JVI Accepted Manuscript Posted Online 18 April 2018
J. Virol. doi:10.1128/JVI.01891-17

Amino acid substitutions within HLA-B*27-restricted T cell epitopes prevent recognition by hepatitis delta virus-specific CD8+ T cells

Hadi Karimzadeh,[a,b] Muthamia M. Kiraithe,[c] Anna D. Kosinska,[a,d] Manuel Glaser,[e] Melanie Fiedler,[b] Valerie Oberhardt,[c] Elahe Salimi Alizei,[c] Maike Hofmann,[c] Juk Yee Mok,[f] Melanie Nguyen,[f] Wim J.E. van Esch,[f] Bettina Budeus,[g] Jan Grabowski,[d,h] Maria Homs,[i] Antonella Olivero,[j] Hossein Keyvani,[k] Francisco Rodríguez-Frías,[i] David Tabernero,[i] Maria Buti,[i] Andreas Heinold,[l] Seyed Moayed Alavian,[m] Tanja Bauer,[a,d] Julian Schulze zur Wiesch,[n] Bijan Raziorrouh,[o] Daniel Hoffmann,[g] Antonina Smedile,[j] Mario Rizzetto,[j] Heiner Wedemeyer,[d,h] Jörg Timm,[p] Iris Antes,[e] Christoph Neumann-Haefelin,[c] Ulrike Protzer,[a,d] Michael Roggendorf [a,b,d,#]

Institute of Virology, Technical University of Munich/Helmholtz Zentrum München, Munich, Germany[a]; Institute of Virology, University Hospital of Essen, University of Duisburg-Essen, Essen, Germany[b]; University Hospital Freiburg, Department of Medicine II – University of Freiburg, Faculty of Medicine, Freiburg, Germany[c]; German Center for Infection Research (DZIF), Munich and Hannover Sites, Germany[d]; Center for Integrated Protein Science Munich at the Department of Biosciences, Technische Universität München, Freising, Germany[e]; Sanquin, Plesmanlaan 125, 1066 CX Amsterdam, the Netherlands[f]; Department of Bioinformatics, University of Duisburg-Essen, Essen, Germany[g]; Department of Gastroenterology, Hepatology and Endocrinology, Hannover Medical School, Hannover, Germany[h]; CIBERehd and Departments of Biochemistry/Microbiology and Hepatology, Vall d'Hebron Hospital, University Autònoma de Barcelona (UAB), Barcelona, Spain[i]; Department of Medical Sciences, University of Turin, Torino, Italy[j]; Department of Virology,

Iran University of Medical Sciences, Tehran, Iran[k]; Institute of Transfusion Medicine, University of Duisburg-Essen, University Hospital, Essen, Germany[l]; Baqiyatallah Research Center for Gastroenterology and Liver Diseases, Baqiyatallah University of Medical Sciences, Tehran, Iran[m]; Department of Medicine, Section Infectious Diseases, University Medical Center Hamburg-Eppendorf, Hamburg, Germany[n]; University Hospital Munich-Grosshadern, Department of Medicine II, Munich, Germany[o]; Institute of Virology, Heinrich-Heine-University, University Hospital, Duesseldorf, Germany[p]

Running Head: Hepatitis delta virus evades CD8+ T cell response

#Address correspondence to Michael Roggendorf
Institute of Virology, Technical University of Munich
Trogerstr. 30, 81675 München, Germany
Tel.: +49-89-4140-7360, Fax: -7444
E-Mail: michael.roggendorf@tum.de

Word count for the abstract: 250
Word count for the text: 6786

Abstract

Virus-specific CD8 T-cell response seems to play a significant role in the outcome of hepatitis delta virus (HDV) infection. However, the HDV-specific T-cell epitope repertoire and mechanisms of CD8 T-cell failure in HDV infection have merely been characterized. We therefore aimed to characterize HDV-specific CD8 T cell epitopes and the impact of viral mutations on immune escape. In this study, we predicted peptide epitopes binding the most frequent HLA types and assessed their HLA binding capacity. These epitopes were characterized in HDV-infected patients by intracellular IFN-γ staining. Sequence analysis of L-HDAg and HLA typing were performed in 104 patients. The impact of substitutions within epitopes on CD8 T cell response was evaluated experimentally and by *in silico* studies. We identified two HLA-B*27-restricted CD8 T-cell epitopes within L-HDAg. These novel epitopes are located in a relatively conserved region of L-HDAg. However, we detected molecular footprints within these epitopes in HLA-B*27-positive patients with chronic HDV infection. The variant peptides were not cross-recognized in HLA-B*27-positive patients with resolved HDV infection, indicating these substitutions represent viral escape mutations. Molecular modeling of HLA-B*27 complexes with the L-HDAg epitope and its potential viral escape mutations indicates that the structural and electrostatic properties of the bound peptides differ considerably at the T-cell receptor interface, which provides a possible molecular explanation for the escape mechanism. This viral escape from the HLA-B*27-restricted CD8 T-cell response correlates to chronic outcome of hepatitis D infection. T-cell failure resulting from immune escape may contribute to the high chronicity rate in HDV infection.

Importance

Hepatitis D virus (HDV) causes severe chronic hepatitis which affects 20 million people worldwide. Only a small number of patients is able to clear the virus, possibly mediated by virus-specific T cell response. Here we performed a systematic screen to define CD8 epitopes and investigated the role of CD8 T cells in outcome of hepatitis D and how they fail in eliminating HDV. Overall the number of epitopes identified was very low as compared to other hepatotropic viruses. We identified, two HLA-B*27-restricted epitopes in patients with resolved infection. In HLA-B*27 positive patients with chronic HDV infection, however, we detected escape mutations within these identified epitopes which could lead to viral evasion from immune responses. These findings support the evidence that HLA-B*27 is important for viral-specific CD8 T cell responses, similar to other viral infections. These results have implications for the clinical prognosis of HDV infection and for vaccine development.

Introduction

It is estimated that 15 to 20 million patients worldwide carry \HDV, a virus that causes the most severe form of viral hepatitis including fulminant hepatitis and a high rate of cirrhosis (1, 2). The HDV virion is composed of a circular single-stranded RNA consisting of around 1,700 bases and depends on hepatitis B virus (HBV) for progeny virus production and spread. The antigenomic open reading frame of HDV encodes the only viral protein, hepatitis delta antigen (HDAg) in two forms, the small and the large hepatitis delta antigen (S- and L-HDAg). L-HDAg contains the protein sequence of S-HDAg but has an additional 19-amino acid (aa) at the C-terminus (3, 4).

Two courses of HDV infection are known (5). HDV and HBV coinfection of patients shows a course similar to acute HBV mono-infection and the majority of the patients clear both HBV and HDV. HDV superinfection in chronic HBV infection in contrast results in a high rate of HDV persistence. Antibodies recognizing both HDV proteins are detected at low titers during acute infection and reach high levels during chronic infection, but they are not able to neutralize the virus (6). The efficacy of current therapies against HDV, e.g. with PEGylated interferon is very limited (7).

Liver damage, after superinfection of HBV carriers with HDV, results in transaminase elevation. This observation as well as the fact that HDV itself is not cytopathic (8) may indicate that cytotoxic T cells (CTL) are involved in the destruction of hepatocytes. T-cell responses against HDV antigens seem weak, but very little is known about T-cell responses controlling HDV infection and T-cell epitopes are poorly characterized. First studies in animal models have shown that a CD8 T-cell response against S- and L-HDAg can be generated. In mice, we demonstrated that immunization with a plasmid DNA vaccine expressing HDAg primed a functional CD4 and CD8 T-cell immune response against both

forms of delta antigens (9). These T cells prevented the development of L-HDAg -expressing tumors in 80–100% of immunized mice (9). Moreover, a DNA prime and adenoviral vector boost vaccination with L-HDAg protected woodchucks from HDV in the setting of a simultaneous infection with WHV and HDV (10). Nisini *et al*. demonstrated a polyspecific but weak CD4 T-cell response to HDAg in patients related to the resolution of HDV-induced disease activity (11). CD8 T-cell responses against two HLA-A*02:01 restricted HDAg epitopes have been described in patients with resolved HDV infection that were absent in patients with an actively replicating chronic infection (12).

In recent years, much has been learned about how human viruses establish chronic and long-lasting infections by escaping the T-cell immune response. For the clinically most relevant chronic viral infections with HIV, HBV and hepatitis C virus (HCV), it has been demonstrated that despite differences in the natural course of infection in humans, failure of T-cell responses is a typical phenomenon (13-15)**.** Viral immune escape variants may contribute to the failure of T-cell responses and allow viral persistence and high-level replication. CD8 T-cell responses are restricted by the host human leukocyte antigen (HLA) class-I molecules with an enormous allelic variation between individuals and across different populations. Since the repertoire of the CD8 T-cell response is dictated by the genetic HLA background of the individual host, diverse peptide epitopes derived from viral proteins maybe presented.

CD8 T cells have been shown to elicit a significant selection pressure in chronic viral infection. In HBV (14), HCV (16) and HIV (17, 18) infection, selection of viral escape variants has been reported. For HDV, a previous study suggested evidence for positive selection within the predicted HDV-epitopes restricted by HLA-A*02:01 (19); however, no in vitro experiments have been carried out confirming that immune escape functionally impair the virus-specific T cell response. The aim of this study therefore was (1) to

128 characterize the variability of the only HDV protein, L-HDAg, in a large cohort of patients,
129 (2) to identify HDAg-specific CD8 T-cell epitopes for frequent HLA alleles by *in silico* and
130 *in vitro* analysis, and (3) to evaluate whether an immune escape of HDV from CD8 T-cell
131 responses by mutation of relevant CD8 epitopes contributes to the persistence of HDV after a
132 superinfection of HBV carriers.

Results

HDV epitope prediction and MHC-binding capability of predicted epitopes in vitro

200 HDV genotype-1 sequences of L-HDAg, retrieved from the GenBank, were submitted to epitope prediction using two different online tools for the prediction of MHC class-I binding epitopes (implemented in the IEDB and SYFPEITHI databases). This prediction on the L-HDAg was done for frequent MHC class-I alleles in the European population (20), as well as HLA-B*27 which, despite its low frequency, has been shown to play a protective role against a number of serious human viral infections such as HCV and HIV (21-23). Two predicted HLA-B*27-restricted epitopes, i.e. L-HDAg$_{99-108}$ (RRDHRRRKAL) and L-HDAg$_{103-112}$ (RRRKALENKK), showed the best percentile scores (Table 1). Although, for the most of the common HLA alleles, no peptides reached average prediction scores which were comparable to the peptide epitopes known from other viruses, we selected 15 peptides with the best predicted binding scores to HLA-A*01, A*02, A*03, A*24, HLA-B*07 and B*27 to determine their binding affinity experimentally.

HLA-binding was assessed by UV-induced peptide exchange and compared to a known high affinity ligand of the respective HLA class-I molecules. None of the synthesized peptides predicted to bind to HLA-A*01, A*02, A*03, A*24 or HLA-B*07 bound with an affinity >25% compared to the respective positive control (Fig. 1). Only the two peptide ligands for HLA-B*27, aa 99-108 (RRDHRRRKAL) and 103-112 (RRRKALENKK), showed high binding affinities of 120 and 111.7%, respectively, relative to that of KRWIILGLNK, a well-known HLA-B*27-restricted HIV epitope which was used as positive control.

Detection of HDV-specific CD8 T cells in patients with resolved HDV infection

To determine if the predicted peptide epitopes would be recognized in HDV infection, we, in a first set of experiments, analyzed PBMCs of one HLA-B*27-positive and three HLA-B*27-

negative patients who had resolved HDV infection. Using an overlapping peptide library spanning the whole L-HDAg (Table 2) divided into 8 peptide pools (A to H), we detected a T-cell response only in the HLA-B*27-positive patient A (Fig. 2). The CD8+ T-cell response, shown by intracellular cytokine staining (ICS), was induced by peptides of pool D (0.53% IFN-γ+ CD8+ T cells compared to 0.04% relative to negative control peptides). Restimulation of the cells with the single peptides of pool D showed that the response was present only after stimulation with peptide D14. This 16-mer peptide (L-HDAg$_{98-113}$ ERRDHRRRKALENKKK) includes the sequences of two HLA-B*27:05 peptide ligands, L-HDAg$_{99-108}$ (RRDHRRRKAL) and L-HDAg$_{103-112}$ (RRRKALENKK), both of which we had predicted and proven to bind to this allele (Table 1, Fig. 1). Four HLA-B*27-negative patients with resolved HDV infection, five HLA-B*27-positive patients with chronic HDV infection (HDV RNA-positive), as well as two HLA-B*27-positive individuals without HBV, HCV and HDV infection (anti-HDV and HDV RNA-negative) used as controls showed no response to the 16-mer or to the single peptide epitopes (Fig. 3).

To confirm the CD8 T-cell response observed in patient A, we were able to gain access to two additional HLA-B*27-positive individuals (patient B and C) with resolved HDV infection (Fig. 2B,C), while PBMCs from patient A were no longer available. Stimulation of PBMCs from both patients B and C with the 16-mer peptide (98-113 ERRDHRRRKALENKKK) resulted in a similar IFN-γ response as had been the case with patient A. Interestingly, patient B recognized the HLA-B*27:05 ligand L-HDAg$_{99-108}$ (RRDHRRRKAL) (Fig. 2B) and patient C L-HDAg$_{103-112}$ (RRRKALENKK) (Fig. 2C) indicating a differential CD8 T-cell response in HLA-B*27-positive individuals. Our HLA-B*27+ patients were recruited from Germany, Italy and Spain, where the HLA-B*27 subtypes (alleles) B*27:05 and B*27:02 are prevalent. Since these two HLA-B*27 subtypes may influence targeting of virus-specific CD8+ T cell epitopes (24), we analyzed

presentation of the epitope L-HDAg$_{103-112}$ by HLA-B*27:05 and B*27:02. Of note, the epitope was presented by both HLA-B*27 subtypes (Fig. 4).

According to prediction and binding assay, in our hand, the two previously described HLA-A*02-restricted peptides (L-HDAg$_{26-34}$ and L-HDAg$_{43-51}$) (12) had low prediction scores and no binding capacities to the HLA-A*02 molecule *in vitro*. Nevertheless, we intended to confirm these two epitopes in 4 HLA-A*02 positive patients with resolved HDV infection. We observed no specific T cell response in PBMCs of all 4 patients after 10 days of culture (Fig. 5).

