The Three-Dimensional Structure of HLA-B27 at 2.1 Å Resolution Suggests a General Mechanism for Tight Peptide Binding to MHC

Dean R. Madden,* † Joan C. Gorge,* †
Jack L. Strominger,* and Don C. Wiley**†
*Department of Biochemistry and Molecular Biology
†Committee on Higher Degrees in Biophysics
‡Howard Hughes Medical Institute
Harvard University
Cambridge, Massachusetts 02138

Summary

Cell surface complexes of class I MHC molecules and bound peptide antigens serve as specific recognition elements controlling the cytotoxic immune response. The 2.1 Å structure of the human class I MHC molecule HLA-B27 provides a detailed composite image of a co-crystallized collection of HLA-B27-bound peptides, indicating that they share a common main-chain structure and length. It also permits direct visualization of the conservation of arginine as an "anchor" side chain at the second peptide position, which is bound in a potentially HLA-B27-specific pocket and may therefore have a role in the association of HLA-B27 with several diseases. Tight peptide binding to class I MHC molecules appears to result from the extensive contacts found at the ends of the cleft between peptide main-chain atoms and conserved MHC side chains, which also involve the peptide in stabilizing the threedimensional fold of HLA-B27. The concentration of binding interactions at the peptide termini permits extensive sequence (and probably some length) variability in the center of the peptide, where it is exposed for T cell recognition.

Introduction

The ability of cytotoxic T cells (CTLs) to detect and destroy diseased cells depends on the cell surface presentation of specific peptide antigens bound to class I major histocompatibility complex (MHC) molecules. Proteolytic fragments of proteins synthesized within the cell are transported into the endoplasmic reticulum, where they associate with class I MHC molecules (reviewed in DeMars and Spies, 1992). These complexes are expressed at the cell surface, "presenting" a sampling of the products of intracellular protein synthesis to T cell receptors (TCRs) on the surface of circulating CTLs (reviewed in Brodsky and Guagliardi, 1991). Different CTLs express TCRs with different binding surfaces, and TCR binding to peptide–MHC complexes on the target cell surface initiates cytosis of the target cell. CTLs reacting to autologous peptide–MHC complexes are thought to be eliminated in the thymus or inactivated, in order to prevent autoimmune responses (reviewed in Davis, 1990). Intracellular synthesis of foreign (viral) or tumor-specific proteins can thus be indicated at the cell surface by class I MHC presentation of peptides to which the CTL population has not been made tolerant, leading to the killing of the diseased cell (reviewed in Townsend and Bocimer, 1989).

Class I MHC molecules are very polymorphic (Parham et al., 1996; Hedrick et al., 1991); however, several amino acids are highly conserved, even in different species (Bjorkman and Parham, 1990). Conserved or nearly conserved MHC side chains are found clustered at either end of a peptide-binding cleft formed by the three-dimensional fold of the α1α2 superdomain (α1α2) (Bjorkman et al., 1987a, 1987b; Saper et al., 1991). Side chains of many of the amino acids that show the greatest variability among alleles also point into the cleft and determine its shape and chemistry (Bjorkman et al., 1987b; Garrett et al., 1989; Saper et al., 1991). Peptide–MHC complexes are extremely stable under physiological conditions, so that only a small fraction of the class I molecules normally found on the cell surface lack bound peptide (Chen and Parham, 1989). Also, at physiological temperatures, bound peptide appears to be required for the expression of stable class I MHC molecules (Townsend et al., 1989). Peptides can be released by denaturation of purified class I MHC complexes and constitute a heterogeneous population 8 or 9 amino acids long, with a pattern of sequence homology at two or three positions (binding motif) characteristic of the MHC antigen involved (Faith et al., 1991; Jardetzky et al., 1991; Van Bleek and Nathenson, 1991; Hunt et al., 1992).

In vitro, extending or shortening a peptide nonamer that binds to H-2Kd reduces the half-life of the resulting peptide–class I complex (Corandolo et al., 1991). However, peptides as long as 13 mers have been found to be associated with class I molecules isolated from cells with antigen-processing defects (Henderson et al., 1992; Wei and Cresswell, 1992).

HLA-B27 is strongly associated with the occurrence of several diseases. Most (90%–95%) of the individuals who suffer from ankylosing spondylitis express HLA-B27, and lower correlations are found with other diseases, including arthritic disorders triggered by enteric bacterial infections (reviewed in Khan, 1987). The association does not appear to be subtype specific and includes HLA-B*2705 (Breur-Vriesendorp et al., 1987). The role of HLA-B27 in causing disease is still unclear, and it remains possible that another gene, tightly linked to HLA-B27, is the causative agent. However, experiments in which transgenic rats expressing high levels of HLA-B27 suffered from skin and joint abnormalities (Hammer et al., 1990) point to a direct role for HLA-B27. Several hypotheses have been advanced for the disease association (reviewed in Benjamin and Parham, 1990), among them the possibility that HLA-B27 may itself, or in complex with an endogenous peptide, serve as the target of an autoimmune response.

The structure of HLA B*2706 (HLA-B27) was determined at 3.0 Å resolution and revealed clear electron den-
Figure 1. Side View of the Superposition of HLA-B27 and HLA-Aw68

High resolution structures were determined by cryocrystallography for HLA-B27 to 2.1 Å (closed bonds) and HLA-Aw68 to 1.9 Å (open bonds) (H.-C. Guo, J. L. S., and D. C. W., unpublished data). α1α2 domain Ca atoms (black) of residues in β sheet or α helix structures were superimposed by the method of least squares to show packing shifts of the α3 domain (blue) and β2m (red). Residues 226 (which influences CD8 binding) [Selte et al., 1990] and 273 (heavy chain C-terminus following papain solubilization) are indicated. The peptide-binding cleft is at the top of the figure, and is viewed from the side, with the α1 α helix running from left to right behind the α2 α helix. The concave surface referred to in the text is located on the lower right surface of the molecule in this view and is formed by the bottom of the peptide-binding groove, the right-hand side of α2, and the front of β2m. Figure produced with program ORTEP (Johnson, 1965), provided by S. J. Remington and modified locally.

sity for nonameric peptides bound in an extended conformation, indicating a likely role of the peptide main chain and of side chains P2, P3, P7, and P9 (where Pi = the i-th peptide residue) in contacting MHC (Madden et al., 1991; Madden and Wiley, 1992). In addition, the clusters of conserved residues at both ends of the MHC cleft form binding pockets for the charged peptide termini. Sequences of individual eluted peptides were consistent with the structural information (Jardetzky et al., 1991).

