HHS Public Access

Author manuscript

Nat Struct Mol Biol. Author manuscript; available in PMC 2017 August 27.

Published in final edited form as:

Nat Struct Mol Biol. 2017 April ; 24(4): 395–406. doi:10.1038/nsmb.3383.

Broad TCR Repertoire And Diverse Structural Solutions To Recognition Of An Immunodominant CD8 T Cell Epitope

InYoung Song[1,2], **Anna Gil**[1], **Rabinarayan Mishra**[1], **Dario Ghersi**[3], **Liisa K. Selin**[1,2,*,4], and **Lawrence J. Stern**[1,2,5,*,4]

[1]Department of Pathology, University of Massachusetts Medical School, Worcester MA 01605

[2]Graduate Program in Immunology and Microbiology, University of Massachusetts Medical School, Worcester MA 01605

[3]School of Interdisciplinary Informatics, University of Nebraska at Omaha, Omaha, NE 68182

[5]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester MA 01605

Abstract

A keystone of antiviral immunity is CD8 T-cell recognition of viral peptides bound to MHC-I proteins. The recognition mode of individual T cell receptors (TCRs) has been studied in some detail, but how TCR variation functions in providing a robust response to viral antigen is unclear. The influenza M1 epitope is an immunodominant target of CD8 T cells helping to control influenza in HLA-A2+ individuals. Here, we show that many distinct TCRs are used by CD8 T cells to recognize HLA-A2/M1, encoding different structural solutions to the problem of specifically recognizing a relatively featureless peptide antigen. The vast majority of responding TCRs target small clefts between peptide and MHC. These broad repertoires lead to plasticity in antigen recognition and protection against T cell clonal loss and viral escape.

*Address correspondence to liisa.selin@umassmed.edu or lawrence.stern@umassmed.edu.
[4]These authors contributed equally to this work

Accession Numbers

LS01-HLA-A2/M1 complex: 5ISZ and LS10-HLA-A2/M1 complex: 5JHD.

Author Contributions

I.Y.S. conceived project, designed experimental approach, performed single-cell sequencing experiments, characterized TCR transfectants, determined crystal structures, and wrote the manuscript. A.G. performed NGS analyses, performed single-cell sequencing experiments, and edited the manuscript. R.M. performed NGS analyses and edited the manuscript. D.G. performed NGS analyses and edited the manuscript. L.K.S. conceived project, designed experimental approach, supervised TCR sequencing analyses, and wrote the manuscript. L.J.S. conceived project, designed experimental approach, supervised cellular and molecular studies, and wrote the manuscript.

Competing Financial Interests Statement

The authors declare no competing financial interests.

Data Availability

Crystallographic structure factors and coordinates are available from the Research Collaboratory for Structural Bioinformatics: Protein Data Bank at http://www.rcsb.org. TCR sequence data are available from the ImmuneACCESS database at http://doi.org/10.21417/B7W88F. Source data for figures 1b–h, 3b, 5b, and 6d–e are available with the paper online.

The importance of T cell immunity to influenza A virus (IAV) is supported by studies in animal models and humans[1,2], and has received increasing attention because a CD8 T cell based vaccine against a conserved epitope potentially could provide broad protection despite viral antigenic shift and drift[3]. The antiviral CD8 T cell response is initiated by interaction between clonally distributed αβ T cell receptor (TCR) heterodimers and viral peptide loaded on MHC-I. TCR genes are assembled by recombination of TRAV (or TRBV) gene segments that encode variable complementarity-determining CDR1 and CDR2 regions, with TRAJ (or TRBD/TRBJ) gene segments that encode hypervariable CDR3 regions. The HLA-A2/M1 epitope, composed of $M1_{58–66}$ (M1), a nonameric peptide from the IAV matrix protein, presented by the common human MHC-I allelic variant HLA-A2*01:01, is a highly conserved immunodominant epitope[4–6] that is abundantly expressed in infected cells[7]. Previous studies of M1-specific CD8 T cell response have suggested that the TCRβ repertoire responding to HLA-A2/M1 is highly biased toward usage of the TRBV19 gene (up to 98%)[8–10], with a highly conserved CDR3β motif, $xR^{98}S^{99}x$[8,9,11]. TCRα bias is less dramatic but preferential usage of TRAV27 and TRAJ42 gene segments has been reported[8,9,12]. As for many viruses that infect hosts chronically or recurrently, IAV infection results in “public” TCRs with identical or near-identical patterns of V-region, J-region, and junctional sequences among HLA-A2-matched but otherwise genetically unrelated individuals.

A crystal structure of HLA-A2/M1 bound to one of these canonical public TCRs (JM22) showed that most of amino acid side chains of M1 were buried in the peptide binding cleft of HLA-A2[13,14]. This ‘featureless’ HLA-A2/M1 complex was recognized mainly by residues from CDR1β, CDR2β and Arg^{98} of the CDR3β $xR^{98}S^{99}x$ motif, explaining the biased selection of TRBV19 and the role of the conserved CDR3β motif, with few MHC or peptide contacts from TCRα side chains[14]. It has been suggested that featureless (or less featured) peptides are more prone to TCR bias than featured peptides, because of a dearth of available recognition modes[15–17]. Direct proof of this concept came from an elegant study[18] where the highly featured PA_{224} epitope from influenza acidic polymerase presented by $H\text{-}2D^b$ was mutated to a more featureless version, inducing a change from a diverse TCR repertoire to a more restricted one. Several studies have suggested that diverse TCR repertoires recognizing virulent virus are correlated with efficient control of viral infection[19–21] and reduction in viral escape[22]. Thus there is a concern about restricted TCR repertoires because of possible loss of protection by either clonal loss or viral escape mutation. In one study, SIV viral load was inversely correlated not with epitope-specific CD8 T cell frequency, recruitment to target organ, multifunctionality, or inability to recognize mutated virus, but rather with the number of public TCR clonotypes[23], implying that the size of the TCR repertoire may be a critical component to understand efficient viral control. Despite the increasing availability of high-throughput TCR sequencing strategies[24] the breadth of TCR responding to human viral infection has been studied only in a few cases at sequence[25,26] or structural levels[27–29] and no study has been reported that combines both aspects.

Here, we systematically examined the HLA-A2/M1-restricted CD8 T cell repertoire by performing comprehensive TCR repertoire analysis on 6 healthy donors using next-generation sequencing (NGS) to obtain unbiased TRBV and TRAV information, identifying tremendous diversity with many hundreds of unique clonotypes in each donor. We evaluated

TCRα and TCRβ chain pairing patterns directly ex vivo using single cell sequencing confirmed by functional analysis in T cells carrying recombinant TCR. We identified a previously unnoticed public TCR that uses TRAV38/J52 and TRBV19/J1-2 genes and sequence motifs in both CDR3α and CDR3β beyond the 'xRSx' motif. In addition, we identified many non-canonical M1-specific TCRs with lower frequency in the HLA-A2/M1-specific CD8 T cell population. X-ray crystal structures of two non-canonical TCRs revealed the structural basis for HLA-A2/M1-recognition without the 'xRSx' motif, and identified unique pockets between the peptide and MHC that appear to be required for recognition of this featureless epitope. Combined with previous work this study now provides the most comprehensive look to date at the breadth of TCR repertoire to a viral antigen, and the structural basis for understanding recognition by thousands of TCRs able to recognize this ubiquitous epitope.

Results

Diversity of CD8 T cell repertoire and dominant usage of TRAV38 in HLA-A2/M1-specific response

We analyzed the TCR repertoires of tetramer-sorted CD8 T cells from peripheral blood mononuclear cells (PBMCs) of 6 healthy adult donors. Since the average frequency of HLA-A2/M1-specific memory CD8 T cells in PBMC of healthy individuals is less than 0.2%[9,30,31], we expanded the antigen-specific population in vitro by M1-peptide stimulation (Fig. 1a, Supplementary Fig. 1a). We then used HLA-A2/M1 dextramer to sort antigen-specific cells before isolating cDNA for NGS analysis of the TRAV and TRBV repertoires. Previous approaches using 5'-RACE PCR[12], or using individual primers for each TRAV or TRBV gene and then subcloning[32], were not designed to exhaustively sample the repertoire, and on average identified 20–100 unique TCRα or TCRβ clonotypes specific to HLA-A2/M1 in any donor, usually examining either TRAV or TRBV repertoire but not both[11,12,32]. Using the ex vivo expansion/NGS strategy we gained a greater appreciation of the complete M1-specific TCR repertoire. We obtained an average of 516 unique TCRα clonotypes (range: 209–1,037) and 432 unique TCRβ clonotypes (range: 150–975) in each individual (Fig. 1b,c, Supplementary Fig. 1b,c), for a total of 2,939 unique TCRα and 2,544 unique TCRβ sequences. These results suggest that previous methods underestimated the diversity of the HLA-A2/M1-specific TCR repertoire (TCRα and TCRβ diversity indices shown in Supplementary Fig. 1j).

Interestingly, even with this much diversity the HLA-A2/M1-specific CD8 T cells from each of the donors predominantly utilized TRBV19, as observed in earlier studies[8–10]. TRBV19 frequencies ranged from 57.1% to 89.5% of all TCRβ sequences read depending on the donor (Fig. 1c). In addition to TRBV19 there was minor usage of other common TRBV genes that differed donor-to-donor with no systematic usage patterns. In contrast to the highly-restricted TRBV usage, many different TRAV genes were utilized in the M1-specific TCR response. Previous studies reported preferential usage of the TRAV27 gene[8,9,12]. We found TRAV27 used commonly in all 6 donors, accounting for up to 47.8% of total TRAV sequences in donor 185, but as low as 2.8% in donor D105 (Fig. 1b). Previous studies with fewer sequences showed higher biased usage of TRAV27 (49–75%)[8,9,12]. In this study

where we sampled ~8000-fold more sequences TRAV27 accounted for only average 16.8% of the HLA-A2/M1-specific TRAV repertoire. In addition to TRAV27, other TRAV genes were commonly detected, including TRAV12, TRAV13, TRAV25, TRAV29, and TRAV38 (Fig. 1b). The TRAV38 gene in particular was found in all six donors and accounted for the most abundant TRAV gene in some donors Although usage of this gene has been observed previously in the HLA-A2/M1 response[12], its public usage and dominance in many donors was not previously appreciated. Together TRAV27 and TRAV38 account for a range of 21 to 53 percent of the overall response depending on the individual (Fig. 1d). Because of the potential for relative TRAV (or TRBV) sequence bias introduced during reverse transcription and amplification steps, we repeated analyses of relative abundance using observed numbers of unique clonotypes instead of observed sequence frequencies (Supplementary Fig.1b–d). In general similar patterns were observed.

TRAV38 gene almost exclusively joins with TRAJ52 forming a novel 15-amino acid long CDR3α

We examined TCRα repertories containing the newly identified dominant TRAV38 gene to assess TRAJ pairing, CDR3α length, and CDR3 sequence composition. TRAV38 mainly rearranged with TRAJ52 in all six donors (Fig. 1e), with the paired frequency from 83.8 to 100%. The length distribution of the TRAV38-containing CDR3α was highly restricted, with 96% from the six donors encoded a 15-mer CDR3α (Fig.1f). In general the CDR3α length distribution for HLA-A2/M1 specific repertoires carrying any TRAV was much broader, with 10–15mer CDR3α all represented (Fig. 1g). However, in the 15-mer group TRAV38/TRAJ52 gene pairs were the most common, ranging from 78.3–98.2%, except in donor 215 (42.8%). Sequence analysis of the HLA-A2/M1-specific TRAV38/TRAJ52 repertoires from all six donors revealed a new CDR3α motif: $\underline{C^{92}A\Phi x_1x_2x_3AGGTSYGKLT}\mathbf{\underline{F}}^{108}$, where C^{92} is the second cysteine of the TRAV gene, F^{108} is the phenylalanine in the characteristic TRAJ gene motif ‘FGxG’, $\underline{\Phi}$ is an aromatic residue, and x_1, x_2, and x_3 represent variable intervening sequences (Fig. 1g). This sequence composition analysis showed that germline-encoded Phe^{94}/Tyr^{94} and Ala^{98} are under selection pressure that may prevent codons for these amino acids from being trimmed off during rearrangement.

These results define a new CDR3α motif in HLA-A2/M1-specific TCRs, consisting of an exclusive paired TRAV38/TRAJ52 together with an extremely restricted CDR3α length and composition. Since this motif constitutes most of HLA-A2/M1-specific TCRs with 15mer CDR3α these particular features may be important in the recognition of HLA-A2/M1.

Two dominant CDR3β motifs of M1-specific TCRs with 10- or 11-amino acid long CDR3β loops: ‘$xG^{98}xY^{100}$’ or ‘$xR^{98}S^{99}x$’

Two conserved TRBV19 CDR3β motifs, ‘$xR^{98}S^{99}x$’ and ‘$xG^{98}xY^{100}$ have been reported for HLA-A2/M1-specific TCRs with 74–84% and 11–13% frequencies respectively[9,11]. Using NGS we identified a strong length bias of CDR3β toward 11-mer at 68% (range: 43–88%) of total TCRβ and 10-mer at 12% (range: 3–23%) (Fig. 1h). Sequence analysis showed that the canonical ‘$xR^{98}S^{99}x$’ motif was restricted to 11-mer CDR3β, with less bias than previously reported[9,11], representing only 50% (range: 15–72%) of total M1-specific TCRs.

Surprisingly, a single motif (xG^{98}xY^{100}) comprised almost all (98±2%) of the TRBV19 10-mers (Fig. 1h). These 'xG^{98}xY^{100}' TCRs represented 11.5% (range: 2.3–25.8%) of the total M1-specific TCR. Our NGS analysis showed that TCRs with these common 'xR^{98}S^{99}x' and 'xG^{98}xY^{100} motifs dominated the M1-specific TRBV19 repertoire in all the donors (Supplementary Fig. 1i). These results suggest that studies focused on just TRBV19 with lower numbers of sequences might overestimate the role of 'xR^{98}S^{99}x' motif while underestimating the truly polyclonal nature of the IAV-MI-specific TCR repertoire and the role of other motifs such as 'xG^{98}xY^{100}.

