Peptide length determines the outcome of TCR/peptide-MHC I engagement

Peptide length determines the outcome of TCR/peptide-MHCI engagement

*Julia Ekeruche-Makinde,1 *John J. Miles,1,2 *Hugo A. van den Berg,3 Ania Skowera,4,5 David K. Cole,1 Garry Dolton,1 Andrea J. A. Schauenburg,1 Mai Ping Tan,1 Johanne M. Pentier,1 Sian Llewellyn-Lacey,1 Kim M. Miles,1 Anna M. Bulek,1 Mathew Clement,1 Tamsin Williams,1 Andrew Trimby,1 Mick Bailey,6 Pierre Rizkallah,1 Jamie Rossjohn,1,7 Mark Peakman,4,5 David A. Price,1 *Scott R. Burrows,2 *Andrew K. Sewell,1 and *Linda Wooldridge1

1Institute of Infection and Immunity, Cardiff University School of Medicine, Cardiff, United Kingdom; 2Cellular Immunology Laboratory, Queensland Institute of Medical Research, Brisbane, Australia; 3Mathematics Institute, University of Warwick, Coventry, United Kingdom; 4Department of Immunobiology, King’s College London, London, United Kingdom; 5National Institute for Health Research Biomedical Research Centre at Guy’s & St Thomas’ National Health Service Foundation Trust and King’s College London, London, United Kingdom; 6Division of Veterinary Pathology, Infection and Immunity, School of Clinical Veterinary Science, University of Bristol, Bristol, United Kingdom; and 7Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Australia

**Key Points**

- MHCI-restricted TCRs exhibit an explicit preference for a single MHCI-peptide length.
- Effective CD8+ T-cell immunity can only be achieved by length-matched Ag-specific T-cell clonotypes.

Introduction

MHC class I (MHCI) molecules present short peptide fragments at the cell surface, enabling inspection of the intracellular proteome by CD8+ T cells. This elegant display system allows clearance of cells that exhibit abnormal biochemistry or express non-self, pathogen-derived proteins. The MHCI complex contains an Ag-binding cleft specifically designed to cradle a single peptide fragment. CD8+ T cells engage pMHC molecules via their clonotypic αβ-TCRs, which are encoded by variable (V), diversity (D), joining (J), and constant (C) gene fragments. These fragments recombine, with additional N-nucleotide insertions/deletions at V(D)J junctions, to produce a theoretical repertoire of 1015-1020 unique receptors.1-3 As a consequence of size limitations and thymic selection, the in vivo TCR repertoire has been estimated at 2.5 × 107 receptors.4,5 The focal point of TCR diversity is manifested in the Ag-binding site, which comprises 6 highly flexible complementarity determining regions (CDRs).6

TCRs are inherently cross-reactive toward pMHC molecules. Indeed, a recent comprehensive analysis of TCR cross-reactivity demonstrated that a single TCR is capable of recognizing more than one million different peptides of defined length presented in the context of a single MHCI restriction element.7 This feature explains how the naive TCR repertoire, which is dwarfed by the potential number of pMHC molecules that could be encountered in nature, can successfully provide broad antigenic coverage. TCR cross-reactivity is essential for normal processes such as positive selection in the thymus and survival of naive T cells in the periphery. A consequence of this cross-reactivity, however, is a high frequency of T cells capable of recognizing distinct peptides in the context of non-self MHC, a phenomenon known as alloreactivity.8 In addition, several studies have demonstrated the induction of autoimmunity with exogenous microbial peptides9 and, more recently, the ability of a single TCR to recognize both pMHC and pMHCII Ags.10-12 The huge scope for CDR loop flexibility on binding to pMHC13,14 and the ability of the TCR to force changes in the shape of the peptide15 provide a structural explanation for the ability of a single TCR to recognize a multitude of different pMHC shapes.15-17


© 2013 by The American Society of Hematology
Initially, it was thought that MHCI presentation was restricted to peptides 8-10 amino acids in length, but it is now established that peptides > 10 amino acids in length play a significant role in immune surveillance by CD8+ T-cells. As such, the TCR repertoire must be able to recognize all possible MHCI-associated peptide lengths (8-14 amino acids) to achieve effective immunity. Given the highly cross-reactive nature of TCR recognition, as well as the inherent bias that TCRs are considered to have for the MHC, it seemed reasonable to assume that TCRs can adapt to and recognize peptides of a different length bound to the same allelotype. To address this assumption, we synthesized 6 combinatorial peptide libraries (CPLs) ranging from 8-13 amino acid residues in length and measured TCR recognition across an array of human MHCI-restricted CD8+ T-cell clones. These αβ-T-cell clones were raised against a range of pMHCI Ags derived from both foreign and self-Ags. Unexpectedly, we found that strict limits exist to TCR plasticity in that each receptor is predisposed to engage peptides of a defined length. Thus, while any given TCR can be capable of recognizing tens of thousands, if not millions, of different peptides, these peptides must be of an explicit length that varies depending on the individual TCR.

