2001) CD8(+) lymphocytes from simian immunodeficiency virus-infected rhesus macaques recognize 14 different epitopes bound by the major histocompatibility complex class I molecule mamu-A*01: implications for vaccine design and testing. J Virol 75(2):738–749

Bertoni R, Sette A, Sidney J, Guidotti LG, Shapiro M, Purcell R, Chisari FV (1998) Human class I supertypes and CTL repertoires extend to chimpanzees. J Immunol 161(8):4447–4455

Bjorkman PJ, Parham P (1990) Structure, function, and diversity of class I major histocompatibility complex molecules. Annu Rev Biochem 59:253–288. doi:10.1146/annurev.bi.59.070190.001345

Boyson JE, Shufflebotham C, Cadavid LF, Urvater JA, Knapp LA, Hughes AL, Watkins DI (1996) The MHC class I genes of the rhesus monkey. Different evolutionary histories of MHC class I and II genes in primates. J Immunol 156(12):4656–4665

Burdo TH, Marcondes MC, Lanigan CM, Penedo MC, Fox HS (2005) Susceptibility of Chinese rhesus monkeys to SIV infection. AIDS 19(15):1704–1706

Carroll TD, Matzinger SR, Genesca M, Fritts L, Colon R, McChesney MB, Miller CJ (2008) Interferon-induced expression of MxA in the respiratory tract of rhesus macaques is suppressed by influenza virus replication. J Immunol 180(4):2385–2395

Cheng Y, Prusoff WH (1973) Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50% inhibition (I50) of an enzymatic reaction. Biochem Pharmacol 22(23):3099–3108

Degenhardt JD, de Candia P, Chabot A, Schwartz S, Henderson L, Ling B, Hunter M, Jiang Z, Palermo RE, Katze M, Eichler EE, Ventura M, Rogers J, Marx P, Gilad Y, Bustamante CD (2009) Copy number variation of CCL3-like genes affects rate of progression to simian-AIDS in rhesus macaques (*Macaca mulatta*). PLoS Genet 5(1):e1000346. doi:10.1371/journal.pgen.1000346

Desrosiers RC (1990) The simian immunodeficiency viruses. Annu Rev Immunol 8:557–578. doi:10.1146/annurev.iy.08.040190.003013

Dzuris JL, Sidney J, Appella E, Chesnut RW, Watkins DI, Sette A (2000) Conserved MHC class I peptide binding motif between humans and rhesus macaques. J Immunol 164 (1):283–291

Evans DT, O'Connor DH, Jing P, Dzuris JL, Sidney J, da Silva J, Allen TM, Horton H, Venham JE, Rudersdorf RA, Vogel T, Pauza CD, Bontrop RE, DeMars R, Sette A, Hughes AL, Watkins DI (1999) Virus-specific cytotoxic T-lymphocyte responses select for amino-acid variation in simian immunodeficiency virus Env and Nef. Nat Med 5(11):1270–1276

Falk K, Rotzschke O, Stevanovic S, Jung G, Rammensee HG (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature 351(6324):290–296. doi:10.1038/351290a0

Gardner MB, Luciw PA (2008) Macaque models of human infectious disease. ILAR J 49(2):220–255

Geisbert TW, Daddario-DiCaprio KM, Geisbert JB, Young HA, Formenty P, Fritz EA, Larsen T, Hensley LE (2007) Marburg virus Angola infection of rhesus macaques: pathogenesis and treatment with recombinant nematode anticoagulant protein c2. J Infect Dis 196(Suppl 2):S372–S381. doi:10.1086/520608

Gulukota K, Sidney J, Sette A, DeLisi C (1997) Two complementary methods for predicting peptides binding major histocompatibility complex molecules. J Mol Biol 267(5):1258–1267. doi:10.1006/jmbi.1997.0937

Haigwood NL (2009) Update on animal models for HIV research. Eur J Immunol 39(8):1994–1999. doi:10.1002/eji.200939576

Hawkins OE, Vangundy RS, Eckerd AM, Bardet W, Buchli R, Weidanz JA, Hildebrand WH (2008) Identification of breast cancer peptide epitopes presented by HLA-A*0201. J Proteome Res 7(4):1445–1457. doi:10.1021/pr700761w

