HLA-B15 Peptide Ligands Are Preferentially Anchored at Their C Termini

Kiley R. Prilliman, Kenneth W. Jackson, Mark Lindsey, Jihua Wang, David Crawford and William H. Hildebrand

*J Immunol* 1999; 162:7277-7284; ;
http://www.jimmunol.org/content/162/12/7277

---

**References**  This article cites 42 articles, 21 of which you can access for free at:
http://www.jimmunol.org/content/162/12/7277.full#ref-list-1

**Subscription**  Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

**Permissions**  Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**  Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Therapies to elicit protective CTL require the selection of pathogen- and tumor-derived peptide ligands for presentation by MHC class I molecules. Edman sequencing of class I peptide pools generates “motifs” that indicate that nonameric ligands bearing conserved position 2 (P2) and P9 anchors provide the optimal search parameters for selecting immunogenic epitopes. To determine how well a motif represents its individual constituents, we used a hollow-fiber peptide production scheme followed by the mapping of endogenously processed class I peptide ligands through reverse-phase HPLC and mass spectrometry. Systematically mapping and characterizing ligands from B*1508, B*1501, B*1503, and B*1510 demonstrate that the peptides bound by these B15 allotypes i) vary in length from 7 to 12 residues, and ii) are more conserved at their C termini than their N-proximal P2 anchors. Comparative peptide mapping of these B15 allotypes further pinpoints endogenously processed ligands that bind to the allotypes B*1508, B*1501, and B*1503, but not B*1510. Overlapping peptide ligands are successful in binding to B*1501, B*1503, and B*1508 because these B15 allotypes share identical C-terminal anchoring pockets whereas B*1510 is divergent in the C-terminal pocket. Therefore, endogenous peptide loading into the B15 allotypes requires that a conserved C terminus be anchored in the appropriate specificity pocket while N-proximal anchors are more flexible in their location and sequence. Queries for overlapping and allele-specific peptide ligands may thus be contingent on a conserved C-terminal anchor. The Journal of Immunology, 1999, 162: 7277–7284.
comprise. However, the extent of deviation around composite motifs remains unrealized, and with accumulating data demonstrating that class I peptide ligands need not fit predicted folded motifs, the search parameters for identifying putative class I ligands become ambiguous.

An ongoing effort in our laboratory is the systematic characterization of endogenously processed class I peptide ligands. Our approach is to map multiple ligands from divergent class I molecules, after which we directly compare the peptide maps (24). The overall goal of comparatively analyzing peptides from different MHC molecules is to prescribe the rules that allow or prohibit natural peptide binding across class I Ag binding polymorphisms. We have sequenced >200 endogenously processed peptide-antibody complexes from a series of four HLA-B15 alleles that differ sequentially in their Ag binding grooves. Analysis of these HLA-B15 ligands depicts the range of functional variability around a peptide motif and demonstrates the key role played by a ligand’s C terminus during endogenous peptide loading.

Materials and Methods

Transfectants, cell culture, and purification of soluble HLA (sHLA)

Transfectants were established and cultured as previously described with some modifications (24). Briefly, the cloned coding regions of the four alleles B*1501, B*1503, B*1508, and B*1510, which differ by up to 7 aa within their Ag binding grooves (Table 1), were truncated by RT-PCR using primers 5PX1 (5‘-GAGGAGCTATCCAGAATCTTCCCCAACAGCCCGAG-3‘) and 3PEI (5‘-CCCGGAATTTCTCATCTCGAGGTTAGG-3‘). This yielded the products that terminated at codon 300 (inserted stop shown in bold italic) and were flanked by 5‘ EcoRI restriction sites (underlined). The products were subcloned into the pcDNA3.1(−) expression vector (Invitrogen, Carlsbad, CA), verified by DNA sequencing, and electroporated into the class I-negative cell line 721.221 (25), which was then selected with 1.5 mg/ml puromycin (Mediatech, Herndon, VA). Transfectants were screened by ELISA using W6/32 as the primary Ab and an anti-β2m, HRP-conjugated secondary Ab (Dako, Carpinteria, CA), and positive clones were subcloned by limiting dilution. Positive wells were subcloned by limiting dilution. Positions of identity with the consensus (italics) drawn from B*1501 are indicated by dashes (−).

