OPEN

Crystal structure of suboptimal viral fragments of Epstein Barr Virus Rta peptide-HLA complex that stimulate CD8 T cell response

Xuelu Huan[1], Ziyi Zhuo[1], Ziwei Xiao[1] & Ee Chee Ren[1,2]*

Peptides presented by Human leukocyte antigen (HLA) class-I molecules are generally 8–10 amino acids in length. However, the predominant pool of peptide fragments generated by proteasomes is less than 8 amino acids in length. Using the Epstein - Barr virus (EBV) Rta-epitope (ATIGTAMYK, residues 134–142) restricted by HLA-A*11:01 which generates a strong immunodominant response, we investigated the minimum length of a viral peptide that can constitute a viral epitope recognition by CD8 T cells. The results showed that Peripheral blood mononuclear cells (PBMCs) from healthy donors can be stimulated by a viral peptide fragment as short as 4-mer (AMYK), together with a 5-mer (ATIGT) to recapitulate the full length EBV Rta epitope. This was confirmed by generating crystals of the tetra-complex (2 peptides, HLA and β2-microglobulin). The solved crystal structure of HLA-A*11:01 in complex with these two short peptides revealed that they can bind in the same orientation similar to parental peptide (9-mer) and the free ends of two short peptides acquires a bulged conformation that is directed towards the T cell receptor. Our data shows that suboptimal length of 4-mer and 5-mer peptides can complement each other to form a stable peptide-MHC (pMHC) complex.

Human leukocyte antigen class-I (HLA-I) molecules are responsible for presenting endogenous peptides mainly generated by proteolytic machinery on the surfaces of cytotoxic CD8+ T lymphocytes[1–3]. The peptide binds in a specific groove of HLA-I molecules called as peptide binding groove, which is lined by the conserved sequence of polymorphic residues to stabilize the peptide[4,5]. The peptide is further held in pockets (A - F) present along the groove which determines the specificity of the HLA-I molecules[6]. In particular, peptide binds to pocket-B and pocket-F with the help of anchor residues P2 and PΩ, respectively[6–8]. The ends of the peptide binding groove are closed which is suitable to optimally accommodate peptide repertoire of 8–10 amino acids in length[4,9]. In recent years, some longer peptides (>11-mers) restricted by HLA-I have emerged, where these longer peptides adopt a bulged conformation and are known to comprise 10% of the HLA-I restricted peptide repertoire[10–16]. On the other hand, there are very few instances where shorter peptides (<8-mer) are shown to occupy the peptide binding groove[17–19]. Furthermore, two-thirds of the peptides generated by proteasome are less than 8-mer in length and only less than 15% of the proteasome breakdown products fall within an optimal range of 8–10 mer[20,21]. We have previously determined the occupancy of peptide binding groove for several antigens using synthetic peptides with serial truncation of peptides from both the N- and C-terminal ends[22]. In addition, there are reports demonstrating that non-contiguous peptide presentation by both MHC class I[23] and MHC class II[24] bring about an added intricacy in the production of T cell epitopes.

In the present study, Epstein-Barr Virus (EBV) was chosen as a viral model to analyse the immunogenicity of peptide fragments of <8-mer, as EBV is widely circulated in the world population with over 90% of adults being infected[25,26] and is involved in the manifestation of several types of malignancies[25,27]. We employed a well characterized HLA-A*11:01-EBV restricted peptide ATIGTAMYK derived from BRLF1 protein product Rta (134–142)[28] to investigate the stability of HLA-A*11:01 molecule in complex with shorter peptides (<8-mer). Rta serves as a transcriptional activator in both B lymphocytes and epithelial cells and is further involved in a switch from latent to the lytic phase[29]. The short peptide fragments derived from the N- and C-terminal regions

[1]Singapore Immunology Network, 8A Biomedical Grove, #03-06 Immunos, Singapore, 138648, Singapore. [2]Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University of Singapore, 5 Science Drive 2, Singapore, 119260, Singapore. *email: ren_ee_chee@immunol.a-star.edu.sg

of ATIGTAMYK (Rta134–142) were synthesized. These shorter peptides, which together make up the complete sequence of the parental peptide, were used alone or in combination to elicit a recall immune response by cytotoxic T lymphocytes. The result shows that short truncated peptides when co-occupying the binding groove complement each other in stabilizing the HLA-I complexes and can evoke antigen specific CD8+ T cell response. On the contrary, short truncated peptides alone forms unstable HLA-I complexes and evoke insignificant CD8+ T cell response. Furthermore, the crystal structure of HLA-A*11:01 in complex with two short truncated peptides at 1.66 Å demonstrates that they can bind in the same orientation similar to the parent peptide.

Materials and Methods

Protein expression and purification.   Soluble HLA-A*11:01 (residues 1–274) and full length β2-microglobulin (residues 1–99) with pET30a vectors were expressed in *Escherichia coli* as inclusion bodies using the pET prokaryotic expression system (Novagen Inc.). They were refolded with ATIGTAMYK (9-mer) peptide and were also refolded separately together with two short peptides namely, ATIGT (5-mer) and AMYK (4-mer), and were purified as previously described[30]. Briefly, the refolded peptide-HLA-A*11:01 (pHLA) complexes were dialyzed overnight at 4 °C in 10 mM Tris-HCl (pH 8.0) buffer. The dialyzed pHLA complexes were purified by anion exchange chromatography using HiPrep DEAE 16/10 column (GE Healthcare), and the pHLA complexes were eluted with buffer containing 10 mM Tris-HCl, pH 8.0 and 1 M NaCl. The pHLA complexes were further purified by size exclusion chromatography using HiLoad 16/60 Superdex 75 preparatory-grade column (GE Healthcare) and the pHLA complexes were eluted using buffer 10 mM Tris-HCl, pH 8.0. Finally, the peak corresponding to pHLA complexes were combined and concentrated to 10 mg/ml using Vivaspin centrifugal concentrators (Sartorius).

HLA-typing, isolation and *in-vitro* culture of human PBMCs.   Blood cones from healthy donors were obtained with ethical approval in accordance with Singapore Health Sciences Authority HSA-IRB-201306-5. Donors were recruited with written and informed consent under a Singhealth CIRB Ref: 2017/2512. All methods were performed in accordance with the relevant guidelines and regulations. PBMCs were isolated with Ficoll-Paque PLUS (GE Healthcare) according to manufacturer's instructions. HLA of donors were identified using sequence based typing as previously described[31]. For 14-day cultures, PBMCs (1×10^6 cells/ml) were incubated with 10 µM peptides (GenScript) and maintained in RPMI 1640 supplemented with 5% human AB serum (Sigma), 100 IU/ml penicillin-streptomycin (Gibco), 2 mM L-glutamine (Sigma), 10 mM HEPES (Sigma) and supplemented with 25 U/ml rIL2 (R&D Systems). Culture medium with rIL2 was replenished every 2–3 days from the fifth day onwards.