Molecular footprints for viral immune escape in HLA-B*27 epitopes L-HDAg$_{99-108}$ and L-HDAg$_{103-112}$

Several studies have demonstrated that viral mutations leading to sequence variation within T-cell epitopes are selected as a result of the immune pressure. These mutations can result in a failure of T-cell recognition and are maybe responsible for viral persistence. We therefore compared the coding sequence for L-HDAg by direct sequencing in a cohort of 104 patients with chronic HDV infection who were HLA-typed. Sequence alignment and subsequent phylogenetic analysis indicated that all isolates were HDV genotype 1 (data not shown). Within the 214 aa L-HDAg, we observed a high variability, e.g. at aa positions 9, 121 and 191, whereas aa 10, 99 and 115 were conserved in all isolates (data not shown). These analyses indicate that the candidate HLA-B*27-restricted epitopes (aa 99-108 and 103-112) are located in a relatively conserved region of the L-HDAg.

When comparing HLA-B*27-positive and -negative patients, we found a higher rate of substitutions within the verified HLA-B*27 epitopes L-HDAg$_{99-108}$ and L-HDAg$_{103-112}$ in HDV isolates from the six HLA-B*27 positive individuals (Fig. 6). Sequence analysis indicated that two amino acid substitutions, R105K and K106M, were significantly enriched

in isolates from HLA-B*27-positive patients (p=0.002). Two additional residues within this region, L-HDAg 100 and 112, showed some variations in both HLA-B*27 positive and negative patients which were not only present in genotype 1, but also in all 8 genotypes of HDV following comprehensive analysis of 621 available isolates from the GenBank (Fig. 7). We also analyzed amino acid sequences upstream of the L-HDAg$_{99\text{-}108}$ and downstream of the L-HDAg$_{103\text{-}112}$ in isolates from HLA-B*27 positive versus HLA-B*27 negative patients with chronic HDV infection. These analyses indicated no enrichment of mutations in the flanking regions of these two epitopes as described elsewhere (25). The L-HDAg sequence analyses was also extended to other 13 predicted epitopes restricted by HLA-A*01, A*02, A*03, A*24 or HLA-B*07 (Table 1). Importantly, for none of these candidate epitopes, a higher frequency of viral mutations was observed in isolates from patients positive for the respective HLA type compared to patients negative for these alleles. These results are consistent with those of prediction and binding assay.

Notably, two of the three HLA-B*27-positve patients (No.2 and 6 in Fig. 6) with wild type epitopes, resolved the HDV infection in the follow-up after more than 3 years of persistent infection.

To identify additional minor variants, next-generation sequencing was performed in HDV isolates from the HLA-B*27-positive patients and the six randomly selected HLA-B*27-negative patients. The analyses of 43,724 validated sequences demonstrated that the substitutions identified by direct sequencing were present in 100% of the haplotypes 4/5 HLA-B*27-positive, but 0/6 HLA-B*27-negative patients. Additional low frequency substitutions (0.3-0.8% of the sequences) in L-HDAg epitope coding region were observed, just in HLA-B*27-positive cases (Table 3). Haplotypes at position 112 showed a polymorphism K/R which was observed in both HLA-B*27-positive and negative cases

(Fig. 6). In two HLA-B*27-negative cases a mixture between 112K and R was observed in minor proportions (Table 3). Thus, distinct molecular footprints were identified in the L-HDAg coding sequence of HDV isolates in HLA-B*27-positive patients.

Impact of footprints within HLA-B*27 epitopes L-HDAg$_{99-108}$ and L-HDAg$_{103-112}$ on the CD8 T cell response

To investigate whether the amino acid variations detected in our patient cohort have an impact on the T-cell response, we tested the variant peptides for cross-recognition in the HLA-B*27+ patients with resolved HDV infection (Fig. 8). Further characterization of L-HDAg$_{103-112}$ using PBMCs from patient C indicated equal recognition of the 10-mer epitope and the corresponding 9-mer (L-HDAg$_{104-112}$: RRKALENKK) (Fig. 8A). Variation at the position 112 (K112R), observed in both groups of HLA-B*27-positive and -negative patients, did not influence CD8 T-cell responses. Therefore, we investigated the impact of R105K and K106M substitutions in both 10-mer and 9-mer epitopes with lysine (K) at position 112. While R105K substitutions affected the corresponding T-cell response in none of the epitopes, the K106M substitution completely impaired the specific T-cell response to 10-mer and 9-mer epitopes. Equivalent results were obtained when IFN-γ ELISpot instead of intracellular cytokine staining was performed as read-out (Fig. 9). This indicated that for patients recognizing the epitope L-HDAg$_{103-112}$ only the variation K106M is able to allow HDV to escape virus-specific immunity.

In addition to the variations R105K and K106M within the L-HDAg$_{99-108}$ epitope, we also observed variations at position 100, the second amino acid in that epitope, throughout all studied isolates (R**R**DHRRRKAL, R**Q**DHRRRKAL, R**E**DHRRRKAL and R**K**DHRRRKAL) (Fig. 6). To address the impact of these amino acid substitutions on T-cell recognition, we tested the peptide variants in HLA-B*27-positive patient B, with resolved

HDV infection, for cross-recognition (Fig. 8B). This assay confirmed that the L-HDAg$_{99-108}$ induced CD8 T-cell responses irrespective of aa 100 variations (R[R/Q/E/K]DHRRRKAL) (Fig. 8B, 4 left panels). However, IFN-γ production was not observed when PBMCs were stimulated with peptides variant either at positions 105 (R105K) or at position 106 (K106M) (Fig. 8B, 4 right panels). This matched our finding that amino acid substitutions R105K and K106M were exclusively observed in HLA-B*27-positive patients with chronic HDV infection, whereas substitutions at position 100 were observed in patients with diverse HLA haplotypes (Fig. 6) and are also prevalent in other HDV genotypes (Fig. 7). Taken together, our data provide strong evidence that two residues, 105 and 106, within the novel HLA-B*27-restricted CD8 T cell epitope L-HDAg$_{99-108}$ are selected under immune pressure by the CD8 T cell-response and the variants confer an immune escape to HDV.

***In silico* structural analysis of six representative peptide variants bound to HLA-B*27:05**

To gain a detailed molecular understanding of the HLA-B*27:05 binding mode of the new-identified epitopes and their T-cell activation properties, we performed a molecular dynamics (MD)-based analysis of the binding properties of six representative peptides bound to HLA-B*27:05: the epitope L-HDAg$_{99-108}$ (RRDHRRRKAL), the naturally occurring L-HDAg peptide variants mutated at amino acid position 100 (Q/E/K100), and the immune escape variants R105K and K106M (Table 4, Figs. 10-12).

In Table 4 the experimental binding and T-cell activation data are provided for these peptides together with the calculated interaction energies. The experimental MHC binding data suggests that R100 is the strongest binder among all peptides varying in position 100 (see column 3 of Table 4). This is also reflected by the calculated interaction energies for the Q100 and E100 variants (see column 5 of Table 4), but not for K100. Thus, two further

analyses of the theoretical results were performed: first, a visual analysis of the structural binding properties (see Fig. 10 and Fig. 12), which showed that all variants (Q/E/K100) fit less well into the binding site compared to R100. Second, a quantitative analysis of the hydrogen bond network between the peptides and the HLA-B*27:05-binding site (data not shown) which substantiated the visual observations. The area of the HLA-B*27:05 binding site in which residue 100 is located, consists of a very deep pocket with two hydrogen bond acceptors/donors, E69 and T48, at its bottom (Fig. 10A, blue inlay). This pocket is optimally filled by the large R residue, which forms three to four stable hydrogen bonds with E69 and T48, while K, only containing one charged group, interacts alternatively with E69 or T48, its side chain flipping back and forth between the two residues during the simulation indicating instable binding. Q and E are simply too small to form stable interactions within the pocket as their side chains cannot reach the bottom of the binding site. Therefore, our computational analyses can provide an explanation for the experimentally observed strong binding of the R100 wild type and is consistent with the experimental HLA-B*27:05 binding data.

To estimate the influence of the R105K and K106M mutations on the T-cell response we investigated their effect on the shape and electrostatic potential of the solvent-exposed surface region of the peptides, as this is the region to which the TCR binds (see Fig. 10 and Fig. 12). In addition, the same visual and a quantitative hydrogen bond analyses were performed as for the R100 variants (Fig. 10 and 11). For the L-HDAg$_{99\text{-}108}$ epitope (Fig. 10C) the TCR interface is dominated by R105 and R104 forming a very characteristic steric and electrostatic fingerprint, which seems important for TCR recognition as for both variants, which do not show any T-cell response, large differences were observed at the TCR interaction region (Fig. 10D,H). The R105K mutation directly leads to strong changes in the electrostatic potential of the peptide at the TCR interaction area (Fig. 10D). In the K106M variant the effects are more complex and are caused indirectly by a different hydrogen bond

pattern within the peptide, which leads to considerable changes in the conformation and thus the shape of the peptide's surface (Fig. 10F,H and 11). The reason behind this different behavior of the mutant is analogous to the observations made for the variations at position 100 as discussed above, namely that the native residue K106 fits optimally into its MHC binding pocket, whereas the M mutant does not. Thus K106 forms stable hydrogen bonds with the pocket residues D101 and D98 (Fig. 10E) leading to an optimal peptide-MHC hydrogen bond network, which is stable throughout the simulation (Fig. 11A). In contrary M106 cannot interact via hydrogen bonding with the negatively charged D101/D98 (Fig. 10F) and thus moves to the edge of the binding pocket during the simulation. This leads to large structural rearrangements within the peptide and a final, alternative peptide conformation stabilized by strong intra-peptide hydrogen bonds between peptide residues R103 and D101 (Fig. 10F and 11B). Due to this alternative conformation the peptide-TCR interaction surface differs considerably from the one of the original K106 epitope (Fig. 10H versus 10G), providing a potential explanation for the missing T-cell response to the K106M variant and which is consistent with the effect observed for R105K.

Importantly, both variants (R105K and K106M) still show strong MHC binding (Table 4). For R105K this is directly reflected by the MHC-peptide interaction energy (Table 4) and is not surprising as residue 105 is located outside the MHC binding site and thus does not contribute to MHC-peptide binding. For variant K106M the calculated interaction energy is rather low, which is in contrast to the experimental MHC binding data. This, however, can be explained by the above discussed conformational rearrangement of the peptide due to the mutation leading to a very strong stabilization of the internal peptide conformation and thus a very favorable internal energy of the peptide (data not shown), which comes at the cost of less strong MHC interactions. Interestingly, this alternative binding mode of a very rigid peptide conformation interacting moderately with the MHC binding site seems to lead also to

very stable bound complexes (see experimental data in Table 4). Thus concluding, our modeling results indicate that the missing T-cell response to the R105K and K106M variants of the HLA-B*27:05 epitope L-HDAg$_{99-108}$ is caused by structural and electrostatic differences at the TCR interface, which are not directly correlated to the peptides' HLA-B*27:05 binding strength.

Discussion

In this study, we identified two overlapping HLA-B*27-restricted CD8 T-cell epitopes, L-HDAg$_{99-108}$ and L-HDAg$_{103-112}$, in a conserved region of the only HDV antigen. To both epitopes, a specific CD8 T-cell response was detected in HLA-B*27-positive patients who were able to resolve their HDV infection. Molecular footprinting using a collection of HDV isolates from patients with chronic hepatitis D provided strong evidence that variations at aa 105 and 106, common to both CD8 T-cell epitopes, are selected. Functional T-cell analysis confirmed that aa 105 and 106 variants confer an immune escape for HDV, and *in silico* structural modelling of the HLA-B*27 epitope L-HDAg$_{99-108}$ revealed that both the R105K as well as the K106M variant dramatically alter the TCR interface explaining the immune escape.

Antibodies directed against the HDV nucleoprotein have no neutralizing capacity and are not effective in controlling HDV infection (6). Therefore, T-cell response plays a major role in eliminating HDV which occurs in about 10-20% of HBV carriers who have been superinfected (26). On the other hand, about 80-90% of superinfected patients are not able to eliminate the virus and develop chronic HDV infection. This may be due to primary or secondary failure of the T-cell response which in turn may be caused by T-cell exhaustion, an immune escape of the HDV variants, or simply by an absence of epitopes within HDV proteins to certain HLA alleles.

HDV has a very short genome of 1.7 kb and only one ORF expressing two isoforms of one protein (S-HDAg and L-HDAg). The 214 aa L-HDAg comprises all potential epitopes that may be recognized by T-cell responses. In our analysis, the number of peptides predicted to bind any of the major HLA class-I molecules was significantly lower than that of proteins from other viruses such as HAV, HBV and HCV (27-29). The lack of epitopes may be explained by the fact that the HDV protein could have been derived from a host protein (30) against which self-reacting T cells are deleted during the selection of T cells in the thymus. Of note is that we did not detect high-affinity HDV epitopes binding to HLA alleles other than HLA-B*27, although we used various approaches. In HLA-binding analyses of the predicted peptide epitopes, we only observed low-affinity binding to e.g. HLA-A*01:01, HLA-A*02 or HLA-B*07:02 alleles with a high prevalence in western countries. Low-affinity binding indicates that these peptides will probably be outcompeted by better binding peptides. A previous study by Huang *et.al* described two HLA-A*02 epitopes in Asian patients (12) that we could not confirm these findings in 4 HLA-A*02-positive patients with resolved HDV infection in our study. This may be due to differences in the aa sequences of genotype-1 isolates prevalent in East Asia. Moreover, these HLA-A*02:01-restricted HDV epitopes are located in a highly variable region of L-HDAg which is not conserved in most European HDV isolates. HDV only contains few highly conserved regions in its genome. Among these is the region around the two identified HLA-B*27 epitopes which is conserved throughout all isolates from all genotypes. In fact, our comprehensive analyses of 545 sequences of L-HDAg from all eight HDV genotypes indicated that the region restricted by HLA-B*27 is conserved within all genotypes, implying a possible cross-genotypic HDV-specific T-cell response. This is maybe a unique feature of HDV and stands in contrast with earlier findings made in studies of HCV-specific T-cell response where the protective antiviral effect of HLA-B*27 was shown to be restricted to HCV genotype-1, and not to any

other prevalent genotypes such as 3a (31). The low substitution rate in this region may thus be due to viral fitness costs of mutations in this region.

