

SCIENTIFIC REPORTS



OPEN

Identification of putative unique immunogenic ZIKV and DENV1-4 peptides for diagnostic cellular based tests

Aaron L. Oom, Davey Smith & Kevan Akrami

Since the re-emergence of Zika virus in 2014 and subsequent association with microcephaly, much work has focused on the development of a vaccine to halt its spread throughout the world. The mosquito vector that transmits this virus is widespread and responsible for the spread of other arboviridae including Dengue. Current diagnostic methods rely on serologic testing that are complicated by cross reactivity and therefore unable to distinguish Zika from Dengue infection in the absence of virus isolation. We performed an *in silico* analysis to identify potential epitopes that may stimulate a unique T-lymphocyte response to distinguish prior infection with Zika or Dengue. From this analysis, we not only identified epitopes unique to Zika and Dengue, but also identified epitopes unique to each Dengue serotype. These peptides contribute to a pool of peptides identified for vaccine development that can be tested *in vitro* to confirm immunogenicity, absence of homology and global population coverage. The current lack of accurate diagnostic testing hampers our ability to understand the scope of the epidemic, implications for vaccine implementation and complications related to mono-infection and co-infection with these two closely related viruses.

Zika virus (ZIKV) is a member of the Flavivirus genus related to Dengue virus (DENV) and was first isolated in 1947 from a Rhesus macaque in the Zika forest of Uganda¹. ZIKV infection commonly presents sub-clinically with only about 20% of infected acknowledging symptoms, and when symptoms are present they are usually fever, rash, and constitutional symptoms similar to those seen during DENV infections. These acute clinical similarities are confounded by a high coincidence of geographical ranges of the two infections² thanks to a shared mosquito vector, *Aedes aegypti*³.

Efforts to distinguish ZIKV and DENV infections, especially after the acute phase, are complicated by the large degree of genetic similarity and serologic responses between the two viruses⁴⁻⁹. Many serologic tests are available for ZIKV and DENV, including IgM antibody capture-ELISA (MAC-ELISA) and Plaque Reduction Neutralization Test (PRNT), but these platforms often cannot reliably distinguish between the two viruses due to antibody cross-reactivity and low sensitivity¹⁰. Cross-reactivity has hindered accurate epidemiological tracking of ZIKV. A commercially available ELISPOT for DENV has demonstrated increased sensitivity when compared to serologic or antigen assays though did not assess for potential cross-reactivity with ZIKV¹¹. This uncertainty makes clinical and epidemiological studies of ZIKV and DENV in human populations difficult to understand. Prior work with H1N1 [ref. 12], Hepatitis B¹³, and Mycobacterial infections^{14,15} has demonstrated the utility of epitope recognition in the development of novel diagnostics and vaccines. Therefore, we propose that detection of a specific lingering cytotoxic T-lymphocyte (CTL) memory response against Zika and/or Dengue, unlike antibody responses, would allow these infections to be distinguished. To this end we performed an *in silico* analysis of ZIKV non-structural protein 1 (NS1) and DENV non-structural protein 2a (NS2a) to determine virus-specific epitopes that could be used in the development of novel diagnostic tests to distinguish these infections. These proteins were chosen as previous data has suggested that DENV NS2a will have minimal ZIKV CTL cross reactivity^{16,17} and that ZIKV NS1 is largely ZIKV specific^{10,18,19}.

University of California, San Diego, Department of Infectious Disease, 9500 Gilman Drive #0711, La Jolla, CA, 92093-0711, USA. Correspondence and requests for materials should be addressed to K.A. (email: kakrami@ucsd.edu)

Results

Assays that can detect human CTL immune response to pathogens require specificity to unique pathogen proteins that are presented to CTLs *in vivo*. To be useful across the entire circulating pathogen population, these assays will need to use relatively conserved peptides in the pathogen of interest. This may be especially difficult for viruses, which cover a large range of genetic diversity. To tackle the challenge of creating a CTL assay that can distinguish ZIKV from DENV, we first collected all protein sequences of these viruses that are likely to be presented to the human CTL immune system during acute infection. We focused on NS1 for ZIKV and NS2a for DENV because these regions were found in previous studies to likely be immunogenic and unique regions with little cross-reactivity between ZIKV and DENV CTL responses^{10,16–19}. We then analyzed sequences for conservation across all known isolates for ZIKV and all DENV sequences available since 2010, which would increase the fidelity of any potential diagnostic assay. We used the following criteria to evaluate which peptides were most likely to generate a pathogen specific CTL response across the human population according to predicted ability to: (1) bind to 3 or more MHC-Class I alleles, (2) bind to 2 or more MHC-II alleles, (3) have a probability score of at least 0.8 (near match or perfect match) for natural processing of peptides and subsequent MHC-I binding in the cell, (4) be immunogenic according to which amino acids are present at positions 4–6 and the presence of large and aromatic side chains in 9-mers, (5) have population coverage of at least 60%, (6) dock in the binding groove of the MHC molecule, and (7) be specific to DENV or ZIKV (i.e. absence of homology). While not all predicted epitopes met the above criteria, they may be used in combination to form a ZIKV and DENV peptide pool to be tested for CTL response *in vitro*.

For the ZIKV NS1 protein sequence, we used a 2015 Brazilian isolate retrieved from the Protein Database (5GS6 [ref. 20]). To ensure that this sequence was representative of all known isolates, the sequence was inputted in a BLASTp analysis against all available full-length ZIKV protein sequences from the NCBI's Virus Variation Zika Database (182 sequences)²¹; both African and Asian clade samples were present to account for any potential clade differences within NS1. Every subject sequence covered residues 7 to 357 (C-terminus) and differed by only 1 residue at most against all ZIKV NS1 regions.

For the DENV NS2a, a consensus amino acid sequence from all 4 serotypes was constructed using DENV NS2a sequences collected between 2010 and 2016 from human hosts (563 sequences) obtained from the NIAID Virus Pathogen Database and Analysis Resource (ViPR)²². When analyzing this sequence using BLASTp, we found that the consensus sequence was biased towards DENV1 NS2a with 74–82% sequence similarity to DENV1 NS2a. When compared with the other serotypes (DENV2–4), sequence similarity ranged from 45–58%. The higher prevalence of DENV1 sequences in viral databases over other serotypes suggests that DENV1 may be over-represented when constructing a consensus sequence. Therefore, we chose to examine each DENV serotype in addition to the consensus sequence.

