Graphical representations of the class I MHC cleft

James L. Cornette,* Benjamin L. King,† Michael D. Silverman,† and Charles DeLisi†

**Department of Mathematics, Iowa State University, Ames, IA USA †Biomolecular Engineering Research Center, Department of Biomedical Engineering, Boston University College of Engineering, Boston, Massachusetts, USA*

We describe computer graphics and computer aided manufacture of three-dimensional models designed specifically to elucidate the cleft in the class I human leukocyte antigen. The models evolve from computer graphical representations and provide a geometrically and chemically concise and detailed view of the antigen binding site. The techniques provide a new approach to representations of binding sites. The model provides sufficient detail to support binding specificities analysis of active sites involved in protein and DNA binding.

Keywords: protein binding; DNA binding; major histocompatibility complex

INTRODUCTION

Among the crucial regulatory elements of the mammalian immune response is a network of molecules encoded in the major histocompatibility gene complex (MHC), including the class I and II MHC products. Class I molecules are found on the surface of virtually every cell type, playing a central role in regulating the levels of cytotoxic and other CD8 T cells directed against infection. The closely related class II products are confined to immune cells and regulate the activity of T helper and other CD4 cells.[1,2] Under normal conditions, class I and II both migrate from the cell surface to its interior, where they bind to proteolytically digested protein fragments having approximately 9 (for class I)[3] and 13 to 17 (for class II)[4] amino acids. The complexes then cycle to the surface where, if the peptide is from a foreign protein, they are recognized by T cells of the appropriate type. Thus the binding site of MHC molecules and the structure of the bound segment are of extraordinary interest and importance in understanding the immune system.

The structures of two alleles of the class I antigen have been determined by X-ray crystallography and placed in the Brookhaven Protein Data Bank.[5-9] Both structures reveal a well-defined cleft that is considered to be the binding site of antigenic peptides, bounded on two sides by α-helices and below by a β-sheet.

A major theme of the descriptions of the clefts is the presence of pockets on its floor and sides, presumed to be locations of side chains of bound antigenic peptides.[7,10] The physical and chemical characteristics of the cleft determine which peptides will bind to the MHC molecule. These properties are the subject of intense study, and we have developed both computerized and three-dimensional acrylic modeling techniques directed specifically at their quantitative representation. We expect the techniques to be generally useful for representing active sites of proteins, and for facilitating drug design.

The core of our process is to locate a 0.2-Å mesh grid surrounding the cleft, with one axis perpendicular to the β-sheet and another axis as the long axis of the cleft. Each grid point is then identified as being inside an atom of the class I molecule, potentially inside a water molecule that does not intersect any atom of the class I molecule, or neither of the two. The grid points that are potentially inside a water molecule form the *water accessible volume,* and a component of that volume is the cleft region available to a peptide. The surface of that component marks the interface between peptide and class I molecule, and the class I atoms on and near that surface determine the binding potential.

We have developed graphics that display cross sections of the cleft with the identities and positions of class I atoms near the cleft clearly marked. Using the grid points in the surface of the cleft, we display two-dimensional maps of the surface with various atom characteristics and locations on the surface. These cross sections and surface pictures are particularly useful in comparing different structures. The class I molecules are very similar in backbone conformations, so that they and their clefts can be easily aligned, although the clefts have notably different geometry and surface features.

Our ultimate construction has been the use of computer aided manufacturing to cut outlines of the cross sections in 2-mm plastic sheets that we assembled into a novel three-dimensional "hard copy" model of the cleft. This model is relatively easy to construct, very accurate in its location of atoms, and provides insights not seen from a previously constructed ball-and-stick model of the class I molecule, or from space-filling models.

Color Plates for this article are on page 187.

Address reprint requests to Dr. Cornette at the Department of Mathematics, Iowa State University, Ames, IA 50011, USA.

Received 20 August 1992; accepted 2 February 1993

MATERIALS AND METHODS

We have used the Brookhaven Protein Data Base[8,9] entries 3HLA (human leukocyte antigen-A2, HLA-A2)[5,6] and 2HLA (HLA-Aw68)[7] which differ from each other in 12 of the first 182 amino acids (which encompass the binding cleft), 11 of which have atoms exposed to the surfaces of at least one of the clefts. All of our graphics have been developed on Silicon Graphics Iris work stations.

Coordinate System

For both class I molecules, we locate a rectangular *XYZ* coordinate system in the cleft based on certain α-carbon coordinates. Generally, the *Y*-axis is perpendicular to the β-sheet underlying the cleft, and the *Z*-axis is the long axis of the cleft. Specifically, compute the centroid, C_1, of the α-carbons of residues 50–84 (contains the α_1 domain helices on one side of the cleft), and the centroid, C_2 of the α-carbons of residues 138–173 (contains the α_2 domain helices on the other side of the cleft). The origin of an initial coordinate system is midway between C_1 and C_2. Compute the plane that is a least-squares fit to all the α-carbons in residues 50–84 and residues 138–173. The *Y*-axis is perpendicular to this plane, oriented so that the α-carbon of the β-sheet residue 97 has negative *Y*-coordinate. The *Z*-axis is chosen so that the *ZY* plane is parallel to the line joining the two centroids of, respectively, the α-carbons of residues 60–64 and the α-carbons of residues 78–82, and oriented so that the α-carbon of residue 82 has positive *Z*-coordinate. The *X*-axis is chosen to form a right-handed system. From this initial coordinate system, translate the origin to the point (0.0, 0.0, −14.0) to obtain a coordinate system well oriented and located with respect to the clefts in both HLA-A2 and HLA-Aw68. The origin is at the "left" end of the cleft, near CE3 of Trp 167 in both molecules, and the point (0.4, 0.0, *z*) between $z = 2$ and $z = 30$ lies entirely within the cleft of both molecules. Cross sections of the cleft are labeled according to their *Z*-coordinate, so that cs5, for example, shows the $z = 5$ cross section. The use of the averages of α-carbon coordinates makes location of this coordinate system robust to small perturbations of atom coordinates between different class I molecules, thus providing a common reference frame for comparisons of these structures.