Hickman HD, Batson CL, Prilliman KR, Crawford DL, Jackson KL, Hildebrand WH (2000) C-terminal epitope tagging facilitates comparative ligand mapping from MHC class I positive cells. Hum Immunol 61(12):1339–1346

Hickman-Miller HD, Bardet W, Gilb A, Luis AD, Jackson KW, Watkins DI, Hildebrand WH (2005) Rhesus macaque MHC class I molecules present HLA-B-like peptides. J Immunol 175 (1):367–375. doi:175/1/367

ILAR (2003) Demands for rhesus monkeys in biomedical research: a workshop report. ILAR J 44(3):222–238

Joag SV, Stephens EB, Adams RJ, Foresman L, Narayan O (1994) Pathogenesis of SIVmac infection in Chinese and Indian rhesus macaques: effects of splenectomy on virus burden. Virology 200 (2):436–446. doi:10.1006/viro.1994.1207

Kaizu M, Borchardt GJ, Glidden CE, Fisk DL, Loffredo JT, Watkins DI, Rehrauer WM (2007) Molecular typing of major histocompatibility complex class I alleles in the Indian rhesus macaque which restrict SIV CD8+ T cell epitopes. Immunogenetics 59 (9):693–703. doi:10.1007/s00251-007-0233-7

Kanthaswamy S, Smith DG (2004) Effects of geographic origin on captive *Macaca mulatta* mitochondrial DNA variation. Comp Med 54(2):193–201

Karl JA, Wiseman RW, Campbell KJ, Blasky AJ, Hughes AL, Ferguson B, Read DS, O'Connor DH (2008) Identification of MHC class I sequences in Chinese-origin rhesus macaques. Immunogenetics 60(1):37–46. doi:10.1007/s00251-007-0267-x

Kestler H, Kodama T, Ringler D, Marthas M, Pedersen N, Lackner A, Regier D, Sehgal P, Daniel M, King N et al (1990) Induction of AIDS in rhesus monkeys by molecularly cloned simian immunodeficiency virus. Science 248(4959):1109–1112

Kindt TJ, Hirsch VM, Johnson PR, Sawasdikosol S (1992) Animal models for acquired immunodeficiency syndrome. Adv Immunol 52:425–474

Knapp LA, Lehmann E, Piekarczyk MS, Urvater JA, Watkins DI (1997) A high frequency of Mamu-A*01 in the rhesus macaque detected by polymerase chain reaction with sequence-specific primers and direct sequencing. Tissue Antigens 50(6):657–661

Kubo RT, Sette A, Grey HM, Appella E, Sakaguchi K, Zhu NZ, Arnott D, Sherman N, Shabanowitz J, Michel H et al (1994) Definition of specific peptide motifs for four major HLA-A alleles. J Immunol 152(8):3913–3924

Larsen T, Stevens EL, Davis KJ, Geisbert JB, Daddario-DiCaprio KM, Jahrling PB, Hensley LE, Geisbert TW (2007) Pathologic findings associated with delayed death in nonhuman primates experimentally infected with Zaire Ebola virus. J Infect Dis 196 (Suppl 2):S323–S328. doi:10.1086/520589

Ling B, Veazey RS, Luckay A, Penedo C, Xu K, Lifson JD, Marx PA (2002) SIV(mac) pathogenesis in rhesus macaques of Chinese and Indian origin compared with primary HIV infections in humans. AIDS 16(11):1489–1496

Ling B, Veazey RS, Hart M, Lackner AA, Kuroda M, Pahar B, Marx PA (2007) Early restoration of mucosal CD4 memory CCR5 T cells in the gut of SIV-infected rhesus predicts long term non-progression. AIDS 21(18):2377–2385. doi:10.1097/QAD.0b013e3282f08b32

Loffredo JT, Sidney J, Wojewoda C, Dodds E, Reynolds MR, Napoe G, Mothe BR, O'Connor DH, Wilson NA, Watkins DI, Sette A (2004) Identification of seventeen new simian immunodeficiency virus-derived CD8+ T cell epitopes restricted by the high frequency molecule, Mamu-A*02, and potential escape from CTL recognition. J Immunol 173(8):5064–5076

Loffredo JT, Sidney J, Piaskowski S, Szymanski A, Furlott J, Rudersdorf R, Reed J, Peters B, Hickman-Miller HD, Bardet W, Rehrauer WM, O'Connor DH, Wilson NA, Hildebrand WH, Sette A, Watkins DI (2005) 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. J Immunol 175(9):5986–5997