Table 1. Ag binding groove-localized amino acid substitutions that distinguish the B*1508, B*1501, B*1503, and B*1510 allotypes

<table>
<thead>
<tr>
<th>Residue Position</th>
<th>24</th>
<th>45</th>
<th>46</th>
<th>63</th>
<th>67</th>
<th>116</th>
<th>156</th>
</tr>
</thead>
<tbody>
<tr>
<td>consensus</td>
<td>A</td>
<td>M</td>
<td>A</td>
<td>E</td>
<td>S</td>
<td>S</td>
<td>W</td>
</tr>
<tr>
<td>B*1508</td>
<td>−</td>
<td>−</td>
<td>A</td>
<td>E</td>
<td>S</td>
<td>−</td>
<td>L</td>
</tr>
<tr>
<td>B*1501</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>B*1503</td>
<td>S</td>
<td>E</td>
<td>E</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>L</td>
</tr>
<tr>
<td>B*1510</td>
<td>S</td>
<td>E</td>
<td>E</td>
<td>N</td>
<td>C</td>
<td>Y</td>
<td>L</td>
</tr>
</tbody>
</table>

* Residues are numbered from the N terminus of the mature class I heavy chains. Positions of identity with the consensus (italics) drawn from B*1501 are indicated by dashes (−).

Isolation, reverse-phase HPLC (RP-HPLC) separation, and Edman sequencing of peptides

Peptides were extracted from eluted sHLA by pooling the fractions of UV absorbance at 280 nm (A280), adding glyclal acetic acid to 10%, and heating for 10 min at 70°C. Upon cooling, peptides were separated from heavy chains, β2m, and BSA through a stirred cell with a 3-kDa YM membrane (Amicon, Beverly, MA). Isolated peptides were purified of salts and free amino acids by a step RP-HPLC gradient of 0–100% acetonitrile with trifluoroacetic acid (TFA) (0.06%) in 1 min, held at 100% for 10 min before return to 0% in 1 min on a 2.1 × 100 mm C18 column (Vydac, Hesperia, CA), during which the region corresponding to A214 was manually collected. While 1/100th of this was used for pooled Edman sequencing, aliquots approximating 400 μg of extracted peptides as based upon ELISA estimates of 2% (24) and an assumed handling loss of up to 50% (27) were next injected in a volume of 100 μl of 10% acetic acid and directly analyzed by RP-HPLC on a 2.1 × 100 mm C18 column (Vydac, Hesperia, CA) using a gradient of 2–10% acetonitrile with 0.06% TFA in 0.02 min and 10–60% of the same buffer in 60 min at 180 μl/min with fractions collected automatically every minute. The dye methyl violet base B (Aldrich, Milwaukee, WI) was added in a quantity of 500 ng to the 100-μl injection aliquots before this final RP-HPLC separation to control for gradient consistency between runs. Of the resulting fractions, 1/9th of each was removed and stored at −20°C for later Edman sequencing as desired; the remaining 160 μl of each was concentrated by speed-vac to remove TFA and resuspended in 25 μl of 0.1% acetic acid in 50% methanol for nanoelectrospray mass spectrometry (NanoES-MS). Edman degradation, which was performed on peptide pools to obtain motifs and on fractions accordingly (21), was conducted for 14 cycles on a 492A pulsed liquid phase sequence analyzer (Perkin-Elmer Applied Biosystems Division, Norwalk, CT) without cysteine derivatization, and raw data analysis was performed as previously described (28).

NanoES-MS ion map generation, comparison, and nanoelectrospray tandem mass spectrometry (NanoES-MS/MS) of potential ligand matches

To generate ion maps, 1–2 μl from each RP-HPLC was loaded into a gold/palladium alloy-coated borosilicate glass NanoES capillary tip (Prota- onda, Odense, Denmark) by gentle centrifugation and subjected to NanoES-MS on an API III triple quadrupole mass spectrometer (PE SCIEX, Foster City, CA) equipped with a collision cell upgrade (29) and a NanoES ionization source inlet (Prota- onda). Approximately 20–30 scans were collected with the following instrument settings: polarity, positive; each was 80 s, 1000 V; N2 curtain gas, 0.6 ml/min; step size, 0.2 atomic mass unit (amu); dwell time, 1.5 ms, and mass range, 325–1400 m/z. Spectra generated from the total ion chromatograms of each fraction were then compared with one another following baseline subtraction (spectral window width, 32 amu) and Kalman smoothing (Kalman gain, 2). Ion matches between spectra were located through datacentroiding (minimum distance, 0.3 amu; minimum width, 0.5 amu) and/or direct visual assessment, the latter proving beneficial in continuous matches of comparative intensity extremes. The states of identified matches were then evaluated by collisional dissociation of candidate ions via NanoES-MS/MS. Typical NanoES-MS/MS scans involved gaiting for an ion with the first quadrupole and scanning a range with the third quadrupole of 30–1000 m/z using a step size of 0.3 amu and a dwell time of 2 ms with lysine underivatized; the collision gas (Ar) was adjusted in each case to optimize fragmentation for the ion examined. NanoES-MS/MS data was evaluated and interpreted using the Predict Sequence algorithm (Bio-MultiView software, PE SCIEX) as well as PeptideSearch 3.0.2 (30) in instances of low ligand ionization/concentration or poor fragmentation. Advanced BLAST searches (31) were performed against databases available through the National Center for Biotechnology Information (National Institutes of Health, Bethesda, MD) web server to identify homology with currently cataloged sequences.