Preparation of peptide-HLA tetramers and cell staining.   The peptide-HLA tetramers were prepared as previously described[32]. Briefly, ATIGTAMYK (EBV Rta134–142) peptide was refolded with HLA-A*11:01 heavy chain and β2-microglobulin in refolding buffer for 72 h. Refolded complex was subsequently dialyzed against 10 mM Tris (pH 8.0) at 4 °C overnight. Following purification via anion exchange chromatography using HiPrep DEAE 16/10 column (GE Healthcare) equilibrated with 10 mM Tris-HCl (pH 8.0) and gel filtration with a HiLoad 16/60 Superdex 75 preparatory-grade GF column (GE Healthcare), the pHLA monomeric complexes were biotinylated by recombinant BirA enzymes. Assembly of tetrameric pHLA complexes was carried out by the stepwise addition of streptavidin-phycoerythrin (PE) (Invitrogen) or streptavidin-allophycocyanin (APC) (BioLegend) at a molar ratio of 4:1. Cells from the 14-day cultures were first harvested and then washed with PBS. Subsequently, they were stained with 12 µg/ml PE-conjugated pHLA tetramer for 20 min, and BV421-conjugated anti-human CD8 (BD Biosciences) for 15 min. Cells were washed again with PBS before analysis with LSR II flow cytometer (BD Biosciences). Data analyses were performed using FlowJo (Tree Star Incorporated).

Thermal stability assay.   To evaluate the difference in stability between HLA-A*11:01 bound to 9-mer and dual peptides (5-mer + 4-mer), a thermal shift assay was performed as described previously[33,34]. The thermal stability assay was performed in a real time PCR system LightCycler II 480 (Roche) and the fluorescent dye SYPRO Orange (Molecular Probes) was used to monitor the unfolding of pHLA complexes. Each HLA complexes or control was set in duplicates and the analysis was repeated independently on 2 different batches for the pHLA complexes at a concentration of 20 µM in Tris-HCl buffer (10 mM, pH 8.0), and temperature was increased from 20 °C to 95 °C with continuous ramp rate of 0.04 °C/s with 15 acquisitions/ °C. The fluorescence intensity was measured with excitation at 490 nm and emission at 575 nm. The raw data were normalized and plotted as negative first derivatives of fluorescence change against temperature, whereby, the melting temperature (Tm) is defined as the minima in the negative peak (-dRFU/dT).

Crystallization.   The crystal screening for purified pHLA complexes was initiated using commercial kits (Hampton Research). The pHLA complexes and crystallization solution drops were dispensed in equal volumes (1:1) by employing Phoenix protein crystallization robot (Art Robbins Instruments) in 96 well plates (Art Robbins Instruments) and plates were incubated at 16 °C. After the initial hits were obtained, the crystal drops for pHLA complexes were set using hanging drop vapour diffusion method in 24 well plates (Hampton Research) and plates were incubated at 16 °C. The crystals for HLA-A*11:01-ATIGTAMYK (9-mer) complex were formed in condition containing 0.2 M lithium sulfate, 0.1 M Tris pH 8.5, 30% PEG 4000. For the HLA-A*11:01-ATIGT (5-mer) + AMYK (4-mer) complex, crystals were formed in condition containing 0.2 M sodium malonate pH 5.0, 20% PEG 3350. The pHLA complex crystals were harvested by soaking in a reservoir solution supplemented with 20% glycerol as a cryo-protectant, and stored in liquid nitrogen before use.

Data collection, processing and structure determination.   X-ray diffraction data were collected at the Taiwan Photon Source at National Synchrotron Radiation Research Centre at TPS-05A beamline using the

| Parameter	HLA-A*11:01-ATIGTAMYK[a] (6JOZ)	HLA-A*11:01-ATIGT + AMYK[a] (6JP3)
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, b, c (Å)	51.40, 72.93, 125.70	65.50, 90.53, 91.24
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution range	30.00–1.35 (1.40–1.35)[b]	50.00–1.66 (1.72–1.66)
Rmeas (%)	8.0 (56.1)	7.0 (85.8)
I/σI	26.2 (3.2)	31.3 (2.4)
CC1/2 (%)	91.8 (72.9)	95.1 (82.4)
Completeness	98.4 (96.5)	99.0 (98.1)
Redundancy	4.4 (2.5)	7.2 (6.7)
Refinement		
No. of reflections	96217	60966
Rwork/Rfree	16.0/18.9	16.6/19.4
No. of atoms		
Protein	3100	3146
Peptide	66	66
Glycerol	42	24
Water	518	447
B-factors (Å^2)		
Protein	17.12	16.74
Peptide	13.36	21.37
Glycerol	24.8	33.50
Water	28.92	30.44
R. m. s. deviations		
Bond lengths (Å)	0.030	0.026
Bond angles (°)	2.492	2.296
Ramachandran plot		
Most favoured (%)	98.11	97.82
Outliers (%)	0	0

Table 1. Data collection and refinement statistics. [a]Diffraction data from one crystal was merged into a complete dataset. [b]Numbers in parentheses correspond to the last resolution shell. r. m. s. deviations: root-mean-square-deviation.

CCD detector MX300-HS. Diffraction data were indexed, integrated and scaled using HKL2000 suite of programs[35]. Both structures were solved by molecular replacement using Phaser[36] and the previously determined HLA-A*11:01 molecule was used as a search model (Pdb code: 1X7Q)[37] with peptide and water molecules removed. The model building was performed manually with program COOT[38] and the refinement was carried out using Refamc5[39]. The quality of the final model was validated using PROCHECK[40] and figures were prepared with PyMol[41]. The structure refinement statistics is reported in Table 1. The atomic coordinates and structure factors have been deposited in the Protein Data Bank, and the accession code for HLA-A*11:01-ATIGTAMYK is 6JOZ and for HLA-A*11:01-ATIGT + AMYK is 6JP3.

Accession codes. The atomic coordinates and structure factors for the reported crystal structures have been deposited in the Protein Data Bank under accession codes 6JOZ (HLA-A*11:01-ATIGTAMYK) and 6JP3 (HLA-A*11:01-ATIGT + AMYK).