Loffredo JT, Maxwell J, Qi Y, Glidden CE, Borchardt GJ, Soma T, Bean AT, Beal DR, Wilson NA, Rehrauer WM, Lifson JD, Carrington M, Watkins DI (2007) Mamu-B*08-positive macaques control simian immunodeficiency virus replication. J Virol 81(16):8827–8832. doi:10.1128/JVI.00895-07

Loffredo JT, Sidney J, Bean AT, Beal DR, Bardet W, Wahl A, Hawkins OE, Piaskowski S, Wilson NA, Hildebrand WH, Watkins DI, Sette A (2009) Two MHC class I molecules associated with elite control of immunodeficiency virus replication, Mamu-B*08 and HLA-B*2705, bind peptides with sequence similarity. J Immunol 182(12):7763–7775. doi:10.4049/jimmunol.0900111

Ma X, Tang LH, Qu LB, Ma J, Chen L (2009) Identification of 17 novel major histocompatibility complex-A alleles in a population of Chinese-origin rhesus macaques. Tissue Antigens 73(2):184–187. doi:10.1111/j.1399-0039.2008.01168.x

Maryanski JL, Pala P, Corradin G, Jordan BR, Cerottini JC (1986) H-2-restricted cytolytic T cells specific for HLA can recognize a synthetic HLA peptide. Nature 324(6097):578–579. doi:10.1038/324578a0

McKinney DM, Erickson AL, Walker CM, Thimme R, Chisari FV, Sidney J, Sette A (2000) Identification of five different Patr class I molecules that bind HLA supertype peptides and definition of their peptide binding motifs. J Immunol 165(8):4414–4422

Monceaux V, Viollet L, Petit F, Cumont MC, Kaufmann GR, Aubertin AM, Hurtrel B, Silvestri G, Estaquier J (2007) CD4+ CCR5+ T-cell dynamics during simian immunodeficiency virus infection of Chinese rhesus macaques. J Virol 81(24):13865–13875. doi:10.1128/JVI.00452-07

Mothe BR, Sidney J, Dzuris JL, Liebl ME, Fuenger S, Watkins DI, Sette A (2002) Characterization of the peptide-binding specificity of Mamu-B*17 and identification of Mamu-B*17-restricted epitopes derived from simian immunodeficiency virus proteins. J Immunol 169(1):210–219

Mothe BR, Weinfurter J, Wang C, Rehrauer W, Wilson N, Allen TM, Allison DB, Watkins DI (2003) Expression of the major histocompatibility complex class I molecule Mamu-A*01 is associated with control of simian immunodeficiency virus SIVmac239 replication. J Virol 77(4):2736–2740

O'Connor DH, Mothe BR, Weinfurter JT, Fuenger S, Rehrauer WM, Jing P, Rudersdorf RR, Liebl ME, Krebs K, Vasquez J, Dodds E, Loffredo J, Martin S, McDermott AB, Allen TM, Wang C, Doxiadis GG, Montefiori DC, Hughes A, Burton DR, Allison DB, Wolinsky SM, Bontrop R, Picker LJ, Watkins DI (2003) Major histocompatibility complex class I alleles associated with slow simian immunodeficiency virus disease progression bind epitopes recognized by dominant acute-phase cytotoxic-T-lymphocyte responses. J Virol 77(16):9029–9040

Otting N, Heijmans CM, Noort RC, de Groot NG, Doxiadis GG, van Rood JJ, Watkins DI, Bontrop RE (2005) Unparalleled complexity of the MHC class I region in rhesus macaques. Proc Natl Acad Sci USA 102(5):1626–1631. doi:10.1073/pnas.0409084102

Otting N, de Vos-Rouweler AJ, Heijmans CM, de Groot NG, Doxiadis GG, Bontrop RE (2007) MHC class I A region diversity and polymorphism in macaque species. Immunogenetics 59(5):367–375. doi:10.1007/s00251-007-0201-2

Otting N, Heijmans CM, van der Wiel M, de Groot NG, Doxiadis GG, Bontrop RE (2008) A snapshot of the Mamu-B genes and their allelic repertoire in rhesus macaques of Chinese origin. Immunogenetics 60(9):507–514. doi:10.1007/s00251-008-0311-5