We demonstrated a HLA-B*27-restricted HDV-specific CD8 T-cell response in some patients with resolved HDV infection. However, other HDV-infected HLA-B*27-positive patients fail to clear HDV and develop a persistent infection. We therefore hypothesized that viral variants emerged which cannot be targeted by the relevant virus-specific CD8 T cells. We addressed this possibility by analyzing a multicentric cohort of patients in which we observed amino acid substitutions in both L-HDAg$_{99\text{-}108}$ and L-HDAg$_{103\text{-}112}$ epitopes abrogating the respective CD8 T-cell response. This sequence variation was absent in HLA-B*27-negative patients. These findings were strengthened by deep sequencing, as we detected minor variations at this location only in the HLA-B*27-positive patients, in addition to those mutations found by conventional sequencing. Further longitudinal studies will be of interest to determine whether such minor species may out-grow the wild-type virus. Classical escape mutations are known as mutations occurring at anchor positions impairing HLA binding capacity and have been detected in other viruses, e.g. HBV and HCV. It is noteworthy that none of the escape mutations described here occur at HLA-B*27 anchor residues. Interestingly, one of the described variants of the L-HDAg$_{103\text{-}112}$ epitope demonstrated a lysine substitution to methionine at residue 106 (the forth residue of this epitope) which was also reported in a HLA-B*27-restricted HCV epitope abrogating the T-cell response to HCV (31); this may indicate the role of methionine in the T-cell interface of HLA-B*27 epitopes regardless of their origin. Our study is the first reporting of impairment of HDV-specific CD8 T-cell response due to escape mutation, suggesting that genetic diversity of HDV, at least in part, is driven by immune pressure.

Molecular modeling studies of the bound complexes of the HDV-specific epitopes of L-HDAg$_{99\text{-}108}$ and its potential immune escape variants to HLA-B*27 provide a potential

structural explanation why variations at the aa positions 105 and 106 of the peptide have a strong impact on TCR binding, whereas aa 100 variations do not. Analysis of the naturally occurring aa 100 variants showed that these mutations do not affect the TCR-binding surface, but can significantly change the binding pattern with the HLA-B*27 molecule and thus lead to less strong interactions. This is reasonable as aa 100 is located in a deep binding pocket of the MHC, four residues away from the two solvent-exposed arginine residues (R104 and R105) crucial for TCR binding. These observations are in accordance with our experimental T-cell analyses, in which aa 100 variant peptides induced a T cell response comparable to that of wild-type peptides.

In silico analysis also provided an explanation for the importance of the surface-exposed peptide residues R104 and R105 for TCR binding. This was experimentally confirmed by the lacking T-cell response to the R105K and K106M variants, supporting the notion that R104 and R105 are crucial for T-cell recognition. Our computational results indicate that the escape mechanism behind the corresponding variants is based on mutation-induced structural (K106M) and electrostatic (R105K) changes in the surface region of the bound peptide, which is essential for T-cell recognition. Thus, structural integrity of the TCR interface seems to be more important than the MHC-binding affinity of the variant. This conclusion is supported by the notion that no correlation between the experimental or calculated MHC-binding properties of the peptides and their ability to activate T-cell responses could be found.

It is important to point out that viral escape may not be the only mechanisms contributing to virus-specific CD8+ T-cell failure during persistent infection. Indeed, Huang et al. demonstrated HLA*02-restricted HDV-specific CD8+ T-cell responses in two patients with long-term negative HDV RNA (and thus probably recovered HDV infection), but were not able to demonstrate HDV-specific responses in persistently infected patients, indicating that

HDV-specific CD8+ T cells were either not primed in these patients, or had undergone T-cell exhaustion and deletion (12). In line with these previous results, we also failed to identify HLA-B*27-restricted HDV-specific CD8+ T-cells after antigen-specific expansion in five patients with chronic HDV infection (Fig. 3). Even after performing peptide/MHC multimer bead-based enrichment of HDV-specific CD8+ T-cells from five patients with chronic HDV infection, a technique well established in our laboratory for enrichment of low-frequency populations for virus-specific CD8+ T cells (32, 33), we were not able to detect HDV-specific CD8+ T cells in chronically infected patients (Fig. 13). These data support the hypothesis that HDV-specific CD8+ T cells undergo final exhaustion and deletion in patients with chronic HBV/HDV co-infection. In addition, sequestration of HDV-specific CD8+ T cells into the liver may be an alternate explanation for their absence in peripheral blood. In contrast, using this technique, HDV-specific CD8+ T cells were readily detectable in a patient with cleared HDV infection but persistent low-level HBV viremia. These HDV-specific CD8+ T cells were antigen-experienced (not naïve, as indicated by CCR7 and CD45RA costaining), displayed high expression of PD1 but varying expression levels of CD127.

Overall, our results indicate that the number of potential epitopes to the frequent HLA alleles in the 214aa of the single protein of HDV is very low; only two HLA-B*27-restricted epitopes could be defined so far, whereas for all major HLA alleles present in the European population no epitope was so far identified. The low number of epitopes and T-cell failure resulting from immune escape may contribute to the high chronicity rate in HDV infection, in addition to the immunoinhibitory environment established by persistent HBV infection that persistent HDV infection is tied to. The hypothesis that HLA-B*27 maybe a protective allele against HDV, as shown in other infections like HIV and HCV, has to be verified by initiation of a prospective study in a larger cohort with patients being HLA-B*27 positive or negative.

452 These data also imply that the development of effective antiviral approaches e.g.
453 immunotherapy, against HDV will be a challenging task.

Materials and Methods

Patient cohorts and samples

Chronically HDV infected patients (n=104, anti-HDV+, HDV-RNA±) were recruited from 8 different medical centers located in Germany (Essen, Freiburg, Hannover and Munich), Spain, Italy and Iran. HLA backgrounds as well as L-HDAg sequences were determined. HLA class-I typing was performed by a Luminex™ Polymerase Chain Reaction-Sequence Specific Oligonucleotide Probe (PCR-SSO) using the LABType™ SSO Kits (One Lambda, Canoga Park, CA, USA) (20) from genomic DNA extracted from blood samples, using a column purification kit (QIAGEN, Hilden, Germany). Two-digit HLA-typing and, in selected patients, high-resolution typing was done to define four-digit HLA types or subgroups.

HDV phylogenetic analysis and epitope prediction

Total RNA was extracted from the patient's serum samples and was reverse transcribed into cDNA by reverse transcriptase from the Moloney Murine Leukemia Virus (Promega, Madison, Wisconsin, USA) at 37°C for 1 hour using HDV-specific primer, 771R (Table 5). In order to obtain sequences of the L-HDAg encoding region, cDNA products from the reverse transcription step was amplified by *Pfu* DNA polymerases (Promega, USA) in a two-step nested PCR. PCR steps using HDV-specific primers 891-F, 339-R, 912-F and 1674-R were performed under the following thermal profile: 94°C for 10 minutes and a 35 cycle of 94°C for 30 seconds, 54°C for 45 seconds, and 72°C for 90 seconds and then an elongation step at 72°C for 7 minutes. Finally, PCR products were purified via QIAquick® Gel Extraction (QIAGEN) and were directly sequenced on an ABI 3730xl DNA Analyzer using internal primers, 912-F and 1674-R. All sequences were submitted to the GenBank (accession numbers: MF175257- MF175360)

HDV sequences from this study and NCBI GenBank were aligned using ClustalX2 software (34). A maximum-likelihood phylogenetic tree was constructed under the Tamura-Nei substitution model using MEGA software v6 (35). Reliability of pairwise comparison and phylogenetic tree analysis was assessed by 1,000 replicates bootstrapping. Genotypes of HDV strains were determined by comparison to full-length L-HDAg reference sequences from GenBank.

Next generation sequencing was performed in 11 selected samples to identify variations at a detection level of 0.1% and to pin down variations in the region flanked from nt position 956 to 1,360 by ultra-deep pyrosequencing (UDPS) (primer sequences are listed in Table 5). The amplicons had a length of 515bp and were isolated from 0.9% agarose gel with the QIAquick Extraction Kit (QIAquick Spin Handbook, QIAGEN, Hilden, Germany) and quantified using Quan-iTPicogreens DNA reagent (Invitrogen). UDPS was performed with the 454/GS-Junior platform (Roche, Branford, CT 06405, USA), using titanium chemistry (GS-Junior Titanium Sequencing Kit). The FastA file from the GS-Junior was demultiplexed and filtered as recently described (36). The specific region analyzed by UDPS (after discarding primer sequences) covered the region from nt 977 to 1,338 of the HDV genome (362 nt analyzed). All sequences were submitted as a new bioproject to the GenBank (accession number: PRJNA434630). FastA reads were filtered and analyzed as recently described (37).

Two hundred L-HDAg sequences of HDV genotype 1 from GenBank were applied for epitope prediction using Immune Epitope database (IEDB; http://www.iedb.org) and SYFPEITHI database (http://www.syfpeithi.de) algorithms (38-40) to predict potential binders to the most common alleles in Europeans (41): HLA-A*02, *01, *03, *11, *24 and HLA-B*35, *51, *07, *18, *08, *27.

Analysis of HLA-binding capability of candidate peptides

Binding affinity of the predicted epitopes was analyzed with the UV-mediated peptide exchange assay using PeliChange™ p*HLA-A*01:01, A*02:01, A*03:01, A*24:02, B*07:02 and B*27:05 kits (Sanquin, Amsterdam, The Netherlands) as described before (27, 42). Briefly, peptide-exchange reactions were performed by the exposure during 30 minutes of conditional pHLA complexes (0.53 µM) to long wavelength UV using a 366-nm UV lamp (CAMAG, Muttenz, Switzerland), in the presence or absence of the indicated peptide (50 µM). Subsequently, peptide-exchange efficiency was analyzed using the HLA class I ELISA which detects beta-2 microglobulin of peptide-stabilized HLA class I complexes in a proper diluted exchange–reaction mixture. To this end streptavidin (2 µg/mL) was bound onto polystyrene microtiter wells (Nunc Maxisorp). After washing and blocking, HLA complex present in exchange reaction mixtures or controls was captured by the streptavidin on the microtiter plate via its biotinylated heavy chain (incubation for 1h at 37°C). Non-bound material was removed by washing. Subsequently, horseradish peroxidase (HRP)-conjugated antibody to human beta2-microglobulin (0.6 µg/mL) was added (incubation for 1h at 37°C). After removal of non-bound HRP conjugate by washing, an ABTS substrate solution was added to the wells. The reaction was stopped after 8 minutes (incubation at room temperature) by the addition of a stop solution and read in an ELISA reader at 414 nm. Every peptide was independently exchanged twice. Every exchange mixture was measured in duplo in the HLA class I ELISA. The absorbances of all peptides were normalized to the absorbance of a known HLA allele-specific ligand with high affinity for each corresponding allele (represents 100%). Negative controls included an HLA allele-specific non-binder and UV-irradiation of the conditional HLA class I complex in the absence of a rescue peptide.

Polyclonal antigen-specific expansion of T-cells

Patient PBMCs were obtained by standard Ficoll density centrifugation (Biocoll; Biochrom, Germany). Freshly isolated PBMCs (2×10^6) were cultured in 1ml complete medium (RPMI 1640 supplemented with 10% fetal calf serum, 100 U/ml penicillin, 100 g/ml streptomycin) containing anti-CD28 antibody (0.5 µg/ml; BD Biosciences). Cells were stimulated with 31 synthetic overlapping 16-mer peptides (EMC Microcollections, Tübingen, Germany; final concentration 10µg/ml per peptide) corresponding to the complete sequence of L-HDAg divided into 8 pools (pool A-H) or individual 10-mer or 11-mer peptides. On Day 2, 1ml complete medium and interleukin-2 (10 U/ml; Roche, Basel, Switzerland) was added. On Day 10 and 12, the cells were tested for IFN-γ secretion by intracellular cytokine staining after re-stimulation with the peptide pools (Day 10) or individual peptides from the positive pool (Day 12) as described (43). Prior to staining, cells were cultured for 5h in the presence of peptides (10µg/ml), and 5µg/ml of Brefeldin A (Sigma-Aldrich, St. Louis, Missouri, United States). Cells were stained using allophycocyanin (APC)-conjugated anti-CD8 and phycoerythrin (PE)-conjugated anti-CD4 and at 4°C for 15 min. After permeabilization, cells were stained with fluorescein isothiocyanate (FITC)-conjugated anti-IFN-γ mAb (all reagents from BD Biosciences) and analyzed by flow cytometry on a FACS Calibur instrument (BD Biosciences). Dead cells were stained using Viaprobe reagent and excluded from analyses. Data files were analyzed with FlowJo 7.6.5 software (Tree Star, Ashland, Oregon).

Computational modeling of HLA-B*27 binding

To investigate the binding properties of selected L-HDAg-derived peptides to HLA-B*27:05 on the structural level, *in silico* modeling studies were performed. The investigated peptide/HLA-B*27:05 complexes were modeled on the basis of the crystal structure of the peptide/HLA-B*27:05 complex with the PDB ID 4G9D (44). Only the α1 and α2

subdomains of the HLA-B*27:05 protein (residues 25 to 206), forming the peptide-binding cleft, were considered during modeling. The peptide sequence was mutated using the IRECS algorithm (45, 46). To allow the HLA-B*27:05 peptide-binding cleft to adapt to the mutated peptide, HLA-B*27:05 side chains were subsequently also remodeled with IRECS. The modeled peptide/HLA-B*27:05 complexes were energy minimized in a neutralized, rectangular box of TIP3P (47) water molecules with a minimum solute distance of 12 Å to the box boundary. The general system setup and the simulations were conducted with AmberTools 14 and the Amber 14 software package (48). The ILDN-corrected Amber ff99SB force field potential energy function and parameter set (49) were chosen as interaction potential. MD simulations were performed using a time step of 1 fs. The SHAKE algorithm (50) was applied to all bonds involving hydrogen atoms. Long-range electrostatic interactions were computed using the Particle Mesh Ewald (PME) method (pmemd.cuda_SPFP) (51). A cutoff of 14 Å was used for the computation of non-bonded interactions. Temperature and pressure were controlled applying the Berendsen thermostat and barostat with default settings (52). The systems were stepwise heated up to 300 K, while gradually decreasing solute atom restraints, using the NVT ensemble. Afterwards NPT-based molecular dynamics simulations were performed for 20 ns for each system.

Trajectory processing and analysis was done with cpptraj (53). MD frames for analysis were extracted every 20 ps from the last 5 ns of the simulations. Representative conformations were obtained by conformational clustering using the average-linkage method with a cutoff of 2 Å, hydrogen bonds and peptide/HLA-B*27:05 interaction energies were calculated applying default settings of cpptraj. Electrostatic potentials were calculated using the PDB2PQR server (Version 2.1.1) and APBS (Version 1.2.1) (54, 55), keeping default settings. Mapping of electrostatic potentials to the solvent accessible surface areas of the

peptide ligands was performed with VMD (Version 1.9) (56). Figures of peptide/HLA-B*27:05 complexes were also generated with VMD (Version 1.9).