Results
Functional validity of N- and C-terminal truncated peptides. Previously, we have shown using short truncated peptides (SSCSSC, 6-mer and SSCPLSK, 7-mer) derived from EBV latent membrane protein-2 epitope, SSCSSCPLSK (LMP2, 340–349) restricted by HLA-A*11:01 to elicit a strong cytotoxic T lymphocytes when presented alone or in combination[22]. Notwithstanding the fact that EBV has the ability to persist and prevail in latent form, however, it will be interesting to investigate how the short peptides originating from the lytic cycle of EBV will behave. Therefore, to evaluate the functionality of the shorter truncated peptides (ATIGT-5-mer and AMYK-4-mer) deduced from the lytic phase of EBV-Rta epitope (ATIGTAMYK, 134–142), PBMCs from two HLA-A*11:01 donors were pulsed with these peptides singly or in combination, and were cultured for 14-days. As shown in Fig. 1a,b, tetramer staining analyses revealed that the truncated peptides were able to stimulate a higher percentage of antigen-specific CD8$^+$ T cells in PBMCs when sensitized in combination by two truncated peptides (ATIGT + AMYK, 12.0% for donor 1 and 4.57% for donor 2). The parental 9-mer peptide (ATIGTAMYK) was able to stimulate 29.0% and 1.62% for donor 1 and donor 2 respectively. In comparison, when PBMCs were
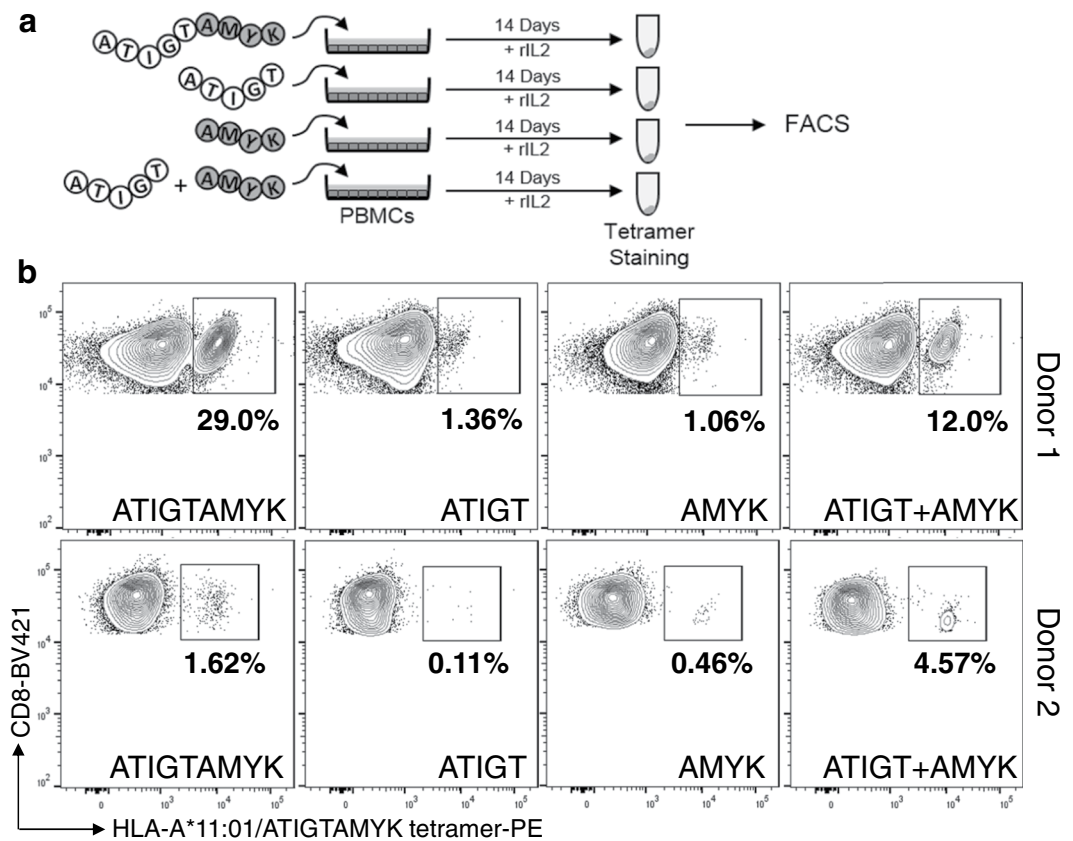
Figure 1. *In-vitro* priming and recall stimulation of human PBMCs using full length and truncated peptides. (**a**) PBMCs were stained with HLA-A*11:01 tetramers formed by using the peptides ATIGTAMYK, ATIGT or AMYK separately, or in combination (ATIGT + AMYK). The plots (**b**) of two different donors represent the frequency of specific CD8+ T cell responses for each tetramer and the values are highlighted in each plot.

treated separately with shorter truncated peptides it yielded a lower percentage of antigen-specific CD8+ T cells (ATIGT, 1.36% and AMYK, 1.06% for donor 1 and ATIGT, 0.11% and AMYK, 0.46% for donor 2). This provides strong evidence of the presence of specific CD8+ T cells from healthy human donor that efficiently recognize the combination of truncated peptides loaded on HLA-A*11:01.

Crystal structure of HLA-A*11:01-ATIGTAMYK and HLA-A*11:01-ATIGT + AMYK. To elucidate the structural basis for the two short peptides occupying the same peptide binding groove, the crystal structure of HLA-A*11:01 in complex with two short peptides, namely ATIGT (5-mer) and AMYK (4-mer) was determined at 1.66 Å, and was compared with the structure of HLA-A*11:01 in complex with the 9-mer full-length ATIGTAMYK peptide resolved at 1.35 Å. The overall structure of both pHLA complexes is similar to conventional MHC class I molecules (Fig. 2a)[3,42]. Clear electron densities were observed for the two short peptides ATIGT and AMYK, and also for the full length peptide, ATIGTAMYK within the peptide binding groove of HLA-A*11:01 (Fig. 2b,c).