Ouyang D, Xu L, Dai Z, Shi H, Zhang G, Zheng Y, He X (2008) Identification of major histocompatibility complex class I alleles in Chinese rhesus macaques. Acta Biochim Biophys Sin Shanghai 40(11):919–927

Parham P, Adams EJ, Arnett KL (1995) The origins of HLA-A, B, C polymorphism. Immunol Rev 143:141–180

Patterson JL, Carrion R Jr (2005) Demand for nonhuman primate resources in the age of biodefense. ILAR J 46(1):15–22

Persidsky Y, Fox H (2007) Battle of animal models. J Neuroimmune Pharmacol 2(2):171–177. doi:10.1007/s11481-006-9046-y

Pinilla C, Martin R, Gran B, Appel JR, Boggiano C, Wilson DB, Houghten RA (1999) Exploring immunological specificity using synthetic peptide combinatorial libraries. Curr Opin Immunol 11(2):193–202

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4(4):406–425

Satkoski J, George D, Smith DG, Kanthaswamy S (2008) Genetic characterization of wild and captive rhesus macaques in China. J Med Primatol 37(2):67–80. doi:10.1111/j.1600-0684.2007.00228.x

Schneidewind A, Brockman MA, Sidney J, Wang YE, Chen H, Suscovich TJ, Li B, Adam RI, Allgaier RL, Mothe BR, Kuntzen T, Oniangue-Ndza C, Trocha A, Yu XG, Brander C, Sette A, Walker BD, Allen TM (2008) Structural and functional constraints limit options for cytotoxic T-lymphocyte escape in the immunodominant HLA-B27-restricted epitope in human immunodeficiency virus type 1 capsid. J Virol 82(11):5594–5605. doi:10.1128/JVI.02356-07

Sette A, Sidney J (1999) Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics 50(3–4):201–212

Sette A, Sidney J, del Guercio MF, Southwood S, Ruppert J, Dahlberg C, Grey HM, Kubo RT (1994a) Peptide binding to the most frequent HLA-A class I alleles measured by quantitative molecular binding assays. Mol Immunol 31(11):813–822

Sette A, Vitiello A, Reherman B, Fowler P, Nayersina R, Kast WM, Melief CJ, Oseroff C, Yuan L, Ruppert J et al (1994b) The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes. J Immunol 153(12):5586–5592

Sette A, Sidney J, Livingston BD, Dzuris JL, Crimi C, Walker CM, Southwood S, Collins EJ, Hughes AL (2003) Class I molecules with similar peptide-binding specificities are the result of both common ancestry and convergent evolution. Immunogenetics 54(12):830–841. doi:10.1007/s00251-002-0530-0

Sette A, Sidney J, Bui HH, Del Guercio MF, Alexander J, Loffredo J, Watkins DI, Mothe BR (2005) Characterization of the peptide-binding specificity of Mamu-A*11 results in the identification of SIV-derived epitopes and interspecies cross-reactivity. Immunogenetics 57:53–68

Sidney J, del Guercio MF, Southwood S, Engelhard VH, Appella E, Rammensee HG, Falk K, Rotzschke O, Takiguchi M, Kubo RT et al (1995) Several HLA alleles share overlapping peptide specificities. J Immunol 154(1):247–259

Sidney J, Grey HM, Southwood S, Celis E, Wentworth PA, del Guercio MF, Kubo RT, Chesnut RW, Sette A (1996a) Definition of an HLA-A3-like supermotif demonstrates the overlapping peptide-binding repertoires of common HLA molecules. Hum Immunol 45(2):79–93. doi:0198-8859(95)00173-5

Sidney J, Southwood S, del Guercio MF, Grey HM, Chesnut RW, Kubo RT, Sette A (1996b) Specificity and degeneracy in peptide binding to HLA-B7-like class I molecules. J Immunol 157 (8):3480–3490

Sidney J, Southwood S, Mann DL, Fernandez-Vina MA, Newman MJ, Sette A (2001a) Majority of peptides binding HLA-A*0201 with high affinity crossreact with other A2-supertype molecules. Hum Immunol 62(11):1200–1216. doi:S0198-8859(01)00319-6