How a cell infected with either ZIKV or DENV processes peptides prior to the presentation and binding of peptides to MHC molecules is important to determine the potential for an epitope to interact with the CTL receptor. Following identification of the consensus amino acid sequences for ZIKV and DENV (all 4 serotypes and serotype specific), we next evaluated each of the peptides for their ability to undergo natural processing and bind to MHC I alleles. We then compared the potential immunogenicity of candidate peptides based on their biochemical structure (i.e. presence of aromatic rings)²³ as assessed by the IEDB immunogenicity tool.

MHC Peptide Processing and Binding. First, we evaluated the ability of potential epitopes in the ZIKV consensus strain 5GS6 NS1 to bind to MHC class I and II following MHC natural processing (Fig. 1a–c and Table 1). This provided us with 2 potential ZIKV peptides to then evaluate for immunogenicity, population coverage, docking and cross-reactivity with other flaviviridae including DENV (Table 1).

Second, we evaluated the consensus sequence of DENV NS2a and found three peptides that were predicted to bind class I and II molecules though were not predicted to be naturally processed (Fig. 2a–c and Table 2). Given the lack of an epitope in the DENV NS2a consensus sequence that fulfilled the initial binding and processing criteria, we focused on the identification of serotype specific peptides that could be used to identify prior DENV infection in a serotype specific way (Fig. 2a–c). Serotype specific identification in clinical practice is rarely performed as it requires isolation and molecular typing of DENV during acute infection. Therefore, we chose to identify potential peptides that may distinguish the serotype of prior infection. For DENV1, a single potential peptide, MLMTGTLAV/MLMTGTLAVF, was predicted to meet our MHC class I binding and natural processing criteria. For DENV2, we found one epitope, ILLVAVSFV, that fulfilled all criteria and another that met natural processing and MHC-I binding but not MHC-II binding, TMTDDIGMGV. For DENV3 and DENV4, we found 3 (HMIAGVFFTF, MIAGVFFTF/MIAGVFFTFV, LLLSGQITW) and 2 (KMSPGYVLGV, STMSLVMWV/STMSLVMWV) epitopes respectively that met our binding and processing criteria.

Immunogenicity and Homology. We then tested the epitopes for ZIKV and DENV identified above for sequence homology using BLASTp and immunogenicity using the IEDB immunogenicity tool, which has been validated for 9-mers that bind MHC-I alleles²³. For ZIKV NS1 we found one peptide, GVFHTSVWLK, predicted to be a MHC class I and II binder, naturally processed, and immunogenic. Following BLASTp search this epitope demonstrated homology with DENV serotypes 1, 3, and 4 (Table 1). From this, it became apparent that satisfaction of all of our criteria above for a unique ZIKV peptide without homology with DENV will be difficult to identify. Therefore, we repeated the binding analysis for ZIKV NS1 (requiring only MHC I binding of 3 or greater alleles and at least 0.8 probability of natural processing without MHC II binding), immunogenicity and homology testing. From this analysis, the epitope, KSYFVRAAK, appeared to be the best candidate peptide for eliciting a CTL response specific to ZIKV infections and distinct from DENV infections. Our revised search criteria also yielded the peptide, FLVEDHGFV, with immunogenicity and homology to a sequence from all 4 DENV

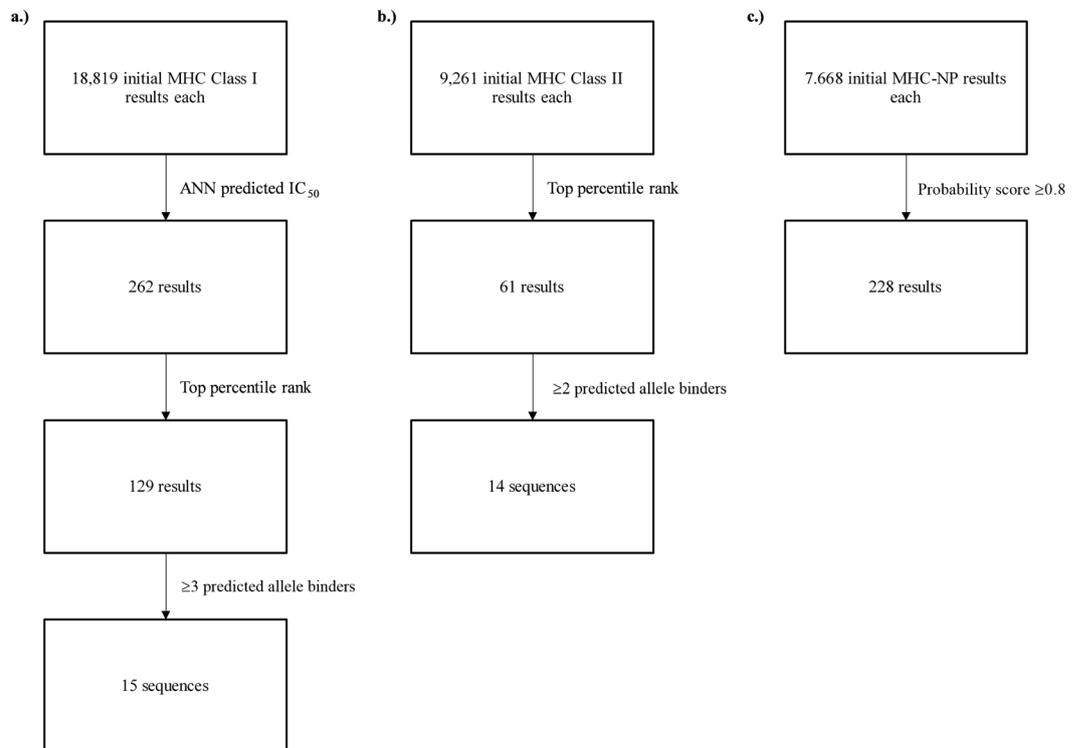


Figure 1. Workflow for ZIKV peptide antigenicity predictions. Briefly, chain A of 5GS6 from the Protein Database was submitted to the IEDB MHC class I and II binding tools and MHC-NP tool. (a) Class I results were first filtered by ANN predicted IC_{50} levels as recommended in Paul *et al.*³⁹ followed by filtering for those results in the top 1 percentile rank. Subsequent results were then sorted by sequence and those predicted to bind ≥ 3 class I alleles were chosen for further analysis. (b) Class II results were first filtered for those results in the top 1 percentile rank and then sorted by sequence with those predicted to bind ≥ 2 class II alleles chosen for further analysis. (c) MHC-NP initial results were filtered for those hits with a probability score of ≥ 0.8 . Cross-referencing of class II and MHC-NP data against class I data was performed to generate putative peptides (Table 1).