The crystallographic analysis of a population of peptides bound to a class I MHC antigen (analogous to the sequencing of pools of eluted peptides [Falk et al., 1991]) detects amino acid structural diversity in that population and therefore complements the study of crystals of class I MHC proteins with a single peptide bound (M. L. Silver, H.-C. Guo, J. L. S., and D. C. W., submitted; Fremont et al., 1992; Zhang et al., 1992). Using X-ray cryocrystallography, we have collected data to 2.1 Å resolution on HLA-B27 crystallized from material containing a naturally processed peptide population. A single, dominant main-chain conformation is found for peptides bound to HLA-B27, with the N- and C-termini precisely positioned by numerous hydrogen bonds at both ends of the cleft. These hydrogen bond networks, which link the peptide to 9 conserved residues located on both sides and on the bottom of the binding cleft, may contribute substantially to the peptide-dependent stability of the HLA molecule. The conservation of many of the MHC residues that sequencer the termini of bound peptides suggests that these interactions are a general feature of tight peptide binding to most class I molecules. High resolution refinement also reveals electron density for several peptide side chains and for 15 water molecules that could not be unambiguously located at lower resolution. A detailed analysis of the fit between observed electron density and peptide side-chain models offers structural evidence for conservation of peptide side chains that bind MHC and variability of peptide side chains involved in TCR discrimination between different antigens. At 2.1 Å resolution, electron density for the P2 peptide side chain is extremely well-fit by arginine, providing structural evidence that arginine is present at P2 in essentially all endogenous peptides bound to HLA-B27. In the P2-bind-
ing pocket, a set of side chains common to the B27 antigens may explain the strict selection for arginine at this position, which could be the unique feature of HLA-B27 involved in disease association.

Results

The Conformation of HLA-B27

HLA-B27 adopts a conformation similar to those of other class I MHC antigens studied, with an extended peptide-binding site in α1α2 and immunoglobulin constant domain folds for the α3 domain (α3) and β2-microglobulin (β2m). The largest structural differences between I A-A*0201 (HLA-A2) (Saper et al., 1991) or HLA-A*3801 (HLA-Aw68) (H.-C. Guo, J. L. S., and D. C. W., unpublished data) involve minor packing rearrangements of α3 and β2m relative to α1α2, and shifts in the conformations of loops (Figure 1). The root mean square (RMS) difference between Cα atoms following least-squares superposition between HLA-B27 and HLA-A2 is 0.84 Å, and between HLA-B27 and HLA-Aw68 is 0.88 Å. Residues of α1α2 that are located in secondary structure elements can be superimposed with RMS deviations of only 0.42–0.46 Å, approaching the level of coordinate error. Given the α1α2 superposition, however, RMS Cα differences for residues in α3 and in loops are between 1.2 and 1.7 Å (Figure 1). This pattern of differences is also found between HLA-A2 and HLA-Aw68 (H.-C. Guo, J. L. S., and D. C. W., unpublished data), between human and murine class I antigens (Fremont et al., 1992; Zheng et al., 1992), and between the structures of HLA-B27 determined at room temperature and at −160°C (overall Cα RMS difference is 0.57 Å). Similarly small temperature-dependent structural differences have been found in studies of crystals of ribonuclease A (Titon et al., 1992). The observed shifts in domain packing appear to reflect the relatively flexible association of α3 with α1α2, which involves only a small number of interdomain contacts (Saper et al., 1991). One additional α3–α1α2 contact is found in HLA-B27, involving a hydrogen bond from the HLA-B locus-specific residue Arg-239 (α3) to the main-chain carbonyl oxygen of Arg-48 (α1), but this does not require a major adjustment of the main chain at either residue.

The packing of α3 and β2m against the underside of the α1α2 β sheet does not completely cover the sheet and leaves the molecule with a concave surface formed by the underside of the peptide-binding domain and the sides of the immunoglobulin-like domains (Saper et al., 1991) (Figure 1). Part of this surface could have a role in the reported formation of class I tetramers on the cell surface (Krishna et al., 1992). Also located on the concave surface is the region of α3 where clusters of CD8-binding residues have been identified (Potter et al., 1989; Sall et al., 1989, 1990; Connolly et al., 1990), including residues in the 220–229 loop. These residues exhibit some of the weakest electron density in the HLA-B27 structure, are characterized by average atomic B factors greater than 45 Å² (maximum for the molecule = 57.5 Å² at Glu-222), and vary greatly between HLA-B27 and HLA-Aw68 (which has a crystal contact in this region and which also does not bind CD8 (Sall et al., 1989)) (Figure 1). These residues may be relatively disordered in solution.

A Consensus Structure for a Population of HLA-B27-Bound Peptides

The peptide electron density found in the HLA-B27 binding site, before the inclusion of any peptide model in the crystallographic calculations (Figure 2A), appears to represent a crystallographic average of a population of bound peptides. A heterogeneous pool of peptides is bound to human lymphoblastoid cell L G2-derived HLA-B27 (Jardetzky et al., 1991) similar (see Experimental Procedures) to that used for crystallization. There is a large volume of solvent directly above the bound peptide in HLA-B27 crystals, and it therefore seems likely that sequence, and at least limited length, heterogeneity could be accommodated without disrupting the crystal lattice. Since electron density detected by X-ray diffraction represents a superposition of the structures of the individual molecules found in the diffracting volume of the crystal, the peptide electron density seen here should represent a composite of the binding conformations of the different peptides in the cocystalized population. Common to all peptides sequenced from the L G2-derived HLA-B27 are main-chain atoms for a peptide nonamer and an arginine side chain at position P2. The remaining peptide side chains are variable (Jardetzky et al., 1991). The peptide electron density observed here is clearest for the atoms common to all peptides, weaker for peptide side chains that show limited sequence diversity (P3 and P9), and poor or absent for the most variable amino acids, as would be expected for an average of peptides in which common elements adopt a common threedimensional structure.

The same general binding conformation appears to be adopted by most of the peptides that bind HLA-B27. Following refinement of a peptide model with the sequence ARAAARAAARAAAR (ARA4), and its inclusion in the HLA-B27 model used to generate starting phases for iterative map averaging, the mean real-space correlation (a measure of the fit of the model to the observed electron density) is 0.90 for the peptide main-chain atoms, compared with 0.91 for HLA residues. The average real-space fit to an iteratively averaged map calculated before introduction of any peptide model (Figure 2A) is 0.85. For the arginine side-chain atoms (Figure 2B), the real-space fit following peptide refinement is 0.95, equivalent to the best 8% of HLA side chains. Crystallographic refinement also indicates high level occupancy of the peptide-binding site. The peptide occupancy refines to 99%, and individual atomic temperature factors, which would be sensitive to static disorder among different bound peptides, refine to an average of 27 Å², comparable with an average of 24 Å² for HLA-B27 as a whole.