Results and Discussion

More than 200 individual ligands extracted from the four distinct HLA-B15 allotypes B*1508, B*1501, B*1503, and B*1510 were characterized in this study. Among our goals in examining these ligands was to test the hypothesis that endogenous peptides will overlap across polymorphisms in the Ag binding groove. Support
for this hypothesis comes primarily from in vitro peptide binding and/or competition assays (32–35), though a handful of endogenously loaded peptide overlaps have also been observed (28, 36–38). The ligands characterized here were those ion masses found in multiple B15 allotypes as demonstrated for one ion in Fig. 1a. Selected ions were then dissociated by NanoES-MS/MS, and the resulting fragment information was compared and interpreted (Fig. 1, b and c) to determine whether the ligands were sequence-identical matches.

Length and sequence variability exist among divergent B15 ligands

Individual peptide ligands characterized from the four B15 allotypes are shown in Fig. 2. The number of ligands for which either complete or partial sequences were obtained here was as follows: B*1508, n = 80; B*1501, n = 101; B*1503, n = 36; and B*1510, n = 35. The pooled motifs of peptides extracted respectively from B*1508, B*1501, B*1503, and B*1510 describe nonamers with various P2 and P9 dominant anchors and P3 auxiliary anchor preferences (Fig. 2), while individual ligand sequences range from 7 to 12 aa in length and demonstrate greater sequence heterogeneity at P2 than at their C-terminal anchors (Figs. 2 and 3). For allotypes B*1508, B*1501, and B*1503, a dominant C terminus was especially prominent, while B*1510 exhibited a P2 anchor nearly as strong as its C-terminal anchor (Fig. 3b). The C-terminal anchor was thus dominant among ligands bound by all four of the B15 allotypes characterized here.

In terms of length heterogeneity, the endogenous peptides eluted from B*1508, B*1501, B*1503, and B*1510 varied in length from 7 to 12 aa. A length breakdown of the peptides demonstrates that 7% are heptamers, 20% are octamers, 50% are nonamers, 20% are decamers, 2% are undecamers, and 1% are dodecamers. Further emerging from the length characterization of individual ligands is the observation that peptides bound by each of the B15 molecules respectively span ranges of 4 to 6 aa in length. For example, peptides eluted from B*1501 and B*1508 were 7–11 aa in length, those from B*1510 were 7–10 aa in length, and those from B*1503 were 7–12 aa in length (Fig. 3a). Coupling this length variability with P2 and C-terminal sequence flexibility leaves ~29% of the endogenously loaded peptides characterized in Fig. 2 as “ideal”
nonamers with both P2 and P9 anchors in concordance with their pooled Edman motifs.

Overlapping peptide ligands exhibit conserved C termini. Through the comparative mapping and sequencing approach outlined (Fig. 1), 19 peptide overlaps across B*1508 and B*1501 were defined. Similarly, seven ligands overlapping across B*1501 and B*1503 were identified, and four ligands were found to overlap across B*1508, B*1503, and B*1501. A conservative estimate is that our ion maps for each of the B15 allotypes represent 2000 individual peptides per molecule, yet B*1510 was not observed to share overlaps with the other three B15 allotypes in our ion maps (Fig. 4).

The data suggests that overlapping ligands bind across divergent B*1508, B*1501, and B*1503 binding grooves but are not permitted to do so by B*1510. This pattern emphasizes a dominant role for C-terminal anchors in natural peptide binding. Fig. 4 depicts, in the context of the class I peptide binding cleft, the locations of polymorphisms that individuate the four allotypes and highlights the anchoring residues for the peptide overlaps as indicated by their presenting molecule’s respective motif (Fig. 2).

The overlapping ligands exhibit conserved C termini. Through the comparative mapping and sequencing approach outlined (Fig. 1), 19 peptide overlaps across B*1508 and B*1501 were defined. Similarly, seven ligands overlapping across B*1501 and B*1503 were identified, and four ligands were found to overlap across B*1508, B*1503, and B*1501. A conservative estimate is that our ion maps for each of the B15 allotypes represent 2000 individual peptides per molecule, yet B*1510 was not observed to share overlaps with the other three B15 allotypes in our ion maps (Fig. 4).

