
Donald F. Hunt, Robert A. Henderson, Jeffrey Shabanowitz, Kazuyasu Sakaguchi, Hanspeter Michel, Noelle Sevilir, Andrea L. Cox, Ettore Appella and Victor H. Engelhard

*average

Why The JI? Submit online.

- **Rapid Reviews! 30 days** from submission to initial decision
- **No Triage!** Every submission reviewed by practicing scientists
- **Fast Publication!** 4 weeks from acceptance to publication

**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
Characterization of Peptides Bound to the Class I MHC Molecule HLA-A2.1 by Mass Spectrometry

DONALD F. HUNT,* ROBERT A. HENDERSON,* JEFFREY SHABANOWITZ,* KAZUYASU SAKAGUCHI,* HANSPETER MICHEL, NOELLE SEVILIR, ANDREA L. COX, ETTORE APPELLA, VICTOR H. ENGELHARD

Antigens recognized by T cells are expressed as peptides bound to major histocompatibility complex (MHC) molecules. Microcapillary high-performance liquid chromatography–electrospray ionization–tandem mass spectrometry was used to fractionate and sequence subpicomolar amounts of peptides isolated from the MHC molecule HLA-A2.1. Of 200 different species quantitated, eight were sequenced and four were found in cellular proteins. All were nine residues long and shared a distinct structural motif. The sensitivity and speed of this approach should enhance the analysis of peptides from small quantities of virally infected and transformed cells as well as those associated with autoimmune disease states.

Cytotoxic T lymphocytes (CTLs) are a part of the immune system concerned with recognition of host cells that express new antigens as a result of viral infection or transformation. CTLs do not recognize new antigens directly, but only as short peptides bound to a deep cleft in class I molecules of the MHC (1–3). Newly synthesized viral and cellular proteins are fragmented into 9–11-residue peptides in the cytoplasm, transported to the endoplasmic reticulum where they bind to class I molecules, and then expressed on the cell surface (4–7). Each of the allelic forms of the class I MHC molecule binds to a complex mixture of structurally distinct peptides (8, 9). Information on the nature of these peptides has been obtained from studies with synthetic peptides (10–12) and from Edman degradation applied to unfracti cated mixtures of peptides extracted from five different class I MHC molecules (8). Sequences of 11 peptides extracted from HLA-B27 were identified after high-performance liquid chromatography (HPLC) fractionation and Edman degradation (9). Because HPLC was unable to completely resolve the complex mixture, this analysis could only be applied to the few fractions that contained one or two dominant peptides. Declining PTH (phenylthiohydantoin)–amino acid yields made it difficult to determine the exact number of residues in several peptides.

We have applied microcapillary HPLC–
electrospray ionization–tandem mass spectrometry to circumvent the above problems. In a matter of hours, this technique determines the molecular mass and therefore maximum length of each peptide component, and the approximate number and quantity of individual peptides. Sequence information can also be obtained on subpicomolar amounts of peptides. We analyzed the naturally processed peptide bound to HLA-A2.1, one of the most widely distributed class I molecules within the human population. The three-dimensional structure of this molecule allows modeling of the complex (2).

HLA-A2.1 molecules were purified by immunoprecipitation from the human B lymphoblastoid cell line CIR-A2.1. The associated peptides were released by acid extraction and separated from HLA-A2.1 and a
tbody by filtration. A "mock" extract was prepared from an equal number of Hmy2.C1R (CIR) cells, which do not express HLA-A or -B products (13). Since CIR-A2.1 is a transfected CIR, these cell lines should be identical except for the expression of HLA-A2.1. Extracts were fractionated by microcapillary reversed-phase HPLC and eluted directly into the electrospray ion source of a triple quadrupole mass spectrometer (14). A broad band of material was observed in the extract from CIR-A2.1 (Fig. 1A), but not in the extract from CIR (Fig. 1B). If all HLA-A2.1 molecules contain bound peptide, the total peptide yield from 10^6 CIR-A2.1 cells should be roughly 20 pmol; we estimate that our total yield of peptide was about 70 to 80% of the amount expected. This provides evidence that the material isolated for CIR-A2.1 cells by this method is specifically associated with the HLA-A2.1 molecule.

Material with a nominal mass-to-charge ratio (m/z) of 1121 was present in the extract from CIR-A2.1 cells but not that from CIR cells (Fig. 1 inset). By a similar analysis, approximately 200 distinct ions having a signal-to-noise ratio (S/N) of at least 2 were observed in the material isolated from CIR-A2.1 cells. Of these, 47% were singly charged, 45% were doubly charged, and 8% were triply charged. Calculated molecular masses ranged between 737 and 1370. Known quantities of synthetic peptides added to cell extracts allowed quantitation of peptides, and defined the lower limit of detection (S/N = 2) as 30 fmol. Of the 200 peptides detected, 10% were present at 150 to 600 fmol per 10^8 cells.