Some electron density is also found for several peptide side chains at positions of sequence variability (Figure 2A). Throughout refinement of the ARA4 peptide model, several of these putative peptide side chains were among the strongest (>4σ) positive features of difference electron density maps (Fobs-Fcalc, see Table 2 legend), which reveal electron density not accounted for by the model. A second
peptide model, with sequence RRIKAITLK, was built and refined, yielding a greater than 0.70 real space fit to iteratively averaged electron density calculated before introduction of any peptide model for only three side chains: P2-Arg (0.79), P3-Ile (0.71), and P9-Lys (0.72) (see Figure 2A). These three side chains have extensive interactions with HLA-B27 residues and have been identified as anchor residues (Jardetzky et al., 1991; Madden et al., 1991). The remaining peptide side chains have correlation coefficients equivalent to the worst four HLA side chains. Still, for several of these side chains (e.g., P1, P7, and P8), the 2.1 Å electron density maps (Figure 2A) are considerably clearer than at 3.0 Å, establishing side-chain orientations that had been tentative. Thus, the observed pattern of side-chain three-dimensional structural variability matches the primary sequence variability found in the peptide population (Jardetzky et al., 1991).

**Peptide–MHC Binding Interactions**
Bound peptide adopts an extended conformation running the length of the binding site (Figure 3). Most HLA-B27 contacts are concentrated in the narrow ends of the binding site where the peptide amino and carboxyl termini are held tightly (Figure 3A). In contrast, in the wide central region the peptide main chain arches up away from the floor of the cleft, and 12 water molecules bind in the site.
underneath and alongside the peptide (Figure 3B). Weak electron density located in this region of the cleft indicates that additional water molecules may also be loosely bound here.

Interactions between the bound peptide and HLA-B27 are extensive. The peptide model RRIKA\textsc{it}LK contacts 87 atoms in 31 HLA-B27 side chains, 10 of them conserved among different class I molecules (Figure 3A); 773 Å² of HLA-B27 and 1230 Å² of peptide solvent-accessible surface area are buried upon binding. Fifteen hydrogen bonds are made from the peptide directly to HLA-B27, and seven further hydrogen bonds are made to water molecules in the site, which in turn hydrogen-bond to HLA-B27 (Table 1). Most peptide-MHC interactions involve atoms common to the peptides known to bind HLA-B27 (Jardetzky et al., 1991). Of the 22 hydrogen bonds formed by peptide model RRIKA\textsc{it}LK (Figure 4B), 19 would be formed by atoms found in the minimal binding peptide ARA\textsc{v}. Of the 2033 Å² of solvent-accessible surface area buried by the binding of RRIKA\textsc{it}LK to HLA-B27, 71% would be buried by the binding of ARA\textsc{v}, and 46% would be buried by the binding of only 5 alanine residues at P1–P3 and P8–P9. Replacing alanine at P2 with arginine would contribute a further 9% of buried surface and four additional hydrogen
Table 1. Peptide Hydrogen Bonds to HLA-B27

<table>
<thead>
<tr>
<th>Peptide Residue</th>
<th>Atom</th>
<th>Distance (Å)</th>
<th>Hydrogen Bond Partner Atom</th>
<th>Residue</th>
<th>Distance (Å)</th>
<th>Second-Shell Hydrogen Bond Partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg-1 N</td>
<td>N</td>
<td>2.6</td>
<td>O₁</td>
<td>Tyr-171</td>
<td>2.6</td>
<td>Tyr-59</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>2.8</td>
<td>O₁</td>
<td>Tyr-7</td>
<td>2.5</td>
<td>Wat-481</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>2.8</td>
<td>O₁</td>
<td>Tyr-159</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Nc</td>
<td>2.9</td>
<td>O₂</td>
<td>Glu-163</td>
<td>3.1</td>
<td>Arg-62</td>
</tr>
<tr>
<td>Arg-2 N</td>
<td>N</td>
<td>3.0</td>
<td>O₁</td>
<td>Glu-63</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>3.1</td>
<td>Nh₁</td>
<td>Arg-62</td>
<td>3.1</td>
<td>Glu-63</td>
</tr>
<tr>
<td></td>
<td>Nh₁</td>
<td>3.2</td>
<td>O₂</td>
<td>Wat-455</td>
<td>2.7</td>
<td>Tyr-99</td>
</tr>
<tr>
<td></td>
<td>Nh₂</td>
<td>2.8</td>
<td>O₁</td>
<td>Wat-455</td>
<td>2.8</td>
<td>His-9</td>
</tr>
<tr>
<td></td>
<td>Oy₁</td>
<td>2.9</td>
<td>O</td>
<td>Wat-455</td>
<td>2.8</td>
<td>Wat-456</td>
</tr>
<tr>
<td></td>
<td>Oy₂</td>
<td>2.9</td>
<td>O</td>
<td>Wat-453</td>
<td>3.0</td>
<td>Wat-455</td>
</tr>
<tr>
<td>Ile-3 N</td>
<td>N</td>
<td>3.0</td>
<td>O₁</td>
<td>Tyr-99</td>
<td>2.7</td>
<td>Wat-456</td>
</tr>
<tr>
<td>Lys-4 N</td>
<td>N</td>
<td>2.8</td>
<td>O</td>
<td>Wat-455</td>
<td>2.8</td>
<td>Lys-163</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>3.1</td>
<td>O</td>
<td>Wat-457</td>
<td>2.6</td>
<td>Wat-457</td>
</tr>
<tr>
<td>Thr-7 O</td>
<td>O</td>
<td>2.9</td>
<td>O</td>
<td>Wat-453</td>
<td>3.0</td>
<td>Wat-454</td>
</tr>
<tr>
<td></td>
<td>Oy₁</td>
<td>2.9</td>
<td>O</td>
<td>Wat-453</td>
<td>3.2</td>
<td>Wat-454</td>
</tr>
<tr>
<td>Leu-6 O</td>
<td>O</td>
<td>2.8</td>
<td>Ne₁</td>
<td>Trp-147</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Lys-9 N</td>
<td>N</td>
<td>2.9</td>
<td>Oβ₁</td>
<td>Asp-77</td>
<td>2.0</td>
<td>Wat-451</td>
</tr>
<tr>
<td></td>
<td>OT₁⁻</td>
<td>2.6</td>
<td>O</td>
<td>Wat-451</td>
<td>2.7</td>
<td>Thr-80</td>
</tr>
<tr>
<td></td>
<td>OT₁⁺</td>
<td>2.7</td>
<td>O</td>
<td>Wat-451</td>
<td>2.0</td>
<td>Asp-77</td>
</tr>
<tr>
<td></td>
<td>OT₂⁻</td>
<td>2.6</td>
<td>O₁</td>
<td>Wat-452</td>
<td>2.9</td>
<td>Lys-146</td>
</tr>
<tr>
<td></td>
<td>OT₂⁺</td>
<td>3.2</td>
<td>Oy₁</td>
<td>Trp-84</td>
<td>3.0</td>
<td>Lys-146</td>
</tr>
<tr>
<td></td>
<td>NC</td>
<td>2.3</td>
<td>Oγ₂</td>
<td>Thr-143</td>
<td>2.7</td>
<td>Tyr-128</td>
</tr>
<tr>
<td></td>
<td>Nc</td>
<td>2.0</td>
<td>Asp-118</td>
<td>2.0</td>
<td>Wat-456</td>
<td></td>
</tr>
</tbody>
</table>