The data suggests that overlapping ligands bind across divergent B*1508, B*1501, and B*1503 binding grooves but are not permitted to do so by B*1510. This pattern emphasizes a dominant role for C-terminal anchors in natural peptide binding. Fig. 4 depicts, in the context of the class I peptide binding cleft, the locations of polymorphisms that individuate the four allotypes and highlights the anchoring residues for the peptide overlaps as indicated by their presenting molecule’s respective motif (Fig. 2).

Bolding the amino acids of these overlapping ligands, which are in agreement with the motifs, underscores the trend whereby a C-terminal anchor sequence is conserved in overlaps while the N-proximal anchor is more flexible in its location and/or sequence. A lack of overlaps with B*1510 could be explained by the S3Y substitution at P116 (Table I). P116 participates in the C, E, and F Ag binding groove specificity pockets (41) and leaves valine (V) and leucine/isoleucine (X) as the preferred C-terminal anchors among most B*1510 ligands characterized (Fig. 2). Thus, the conserved C-terminal anchors that facilitate the occurrence of B*1508, B*1501, and B*1503 overlaps may fail to preferentially interact with the B*1510 C-terminal specificity pockets.

Length variation among overlapping ligands tends to mimic that observed among the total ligands characterized, previously discussed from Fig. 3. Notable exceptions were that overlaps between B*1501 and B*1503 were primarily octamers and that only one overlapping ligand was longer than 10 aa in length (Fig. 4). However, a limited number of B*1503 overlaps identified to date makes trends difficult to clearly establish regarding this B15 allotype.

Distilling the data from the 30 overlapping ligands so far identified leads us to suggest a model whereby the majority of ligands are first anchored into the binding groove by their C terminus. In order for a peptide to stably remain in the groove, we propose that following rigid initial anchoring at the C terminus a ligand might be further tethered into the class I Ag binding cleft at a more elastic N-proximal position. Such could argue the case for peptide ligand NQZHGSAEY, a nonamer that overlaps across B*1508, B*1501,
and B*1503 (Fig. 1). According to this model (Fig. 5), a C-terminal tyrosine anchors NQZHGSAEY into all three B15 allotypes, while a glutamine at P2 anchors the peptide into B*1501 and B*1503 and a glutamine/lysine (Z; most likely a lysine based upon motif assignments and fractional Edman sequencing data) at P3 provides additional anchoring for B*1501 and serves as the sole N-proximal anchor for B*1508. Therefore, length and N-proximal specificity would appear to play more secondary roles in endogenous ligand binding of B15 peptide epitopes.

C-proximal auxiliary anchors are not apparent in B15 pooled motifs

Edman-derived motifs for the four B15 allotypes shown in Fig. 2 illustrate P2 and P9 dominant anchors and P3 auxiliary anchors, but these motifs fail to capture trends for auxiliary anchoring at the C-proximal regions of endogenously bound ligands. Specifically, additional anchor preferences occur at both P7 and P8 in the case of nonamers, octamers, and heptamers, while in the cases of ligands longer than nonamers the two residues internal to the C terminus, which we refer to as C\textsubscript{2} and C\textsubscript{3}, also act as auxiliary anchors. A glance through the individual B15 ligands shows that the amino acids threonine (T), serine (S), valine (V), and glutamine/lysine (Z) predominate at the C-proximal ends (involving either P7/P8 or C\textsubscript{2}/C\textsubscript{3}) of these ligands (Fig. 2). An example of how C-proximal auxiliary anchors impact endogenous ligand binding is that more than half of the peptides that overlap across the B*1508 and B*1501 Ag binding grooves bear threonine at P7 or P8, and in four cases the peptides that bind B*1508/1501 or B*1508/1501/1503 are heptamers with threonine occupying P7, their C-terminal positions (Fig. 4). The role of threonine as a C-proximal/C-terminal auxiliary anchor is best illustrated with the B*1508/1501/1503 overlapping heptamer CPLSCFT, where threonine provides a C-terminal anchor not evident in the pooled motif.