TRBV19 with 10-mer xGxY CDR3β motif paired exclusively with novel TRAV38-TRAJ52 with 15mer motif

To obtain information about which TCRα paired with which TCRβ, we used a single-cell sequencing approach in which rearranged TCRα and TCRβ genes were amplified from individual HLA-A2/M1-dextramer+ CD8 T cells sorted directly ex vivo from PBMC (Supplementary Fig. 2a). TCR genes from individual cells were sequenced using a previously described[19] nested PCR strategy. Of the 82 productive TCRβ genes sequenced, 33 TCRβ paired with multiple TCRα, 15 paired with non-productive single TCRα genes, and 34 paired with productive single TCRα genes (Supplementary Table 1). TRAV and TRBV gene usage patterns (Supplementary Fig. 2b) generally were similar to those observed by NGS of dextramer+ T cells from the same donor (D085). Of the 34 single TCRα/TCRβ pairs, twenty-three had TRBV19 genes containing the 11mer 'xRSx' motif ('xR^{98}A^{99}x' in some cases) paired with 9 different TRAV genes (Group I, Supplementary Table 1). All of these TCR utilized TRAJ42 or TRAJ37 gene segments.

Surprisingly, all of the TCRβ chains with the CDR3 10mer 'xG^{98}xY^{100}' motif paired exclusively with the novel TCRα composed of the paired TRAV38/TRAJ52 and the novel 15mer CDR3α motif (Group II, Supplementary Table 1). The greater abundance of TRBV19 clones with the xR^{98}S^{99}x' motif might relate to their ability to successfully pair with many different TRAV as compared to the stringent pairing requirement for those with xG^{98}xY^{100}. The structural basis for this unusual paired selection will be described below.

Featureless surface presented by HLA-A2/M1

HLA-A2/M1 has been considered a relatively featureless ligand[13,33] because most of M1 side chains are buried within the peptide binding pocket of HLA-A2, as compared to other MHC-I/peptide complexes such as HLA-A2/RT or HLA-A2/tax where one or more peptide side chains are largely exposed (Supplementary Fig. 3a,b). The solvent accessible area of M1-peptide bound to HLA-A2 is 248 Å^2, the smallest among 107 structures of free peptide-HLA-A2 complexes deposited in PDB (Supplementary Fig. 3c), although the total peptide buried surface area (BSA) reflecting MHC-peptide interaction is slightly above the average (Supplementary Fig. 3d). It has been suggested that TCR specific for HLA-A2/M1 would necessarily be highly restricted because of the limited ways available to recognize a featureless peptide[13]. As we observed a broad response to HLA-A2/M1 including TCRs with different recognition motifs, we were interested in how the newly identified TCRs could specifically recognize the rather featureless M1 ligand.

Characterization of TCR proteins with sequences identified by single-cell sorting

To validate the TCRα/β chain-pairing information from single cell PCR, thirteen representative full-length TCRα/TCRβ chain genes were assembled by overlapping PCR and expressed in TCRα/β deficient Jurkat T cells expressing human CD8α (J76-CD8). We selected seven TCRαβ pairs from Group I containing the canonical 'xRSx' motif (LS02–LS06, LS08, LS13), three from Group II utilizing the newly identified dominant public TCR encoded by TRAV38/TRAJ52/TRBV19/TRBJ1-2 (LS10, LS11, LS12), two from Group III with unique 11mer TRBV19 CDR3β sequences with aromatic amino acids in place of R^{98} in the 'xR^{98}S^{99}x' (LS01, LS07), and one TCR from Group IV (LS09) representing the set of TCR lacking CDR3 homologies among different individuals (Supplementary Table 1). Each TCRα/β pair was expressed at the cell surface (Supplementary Fig. 4a), each could bind HLA-A2/M1 dextramer (Fig. 2a), and each could initiate T cell signaling responses as measured by CD69 upregulation after stimulation with HLA-$A2^+$ cells pulsed with M1 peptide but not with control peptides (Supplementary Fig. 4d). These data support the reliability of chain pairing information from single cell PCR and the functional competence of the resultant TCR proteins.

We chose for further characterization LS01 and LS10 as representatives of Groups II and III respectively and compared these to JM22, the prototypical canonical Group I public TCR recognizing HLA-A2/M1[9]. We evaluated the relative HLA-A2/M1 binding activity of these TCRs after stable expression in J76-CD8. Concentration-dependent tetramer binding was observed for JM22, LS10, and LS01 (Supplementary Fig. 4b), with similar half-maximal concentrations (Fig. 2b) but with different maximum binding levels, consistent with the respective TCR expression levels (Supplementary Fig. 4c). We compared the relative functional sensitivity of the TCRs in response to stimulation by peptide-pulsed antigen presenting cells (Supplementary Fig. 4d). EC_{50} values for CD69 upregulation were similar for JM22, LS10, and LS01 (Fig. 2c). Finally, we prepared soluble TCR and MHC proteins and evaluated binding directly using surface plasmon resonance (Fig. 2d). Apparent K_d values for LS01 and LS10 (32 and 30 μM, respectively) were in the range previously observed for agonist MHC-peptide[34,35] but somewhat weaker than reported previously for JM22 (5 μM)[14]. To validate the somewhat weaker activation response for JM22 despite its apparent higher affinity we evaluated also the LS06 TCR (LS06 is almost identical to JM22 but with Thr^{97}β in place of Ser^{97}β at a position that doesn't contact any MHC-peptide residues), which exhibited EC_{50} for CD69 upregulation essentially identical to JM22. Overall LS01, LS10, and JM22 had similar binding and activation characteristics despite using different recognition motifs.

Crystal structures of non-canonical TCRs LS10 and LS01

To investigate how LS01 and LS10 TCRs recognize HLA-A2/M1 without the canonical CDR3β 'xRSx' motif, and to discover why specific CDR3α sequences are required, we determined the X-ray crystal structures of these TCRs bound to HLA-A2/M1. Data collection and refinement statistics are shown in Table 1, with representative omit maps shown in Supplementary Fig. 5a. Except for a few disordered loops, the TCR, MHC, and peptide all were defined well by the available data with clear electron density and good geometry in the final models (See Methods). Overall the structures of LS01 and LS10 bound

to HLA-A2/M1 showed that they docked similarly to JM22 (Fig. 3a). All three TCRs used a conventional diagonal binding mode, with crossing angles[36] 70–85° and CDR3α and CDR3β loops centered over the peptide (Fig. 3c–d). The interaction between TCR and MHC/peptide buried similar amounts of surface area (Fig. 3b). Footprints of the TCRs on the MHC/peptide were broadly similar, with a greater contribution of TCRα for both LS01 and LS10, mostly due to CDR1α in LS01 (blue in Fig. 3d) and to CDR3α in LS10 (light green in Fig. 3d). The HLA-A2/M1 component of the LS01/HLA-A2/M1 complex overall is essentially identical to the structure of free unliganded HLA-A2/M1 (Fig. 3c), except for a few key residues such as Arg^{65MHC} and Gln^{155MHC} that are members of the 'restriction triad'[35]. Differences between the crystal structures and the roles of the various CDR loops and TCR, MHC, and peptide structural rearrangements are discussed in detail in sections that follow.

LS10: a new mode of HLA-A2/M1 recognition

LS10 uses a novel mechanism to recognize HLA-A2/M1, involving a conformational change in the M1 peptide induced by interactions with the long CDR3α loop encoded mostly by TRAJ52. In complex with LS10, M1 bound to HLA-A2 adopts a conformation different from those observed in free HLA-A2/M1 and in HLA-A2/M1 of the JM22-bound complex, with changes concentrated in the center of the peptide (Fig. 4a). The Phe^{5}-p side chain adopts a different rotamer and moves towards the MHC α2 helix, with the Cα atom and nearby main chain moving by ~1Å and the side chain phenyl ring moving by almost 5Å. These changes appear to be induced by interaction with $Ala^{98}α$ and $Tyr^{103}α$ from LS10 CDR3α loop (Fig. 4b). These residues pack against Phe^{5}-p side chain, but would clash severely if the Phe^{5}-p were to retain the original conformation as observed in the unliganded HLA-A2/M1 structure.

The motion of the Phe^{5}-p side chain fills a notch in the unliganded structure lined by Phe^{5}-p and Phe^{7}-p from M1 and Ala^{150MHC}, Val^{152MHC}, Gln^{155MHC} from HLA-A2 α2 helix (dotted circle in Fig. 4c). This notch is believed to play a key role in recognition of HLA-A2/M1 by canonical TCR, with the side chain of the conserved $Arg^{98}β$ from the 'xRSx' motif inserting into the notch, as observed in the JM22/HLA-A2/M1 complex (compare dotted circle in Fig. 4c and yellow surface in Fig. 4f). The motion of Phe^{5}-p into the notch opens up a shallow hydrophobic pocket (dotted circle in Fig. 4d). In the LS10 complex, this new pocket becomes occupied by $Ala^{98}α$ and $Gly^{99}α$ and covered by the side chain of $Tyr^{103}α$, all from CDR3α (Fig. 4e). In this region, CDR3α is closely apposed to the corresponding CDR3β loop from the other TCR subunit (Fig. 4g), and $Gly^{98}β$ from the CDR3β 'xG^{98}xY^{100}' motif lodges between the phenyl ring of the displaced Phe^{5}-p and the phenyl ring of $Y^{103}α$ from CDR3α (Fig. 4e), leaving no room for a side chain at position 98 of CDR3β. The tight packing of these four residues ($Ala^{98}α$, $Tyr^{103}α$, $Gly^{98}β$, Phe^{5}-p) helps explain strict pairing and sequence requirements of the 15-mer CDR3α 'CAΦxxx$\underline{A^{98}}$GGTS$\underline{Y^{103}}$GKLTF' motif and the 10-mer CDR3β 'x$\underline{G^{98}}$xY^{100}' motif. The side chain of $Tyr^{100}β$, the other component of the 'xG^{98}xY^{100}' CDR3β motif, packs against Gln^{155MHC} in the HLA-A2 α2 helix (Fig. 4g). Gln^{155MHC} has been referred to previously as a "gatekeeper"[37] and it regulates access to the notch by $Arg^{98}β$ from the 'xR^{98}Sx' motif in canonical TCR recognition of HLA-A2/M1[13].

We validated the importance of these interactions by mutagenesis (see Supplementary Notes). We investigated also the role of other residues in the long 15-mer CDR3a motif, including Gly^{99}α, Gly^{100}α, and Gly^{104}α which appear to be required for formation of two hairpin loops in CDR3α loop structure that allow Tyr^{103}α to pack against Ala^{98}α on the side of the CDR3α loop (Fig. 4h and Supplementary Notes), Tyr^{94}α encoded by the extreme 3' end of the TRAV38 gene, which nestles into a hydrophobic pocket formed by TRAV38 residues from the framework region and Thr^{107}α from TRAJ52 (Fig. 4i and Supplementary Notes), and features that stabilize the convoluted structure adopted by the long 15-residue TRAJ52 gene segment in LS10 TCR as compared to other TCR utilizing this gene segment (Supplementary Fig. 6f and Supplementary Notes).

Overall, LS10 recognizes the relatively featureless M1 peptide by inducing a peptide conformational change that opens up a small pocket able to be accessed by residues on the sides of a long CDR3α loop, with the interaction dependent on TRAV-TRAJ pairing and selection of a particular 10-mer CDR3β xGxY motif.

LS01: another solution to recognizing HLA-A2/M1

Among TRBV19-containing TCRs that recognize HLA-A2/M1, a substantial fraction with 11-residue long CDR3β did not contain the predominant 'xRSx' motif, and instead often encoded hydrophobic residues such as Phe, Tyr, or Leu in place of Arg^{98}β. Frequencies of this type of TCRs in 6 donors investigated ranged from 1% to 4% of total M1-specific TCRs (Fig. 5a). Previous work on the canonical JM22 revealed that Arg^{98}β of 'xR^{98}S^{99}x' was critical for M1 interaction, and mutations of Arg^{98} to histidine and alanine were not tolerated[14]. We tested whether other amino acids could replace Arg^{98}β of JM22, substituting it with residues that retain charge, hydrogen-bonding, or non-polar characteristics of the arginine side chain. Mutated JM22 TCRs were expressed transiently in J76-CD8 cells and assessed for HLA-A2/M1-tetramer binding (Fig. 5b). Arg^{98}β in JM22 was highly resistant to mutation, with Lys^{98}, Gln^{98}, Phe^{98}, and Tyr^{98} substitutions all leading to complete loss of tetramer binding. These results raised the question of how Group III TCR like LS01, which have an aromatic group at position 98, are able to recognize HLA-A2/M1.

In the crystal structure of LS01 bound to HLA-A2/M1, the CDR3β loop lies above and between the MHC α2 helix and the Phe^{5}-p / Phe^{7}-p region of M1 (Fig. 5c). Residues Ile^{97}β and Phe^{98}β make main-chain van der Waals contacts and water-mediated hydrogen bonds with M1 main chain, and residues Gln^{100}β and Arg^{101}β make side chain contacts with MHC side chains (Fig. 5c). Prominently, Phe^{98}β inserts its phenyl ring into the hydrophobic notch formed by Phe^{5}-p, Phe^{7}-p of M1-peptide and the side of the MHC-I α2 helix between Val^{153MHC} and the gatekeeper Gln^{155MHC} (Fig. 5c). This is the same site as occupied by Arg^{98}β of JM22, and by the displaced Phe^{5}-p of LS10 as described above. In the unliganded HLA-A2/M1 structure, Gln^{155MHC} adopts a rotameric conformation that partially blocks access to the hydrophobic notch. In the LS01-bound structure, residue Gln^{155MHC} was displaced from its usual position in the unliganded structure by interactions with CDR3β residues Gln^{100}β and Arg^{101}β (Fig. 5c). These interactions allow access to the hydrophobic notch, which becomes occupied by the side chain of Phe^{98}β. Mutagenesis experiments showed that each of the residues Phe^{98}β, Gln^{100}β, and Arg^{101}β was essential for M1-

recognition (Supplementary Fig. 6e). Considering that all three residues Phe^{98}β, Gln^{100}β, Arg^{101}β are encoded by non-templated sequences and mutation of any of them almost completely abrogates M1 recognition, it is reasonable that TCRs with non-RS motifs are found at lower frequency in the overall M1-specific TCR repertoire than are RS-motif containing TCRs where only Arg^{98}β is required in CDR3β[14].