### Methods

**Cells**

The following HLA A*0201-restricted CD8+ T-cell clones were used in this study: 1E6, specific for the preproinsulin-derived epitope ALWGPDPAAA (residues 15-24)\(^{21}\); ILA1, specific for the human telomerase reverse transcriptase (hTERT)-derived epitope ILAKFLHVL (residues 540-548)\(^{20}\); and MEL5, specific for the Melan-A–derived epitope ELAGIGILTV (residues 26-35)\(^{21}\). In addition, the following HLA A*0201-restricted CD8+ T-cell clones were used: SB14, specific for the EBV-derived EBNA-1 epitope HPVGEADYFEY (residues 407-417)\(^{22}\); SB10, specific for the HCMV-derived pp65 epitope CPSQEPMSIYV (residues 103-114)\(^{23}\); and SB27, specific for the EBV-derived BZLF1 epitope LPEPLPQQGQTAY (residues 52-64).\(^{24}\) The reaction parameters for these CD8+ T-cell clones are detailed in Table 1. CIR-HLA A*0201 target cells and T2-HLA B*3508 target cells were generated as described previously.\(^{25}\)

**Sizing scan**

For T2 cell-based pMHCI-binding assays, the T-lymphoblastoid hybrid cell line 0.174 × CEM.T2 was used. These cells, referred to as T2 cells, lack the transporter associated with Ag processing (TAP); therefore, the addition of exogenous binding peptide is required for stable cell-surface expression of MHCI. A total of 0.5 × 10⁶ T2 cells were incubated in RPMI 1640 medium (Life Technologies) with various concentrations of the indicated peptide or peptide mixture for 14-16 hours at 26°C, then for 2 hours at 37°C, before staining for HLA A*0201 surface expression with the PE-labeled mAb BB7.2 (BD Biosciences). Duplicate samples for each condition were acquired using a FACSCanto II flow cytometer (BD Biosciences). Data were analyzed with FlowJo Version 9.5.3 software (Treestar).

#### Table 1. Human CD8+ T-cell clones used in this study

<table>
<thead>
<tr>
<th>Clone</th>
<th>Specificity</th>
<th>HLA restriction</th>
<th>Minimal epitope</th>
<th>Length</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILA1</td>
<td>Telomerase</td>
<td>A*0201</td>
<td>ILAKFLHVL</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>MEL5</td>
<td>Melan-A</td>
<td>A*0201</td>
<td>ELAGIGILTV</td>
<td>10</td>
<td>21</td>
</tr>
<tr>
<td>1E6</td>
<td>Preproinsulin</td>
<td>A*0201</td>
<td>ALWGPDPAAA</td>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>SB14</td>
<td>EBV</td>
<td>B*3508</td>
<td>HPVGEADYFEY</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>SB10</td>
<td>HCMV</td>
<td>B*3508</td>
<td>CPSQEPMSIYV</td>
<td>12</td>
<td>23</td>
</tr>
<tr>
<td>SB27</td>
<td>EBV</td>
<td>B*3508</td>
<td>LPEPLPQQGQTAY</td>
<td>13</td>
<td>24</td>
</tr>
</tbody>
</table>

**Peptide titration assays: MIP1β ELISA and quantification of T-cell cross-reactivity**

Six × 10⁴ target cells were pulsed with peptide at the indicated concentrations for 2 hours at 37°C. Three × 10⁴ clonal CD8+ T cells were added and the assay was incubated overnight at 37°C. Subsequently, the supernatant was assayed for MIP1β by ELISA (R&D Systems). Functional sensitivity is expressed by the pEC\(_{50}\) for each peptide with respect to the TCR. This is defined as minus 1 times the base-10 logarithm (\(\log_{10}\)) of the 50% efficacy concentration (EC\(_{50}\); a greater functional sensitivity is indicated by a larger pEC\(_{50}\) value, which was estimated as described previously.\(^{2}\) T-cell cross-reactivity was quantified by CPL-based importance sampling as described previously.\(^{2}\)

**pMHCI staining and flow cytometry**

Soluble biotinylated pMHCI monomers were produced as described previously.\(^{21}\) Tetrameric pMHCI reagents (tetramers) were constructed by the addition of PE-conjugated streptavidin (Life Technologies) at a pMHCI:streptavidin molar ratio of 4:1. A total of 5 × 10⁴ CLIP tetramers, PE-conjugated streptavidin (Life Technologies) at a pMHCI:streptavidin molar ratio of 4:1. A total of 5 × 10⁴ CLIP tetramers, were incubated overnight at 37°C. Subsequently, the supernatant was assayed for MIP1β by ELISA (R&D Systems). Functional sensitivity is expressed in positional scanning format (Pepscan). CPL parameters are detailed in Table 3. For CPL screening, CD8+ T-cell clones were washed and rested overnight in RPMI 1640 medium containing 100 U/mL of penicillin, 100 μg/mL of streptomycin, 2 mM-L-glutamine, and 2% heat-inactivated FCS (Life Technologies). In 96-well, U-bottom plates, 6 × 10⁴ target cells were incubated with various library mixtures (at 100μg/mL) in duplicate for 2 hours at 37°C. After peptide pulsing, 3 × 10⁴ clonal CD8+ T cells were added and the assay was incubated overnight at 37°C. The supernatant was assayed by MIP1β ELISA (R&D Systems).