Hydrogen bonds (<3.2 Å) from peptide model RRIKAIILK to HLA-B27 and bound water molecules (target atoms). "Second-shell" hydrogen bonds are formed from target atoms to other atoms in the structure. Underscored: MHC side chains conserved in most class I molecules. Nomenclature for side-chain atoms described in Schulz and Schirmer (1979).

* OT₁⁻ and OT₂⁻: terminal peptide carboxylate oxygen atoms.

Figure 4. Peptide Contacts with HLA-B27

(A) Peptide (RRIKAIILK) solvent accessible surface area buried by HLA-B27 binding, using a 1.4 Å probe radius. The portion of peptide surface buried by residues conserved among most class I MHC molecules is shown in black.

(B) Number of hydrogen bonds formed between peptide and HLA-B27 atoms, as a function of peptide sequence number. Hydrogen bonds contributed by peptide main-chain and side-chain atoms are tabulated separately.

The binding potential of the peptide termini, and in particular of main-chain atoms, is almost fully exploited. Hydrogen bonds to the peptide involve 5 of the 8 possible main-chain proton donor or acceptor sites of P1–P2 and 6 of the 8 potential sites of P8–P9. The main-chain atoms of P1 and P2 are completely inaccessible to solvent when bound to HLA-B27, while those of P9 are 95% buried. The remaining 5% of the P9 main-chain surface is occupied by tightly bound water molecules that form hydrogen bonds to HLA-B27 as well.

The same measures that suggest a substantially shared conformation for all peptides also indicate that deviations are concentrated in the central few peptide residues, as
would be expected given their small number of binding interactions. Only for residues P4–P8 does the mean temperature factor of each residue exceed the peptide average of 27 Å², reaching a maximum of 37 Å² at P6. Peptide residues P5–P7 have the worst main-chain real-space correlation coefficients (0.76–0.79) against the iteratively real-space averaged electron density map calculated before introduction of any peptide model. Following refinement, and inclusion of the peptide model in map phases, the real-space fit for these residues improved to 0.84–0.87, but remains the lowest in the model. This is the part of the peptide that arches up away from the floor of the peptide-binding cleft, above water molecules that are observed to form a bridge between the peptide and the MHC molecule (see Figure 3B). The side chain of P7, a position of sequence variability that points downward into the cleft, is located at the edge of the solvent-filled volume under the peptide main-chain arch.

Seven HLA-B27 subtypes have been identified, and their sequences differ at 11 positions (López de Castro, 1989; Choo et al., 1991). Of these 11 side chains, 9 would contact the peptide model presented here and therefore may influence the selection of peptides that bind to a given subtype, as suggested by recent studies with allorreactive CTL recognition of different HLA-B27 subtypes (López et al., 1992; Villadangos et al., 1992).

Conserved Binding Sites for the Peptide Termini

At the two ends of the HLA-B27 cleft, there are binding sites, conserved in other class I molecules, for the termini of bound peptide nonamers. At the peptide N-terminus (Figure 5A and Table 1), the P1 carbonyl oxygen forms hydrogen-bonds to conserved Tyr-159, while the positively charged peptide amino group forms a pentagonal hydrogen bond network with conserved residues Tyr-7, Tyr-59, and Tyr-171, and with a bound water molecule (461) (Figure 5A). Water 461 forms two weak, long-range hydrogen bonds with Glu-63 (3.3 Å) and Glu-45 (3.4 Å) that may help to neutralize the buried positive charge on the peptide N-terminus. Disruption of this network by Phe substitutions for Tyr-7 or Tyr-171 greatly reduces CTL recognition of a known peptide antigen (Lalaton et al., 1992). Further hydrogen bonds are made to the main chain of peptide residues P2–P4 (Table 1). A similarly complex hydrogen bond network is centered on the peptide C-terminus (Figure 5B and Table 1). One peptide-terminal carboxylate oxygen forms hydrogen bonds to conserved residues Tyr-84 and Thr-143, and via Thr-143, to conserved Tyr-123. The other carboxylate oxygen hydrogen-bonds to two water molecules, one of which mediates hydrogen bonds to Thr-60 and Asp-77, while the second hydrogen-bonds to conserved Lys-146. Since Lys-146 is involved in a crystal contact to Asp-196 of an adjacent molecule, its position may differ from that in solution, where it could directly interact with the peptide. The hydrogen bond from the carbonyl oxygen of P8 to the pyrrole nitrogen of conserved Trp-147 is the only conserved hydrogen bond not made by P1 or P9 polar groups. Model-building experiments suggest that binding of the peptide carboxylate and penultimate carbonyl bond may orient the P9 side chain down into the cleft, explaining the frequency of a C-terminal anchor residue (Falk et al., 1991). In known HLA-B27 peptide sequences (Jardetzky et al., 1991), P9 shows a preference for basic side chains, but also tolerates hydrophobic side chains. Just beneath the P9 Cα, the surface of the pocket is contributed by hydrophobic atoms of side chains Leu-81, Tyr-123, and Thr-143 and can therefore accept either a hydrophobic P9 side chain or the hydrophobic portion of a longer basic side chain. The carboxylate groups of side chains Asp-74, Asp-77, and Asp-116 are clustered around the end of the pocket, positioned to neutralize the positive charge of a P9 Arg or Lys side chain. However, because these carboxylates can either form salt bridges with other MHC side chains or else contact solvent even in the presence of a bound peptide, a basic side chain at P9 may not be strictly required.

P2 Arginine-Binding Pocket

The P2 arginine side chain is located in a deep pocket between the floor of the cleft and the aβ helix that is also found in HLA-A2 and HLA-Aw68 near residue 45 ("45" or "B" pocket) (Garrett et al., 1989; Saper et al., 1991). Four polymorphic MHC residues (9, 24, 45, and 67) are positioned with side chains pointing into the end of the pocket. Together with a water molecule (456) bound in the site, His-9, Thr-24, and Glu-45 form a planar network of hydrogen bonds with the P2 arginine guanidinium group (Figure 6 and Table 1). The γ-sulfhydryl of Cys-67 is poised 3.6 Å above the plane of the guanidinium group, almost directly above its central carbon (Cε) (Figure 6).