We examined the LS01/HLA-A2/M1 structure to understand the role of TCRα. The alpha subunit of LS01 makes extensive interactions with HLA-02/M1, contacting both the peptide and MHC α2 helix adjacent to the region contacted by CDR3β but closer to the peptide N-terminus. Prominent contacts are made by Tyr^{31}α from CDR1α and by Asn^{95}α from CDR3α, which together insert into a cleft between the peptide and the MHC α2-helical region (Fig. 5d and Supplementary Notes). Nearby, Asn^{95}α forms hydrogen bonds with Glu^{166MHC} of the α2 helix and with Thr^{94}α of CDR3α, providing additional stabilization to the intricate network of interactions in this region (Fig. 5e and Supplementary Notes). The requirement for specific TCRα interactions from both germline encoded CDR1α and CDR3α sequences in addition to CDR3β aromatic residue at position 98 could provide a clue about the infrequent usage for Group III TCR such as LS01 as compared to canonical ‘xRSx’ motif-containing TCR.

Comparison of binding strategies used by LS01, LS10, and JM22 TCRs to recognize HLA-A2/M1

The LS01, LS10, and JM22 TCRs use different strategies to sense the notch between M1 and MHC α2 helix near Phe^{5}-p. All three TCR use TRBV19, like most TCR recognizing HLA-A2/M1. Although they thus share identical CDR1β and CDR2β sequences, there are subtle differences in hydrogen bonding, van der Waals contacts, and water usage (Fig. 6a), resulting in different energetic contributions of CDR1β/2β residues (see Supplementary Notes). Despite these differences, the overall location of CDR1β/2β loops is preserved, in particular with Ile^{53}β over the MHC-peptide. Pivoting around this residue, the CDR3 loops of the three TCRs position into different locations, recognizing different aspects of the HLA-A2/M1 complex (Fig. 6b). LS01 and JM22 sense the common pocket using Phe^{98}β or Arg^{98}β of CDR3β, respectively. LS10, on the other hand, recognizes the new notch originally taken by Phe^{5}-p using Ala^{98}α and Tyr^{103}α of CDR3α. Cross-sectional views of the cleft between Phe^{5}-p, Phe^{7}-p and α2 helix (dashed lines in Fig. 6b) show that three TCRs fill up this notch using different residues (Fig. 6c). In addition, interactions with gatekeeper Gln^{155MHC} are distinct among three TCRs. In LS01, CDR3β residues Gln^{100}β and Arg^{101}β and CDR1α Tyr^{31}α flip the side chain of Gln^{155MHC} towards the α2 helix, opening up the pocket adjacent to Phe^{7}-p for Phe^{98}β (Fig. 5c, Supplementary Fig. 6g). In JM22, CDR3 Arg^{98}β itself pushes away the gatekeeper to enter the cleft with the aid of Ser^{100}β and main chain of CDR3α. In LS10, Tyr^{103}α, Tyr^{100}β and main chain of CDR3β participate in stabilizing the gatekeeper to create a new cleft for Ala^{98}α/Gly^{99}α.

To further compare interaction patterns of three TCRs for HLA-A2/M1, we derived contact maps showing where TCRα and TCRβ residues contact peptide or MHC (Supplementary Fig. 6j). These show that the LS01 CDR1α contacts more peptide and MHC-α2 than does CDR1α of LS10 or JM22, and that the LS10 CDR3α interacts more with MHC-α1/α2

helices and peptide than do CDR3α of the other TCRs. In contrast to these TCRα chain interaction differences, the similar distribution of contact residue pairs for TCRβ chain interactions illustrates how similarly CDRβ loops of the three TCR interact with MHC helices and peptide (Supplementary Fig. 6j).

Discussion

Previous research has suggested that TCRs utilize a narrow sequence repertoire to recognize HLA-A2/M1[8,9,13]. In some other systems, constrained repertoires have been associated with poor protection/prognosis[19–22]. By deep sequencing TCRα and TCRβ genes from HLA-A2/M1-tetramer sorted cells from influenza-immune donors we obtained a more comprehensive and accurate picture of the diversity of TCRs responding to this antigen than in previous studies[8–12]. We found that the TCR repertoire recognizing HLA-A2/M1 is substantially broader than previously appreciated, with most donors having several hundred different TCRα and TCRβ sequences used by CD8 T cells in resting memory. These numbers are in line with estimates of the precursor frequency of naïve T cells recognizing various antigenic peptides in other mouse and human models[38–40]. Overall diversity measures for this repertoire were at the upper end of the range of values previously reported for other viral epitope-specific responses[19,32,41,42]. As previously reported, a single public TRBV gene segment (TRBV19) with restricted CDR3β motif (xRS/Ax) (Group I) dominated the HLA-A2/M1 response, representing in this cohort ~50% (range 15–72%) of the responding CD8 T cells and 12–55% of the unique sequences. This is a somewhat lower proportion than previous estimates of 74–84% of the overall response[9,11] Single-cell PCR and functional analysis of recombinant TCR transfected into T cells showed that these TRBV19 xRS/Ax chains can pair with many different TCRα, greatly enhancing their chances of selection. We identified a second dominant public TCR with TRBV19 CDR3β motif, xGxY, which represented ~12% (range 2–26%) of the overall HLA-A2/M1 response (Group II), and which was highly restricted to pairing to TRAV38-TRAJ52 TCRα chains with a 15mer CDR3α motif. Another set of TCRs representing 1–4% (Group III) used 11-mer CDR3β but with a hydrophobic residue in place of Arg^{98} in the xRS/Ax motif. Overall Group I TCRs (range 27–87%), Group II TCRs (2.6–40%) and Group III TCRs (1.2–4.8%) comprise most of the TRBV19-restricted response to this important antigen (Fig. 6d).

Together groups I, II and II comprise ~65% of the overall CD8 T cell response to HLA-A2/M1 and represent 45% of the sequence diversity. The remaining 35% of the TCR repertoire (Group IV, 13–60%) includes TRVB19 and non-TRVB19 TCR with many different TRAV and was highly private and diverse without any obvious CDR3 motif. This would suggest that there may be many other structural solutions to recognizing the relatively featureless HLA-A2/M1 complex. The combination of both conserved public and diverse private components to even a single antigen-specific TCR repertoire maybe a basic principle for TCR repertoire structure.

The crystal structures reported here combined with previous work reveal the structural basis for recognition of HLA-A2/M1 by Group I (exemplified by JM22[13,14]), Group II (LS10), and Group III (LS01) TCRs, altogether comprising the majority of the overall HLA-A2/M1 specific response to this immunodominant antigen in our cohort. These structures show that

there are many ways to recognize a featureless peptide. The different TCRs find different solutions to specifically binding HLA-A2/M1, but all utilize a small niche between the peptide and MHC α2 helix. It is tempting to speculate that other TCRs recognizing HLA-A2/M1 might also target these same pockets, and that conformational flexibility in this region might be a defining feature of this MHC-peptide complex, allowing it to be specifically recognized despite the absence of overt structural features that differentiate it from HLA-A2 complexes carrying other peptides. The configuration of the M1 peptide in the LS10-bound complex is nearly identical to that of M1 bound to HLA-C*08 (Supplementary Fig.5b,[43]), and we suggest that HLA-C*08/M1 restricted TCRs, recently shown to be elicited by IAV infection in HLA-C*08+ donors[43], might recognize their ligand utilizing the same cleft as does LS-10.

Experimental equilibrium dissociation constants (Kd) measured for soluble TCR binding to immobilized HLA-A2/M1 were similar for LS10 and LS01, but approximately 5-fold lower for JM22 and the related LS06 (Fig 2d). Examination of the corresponding crystal structures did not reveal any obvious candidates for these differences: buried surface area, number of hydrophobic and total contacts, and predicted interaction energy all were similar for the three complexes or greater for the weaker-binding LS10 and LS01. One difference is the larger number of interfacial hydrogen bonding interactions for JM22, seven, as compared to one or two for the other TCRs. While buried hydrogen bonds are not expected to contribute significantly to the overall binding energy, differential organization of bound solvent in the TCR-peptide-MHC interface has been implicated as playing a role in affinity determination for these complexes[44], and interfacial hydrogen bonding might influence this. We did not obtain detailed structural information for non-bound LS10 or LS01 TCR, but binding-induced conformational changes have been suggested to play a role in affinity determination through entropic effects for JM22[14].

If there are many ways to recognize HLA-A2/M1, why is the repertoire so biased towards TCR having TRBV19 with the 11mer ‘xR(S/A)x’ motif? It has been suggested that these public TCR could represent clonotypes present at high frequency in the naïve precursor pool as a result of bias in the recombination machinery[45] or convergent recombination of key contract sites[46]. While convergent recombination could explain the high abundance of group I TCR and the relatively lower abundance of group II TCR, it cannot explain the much lower abundance of group III TCR, since the group I motif ‘xR(S/A)x’ can be encoded by a similar frequency of random sequences (1.6%) as the group III motif ‘x(F/Y/W)(S/A)x’ (1.3%). Using the structural analyses reported here, we can see that TCR immunodominance patterns seem to scale with number of specific interactions required, and this might provide an alternate explanation for the observed abundance patterns. Group I TCRs are present at the highest frequency, and require only TRBV19 and ‘xRSx’-containing 11mer-CDR3β, with wide latitude in TRAV gene usage and CDR3α sequences. Group II TCR require both TRBV19 with 10-mer ‘xGxY’ and also TRAV with a highly constrained 15-mer CDR3α motif. Group III TCR are present at even lower frequency, and require not only a hydrophobic bulky residue in place of Arg[98] in xRSx, but also a constellation of residues from CDR3β and TCRα. A plot of group frequencies versus the number of TCR residues that have side chains making contact with peptide-MHC is shown in Fig. 6e. It appears that

TCRs able to find simpler solutions to recognizing HLA-A2/M1 by involving fewer specific amino acids are easier to evolve and come to dominate the memory pool.

There have been few reports where multiple TCRs recognize the same peptide-MHC complex[27–29,47]. The HLA-A2/M1-TCR structures reported here are unique in that they enable us to infer the recognition mechanisms of a substantial portion of the natural viral epitope-specific TCR repertoire from any HLA-A2 individual. By contrast, previous studies showed structures of either TCRs against an artificial ligand from genetically different mice[47], different individuals[27], or TCRs that reflect limited repertoire to the viral epitopes at the individual or population level[28,29]. Deep sequencing of TRBV repertoire has been examined for some viral epitopes[25] and TRAV analyses have been examined in CMV-specific responses of 2 donors[26]. However, there are no reports where human viral antigen-specific TCR repertoires have been characterized for both TRBV and TRAV sequences by NGS sequencing. The analyses reported here reveal a combination of a highly diverse public and private repertoire that may be the prototype of a highly successful and resilient response, likely to be present in all HLA-A2+ individuals. Increasing evidence suggests that an antigen-specific TCR repertoire organization with focused diversity, i.e. with dominant public clonotypes combined with an underlying highly diverse private component, may be more common for many antigens than previously thought. For instance, TCR repertoires to two featured epitopes, one from CMV pp65 in HLA-A2+ donors and one from EBV EBNA 3A in HLA-B8+ donors, were once considered highly public oligoclonal responses, but now have been shown by NGS to also contain an underlying polyclonal repertoire[25–27]. Based on our observations of the repertoire responding to HLA-A1/M1, we would expect that the dominant public clonotypes might represent T cell clones preferentially selected by convergent recombination and/or that use a small number of TCR residues to contacting peptide-MHC. Similarly, diverse private clonotypes might represent T cell clones that have more stringent contact requirements, which could be fulfilled in many different ways with various TCRα and TCRβ sequences. Contrary to our expectations, the relatively unfeatured M1 peptide was recognized by many different TCR using different recognition strategies. Other more featured peptides likely also are recognized in different ways by different TCR, and the relationship between TCR dominance patterns and peptide-MHC contacts observed for TCR responding to HLA-A2/M1 might hold for many other antigens.

A highly diverse repertoire, such as the one described here recognizing HLA-A2/M1, should allow resilience against loss of individual clonotypes with aging[32] and against skewing of the response after infection with a cross-reactive pathogen[48,49]. The large number of HLA-A2/M1-specific clonotypes contributes to the overall memory T cell pool, enhancing the opportunity for protective heterologous immunity now recognized to be an important aspect of immune maturation[50,51]. A large pool of TCR clonotypes responding to HLA-A2/M1 could provide increased resistance to viral drift, although the xRS/A-containing JM22 TCR recently has been shown to be able to recognize M1 variants from circulating IAV strains[52]. Finally, it is possible that different TCRs activate antigen specific cell functions differently, leading to a more functionally heterogeneous and more complete pool of memory cells[53]. A better understanding of TCR repertoires is becoming increasingly important as suggested by reports that the diversity index of mucosal resident T cell repertoire predicts clinical

prognosis in gastric cancer[54]. Ideally, vaccines would be able to induce dominant public as well as diverse private responses, in order to provide a resilient repertoire of memory cells.

Combining structural and sequence information, we now have the most comprehensive and highly detailed view of CD8 T cell recognition for any known antigen. Thousands of different TCR sequences representing the bulk of the public HLA-A2/M1-restricted CD8 T cell response, can be understood in terms of the interactions identified in the structural prototypes JM22, LS10, and LS01. The remaining idiosyncratic repertoire includes highly diverse TCR sequences that provide resiliency against clonal loss, diversion, and pathogen variation.

Online Methods

Study population

Blood samples were collected from six HLA-A201 donors. Donor 185, 215, 240 and 264 are healthy IAV immune donors between age 18–20, and D085 and D105 were middle-aged donors. All donors are volunteers from the University of Massachusetts (UMass) Student Health Services (Amherst, Massachusetts, USA) or UMass Medical Center (Worcester, Massachusetts). HLA status was assessed using an HLA-A2-specific mAb (BB7.2; BD Biosciences, San Jose, CA). Donors of this age are assumed to have been exposed to influenza A virus, and all exhibited positive staining with HLA-A2/M1 tetramers indicating that they had been previously exposed[30]. The influenza diagnosis was confirmed based on the symptoms during hospital visit. The Institutional Review Board committee from UMass Medical School in Worcester approved this study, and all donors participating in this study gave informed consent.