**Crystallization, structure determination, and refinement**

HLA A*0201 in complex with ILAFLHLRL (A2-ILA) was refolded as described previously\(^{21}\) and then concentrated to 10 mg/mL in 10mM MES containing 100 U/mL of penicillin, 100 μg/mL of streptomycin, 2 mM-L-glutamine, and 2% heat-inactivated FCS (Life Technologies). In 96-well, U-bottom plates, 6 × 10⁴ target cells were incubated with various library mixtures (at 100μg/mL) in duplicate for 2 hours at 37°C. After peptide pulsing, 3 × 10⁴ clonal CD8+ T cells were added and the assay was incubated overnight at 37°C. The supernatant was assayed by MIP1β ELISA (R&D Systems).
**Results**

**MHCI-restricted T-cell clones exhibit explicit peptide length recognition footprints**

To determine the extent of peptide length cross-reactivity displayed by TCRs, we examined the functional recognition of a custom-built “sizing scan” comprising random peptide libraries of different lengths (X8, X9, X10, X11, X12, and X13, where X is any of the 19 proteogenic L-amino acids excluding cysteine; Table 2) by a panel of CD8+ T-cell clones raised against peptides ranging from 9-13 residues (Table 1). This spectrum of recognition specificities permitted a broad overview of TCR cross-reactivity. The sizing scan suggested that each CD8+ T-cell clone exhibited a preference for MHCI-peptide length (Figure 1 and supplemental Figure 1, available on the Blood Web site; see the Supplemental Materials link at the top of the online article). This MHCI-peptide length preference was not random, but reflected the length of the peptide sequence against which each clone had been raised. Accordingly, the ILA1, 1E6, SB14, SB10, or SB27 CD8+ T cells (Table 1) were pulsed in duplicate with the following “sizing scan” mixtures (1mM) for 2 hours at 37°C: X8, X9, X10, X11, X12, and X13 (where X is any of the 19 proteogenic L-amino acids excluding cysteine). Subsequently, 3 × 10^4 ILA1, 1E6, SB14, SB10, or SB27 CD8+ T cells (Table 1) were added and incubated overnight. The supernatant was then harvested and assayed for MIP1α by ELISA. Values are expressed as a percentage of the highest response, where black represents 100% of the highest response and white represents 0% of the highest response. The raw data used to construct this figure can be viewed in supplemental Figure 1.

Although the sizing scans indicated that individual TCRs can exhibit an MHCI-peptide length preference, more detailed data can be obtained by performing positional scanning CPL scans across multiple peptide lengths. Therefore, we examined MHCI-peptide length preference by comparing the functional recognition of 8-13mer CPLs by 4 different CD8+ T-cell clones (ILA1, 1E6, MEL5, and SB27; Table 3). Figure 2A through D show, in heat map form, the detailed TCR recognition footprints of all 4 CD8+ T-cell clones (raw data are shown in supplemental Figures 2-5). ILA1, 1E6, SB14, SB10, and SB27 clones demonstrated a recognition preference for 9mer, 10mer, 11mer, 12mer, and 13mer peptides (Figure 1), respectively.

<table>
<thead>
<tr>
<th>CPT ID</th>
<th>Sequence of position one sublibrary</th>
<th>Total no. of peptides in library (a + 19) × 19</th>
<th>No. of sublibraries</th>
<th>No. of peptides in each sublibrary (19)</th>
<th>Concentration of each peptide in sublibrary</th>
</tr>
</thead>
<tbody>
<tr>
<td>8mer</td>
<td>OXXXXXXX</td>
<td>2.4 × 10^{10}</td>
<td>160</td>
<td>8.9 × 10^8</td>
<td>1.1 × 10^{-13} M</td>
</tr>
<tr>
<td>9mer</td>
<td>OXXXXXXX</td>
<td>4.8 × 10^{11}</td>
<td>180</td>
<td>1.7 × 10^{10}</td>
<td>5.9 × 10^{-13} M</td>
</tr>
<tr>
<td>10mer</td>
<td>OXXXXXXX</td>
<td>9.4 × 10^{12}</td>
<td>200</td>
<td>3.2 × 10^{11}</td>
<td>3.1 × 10^{-13} M</td>
</tr>
<tr>
<td>11mer</td>
<td>OXXXXXXX</td>
<td>1.8 × 10^{13}</td>
<td>220</td>
<td>6.1 × 10^{12}</td>
<td>1.6 × 10^{-13} M</td>
</tr>
<tr>
<td>12mer</td>
<td>OXXXXXXX</td>
<td>3.6 × 10^{14}</td>
<td>240</td>
<td>1.2 × 10^{14}</td>
<td>8.6 × 10^{-13} M</td>
</tr>
<tr>
<td>13mer</td>
<td>OXXXXXXX</td>
<td>7.1 × 10^{15}</td>
<td>260</td>
<td>2.2 × 10^{15}</td>
<td>4.5 × 10^{-23} M</td>
</tr>
</tbody>
</table>

n indicates the number of degenerate positions; O, fixed sequence position (1 of the 20 proteogenic L-amino acids; O is moved systematically through the peptide backbone in a full CPT); X, degenerate position (1 of 19 proteogenic L-amino acids, excluding cysteine); and a, full peptide length.

*When mixtures are used at a concentration of 100μM.
MEL5, and SB27 exhibited a large number of responses to peptide mixtures from the corresponding length-matched CPLs, in accordance with the results obtained from the sizing scan analysis (Figure 1). In stark contrast, CPL scans of other lengths elicited a very low number of positive responses, with a large number of mixtures failing to elicit outputs above background levels (Figure 2A-D). Some functional responses could be detected at the nonpreferred peptide length for all CD8\(^+\)/H11001 T-cell clones, although it was not possible to identify responses at all peptide positions at the nonpreferred lengths (with the possible exception of the 8mer CPL scan on MEL5; Figure 2C). Therefore, detailed CPL scans of multiple CD8\(^+\)/H11001 T-cell clones suggest that every TCR is characterized by a defined MHCI-peptide length recognition footprint.