The side chains that form the P2-binding pocket are shared by all seven known HLA-B27 subtypes (Buxton et al., 1992). In addition, the constellation of amino acids at positions 9, 24, 45, and 67 is not found in other sequenced class I histocompatibility antigens. Thirteen sequenced class I antigens other than HLA-B27 have Glu-45, but differ at 67 or 9; eight have Phe or Tyr-67; and five have Tyr-9. Model-building experiments indicate that these changes at 9 or 67 would alter the properties of the site. Phe-67 or Tyr-67 would substantially fill the P2-binding pocket, leaving in its place a small hydrophobic pocket and blocking the binding of longer side chains, such as arginine. Replacing His-9 with a tyrosine side chain would displace a water molecule (456), which hydrogen-bonds to the P2 Nε1 atom, as well as to His-9, Tyr-99, and water 458 in a tetrahedral arrangement (Figure 6). Although this substitution is less dramatic than the Cys-67 replacements, the Tyr-9 hydroxyl group appears unable to form a hydrogen bond to the P2 arginine with acceptable geometry (C-O...N angle > 90°) to replace that made by water 456. Supporting the idea that Tyr-9 would affect peptide binding, Buxton et al. (1992) find that a His-9-Tyr site-directed mutant of HLA-B27 shows dramatically reduced binding of an influenza A nucleoprotein (383–394) decaamer peptide (Huett et al., 1990). Determination of peptide-binding motifs for HLA molecules that share Glu-45 with HLA-B27, but have Tyr-9, Phe-67, or Tyr-67 (instead of His-9 and Cys-67) would test the proposed uniqueness of the HLA-B27 P2 arginine anchor. Also located in the vicinity of the P2-binding pocket is the side chain of Lys-70,
the single side chain common to the seven known HLA-B27 subtypes and not found in the other sequenced class I molecules (Choo et al., 1991). It is oriented away from the pocket, forming a salt bridge with Asp-74.

**Discussion**

A given class I MHC protein is able to bind and present to CTLs a diverse population of peptides of widely varying
sequence. The key to both the flexibility and the long half-life of this interaction appears to lie in the concentration of peptide–MHC binding energy at the ends of the cleft, involving the peptide main chain and only a few peptide side chains. Extensive van der Waals contacts and hydrogen bonds are formed between HLA-B27 and the main chain of 5 peptide amino acids (P1–P3 and P9–P8) (Figures 4 and 5, Table 1), accounting for much of the overall peptide–MHC interaction. Conversely, the binding potential of these few peptide atoms is almost completely exploited, with most of their solvent-accessible surface area buried and three quarters of all possible hydrogen bonds made. These contacts do not depend on the peptide sequence, except that it must be possible for the peptide to adopt a conformation placing its termini in the appropriate binding sites. Supplementing the peptide-terminal binding interactions, and perhaps also important for the correct orientation of the peptide termini, are contacts of mostly polymorphic MHC side chains with a few peptide (anchor) side chains, which therefore can vary among different class I molecules. In HLA-B27, side chains P4–P6 and P8 appear to make minor contributions to the total binding energy, so that peptide binding should be relatively indifferent to their identities, as long as they can be sterically accommodated. Since these side chains all point toward the solvent, even the steric constraint is probably minimal.

The clarity of the main-chain electron density found for a nonameric peptide in the HLA-B27 peptide-binding cleft indicates that nonamers dominate the collection of endogenous peptides presented by HLA-B27 and also that most of these nonamers adopt similar main-chain conformations, despite sequence heterogeneity. It is a theoretical possibility that some HLA-B27 molecules, bound to particular peptides, could have been excluded from crystallization by unfavorable lattice contacts and therefore prevented from contributing to the composite peptide electron density. In this case, however, no crystal-packing interactions involving the peptide are observed, and the volume directly above the bound peptide is occupied by solvent, so that the observed image probably reflects a representative sampling of the peptide population. Furthermore, as seen in Figure 2, the details of the composite image are those expected for a collection of peptides defined by elution and sequencing (Jarcietzky et al., 1991). The predominance of a standard nonamer main-chain conformation within this population suggests peptide sequence variation can be presented to TCRs directly through side-chain shape and chemistry. In contrast, indirect presentation of peptide sequence differences, either by adjustment of the peptide main-chain conformation, or by adjustment of the conformation of MHC side chains, appears to have a limited role in HLA-B27, given the well-defined electron density observed for these features of the peptide-binding cleft. Possibly stabilizing the overall conformation of the peptide is a cluster of water molecules bound between H1 A R97 and the peptide main chain in the center of the cleft. These bound water molecules could presumably be displaced in some cases by peptides that cannot be accommodated in the standard "consensus" conformation. Adjustments in the center of the cleft are observed for other class I molecules as a result of peptide length variation and the presence of central peptide anchor side chains (see below). Even for a given class I molecule, the amount of peptide length and sequence variability may reflect, in addition to the requirements of MHC binding, proteolytic and transport processes (DeMars and Spies, 1992), as sug-
ggested by experiments with the rat class I MHC molecule RT1-A* (Powieis et al., 1992).

Since the binding interactions of nonamer peptides are concentrated at the peptide termini, longer peptides capable of tight MHC binding are likely to conserve these interactions by accommodating increased length as "insertions" between the termini, rather than losing these interactions by "extensin" out the ends of the binding cleft. This would conserve the bound positions of the charged peptide termini in the conserved sites evolved for that purpose and would concentrate the structural differences of the peptides in the likely region of TCR binding. Heterogeneity in the length of optimal peptide binding HLA-B27 has been observed, with the identification of a decamer minimal epitope for human immunodeficiency virus GAG p24 265-274 (sequence: KRWILGLNK) (S. Rowland-Jones and A. McMichael, personal communication). This peptide has Arg at P2 and has Lys at its C-termi-

nus, matching the HLA-B27 C-terminal preference for basic amino acids (Jardetzky et al., 1991). Sequences of longer (>9 mer) peptides eluted from class I molecules expressed in mutant cell lines (Henderson et al., 1992; Wei and Cresswell, 1992) often contain anchor residues appropriate to the given class I molecule and correctly spaced relative to the N- and C-termini (but separated by extended central regions), as would be expected for an insertion model. The structure of HLA-Aw68, cocystal-

lized with a heterogeneous population of bound peptides, shows electron density that is clear at the peptide termini but weak for the central peptide residues, and peptides eluted from this class I antigen show greater length heterogeneity than was seen for HLA-B27 (H.-C. Guo, T. S. Jardetzky, W. S. Lane, J. L. S., and D. C. W., submitted; Jardetzky et al., 1991). Finally, length accommodation in the center of the cleft is also seen in H-2K, where an octamer and a nonamer peptide bind with essentially the same N- and C-terminal interactions (Fremont et al., 1992; Matsumura et al., 1992). In H-2K, a central bulge of 1 residue is observed even though bound peptides have a Tyr anchor in the center (PS in an octamer) (Falk et al., 1991).