Conformation of peptides in the peptide binding groove. To understand the structural differences between two pHLA complexes, the overall Cα atoms of HLA-A*11:01 in complex with dual peptide (ATIGT + AMYK) was superimposed with HLA-A*11:01-ATIGTAMYK complex and showed an overall r. m. s. d. of 0.38 Å. These small perturbations in two structures can be attributed to conformational flexibility upon peptide binding. However, overlaying only the Cα-atoms of 5-mer peptide (ATIGT) and 4-mer peptide (AMYK) with 9-mer peptide (ATIGTAMYK) showed an r. m. s. d. of 2.05 Å and 0.78 Å, respectively. The 9-mer peptide is seated in a more extended conformation and is held in peptide binding groove in a canonical fashion by primary anchor residues, namely P2-Thr (P2 anchor residue) and P9-Lys (PΩ anchor residue) (Fig. 3a,b). In contrast, the two short peptides ATIGT and AMYK adopted a bulged conformation, whereby, the N-terminal peptide ATIGT contributing P2 anchor residue (P2-Thr) and C-terminal peptide AMYK contributing PΩ anchor residue (P9-Lys) (Fig. 3a,b). The crystal structure of HLA molecules with two short peptides, HLA-A*11:01-ATIGT + AMYK is also compared with HLA molecules with longer peptides (>11-mer), HLA-B*35:01-LPAVVGLSPGEQEY (PDB code: 1XH3), HLA-A*02:01-ALQDAGDSSRKEYFI (PDB code: 4U6X), and HLA-A*02:01-FLNKDLEVDGHFVTM (PDB code: 4U6Y)[12,43]. Several structural studies had previously demonstrated that long peptides (>11-mer) binding to HLA class I alleles displayed high flexibility in
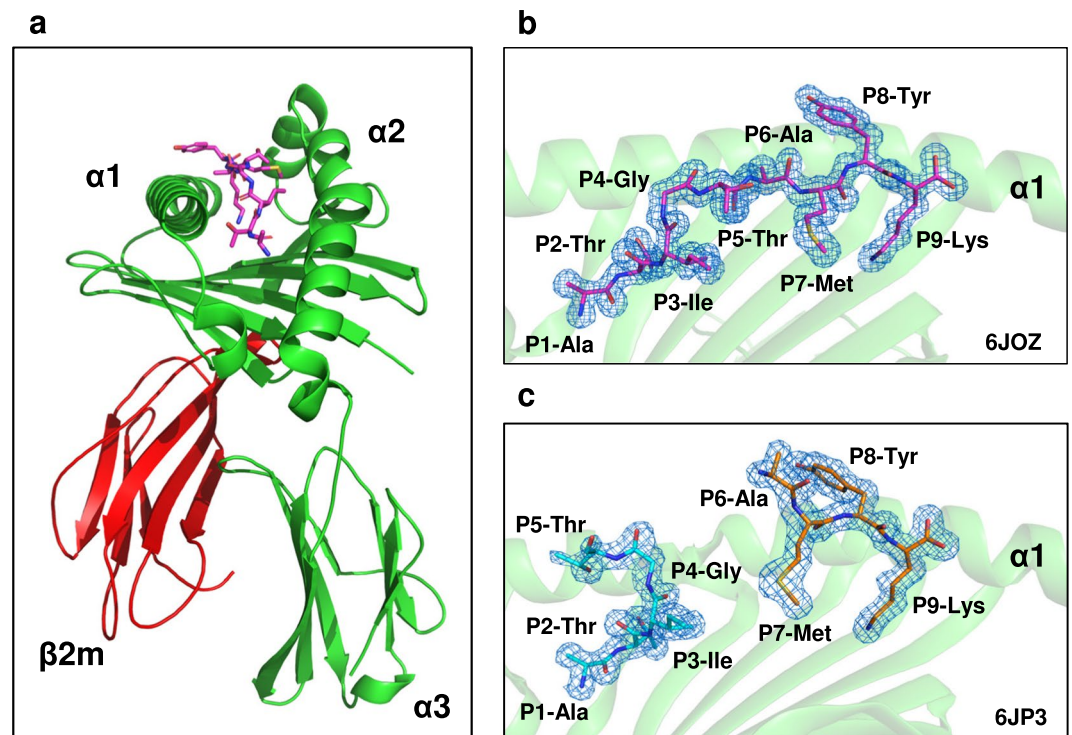
Figure 2. Solved structures and electron density maps of ATIGTAMYK and ATIGT + AMYK peptides bound to HLA-A*11:01. **(a)** The overall structure of HLA-A*11:01 in complex with ATIGTAMYK showing heavy chains (green), β2-microglobulin (red) and peptide (magenta). The panel (b,c) shows electron density map for ATIGTAMYK peptide (magenta) and two short truncated peptides namely ATIGT (cyan) and AMYK (orange) in complex with HLA-A*11:01, respectively. The 2Fo - Fc map is contoured at 1 σ in blue color. The HLA-A*11:01 is represented as a green schematic and for clarity only helix-1 (α1) is shown and the peptides are represented in stick.

the super bulged central region. These results suggest that the free ends of two short peptides acquire a bulged conformation that is directed towards the T cell receptor (Supplementary Fig. 1).

Furthermore, the solvent accessible surface area for different peptide residues was calculated with PDBePISA server using a 1.4 Å radius probe (Fig. 4). For HLA-A*11:01-ATIGTAMYK complex, residues P1-Ala (solvent exposure: 0.4 Å^2), P2-Thr (0.0 Å^2), P3-Ile (9.3 Å^2), P7-Met (0.6 Å^2) and P9-Lys (8.1 Å^2) show the lowest degree of solvent exposure and are therefore, buried deep in the peptide binding groove. Whereas, residues P4-Gly (49.3 Å^2), P5-Thr (33.4 Å^2) and P8-Tyr (59.3 Å^2) display the highest degree of solvent exposure and are oriented toward the exterior of the binding groove. However, P6-Ala (20.0 Å^2) exhibit an intermediate level of solvent exposure and is considerably buried in the peptide binding groove.

Similarly, for HLA-A*11:01-ATIGT + AMYK structure, peptide residues P1-Ala (0.4 Å^2), P2-Thr (0.0 Å^2), P7-Met (4.3 Å^2) and P9-Lys (2.9 Å^2) are minimally exposed to the solvent and resides within the interior of the peptide binding groove. While peptide residues P3-Ile (22.5 Å^2) and P4-Gly (57.9 Å^2) display a moderate level of solvent exposure and are markedly buried within the peptide binding groove. In comparison to 9-mer structure, the peptide residues P5-Thr (89.5 Å^2), P6-Ala (96.5 Å^2) and P8-Tyr (97.4 Å^2) of dual peptide structure show maximum amount of solvent exposure and their side chains are pointed outward of the peptide binding groove. Taken together, two short peptides can occupy the same peptide binding groove such that they adopt a bulged conformation similar to HLA molecules with longer peptides (>11-mer)[13].