Peptide synthesis

Following HLA-A2 specific peptides were synthesized by 21st Century Biochemicals (Marlboro, MA) and purified to 90% purity; IAV $M1_{58–66}$ (GILGFVFTL), two EBV peptides $BMLF1_{280–288}$ (GLCTLVAML) and $BRLF1_{109–117}$ (YVLDHLIVV), human tyrosinase $peptide_{369–377}$ (YMDGTMSQV), and vaccinia virus MVA090 (KLTFLVEV)

Blood preparation and bulk CD8 T cell culture

PBMCs were isolated from fresh blood sample using Ficoll Paque plus (Amersham Biosciences, Piscataway, NJ). CD8 T cells were purified from PBMCs by positive selection using human CD8 specific MicroBeads (Miltenyi Biotech, MA). CD8 T cells (2.5×10^5/ml) were stimulated with peptide-pulsed (1 μM) irradiated TAP-deficient T2 cells (5×10^4/ml) (CRL-1992; ATCC). T cell lines were fed every 3–4 days with AIM-V medium supplemented with 14% human serum, 16% of MLA-144 culture supernatant, 10U/ml rIL-2, 1% L-glutamine, 0.0005% β-mercaptoethanol (Sigma-Aldrich) and 1% HEPES (HyClone)[30]. At the end of each week T cells were counted and re-stimulated with peptide pulsed irradiated T2 cells for a total period of three weeks. In previous experiments using this in vitro protocol to expand antigen-specific CD8 T cells, TCR repertoires were comparable to those observed by tetramer sorting directly ex vivo from PBMCs[55]. Because the expansion step might introduce some skewing of the repertoire[10] we evaluated the

relative frequency of 18 TRBV families directly ex vivo, by antibody staining flow cytometry, and after in vitro expansion, by NGS sequencing. No significant skewing was detected (Supplementary Figure 1a).

Staining and sorting of CD8 T cells

Positive staining with HLA-A2/M1-dextramer (Immundex USA) was used as an indication that these individuals had been exposed to influenza virus. Magnetic-bead purified CD8 T cells from Donor 085 were stained with HLA-A2/M1-dextramer, anti-CD8/anti-CD3. M1/CD3/CD8 positive cells were directly sorted to a 96 well plate (Biorad) with Aria II flow cytometer (BD). The plate was kept at −80°C until further processing. The same staining procedure was applied for bulk M1-specific CD8 T cell isolation from cultured CTL and sorted cells were sorted in tubes containing PBS buffer with 2% fetal bovine serum. After quick spin, cells were resuspended in lysis buffer (RNeasy kit, Qiagen) and stored in −80°C for RNA isolation.

TCR Vβ Analysis ex vivo with mAb

Sorted CD8 T-cells from fresh PBMCs directly *ex vivo* were incubated for 20 min. with IAV-M1-specific tetramer, which was then washed off. An additional 20 min. incubation was performed with 24 TCR Vβ antibodies which cover >70% of commonly used human Vβ (IO Test Beta Mark TCR Vbeta Repertoire Kit, Beckman Coulter, Fullerton, CA). Samples were read on LSRII (Beckman Coulter, Fullerton, CA). IMGT TCR gene nomenclature was used to define TCR Vβ types.

RNA isolation, cDNA synthesis and NGS of bulk M1-specific cells

Total RNA was isolated from lysates of cultured and sorted CD8 T cells using RNeasy Mini kit (Qiagen) following the manufacturer's recommendations. TCRα and TCRβ CDR3 regions were amplified and sequenced from cDNA reverse transcribed from 100 ng-300 ng of total RNA samples (SuperScript VILO cDNA Synthesis Kit, Invitrogen). Amplification and high-throughput sequencing of CDR3 regions were performed on the ImmunoSEQ platform at Adaptive Biotechnologies as previously described[56]. The platform uses a panel of multiplexed TRV and TRJ primers, selected to reduce differential amplification bias of TCRα and TCRβ sequences[56]. The use of cDNA as a source of TCRα and TCRβ templates potentially can introduce bias due to differential mRNA expression, reverse transcription, or PCR amplification. These effects generally are expected to introduce errors in relative abundance calculations of ~2-fold or less[57,58], although a recent analysis has indicated the potential for much greater skewing using a different reverse transcription protocol[59]. Such skewing would affect the total sequence numbers reported here but not the numbers of unique clonotypes

Multiplex nested single cell RT-PCR

Single cells sorted into 96-well plates were subjected to cDNA synthesis using SuperScript VILO cDNA Synthesis Kit (Invitrogen) in 2.5 μl reaction mixture containing 0.1% Trition X-100 (Sigma). Multiplex V-gene specific primer sets were used in two rounds of PCR to amplify CDR3α/β from single cells as previously described[19]. CDR3 amplicons were

purified (ExoSAP-IT) and sequenced with primers recognizing constant regions of TRAC and TRBC[19]. Sanger DNA sequencing was performed by Genewiz (Cambridge, MA).

Identifying CDR3 sequences and NGS analysis

The TCRα and TCRβ CDR3 sequences were identified according to the definition founded by the International ImMunoGeneTics collaboration[60]. NGS data were analyzed using ImmunoSEQ Analyzer 2.0 provided by Adaptive Biotechnologies (http://www.adaptivebiotech.com/immunoseq/analyzer). Single cell CDR3 sequences were analyzed by IMGT/V-QUEST[61]. Only productively rearranged TCRα and TCRβ sequences without stop codon were used for repertoire analysis including sequence composition and gene frequency analyses. V gene frequencies, CDR3 length and V/J gene pairing were analyzed using subprograms of ImmunoSEQ Analyzer software and further processed by Microsoft Excel. Conserved motifs in CDR3 were assessed using Weblogo software (http://weblogo.berkeley.edu/logo.cgi). Clonotype diversity for each donor was evaluated using Shannon diversity index $H=-\Sigma p_i \ln(pi)$ and Simpson diversity index $SI=1-\Sigma p_i^2$, where p is the clonotype fractional abundance.

Cloning of M1-specific full length TCRα and TCRβ chain

The multiplex nested single cell PCR strategy described above identified paired CDR3α/β sequences and information on V gene usage, but for cloning full-length TCRα and TCRβ we needed to isolate V gene sequences upstream of the CDR3 and C gene sequences downstream. Individual full-length TCRα and TCRβ chain genes were prepared by ligating 5' fragments that cover signal sequence of V gene to part of CDR3 and 3' fragments corresponding CDR3 and termination codon of TRC gene. As template DNA, cDNA from remaining PBMC after CD8 T cell separation was used for 5' and 3' fragment amplification. Primers were designed in a way that 3' end of amplified 5' fragment should overlap 5' end of amplified 3' fragment. More specifically, for 5' fragment amplification, the forward primers included *Nde* I and *Eco*R I for TCRα and TCRβ respectively and sequences that could anneal to signal sequence of each V gene and the reverse primer contained sequence 5' of CDR3 that could anneal to the 3' end of germline V gene. For 3' fragment amplification, the forward primers were designed to have sequence 3' of CDR3 and germline J gene for annealing and the reverse primers included sequence that should anneal to 3' end of C gene and *Bam*H I and *Bsp*E I for TCRα and TCRβ respectively. The 5' fragments and 3' fragments were amplified using Phusion enzyme (NEB) and were gel-purified for subsequent overlapping PCR.

Construction of TCRα/β expression vector

Full-length TCRα and TCRβ from overlapping PCR were cloned into mammalian expression vector containing eGFP protein (pEF1-IRES-eGFP, CLONTECH). The "self-cleaving" 2A sequence of foot and mouth disease virus was inserted between TCRα and TCRβ to express two chains simultaneously[62]. *Eco*RI, *Bsp*EI, *Nde*I and *Bam*HI enzyme sites were used to insert TCRβ, 2A sequence, and TCRα (*Eco*RI-TCRβ-*Bsp*EI-2A-*Nde*I-TCRα-*Bam*HI) into the vector. To generate mutants of LS01 and LS10, *Dpn*I mediated site-directed mutagenesis method was applied. Full-length sequencing of inserts of wild type and mutants was performed to confirm the absence of unwanted mutation.

Transient expression of TCR in TCR-deficient J76CD8

TCRα/β expression vectors were transferred via electroporation into TCR deficient and CD8α expressing Jurkat cells (J76-CD8α)[63] kindly provided by Dr. Wolfgang Uckert (Max Delbruck center, Berlin). J76-CD8α cells were maintained in RPMI1640 media containing 10% FBS in early log phase for transfection. Four million J76-CD8 cells (10^7/ml) were harvested and mixed with 10 – 15 μg of expression vector plasmid in a 4 mm gap cuvette (Biorad) for electroporation. Transfection was performed using BTX electroporator (260 V, 1050 μF). Thirty-six hours after transfection, cells were harvested for MHC-tetramer staining.

Construction of TCR-expressing stable cell lines

For T cell activation (CD69 upregulation) assays, stable cell lines individually expressing the 13 different TCRα/β pairs were constructed. Linearized TCR expression vectors were transferred into J76-CD8α through electroporation as described above. Two days after transfection, cells were transferred and selected in 400 μg/ml of G418-containing media for 3 weeks. For LS01, LS06, LS10 or JM22, G418-resistant cells were further diluted and transferred into 96 well plates at a density of 0.5 cell/well for single cell cloning. G418 resistant single cell clones were first selected against GFP expression to choose expression vector containing clones, then further screened for TCR expression.

Preparation of peptide-HLA-A2 monomer/tetramer

We prepared peptide-HLA-A2 monomers by folding urea-solubilized bacterially-expressed inclusion bodies of HLA-A2 heavy chain and human β2-microglobin in the presence of 5mg/1L of M1 peptide or control peptide MVA090 as described[64]. In some cases the modified HLA-A2 heavy chain with a free c-terminal cysteine at position 282 was used to add biotin using thiol chemistry for tetramer preparation and Biacore experiments. The folding mixture was filtered with 0.2 μM filter unit (Corning) and buffer exchanged with 10mM Tris-Cl (pH 8.0) using a tangential flow concentrator. Folded HLA-A2/peptide complexes were isolated from the buffer-exchanged folding mixture by a series of chromatography steps consisting of Hitrap Q and Mono Q ion exchange and S-200 gel-filtration columns (GE healthcare). For Biacore experiments and tetramer preparation, purified cysteine-containing HLA-A02-peptide monomers were reduced with 5mM DTT and biotinylated using EZ-Link® Maleimide-PEG2-Biotin (Thermo Scientific). Biotinylated M1-monomers were multimerized by R-phycoerythrin (PE)-labeled streptavidin by mixing at final 5:1 (monomer: streptavidin) ratio.

Staining of TCR-expressing J76-CD8α and flow cytometry

TCR-expressing J76-CD8 cells were washed twice with FACS buffer (PBS with 2% BSA, 0.02% azide) and 0.2–0.5 million cells were stained in 100 μl staining solution containing LIVE/DEAD® Fixable Violet Dead Cell Stain, anti-TCR antibody (BioLegend, IP26) and HLA-A2/M1 dextramer or control dextramer (HLA-A2/BRLF1 dextramer, Immudex) for 30 min at room temperature. For dose dependent tetramer staining, increasing amount of HLA-A2/peptide tetramers (300 nM, 120 nM, 48 nM, 19.2 nM, 7.68 nM, 3.07 nM) were added instead of dextramers. Stained cells were washed three times with FACS buffer for flow

cytometry. Samples were analyzed using a LSRII flow cytometer (BD Biosciences) and FlowJo software (Tree Star).

CD69 up-regulation assay

For assessing early TCR-driven signaling, J76-CD8 stably expressing TCRs were stimulated with peptide-pulsed HLA-A2-tranfected T2 cells[65]. T2-A2 cells were loaded with a single dose (1 μM) M1-peptide or with varying concentration of M1-peptide (10^{-12} M to 10^{-7} M) for 1 hr at 37 °C. Unloaded excess M1-peptide was removed by washing cells twice with PBS. M1-loaded T2-A2 cells (0.1×10^6 cells) were incubated with TCR expressing J76-CD8 cells (0.4×10^6 cells) at 37°C in 24-well dishes for 12 hr. After 12 hr incubation, cells were cooled on ice and harvested for FACS staining. 0.2–0.5 million cells were stained with anti-CD69 antibody (BioLegend, FN50), anti-TCR antibody BioLegend, IP26), and LIVE/DEAD® Fixable Violet dye for 30 min 4°C. Stained cells were washed with PBS and analyzed by FACS as described above.

Preparation of soluble TCR

For Biacore experiments and crystallization, extracellular portions of TCRα and TCRβ chains of LS01 and LS10 were engineered as stable soluble TCRs (sTCR) by introducing a interchain disulfide as previously described[66]. Engineered TCRα and TCRβ chains were expressed as inclusion bodies. Urea-solubilized TCRα and TCRβ inclusion bodies were mixed and folded by dilution and dialysis as described[67]. Dialyzed folded sTCRs were purified by successive chromatography using Hitrap Q (5ml), Mono Q (5ml) and Superdex 26/600 (320ml) columns. Final gel-filtration chromatography was used to exchange buffers for crystallization (10 mM Tris-Cl pH 8.0, 50 mM NaCl, 1 mM EDTA, 0.02% NaN3) or Biacore experiments (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% Tween20).

Surface plasmon resonance analysis

Neutravidin (~ 3000 RU) was immobilized in four flow-cells of CM5 chip at pH 5.5 via amine coupling reaction in a Biacore 3000 instrument (BIAcore AB). Biotinylated HLA-A2/M1 was bound to immobilized neutravidin in flow cell 3 and 4 to achieve ~1200 RU. Neutravidin alone in flow cell 1 and HLA-A2/ MVA090-bound neutravidin in flow cell 2 were used as negative controls. Solutions of LS01and LS10 sTCR (soluble TCR, 2-fold dilutions from 150 μM to 1.17 μM) were injected over the whole cell (flow cells 1 to 4) at 5 μL/min for 3 min to allow equilibrium binding followed by return to running buffer for dissociation. Running buffer was injected at 50 μl/min for 0.25 min to regenerate cells. The increased RU signal from specific binding of LS01 and LS10 to immobilized HLA-A2/M1 was calculated by subtracting RU from flow cell immobilized with irrelevant peptide-HLA-A2 complex and plotted against sTCR concentration to calculate equilibrium dissociation constants (Kd) using Prism 6 software (GraphPad Software, Inc.)