Study approval

Local ethics committees had approved this study according to the 1975 Declaration of Helsinki guidelines. The main approaches for "analyzing the HDV-specific T cell immune responses" in HDV infected patients were approved by the Ethics Committee of the faculty of Medicine of the Duisburg-Essen University (10-4472). The protocol of isolating and analyzing the Genomic DNA (genome-wide association study) was also approved by the Ethics Committee of Medical School of Hannover (OE 9515/No. 3388 and No. 5292M). And the Ethics Committee of the Albert-Ludwigs-Universität Freiburg, Nr. 369-15, approved the protocol for investigation of "Viral escape from dominant virus-specific CD8+ T cell response in HBV mono-infection and HBV/hepatitis D virus co-infection and strategies to restore antiviral CD8+ T cell response". All the patients participated in this study were interviewed and informed about the concept of the study and possible side effects before they accept to participate by giving a written, informed consent.

Acknowledgments

This work was financially supported by internal funds of the Medical Faculty of the University of Duisburg-Essen / Technical University of Munich and by Registry Grant from European Association for the Study of the Liver (EASL). HKa was supported by a grant from the German Academic Exchange Service (DAAD). The Deutsche Forschungsgemeinschaft (DFG) financed parts of the study by support to HKa and BR (SFB-TR179/TP03), UP (SFB-TR179/TP14 and TRR36/B13), MG (SFB 1035/A10), CNH (SFB-TRR-179/TP 2), DH and JT (TRR60/B1). Instituto de Salud Carlos III, grant PI14/01416 co-financed by European Regional Development Fund (ERDF) also supported this study.

We thank all the people who volunteered in this research and the staff working in the diagnostic sections, institutes of virology, University Hospital Essen and TU Munich, Germany, for their close collaborations recruiting the samples and technical assistance. We are also grateful to Prof. Dr. Ulrich Spengler, Dr. Svenja Hardtke and Thomas Schirdewahn for providing information from patients including those with resolved HDV infection. The authors appreciate the help and support of the German Center for Infection Research (DZIF) regarding this work.

HKa, MRo, MRi and JT assisted with proof of concept and design; HKa, ADK, MMK, MF, VO, ESA, JYM, MN, JG, AO, AH, MHom and FRF conducted various aspects of experimental work; BB and DH prepared statistical analysis; MG and IA performed the molecular modeling studies; HKa, ADK, MHof, CNH, FRF, DT, WJEvE, AS, TB, BR, HKe, JSzW, MB, SMA HW, UP and MRo analysed and interpreted the data; HKa, ADK, MHom, IA, UP and MRo wrote the manuscript; CNH, MF, MB, TB, BR, DH, HW, MRi, MF, UP and JT provided critical revision for intellectual content; MF, HW, MH and MB facilitated

609 recovered patients recruitments; MF, HKa, JG, MRo, HKe, JSzW and SMA handled ethics
610 and sample recruitment; CNH, UP, IA and MRo supervised the study.
611

References

1. Rizzetto M. 2015. Hepatitis D Virus: Introduction and Epidemiology. Cold Spring Harb Perspect Med 5:a021576.
2. Gomes-Gouvea MS, Soares MC, Bensabath G, de Carvalho-Mello IM, Brito EM, Souza OS, Queiroz AT, Carrilho FJ, Pinho JR. 2009. Hepatitis B virus and hepatitis delta virus genotypes in outbreaks of fulminant hepatitis (Labrea black fever) in the western Brazilian Amazon region. J Gen Virol 90:2638-43.
3. Weiner AJ, Choo QL, Wang KS, Govindarajan S, Redeker AG, Gerin JL, Houghton M. 1988. A single antigenomic open reading frame of the hepatitis delta virus encodes the epitope(s) of both hepatitis delta antigen polypeptides p24 delta and p27 delta. J Virol 62:594-9.
4. Wang JG, Cullen J, Lemon SM. 1992. Immunoblot analysis demonstrates that the large and small forms of hepatitis delta virus antigen have different C-terminal amino acid sequences. J Gen Virol 73 (Pt 1):183-8.
5. Rizzetto M, Smedile A. 2002. Hepatitis D, p 863-875. Lippincott, Williams and Wilkins, Philadelphia, PA.
6. Rizzetto M, Hoyer BH, Purcell RH, Gerin JL. 1984. Hepatitis delta virus infection, p 371-379. Grune & Stratton, Orlando, FL.
7. Wedemeyer H, Yurdaydin C, Dalekos GN, Erhardt A, Cakaloglu Y, Degertekin H, Gurel S, Zeuzem S, Zachou K, Bozkaya H, Koch A, Bock T, Dienes HP, Manns MP. 2011. Peginterferon plus adefovir versus either drug alone for hepatitis delta. N Engl J Med 364:322-31.
8. Wang D, Pearlberg J, Liu YT, Ganem D. 2001. Deleterious effects of hepatitis delta virus replication on host cell proliferation. J Virol 75:3600-4.

636 9. Mauch C, Grimm C, Meckel S, Wands JR, Blum HE, Roggendorf M, Geissler M.
637 2001. Induction of cytotoxic T lymphocyte responses against hepatitis delta virus
638 antigens which protect against tumor formation in mice. Vaccine 20:170-80.
639 10. Fiedler M, Kosinska A, Schumann A, Brovko O, Walker A, Lu M, Johrden L, Mayer
640 A, Wildner O, Roggendorf M. 2013. Prime/boost immunization with DNA and
641 adenoviral vectors protects from hepatitis D virus (HDV) infection after simultaneous
642 infection with HDV and woodchuck hepatitis virus. J Virol 87:7708-16.
643 11. Nisini R, Paroli M, Accapezzato D, Bonino F, Rosina F, Santantonio T, Sallusto F,
644 Amoroso A, Houghton M, Barnaba V. 1997. Human CD4+ T-cell response to
645 hepatitis delta virus: identification of multiple epitopes and characterization of T-
646 helper cytokine profiles. J Virol 71:2241-51.
647 12. Huang YH, Tao MH, Hu CP, Syu WJ, Wu JC. 2004. Identification of novel HLA-
648 A*0201-restricted CD8+ T-cell epitopes on hepatitis delta virus. J Gen Virol 85:3089-
649 98.
650 13. Neumann-Haefelin C, Oniangue-Ndza C, Kuntzen T, Schmidt J, Nitschke K, Sidney
651 J, Caillet-Saguy C, Binder M, Kersting N, Kemper MW, Power KA, Ingber S, Reyor
652 LL, Hills-Evans K, Kim AY, Lauer GM, Lohmann V, Sette A, Henn MR, Bressanelli
653 S, Thimme R, Allen TM. 2011. Human leukocyte antigen B27 selects for rare escape
654 mutations that significantly impair hepatitis C virus replication and require
655 compensatory mutations. Hepatology 54:1157-1166.
656 14. Kefalakes H, Budeus B, Walker A, Jochum C, Hilgard G, Heinold A, Heinemann
657 FM, Gerken G, Hoffmann D, Timm J. 2015. Adaptation of the hepatitis B virus core
658 protein to CD8(+) T-cell selection pressure. Hepatology 62:47-56.
659 15. Fitzmaurice K, Petrovic D, Ramamurthy N, Simmons R, Merani S, Gaudieri S, Sims
660 S, Dempsey E, Freitas E, Lea S, McKiernan S, Norris S, Long A, Kelleher D,

Klenerman P. 2011. Molecular footprints reveal the impact of the protective HLA-A*03 allele in hepatitis C virus infection. Gut 60:1563-71.

16. Timm J, Walker CM. 2014. Mutational escape of CD8+ T cell epitopes: implications for prevention and therapy of persistent hepatitis virus infections. Med Microbiol Immunol doi:10.1007/s00430-014-0372-z.

17. Phillips RE, Rowland-Jones S, Nixon DF, Gotch FM, Edwards JP, Ogunlesi AO, Elvin JG, Rothbard JA, Bangham CR, Rizza CR, et al. 1991. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. Nature 354:453-9.

18. Goulder PJ, Phillips RE, Colbert RA, McAdam S, Ogg G, Nowak MA, Giangrande P, Luzzi G, Morgan B, Edwards A, McMichael AJ, Rowland-Jones S. 1997. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. Nat Med 3:212-7.

19. Wang SY, Wu JC, Chiang TY, Huang YH, Su CW, Sheen IJ. 2007. Positive selection of hepatitis delta antigen in chronic hepatitis D patients. J Virol 81:4438-44.

20. Heinemann FM. 2009. HLA Genotyping and Antibody Characterization Using the Luminex Multiplex Technology. Transfus Med Hemother 36:273-278.

21. Fitzmaurice K, Hurst J, Dring M, Rauch A, McLaren PJ, Gunthard HF, Gardiner C, Klenerman P. 2015. Additive effects of HLA alleles and innate immune genes determine viral outcome in HCV infection. Gut 64:813-9.

22. McLaren PJ, Carrington M. 2015. The impact of host genetic variation on infection with HIV-1. Nat Immunol 16:577-83.

23. Neumann-Haefelin C, McKiernan S, Ward S, Viazov S, Spangenberg HC, Killinger T, Baumert TF, Nazarova N, Sheridan I, Pybus O, von Weizsäcker F, Roggendorf M, Kelleher D, Klenerman P, Blum HE, Thimme R. 2006. Dominant influence of an

HLA-B27 restricted CD8+ T cell response in mediating HCV clearance and evolution. Hepatology 43:563-572.

24. Nitschke K, Barriga A, Schmidt J, Timm J, Viazov S, Kuntzen T, Kim AY, Lauer GM, Allen TM, Gaudieri S, Rauch A, Lange CM, Sarrazin C, Eiermann T, Sidney J, Sette A, Thimme R, Lopez D, Neumann-Haefelin C. 2014. HLA-B*27 subtype specificity determines targeting and viral evolution of a hepatitis C virus-specific CD8+ T cell epitope. J Hepatol 60:22-9.

25. Shirvani-Dastgerdi E, Amini-Bavil-Olyaee S, Alavian SM, Trautwein C, Tacke F. 2014. Comprehensive analysis of mutations in the hepatitis delta virus genome based on full-length sequencing in a nationwide cohort study and evolutionary pattern during disease progression. Clin Microbiol Infect doi:10.1016/j.cmi.2014.12.008.

26. Farci P, Niro GA. 2012. Clinical features of hepatitis D. Semin Liver Dis 32:228-36.

27. Schulte I, Hitziger T, Giugliano S, Timm J, Gold H, Heinemann FM, Khudyakov Y, Strasser M, Konig C, Castermans E, Mok JY, van Esch WJ, Bertoletti A, Schumacher TN, Roggendorf M. 2011. Characterization of CD8+ T-cell response in acute and resolved hepatitis A virus infection. J Hepatol 54:201-8.

28. Brinck-Jensen NS, Vorup-Jensen T, Leutscher PD, Erikstrup C, Petersen E. 2015. Immunogenicity of twenty peptides representing epitopes of the hepatitis B core and surface antigens by IFN-gamma response in chronic and resolved HBV. BMC Immunol 16:65.

29. Wong DK, Dudley DD, Afdhal NH, Dienstag J, Rice CM, Wang L, Houghton M, Walker BD, Koziel MJ. 1998. Liver-derived CTL in hepatitis C virus infection: breadth and specificity of responses in a cohort of persons with chronic infection. J Immunol 160:1479-88.

30. Brazas R, Ganem D. 1996. A cellular homolog of hepatitis delta antigen: implications for viral replication and evolution. Science 274:90-4.

31. Neumann-Haefelin C, Timm J, Schmidt J, Kersting N, Fitzmaurice K, Oniangue-Ndza C, Kemper MN, Humphreys I, McKiernan S, Kelleher D, Lohmann V, Bowness P, Huzly D, Rosen HR, Kim AY, Lauer GM, Allen TM, Barnes E, Roggendorf M, Blum HE, Thimme R. 2010. Protective effect of human leukocyte antigen B27 in hepatitis C virus infection requires the presence of a genotype-specific immunodominant CD8+ T-cell epitope. Hepatology 51:54-62.

32. Schmidt J, Neumann-Haefelin C, Altay T, Gostick E, Price DA, Lohmann V, Blum HE, Thimme R. 2011. Immunodominance of HLA-A2-restricted hepatitis C virus-specific CD8+ T cell responses is linked to naive-precursor frequency. J Virol 85:5232-6.

33. Nitschke K, Flecken T, Schmidt J, Gostick E, Marget M, Neumann-Haefelin C, Blum HE, Price DA, Thimme R. 2015. Tetramer enrichment reveals the presence of phenotypically diverse hepatitis C virus-specific CD8+ T cells in chronic infection. J Virol 89:25-34.

34. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947-8.

35. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725-9.

36. Quer J, Gregori J, Rodriguez-Frias F, Buti M, Madejon A, Perez-Del-Pulgar S, Garcia-Cehic D, Casillas R, Blasi M, Homs M, Tabernero D, Alvarez-Tejado M, Munoz JM, Cubero M, Caballero A, delCampo JA, Domingo E, Belmonte I, Nieto L, Lens S, Munoz-de-Rueda P, Sanz-Cameno P, Sauleda S, Bes M, Gomez J, Briones C,

Perales C, Sheldon J, Castells L, Viladomiu L, Salmeron J, Ruiz-Extremera A, Quiles-Perez R, Moreno-Otero R, Lopez-Rodriguez R, Allende H, Romero-Gomez M, Guardia J, Esteban R, Garcia-Samaniego J, Forns X, Esteban JI. 2015. High-Resolution Hepatitis C Virus Subtyping Using NS5B Deep Sequencing and Phylogeny, an Alternative to Current Methods. J Clin Microbiol 53:219-26.

37. Homs M, Rodriguez-Frias F, Gregori J, Ruiz A, Reimundo P, Casillas R, Tabernero D, Godoy C, Barakat S, Quer J, Riveiro-Barciela M, Roggendorf M, Esteban R, Buti M. 2016. Evidence of an Exponential Decay Pattern of the Hepatitis Delta Virus Evolution Rate and Fluctuations in Quasispecies Complexity in Long-Term Studies of Chronic Delta Infection. PLoS One 11:e0158557.

38. Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, Lundegaard C, Sette A, Lund O, Bourne PE, Nielsen M, Peters B. 2012. Immune epitope database analysis resource. Nucleic Acids Res 40:W525-30.

39. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S. 1999. SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 50:213-9.

40. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. 2008. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. Nucleic Acids Res 36:W509-12.