NS1 Residues	Sequence	HLA Class	Flaviviridae Homology (E-value <1)?	World Coverage	South America Coverage	Immunogenicity
16–24/25	KETRCGTGV/KETRCGTGVF*	I	DENV2-4 NS1	20.88%	10.34%	0.09664/0.13444
125–133	KSYFVRAAK*	I	None	38.48%	28.66%	0.2614
158–167	FLVEDHGFV**@	I	DENV1-4 NS1	41.35%	21.85%	0.31775
166–175	GVFHTSVWLK*	I & II	DENV1&3-4 Peptidase S7, NHUV glycoprot.	69.67%	78.11%	0.16657
169–177	HTSVWLKVR**	I & II	DENV1-3 NS1	58.80%	79.86%	0.03522
1059	KGPWHSEEL	I	None	Dar <i>et al.</i> ²⁶ Polyprotein position used in residue numbering All peptides bind >3 alleles		
878	VQLTVVVGVS	II	None			
878	VQLTVVVGSVKNPM	II	None			
965	VREDYSLEC	II	None			
982	VKGKEAVHS	II	None			
1004	WRLKRAHLI	II	None			
1045	LSHHNTREG	II	None			
1124	WYGMEIRPR	II	WNV NS1			
80–88	ILEENGVQL	I	None	Dikhit <i>et al.</i> ²⁵ NS1 protein position used in residue numbering All peptides bind >3 alleles		
87–95	QLTVVVGVS	I	None			
364–372	SLGVLVILL	I	None			
366–374	GVLVILLMV	I	None			
370–378	ILLMVQEGL	I	None			

Table 1. Predicted ZIKV NS1 epitopes from consensus ZIKV NS1 sequence, Dar *et al.*, and Dikhit *et al.* *Exact MHC-NP Matches. **Near MHC-NP Matches. @Due to high immunogenicity and DENV1-4 homology, this peptide may serve as a positive control for DENV/ZIKV infection. Immunogenicity >0 is the threshold for predicted immunogenicity.

serotypes in their NS1 region. Taken together, it is likely that this peptide may be used in a CTL based assay to identify those infected with either ZIKV and/or DENV.

We then evaluated the DENV NS2a potential binders for immunogenicity and homology with ZIKV. For the DENV NS2a consensus sequence, as above, we could not identify peptides that satisfied all 3 binding criteria. Therefore, we identified three peptides predicted to bind class I and II molecules but were not MHC-NP matches for further analysis. Of these three peptides, YQLWTTLLSL demonstrated high predicted immunogenicity and no significant homology detected against ZIKV sequences, suggesting a potential role as an epitope to detect prior infection with any of the 4 DENV serotypes. For the DENV1 NS2a predicted binding epitope, we found a positive immunogenicity score (Table 2) with no significant homology to ZIKV or DENV2-4 sequences, which suggests that this may be a unique peptide able to identify prior infection with DENV1. For DENV2, both sequences from the binding analysis were predicted to have positive immunogenicity scores and no significant homology to ZIKV or DENV1,3-4 sequences. For DENV3, of the three matches, two sequences were predicted to have highly positive immunogenicity scores. Interestingly, both of the hits (HMIAGVFFTF and MIAGVFFTF/MIAGVFFTFV) start around residues 31–33 of the NS2a protein, a small sequence starting range that appears to be antigenic in DENV1 and DENV2 as well, but not in DENV4 (Table 2). MIAGVFFTF/MIAGVFFTFV showed no significant homology to ZIKV or DENV1-2,4 sequences. For DENV4, only one of the two predicted binders, KMSPGYVLGV, was predicted to have a positive immunogenicity score (Table 2). Further, KMSPGYVLGV showed no significant homology to ZIKV or DENV1-3 sequences. In summary, our analysis revealed peptides for each DENV serotype that are immunogenic and unique for each serotype without apparent homology to ZIKV or other flaviviridae. Further, the ZIKV analysis revealed not only several immunogenic peptides that may be capable of distinguishing prior ZIKV from DENV infection, but also one immunogenic peptide that demonstrates shared homology between ZIKV and DENV NS1, which may trigger a CTL response in samples from those previously infected with ZIKV and/or DENV.

Predicted Population Coverage. Given the global scope of populations at risk for these arboviral infections, our epitopes must be capable of binding to a broad range of HLA alleles and therefore offer wide coverage of the human population. Thus, we evaluated the population coverage for our promising ZIKV and DENV peptides described above using the IEDB coverage tool (Tables 1 and 2). It appeared that ZIKV peptide, KSYFVRAAK, might have quite low coverage globally (38.5%) and in South America (28.7%), which suggests that multiple ZIKV peptides may be needed in the potential assay for optimal population coverage. For the DENV NS2a consensus sequence, YQLWTTLLSL was predicted to have world coverage of 78.77% and South American coverage of 86.15%. When evaluated by serotype, the highest predicted South American population coverage was 67.85% (DENV1), 57.84% (DENV2), 89.85% (DENV3), and 30.96% (DENV4). The highest predicted global coverage for each serotype was 86.75% (DENV1), 72.68% (DENV2), 94.28% (DENV3), and 51.64% (DENV4).

Docking Analysis. To further assess whether the identified epitopes would bind to HLA molecule grooves and thereby be presented to CTLs, we performed docking simulations of our predicted peptides. Docking analysis was performed for ZIKV epitopes (Fig. 3a–c) and DENV epitopes (Fig. 3d–h) to determine whether they are predicted to dock within the binding groove of HLA alleles. Both KSYFVRAAK (Fig. 3a) and FLVEDHGFGV (Fig. 3b) displayed typical or near typical binding, with FLVEDHGFGV predicted to have the most prominent arching, as has been previously described in MHC class I binding²⁴. This arching leads to exposure of the internal residues, crucial for MHC class I immunogenic responses²³, which may explain the high predicted immunogenicity of FLVEDHGFGV (Table 1). Conversely, the presentation of the terminal residues of GVFHTSVWLK (Fig. 3b) may explain its poor predicted immunogenicity (Table 1). We then evaluated whether the following predicted DENV epitopes were capable of docking within the binding groove of HLA alleles: MLMTGTLAVF, ILLVAVSFV, MIAGVFFTFV, MAIFIEEVMR, and MMLKLVNTNF (Fig. 3d–h). Similar to the ZIKV predictions, MIAGVFFTFV (Fig. 3f) and MAIFIEEVMR (Fig. 3g) appeared to best display internal residues integral to a class I immunogenic response, and correlate with high predicted immunogenicity (Table 2). The presentation of the C-terminal end of MMLKLVNTNF (Fig. 3h) may then explain its poor predicted immunogenicity (Table 2). The average predicted immunogenicity of the remaining Dengue peptides may be due to the varying degrees of presentation of internal residues (Fig. 3d and e). From the PatchDock analysis of the DENV4 peptide KMSPGYVLGV, there were no top results that predicted binding and/or interaction of the peptide within the HLA-A*02:01 binding groove. Docking only appeared at locations on the molecule outside of the binding groove.