MHCI-peptide length preference is not determined by MHCI binding

HLA A24, B7, and B35 molecules are capable of binding peptides of different lengths (8-14 amino acids).\(^{29,30}\) However, limited data are available for the peptide length binding specificity of HLA A*0201 and this has not been assessed previously using unbiased peptide mixtures. Therefore, we examined the ability of HLA A*0201 to bind sizing scan mixtures of different lengths (Table 2). To accomplish this, we measured the peptide-dependent stability of HLA A*0201 on the surface of T2 cells by flow cytometry. These experiments demonstrated that HLA A*0201 is capable of binding to all sizing scan mixtures that span 8-13 amino acids in length (Figure 3 and supplemental Figure 6). Importantly, under conditions that were identical to those used in the sizing scan assays (Figure 1 and supplemental Figure 1), sizing scan mixtures of all lengths bound efficiently to cell-surface HLA A*0201 (Figure 3). In addition, levels of binding for all peptide lengths were similar to those observed for the immunodominant influenza A matrix protein (M1)–derived epitope GILGFVFTL (residues 58-66; Figure 3), which is known to bind strongly to HLA A*0201.\(^{31,32}\) These data suggest that the peptide length preference exhibited by each individual TCR is not determined by differential MHCI-peptide binding. Instead, the peptide length preference of individual CD8\(^+\) T-cell clones appears to be determined strictly by the TCR.

![Figure 2](image-url)
Peptides designed from CPL scan data confirm MHC-I-peptide length preference

Next, we designed peptide sequences based on the preferred amino acids at each peptide position as revealed by CPL scan analysis (Figure 2) for the CD8⁺ T-cell clones 1E6, MEL5, and SB27. Previously, we demonstrated that > 6 x 10⁶ decameric peptides can act as substantially more potent agonists for the 1E6 TCR compared with the index peptide.⁷ Remarkably, in contrast, activation of 1E6 CD8⁺ T cells with a 9mer CPL scan yielded a very low number of positive responses relative to the 10mer CPL scan (Figure 2B). To study this observation in more detail, we designed panels of 9mer and 10mer peptides from combinations of the amino acids that elicited the strongest responses at each position observed in the corresponding CPL scans. The 9mer peptides tested on 1E6 were not recognized (Figure 4A-B). In contrast, the majority of 10mer peptides were recognized efficiently, with 13 of 17 acting as more potent agonists for this clone compared with the index peptide (Figure 4C-D). These data demonstrate that the 1E6 TCR exhibits a stringent preference for 10mer peptides compared with peptide sequences that are 9 amino acids in length, thereby verifying the CPL scan data (Figure 2B).
The MEL5 CD8+ T-cell clone also exhibited a strong preference for 10mer peptides in the corresponding CPL scan (Figure 2C). In addition, the best-recognized CPL sublibrary did not always correspond to the amino acid residue present in the index ELA-GIGILTV (ELA) peptide (Figure 2C). Multiple substitutions at position one (P1) of the index ELA peptide produced more potent agonists (Figure 5A), which is consistent with the structure of the TCR/pMHC complex (supplemental Figure 7). Here, the CDR1 loop of the MEL5 TCR lies closely over the N-terminus of the ELA peptide, offering opportunities for additional atomic contacts. A strong preference for P3-Thr and P8-Ile was also observed (Figure 5B). Single, double, and triple substitutions in the ELA peptide yielded a set of very strong agonists, the majority of which were considerably more potent than the index ELA peptide (Figure 5A-B). These results corroborate the CPL scan data demonstrating that a large number of 10mer peptides can be recognized efficiently by the MEL5 TCR. In contrast, 9mer and 11mer sequences designed from the low responses observed in the corresponding CPL scans exhibited poor activation potency, with the exception of one 9mer peptide, FWLLPAWAL, which activated MEL5 CD8+ T cells with intermediate potency (Figure 5C). pMHC1 tetramer staining with peptide variants was consistent with the activation data (Figure 5D). Therefore, the MEL5 TCR preferentially engages 10mer peptides.

It is now widely accepted that longer peptides play an important role in MHC1-restricted immunity. The CD8+ T-cell clone SB27, which recognizes the EBV BZLF1-derived 13mer peptide LPEPLPQGQLTAY (LPE), epitomizes this phenomenon. Activation of SB27 with 13mer CPL screen mixtures yielded a very large number of responses, higher than any other CPL scan performed in the present study (Figure 2D). All 13mer peptides in an extensive panel designed from the corresponding scan data were recognized efficiently (Figure 6C-D). It was also clear from the CPL scan data that the SB27 TCR exhibited a strong preference for 13mer peptides (Figure 2D). We examined a panel of 9mer peptides designed from the corresponding scan data and tested them for recognition by SB27. As expected from preceding length recognition data, activation was poor and approximately 10 000-fold lower compared with the index 13mer peptide (Figure 6A-B). This result implies that TCRs specific for longer peptides are not capable of engaging shorter peptides efficiently. Immunity to longer peptides may therefore require mobilization of specific portions of the CD8+ T-cell compartment that exhibit appropriate length specificity.