Crystallization and data collection

Purified sTCR and HLA-A2/M1 (without free C-terminal cysteine) were mixed at final concentration of 10–15 mg/ml at 1:1 molar ratio overnight to preform TCR-MHC-peptide

ternary complex. All crystallization conditions were set up with sitting drop vapor diffusion technique in 96 well plates by mixing 0.5 μl of reservoir buffer and 0.5 μl of protein mixture at room temperature. The crystallization plates were stored at room temperature or 4°C. One buffer condition (14% (w/v) PEG 4000, 100 mM Na-HEPES pH 7.0, 200 mM ammonium sulfate) gave crystal plates of LS01 /HLA-A2/M1 complex grown at 4°C. Crystals of LS10/HLA-A2/M1 complex grew at 21°C as plates in a buffer (10% (w/v) PEG 8000, 100 mM Tris-Cl pH 7.0, 200 mM magnesium chloride) from Wizard Classic crystallization screen (Rigaku). Crystals were briefly soaked in 1:1 mixture of saturated sucrose and reservoir buffer for cryoprotection and flash-frozen in liquid nitrogen and sent to LRL-CAT beamline at the **Advanced Photon Source (Argonne, IL USA)**. Two data sets (2.06 Å and 2.48 Å for LS01 and LS10, respectively) were collected with 0.979 Å wavelength radiation and MAR-165 charge-coupled device (CCD) detector. Data sets were indexed, integrated, and scaled with iMOSFLM and SCALA[68–70]. LS01-complex crystal belonged to P21 space group and LS10-complex crystal to P1 space group (two molecules per asymmetric unit). Detailed data collection statistics and unit cell parameters are shown in Table 1.

Structure determination and refinement

Both TCR structures were determined by molecular replacement using Phaser[71,72] and a sculpted JM22/HLA-A2/M1 structure (PDB:1OGA) as an initial search model with separate two search ensembles, where MHC/peptide was first located and followed by TCR. LS01 complex contained one pMHC and one TCR in asymmetric unit and LS10 complex contained two pMHC and two TCRs in asymmetric unit. After one round of rigid body refinement using separate structural domains (HLA-A2 α1 and α2, α3, β2m, peptide, Vα, Cα, Vβ and Cβ) of output model from molecular replacement, models were built and refined by AutoBuild component of PHENIX[73]. Composite omit maps of both models clearly showed densities for peptide, MHC, and TCR. Autobuilt models were further refined with several rounds of manual model building using the software COOT[74] and automated refinement cycles using the phenix.refine program with diverse parameter adjustments (XYZ coordinates, real-space, rigid body, simulated annealing (torsion), individual restrained B-factor, and NCS restrain for LS10, simulated annealing. The final LS01/HLA-A2/M1 model had 96.9% of residues in favored regions of the Ramachandran plot, 3.1% in allowed regions and no outliers, and included MHC heavy chain residues 1–275, β2-microglobulin residues 2–99, the complete M1 peptide residues 1–9, TCRα subunit residues 2–201, and TCRβ subunit residues 3–243. Side chains of residues of following residues were modeled but with lower confidence: 222–226 of HLA-A2; 58, 83, 127–131, 142 of TCRα; 43, 118, 180–185, 219–222 of TCRβ. Side chains of residues 126, 129 of TCRα and 118, 219, 220, 222 of TCRβ were not resolved. The final LS10/HLA-A2/M1 model has two molecules in an asymmetric unit with 97.1% of residues in favored regions of the Ramachandran plot, 2.9% in allowed regions and no outliers, and included two MHC heavy chain residues of 1–276, two β2-microglobulin residues 0–99, the two complete M1 peptide residues of 1–9, two TCRα subunit residues of 3–132,138–208, and two TCRβ subunit residues of 4–242. Residues 133–137 of two TCRα chains were not built and following side chains were not resolved: two residue 48 of b2-microglobulin; two 85, 115,117 residues of TCRβ; residue 268 of MHC heavy chain; residue 137 of TCRα. Following loops were modeled but with lower confidence: 190–200, 218–228, 246–258, 272–276 of MHC heavy chains; 54–62,

154–158 of TCRα chains; 179–183, 215–230 of TCRβ chains. Also, electron density of side chains of following residues were very weak: residue 268 of MHC heavy chains; residue 19,44,58, 6,74,75 of β2-microglobulins; residue 54, 121, 171, 187 of TCRα chains, 14,16,17,72, 86, 163 of TCRβ chains. Diffraction data and coordinates were deposited with the PDB with accession codes 5ISZ and 5JHD.

Structure analysis

Pymol (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC) was used for graphical representation of TCR footprint and interacting residues. BSA between MHC/peptide and TCR and predicted interfacial binding energy was analyzed using PISA server (http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver). Individual contribution of TCRα, TCRβ, α1, α2 of MHC, and peptide to BSA was plotted as bar graphs. Residue-residue interactions were assessed using contact map analysis server (http://ligin.weizmann.ac.il/cma/). BSA's of HLA-A2 bound peptides were analyzed by PISA server using following PDB deposited models: 1B0G, 1DUZ, 1EEY, 1EEZ, 1HHG, 1HHH, 1HHI, 1HHJ, 1HHK, 1I1F, 1I1Y, 1I4F, 1I7R, 1I7T, 1I7U, 1IM3, 1JF1, 1JHT, 1P7Q, 1QEW, 1QR1, 1S8D, 1S9W, 1S9X, 1S9Y, 1T1W, 1T1X, 1T1Y, 1T1Z, 1T20, 1T21, 1T22, 1TVB, 1TVH, 2AV1, 2AV7, 2C7U, 2CLR, 2GIT, 2GT9, 2GTW, 2GTZ, 2GUO, 2V2W, 2V2X, 2VLL, 2X4O, 2X4R, 2X4P, 2X4S, 2X4T, 2X4U, 2X70, 3BGM, 3BH8, 3BH9, 3BHB, 3D25, 3FQR, 3FQT, 3FQU, 3FQW, 3FQX, 3FT2, 3FT3, 3FT4, 3GIV, 3GSO, 3GSQ, 3GSR, 3GSU, 3GSV, 3GSW, 3GSX, 3H7B, 3H9H, 3HPJ, 3I6G, 3I6K, 3IXA, 3KLA, 3MGO, 3MGT, 3MR9, 3MRB, 3MRC, 3MRD, 3MRF, 3MRG, 3MRH, 3MRI, 3MRJ, 3MRK, 3MRL, 3MRM, 3MRN, 3MRO, 3MRP, 3MRQ, 3MRR, 3MYJ, 3O3A, 3O3B, 3O3D, 3O3E, 3PWJ, 3PWL, 3PWN, 3QFD, 3REW, 3TO2, 3UTQ, 3V5D, 3V5H, 3V5K, 4E5X, 4GKN, 4GKS, 4I4W, 4JFO, 4JFP, 4JFQ, 4K7F, 4NNX, 4NNY, 4NO2, 4NO3, 4NO5, 4UQ3, 4WJ5. Existence of interaction between Arg^{157MHC} and TCRs in 35 TCR-HLA-A2/peptide complexes were examined in following models: 1AO7, 1BD2, 1LP9, 1OGA, 1QRN, 1QSE, 1QSF, 2BNQ, 2BNR, 2GJ6, 2J8U, 2JCC, 2UWE, 2VLJ, 2VLK, 2VLR, 3D39, 3D3V, 3GSN, 3H9S, 3HG1, 3O4L, 3PWP, 3QDG, 3QDJ, 3QDM, 3QEQ, 3QFJ, 3UTS, 4EUP, 4FTV, 4L3E, 4QOK, 5D2L, 5D2N.

Statistical analyses

Non-linear least squares curve fitting and statistical analyses were performed using Prism version 6 (Graphpad Software Inc.). Statistical tests included Student's T test, linear regression, Pearson correlation co-efficient, repeated measures two-way ANOVA, and Wilcoxon matched pairs signed rank test.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH grants AI038996 (LJS), AI49320 (LKS), and AI109858 (LJS, LKS) and a Nebraska Research Initiative Grant (DG). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the Lilly Research Laboratories

Collaborative Access Team (LRL-CAT) beamline at Sector 31 of the Advanced Photon Source was provided by Eli Lilly Company, which operates the facility. We thank J. Birtley and Z. Maben for assistance with crystallization, freezing and shipping crystals, P. Trehn for advice on model analysis, W. Uckert for TCRαβ-negative Jurkat J76 transfected with human CD8α and P. Thomas for technical advice on single cell PCR.

References

1. La Gruta NL, Turner SJ. T cell mediated immunity to influenza: mechanisms of viral control. Trends Immunol. 2014; 35:396–402. [PubMed: 25043801]
2. Guo H, Lambert K, Takimoto T, Topham DJ. T Cell-Mediated Protection against Lethal 2009 Pandemic H1N1 Influenza Virus Infection in a Mouse Model. J. Virol. 2011; 85:448–455. [PubMed: 20980523]
3. Tan ACL, et al. The design and proof of concept for a CD8+ T cell-based vaccine inducing cross-subtype protection against influenza A virus. Immunol. Cell Biol. 2013; 91:96–104. [PubMed: 23146941]
4. Gotch F, McMichael A, Smith G, Moss B. Identification of viral molecules recognized by influenza-specific human cytotoxic T lymphocytes. J. Exp. Med. 1987; 165:408–416. [PubMed: 3029268]
5. Gotch F, Rothbard J, Howland K, Townsend A, McMichael A. Cytotoxic T lymphocytes recognize a fragment of influenza virus matrix protein in association with HLA-A2. Nature. 1987; 326:881–882. [PubMed: 2437457]
6. Assarsson E, et al. Immunomic Analysis of the Repertoire of T-Cell Specificities for Influenza A Virus in Humans. J. Virol. 2008; 82:12241–12251. [PubMed: 18842709]
7. Keskin DB, et al. Physical detection of influenza A epitopes identifies a stealth subset on human lung epithelium evading natural CD8 immunity. Proc. Natl. Acad. Sci. 2015; 112:2151–2156. [PubMed: 25646416]
8. Moss, Pa, et al. Extensive conservation of alpha and beta chains of the human T-cell antigen receptor recognizing HLA-A2 and influenza A matrix peptide. Proc. Natl. Acad. Sci. U. S. A. 1991; 88:8987–8990. [PubMed: 1833769]
9. Lehner PJ, et al. Human HLA-A0201-restricted cytotoxic T lymphocyte recognition of influenza A is dominated by T cells bearing the Vβ17 gene segment. J. Exp. Med. 1995; 181:79–91. [PubMed: 7807026]
10. Naumov YN, et al. Complex T cell memory repertoires participate in recall responses at extremes of antigenic load. J. Immunol. 2006; 177:2006–2014. [PubMed: 16849515]
11. Naumov YN, Hogan KT, Naumova EN, Pagel JT, Gorski J. A class I MHC-restricted recall response to a viral peptide is highly polyclonal despite stringent CDR3 selection: implications for establishing memory T cell repertoires in ‘real-world’ conditions. J. Immunol. 1998; 160:2842–2852. [PubMed: 9510187]
12. Naumov YN, et al. Multiple glycines in TCR α-chains determine clonally diverse nature of Human T cell memory to influenza A virus. J. Immunol. 2008; 181:7407–7419. [PubMed: 18981164]
13. Stewart-Jones GBE, McMichael AJ, Bell JI, Stuart DI, Jones EY. A structural basis for immunodominant human T cell receptor recognition. Nat. Immunol. 2003; 4:657–663. [PubMed: 12796775]
14. Ishizuka J, et al. The Structural Dynamics and Energetics of an Immunodominant T Cell Receptor Are Programmed by Its Vβ Domain. Immunity. 2008; 28:171–182. [PubMed: 18275829]
15. Turner SJ, Doherty PC, McCluskey J, Rossjohn J. Structural determinants of T-cell receptor bias in immunity. Nat. Rev. Immunol. 2006; 6:883–894. [PubMed: 17110956]
16. Motozono C, et al. Molecular Basis of a Dominant T Cell Response to an HIV Reverse Transcriptase 8-mer Epitope Presented by the Protective Allele HLA-B*51:01. J. Immunol. 2014; 192:3428–3434. [PubMed: 24600035]
17. Sun X, et al. Superimposed Epitopes Restricted by the Same HLA Molecule Drive Distinct HIV-Specific CD8+ T Cell Repertoires. J. Immunol. 2014; 193:77–84. [PubMed: 24899498]
18. Turner SJ, et al. Lack of prominent peptide-major histocompatibility complex features limits repertoire diversity in virus-specific CD8+ T cell populations. Nat. Immunol. 2005; 6:382–389. [PubMed: 15735650]