41. Muller CR, Ehninger G, Goldmann SF. 2003. Gene and haplotype frequencies for the loci hLA-A, hLA-B, and hLA-DR based on over 13,000 german blood donors. Hum Immunol 64:137-51.

42. Rodenko B, Toebes M, Hadrup SR, van Esch WJ, Molenaar AM, Schumacher TN, Ovaa H. 2006. Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. Nat Protoc 1:1120-32.

43. Kern F, Surel IP, Brock C, Freistedt B, Radtke H, Scheffold A, Blasczyk R, Reinke P, Schneider-Mergener J, Radbruch A, Walden P, Volk HD. 1998. T-cell epitope mapping by flow cytometry. Nat Med 4:975-8.

44. Ladell K, Hashimoto M, Iglesias MC, Wilmann PG, McLaren JE, Gras S, Chikata T, Kuse N, Fastenackels S, Gostick E, Bridgeman JS, Venturi V, Arkoub ZA, Agut H, van Bockel DJ, Almeida JR, Douek DC, Meyer L, Venet A, Takiguchi M, Rossjohn J, Price DA, Appay V. 2013. A molecular basis for the control of preimmune escape variants by HIV-specific CD8+ T cells. Immunity 38:425-36.

45. Hartmann C, Antes I, Lengauer T. 2007. IRECS: a new algorithm for the selection of most probable ensembles of side-chain conformations in protein models. Protein Sci 16:1294-307.

46. Hartmann C, Antes I, Lengauer T. 2009. Docking and scoring with alternative side-chain conformations. Proteins 74:712-26.

47. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. J Chem Phys 79.

48. Case DA, Babin V, Berryman JT, Betz RM, Cai Q, Cerutti DS, Cheatham TE, III, Darden TA, Duke RE, Gohlke H, Goetz AW, Gusarov S, Homeyer N, Janowski P, Kaus J, Kolossváry I, Kovalenko A, Lee TS, LeGrand S, Luchko T, Luo R, Madej B, Merz KM, Paesani F, Roe DR, Roitberg A, Sagui C, Salomon-Ferrer R, Seabra G, Simmerling CL, Smith W, Swails J, Walker RC, Wang J, Wolf RM, Wu X, Kollman PA. 2014. AMBER 14. University of California, San Francisco.

49. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78:1950-8.

50. Ryckaert J-P, Ciccotti G, Berendsen HJC. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. J Comput Phys 23:327-341.

51. Salomon-Ferrer R, Götz AW, Poole D, Grand SL, Walker RC. 2013. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. J Chem Theory Comput 9:3878-3888.

52. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. 1984. Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684-3690.

53. Roe DR, Cheatham TEI. 2013. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J Chem Theory Comput 9:3084-3095.

54. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A 98:10037-41.

55. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. 2004. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res 32:W665-7.

56. Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. J Mol Graph 14:33-8, 27-8.

57. Trolle T, Metushi IG, Greenbaum JA, Kim Y, Sidney J, Lund O, Sette A, Peters B, Nielsen M. 2015. Automated benchmarking of peptide-MHC class I binding predictions. Bioinformatics 31:2174-81.

Figures and Figure legends

Fig. 1: MHC binding assay of the highest ranking predicted L-HDAg epitopes. Binding to the indicated HLA molecule was assessed by UV-induced ligand exchange. A total of 15 different predicted epitopes were tested for binding HLA-A*01:01 (n=2), HLA-A*02:01(n=3), HLA-A*03:01(n=2), HLA-A*24:02(n=3), HLA-B*07:02 (n=3) and HLA-B*27:05(n=2). L-HDAg$_{198-206}$ was tested with HLA-A*2:01 and *24:02. Binding of predicted peptide epitopes is given as % binding of indicated epitopes with high binding-affinities for the respective HLA molecule. Mean±SD is shown. Negative controls included exchange of a non-binder and UV illumination in the absence of any peptide.

Fig. 2: Peptide recognition by T cells from patients with resolved HDV infection. PBMCs were stimulated with indicated peptides after *in vitro* expansion and surface staining for T-cell markers. ICS was performed and analyzed by flow cytometry. (**A**) CD8 and IFN-γ staining after stimulation with pools A to H (upper and middle panel) of overlapping 16-mer peptides spanning the whole L-HDAg and indicated single peptides from pool D (lower panel). Peptide D14 was L-HDAg$_{98-113}$ (ERRDHRRRKALENKKK). (**B,C**) Representative dot plots of CD8 T-cell response of patient B (**B**) and patient C (**C**) to the 16-mer peptide D14 as well as the two indicated 10-mer candidate epitopes L-HDAg$_{99-108}$ (RRDHRRRKAL) and L-HDAg$_{103-112}$ (RRRKALENKK) within the 16-mer peptide. Negative controls have been treated exactly as the other samples from the respective patients, but were not stimulated with peptide prior to intracellular interferon gamma staining.

Fig. 3: HLA-B*27-restricted epitopes recognition by HLA-B*27 negative patients with resolved infection, HLA-B*27 positive patients with chronic HDV infection and HLA-B*27 positive healthy individuals. (A) PBMCs from four HLA-B*27 negative patients with

resolved HDV infection (rHDV) were stimulated with the pool of peptides including the 16-mer peptide (D14) spanning both identified HLA-B*27-restricted epitopes and evaluated for interferon-gamma production by CD8 T cells. **(B)** PBMCs from five HLA-B*27 positive patients with chronic HDV infection (cHDV) were expanded with 16-mer peptide spanning both HLA-B*27-restricted epitopes as well as single 10-mer epitopes (cHDV 1-3) or only 10-mers (cHDV4-5). Interferon-gamma production was evaluated by surface and intracellular staining after stimulation. **(C)** The same was performed on PBMC samples from 2 HLA-B*27 positive individuals without HBV, HCV and HDV infection (Healthy 1-2). Negative controls have been treated exactly as the other samples from the respective patients, but were not stimulated with peptide prior to intracellular interferon gamma staining. Positive controls were stimulated with PMA/ionomycin prior to intracellular interferon gamma staining.

Fig. 4: The epitope HLA-B*27 L-HDAg$_{103-112}$ is restricted by the HLA-B*27 subtypes HLA-B*27:05 (HLA-B*27 prototype) as well as HLA-B*27:02 (minor Mediterranean HLA-B*27 subtype). PBMC were expanded in the presence of peptide RRRKALENKK for 10 days. Then, intracellular interferon-gamma staining was performed after direct stimulation with peptide (left column) or stimulation with EBV-immortalized B cell lines that were positive for HLA-B*27:05, HLA-B*27:02, or negative for HLA-B*27, and were loaded with peptide overnight and then washed extensively (lower line) or unloaded with peptide (upper line). Cells were also stimulated with PMA/ionomycin as positive controls.

Fig. 5: T cell response analysis of HLA-A*02 positive patients to previously described HLA-A*02-restricted HDV epitopes: PBMCs from 4 HLA-A*02 positive patients with resolved HDV infection were stimulated with previously described HLA-A*02-restricted HDV epitopes, L-HDAg$_{43-51}$ (KLEDENPWL) and L-HDAg$_{26-34}$ (KLEDLERDL), in a 10 day culture. T-cell surface marker staining as well as ICS were performed and analyzed by flow cytometry. Percentage of IFN-γ+/CD8+ T cells are indicate in each panel. Patient number 4

was not stimulated with L-HDAg$_{26-34}$ (KLEDLERDL). Negative controls have been treated exactly as the other samples from the respective patients, but were not stimulated with peptide prior to intracellular interferon gamma staining. Positive controls were stimulated with PMA/ionomycin prior to intracellular interferon gamma staining.

Fig. 6: Viral sequence variations within the HLA-B*27-restricted epitopes. Amino acid sequences spanning the two overlapping HLA-B*27-restricted epitopes L-HDAg$_{99-108}$ (RRDHRRRKAL) and L-HDAg$_{103-112}$ (RRRKALENKK) were deducted from HDV-isolate sequence analyses. Amino acid sequence alignments were sorted by HLA-B*27 haplotype. The first six lines show the sequence data from individual HLA-B*27-positive patients. The lower part shows the frequency of respective amino acid variations found in 98 HLA-B*27-negative patients. Apart from variability at aa 100 and 112 present in both HLA-B*27 positive and negative patients, variations within the two epitopes was observed in 6 out of 98 HLA-B*27 negative patients versus 3 out of 6 in HLA-B*27 positive patients. Fisher's exact test was used to make contingency tables and calculate the *p* values for each amino acid-HLA-B*27 correlation. *P* values smaller than 0.05 were considered as statistical significant. Substitutions at residues 105 and 106, reflecting immune escape mutations inside these epitopes, are indicated by a darker color.

Fig. 7: Amino acid substitutions at aa 100 of L-HDAg in between HDV genotype 1-8. List of the most frequent viral protein sequence variations in the region binding to HLA-B*27 within HDV genotype 1 and other HDV genotypes (2-8).

Fig. 8: Functional impact of aa variations in L- HDAg$_{103-112}$ and L- HDAg$_{99-108}$ on HDV-specific CD8 T-cell response. (**A**) 10-mer L-HDAg$_{103-112}$ (RRRKALENKK) or 9-mer L-HDAg$_{104-112}$ (RRKALENKK) peptide were used to stimulate IFN-γ responses of CD8+ T cells from a donor with resolved HDV infection (patient C). The functional impact of K112R, R105K and K106M amino acid substitutions on the T-cell response was analyzed by flow

cytometry using respective 10-mer (left) and 9-mer (right) peptides for stimulation. (**B**) Either highly prevalent L-HDAg$_{99-108}$ R**Q**DHRRRKAL and R**R**DHRRRKAL variants of HDV or rare Q/R100K, Q/R100E variants were used to stimulate IFN-γ responses of CD8+ T cells from a donor with resolved HDV infection (patient B) (4 left panels). The functional impact of potential immune escape variants R105K and K106M (indicated in red) on the T-cell response was analyzed by flow cytometry (4 right panels). The negative control has been treated exactly as the other samples from the respective patients, but was not stimulated with peptide prior to intracellular interferon gamma staining.

Fig. 9: Mutation K106M leads to viral escape.

PBMCs from patient C two and three years after HDV clearance (upper panels) were expanded in the presence of peptide L-HDAg$_{104-112}$ for 10 days and were then analyzed by intracellular interferon-gamma staining after stimulation with peptide L-HDAg$_{104-112}$ as well as the two viral variants K106M (■) and R105K (▲) in increasing concentrations.

The ELISpot was also performed after stimulation with wild-type or variant peptide in increasing peptide concentrations (lower panel). Tests two years after HDV clearance were performed in unicates due to limitation of cell numbers; tests three years after HDV clearance were performed in triplicates (data are presented as mean ± standard deviation). For statistical analysis of responses against variant *versus* consensus, 2-tailed Student's t test was performed. A p value less than 0.05 was considered significant. * $p<0.05$; *** $p<0.001$.

Fig. 10: Representative structures of L-HDAg$_{99-108}$ peptide epitopes carrying natural variants R/Q100 or immune escape variants R105K and K106M bound to HLA-B*27:05. For better visualization, all bound peptide ligands are oriented from C- to N-terminus (left to right side of the pictures). The HLA-B*27:05 molecule is shown in cartoon representation in gray and important residues are given in atom type colored licorice representation (cyan carbon atoms). The peptides are shown in licorice (orange)

representation or alternatively as a solvent-accessible surface area colored according to the electrostatic potential (i.e. blue represents positively charged, red negatively charged and white neutral areas of the surface). Hydrogen bonds are shown in magenta. (**A,B**) Structures of the bound R100 (**A**) and Q100 (**B**) L-HDAg$_{99\text{-}108}$ peptides undergoing alternative binding site interactions; blue inlays show a magnification. (**C,D**) Structures of (**C**) the bound wild-type L-HDAg$_{99\text{-}108}$ peptide and (**D**) its variant R105K. The residues R105 and R104 are highlighted in a yellow box. (**E-H**) Structures of the bound wild-type (**E,G**) and variant K106M (**F,H**) peptide that loses its interaction to one of the HLA-B*27:05 α-helices. The red inlay provides a magnification of the binding site interactions. (**G**) and (**H**) give a structural representation of the MHC binding site with the bound peptide shown as surface representation.

Fig. 11: Inter- and intramolecular hydrogen bonds of the L-HDAg$_{99\text{-}108}$ epitope (A) and the K106M variant of the L-HDAg$_{99\text{-}108}$ epitope (B) monitored over the simulation time. A vertical black line indicates that at least one hydrogen bond is formed at a specific timeframe during the simulation between the functional groups of the two amino acid residues given at the Y-axis. The percentage of simulation time during which a hydrogen bond exists is given on the right side. L-HDAg$_{99\text{-}108}$ epitope and K106M variant amino acid residues are depicted in orange, MHC binding pocket amino acid residues are shown in cyan.

Fig. 12: Representative structures of L-HDAg$_{99\text{-}108}$ peptide epitopes carrying the natural variants K/E100 bound to HLA-B*27:05. For better visualization, all bound peptide ligands are oriented from the C- to N-terminus (left to right side of the pictures). The MHC molecule is shown in cartoon representation in gray and important residues are given in atom type colored licorice representation (cyan carbon atoms). The peptides are shown in licorice (orange) representation. Hydrogen bonds are shown in magenta. The blue inlays provide a

magnification of the binding site interactions. (**A**) Structure of the bound K100 peptide variant. (**B**) Structure of the bound E100 peptide variant.

Fig. 13: *Ex vivo* peptide/tetramer bead-based enrichment of L-HDAg104-112-specific CD8+ T cells in patients with resolved versus chronic HDV infection

(**A**) L-HDAg104-112 peptide/HLA-B*27:05 tetramer bead-based ex vivo enrichment of HDV-specific CD8+ T cells was performed from five patients chronic HBV/HDV co-infection as well as one patient with chronic HBV infection and resolved HDV infection (methods, see Schmidt et al., J Virol 2011:85:5232-6). Using this extremely sensitive method, we were able to detect HDV-specific CD8+ T cells from a patient with chronic HBV infection and resolved HDV infection. In contrast, we were not able to enrich HDV-specific CD8+ T cells from the five patients with chronic HBV/HDV co-infection. The autologous sequences are also displayed (ND: viral sequencing not done due to low viral load).

(**B**) *Ex vivo* enriched HDV-specific CD8+ T cells from the patient with chronic HBV infection and resolved HDV infection were further analyzed for expression of PD1 and CD127. Naïve CD8+ T cells were excluded by co-staining for CCR7 and CD45RA.

Tables

Table 1. List of best predicted epitopes on L-HDAg for the selected HLA alleles. The predicted epitopes are determined for HLA-A*01, A*02, A*03, A*24, B*07, and B*27 using two different algorithms.