Discussion

Current efforts to identify unique ZIKV epitopes for diagnostic and vaccine testing purposes have done little to consider cross-reactivity with DENV. It is clear, however, that distinguishing DENV and ZIKV past infections would be important for research purposes to better understand the long-term consequences of each infection and co-infection. Definitive associations may be difficult in the absence of a method to properly differentiate and diagnose the two infections. Ideally, a panel of peptides can be used in CTL assays that can distinguish prior ZIKV and DENV infection (including which serotype). Our *in silico* analysis identified several epitopes that may help distinguish prior ZIKV infection from DENV by measurement of lymphocyte activation in an ELISPOT assay. Moving forward, all identified peptides (Tables 1 and 2) will need to be tested *in vitro* to confirm whether these predicted peptides can stimulate a CTL response that is specific to each virus (ZIKV and DENV1-4).

One potential limitation of our peptide pool is relatively low predicted population coverage for ZIKV peptides. Therefore, a single assay with a combination of peptides will be needed to adequately cover the broad population at risk for ZIKV infection. To this end, we have included peptides proposed by Dikhit *et al.*²⁵ and Dar *et al.*²⁶ from similar *in silico* analyses (Table 1) that may be included in a peptide pool to diagnose prior ZIKV

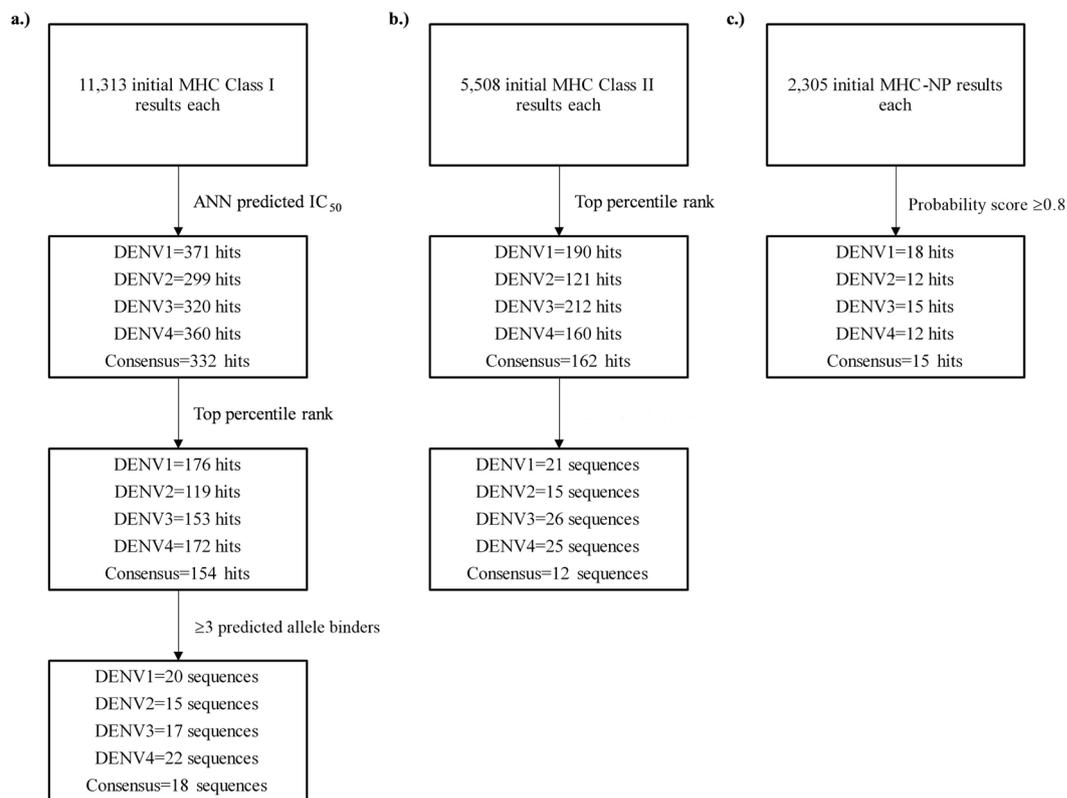


Figure 2. Workflow for DENV peptide antigenicity predictions. Briefly, each serotype NS2a consensus sequence as well as the DENV NS2a consensus sequence were submitted to the IEDB MHC class I and II binding tools and MHC-NP tool. **(a)** Class I results were first filtered by ANN predicted IC₅₀ levels as recommended in Paul *et al.*³⁹ followed by filtering for those results in the top 1 percentile rank. Subsequent results were then sorted by sequence and those predicted to bind ≥3 class I alleles were chosen for further analysis. **(b)** Class II results were first filtered for those results in the top 1 percentile rank and then sorted by sequence with those predicted to bind ≥3 class II alleles chosen for further analysis. **(c)** MHC-NP initial results were filtered for those hits with a probability score of ≥0.8. Cross-referencing of class II and MHC-NP data against class I data was performed to generate putative peptides (Table 2).