19. Wang GC, Dash P, McCullers JA, Doherty PC, Thomas PG. T cell receptor αβ diversity inversely correlates with pathogen-specific antibody levels in human cytomegalovirus infection. Sci. Transl. Med. 2012; 4:128ra42.
20. Messaoudi I, Guevara Patiño JA, Dyall R, LeMaoult J, Nikolich-Zugich J. Direct link between mhc polymorphism, T cell avidity, and diversity in immune defense. Science. 2002; 298:1797–1800. [PubMed: 12459592]
21. Yager EJ, et al. Age-associated decline in T cell repertoire diversity leads to holes in the repertoire and impaired immunity to influenza virus. J. Exp. Med. 2008; 205:711–723. [PubMed: 18332179]
22. Price, Da, et al. T Cell Receptor Recognition Motifs Govern Immune Escape Patterns in Acute SIV Infection. Immunity. 2004; 21:793–803. [PubMed: 15589168]
23. Vojnov L, et al. GagCM9-Specific CD8+ T Cells Expressing Limited Public TCR Clonotypes Do Not Suppress SIV Replication In Vivo. PLoS One. 2011; 6:e23515. [PubMed: 21887264]
24. Woodsworth DJ, Castellarin M, Holt RA. Sequence analysis of T-cell repertoires in health and disease. Genome Med. 2013; 5:98. [PubMed: 24172704]
25. Klarenbeek PL, et al. Deep Sequencing of Antiviral T-Cell Responses to HCMV and EBV in Humans Reveals a Stable Repertoire That Is Maintained for Many Years. PLoS Pathog. 2012; 8:e10002889.
26. Link CS, et al. Abundant CMV reactive clonotypes in the CD8(+) T cell receptor alpha repertoire following allogeneic transplantation. Clin. Exp. Immunol. 2016; 184:389–402. [PubMed: 26800118]
27. Gras S, et al. A Structural Basis for Varied αβTCR Usage against an Immunodominant EBV Antigen Restricted to a HLA-B8 Molecule. J. Immunol. 2012; 188:311–321. [PubMed: 22140258]
28. Liu YC, et al. Highly Divergent T-cell Receptor Binding Modes Underlie Specific Recognition of a Bulged Viral Peptide bound to a Human Leukocyte Antigen Class I Molecule. J. Biol. Chem. 2013; 288:15442–15454. [PubMed: 23569211]
29. Yang X, et al. Structural basis for clonal diversity of the public T cell response to a dominant human cytomegalovirus epitope. J. Biol. Chem. 2015; 290:29106–29119. [PubMed: 26429912]
30. Clute SC, et al. Cross-reactive influenza virus-specific CD8+ T cells contribute to lymphoproliferation in Epstein-Barr virus-associated infectious mononucleosis. J. Clin. Invest. 2005; 115:3602–3612. [PubMed: 16308574]
31. Tan ACL, La Gruta NL, Zeng W, Jackson DC. Precursor frequency and competition dictate the HLA-A2-restricted CD8+ T cell responses to influenza A infection and vaccination in HLA-A2.1 transgenic mice. J. Immunol. 2011; 187:1895–1902. [PubMed: 21765016]
32. Gil A, Yassai MB, Naumov YN, Selin LK. Narrowing of human influenza A virus-specific T cell receptor alpha and beta repertoires with increasing age. J. Virol. 2015; 89:4102–4116. [PubMed: 25609818]
33. Davis MM. The problem of plain vanilla peptides. Nat. Immunol. 2003; 4:649–650. [PubMed: 12830143]
34. Bridgeman JS, Sewell AK, Miles JJ, Price Da, Cole DK. Structural and biophysical determinants of αβ T-cell antigen recognition. Immunology. 2012; 135:9–18. [PubMed: 22044041]
35. Gras S, et al. A structural voyage toward an understanding of the MHC-I-restricted immune response: lessons learned and much to be learned. Immunol. Rev. 2012; 250:61–81. [PubMed: 23046123]
36. Buslepp J, Wang H, Biddison WE, Appella E, Collins EJ. A correlation between TCR Vα docking on MHC and CD8 dependence: implications for T cell selection. Immunity. 2003; 19:595–606. [PubMed: 14563323]
37. Tynan FE, et al. T cell receptor recognition of a 'super-bulged' major histocompatibility complex class I–bound peptide. Nat. Immunol. 2005; 6:1114–1122. [PubMed: 16186824]
38. Jenkins MK, Moon JJ. The Role of Naive T Cell Precursor Frequency and Recruitment in Dictating Immune Response Magnitude. J. Immunol. 2012; 188:4135–4140. [PubMed: 22517866]
39. Neller MA, et al. Naive CD8 + T-cell precursors display structured TCR repertoires and composite antigen-driven selection dynamics. Immunol. Cell Biol. 2015; 93:625–633. [PubMed: 25801351]
40. Tscharke DC, Croft NP, Doherty PC, Gruta NLLa. Sizing up the key determinants of the CD8 + T cell response. Nat. Rev. Immunol. 2015; 15:705–716. [PubMed: 26449178]

41. Costa AI, et al. Complex T-Cell Receptor Repertoire Dynamics Underlie the CD8 T-Cell Response to HIV-1. J. Virol. 2015; 89:110–119. [PubMed: 25320304]
42. Koning D, et al. CD8+ TCR repertoire formation is guided primarily by the peptide component of the antigenic complex. J. Immunol. 2013; 190:931–939. [PubMed: 23267020]
43. Choo JAL, Liu J, Toh X, Grotenbreg GM, Ren EC. The immunodominant influenza A virus M158-66 cytotoxic T lymphocyte epitope exhibits degenerate class I major histocompatibility complex restriction in humans. J. Virol. 2014; 88:10613–10623. [PubMed: 24990997]
44. Madura F, et al. T-cell receptor specificity maintained by altered thermodynamics. J. Biol. Chem. 2013; 288:18766–18775. [PubMed: 23698002]
45. Yassai M, et al. Naive T Cell Repertoire Skewing in HLA-A2 Individuals by a Specialized Rearrangement Mechanism Results in Public Memory Clonotypes. J. Immunol. 2011; 186:2970–2977. [PubMed: 21282510]
46. Venturi V, Price DA, Douek DC, Davenport MP. The molecular basis for public T-cell responses? Nat. Rev. Immunol. 2008; 8:231–238. [PubMed: 18301425]
47. Dai S, et al. Crossreactive T Cells Spotlight the Germline Rules for αβ T Cell-Receptor Interactions with MHC Molecules. Immunity. 2008; 28:324–334. [PubMed: 18308592]
48. Selin LK, et al. Heterologous immunity: Immunopathology, autoimmunity and protection during viral infections. Autoimmunity. 2011; 44:328–347. [PubMed: 21250837]
49. Liisa KS, et al. Memory of mice and men: CD8+ T-cell cross-reactivity and heterologous immunity. Immunol. Rev. 2006; 211:164–181. [PubMed: 16824126]
50. Che JW, Kraft ARM, Selin LK, Welsh RM. Regulatory T Cells Resist Virus Infection-Induced Apoptosis. J. Virol. 2015; 89:2112–2120. [PubMed: 25473049]
51. Benn CS, Netea MG, Selin LK, Aaby P. A Small Jab - A Big Effect: Nonspecific Immunomodulation By Vaccines. Trends Immunol. 2013; 34:431–439. [PubMed: 23680130]
52. Valkenburg SA, et al. Molecular basis for universal HLA-A*0201-restricted CD8+ T-cell immunity against influenza viruses. Proc. Natl. Acad. Sci. U. S. A. 2016; 113:4440–4445. [PubMed: 27036003]
53. Han A, Glanville J, Hansmann L, Davis MM. Linking T-cell receptor sequence to functional phenotype at the single-cell level. Nat. Biotechnol. 2014; 32:684–692. [PubMed: 24952902]
54. Jia Q, et al. Diversity index of mucosal resident T lymphocyte repertoire predicts clinical prognosis in gastric cancer. Oncoimmunology. 2015; 4:e1001230. [PubMed: 26137399]

References

55. Clute SC, et al. Broad Cross-Reactive TCR Repertoires Recognizing Dissimilar Epstein-Barr and Influenza A Virus Epitopes. J. Immunol. 2010; 185:6753–6764. [PubMed: 21048112]
56. Robins HS, et al. Comprehensive assessment of T-cell receptor β-chain diversity in αβ T cells. Blood. 2009; 114:4099–4107. [PubMed: 19706884]
57. Carlson CS, et al. Using synthetic templates to design an unbiased multiplex PCR assay. Nat. Commun. 2013; 4:2680. [PubMed: 24157944]
58. Bolotin DA, et al. Next generation sequencing for TCR repertoire profiling: Platform-specific features and correction algorithms. Eur. J. Immunol. 2012; 42:3073–3083. [PubMed: 22806588]
59. Best K, Oakes T, Heather JM, Shawe-taylor J, Chain B. Computational analysis of stochastic heterogeneity in PCR amplification efficiency revealed by single molecule barcoding. Sci. Rep. 2015:14629. [PubMed: 26459131]
60. Monod MY, Giudicelli V, Chaume D, Lefranc MP. IMGT/JunctionAnalysis: The first tool for the analysis of the immunoglobulin and T cell receptor complex V-J and V-D-J JUNCTIONs. Bioinformatics. 2004; 20:i379–i385. [PubMed: 15262823]
61. Brochet X, Lefranc MP, Giudicelli V. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. Nucleic Acids Res. 2008; 36:503–508.
62. Szymczak AL, et al. Correction of multi-gene deficiency in vivo using a single ‘self-cleaving’ 2A peptide-based retroviral vector. Nat. Biotechnol. 2004; 22:589–594. [PubMed: 15064769]

63. Liang X, et al. A Single TCR aChain with Dominant Peptide Recognition in the Allorestricted HER2/neu-Specific T Cell Repertoire. J. Immunol. 2010; 184:1617–1629. [PubMed: 20042572]

64. Garboczi DN, Hung DT, Wiley DC. HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in Escherichia coli and complexed with single antigenic peptides. Proc. Natl. Acad. Sci. 1992; 89:3429–3433. [PubMed: 1565634]

65. Salter RD, Cresswell P. Impaired assembly and transport of HLA-A and -B antigens in a mutant TxB cell hybrid. EMBO J. 1986; 5:943–949. [PubMed: 3522223]

66. Boulter JM, et al. Stable, soluble T-cell receptor molecules for crystallization and therapeutics. Protein Eng. Des. Sel. 2003; 16:707–711.

67. Willcox BE, et al. Production of soluble αβ T-cell receptor heterodimers suitable for biophysical analysis of ligand binding. Protein Sci. 1999; 8:2418–2423. [PubMed: 10595544]

68. Winn MD, et al. Overview of the CCP 4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. 2011; 67:235–242. [PubMed: 21460441]

69. Evans P. Scaling and assessment of data quality. Acta Crystallogr. Sect. D Biol. Crystallogr. 2006; 62:72–82. [PubMed: 16369096]

70. Battye TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW. iMOSFLM : a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. Sect. D Biol. Crystallogr. 2011; 67:271–281. [PubMed: 21460445]

71. McCoy AJ, et al. Phaser crystallographic software. J. Appl. Crystallogr. 2007; 40:658–674. [PubMed: 19461840]

72. Morin A, et al. Collaboration gets the most out of software. Elife. 2013; 2:e01456. [PubMed: 24040512]

73. Adams PD, et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr. 2010; 66:213–221. [PubMed: 20124702]

74. Emsley P, Cowtan K. Coot : model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. 2004; 60:2126–2132.

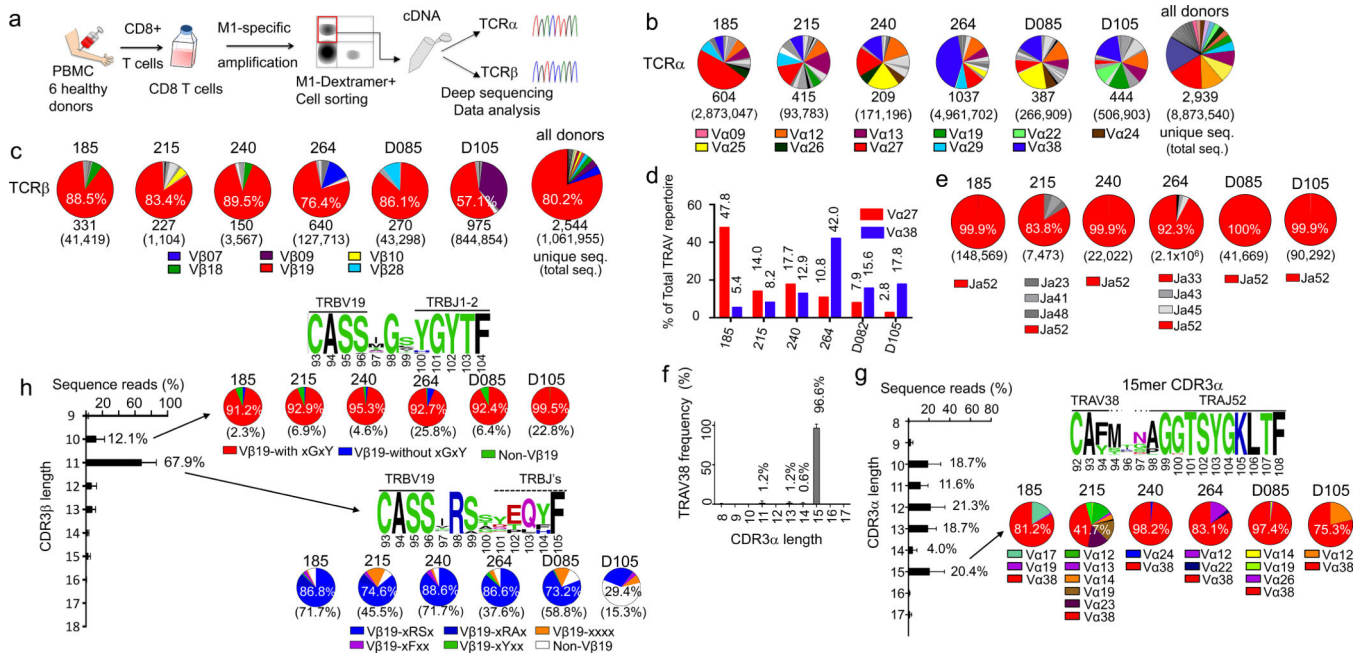

Figure 1. Diversity of CD8 T cell repertoire, dominant usage of TRAV38, and CDR3α and CDR3β sequence motifs in the HLA-A2/M1-specific response

(a) Schematic diagram. CD8 T cells isolated from donor PBMC were expanded using M1-pulsed HLA-A2+ presenting cells, and cDNA of M1-dextramer positive CD8 T cells sorted by FACS was analyzed by NGS. (b,c) HLA-A2/M1-specific TCR repertoires were analyzed for 6 healthy donors (185, 215, 240, 264, D085, D105). Frequency of each TRAV (b) and TRBV (c) in total HLA-A2/M1- specific TCR repertoire is shown in pie charts. Number of unique sequences found for each donor is shown with total number of sequences in parentheses. (d) Frequency of two most common TRAV genes in M1-specific TCRα repertoire. (e) Frequency of TRAJ gene usage among TRAV38-containing TCRαin the 6 donors. TRAJ52 is used almost exclusively. (f) Average frequency of TRAV38-containing TCRs for various CDR3α lengths. (g) Average frequencies of different CDR3α lengths within the overall CD8 T cell population recognizing HLA-A2/M1 are shown in bar graph at the left. TRAV gene usage for TCR with 15-mer CDR3α is shown for each donor in pie charts. Major TRAV genes (Vα) from each donor are labeled below charts. Numbers in pie charts represent percentage of TRAV38 in total M1-specific TCRs with 15-mer CDR3α. Amino acid compositions of 15-mer CDR3α in TRAV38-containing TCR from six donors were analyzed and depicted as sequence logo above pie charts. Solid lines above logo indicate germline-encoded sequences. Dotted lines show partially conserved sequence. (h) Average frequencies of different CDR3β lengths within the overall CD8 T cell population recognizing HLA-A2/M1 are shown in bar graph at the left, with TRBV gene usage and sequence logos for TCR with 10-mer and 11-mer CDR3β. The number in parentheses represents the individual's frequency of either 10-mer or 11-mer clonotypes. In panels (b–h), TRAV and TRBV genes are abbreviated as Vα and Vβ. In panel (f–h) the error bars represent standard deviation of the mean for six donors.