### Peptide length has a profound effect on TCR degeneracy

We next examined the impact of peptide length on the level of TCR degeneracy using a previously verified method called CPL-based importance sampling, which incorporates a sampling bias toward strong agonists as indicated by CPL scan data and thus allows an estimation of the number of peptides of a defined length that can be recognized by a given TCR on the basis of a sample of manageable size (in this case, 30 peptides). To correct the observed distribution for this bias, the observations are weighted by the reciprocal of the original sampling probability, giving estimates of numbers of peptides of a defined length that can be recognized at varying functional sensitivities. Figures 2 and 4 show that the 1E6 TCR demonstrates a strict preference for 10mer peptides. Indeed, when we previously performed CPL-based importance sampling using this 10mer scan, we found a large number of robustly recognized peptides, with more than 25% being recognized at higher functional sensitivity than the index peptide sequence. Responses to scan mixtures at all other peptide lengths were either entirely absent or very poor (Figure 2). To characterize the diminished efficacy of the interaction with 11mers and 12mers, CPL-based importance sampling was performed. In stark contrast to previous results obtained using 10mer CPL scan data, sampled 11mer peptides were all recognized with > 1000-fold lower efficiency compared with the index peptide (Figure 7A and supplemental Table 1). Similarly, sampled 12mer peptides were also poorly recognized by the 1E6 TCR (Figure 7B and supplemental Table 1).

Whereas CPL-based importance sampling performed at the preferred peptide length yielded small peptide libraries that were highly enriched with agonists, the extent of enrichment observed at the nonpreferred peptide lengths was lower, possibly because CPL scans performed at these peptide lengths contain less information. Alternatively, the degree of enrichment may be of comparable magnitude in all cases, but the occurrence rate of strong agonists (determined by both TCR degeneracy and degree of enrichment) remains too low to render agonists readily detectable in the samples. Degeneracy curves for the 1E6 TCR are shown in Figure 7C; the abscissa is the functional sensitivity of peptides relative to the index peptide (ie, peptides with a value < 0, 0, or > 0 are either poor agonists compared with index, equivalent to index, or better agonists than index, respectively), whereas the ordinate axis is the number of peptides that exhibit a functional sensitivity of at least the abscissa. The degeneracy curves suggest that no 11mer or 12mer peptides were recognized with a functional sensitivity equivalent to index, indicating that peptide length has a major impact on the degeneracy profile of a TCR.

### Different length peptide agonists can occur but are often poorly recognized and distinct in sequence

We have shown that Ag recognition by MHC1-restricted TCRs is constrained by peptide length, and that efficient recognition of peptides with alternative lengths is less likely. Interestingly, we identified one 12mer peptide that could be recognized by 1E6, although the efficiency of recognition was approximately 10-fold lower compared with the index peptide (Figure 7B). The sequence of this 12mer peptide (MMLWLRWIDVPT) was entirely different from the index sequence (ALWGDPDAAA) for the 1E6 TCR. Furthermore, 2 × 8mer epitopes and 1 × 9mer epitope were recognized by the MEL5 TCR (Figure 5C-F). Again, the sequences of these peptides (WLLPAWGV, WLLPTWGV, and FWLLPAWAL) were entirely different from the index sequence for the MEL5 TCR (ELAGIGILTV). Therefore, although rare, recognition of different length peptides is possible. However, the eliciting peptide sequence can be distinct from that of the index peptide and recognition typically occurs with reduced levels of functional sensitivity.

### Peptide length has a major impact on the complexity of antigenic structure

To examine the underlying molecular mechanism of peptide length specificity exhibited by MHC1-restricted TCRs, we conducted a structural comparison of pMHC1 complexes recognized by the T cells used in this study. The atomic structure of the nonligated ELAGIGILTV (10mer), ALWGDPDAAA (10mer), HPVGEADY- FEY (11mer), CPSQEPSMSIYYV (12mer), and LPEPLPQGQLTAY (13mer) peptides, in complex with their corresponding MHC1s, have been solved previously. We also included previously solved structures of an 8mer and 14mer peptide in complex with MHC1 (Figure 8). To complete our structural comparison of peptides
Figure 5. The MEL5 TCR exhibits a strict preference for 10mer peptides. C1R-HLA A*0201 cells (6 × 10⁴) were pulsed with the indicated 10mer (A-B), 9mer (C), 11mer (D), or 8mer (E) peptides at the concentrations depicted for 2 hours at 37°C. Subsequently, 3 × 10⁴ MEL5 CD8⁺ T cells were added and incubated overnight. The supernatant was then harvested and assayed for MIP1β by ELISA. Error bars represent SDs. (D,F) MEL5 CD8⁺ T cells (5 × 10⁴) were incubated with PE-conjugated HLAA*0201 tetramer (25 μg/mL) folded around ELAGIGILTV (black), FWLLPAWAL (red), FWLLGAWAL (blue), FFAGGIGIRTI (cyan), FLAGGIGIRTI (green), WLLPAWG (yellow), or WLLPTWG (pink) for 15 minutes at 37°C and then stained with 5 μL of 7-aminoactinomycin D for 30 minutes at 4°C, washed twice, and resuspended in PBS. Negative control staining is shown in dark purple.
corresponding to the CD8+ T-cell clones included in this study, we solved the atomic structure of HLA A*0201-ILAKFLHRL (9mer), recognized by ILA1, to a resolution of 1.8Å (supplemental Table 2). The overall structure of HLAA*0201-ILAKFLHRL was similar to previously solved pMHCI structures, with the peptide anchored at position 2 and the peptide C-terminus, enabling the central residues to bulge out away from the groove for potential TCR interactions. The analysis demonstrated that the positions of the anchor residues for each of the different length peptides were superimposable (Figure 8). Therefore, the distance between peptide residue 2, and the C-terminus of every peptide investigated, irrespective of differences in length, was very similar (approximately 20Å). Consequently, the only option available to accommodate extra residues as MHCI-bound peptides increase in length was an incremental increase in: (1) the degree of backbone “bulging” (ie, the height of the central peptide bulge away from MHCI groove; Figure 8A), (2) sideways displacement of the peptide toward the MHCI α1 or α2 helices (Figure 8B), or (3) a combination of both. For example, the central residues in the LPEPLPQQLTAY 13mer peptide bulged out of the groove by approximately 20Å, compared with only approximately 8Å for the ILAKFLHRL (9mer) peptide. The increasingly complex antigenic shape of the longer peptides revealed by this structural meta-analysis may restrict some TCRs from mediating stabilizing interactions with the MHCI surface, as described previously.24 This notion is further supported by our observation that individual TCRs are hard-wired to recognize peptides of a specific length and presumably cannot adapt to large structural variations in peptide conformation. Therefore, antigenic complexity (degree of peptide backbone “bulging”/sideways displacement) is determined by peptide length (in addition to peptide sequence), which in turn determines the outcome of TCR/peptide-MHCI engagement.