The ability of bound peptide to stabilize the three-

dimensional fold of class I molecules (Townsend et al., 1989) appears to depend on the peptide termini, "knitting together" conserved residues on the two α helices and the β sheet that form the cleft (Figure 5). For example, the first peptide amino acid forms hydrogen bonds with 4 conserved residues: 2 on the α2 helix (Tyr-159 and Tyr-171), 1 on the β sheet floor of the cleft (Tyr-79), and (through Tyr-171) 1 on the α1 helix (Tyr-59) (Figure 5A). Similarly, at the C-terminus of the peptide, the terminal carbohydrate forms hydrogen bonds with conserved residues of both α helices forming the sides of the cleft (Tyr-24, Thr-143, and Lys-140) and the β sheet (Tyr-123, via Thr-143). The penultimate peptide residue makes one hydrogen bond to a conserved residue on the α2 helix (Trp-147) (Figure 5B). Peptide contacts with polymorphic MHC residues also involve both helices and the sheet (Table 1). Furthermore, the first few and last few peptide residues are almost completely enfolded by HLA-B27 (Figure 4). As a result, these peptide amino acids participate in the three-dimensional structure of HLA-B27 as completely as most heavy chain or β2m residues.

If the disease association of HLA-B27 does require the presentation by HLA-B27 of a proposed "arthritogenic peptide" (Benjamin and Parham, 1990), then it is possible that the specificity of the association will involve the P2-

arginine-binding pocket. The potentially unique and absolute requirement for arginine at this position appears to be determined by the presence of a planar network of hydrogen bond acceptors at the end of this pocket (including the negatively charged side chain of Glu-45) that would complement an arginine guanidinium group, but not a tetraheiral lysine ε amino group. This arginine-binding network is common to the known HLA-B27 antigens and is not found in the sequences of other known class I human antigens. If HLA-B*2703 is not associated with ankylosing spondylitis, as speculated by Hill et al. (1991), then this may be a result of changing the otherwise conserved MHC residue Y59 to histidine, which alters subtype-specific CTL recognition (Calvo et al., 1990; Villadangos et al., 1992).

The conservation of many MHC side chains that contact main-chain atoms at the ends of the peptide suggests that these interactions, and their implications for peptide presentation (discussed above), represent a general phenomenon. Despite extensive polymorphism in the HLA peptide-binding site, similar binding to the main chain of the first few and last few peptide residues is likely to be observed in many class I molecules. MHC interactions with the termini of peptides bound to HLA-Aw68 (M. L. Silver, H.-C. Guo, J. L. S., and D. C. W., submitted; H.-C. Guo, T. S. Jardetzky, W. S. Lane, J. L. S., and D. C. W., submitted) and single peptides bound to H-2K (Fremont et al., 1992; Zhang et al., 1992) are similar to those seen for HLA-B27. For peptide side chains, by contrast, few conserved interactions can be identified, and the binding of peptides to other class I antigens with different anchor positions (Falk et al., 1991) may cause structural adaptations in addition to those already observed.

**Experimental Procedures**

**Protein Purification and Crystallization**

HLA-B27 was purified from LC-2 cells as described (Turner et al., 1979; Parham et al., 1977; López de Castro, 1984), except that prior to papain cleavage, protein was released from cell membrane by incubation for 30 min in the standard hypotonic lysis buffer (200 μM dithiothreitol, 100 μM phenylmethylsulfonyl fluoride [PMSF] ~ 10 mg/ml in ethanol), 10 mM Tris [pH 8.0] supplemented with 4% deoxycyclate. After high performance liquid chromatography gel filtration purification was performed (S7800 column, 100 mM Na,[N-morpholino]ethanesulfonic acid [MES], 0.1% sodium azide [pH 6.5]). Fractions were concentrated in a Centricron 30 filter to a final concentration of 4.3 mg/ml in 25 mM MES, 0.1% sodium azide (pH 5.5).

Microseeds for crystallization were prepared by serial dilution of crystal HLA-B27 crystals (Gorga et al., 1992) in 24% (w/v) polyethy-

lene glycol 3350, 250 mM sodium acetate, 100 mM Tris, 25 mM MES, 0.1% sodium azide (pH 8.5). Hinging drops (2 μl of microseeds and 2 μl of protein solution) were equilibrated by vapor diffusion as described (Gorga et al., 1992). Crystals were harvested to modified molybdenum well solution (25% versus 24% polyethylene glycol 3350) in 50 μl dialysis buttons, equilibrated at room temperature for 12 hr against harvest buffer containing 16% glycerol as cryoprotectant (Tang, 1990), and then equilibrated for 6 hr against harvest buffer with 20% glycerol.
Table 2. Data Quality and Refinement Statistics

<table>
<thead>
<tr>
<th>Resolution (Å)</th>
<th>R_{	ext{merge}}</th>
<th>Fraction Complete</th>
<th>F_{c} &gt; 0.6σ</th>
<th>F_{c} &gt; 3σ</th>
<th>R_{	ext{free}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0–5.50</td>
<td>0.070</td>
<td>0.95</td>
<td>0.363</td>
<td>0.359</td>
<td>0.389</td>
</tr>
<tr>
<td>5.50–4.75</td>
<td>0.050</td>
<td>0.95</td>
<td>0.169</td>
<td>0.168</td>
<td>0.248</td>
</tr>
<tr>
<td>4.73–3.73</td>
<td>0.067</td>
<td>0.95</td>
<td>0.147</td>
<td>0.144</td>
<td>0.206</td>
</tr>
<tr>
<td>3.75–0.01</td>
<td>0.070</td>
<td>0.94</td>
<td>0.104</td>
<td>0.102</td>
<td>0.251</td>
</tr>
<tr>
<td>3.31–3.01</td>
<td>0.079</td>
<td>0.93</td>
<td>0.160</td>
<td>0.160</td>
<td>0.205</td>
</tr>
<tr>
<td>3.01–2.77</td>
<td>0.091</td>
<td>0.93</td>
<td>0.211</td>
<td>0.201</td>
<td>0.287</td>
</tr>
<tr>
<td>2.77–2.50</td>
<td>0.110</td>
<td>0.94</td>
<td>0.228</td>
<td>0.215</td>
<td>0.292</td>
</tr>
<tr>
<td>2.50–2.24</td>
<td>0.131</td>
<td>0.94</td>
<td>0.242</td>
<td>0.224</td>
<td>0.308</td>
</tr>
<tr>
<td>2.44–2.13</td>
<td>0.157</td>
<td>0.93</td>
<td>0.264</td>
<td>0.231</td>
<td>0.296</td>
</tr>
<tr>
<td>2.31–2.00</td>
<td>0.100</td>
<td>0.91</td>
<td>0.260</td>
<td>0.250</td>
<td>0.300</td>
</tr>
<tr>
<td>2.20–2.10</td>
<td>0.243</td>
<td>0.81</td>
<td>0.277</td>
<td>0.243</td>
<td>0.310</td>
</tr>
<tr>
<td>1.50–1.20</td>
<td>0.078</td>
<td>0.92</td>
<td>0.203</td>
<td>0.191</td>
<td>0.287</td>
</tr>
<tr>
<td>1.00–0.20</td>
<td>0.077</td>
<td>0.92</td>
<td>0.229</td>
<td>0.218</td>
<td>0.286</td>
</tr>
</tbody>
</table>