The orientation of nonamer and N- and C-terminal truncated peptides in HLA-A*11:01 groove.
The overall orientation of 5-mer peptide is very similar to 9-mer peptide in the peptide binding groove. Nonetheless, P5-Thr of N-terminal peptide ATIG<u>T</u> has flipped outwards by 6.8 Å (Fig. 3c) in comparison to P5-Thr of ATIG<u>T</u>AYMK. Instead of forming a direct hydrogen bond of Gln55 on α2 helix, P5-Thr of N-terminal peptide ATIG<u>T</u> flipped over and is held by forming a direct hydrogen bond with Asn66 on α1 helix. Similarly, overall orientation of the C-terminal peptide A<u>M</u>YK is very similar to 9-mer peptide except that P6-Ala of A<u>M</u>YK has moved outward of the peptide binding groove by 6.5 Å (Fig. 3d) when compared with corresponding P6-Ala of ATIGT<u>A</u>MYK, pointing its side chain outward of the peptide binding groove. Likewise, P7-Met of A<u>M</u>YK has moved to the exterior by 2.5 Å in correlation to the equivalent P7-Met of ATIGTA<u>M</u>YK. While same with P7-Met of ATIGTA<u>M</u>YK, P7-Met of A<u>M</u>YK is stably held by a direct hydrogen bond forming with Gln156 on α2 helix. Furthermore, the head of P8-Tyr of AM<u>Y</u>K peptide is slanted at an angle by 3.8° in contrast to the corresponding P8-Tyr of 9-mer peptide, where the head of P8-Tyr in 9-mer is parallel to the peptide binding groove helices (Fig. 3d). As a result, unlike the 9-mer peptide, which seated in a more extended conformation and held in
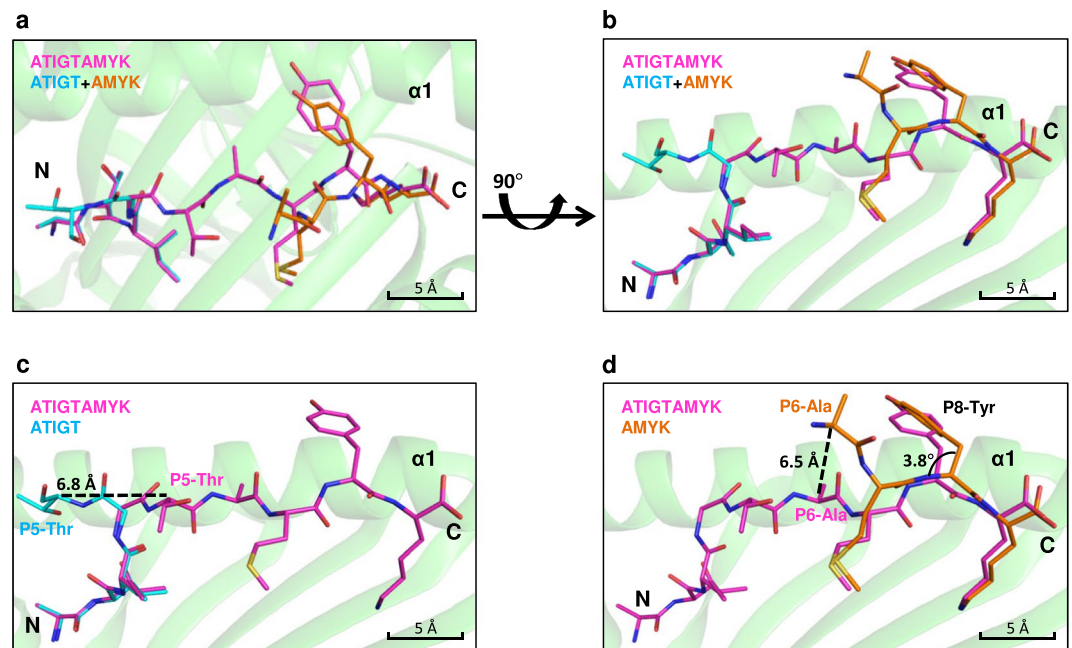
Figure 3. Structural comparison of peptides in the binding groove of HLA-A*11:01. Superposition of peptides showed that short truncated peptides (ATIGT + AMYK) adopted bulged conformation as compared to an extended conformation of parental peptide, ATIGTAMYK. The conformation of the peptides is viewed either from the top (**a**) or via the side (**b**) such that α2 helix is omitted for clarity. The residues, P5-Thr of N-terminal short peptide ATIGT and P6-Ala of C-terminal peptide A̲MYK have flipped outwards by 6.8 Å (**c**) and 6.5 Å (**d**), respectively. Furthermore, the ring of P8-Tyr of AM̲YK (C-terminal, 4-mer) has slanted outward by 3.8° in comparison to P8-Tyr of 9-mer peptide (**d**). Peptides are shown as sticks, using the same color scheme. The distances between equivalent residues are determined from the respective peptide Cα atoms.
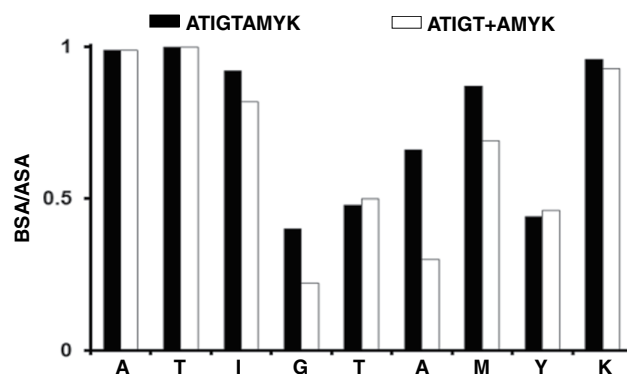
Figure 4. Ratios of the buried surface area (BSA) to the accessible surface area (ASA) of HLA-A*11:01 when bound to ATIGTAMYK (solid bars) or ATIGT + AMYK (open bars) respectively.

peptide binding groove in a canonical fashion, the free ends of two short peptides acquires a bulged conformation that is directed towards the T cell receptor.