Serotype	Residues	Sequence	MHC II Binding?	ZIKV Homology (E-value <1)?	Other DENV Homology (E-value <1)?	World Coverage	S. America Coverage	Immunogenicity
1	31–39/40	MLMTGTLAV/MLMTGTLAVF**	Yes, <3 alleles	None	None	86.75%	67.85%	0.0651/0.09956
2	33–41	ILLVAVSFV**	Yes	None	None	72.68%	57.84%	0.04368
	63–72	TMTDDIGMGV*	No	None	None	41.35%	21.85%	0.08962
3	31–40	HMIAGVFFTF*	Yes	None	None	84.20%	83.18%	0.38745
	32–40/41	MIAGVFFTF/MIAGVFFTFV*	Yes	None	None	94.28%	89.85%	0.31868/0.41236
4	81–90	KMSPGYVLGV*	Yes	None	None	51.64%	30.96%	0.01384
Consensus	15–23/24	MAIFIEEVM/MAIFIEEVMR	Yes	None		44.73%	31.63%	0.49347/0.41459
	42–50/51	LLIMGQLTW/LLIMGQLTWR*	Yes, <3 alleles	NS4b		66.08%	78.82%	–0.19622/–0.06156
	86–94/95	MFAVGLLLR/MFAVGLLLRK	Yes	None		66.08%	78.82%	0.06096/0.08688
	132–140	MMLKLVTFN*	Yes, <3 alleles	None		34.12%	14.89%	–0.16356
	143–152	YQLWTLLSL	Yes	None		78.77%	86.15%	0.17843
	170–178	MVLAVVSLF**	Yes, <3 alleles	None		43.07%	27.68%	–0.03127

Table 2. Predicted DENV NS2a epitopes for DENV1–4 and consensus DENV NS2a sequence. *Exact MHC-NP match. **Near MHC-NP match. Immunogenicity >0 is the threshold for predicted immunogenicity.

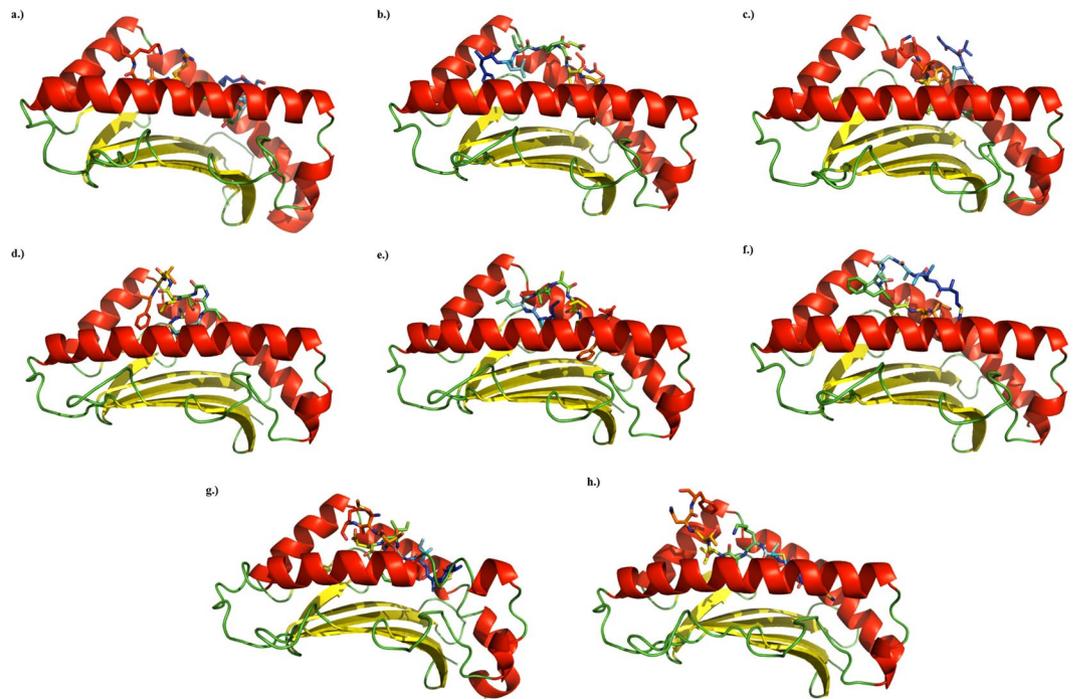


Figure 3. Top PatchDock predictions of epitope binding to MHC class I alleles as rendered by PyMOL Molecular Graphics System. Epitopes are shown as sticks with rainbow coloring (blue to red, N-terminus to C-terminus). HLA allele binding grooves are shown as cartoon with β -sheets in yellow, α -helices in red, and loops in green. DENV4 peptide KMSPGYVLGV did not bind within the binding groove of HLA-A*02:01 in any of the top ten PatchDock predictions. (a) ZIKV peptide KSYFVRAAK bound to HLA-A*03:01. (b) ZIKV peptide FLVEDHGFGV bound to HLA-A*02:01. (c) ZIKV peptide GVFHTSVWLK bound to HLA-A*03:01. (d) DENV1 peptide MLMTGTLAVF bound to HLA-A*02:01. (e) DENV2 peptide ILLVAVSFV bound to HLA-A*02:01. (f) DENV3 peptide MIAGVFFTFV bound to HLA-A*02:01. (g) DENV peptide MAIFIEEVMR bound to HLA-B*35:01. (h) DENV peptide MMLKLVTNF bound to HLA-B*15:01.

infections. These studies, while similar in their approach to the one used here, have a major limitation in their focus on vaccine development that did not consider cross-reactive CTL response with Dengue, which we have addressed (Table 1). A combination of the peptides proposed here—FLVEDHGFGV (ZIKV/DENV control), KSYFVRAAK and GVFHTSVWLK (ZIKV), MLMTGTLAVF (DENV1), ILLVAVSFV (DENV2), MIAGVFFTFV (DENV3), KMSPGYVLGV (DENV4), and MFAVGLLLRK and YQLWTTLLSL (DENV control)—in conjunction with those of others may greatly increase population coverage. Additionally, peptides predicted here to have low immunogenicity (for vaccine purposes) may have a role in a diagnostic setting as a higher concentration of peptide in the assay may overcome low immunogenicity (Table 3). This also speaks to the power of IEDB in peptide discovery for diagnostic purposes as well as vaccine purposes; other services such as Predivac²⁷ are more focused on vaccine generation and have a smaller range of capabilities and freedoms, typically coupling together analyses that IEDB handles independently of one another.

Our predicted peptides also focused on identifying peptides unique to each serotype of DENV, which has not been previously evaluated for diagnostic purposes. The ability to distinguish DENV serotype will be important for prospective studies as it is clear that certain serotypes and the sequence of infection with different serotypes have been more clearly associated with the development of Antibody Dependent Enhancement (ADE)²⁸. The serotype specific analysis of DENV peptides included here may also contribute to prospective studies to determine whether there are clinically significant ADE-mediated effects between individual DENV serotypes and ZIKV, as has been demonstrated *in vitro*^{6,7} and recently in patient plasma samples and mice²⁹.

Interestingly, when we analyzed all the potential peptides for the diagnostic test using the BLASTp, we found that the only peptide with sequence homology to other *Flaviviridae* family members was the ZIKV peptide GVFHTSVWLK, and it was an exact match to the Nhumirim virus (Table 1). However, only the original discoverers have sequenced the virus, collected in 2010 from Brazil, leaving open questions of how representative this particular sequence is of circulating Nhumirim virus³⁰.