Symbols: R_{	ext{merge}} = (Σ |∥∥-<∥>∥∥)Σ |∥∥, R_{	ext{free}} = (Σ |∥∥-<∥>∥∥)Σ |∥∥, h.c. (refinement reflections), R_{	ext{merge}} = (Σ |∥∥-<∥>∥∥)Σ |∥∥, h.c. (refinement reflections), F_{c} > 0.6σ, F_{c} > 3σ, and F_{c} observed and calculated structure factor amplitudes, σ_{F}, standard deviation on F.

Two different papain cleavage protocols (directly from cell membranes [Turner et al., 1975; Parham et al., 1977] or from detergent micelles [Gorga et al., 1987, 1992]) have been used to prepare HLA antigens for crystallization. Cell membrane cleavage of HLA-B27 yielded a heavy chain fragment with 270 or 271 residues (Esquerra et al., 1985), and such material was used in peptide elution experiments (Juretzky et al., 1985). Detergent solubilization before papain cleavage produces material slightly larger by gel electrophoresis (data not shown). Only the detergent preparation yielded diffraction quality HLA-B27 crystals, and electron density is found in these crystals for residues 271–276, forming the seventh strand of the standard immunoglobulin fold for α, and stabilizing a crystal contact.

Data Collection and Processing

A single crystal was mounted in a thin film of harvest buffer, supported by a wire loop. 1 mm in diameter and made of 75 µm diameter nichrome wire (Teng, 1990), and was quickly frozen in -165°C nitrogen gas cooled by a modified Nicotellim Series LT-1 cryostat (Blum, 1980). Data were collected on a X440 area detector (Durbin et al., 1990) using CuKa radiation generated by an Elliot GX-13 (40 Kβ, 60 mA) at a crystal to detector face distance of 103 mm. Collection proceeded for 15 minutes without a noticeable decrease in crystal diffraction power. Structure factor amplitudes were calculated using BUIJDA data reduction software (Blum et al., 1987) and CCP4 programs ROTAVATA/AO/NXVATA (Fox and Holmoc, 1985) and TRUNCATE (Fonset and Wilsen, 1978). The space group is triclinic P1, with two molecules in the asymmetric unit and unit-cell parameters a = 38.1 Å, b = 68.8 Å, c = 81.1 Å, α = 80.3°, β = 86.8°, γ = 89.9° (determined by periodic refinements during data integration). These unit-cell lengths are 1.7%–3.1% shorter than those observed at room temperature (Gorga et al., 1992). Data quality and completeness as a function of resolution are shown in Table 2. A randomly selected 10% of all data were excluded from automated refinement (but included in map calculations) and used to compute a free R factor (R_{free}) (Brünger, 1992) throughout model building and refinement. A conventional crystallographic R factor (R) was computed for reflections included in refinement. No standard deviation cutoff was applied to the refinement set data until the final stage of refinement at 2.1 Å.

Model Building

Starting models were obtained by molecular replacement, searching 20–4 Å data with a model of HLA-A2 (Saper et al., 1991), using the package MDRIT (Fitzgerald, 1988). The two strongest rotation function solutions (Rossmann and Blow, 1982) (7.9° and 7.7°) are related by a monoclinic 2-fold axis close to the α, consistent with the observed pseudo-monoclinic space group (and with a 16.1° self-rotation function solution at Ψ = 90°, φ = 180°, K = 180°). A translation function search (Crouch and Blow, 1967) yielded a single clear solution (0.4°) for the displacement between the two molecules in the crystallographic asymmetric unit (τ = 0.025, 0.205, 0.407). The final solution was rigid body refined (8 domains) (R_{final} = 42.1%) against 10–3 Å data using the program X-PLOR (Brünger, 1989).

To obtain a molecular model of HLA-B27 as independent as possible of the HLA-A2 structure, we calculated 10–3 Å single-domain omit maps. A single domain (a, α, β, γ, or δ) of roughly 60 amino acids was omitted from both of the molecules in the asymmetric unit, and the resulting model was used to provide starting phases for iterative real-space averaging ("phase averaging") of electron density maps by noncrystallographic symmetry (Bricogne, 1976). An atomic model was built for the parts of the omitted domain where clear density was found, using FREEO (Jones, 1972). The process was repeated for each domain. No model for bound peptide was included until the final round of refinement at 2.1 Å. Eighty-nine percent of the atoms in the structure were located in the first set of phases. Successive cycles of X-PLOR positional and simulated annealing refinement, phase averaging, and manual rebuilding continued until the model was 97% complete, with all main-chain atoms located and 36 of 75 side chains omitted. R_{final} was 20.0% and R_{merge} was 37.6% (10–3 Å). In the final stage of model building, the resolution range was 6–2.7 Å, and restrained individual atomic temperature factor refinement was carried out. R_{merge} = 22.9% and R_{final} = 34.9% (6–2.7 Å).