Stability of the HLA-A*11:01-peptide complexes. To further examine the stability of HLA-A*11:01 in complex with two short peptides (ATIGT + AMYK), a thermal stability assay was performed and compared these values to HLA-A*11:01 bound to a canonical 9-mer peptide (ATIGTAMYK). The pHLA complex unfolding was monitored by using fluorescent dye Sypro orange. The data for each pHLA complex was recorded with two different protein batches, resulting in an average melting temperature (Tm) of 33 °C for HLA-A*11:01-ATIGT + AMYK and 70 °C for HLA-A*11:01-ATIGTAMYK (Fig. 5). The higher Tm-values for 9-mer peptide clearly demonstrate the preference for canonical mode of peptide binding to HLA molecules. However, it is possible for two short peptides to bind to the same peptide binding groove of HLA molecule in a non-canonical manner, but with lesser stability.
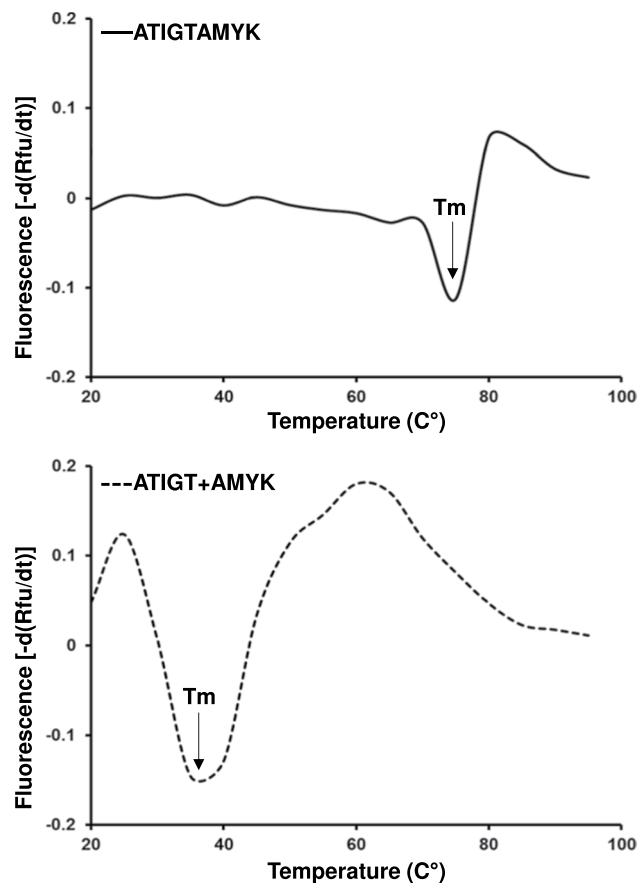
Figure 5. Thermal stability of the HLA-A*11:01-ATIGTAMYK (solid line) and HLA-A*11:01-ATIGT + AMYK (dashed line) was 70 °C and 33 °C, respectively.
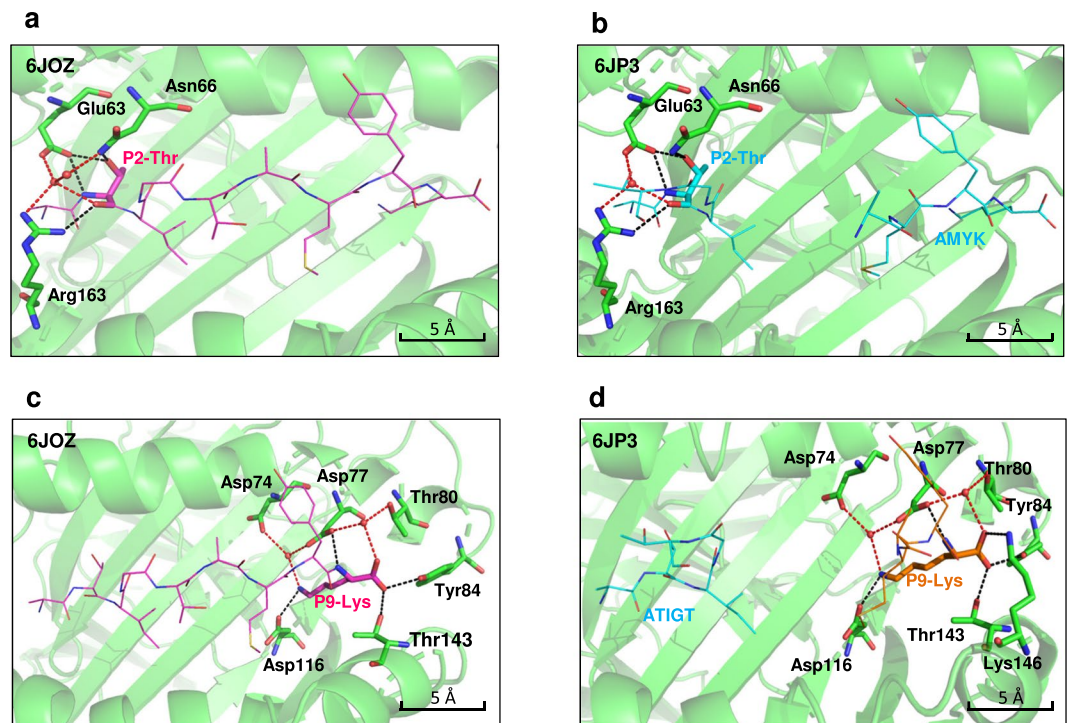
Peptide-binding interactions at P2 and PΩ anchor residues.

The various peptides in the study, namely, ATIGTAMYK (9-mer), ATIGT (5-mer) and AMYK (4-mer) are held in the peptide binding groove by forming stable interactions with HLA-A*11:01 molecule essentially by primary anchor residues P2-Thr and P9-Lys, either directly or mediated via water molecules. Briefly, the pocket B of peptide binding groove of HLA-A*11:01 where primary anchor P2 is seated is lined with polymorphic residues which are suitable to hold small aliphatic amino acids like valine, leucine, serine as well as threonine[44]. The side chain of P2-Thr of 9-mer forms direct hydrogen bond with Glu63 and Asn66, as well as through its main chain amino group it is directly hydrogen bonded to Glu63. Additionally, the main chain carbonyl group of P2-Thr is stably held by forming a direct hydrogen bond with Arg163. Furthermore, P2-Thr interacts with Glu63, Asn66 and Arg163 via framework of two water molecules (Fig. 6a). On the other hand, P2 anchor residue threonine of N-terminal peptide ATIGT in dual peptide structure is stabilised by similar bonding pattern as P2-Thr of 9-mer structure, except that it does not make any water mediated interaction with Asn66 (Fig. 6b).

The F-pocket of peptide binding groove of HLA-A*11:01 is characterized by negatively charged backdrop made up of Asp74, Asp77 and Asp116 residues, and which is suitable to accommodate positively charged residues, lysine or arginine. This is evident from a salt-bridge contact between the side chain Nζ-ammonium group of P9-Lys and Asp116. The side chain of P9-Lys also bonds with Asp74 and Asp77 via water molecules. Additionally, main chain amino group of P9-Lys makes a direct hydrogen bond with Asp77 and carboxylate of P9-Lys forms direct hydrogen bond with Tyr84 and Thr143, and also interacts with Asp77 and Thr80 via water molecules (Fig. 6c). In comparison with the 9-mer structure, the C-terminal peptide AMYK in a dual peptide structure is held stably in a similar manner in the F pocket by corresponding PΩ anchor residue P9-Lys except that carboxylate makes an additional direct hydrogen bond with Lys146 (Fig. 6d).