HLA:	Position:	Sequence:	Length	Score (IEDB)*	Score SYFPEITHI*
A*01	75-83	RTDQMEVDS	9	3.2	17
	94-102	FTDKERQHD	9	1.3	17
A*02	43-51	KLEEDNPWL	9	1.3	23
	143-152	RVAGPPVGGV	10	4.2	23
	198-206	ILFPSDPPF	9	0.8	16
A*03	53-61	NIKGILGKK	9	2.7	22
	89-97	PLRGGFTDK	9	2.1	28
A*24	49-57	PWLGNIKGI	9	3.9	13
	192-200	RGFPWDILF	9	1.1	13
	198-206	ILFPSDPPF	9	2.4	11
B*07	68-76	APPAKRART	9	1.9	21
	71-79	AKRARTDQM	9	3.1	12
	168-176	VPSMQGVPE	9	3.0	16
B*27	99-108	RRDHRRRKAL	10	0.35	24
	103-112	RRRKALENKK	10	0.20	29

***** The two algorithems (IEDB recommended and SYFPEITHI) are described in detail elsewhere (39, 57)

Table 2. Library of HDV peptides. Overlapping 16-mer peptides spanning the whole HDAg open reading frame of 214 amino acids which were used for T cell stimulation.

Pool	Peptide name	Position	Sequences
A	A1	1-16	MSRSESRKNRGGREEL
A	A2	2-17	SRSESRKNRGGREELL
A	A3	10-23	RGGREELLEQWVAGRK
A	A4	18-33	EQWVAGRKKLEELERD
B	B5	26-41	KLEELERDLRKTKKKL
B	B6	34-49	LRKTKKKLKKIEDENP
B	B7	42-57	KKIEDENPWLGNIKGI
B	B8	50-65	WLGNIKGILGKKDKDG
C	C9	58-73	LGKKDKDGEGAPPAKR
C	C10	66-81	EGAPPAKRARTDQMEV
C	C10a	66-81	EGAPPAKRARTDRMEV
C	C11	74-89	ARTDQMEVDSGPGKRP
D	D11a	74-89	ARTDRMEVDSGPGKRP
D	D12	82-97	DSGPGKRPLRGGFTDK
D	D13	90-105	LRGGFTDKERRDHRRR
D	D14	98-113	ERRDHRRRKALENKKK
E	E15	106-121	KALVNKKKQLSAGGKN
E	E16	114-129	QLSAGGKNLSKEEEEE
E	E17	122-137	LSKEEEEELRRLTEED
E	E18	130-145	LRRLTEEDERRERRVA
F	F19	138-153	ERRERRVAGPPVGGVN
F	F20	146-161	GPPVGGVNPLEGGSRG
F	F21	154-169	PLEGGSRGAPGGGFVP
F	F22	162-177	APGGGFVPNLQGVPES
G	G23	170-185	NLQGVPESPFSRTGEG
G	G23a	170-185	NLQGVPESPFARTGEG
G	G24	178-193	PFSRTGEGLDIRGNQG
G	G24a	178-193	PFARTGEGLDIRGNQG
H	H25	186-201	LDIRGNQGFPQDTLFP
H	H26	194-209	FPQDTLFPADPPLSPQ
H	H27	199-214	LFPADPPLSPQSCRPQ

Table 3. Deep sequencing analysis of 11 samples of patients (5 HLA-B*27 positive and 6 HLA-B*27 negative) using UDPS. The master sequence corresponding to epitope region (HDAg codons 99 to 112) and the corresponding polymorphisms and their proportions are indicated in bold. The amino acid position affected by polymorphism, underlined the substitution in relation to the canonical epitope sequence observed in master sequence.

	Patients	99-112 seq.	Epitope polymorphism	# of Haplotypes	# of read
HLA-B*27 Pos.	H09	RQDHRRRKALEN**KS**	K 111R 0.8%; K 112S 100%	14	3428
	H20	RQDHRRRKALENKR		26	3346
	S17	RQDHRR**K**KALEN**K**K	R105K 100%; K111R 0.8%	29	4244
	H36	R**Q**DHRRRKALENK**E**	Q100R 0.8%; K112E 100%	17	3959
	It06	RQDHRR**RM**ALENKK	K106M 100%; R105G 0.3%	15	2341
HLA-B*27 Neg.	H26	RQDHRRRKALENKR		9	3447
	H40	RQDHRRRKALENKR		41	2422
	It04	RQDHRRRKALENKR		22	4457
	It20	RQDHRRRKALENKK	K112R 18.7%	46	2487
	S10	RQDHRRRKALENKK	K112R 2%	12	3397
	E18	RQDHRRRKALENKK		19	3277

Table 4. Experimental and calculated MHC binding and TCR activity data.

Peptide	Abbreviation	MHC Binding	T-Cell Activity	E_{int} [kcal/mol][d]
RRDHRRRKAL	WT	127.2%[a]±13.3[b]	0.59%[c]	-686.96±15.39[e]
RQDHRRRKAL	R100Q	61.0%[a]±13.8[b]	0.57%[c]	-668.75±16.85[e]
REDHRRRKAL	R100E	62.6% [a]±12.4[b]	0.52%[c]	-570,54±41.39[e]
RKDHRRRKAL	R100K	40.2% [a]±5.3[b]	0.50%[c]	-727,40±3.17[e]
RRDHRR**K**KAL	R105K	92.7% [a]±3.7[b]	0.08%[c]	-678.31±14.93[e]
RRDHRRR**M**AL	K106M	133.0% [a]±2.7[b]	0.18%[c]	-484.66±24.61[e]

[a] relative to positive control KRWIILGLNK

[b] standard deviation of MHC binding of two replicates

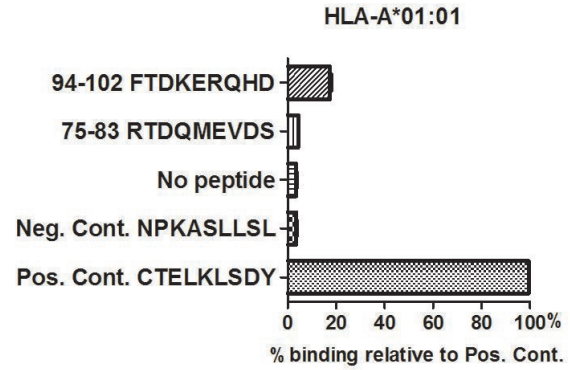
[c] cf. Fig. 10

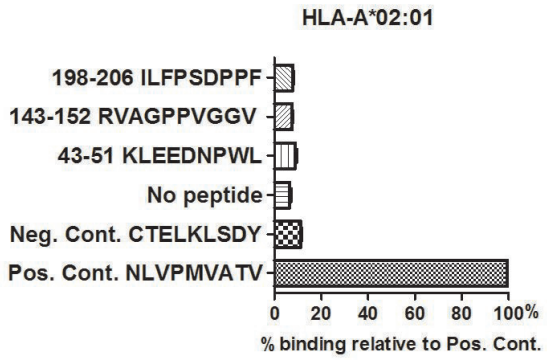
[d] average of the interaction energies of two independent MD simulations

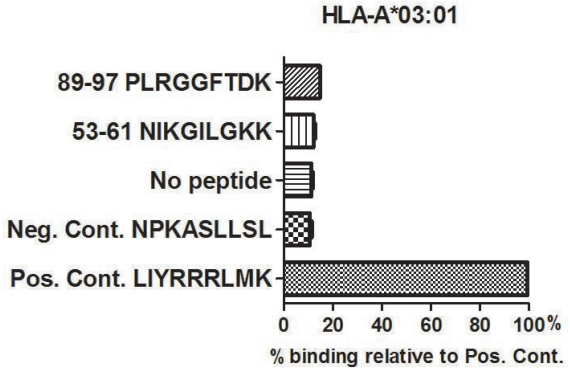
[e] standard deviation of the interaction energies of two independent MD simulations

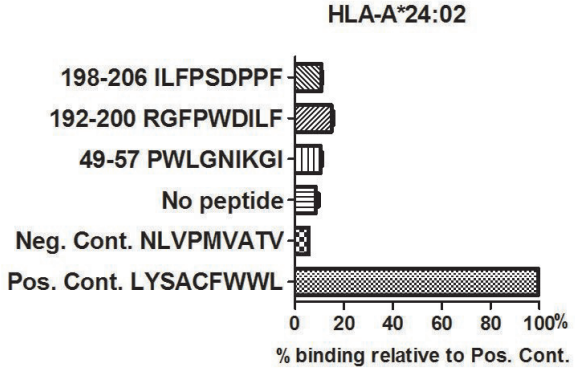
Table 5. List of primers used in L-HDAg amplification for direct and deep sequencing.

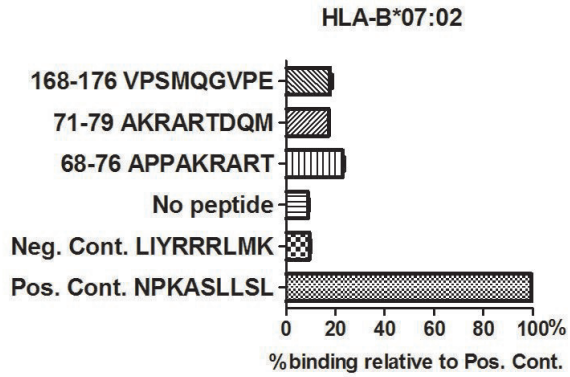
Primers	Sequence	Location
771R	5'-CGGTCCCCTCGGAATGTTG-3'	753-771 nt
339R	5'- GCTGAAGGGGTCCTCTGGAGGTG-3'	319-341 nt
1674R	5'-AGAAAAGAGTAAGAGYACTGAGG-3'	1652-1674 nt
912F	5'-GAGATGCCATGCCGACCCGAAGAG-3'	912-936 nt
891F	5'-AGGTCGGACCGCGAGGAGGT-3'	891-910 nt
M13 HDV 956	5'-GTTGTAAAACGACGGCCAGT TCACTGGGGTCGACAACTCTG-3'	
M13 HDV 1360	5'-CACAGGAAACAGCTATGACC GTAGACTCCGGACCTAGGAAGA-3'	
OligoAMIDM13fw	5'-**CGTATCGCCTCCCTCGCGCCATCAG** *MID*GTTGTAAAACGACGGCCAGT-3'	
OligoBMIDM13rv	5'-**CTATGCGCCTTGCCAGCCCGCTCAG** *MID*CACAGGAAACAGCTATGACC-3'	