Model Refinement

Resolution was extended in several steps to 2.1 Å. Each extension was performed with manual refining of the model against standard difference electron density maps and iteratively real-space averaged maps calculated at the new resolution. This was followed by automated refinement using positional, simulated annealing and restrained individual atomic B factor refinement with X-PLOR. R_{merge} improved throughout resolution extension, even at stages where R_{merge} increased, presumably owing to a reduction in the possibility of "overfitting" (Brünger, 1992). R_{merge} was also used to monitor the usefulness of individual refinement procedures. At 6–2.5 Å the model was complete (R_{merge} = 21.6%, R_{final} = 33.6%). At 6–2.3 Å (R_{merge} = 22.9%, R_{final} = 35.8%) the first solvent molecules were identified, using four criteria: peak electron density greater than four times the RMS variation in electron density in an (F_{o} – F_{c}) map, an electron density map density; presence of hydrogen-bonding partners; not density for an alternate side-chain conformation or part of the peptide electron density; and the presence of a second peak, meeting criteria 1–3 and related to the first by noncrystallographic symmetry. At 6–2.1 Å, with 130 water molecules, R_{merge} = 22.5% and R_{final} = 30.3%. Water molecules were subsequently removed by manual refinement. The noncrystallographic symmetry requirement for solvent molecules was relaxed for 4° R_{merge} peaks, and the threshold for noncrystallographic symmetry-related peaks was reduced to 8° R_{merge}. Two cutoffs were applied to water molecules based on refined temperature factors, above 50 Å² they were removed from the structure, and above 50 Å² they were required to have clear electron density in phase averaged electron density maps. In the final stages of refinement, the map was changed to 5.5 Å, and reflections with structure factor amplitudes F < 3σ_{F} (standard deviation of F) were excluded from X-PLOR refinement calculations.

In addition to difference density maps, the program O (XAV version 5.6 and SGI version 5.7) (Jones et al., 1991) was used to monitor the quality of the model. PROLP identified 2 residues as candidates for peptide bond inversion. One bond (I) was inverted, but the other (29) was not, despite a nonstandard main-chain conformation (φ > 0, Ψ < 0), based on omit map calculations and the identification of a similar conformation (Marquart et al., 1989) in a data base of 40 high resolution refined structures (Jones and Thirup, 1989). The conformations of HLA-B27 side chains were compared with standard side-chain rotamers, using the RSC utility in O and a locally modified library of side-chain conformations containing all standard rotamers (Ponder and Richards, 1987). A few residues, mainly Leu, were adjusted (within the side-chain electron density) to standard conformations. Five side chains in the final model have an RMS deviation from the closest rotamer of more than 2 Å, excluding poorly sampled Thr and Met side.
chains. A third evaluative function, the real-space correlation of the atomic model against electron density maps, was used to identify regions of poor density and/or poor model fit. Residues by residue averages for both side-chain and main-chain atoms were calculated. Residues with a real-space fit of less than 0.8 were inspected using FRIDO. Several residues were rebuilt, and side chains reoriented, followed in each case by refinement of peptide geometry using the REFI facility (Hermans and McQueen, 1974). In a few cases, subsequent positional refinement of the model (X-PLOR) restored the poorer conformation even in regions of clear electron density. In the final stages of refinement, harmonic constraints were removed and all residues were subjected to such refinement. The final real-space fit for protein main-chain atoms was 0.31 and for side-chain atoms, 0.37.

Peptide Model
No model was included for bound peptide until the last stage of refinement and refitting at 2.1-Å resolution. Rmin (6.5-2.1 Å) was 29.6% and Rfree (0.2-1.0 Å) was 30.6%, with 290 water molecules included in the model. Phase-averaged maps showed clear, continuous peptide main-chain density and at most positions side-chain and peptide carbonyl group orientation (Figure 2). Unaveraged electron density maps (15-2.1-Å resolution) showed nearly continuous density in a significant portion. A peptide model was built with sequence ARAAAAAAA and refined using X-PLOR. Rmin dropped by 0.5% before and a further 0.4% following positional and B factor refinement. Additional water molecules were identified in the peptide and several rounds of X-PLOR positional and atomic temperature factor refinement were performed. Water molecules were removed if their refined temperature factors exceeded 50 Å², and additional water molecules were identified. Electron density for the peptide main chain at P5-P7 improved slightly following refinement, but was worse for side chain P1. A second peptide model (sequence: RIKIATTLK) was also built, using the most common side chain found at each position among sequenced peptides and resolving ambiguities by reference to the electron density. This model was also refined. No crystal contacts were made to either peptide model. The closest approach of any symmetry-related atom to the RHIKATTLK model is a water molecule (110G) located 5.1 Å away from the end of the fully extended P1 Arg side chain.

The final model includes 8268 non-hydrogen protein atoms (708 residues) and 450 water molecules. Each molecule in the asymmetric unit consists of residues 1-373 of the HLA-B27 heavy chain, 50 residues for βm and 9 peptide residues (ARAAAAAAA). No density was found for side chains of three HLA-B27 residues (106, 266, βm-80), whose occupancies have been set to 0, nor for the single bound carbohydrate moiety, which has therefore not been modeled. Fifteen water molecules are in each peptide binding site. Agreement of observed and calculated electron density is reported in Table 2. The RMS deviation for bond lengths is 0.01 Å and for angles, 2.8 Å. The average atomic temperature factor is 23.9 Å² for HLA-B27 and 27.0 Å² for the peptide. The RMS difference in temperature factors between bound and unbound atoms is 0.06 Å for the main chain and 1.41 Å for side chains. The coordinate error, as estimated by the method of Luzzati (1952), is 0.25 Å based on Rmin or 0.36 Å based on Rfree. The RMS difference between main-chain atoms of the two HLA complexes in the asymmetric unit is 0.26 Å, and for all atoms is 0.37 Å, and the two molecules can be considered to have the same structure. Solvent-accessible surface area (Lee and Richards, 1971) were calculated using the program ACCESS, written by M. D. Handschumacher and F. M. Richards. van der Waals contacts were also computed with the program AUCAS. MHC atoms were considered to contact the peptide if their accessible surface area differed in the presence or absence of peptide atoms. A 0.01 Å radius probe was used, so that the MHC-accessible surface would approximate the contact surface and would only differ if a peak if the contact surface of a peptide atom came within 0.01 Å of it. Standard van der Waals contact radii were increased by 23% (e.g., carbonyl carbons were increased to 2.1 Å, instead of 1.7 Å). Contacts at this larger distance contribute 50% of the maximum possible interaction energy (which occurs at the standard separation). Hydrogen bond contacts were calculated using the CCP4 program CONTACTS, written by T. Skarzynski. Least-squares superpositioning of coordinate sets used the subroutine OCMFIT, written by P. J. Bernstein (Kabsch, 1978; McLachlan, 1979). Figures 3, 5, and 6 were produced using HYDRAGRAPH.

Acknowledgments
The authors would like to thank Anastasia Haykov, Kristine Svenson, Mia Frayer, Polina Klimovitsky, Constance Barnett, and Richard Crouse for excellent technical support. This work was supported by the National Institutes of Health (D. R. M. and J. L. S.), the Howard Hughes Medical Institute (O. W.), the Medical Foundation/Children's King Trust Fund (J. C. G.), and the Pittsburgh Supercomputer Center.

Coordinates will be deposited with the Brookhaven Protein Data Bank. Until they are processed, they will also be available by e-mail (madison@xtal.harvard.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC Section 1734 solely to indicate this fact.

Received July 20, 1992; revised August 3, 1992.

References