Discussion

Classically, HLA-I molecules present peptides with 8–10 amino acids in length with their N and C termini bound in the A and F-pockets, respectively[3,45]. The cell's proteasome machinery is responsible for the generation of the bulk of peptide source for HLA-I molecules[46] but only 15% of proteasome breakdown product falls within an optimal range of 8–10 residues[20]. Another 10% of the proteasomal products are longer than 11 amino acids, and in the recent years the evidence of longer peptides (>11-mer) occupying the peptide binding groove is emerging[13]. Moreover, the majority two-thirds of all peptides produced by proteasome are shorter than 8 residues[20,21] and fate of these short peptides is not very well documented in terms of their contribution to HLA-I molecule

Figure 6. Interactions stabilizing the HLA-A*11:01 peptide binding groove at P2 and PΩ anchor residues. The main chain of HLA-A*11:01 is depicted schematically as cartoon (light green) with selected HLA-I residues shown as sticks. Panel (a) and (b) shows interactions between P2 anchor residues (P2-Thr) of ATIGTAMYK and N-terminal truncated peptide ATIGT, respectively with HLA-A*11:01. Similarly, panel (c) and (d) details interactions between PΩ anchor residue (P9-Lys) in both the structures, respectively. The hydrogen bonding in indicated as black colored dotted line, whereas, interactions mediated via water molecules are indicated as red colored dotted line. Peptides are displayed as sticks and lines following the same color scheme and only P2 or PΩ anchor residue is labelled. The water molecules are shown in red spheres.

peptidome. The present study has provided functional and structural evidence to display that HLA-I molecules can bind short peptides (<8-mer) and present them to specific cytotoxic T lymphocytes (CTLs).

In this study we chose a very well characterized T cell epitope, ATIGTAMYK (9-mer) originated from the EBV-Rta protein (134–142) to examine the peptides that are competent in evoking T cell responses[32]. Rta (immediate early transactivator) is the product of BRLF1, and is associated with the reactivation of EBV from latency and in impelling the EBV during lytic phase[29]. During the lytic phase, Rta evokes the cellular immune response in EBV carriers. The data presented here demonstrates that synthetic shorter peptides (ATIGT 5-mer; AMYK 4-mer) derived by truncation of parental peptide (ATIGTAMYK 9-mer) when presented alone stimulated a negligible percentage of T cell responses from PBMCs of healthy HLA-A*11:01 positive donors previously exposed to EBV. When these two shorter truncated peptides are presented in combination they elicit a higher CD8+ T cell recall response. Furthermore, the cellular immune response by canonical 9-mer parental peptide (ATIGTAMYK) is relatively lesser when compared with shorter truncated peptides presented in combination. Taken together, shorter peptides can complement each other when presented in tandem and may contribute in stabilizing the peptide-HLA complex. This notion is supported by a recent report suggesting that progressively truncated short peptides from immunogenic EBV antigen SSCSSCPLSK can still bind to HLA-I molecule and elicit CTLs despite having lower affinities[22]. Likewise, short peptides derived from L^d peptide ligands, tum$^-$ and p2Ca are also shown to bind MHC class-I molecule as complementary pairs and stimulate specific CTLs[18].

The crystal structure of HLA-A*11:01 in complex with two short complementary peptides (ATIGT 5-mer and AMYK 4-mer) indicates that short synthetic peptide pairs can bind in similar orientation as the original longer parental peptide (ATIGTAMYK). Interestingly, these short complementary peptide pair adopt a bulged conformation in the centre similar to longer peptides (>11-mer)[13] whereby, the residues in the centre are highly solvent exposed, and could serve as a potential contact point for T cell receptors (TCR). Accordingly, it will be of great interest to know how T-cells can recognize short complementary peptide pairs. The TCR could likely bind in a diagonal orientation as generally accepted[45,47] or it will predominantly focus on peptides[48] and accommodate the dual peptide fragments in a different way.

Our analysis of short synthetic peptides co-occupying the same HLA-I binding groove is of considerable interest in devising the strategy for the development of vaccines and immunotherapy. It demonstrates the flexibility of HLA-I groove to accommodate short peptides of different sizes and as short as 4-mer while maintaining the specificity of T cell recognition. Furthermore, there is a possibility of one of the peptide pair originating from self-antigens, and in combination with non-self-antigens it can give rise to neo-epitopes with far wider

implications on autoimmune disorders. In addition, these short peptide pairs may originate from the same protein or different proteins, or may be non-contiguous as in the case of spliced peptides[23] or hybrid peptides[24]. In summary, we demonstrated that HLA-A*11:01 can bind two short synthetic peptides derived from EBV epitope and this pHLA complex can evoke recall response of specific CTLs *in vitro*. This combine-and-complement pair of peptides can contribute towards the existing pool of HLA-I immunopeptidome giving new insights about peptide presentation, and potentially augmenting the possibilities to generate new vaccine design and immunotherapy.