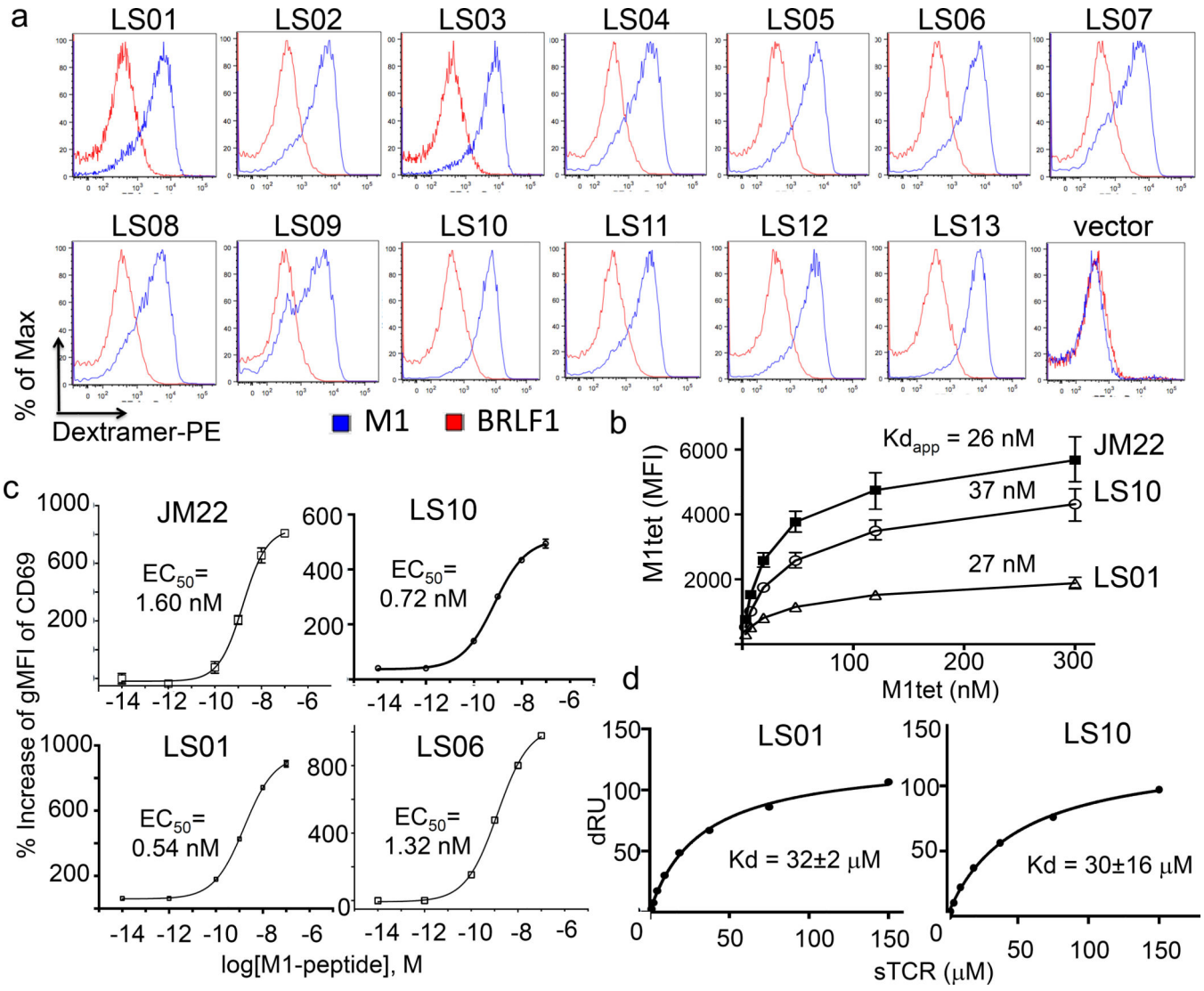

Figure 2. TCRαβ pairs bind HLA-A2/M1 and stimulate T cell signaling

(a) TCRα/β deficient Jurkat cells (J76-CD8) transiently expressing each of 13 paired full-length TCRα/TCRβ (LS01 – LS13) stained with M1-HLA-A2 dextramer or negative control BRLF1-HLA-A2 dextramer show specific binding by cloned TCRαβ. TCR surface expression levels for these transfectants are shown in Supplementary Fig. 4a, and T cell activation levels induced by peptide-pulsed HLA-A2+ presenting cells are shown in Supplementary Figure 4b. (b) Representative TCRs (Group I: JM22, Group II: LS10, Group III: LS01) show dose-dependent HLA-A2/M1 tetramer binding. J76-CD8 stably expressing LS01, LS10 or JM22 were stained with increasing concentrations of M1-tetramer. Plot of geometric mean of fluorescent intensities of bound HLA-A2/M1 tetramer subtracted from empty vector controls against increasing M1-tet concentration is shown, with half maximal binding concentrations (Kd_{app}) indicated. Error bars represent range of two independent duplicate experiments. TCR surface expression levels are shown in Supplementary Fig. 4d, and FACS plots in Supplementary Fig 4c. (c) LS01 and LS10 recognized M1-peptide as efficiently as canonical TCRs (JM22 and LS06). J76-CD8 cells expressing TCRs were

stimulated by T2 cells loaded with increasing M1 peptide, and surface expression of the activation marker CD69 was measured. Half-maximal stimulating concentrations (EC50) are shown. Error bars represent mean of triplicate measurements. d) Soluble LS01 and LS10 TCR proteins bind to immobilized HLA-A2/M1. Increasing concentrations of soluble LS01 and LS10 proteins were flowed over immobilized HLA-A2/M1 in surface plasmon resonance experiments. Increased response units relative to control channel (dRU) are plotted against soluble TCR concentration. Equilibrium binding constants (Kd) from fit to a single-site binding equation are shown with standard deviation of three independent experiments. Measured Kd for LS06 was 2.1±0.2 μM.

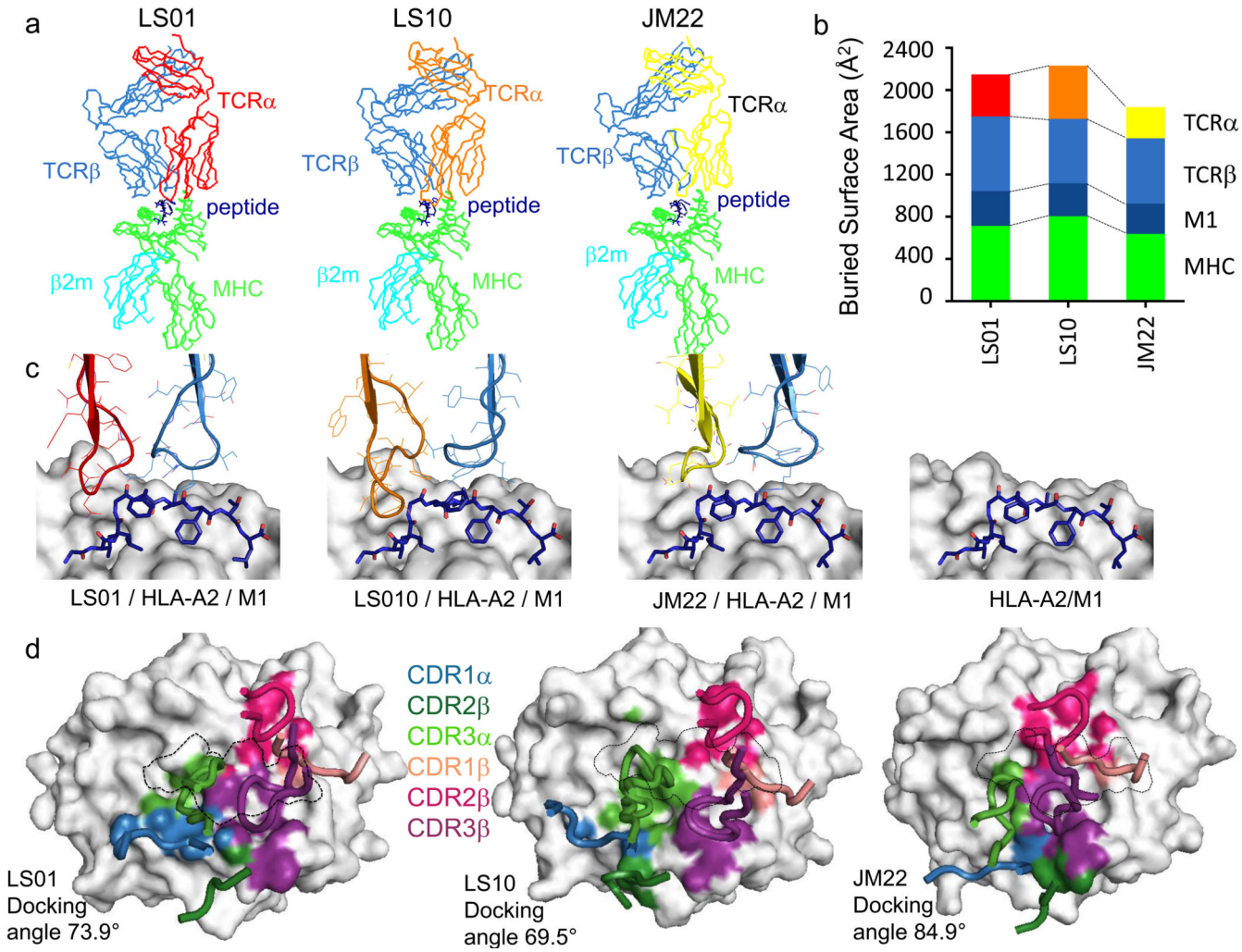

Figure 3. Structural Comparison of three TCRs docking onto HLA-A2/M1

(a) Ribbon diagrams of three TCR binding to HLA-A2/M1. The common TRBV19 TCRβ chain (blue) paired with different TCRα chains (red, orange, yellow) and bound to M1 peptide (dark blue) presented by HLA-A2 heavy chain (green) and beta-2 globulin (β2m) (b) BSAs at the interface between TCRs and HLA-A2/M1 are plotted with contribution of TCR α and β chains, peptide, and MHC colored as in panel a. Total buried surface area was 2137, 2231 and 1838 Å^2 respectively for LS01, LS10, and JM22 (c) CDR3α and CDR3β loops orient over the M1 peptide with different interactions. Unliganded HLA-A2/M1 (PDB 2VLL) is shown in the far right. (d) Surface representations of HLA-A2 are shown with footprints of TCRs colored by CDR loop. Locations of M1 peptide is outlined with dotted lines.

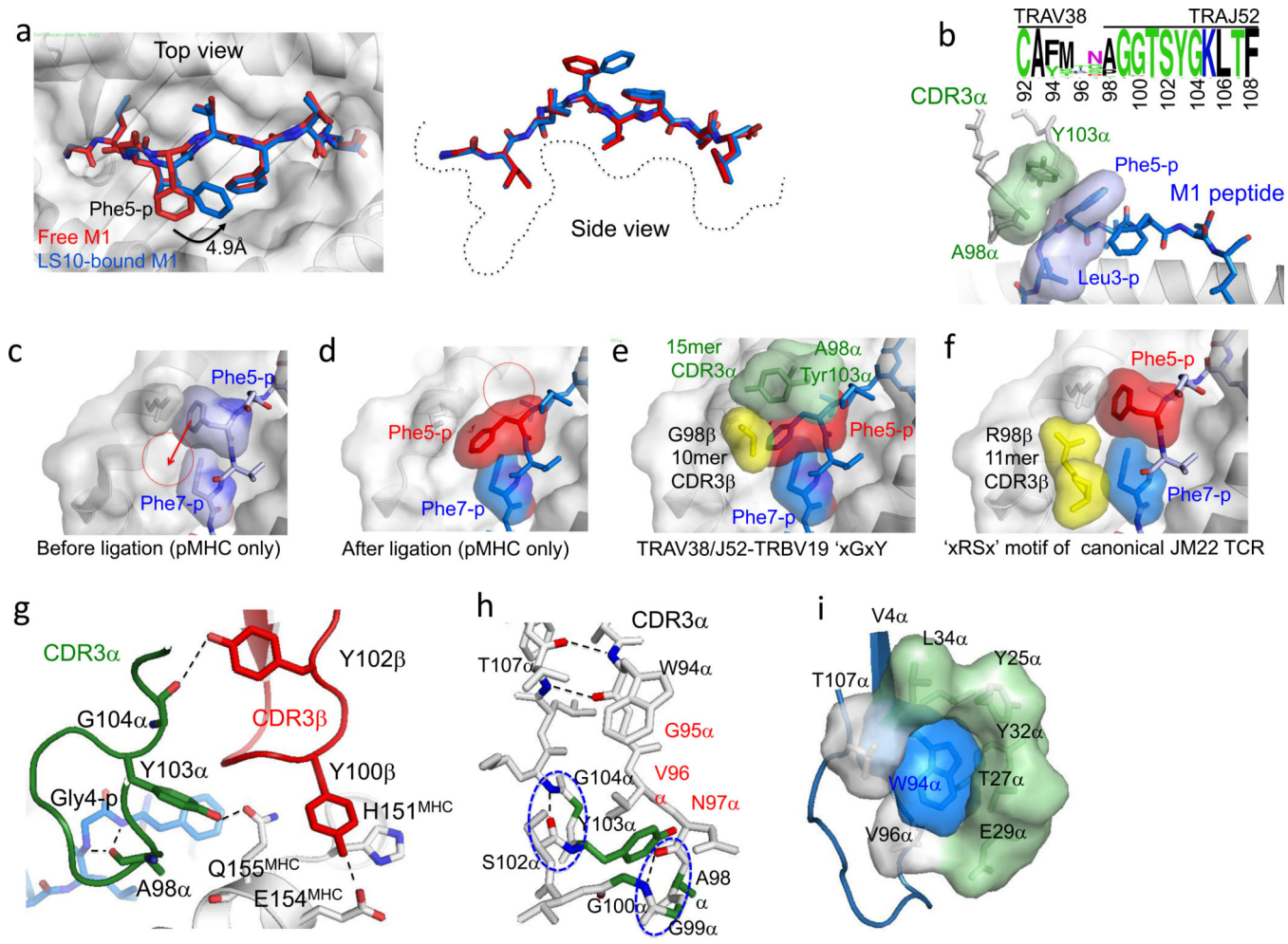

Figure 4. LS10 TCR uses conserved 15-mer CDR3α and 'xGxY' CDR3β motifs to select a M1 peptide conformation with Phe5 -p occupying the notch between peptide and MHC