Figure 6. Poor recognition of shorter peptides by the SB27 TCR. T2-HLA B*3508 cells (6 × 10⁴) were pulsed with the indicated 9mer (A-B) or 13mer (C-D) peptides at the concentrations depicted for 2 hours at 37°C. Subsequently, 3 × 10⁶ SB27 CD8+ T cells were added and incubated overnight. The supernatant was then harvested and assayed for MIP1β by ELISA. Error bars represent SDs.

PEPTIDE LENGTH SPECIFICITY OF HUMAN TCRs 1119
BLOOD, 14 FEBRUARY 2013 • VOLUME 121, NUMBER 7
From bloodjournal.hematologylibrary.org at Cardiff University on July 8, 2013. For personal use only.
Discussion

Individual αβ-TCRs can recognize vast numbers of peptides encapsulated within a singular length. This inherent degeneracy is essential for effective immune coverage against the enormous number of antigenic peptides that could be encountered. Given the remarkable extent of such length-confined TCR plasticity, it seemed likely that TCRs could also cross-react with peptides of different lengths in the context of the same, or even different, presenting MHCI molecules. Indeed, the natural killer T-cell (NKT) TCR can recognize lipid Ags with various length headgroups, namely galcer (1 sugar) and iGb3 (3 sugars). Such peptide length flexibility would further maximize the ability of the numerically limited TCR repertoire to provide broad antigenic coverage. Indeed, the Ag processing/presentation machinery naturally generates MHCI-binding peptides across a range of fragment lengths, typically 8-14 residues. However, the degree to which αβ-TCRs cross-recognize these fragment lengths is ill-defined. In the present study, we used an approach that allowed for a sweeping and inclusive scan of TCR specificity. Surprisingly, we discovered a clear limit to receptor plasticity in terms of a defined peptide length preference. This unexpected finding was shown to be an inherent property of all TCRs examined across numerous biologic specificities and restricting MHCI alleles.

The data presented herein demonstrate that TCR plasticity is delimited by peptide length. Therefore, individual TCRs are predisposed to engage peptides of a defined length and cross-reactivity across peptide length boundaries is uncommon. These findings uncover a novel and fundamental feature of MHCI-restricted T-cell immunity that has wide implications. In particular, they suggest that the CD8+ T-cell repertoire is compartmentalized.

Figure 7. Peptide length has a major impact on TCR degeneracy. (A) 1E6 CD8+ T-cell recognition of a set of 30 11mer peptides sampled by CPL-based importance. (B) 1E6 CD8+ T-cell recognition of a set of 30 12mer peptides sampled by CPL-based importance. Assays were conducted as described in Figure 3. (C) CPL-based importance sampling was used to construct a degeneracy curve that describes the number of 11mer and 12mer peptides that can be recognized by the 1E6 TCR. The ordinate is the estimated number of peptides with a functional sensitivity of at least the value indicated on the abscissa. Error bars represent SDs.
by length preference. Effective MHC-mediated immunity would therefore require the mobilization of length-matched Ag-specific CD8+ T-cell clonotypes from the peripheral repertoire. This is especially pertinent to the design of “superagonists” or heteroclitic variants for the purposes of peptide vaccination because our results established that the TCR repertoire is characterized by a high frequency of alloreactive T cells.8 This population represents a major barrier to transplantation and, as such, there is substantial pressure to elucidate the biochemical and structural basis of this phenomenon. Our data suggest that peptide length will be an important consideration in the identification of alloreactive ligands. Indeed, there is mounting evidence to show that alloreactive T-cell recognition is mediated in a peptide-dependent manner.17,18

Despite clear length preferences, TCRs can recognize peptides of nonpreferred lengths. However, this phenomenon is relatively uncommon and generally suboptimal in terms of CD8+ T-cell activation compared with peptides of preferred length. Length-disparate agonist peptides can be entirely distinct at the primary sequence level with different amino acid residues at every position. For example, the 1E6 TCR is specific for the preproinsulin-derived epitope ALWGDPAAAA (residues 15-24), but can recognize a 12mer peptide with almost no sequence homology (MMLLWRID-VPT). The 1E6 clone was generated from a patient with type 1 diabetes and is a well-documented example of an autoreactive CD8+ T-cell that can kill human pancreatic islet β cells.19,36 This is a particularly interesting result because it suggests that CD8+ T cells of pathogenic relevance in autoimmune disease can be activated by cross-reactive ligands that are not amenable to a priori prediction on the basis of peptide length or sequence.