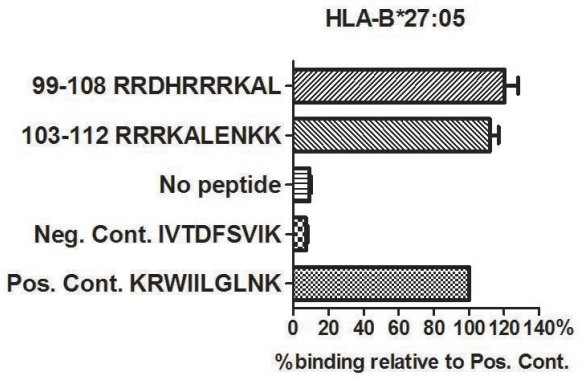


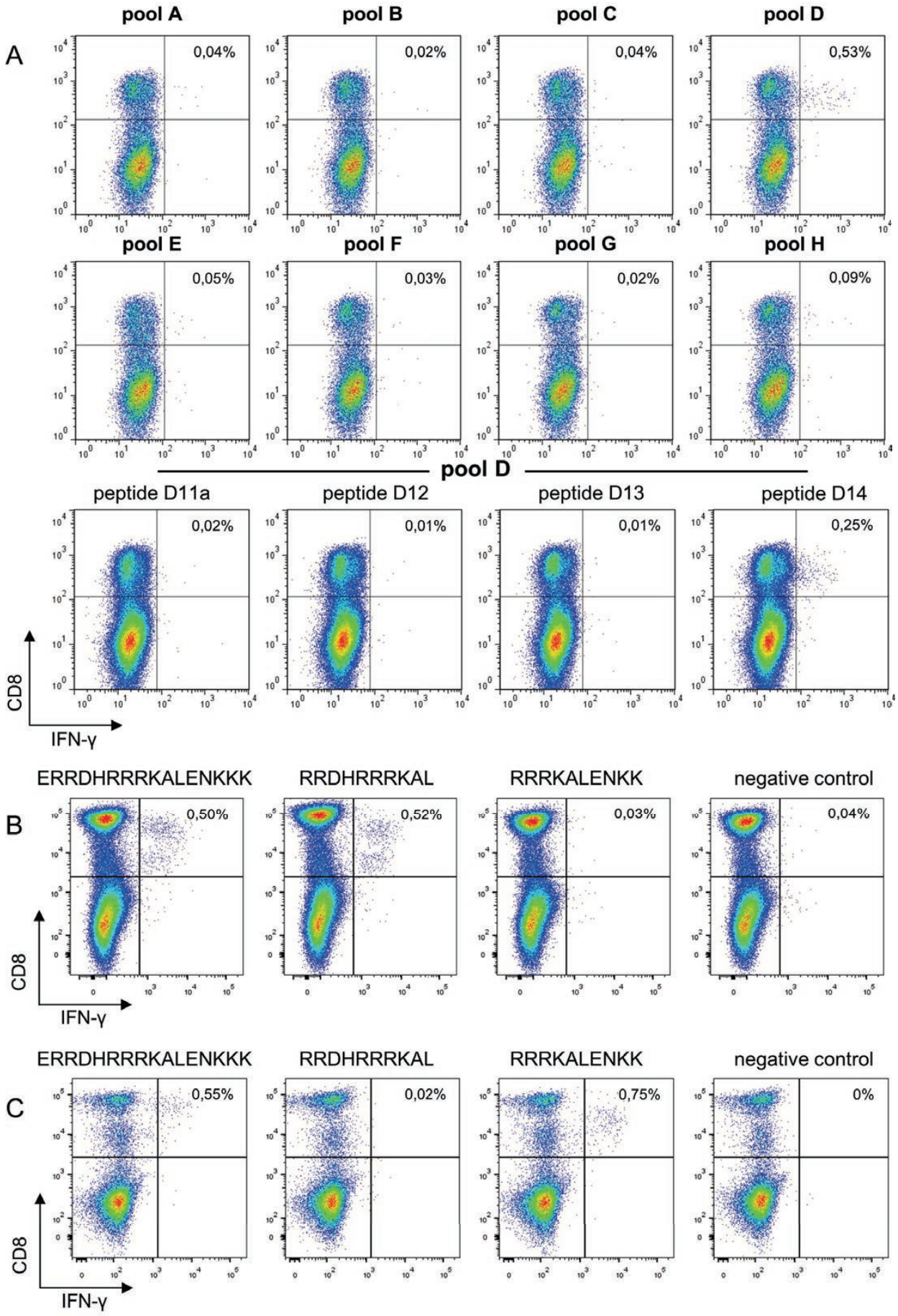


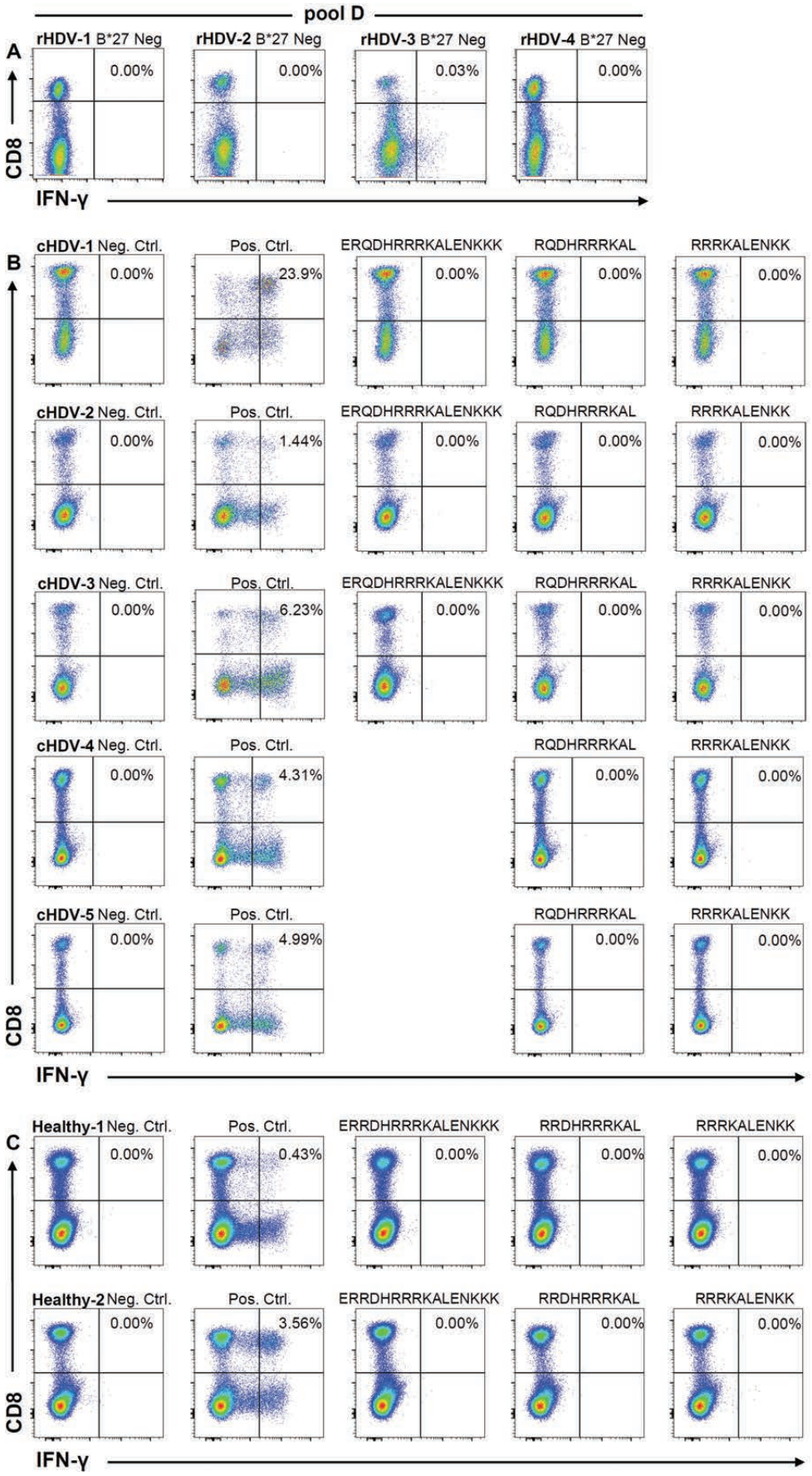


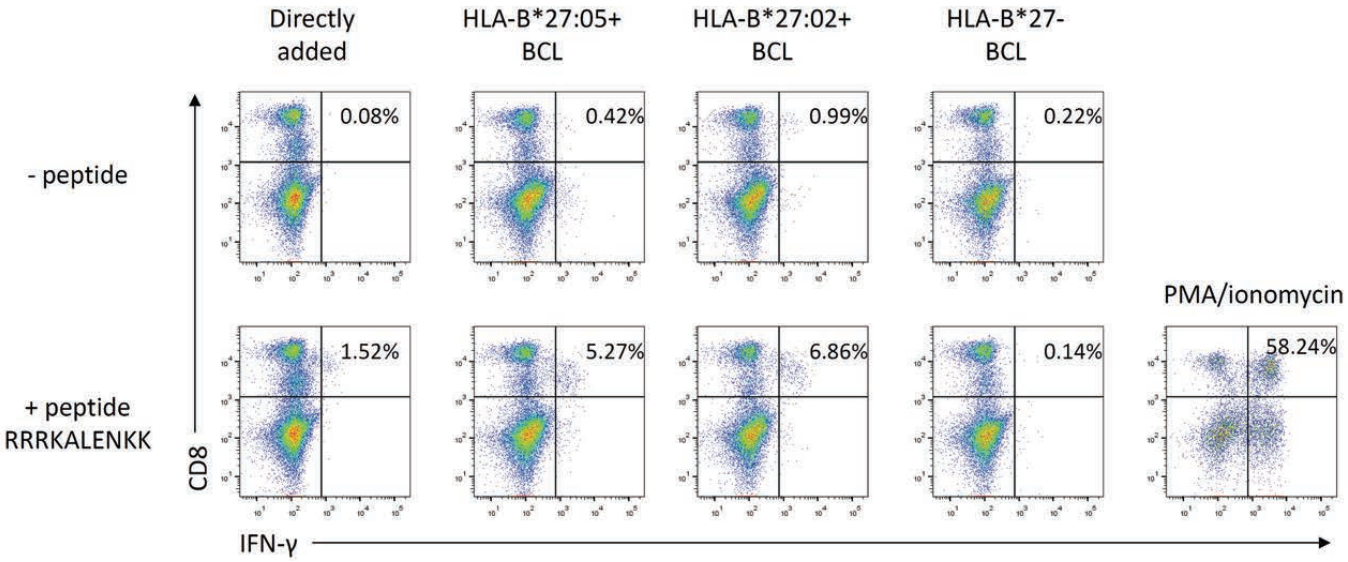


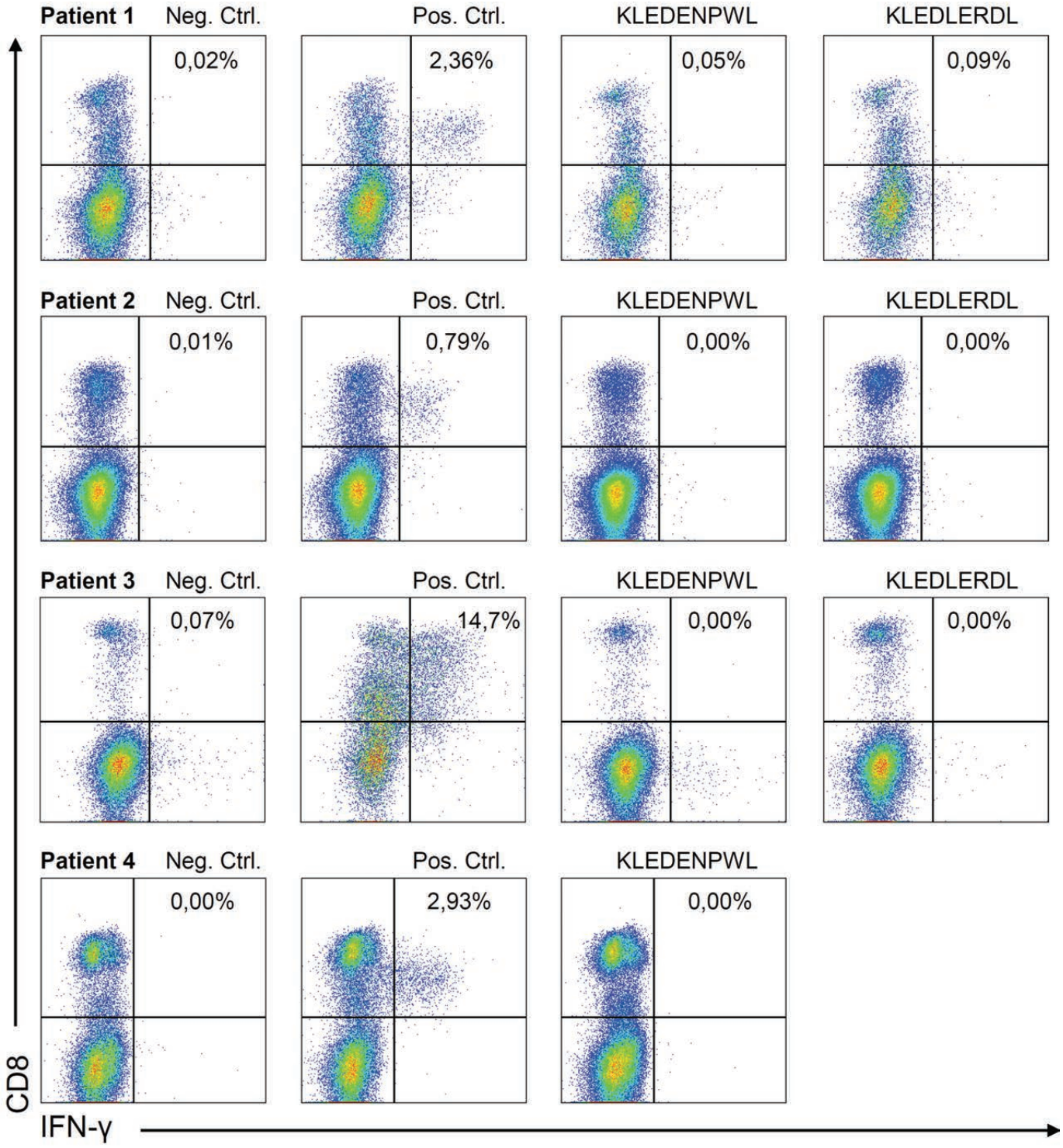


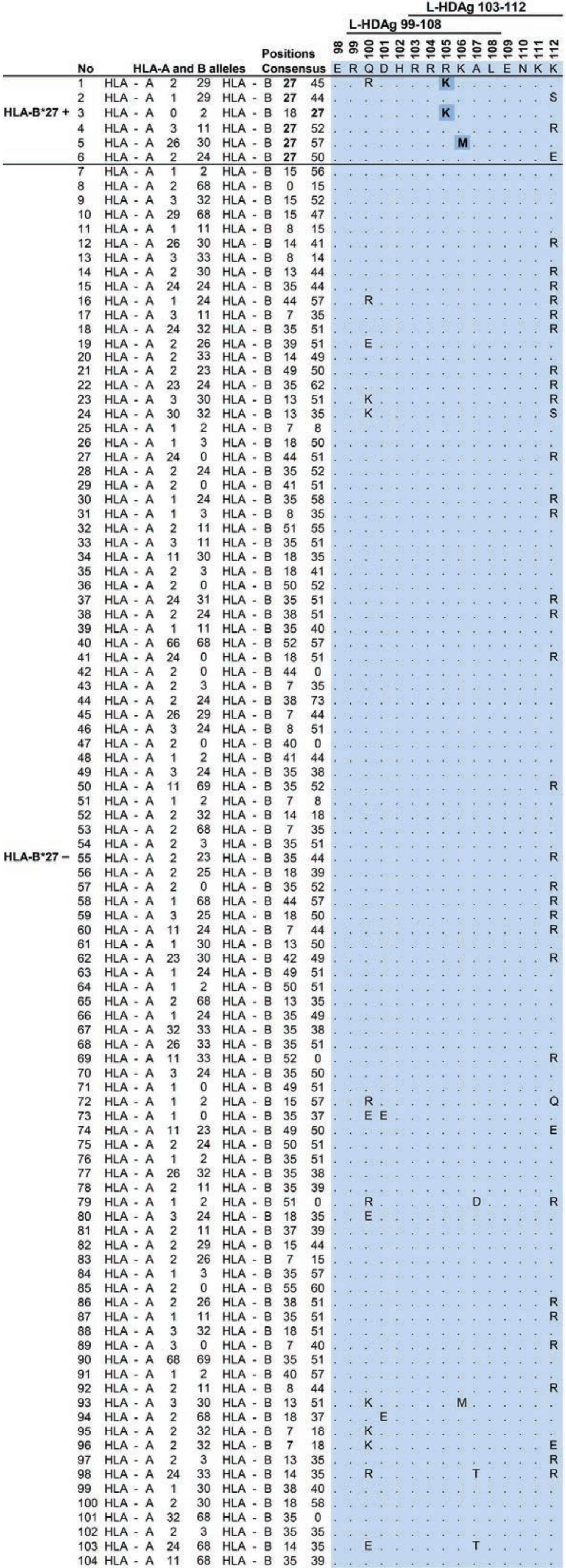










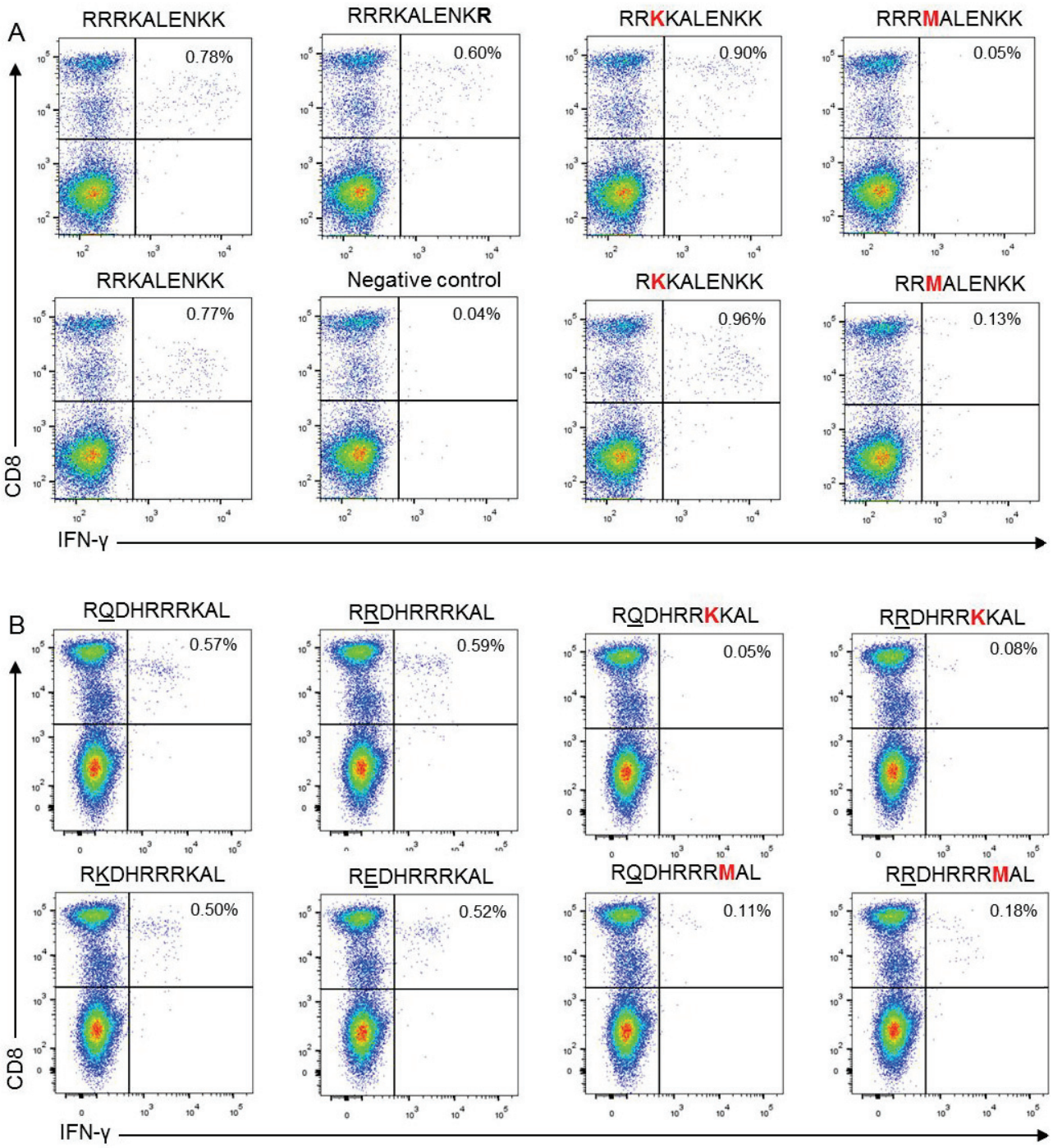


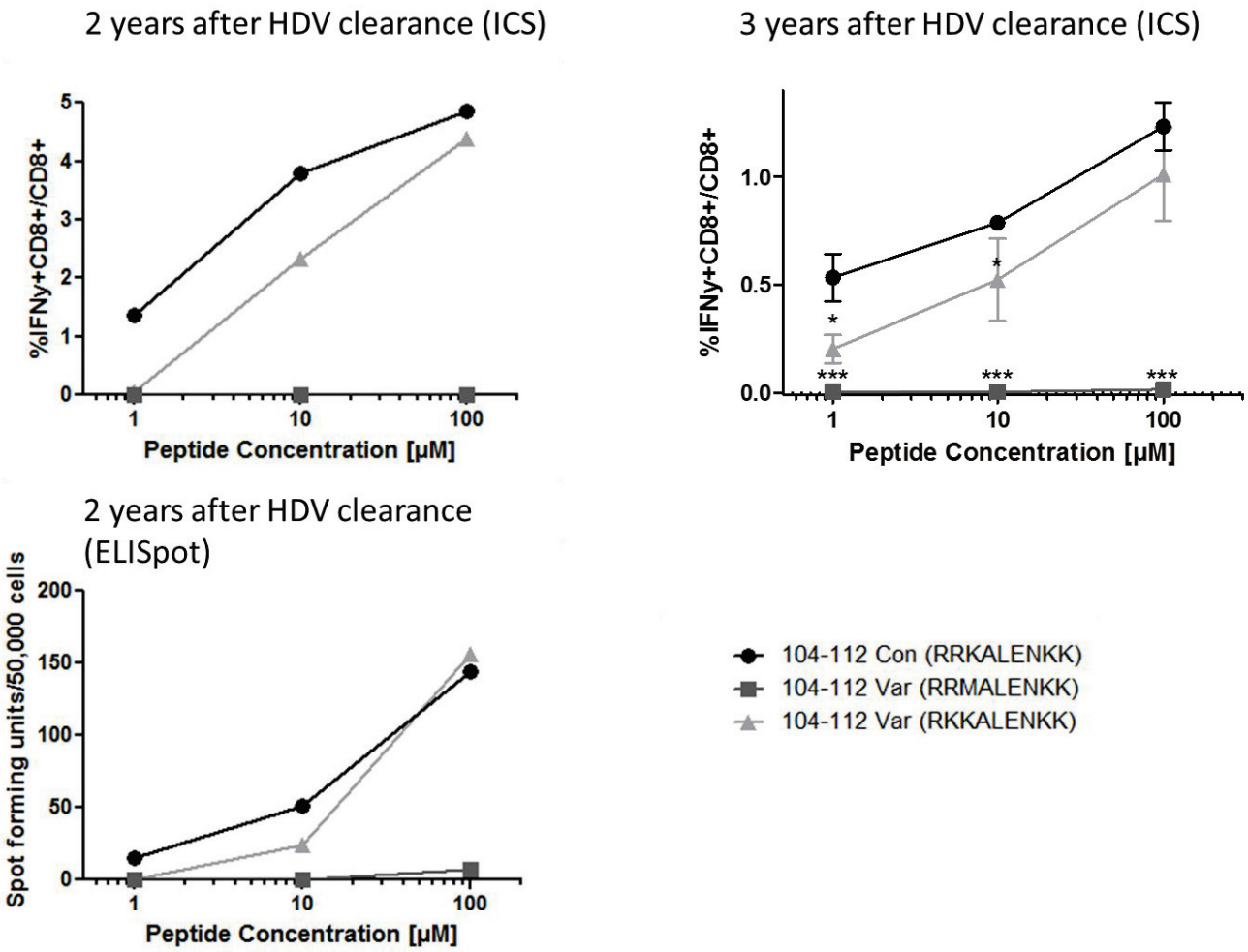
L-HDAg 103-112

L-HDAg 99-108

							Positions	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112
	No	HLA-A and B alleles					Consensus	E	R	Q	D	H	R	R	R	K	A	L	E	N	K	K
	1	HLA - A	2	29	HLA - B	**27**	45	.	.	R	.	.	.	.	**K**	.	.	.	.	.	.	.
	2	HLA - A	1	29	HLA - B	**27**	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S
HLA-B*27 +	3	HLA - A	0	2	HLA - B	18	**27**	.	.	.	.	.	.	.	**K**	.	.	.	.	.	.	.
	4	HLA - A	3	11	HLA - B	**27**	52	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	5	HLA - A	26	30	HLA - B	**27**	57	.	.	.	.	.	.	.	.	**M**	.	.	.	.	.	.
	6	HLA - A	2	24	HLA - B	**27**	50	.	.	.	.	.	.	.	.	.	.	.	.	.	.	E
	7	HLA - A	1	2	HLA - B	15	56	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	8	HLA - A	2	68	HLA - B	0	15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	9	HLA - A	3	32	HLA - B	15	52	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	10	HLA - A	29	68	HLA - B	15	47	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	11	HLA - A	1	11	HLA - B	8	15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	12	HLA - A	26	30	HLA - B	14	41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	13	HLA - A	3	33	HLA - B	8	14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	14	HLA - A	2	30	HLA - B	13	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	15	HLA - A	24	24	HLA - B	35	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	16	HLA - A	1	24	HLA - B	44	57	.	.	R	.	.	.	.	.	.	.	.	.	.	.	R
	17	HLA - A	3	11	HLA - B	7	35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	18	HLA - A	24	32	HLA - B	35	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	19	HLA - A	2	26	HLA - B	39	51	.	.	E	.	.	.	.	.	.	.	.	.	.	.	.
	20	HLA - A	2	33	HLA - B	14	49	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	21	HLA - A	2	23	HLA - B	49	50	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	22	HLA - A	23	24	HLA - B	35	62	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	23	HLA - A	3	30	HLA - B	13	51	.	.	K	.	.	.	.	.	.	.	.	.	.	.	R
	24	HLA - A	30	32	HLA - B	13	35	.	.	K	.	.	.	.	.	.	.	.	.	.	.	S
	25	HLA - A	1	2	HLA - B	7	8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	26	HLA - A	1	3	HLA - B	18	50	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	27	HLA - A	24	0	HLA - B	44	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	28	HLA - A	2	24	HLA - B	35	52	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	29	HLA - A	2	0	HLA - B	41	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	30	HLA - A	1	24	HLA - B	35	58	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	31	HLA - A	1	3	HLA - B	8	35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	32	HLA - A	2	11	HLA - B	51	55	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	33	HLA - A	3	11	HLA - B	35	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	34	HLA - A	11	30	HLA - B	18	35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	35	HLA - A	2	3	HLA - B	18	41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	36	HLA - A	2	0	HLA - B	50	52	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	37	HLA - A	24	31	HLA - B	35	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	38	HLA - A	2	24	HLA - B	38	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	39	HLA - A	1	11	HLA - B	35	40	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	40	HLA - A	66	68	HLA - B	52	57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	41	HLA - A	24	0	HLA - B	18	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	42	HLA - A	2	0	HLA - B	44	0	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	43	HLA - A	2	3	HLA - B	7	35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	44	HLA - A	2	24	HLA - B	38	73	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	45	HLA - A	26	29	HLA - B	7	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	46	HLA - A	3	24	HLA - B	8	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	47	HLA - A	2	0	HLA - B	40	0	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	48	HLA - A	1	2	HLA - B	41	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	49	HLA - A	3	24	HLA - B	35	38	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	50	HLA - A	11	69	HLA - B	35	52	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	51	HLA - A	1	2	HLA - B	7	8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	52	HLA - A	2	32	HLA - B	14	18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	53	HLA - A	2	68	HLA - B	7	35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	