Sources of individual B15 ligands reflect cytosolic Ag processing

Examples of peptides from this study with complete homology to stretches of known proteins are shown in Table II. The majority of ligands are derived from intracellular proteins, which illustrates the endogenous pathway is generating the class I-loaded peptides described here (2). Of the peptide sequences listed, six have also been reported by other laboratories. It is noteworthy that overlaps across other molecules of the HLA-B15 group are evident within this data collection: the B*1510 tapasin\textsubscript{354–362} ligand HHSbGVS has been sequenced from B*1509 extracts (36), and the B*1508/1501 ubiquitin-protein ligase\textsubscript{83–91} -derived ligand ILGP PGbSY was characterized from endogenously bound B*1502 peptides (36). We identified the eIF3-p66\textsubscript{61–69} nonamer SQFGGGGSQY (42, 43) within B*1508, B*1501, and B*1503 extracts. The decamer YMIDPSGVSY, which is homologous to proteasome subunit C8\textsubscript{150–159}, was also previously described as a ligand for B*1502 (36), B*1508 (36), and B*4601 (42); we found it presented by B*1501 and by B*1508. Therefore, overlapping ligands identified here overlap in Ag presentation with the HLA-B15 allotypes characterized by others, and these overlaps are likewise determined by both flexible N-proximal anchor positions and conserved C-terminal anchor/C-proximal auxiliary anchor preferences.

Conclusions

The primary impetus for characterizing class I peptide ligands is to understand the functional impact of class I polymorphism in terms of endogenous ligand presentation. A fundamental realization of
how naturally processed peptides bind both individual and multiple class I molecules can then be translated into peptide-based therapies to elicit protective CTL. Therefore, an accurate interpretation of peptide sequence data, either individual or pooled, will in turn further the selection of optimal viral and tumor-associated ligands for therapeutic applications.

Our examination of numerous B15 ligands enhances understanding the rules that govern class I peptide presentation. Building upon the pooled peptide motif, our data indicates that queries for potential epitopes specific to these allotypes might be optimized in three ways. First, although nonamers represent half the ligand population, peptide epitopes can range in length from 7 to 12 aa.

In summary, divergent HLA-B15 molecules appear to present similar peptide ligands as long as polymorphisms do not alter C-terminal anchoring pockets and while an N-proximal ligand residue can be subsequently anchored within the Ag binding groove. Supporting data suggests that these principles might extend beyond the four B15 allotypes studied here. Specifically, unpublished results by Ghosh and Wiley (noted in Ref. 44) indicate that an octamer has been observed to successfully bind a class I molecule.

FIGURE 4. Overlapping HLA-B15 ligands identified by NanoES-MS mapping and characterized by NanoES-MS/MS. Ribbon diagrams of the class I Ag binding groove show residue substitutions (black, numbered) between B*1508, B*1503, and B*1510 with respect to the ancestral B15 allele B*1501 (top). Shown below the ribbons are three different groups of ligands identified that overlap the divergent binding grooves. The ligands are categorized from top to bottom as follows: peptides that overlap B*1508 and B*1501; peptides that overlap B*1508, B*1501, and B*1503; and peptides that overlap B*1501 and B*1503. No overlapping ligands have been identified to date between these three molecules and B*1510. Ligand residues that coincide with dominant and strong motif amino acids for the given motifs (Fig. 2) are indicated in bold type. X, Z, and dashes (−) are as defined in Fig. 2.

FIGURE 5. N-proximal and C-terminal anchoring of the B*1508/1501/1503-overlapping nonamer NQZHGSAYE. The shared C-terminal anchoring preference for tyrosine in the NQZHGSAYE ligand among B*1508, B*1501, and B*1503 as defined by the respective motifs (Fig. 2) are shaded black, while the varied N-proximal anchoring preferences likewise reflected in the motifs are shaded gray. Residues are numbered sequentially from the N-terminus.
by its C terminus despite being shown through x-ray crystallography to not even reach the N-terminal pocket of the binding groove. In addition, a recently described HIV-gag197–205 CTL epitope presented by murine class I Kd fails to show a motif-preferred tyrosine at P2 and instead associates stably through its C terminus despite being shown through x-ray crystallography to not even reach the N-terminal pocket of the binding groove. We thank Dr. Caroline Thompson (University of Oklahoma Health Sciences Center DNA Sequencing Facility) for the sequencing of class I HLA.

Acknowledgments

We thank Dr. Caroline Thompson (University of Oklahoma Health Sciences Center DNA Sequencing Facility) for the sequencing of class I clones.

References

8. Kataoka, K., O. Rötzschke, S. Stevanovic, G. Jung, and H.-G. Rammensee. 1991. Allele-specific motifs revealed by sequencing of self peptides eluted from other HLA-B, as well as HLA-A and HLA-C, molecules confirm this speculation by exhibiting conserved C termini along with length and N-proximal sequence versatility.