To understand the molecular mechanism that underlies the MHC-peptide length preferences displayed by different TCRs, it is informative to consider the conformations adopted by MHC-bound peptides of different lengths. We show that incremental increases in peptide length result in increased antigenic complexity. Bound peptides of different lengths. We show that incremental increases in peptide length result in increased antigenic complexity. Peptide backbones are shown with ribbon representation and include the 8mer B*3501-VPLRPMTY (orange), 9mer A*0201-LAKFLHRL (red), 10mer A*0201-ELAGIGILTV (green), 10mer A*0201-ALWGDPAAAA (dark green), 11mer B*3508-HPVGEADYFYEY (blue), 12mer B*3508-CPSQEPMISYYV (grey), 13mer B*3508-LPEPLPQGQLTAY (yellow), and 14mer B*3501-LPAVGLSPGEGEY (purple). (A) Side view demonstrating the different peptide bulges out the MHC-I groove depending on the length of the peptide. The scale on the left was calculated using the conserved position of peptide residue 2 from the Cα atom as 0 Å. The MHC α2 helix has been removed for clarity. (B) “Bird’s eye view” demonstrating the different sideways displacement of peptides toward MHC-I α1 or α2 helices depending on the length of the peptide. The scale on the left was calculated using the average central position of all of the peptides as 0 Å. The MHC α1 (above) and α2 helix (below) are shown. PDB codes are 1A1N, 10GA, 2GT9, 2FZ3, 3BW9, 12HL, 1XH3, and 414W.

Figure 8. The degree of MHC-I-peptide backbone “bulging” is proportional to peptide length. Supposition of pMHC structures with increasing epitope length. The peptides have been superimposed on HLA A*0201, which is shown in cartoon representation. MHCIs were aligned using the MHC α1 and β2 domains. Peptide backbones are shown with ribbon representation and include the 8mer B*3501-VPLRPMTY (orange), 9mer A*0201-LAKFLHRL (red), 10mer A*0201-ELAGIGILTV (green), 10mer A*0201-ALWGDPAAAA (dark green), 11mer B*3508-HPVGEADYFYEY (blue), 12mer B*3508-CPSQEPMISYYV (grey), 13mer B*3508-LPEPLPQGQLTAY (yellow), and 14mer B*3501-LPAVGLSPGEGEY (purple). (A) Side view demonstrating the different peptide bulges out the MHC-I groove depending on the length of the peptide. The scale on the left was calculated using the conserved position of peptide residue 2 from the Cα atom as 0 Å. The MHC α2 helix has been removed for clarity. (B) “Bird’s eye view” demonstrating the different sideways displacement of peptides toward MHC-I α1 or α2 helices depending on the length of the peptide. The scale on the left was calculated using the average central position of all of the peptides as 0 Å. The MHC α1 (above) and α2 helix (below) are shown. PDB codes are 1A1N, 10GA, 2GT9, 2FZ3, 3BW9, 12HL, 1XH3, and 414W.

In addition, our findings have implications for T-cell selection, naïve T-cell survival in the periphery, and the identification of alloreactive ligands. TCR cross-reactivity is the molecular basis of a defined length. In terms of non-self pMHC molecules, it is established that the TCR repertoire is characterized by a high frequency of alloreactive T cells.8 This population represents a major barrier to transplantation and, as such, there is substantial pressure to elucidate the biochemical and structural basis of this phenomenon. Our data suggest that peptide length will be an important consideration in the identification of alloreactive ligands. Indeed, there is mounting evidence to show that alloreactive T-cell recognition is mediated in a peptide-dependent manner.17,18

Despite clear length preferences, TCRs can recognize peptides of nonpreferred lengths. However, this phenomenon is relatively uncommon and generally suboptimal in terms of CD8+ T-cell activation compared with peptides of preferred length. Length-disparate agonist peptides can be entirely distinct at the primary sequence level with different amino acid residues at every position. For example, the 1E6 TCR is specific for the preproinsulin-derived epitope ALWGDPAAAA (residues 15-24), but can recognize a 12mer peptide with almost no sequence homology (MMLLWRID-VPT). The 1E6 clone was generated from a patient with type 1 diabetes and is a well-documented example of an autoreactive CD8+ T-cell that can kill human pancreatic islet β cells.19,36 This is a particularly interesting result because it suggests that CD8+ T cells of pathogenic relevance in autoimmune disease can be activated by cross-reactive ligands that are not amenable to a priori prediction on the basis of peptide length or sequence.