The inclusion of molecular structures here offers one potential explanation for the predicted immunogenicity of each peptide. Those peptides that most closely resemble the expected binding of the C- and N-terminal residues in the binding groove and the protruding arch-like presentation of internal residues²⁴ are also those that are predicted to have the highest immunogenicity. These models, while potentially illuminating, have their limits. Even when limiting the HLA allele to only the binding groove as the receptor rather than the full molecule, docking predictions still have varying success in placing the peptide within the binding groove rather than adjacent to it. Whether this is a reflection on the IEDB binding prediction or on the PatchDock model itself is not clear.

Peptide	Residues	Sequence	HLA Class	Flaviviridae Homology (E-value <1)?	World Coverage	S. America Coverage	Immunogenicity
ZIKV/DENV control	125–133	KSYFVRAAK*	I	None	38.48%	28.66%	0.2614
ZIKV	158–167	FLVEDHGFV**	I	DENV1-4 NS1	41.35%	21.85%	0.31775
ZIKV	166–175	GVFHTSVWLK*	I & II	DENV1&3-4 Peptidase S7, NHUV glycoprot.	69.67%	78.11%	0.16657
DENV1	31–39/40	MLMTGTLAV/MLMTGTLAVF**	I & II	None	86.75%	67.85%	0.0651/0.09956
DENV2	33–41	ILLVAVSFV**	I & II	None	72.68%	57.84%	0.04368
DENV3	32–40/41	MIAGVFFTF/MIAGVFFTFV*	I & II	None	94.28%	89.85%	0.31868/0.41236
DENV4	81–90	KMSPGYVLGV*	I & II	None	51.64%	30.96%	0.01384
DENV Control	86–94/95	MFAVGLLLR/MFAVGLLLRK	I & II	None	66.08%	78.82%	0.06096/0.08688
	143–152	YQLWTTLLSL	I & II	None	78.77%	86.15%	0.17843

Table 3. Proposed diagnostic kit peptide summary. *Exact MHC-NP match. **Near MHC-NP match. Immunogenicity >0 is the threshold for predicted immunogenicity.

Validation of both IEDB and PatchDock results will require *in vitro* testing to determine CTL responsiveness to each peptide as well as to evaluate for cross-reactivity.

In areas endemic to ZIKV and DENV, Chikungunya virus has emerged as yet another virus often transmitted through *A. aegypti* that produces clinical symptoms similar to DENV and ZIKV infection³¹. If the current peptides are confirmed to generate specific ZIKV or DENV cytokine release and lead to a diagnostic platform, people with congruent symptoms but with negative results on this test would need to be screened for CHIKV infection. Importantly, while clinical similarities are manifest, CHIKV is of an entirely separate viral family altogether—*Togaviridae*—and molecular misdiagnosis is unlikely. In contrast to the serologic cross reactivity apparent between the flaviviridae, CHIKV triggers an immunologic response that can be distinguished from ZIKV or DENV with serological testing alone.

Here, we have demonstrated the utility of *in silico* analysis to identify a pool of peptides that may be important in the development of a CTL based test to diagnose prior infection with either DENV (including individual serotype) or ZIKV. This diagnostic tool may have implications for Dengue and Zika vaccine implementation and risk of ADE^{32,33}, women of child bearing age regarding the risk of microcephalus and identification of any long-term complications related to co-infection. Such a test will require rigorous evaluation in a clinical setting to determine efficacy and epidemiological power.

Methods

ZIKV NS1 Protein Sequence. The ZIKV NS1 protein sequence used here is a 2015 Brazilian isolate and was retrieved from the Protein Database (5GS6 [ref. 20]).

DENV1-4 NS2a Protein Sequences. For DENV, a consensus amino acid sequence of NS2a from all 4 serotypes was constructed on Geneious version 10.0.6 [ref. 34] at a 25% threshold using DENV NS2a sequences. DENV NS2a sequences collected between 2010 and 2016 from human hosts (563 sequences) were obtained from and aligned by the NIAID Virus Pathogen Database and Analysis Resource (ViPR)²² through the web site at <http://www.viprbrc.org/>. We limited our search from 2010 to 2016, as this represents the consensus sequence of the most likely circulating strains of DENV. We obtained and aligned individual serotype consensus sequences through ViPR followed by generation of a consensus sequence for each serotype at a 50% threshold using Geneious version 10.0.6.

MHC Class I and II Binding. Potential unique epitopes that bind to MHC Class I and II were identified using the Immune Epitope Database and Analysis Resource (IEDB) (<http://www.iedb.org/>) [ref. 35]. IEDB hosts a suite of analysis tools for epitope prediction including MHC class I and II binding predictive analysis through biophysical comparison, rather than sequence homology, of known class I and II epitopes to potential epitopes^{36–38}. The MHC class I binding tool has been well-characterized by the IEDB group with defined IC₅₀ thresholds as determined by the Artificial Neural Network method (ANN) for many of the most common human HLA class I alleles³⁹. Chain A of 5GS6 and the generated DENV NS2a consensus sequences were tested against the MHC class I and II reference allele sets that are designed to cover ~97% and 99% of the world population, respectively^{40,41} (Supplemental Tables S1 and S2). The approach for ZIKV and DENV peptide MHC class I and II binding are detailed as previously described (Figs 1a,b and 2a,b).

MHC Natural Processing. The IEDB MHC Natural Processing (MHC-NP) tool predicts naturally processed epitopes based on physicochemical properties and residue position comparisons to known epitopes⁴². Chain A of 5GS6 and the generated DENV NS2a consensus sequences were submitted to the MHC-NP tool with default parameters against HLA-A*02:01, HLA-B*07:02, HLA-B*35:01, HLA-B*44:03, HLA-B*53:01, and HLA-B*57:01 for predicted 9-mers and 10-mers. Results were filtered for hits with a probability score of 0.8 or greater and cross referenced against MHC I binding data for matching sequences regardless of predicted binding allele. This compensated for the limited allele predictive power of MHC-NP. Exact matches between the data sets as well as near matches—those MHC-NP hits that were 1 residue off from the class I starting position or predicted sequence length (9-mer vs 10-mer)—were noted (Figs 1c and 2c).