(a) Top and side views of HLA-A2/M1 structures before (red) after (blue) LS10 ligation show that M1 undergoes significant movement upon TCR engagement (dotted line: HLA-A2 surface). (b) Tyr103α and Ala98α of CDR3α make close contacts with Phe5-p of M1 in the new conformation. (c–e) Rearrangement of M1 peptide upon LS10 interaction. Top views of HLAA2/ M1 are shown with MHC in grey and peptide colored before (c) and after (d–e) LS10-ligation. Notches between MHC and peptide (dashed circles) are filled with displaced Phe7-p (d) and CDR3α residues. (e) Close packing among Tyr103α (CDR3α), Gly98β (CDR3β) and Phe5-p (M1) is shown. (f) Similar view as (e) but after JM22 binding. (g) CDR3β (red) and CDR3α (green) loops of LS10 are shown near α2-helix (white) and M1-peptide (blue). Hydrogen bonds shown as dotted lines. (h) CDR3α of ligated LS10 adopts a structured configuration with two β-hairpins (dashed lines). Trp94α and Thr107α make two hydrogen bonds (dotted lines). Non-conserved residues (xxx in CAΦxxxAGGTSYGKLTF) are labeled red. (i) Trp94α (blue) of CDR3α is surrounded by TRAV38-specific residues (light green) and CDR3α (grey).

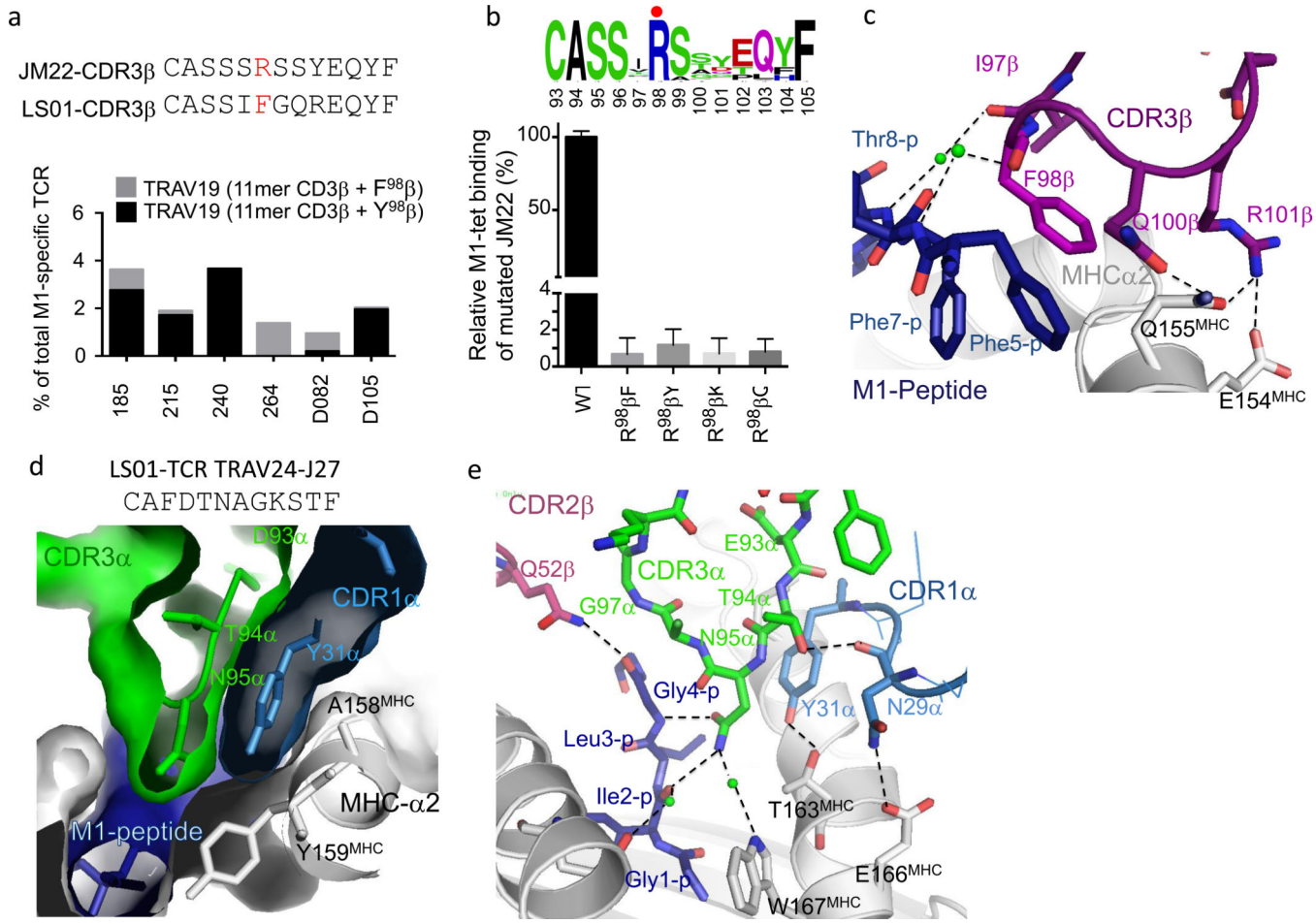

Figure 5. LS01 TCR uses CDR3β Phe98 to occupy the notch between peptide and MHC with additional interactions from CDR1α, CDR3α, and CDR3β

(a) CDR3β sequence comparison of LS01 and JM22. LS01 has Phe98β instead of conserved Arg98β Frequencies of M1-specific TCRs with Phe and Tyr in total M1-specific TCRs with 11mer CDR3β. (b) Substitution of Arg98β of xRSx motif abolishes HLA-A2/M1 tetramer binding. Relative MFI of HLA-A2/M1 tetramer bound to JM22 variants is shown. (c) CDR3β of LS01 (purple) with nearby portions of M1 (blue) and MHC α2-helix (grey). Water molecules are shown in green with hydrogen bonds indicated by dashed lines. (d) CDR1α (blue) and CDR3α (green) interactions in the LS01/HLA-A2/M1 interface. Tyr31α of CDR1α is inserted between CDR3α and MHCα2 helix interacting with Asn95α, Thr94α, Glu93α, Ala158MHC, and Tyr159MHC mainly via van der Waals interaction. (e) Tyr31α and Asn95α are involved in a network of hydrogen bonds with HLA-A2/M1. Error bars in (b) represent range of two independent duplicate samples.

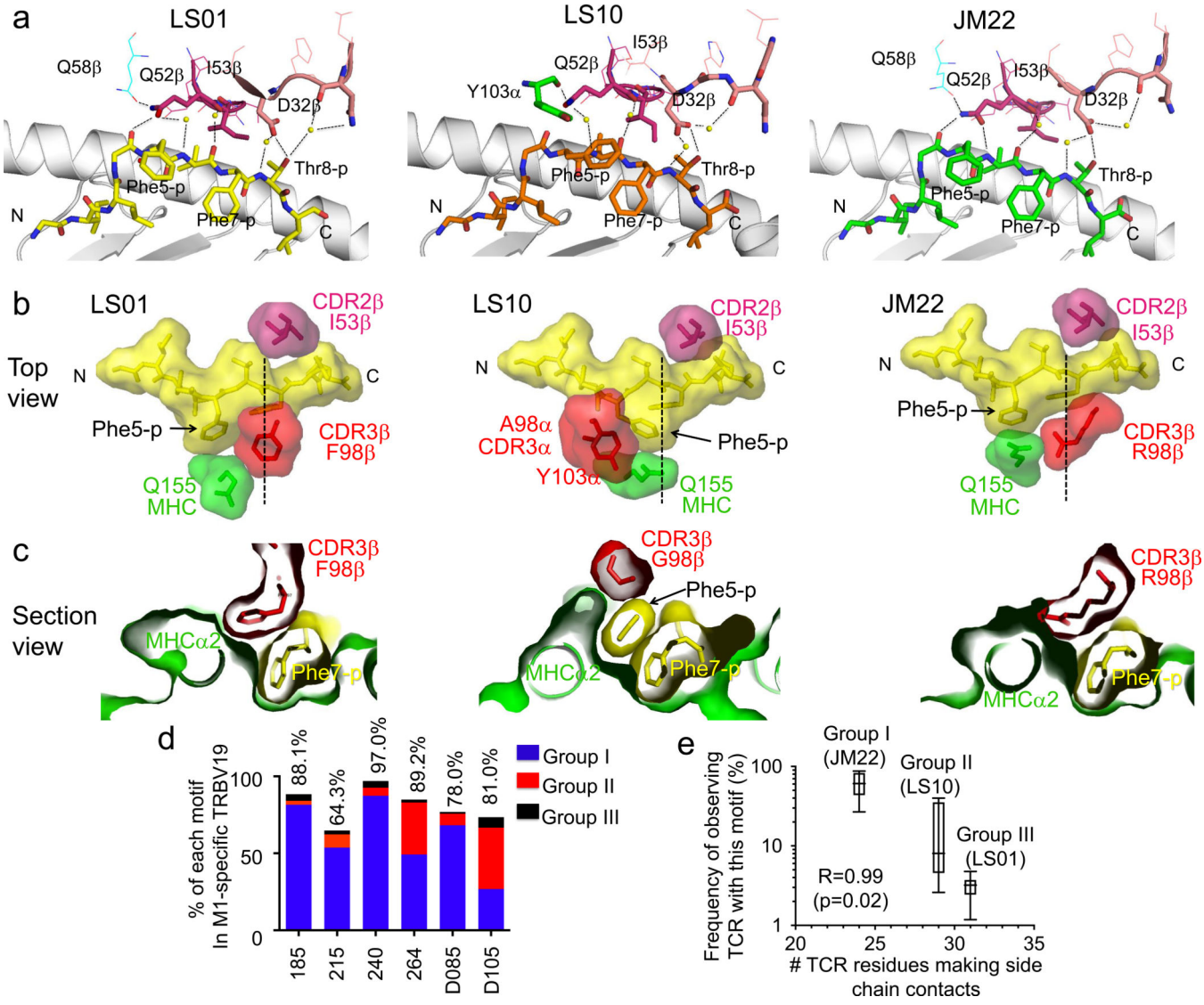

Figure 6. Different structural solutions to high-avidity binding of a featureless peptide

(a) Identical CDR1β and CDR2β sequences from three TRBV19-containing TCR make different interaction with HLA-A2/M1. (b) Top views of three HLA-A2/M1/TCR complexes show overall similarity and fine specificity of M1 recognition. TCR-ligated M1 peptide residues (yellow), Ile53β of CDR2β (pink), critical residues from CDR3α or CDR3β (red), and Gln155MHC of MHC (green) are displayed as surface/stick representations. (c) Sectional views of the three TCR in the pocket region. Sectional lines are shown by dashes in panel (b). (d) Percentage of TRBV19 TCR for each donor with motif from group I, II, or III. Numbers above bars indicate total for three groups. (g) Frequency of group I, II, and III TCRs plotted against numbers of TCR residues making side chain contacts with peptide-MHC in corresponding crystal structure. Box and whisker plot represent mean, quartile and range of frequencies for five donors and R is Pearson correlation coefficient.

Table 1

Data collection and refinement statistics (molecular replacement)

	LS01-HLA-A*02/M1	LS01-HLA-A*02/M1
Data collection		
Space group	P 1 21 1	P 1
Cell dimensions		
a, *b*, *c* (Å)	63.81, 75.46, 121.20	49.71, 101.84, 113.55
α, β, γ (°)	90, 98.01, 90	99.38, 92.60, 103.45
Resolution (Å)	29.14-2.06 (2.13-2.06)[a]	29.05-2.46 (2.55-2.46)
R_{merge}	0.176 (0.507)	0.116 (0.852)
$I / \sigma I$	5.15 (2.64)	8.14 (1.50)
$CC_{1/2}$	0.99 (0.75)	0.99 (0.66)
Completeness (%)	0.80 (0.84)	0.94 (0.67)
Redundancy	5.7 (5.5)	3.9 (3.7)
Refinement		
Resolution (Å)	29.14-2.06	29.05-2.46
No. reflections	56793	72567
R_{work} / R_{free}	0.176/0.201	0.202/0.222
No. atoms (non-H)		
Protein	6623	13293
Ligand/ion	18	40
Water	594	78
B factors		
Protein	38.9	67.5
Ligand/ion	19.5	35.7
Water	37.5	46.8
r.m.s. deviations		
Bond lengths (Å)	0.007	0.005
Bond angles (°)	0.84	0.88

[a]Values in parentheses are for highest-resolution shell.