To understand the molecular mechanism that underlies the MHC-I-peptide length preferences displayed by different TCRs, it is informative to consider the conformations adopted by MHC-bound peptides of different lengths. We show that incremental increases in peptide length result in increased antigenic complexity. Bound peptides of different lengths. We show that incremental increases in peptide length result in increased antigenic complexity. Peptide backbones are shown with ribbon representation and include the 8mer B*3501-VPLRPMTY (orange), 9mer A*0201-LAKFLHRL (red), 10mer A*0201-ELAGIGILTV (green), 10mer A*0201-ALWGDPAAAA (dark green), 11mer B*3508-HPVGEADYFYEY (blue), 12mer B*3508-CPSQEPMISYYV (grey), 13mer B*3508-LPEPLPQGQLTAY (yellow), and 14mer B*3501-LPAVGLSPGEGEY (purple). (A) Side view demonstrating the different peptide bulges out the MHC-I groove depending on the length of the peptide. The scale on the left was calculated using the conserved position of peptide residue 2 from the Cα atom as 0 Å. The MHC α2 helix has been removed for clarity. (B) “Bird’s eye view” demonstrating the different sideways displacement of peptides toward MHC-I α1 or α2 helices depending on the length of the peptide. The scale on the left was calculated using the average central position of all of the peptides as 0 Å. The MHC α1 (above) and α2 helix (below) are shown. PDB codes are 1A1N, 10GA, 2GT9, 2FZ3, 3BW9, 12HL, 1XH3, and 414W.

In addition, our findings have implications for T-cell selection, naïve T-cell survival in the periphery, and the identification of alloreactive ligands. TCR cross-reactivity is the molecular basis of a defined length. In terms of non-self pMHC molecules, it is established that the TCR repertoire is characterized by a high frequency of alloreactive T cells.8 This population represents a major barrier to transplantation and, as such, there is substantial pressure to elucidate the biochemical and structural basis of this phenomenon. Our data suggest that peptide length will be an important consideration in the identification of alloreactive ligands. Indeed, there is mounting evidence to show that alloreactive T-cell recognition is mediated in a peptide-dependent manner.17,18

Despite clear length preferences, TCRs can recognize peptides of nonpreferred lengths. However, this phenomenon is relatively uncommon and generally suboptimal in terms of CD8+ T-cell activation compared with peptides of preferred length. Length-disparate agonist peptides can be entirely distinct at the primary sequence level with different amino acid residues at every position. For example, the 1E6 TCR is specific for the preproinsulin-derived epitope ALWGDPAAAA (residues 15-24), but can recognize a 12mer peptide with almost no sequence homology (MMLLWRID-VPT). The 1E6 clone was generated from a patient with type 1 diabetes and is a well-documented example of an autoreactive CD8+ T-cell that can kill human pancreatic islet β cells.19,36 This is a particularly interesting result because it suggests that CD8+ T cells of pathogenic relevance in autoimmune disease can be activated by cross-reactive ligands that are not amenable to a priori prediction on the basis of peptide length or sequence.

In addition, our findings have implications for T-cell selection, naïve T-cell survival in the periphery, and the identification of alloreactive ligands. TCR cross-reactivity is the molecular basis of a defined length. In terms of non-self pMHC molecules, it is established that the TCR repertoire is characterized by a high frequency of alloreactive T cells.8 This population represents a major barrier to transplantation and, as such, there is substantial pressure to elucidate the biochemical and structural basis of this phenomenon. Our data suggest that peptide length will be an important consideration in the identification of alloreactive ligands. Indeed, there is mounting evidence to show that alloreactive T-cell recognition is mediated in a peptide-dependent manner.17,18
showed no correlation with peptide length specificity. Further studies are required to define a role for TCR usage as a determinant of MHCI-peptide length preference.

In summary, we have shown that MHCI-restricted T cells exhibit a preference for peptide length that is governed by the TCR and that effective recognition of pMHCI Ag can only be achieved by length-matched Ag-specific CD8+ T-cell clonotypes. This novel finding has broad implications for understanding how effective CD8+ T-cell immunity is achieved and the future design of peptide vaccination. In addition, we propose that every TCR is characterized by a unique “peptide-recognition signature” that is defined by: (1) a preference for peptide length, (2) the number of peptides that can be recognized, and (3) the amino acid sequence of these agonist peptides. The peptide recognition signature of an induced CD8+ T-cell response is likely to have a major impact on CD8+ T-cell immunity and could underlie the pathogenesis of many disease states.

Acknowledgments

This work was supported by the Biotechnology and Biological Sciences Research Council (grant BB/H001085/1) and The Wellcome Trust (grant WT079848MA). L.W. is a Wellcome Trust Clinical Intermediate Fellow, D.A.P. is a Medical Research Council Senior Clinical Fellow, and D.K.C. is a Wellcome Trust Career Development Fellow. J.J.M., J.R., and S.R.B. are funded by the National Health and Medical Research Council. J.R. is a National Health and Medical Research Council Australia Fellow. A.S., G.D., M.P., and D.A.P. are funded by the Juvenile Diabetes Research Foundation (grants 7-2005-877, 1-2007-1803, and 17-2009-806). J.E.-M., M.P.T., M.C., and T.W. are funded by The Wellcome Trust. Research by A.S. and M.P. is supported by the National Institute for Health Research Biomedical Research Center based at Guy’s and St Thomas’ National Health Service Foundation Trust and King’s College London. The views expressed are those of the author(s) and not necessarily those of the National Health Service, the National Institute for Health Research, or the Department of Health.

Authorship


Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Dr Linda Wooldridge, Institute of Infection and Immunity, Cardiff University, Henry Wellcome Building, Heath Park, Cardiff, CF14 4XN United Kingdom; e-mail: wooldrigel@cardiff.ac.uk.

References