Population Coverage. Population coverage was determined for promising peptides using the IEDB population coverage tool⁴³. For those peptides that were non-MHC class II predicted binders, only the MHC class I setting was used. The combined MHC class I and II setting was used for peptides predicted to bind both MHC class I and II. Population coverage was calculated for the world population as well as South America. Restricted alleles came from the IEDB MHC class I and II predictive binding tools. For all class II alleles, only the beta chains were used in calculating frequency as a proxy of the alpha/beta combined allele in the IEDB reference set given that alpha chains are not considered to contribute significantly towards the binding.

Immunogenicity Tool. MHC class I and II binding along with natural processing were used to determine potential antigenicity of peptides. To predict immunogenicity of chosen peptides, the IEDB class I immunogenicity tool was used²³. This tool examines immunogenicity of MHC class I and epitope complexes placing emphasis on residues 4–6 and their physicochemical properties. Default settings for the tool were employed.

Protein Basic Local Alignment Search Tool (BLASTp). Promising ZIKV peptides, those that performed well for multiple parameters, were analyzed in a BLASTp query with default parameters against each individual DENV serotype using all available full-length polyprotein sequences on Virus Variation for each collected from 2010–2016 [ref. 21]. Promising DENV peptides, similarly defined, were analyzed in a BLASTp query with default parameters against all available ZIKV full-length polyprotein sequences from Virus Variation²¹. Serotype specific DENV peptides were also analyzed against serotype separated DENV full-length polyprotein sequences. An E-value of less than 1 was used to define significantly homologous sequences in all instances. As an added measure, all putative peptides were analyzed in a BLASTp query with default parameters against all *Flaviviridae* family member sequences in the BLAST database.

Docking Modeling Studies. Predictive docking models were generated using PatchDock^{44,45} to further support predicted antigenicity and immunogenicity in IEDB. Three-dimensional structures of chosen ZIKV and DENV peptides were first generated using the “Natural Peptides Module for Beginners” feature of PEPstrMOD with default settings^{46,47}. PEPstrMOD results were then submitted to PatchDock along with a predicted binding HLA allele using the default settings. HLA alleles were chosen based on availability of high-quality protein crystal structures and predicted population coverage of the allele as demonstrable examples with the greatest possible relevance (114F [ref. 48]; 2XPG [ref. 49]; 1XR8 [ref. 50]; 3LKO [ref. 51]). The highest-ranking result for each peptide that showed predicted binding within the HLA binding groove is shown here. All images were rendered using the PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC.

References

- Dick, G. W., Kitchen, S. F. & Haddock, A. J. Zika virus (I). isolations and serological specificity. *Trans. R. Soc. Trop. Med. Hyg.* **46**, 509–520 (1952).
- Musso, D. & Gubler, D. J. Zika virus. *Clin. Microbiol. Rev.* **29**, 487–524 (2016).
- Iovine, N. *et al.* Coinfection with Zika and Dengue-2 viruses in a traveler returning from Haiti, 2016: clinical presentation and genetic analysis. *Clin. Infect. Dis.* **64**, 72–75 (2016).
- Adiga, R. Phylogenetic analysis of the NS5 gene of Zika virus. *J. Med. Virol.* **88**, 1821–1826 (2016).
- Lanciotti, R. S., Lambert, A. J., Holodniy, M., Saavedra, S. & Signor, LdelC. C. Phylogeny of Zika virus in western hemisphere, 2015. *Emerg. Infect. Dis.* **22**, 933–935 (2016).
- Dejnirattisai, W. *et al.* Dengue virus sero-cross-reactivity drives antibody-dependent enhancement of infection with Zika virus. *Nat. Immunol.* **17**, 1102–1108 (2016).
- Priyamvada, L. *et al.* Human antibody responses after Dengue virus infection are highly cross-reactive to Zika virus. *Proc. Natl. Acad. Sci. USA* **113**, 7852–7857 (2016).
- Bhatt, S. *et al.* The global distribution and burden of dengue. *Nature*. **496**, 504–507 (2013).
- de Teive e Argolo, A. F. L. *et al.* High frequency of pre-existing neutralizing antibody responses in patients with dengue during an outbreak in Central Brazil. *BMC Infect. Dis.* **16**, 546, doi:10.1186/s12879-016-1867-6 (2016).
- Wong, S. J. *et al.* A multiplex microsphere immunoassay for Zika virus diagnosis. *EBioMed.*; j.ebiom.2017.01.008 (2017).
- Piedrahita, L. D. *et al.* Evaluation of commercially available assays for diagnosis of acute dengue in schoolchildren during an epidemic period in Medellín, Colombia. *Am. J. Trop. Med. Hyg.* **95**, 315–321 (2016).
- Lohia, N. & Baranwal, M. Identification of conserved peptides comprising multiple T cell epitopes of Matrix 1 protein in H1N1 influenza virus. *Viral Immunol.* **28**, 570–579 (2015).
- Brinck-Jensen, N.-S., Vorup-Jensen, T., Leutscher, P. D. C., Erikstrup, C. & Petersen, E. Immunogenicity of twenty peptides representing epitopes of the hepatitis B core and surface antigens by IFN- γ response in chronic and resolved HBV. *BMC Immunol.* **16**, doi:10.1186/s12865-015-0127-7 (2015).
- Devasundaram, S., Deenadayalan, A. & Raja, A. *In silico* analysis of potential human T cell antigens from *Mycobacterium tuberculosis* for the development of subunit vaccines against tuberculosis. *Immunol. Invest.* **43**, 137–159 (2014).
- Eirin, M. E. *et al.* Identification and evaluation of new *Mycobacterium bovis* antigens in the *in vitro* interferon gamma release assay for bovine tuberculosis diagnosis. *Tuberculosis* **95**, 795–801 (2015).
- Liu, Y., Liu, J. & Cheng, G. Vaccines and immunization strategies for dengue prevention. *Emerg. Microbes Infect.* **5**, e77, doi:10.1038/emi.2016.74 (2016).
- Mathew, A. *et al.* Dominant recognition by human CD8+ cytotoxic T lymphocytes of dengue virus nonstructural proteins NS3 and NS1.2a. *J. Clin. Invest.* **98**, 1684–1691 (1996).
- Stettler, K. *et al.* Specificity, cross-reactivity, and function of antibodies elicited by Zika virus infection. *Science*. **353**, 823–826 (2016).
- Steinhagen, K. *et al.* Serodiagnosis of Zika virus (ZIKV) infections by a novel NS1-based ELISA devoid of cross-reactivity with Dengue virus antibodies: a multicohort study of assay performance, 2015 to 2016. *Eurosurveillance* **21**, doi:10.2807/1560-7917.ES.2016.21.50.30426 (2016).
- Xu, X. *et al.* Contribution of intertwined loop to membrane association revealed by Zika virus full-length NS1 structure. *EMBO. J.* **35**, 2170–2178 (2016).
- Brister, J. R. *et al.* Virus variation resource – recent updates and future directions. *Nucleic Acids Res* **42**, D660–665, doi:10.1093/nar/gkt1268 (2014).