References

1. Blum, J. S., Wearsch, P. A. & Cresswell, P. Pathways of Antigen Processing. *Annu Rev Immunol.* **31**, 443–473 (2013).
2. Garcia, K. C., Teyton, L. & Wilson, I. A. Structural basis of T cell recognition. *Annu. Rev. Immunol.* **17**, 369–397 (1999).
3. Bjorkman, P. J. *et al.* Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* **329**, 506–512 (1987).
4. Neefjes, J., Jongsma, M. L. M., Paul, P. & Bakke, O. Towards a systems understanding of MHC class i and MHC class II antigen presentation. *Nat. Rev. Immunol.* **11**, 823–836 (2011).
5. Rammensee, H. G., Falk, K. & Rötzschke, O. Peptides Naturally Presented by MHC Class I Molecules. *Annu. Rev. Immunol.* **11**, 213–244 (1993).
6. Matsui, M., Hioe, C. E. & Frelinger, J. A. Roles of the six peptide-binding pockets of the HLA-A2 molecule in allorecognition by human cytotoxic T-cell clones. *Proc. Natl. Acad. Sci. USA* **90**, 674–678 (1993).
7. Madden, D. R. The Three-Dimensional Structure of Peptide-MHC Complexes. *Annu. Rev. Immunol.* **13**, 587–622 (1995).
8. Matsumura, M., Fremont, D. H., Peterson, P. A. & Wilson, I. A. Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science.* **257**, 927–934 (1992).
9. Bouvier, M. & Wiley, D. C. Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. *Science.* **265**, 398–402 (1994).
10. Bade-Doding, C. *et al.* The impact of human leukocyte antigen (HLA) micropolymorphism on ligand specificity within the HLA-B*41 allotypic family. *Haematologica* **96**, 110–118 (2011).
11. Burrows, S. R., Rossjohn, J. & McCluskey, J. Have we cut ourselves too short in mapping CTL epitopes? *Trends Immunol.* **27**, 11–16 (2006).
12. Hassan, C. *et al.* Naturally processed non-canonical HLA-A∗02:01 presented peptides. *J. Biol. Chem.* **290**, 2593–2603 (2015).
13. Josephs, T. M., Grant, E. J. & Gras, S. Molecular challenges imposed by MHC-I restricted long epitopes on T cell immunity. *Biol. Chem.* **398**, 1027–1036 (2017).
14. Probst-Kepper, M. *et al.* An alternative open reading frame of the human macrophage colony-stimulating factor gene is independently translated and codes for an antigenic peptide of 14 amino acids recognized by tumor-infiltrating CD8 T lymphocytes. *J. Exp. Med.* **193**, 1189–1198 (2001).
15. Speir, J. A., Stevens, J., Joly, E., Butcher, G. W. & Wilson, I. A. Two different, highly exposed, bulged structures for an unusually long peptide bound to rat MHC class I RT1-Aa. *Immunity* **14**, 81–92 (2001).
16. Tynan, F. E. *et al.* High resolution structures of highly bulged viral epitopes bound to major histocompatibility complex class I: Implications for T-cell receptor engagement and T-cell immunodominance. *J. Biol. Chem.* **280**, 23900–23909 (2005).
17. Apostolopoulos, V. *et al.* Crystal Structure of a Non-canonical Low-affinity Peptide Complexed with MHC Class I: A New Approach For Vaccine Design. *J. Mol. Biol.* **318**, 1293–1305 (2002).
18. Gillanders, W. E., Hanson, H. L., Rubocki, R. J., Hansen, T. H. & Connolly, J. M. Class I-restricted cytotoxic T cell recognition of split peptide ligands. *Int. Immunol.* **9**, 81–89 (1997).
19. Glithero, A. *et al.* The crystal structure of H-2Db complexed with a partial peptide epitope suggests a major histocompatibility complex class I assembly intermediate. *J. Biol. Chem.* **281**, 12699–12704 (2006).
20. Kisselev, A. F., Akopian, T. N., Woo, K. M. & Goldberg, A. L. The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J. Biol. Chem.* **274**, 3363–3371 (1999).
21. Nussbaum, A. K. *et al.* Cleavage motifs of the yeast 20S proteasome β subunits deduced from digests of enolase. *Proc. Natl. Acad. Sci. USA* **95**, 12504–12509 (1998).
22. Xiao, Z., Ye, Z., Tadwal, V. S., Shen, M. & Ren, E. C. Dual non-contiguous peptide occupancy of HLA class I evoke antiviral human CD8 T cell response and form neo-epitopes with self-antigens. *Sci. Rep.* **7**, 5072 (2017).
23. Liepe, J. *et al.* A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science.* **354**, 354–358 (2016).
24. Delong, T. *et al.* Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion. *Science.* **351**, 711–714 (2016).
25. Ressing, M. E. *et al.* Immune evasion by Epstein-Barr virus. *Curr. Top. Microbiol. Immunol.* **391**, 355–381 (2015).
26. White, D. W., Suzanne Beard, R. & Barton, E. S. Immune modulation during latent herpesvirus infection. *Immunol. Rev.* **245**, 189–208 (2012).
27. Rezk, S. A. & Weiss, L. M. Epstein-Barr virus-associated lymphoproliferative disorders. *Hum. Pathol.* **38**, 1293–1304 (2007).
28. Pepperl, S., Benninger-Döring, G., Modrow, S., Wolf, H. & Jilg, W. Immediate-early transactivator Rta of Epstein-Barr virus (EBV) shows multiple epitopes recognized by EBV-specific cytotoxic T lymphocytes. *J. Virol.* **72**, 8644–8649 (1998).
29. Zalani, S., Holley-Guthrie, E. & Kenney, S. Epstein-Barr viral latency is disrupted by the immediate-early BRLF1 protein through a cell-specific mechanism. *Proc. Natl. Acad. Sci. USA* **93**, 9194–9199 (1996).
30. Garboczi, D. N., Hung, D. T. & Wiley, D. C. HLA-A2-peptide complexes: Refolding and crystallization of molecules expressed in Escherichia coli and complexed with single antigenic peptides. *Proc. Natl. Acad. Sci. USA* **89**, 3429–3433 (1992).
31. Dormoy, A. *et al.* Mono-allelic amplification of exons 2-4 using allele group-specific primers for sequence-based typing (SBT) of the HLA-A, -B and -C genes: Preparation and validation of ready-to-use pre-SBT mini-kits. *Tissue Antigens* **62**, 201–216 (2003).
32. Yu, H. X., Srinivasan, N., Ren, E. E. & Chan, S. H. A11 Tetramer-assisted characterization of Rta-specific CD8+ T-cell responses in healthy virus carriers. *Tissue Antigens* **65**, 539–543 (2005).
33. Illing, P. T. *et al.* Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. *Nature* **486**, 554–558 (2012).
34. Layton, C. J. & Hellinga, H. W. Quantitation of protein-protein interactions by thermal stability shift analysis. *Protein Sci.* **20**, 1439–1450 (2011).
35. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
36. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
37. Blicher, T., Kastrup, J. S., Buus, S. & Gajhede, M. High-resolution structure of HLA-A*1101 in complex with SARS nucleocapsid peptide. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **61**, 1031–1040 (2005).
38. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).

39. Murshudov, G. N. *et al*. REFMAC 5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 355–367 (2011).
40. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
41. Schrödinger, L. The PyMol Molecular Graphics System, Versión 1.8. *San Carlos*, *CA*, *USA*, http://www.pymol.org (2015).
42. Guo, H.-C. *et al*. Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. *Nature* **360**, 364–366 (1992).
43. Probst-Kepper, M. *et al*. Conformational Restraints and Flexibility of 14-Meric Peptides in Complex with HLA-B*3501. *J. Immunol.* **173**, 5610–5616 (2004).
44. Sidney, J. *et al*. Definition of an HLA-A3-like supermotif demonstrates the overlapping peptide-binding repertoires of common HLA molecules. *Hum. Immunol.* **45**, 79–93 (1996).
45. Garboczi, D. N. *et al*. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* **384**, 134–141 (1996).
46. Rock, K. L., York, I. A., Saric, T. & Goldberg, A. L. Protein degradation and the generation of MHC class I-presented peptides. *Adv. Immunol.* **80**, 1–70 (2002).
47. Garcia, K. C. *et al*. An alphabeta T cell receptor structure at 2.5 A and its orientation in the TCR-MHC complex. *Science* **274**, 209–219 (1996).
48. Tynan, F. E. *et al*. T cell receptor recognition of a 'super-bulged' major histocompatibility complex class I–bound peptide. *Nat. Immunol.* **6**, 1114–1122 (2005).

Acknowledgements

Author contributions

X.H. contributed to the protein crystal structure and data analysis. Z.Z. and Z.X. contributed to data generation and functional analyses. E.C.R., Z.X. and X.H. contributed to experimental design, and writing of manuscript. E.C.R. is the senior and corresponding author who led the project.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-019-53201-6.

Correspondence and requests for materials should be addressed to E.C.R.

Reprints and permissions information is available at www.nature.com/reprints.

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