Fig. 1. Total ion chromatograms of HLA-A2.1–associated peptides. CIR-A2.1 (A) or CIR (B) cells (2 x 10^7) were solubilized in 20 mM tris, pH 8.0, 150 mM NaCl, 1% NP-40, 2 mM PMSE, 100 mM iodoacetamide, aprotinin (5 mg/ml), leupeptin (10 mg/ml), pepstatin A (10 μg/ml), EDTA (5 mg/ml), and 0.2% sodium azide. After sample centrifugation at 100,000g for 1 hour, supernatants were immunoprecipitated with the HLA-A2.1-specific monoclonal antibody B7.2 (30) and protein A–Sepharose (Sigma). Precipitates were extracted with 0.2 N acetic acid. The supernatants were brought to pH 2.1 with glacial acetic acid and boiled for 5 min, centrifuged through a MiniSpin unit (Millipore) with a nominal limit of 5 kDa, and concentrated by vacuum centrifugation. Mass spectra were recorded on a Finnigan-MAT TSQ-70 (San Jose, California) triple quadrupole mass spectrometer equipped with an electrospray ion source and a C-18 microcapillary-HPLC column (75 μm by 10 cm) (14). The column was eluted with a gradient of 0 to 80% acetonitrile in 0.5% acetic acid over 10 min at 1 to 2 μl/min. The electrospray needle was operated with a voltage differential of 4 to 5 kV and a sheath flow of 2 to 4 μl/min of a 3:1 mixture of methanol:0.5% acetic acid. Total ion signal from samples equivalent to 10^6 cells was obtained by scanning the range of masses corresponding to m/z values between 300 and 1500 every 2 s and then summing each of the spectra. Insects are plots of ion current for the peptide (M+H)^+ ions at m/z 1121 (nominal mass).

D. F. Hunt, J. Shabanowitz, H. Michel, A. L. Cox, Department of Chemistry, University of Virginia, Charlottesville, VA 22903.
R. A. Henderson, N. Sevill, V. H. Engelhard, Department of Microbiology and Immunology, University of Virginia, Charlottesville, VA 22908.
K. Sakaguchi and E. Appella, Laboratory of Cell Biology, National Cancer Institute, NIH, Bethesda, MD 20892.

*Contribution of these authors were equivalent and their order should be considered arbitrary.

6 MARCH 1992
whereas 90% were present at 30 to 150 fmol. By this measure, individual peptides in the mixture represent between 0.08 and 1.2% of the total peptide bound. Depending on the cell, this would correspond to as few as 100 complexes per cell or as many as 3000. These numbers are in accord with estimates of the number of peptide-MHC complexes required to stimulate T cells (12, 15, 16). However, 67 of the 200 peptides detected carried two-thirds of the ion current observed for this group. Thus, HLA-A2.1 seems to be preferentially occupied by a relatively small set of peptides. On the other hand, the number of peptides detected accounts for only 50% of the total ion current. If the remaining undetected peptides are all present at amounts near that of the background, the total number of different peptides presented by HLA-A2.1 could easily exceed 1000.

Table 1 shows a partial list of data obtained on peptides chosen at random from the more than 200 species observed. Fourteen of the peptides were in extracts of immunoprecipitates from both CIR-A2.1 and another HLA-A2.1-expressing cell line, JY, but not in extracts from CIR-B cells. Five other affinity-purified class I peptides in extracts from the JY cells but were undetected in extracts from either CIR-A2.1 or CIR-B cells. To obtain sequence information, (M+H)+ and (M+2H)+ ions were subjected to collision-activated dissociation (CAD) on the triple quadrupole mass spectrometer (17). Representative results for the (M+H)+ ion at m/z 1121 (peptide 19) are shown in Fig. 2. Complete sequences for 8 peptides in Table 1 have been obtained; partial sequences have been obtained on 11 others. Each of the completely sequenced peptides is a nonamer and contains either Leu or Ile at position 2. Since Ile and Leu are of identical mass, they cannot be differentiated on the triple quadrupole instrument. Leu was considered a "dominant residue" at position 2 in unfraccionated peptide mixtures isolated from HLA-A2.1 (8). For peptide 19, a definitive assignment of Leu at position 2 was established by adding synthetic peptides containing Leu or Ile at this position to the cell extract, and demonstrating that only the Leu-containing peptide coeluted from the microcapillary HPLC column with the naturally processed species. For the 19 peptides for which COOH-terminal information is available, all end in residues with aliphatic hydrocarbon side chains: nine terminete in Val, eight in Leu/Ile, and two in Ala (Table 1 and 18). Only Val was categorized as "dominant" based on the pool sequencing technique (8). No other residues were present at any position in more than 50% of the peptides. None of the sequences in Table 1 contains Met at position 6, or Tyr at position 8, assignments that were considered "strong" in analysis of pooled peptides (8).