22. Pickett, B. E. *et al.* ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic Acids Res.* **40**, D593–598, doi:10.1093/nar/gkr859 (2012).
23. Callis, J. J. A. *et al.* Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput. Biol.* **8**, e1003266, doi:10.1371/journal.pcbi.1003266 (2013).
24. Rossjohn, J. *et al.* T cell antigen receptor recognition of antigen-presenting molecules. *Annu. Rev. Immunol.* **33**, 169–200 (2015).
25. Dikhit, M. R. *et al.* Computational prediction and analysis of potential antigenic CTL epitopes in Zika virus: a first step towards vaccine development. *Infect. Genet. Evol.* **45**, 187–197 (2016).
26. Dar, H. *et al.* Prediction of promiscuous T-cell epitopes in the Zika virus polyprotein: an *in silico* approach. *Asian Pac. J. Trop. Med.* **9**, 844–850 (2016).
27. Oyarzun, P., Ellis, J. J., Boden, M. & Kobe, B. PREDIVAC: CD4+ T-cell epitope prediction for vaccine design that covers 95% of HLA class II protein diversity. *BMC Bioinformatics* **14**, doi:10.1186/1471-2105-14-52 (2013).
28. Soo, K.-M., Khalid, B., Ching, S.-M. & Chee, H.-Y. Meta-analysis of dengue severity during infection by different dengue virus serotypes in primary and secondary infections. *PLoS ONE* **11**, e0154760, doi:10.1371/journal.pone.0154760 (2016).
29. Bardina, S. V. *et al.* Enhancement of Zika virus pathogenesis by preexisting antinflavivirus immunity. *Science*. **356**, 175–180 (2017).
30. Pauvolid-Corrêa, A. *et al.* Nhumirim virus, a novel flavivirus isolated from mosquitoes from the Pantanal, Brazil. *Arch. Virol.* **160**, 21–27 (2015).
31. Sahadeo, N. S. D. *et al.* Understanding the evolution and spread of chikungunya virus in the Americas using complete genome sequences. *Virus Evol.* **3**, 1–10 (2017).
32. Halstead, S. B. Critique of World Health Organization recommendation of a Dengue vaccine. *J. Infect. Dis.* **214**, 1793–1795 (2016).
33. Aguiar, M., Halstead, S. B. & Stollenwerk N. Consider stopping dengvaxia administration without immunological screening. *Expert Rev. Vaccines*, doi:10.1080/14760584.2017.1276831 (2016).
34. Kearse, M. *et al.* Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
35. Vita, R. *et al.* The immune epitope database (IEDB) 3.0. *Nucleic Acids Res.* **43**, D405–412, doi:10.1093/nar/gku938 (2015).
36. Moutaftsi, M. *et al.* A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus. *Nat Biotechnol.* **24**, 817–819 (2006).
37. Wang, P. *et al.* A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Comput. Biol.* **4**, e1000048, doi:10.1371/journal.pcbi/1000048 (2008).
38. Wang, P. *et al.* Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinformatics*. **11**, 568, doi:10.1186/1471-2105-11-568 (2010).
39. Paul, S. *et al.* HLA class I alleles are associated with peptide-binding repertoires of different size, affinity, and immunogenicity. *J. Immunol.* **191**, 5831–5839 (2013).
40. Weiskopf, D. *et al.* Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. *Proc. Natl. Acad. Sci. USA* **110**, E2046–2053, doi:10.1073/pnas.1305227110 (2013).
41. Greenbaum, J. *et al.* Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. *Immunogenetics* **63**, 325–335 (2011).
42. Giguère, S. *et al.* MHC-NP: predicting peptides naturally processed by the MHC. *J. Immunol. Methods*. **400–401**, 30–36 (2013).
43. Bui, H. H. *et al.* Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics*. **7**, 153, doi:10.1186/1471-2105-7-153 (2006).
44. Duhovny D., Nussinov R. & Wolfson H. J. Efficient unbound docking of rigid molecules in *Algorithms in Bioinformatics. WABI 2002. Lecture Notes in Computer Science, vol 2452* (eds. Guigó, R., Gusfield, D.) 185–200 (Springer, 2002).
45. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucl. Acids. Res.* **33**, W363–367 (2005).
46. Singh, S. *et al.* PEPstrMOD: structure prediction of peptides containing natural, non-natural and modified residues. *Biol. Dir.* **10**, 73, doi:10.1186/s13062-015-0103-4 (2015).
47. Kaur, H., Garg, A. & Raghava, G. P. S. PEPstr: a de novo method for tertiary structure prediction of small bioactive peptides. *Protein. Pept. Lett.* **14**, 626–631 (2007).
48. Hillig, R. C. *et al.* High-resolution structure of HLA-A*0201 in complex with a tumour-specific antigenic peptide encoded by the MAGE-A4 gene. *J. Mol. Biol.* **310**, 1167–1176 (2001).
49. McMahon, R. M. *et al.* Structure of Hla-A0301 in complex with a peptide of proteolipid protein: insights into the role of Hla-A alleles in susceptibility to multiple sclerosis. *Acta Crystallogr. Sect. D*. **67**, 447–454 (2011).
50. Roder, G. *et al.* Crystal structures of two peptide-HLA-B*1501 complexes; structural characterization of the HLA-B62 supertype. *Acta Crystallogr. Sect. D*. **62**, 1300–1310 (2006).
51. Gras, S. *et al.* Cross-reactive CD8+ T-cell immunity between the pandemic H1N1-2009 and H1N1-1918 influenza A viruses. *Proc. Natl. Acad. Sci. USA* **107**, 12599–12604 (2010).

Acknowledgements

The authors wish to thank Dr. Sergio Ita for providing technical support in rendering images for publication.

Author Contributions

A.L.O. performed the epitope prediction and molecular docking analyses, K.A. reviewed and critiqued the analyses. A.L.O. and K.A. wrote the main manuscript text. D.S. critically reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-05980-z

Competing Interests: A.L.O. is funded through a fellowship established by Genentech Inc. for the UCSD Biomedical Sciences Ph.D. program. K.A. and D.S. have no conflicts of interest to declare.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the

material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017