54	HLA - A	2	3	HLA - B	35	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HLA-B*27 –	55	HLA - A	2	23	HLA - B	35	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	56	HLA - A	2	25	HLA - B	18	39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	57	HLA - A	2	0	HLA - B	35	52	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	58	HLA - A	1	68	HLA - B	44	57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	59	HLA - A	3	25	HLA - B	18	50	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	60	HLA - A	11	24	HLA - B	7	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	61	HLA - A	1	30	HLA - B	13	50	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	62	HLA - A	23	30	HLA - B	42	49	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	63	HLA - A	1	24	HLA - B	49	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	64	HLA - A	1	2	HLA - B	50	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	65	HLA - A	2	68	HLA - B	13	35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	66	HLA - A	1	24	HLA - B	35	49	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	67	HLA - A	32	33	HLA - B	35	38	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	68	HLA - A	26	33	HLA - B	35	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	69	HLA - A	11	33	HLA - B	52	0	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	70	HLA - A	3	24	HLA - B	35	50	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	71	HLA - A	1	0	HLA - B	49	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	72	HLA - A	1	2	HLA - B	15	57	.	.	R	.	.	.	.	.	.	.	.	.	.	.	Q
	73	HLA - A	1	0	HLA - B	35	37	.	.	E	E	.	.	.	.	.	.	.	.	.	.	.
	74	HLA - A	11	23	HLA - B	49	50	.	.	.	.	.	.	.	.	.	.	.	.	.	.	E
	75	HLA - A	2	24	HLA - B	50	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	76	HLA - A	1	2	HLA - B	35	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	77	HLA - A	26	32	HLA - B	35	38	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	78	HLA - A	2	11	HLA - B	35	39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	79	HLA - A	1	2	HLA - B	51	0	.	.	R	.	.	.	.	.	.	D	.	.	.	.	R
	80	HLA - A	3	24	HLA - B	18	35	.	.	E	.	.	.	.	.	.	.	.	.	.	.	.
	81	HLA - A	2	11	HLA - B	37	39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	82	HLA - A	2	29	HLA - B	15	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	83	HLA - A	2	26	HLA - B	7	15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	84	HLA - A	1	3	HLA - B	35	57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	85	HLA - A	2	0	HLA - B	55	60	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	86	HLA - A	2	26	HLA - B	38	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	87	HLA - A	1	11	HLA - B	35	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	88	HLA - A	3	32	HLA - B	18	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	89	HLA - A	3	0	HLA - B	7	40	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	90	HLA - A	68	69	HLA - B	35	51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	91	HLA - A	1	2	HLA - B	40	57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	92	HLA - A	2	11	HLA - B	8	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	93	HLA - A	3	30	HLA - B	13	51	.	.	K	.	.	.	.	.	M	.	.	.	.	.	.
	94	HLA - A	2	68	HLA - B	18	37	.	.	.	E	.	.	.	.	.	.	.	.	.	.	.
	95	HLA - A	2	32	HLA - B	7	18	.	.	K	.	.	.	.	.	.	.	.	.	.	.	.
	96	HLA - A	2	32	HLA - B	7	18	.	.	K	.	.	.	.	.	.	.	.	.	.	.	E
	97	HLA - A	2	3	HLA - B	13	35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R
	98	HLA - A	24	33	HLA - B	14	35	.	.	R	.	.	.	.	.	.	T	.	.	.	.	R
	99	HLA - A	1	30	HLA - B	38	40	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	100	HLA - A	2	30	HLA - B	18	58	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	101	HLA - A	32	68	HLA - B	35	0	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	102	HLA - A	2	3	HLA - B	35	35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	103	HLA - A	24	68	HLA - B	14	35	.	.	E	.	.	.	.	.	.	T	.	.	.	.	.
	104	HLA - A	11	68	HLA - B	35	39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Consensus	**E**	**R**	**R**	**D**	**H**	**R**	**R**	**R**	**K**	**A**	**L**	**E**	**N**	**K**	**K**	**K**
Genotype 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	K	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	E	.	.	.	.	.	.	.	.	.	.	.	.	.
Genotype 2	.	.	E	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	Q	.	.	.	.	.	.	.	.	Q	.	.	.	.
Genotype 3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.
Genotype 4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Genotype 5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.
Genotype 6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Genotype 7,8	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.





Journal of Virology

JVI

Journal of Virology