Sidney J, Southwood S, Oseroff C, del Guercio MF, Sette A, Grey HM (2001b) Measurement of MHC/peptide interactions by gel filtration. Curr Protoc Immunol Chapter 18:Unit 18 13. doi:10.1002/0471142735.im1803s31

Sidney J, Southwood S, Sette A (2005) Classification of A1- and A24-supertype molecules by analysis of their MHC-peptide binding repertoires. Immunogenetics 57(6):393–408. doi:10.1007/s00251-005-0004-2

Sidney J, Asabe S, Peters B, Purton KA, Chung J, Pencille TJ, Purcell R, Walker CM, Chisari FV, Sette A (2006) Detailed characterization of the peptide binding specificity of five common Patr class I MHC molecules. Immunogenetics 58(7):559–570

Sidney J, Assarsson E, Moore C, Ngo S, Pinilla C, Sette A, Peters B (2008) Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. Immunome Res 4:2. doi:10.1186/1745-7580-4-2

Solomon C, Southwood S, Hoof I, Rudersdorf R, Peters B, Sidney J, Pinilla C, Marcondes MC, Ling B, Marx P, Sette A, Mothe BR (2010) The most common Chinese rhesus macaque MHC class I molecule shares peptide binding repertoire with the HLA-B7 supertype. Immunogenetics. doi:10.1007/s00251-010-0450-3

Southwick CH, Siddiqi MF (1988) Partial recovery and a new population estimate of rhesus monkey populations in India. Am J Primatol 16(3):187–197

Tamura K (1992) The rate and pattern of nucleotide substitution in *Drosophila* mitochondrial DNA. Mol Biol Evol 9(5):814–825

Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. Curr Protoc Bioinformatics Chapter 2:Unit 2 3. doi:10.1002/0471250953.bi0203s00

Townsend AR, Rothbard J, Gotch FM, Bahadur G, Wraith D, McMichael AJ (2006) The epitopes of influenza nucleoprotein recognized by cytotoxic T lymphocytes can be defined with short synthetic peptides. 1986. J Immunol 176(9):5141–5150. doi:176/9/5141

Trichel AM, Rajakumar PA, Murphey-Corb M (2002) Species-specific variation in SIV disease progression between Chinese and Indian subspecies of rhesus macaque. J Med Primatol 31(4–5):171–178. doi:2x003

van der Most RG, Murali-Krishna K, Whitton JL, Oseroff C, Alexander J, Southwood S, Sidney J, Chesnut RW, Sette A, Ahmed R (1998) Identification of Db- and Kb-restricted subdominant cytotoxic T-cell responses in lymphocytic choriomeningitis virus-infected mice. Virology 240(1):158–167

Vitiello A, Yuan L, Chesnut RW, Sidney J, Southwood S, Farness P, Jackson MR, Peterson PA, Sette A (1996) Immunodominance analysis of CTL responses to influenza PR8 virus reveals two new dominant and subdominant Kb-restricted epitopes. J Immunol 157(12):5555–5562

Voss G, Letvin NL (1996) Definition of human immunodeficiency virus type 1 gp120 and gp41 cytotoxic T-lymphocyte epitopes and their restricting major histocompatibility complex class I alleles in simian-human immunodeficiency virus-infected rhesus monkeys. J Virol 70(10):7335–7340

Watkins DI, Burton DR, Kallas EG, Moore JP, Koff WC (2008) Nonhuman primate models and the failure of the Merck HIV-1 vaccine in humans. Nat Med 14(6):617–621. doi:10.1038/nm.f.1759

Wiseman RW, Karl JA, Bimber BN, O'Leary CE, Lank SM, Tuscher JJ, Detmer AM, Bouffard P, Levenkova N, Turcotte CL, Szekeres E Jr, Wright C, Harkins T, O'Connor DH (2009) Major histocompatibility complex genotyping with massively parallel pyrosequencing. Nat Med 15:1322–1326. doi:10.1038/nm.2038

Yant LJ, Friedrich TC, Johnson RC, May GE, Maness NJ, Enz AM, Lifson JD, O'Connor DH, Carrington M, Watkins DI (2006) The high-frequency major histocompatibility complex class I allele Mamu-B*17 is associated with control of simian immunodeficiency virus SIVmac239 replication. J Virol 80(10):5074–5077. doi:10.1128/JVI.80.10.5074-5077.2006