The length of the peptides bound to HLA-A2.1 is identical to that of peptides bound to HLA-B27 (9) and inferred for other class I molecules (12). This could reflect the mechanism by which these peptides were generated, or their ability to bind with particularly high affinity (3, 19). However, the motif for peptides bound to HLA-A2.1 is different from that of peptides bound to HLA-B27. Position 2 was also the most conserved in the peptides isolated from HLA-B27, but was an Arg (9). This residue is suggested to interact with negatively charged Glu9 in the B pocket of the HLA-B27 antigen binding site (3). In HLA-A2.1, this pocket is hydrophobic and is capable of accommodating a Leu or an Ile side chain.
Fig. 3. Synthetic peptides corresponding to sequenced peptides bind to HLA-A2.1. Synthetic peptides are referenced by the numbers in Table 1. JY cells were incubated with the M1 peptide of influenza A (8 μg/mL), in the presence or absence of one of the synthetic peptides (200 μg/mL) as a competitor, in a total volume of 250 μL of RPMI 1640 containing 5% non-radioactive calf serum and 50 μCi of 81Cr for 3 hours. Washed target cells (2 × 10⁶) were incubated in 200 μL with either an M1 peptide-specific, HLA-A2.1-restricted CTL clone 5.2-11 (31) (filled bars) or an HLA-A2.1-specific human allogeneic CTL line (stippled bars) at an effector to target ratio of 1:25:1. Data are the means of duplicate samples with an SD of less than 5% and are representative of at least three independent experiments.

(2). In addition, positively charged residues predominate at position 9 in peptides associated with HLA-B27, and this residue was suggested to interact with Asp¹¹⁰ in the F pocket. In HLA-A2.1, position 116 is a Tyr, rendering the F pocket smaller and more hydrophobic. Based on these considerations and the similarity of the electron density corresponding to peptide in these two molecules (2, 3), we suggest that interactions of Leu or Ile at position 2 with the B pocket, an aliphatic side chain at position 9 with the F pocket, and perhaps a nine-residue length are the primary determinants of peptide binding to HLA-A2.1.

The eight completely sequenced peptides in Table 1 were synthesized in order to demonstrate that the peptides identified actually bound to HLA-A2.1. When no complete match to a sequence in the protein or gene databases was found (below), the mixture of Leu and Ile was incorporated at positions specified by X. Cells of the HLA-A2.1-expressing line JY were incubated with the M₁₇₇₀ peptide of influenza A virus in the absence or presence of a 10- to 20-μg/mL excess of one of the competitor peptides and assayed for their ability to be lysed by an M₁ peptide-specific, HLA-A₂.1-restricted CTL clone. All eight peptides inhibited recognition under these circumstances but had no effect on recognition by a human allogeneic CTL line that is specific for HLA-A₂.1 (Fig. 3). Thus, these synthetic peptides bind to HLA-A₂.1 in vitro, which confirms their association with HLA-A₂.1 in intact cells. Additional peptides were synthesized in which two Gly residues at positions 5 and 6 in peptide 19 were replaced by a single Asn to produce an octamer, or the Thr and Leu in positions 1 and 2 of peptide 19 were substituted with Val and Asn or Val and Asp. None of these peptides blocked recognition, indicating that a nine-residue length, the presence of Leu or Ile at position 2, and perhaps a polar residue at position 1 are important features for peptide binding.

To identify possible precursor proteins for these peptides, gene and protein sequence databases were searched in a manner that allowed X residues to be assigned as either Leu or Ile. Only four peptides gave 100% matches with proteins in the library (Table 1). All of the sequences contain Leu at position 2, although Ile is often found at other positions. A match for peptide 8 occurs in the 65-kD α and β regulatory subunits of human protein phosphatase 2A and in a highly homologous human transformation-associated protein, p61. Both are cytoplasmic proteins and p61 represents only 0.1% to 0.02% of the total cellular protein (20). A match for peptide 14 appears in the human nuclear protein p68, which exhibits RNA-dependent adenine triphosphatase activity and RNA unwinding activity (21). A different sequence from this protein is found in a peptide bound to HLA-B₂.7 (9). A match for peptide 19 is found in TIS21, a protein of unknown function that is a member of the primary response group of genes induced by growth factors (22). Finally, a match for peptide 4 is found in the human protein IF-30, the expression of which is induced by gamma interferon (23). This precursor sequence is unique among all naturally processed peptides described to date in that it is located in the signal peptide domain of the protein. This peptide is also bound to HLA-A².1 molecules expressed on CEM721.174(T2) (24), a murine cell line defective in the normal antigen processing pathway (6, 7). This result suggests a second pathway through which peptides can enter the endoplasmic reticulum and associate with MHC class I molecules.

Peptides derived from the processing of normal cellular proteins located in the cytoplasm have been implicated in the correct folding and intracellular transport of class I molecules (6, 7) and in the epitopes recognized by allogeneic T cells (25-28). The use of tandem mass spectrometry for the direct analysis of these peptides has enhanced our knowledge of peptide-MHC interactions and the cellular processes that regulate formation of these complexes and should allow the direct identification of peptides that are allogeneic T cell epitopes. Peptides associated with class II MHC molecules can also be characterized by the above approach (29). This same methodology should also facilitate identification of peptide antigens associated with viral infection, cellular transformation, and autoimmunity.

REFERENCES AND NOTES

11. O. Rotschke et al., ibid., p. 252.
32. Supported by USPHS grant GM37537 (to D.P.H.) and AI20963 (to V.H.) and by National Science Foundation predoctoral fellowship RCD 9145777 (to N.S.).

5 November 1991; accepted 22 January 1992