We translate and rotate the coordinates of the Protein Data Bank files to these coordinates, and use a 0.2-Å mesh grid covering $-12.0 \le x \le 12.0$, $-10.0 \le y \le 5.0$, and $0.0 \le z \le 32.0$. Our programs use atomic radii as given either in ACCESS, the program of T. J. Richmond which computes accessible surface area of a molecule,[11] or the molecular modeling program CHARMm[12] (the extended atom radii of version 21 in Quanta, the molecular graphics program of Polygen Corporation[13]). The CHARMm radii are about 10% larger than that used by Richmond and produce a slightly smoother surface of the cleft. The figures shown here use the larger radii.

Water Accessible Volume

To identify the water accessible volume of the cleft, and the cleft surface, we first mark as atom points the grid points whose distance from the center of an atom in the oriented molecule is less than the corresponding atomic radius. For future graphics use, the marks are coded to reflect the atom type. At each of the remaining grid points, a sphere of radius 1.4 Å (a water molecule) is centered, and if the sphere does not intersect any atom, every grid point within or on the sphere is marked as a potential cleft point. With this procedure, there are a few grid points marked as potential cleft points that are not between the two helices. In order to exclude these points, we define the cleft to be the component of the potential cleft points that contains (0.0, 0.0, 14.0) (which is midway between the centroids of the two helices and is clearly a cleft point). The cleft contains just those grid points that can be connected to (0.0, 0.0, 14.0) with a chain of adjacent potential grid points.

The surface points of the cleft are next to be distinguished, being those cleft points for which an adjacent grid point is not a cleft point. For manufacturing the three-dimensional model, the surface points are distinguished within each constant *z* cross section with respect to only the points in that cross section (a surface point is a cleft point for which there is an adjacent grid point in the same cross section that is not a cleft point).

Following the lead of Madden et al.,[14] we also distinguished the cleft points that lie within a 3-Å sphere that does not intersect any atom. The points of the cleft *not* inside any of these larger spheres may be considered to be potential pocket points. Some side chains of a peptide bound in the cleft are presumed to fit into the more confined regions of the cleft, surrounded closely by class I atoms, and would not be in the more open regions marked by the 3-Å spheres. We find that some sections of the potential pocket points are too shallow to accommodate side chain atoms, and define pocket points to be potential pocket points lying within a 0.6-Å sphere of potential pocket points.

Both 3HLA and 2HLA data files give locations of water molecules. We mark for later graphics use the grid points within those water molecules. Mostly those molecules are buried within the protein; only three (in HLA-A2) of the water molecules are in the water accessible volume we identify as cleft.

Using a Silicon Graphics Iris we then convert the grid data to pictures of cross sections of the molecule perpendicular to the *Z*-axis. Any atom with center within a distance $D = 0.63$ Å of a cross section is labeled at the center of its intersection with the cross section. Because $D > 0.5$, every atom within the grid is marked on at least one integer-valued *Z*-cross section. We found that values of $D > 0.63$ caused too much overlapping of labels.

Cross Sectional Traces: Three-Dimensional Model and Maps of the Cleft Floor

We trace the surface of the cleft in any *Z* cross section, in the general direction from negative *x* to positive *x*, and write to a data file the "trace sequence" of *XY* coordinates of the surface points in that cross section. This data is used in two ways. First it is sent to a computer that controls a three axis computer numerically controlled milling machine (TMC 1000, Light Machines Corporation, Manchester, NH). Using this machine, we cut 2 mm × 15 cm × 24 cm acrylic

sheets along the trace sequence, maintaining a 1 cm = 1 Å scale in three dimensions. These sheets were then assembled in packets of 5. A color print of the appropriate integer cross section was placed between slices 1 and 2 of the packet to show the identity and locations of atoms surrounding the cleft. The five slices in each packet were held together with four plastic pegs, 1 cm long and 3/16 in diameter, inserted into holes drilled through the five slices and held in place with a two-part resin, Weld-On 40™. The resulting 32 packets were assembled into a complete model of the cleft. The program for cutting slices 2, 3, 4, and 5 of each packet of five included instruction to cut two 1.125 inch diameter circles towards the left and right edges of the slices. After assembly of the packets, two 0.25-inch thick, 1-inch diameter ceramic magnetic discs were placed in the resulting void space. These magnets are quite strong and firmly hold the 32 packets together and in alignment. However, the model also may be separated between any two packets in order to examine the cleft at that cross section.

The trace sequences were also used to map the floor of the cleft. The point of each trace for which $x = 0.4$ Å (with smallest Y-coordinate if there were two or more) was chosen as a base point. (We used $x = 0.4$ because the interval of points $(0.4,0.0,z)$, with $0 \leq z \leq 32$ lies entirely within the clefts of both HLA-A2 and HLA-Aw68.) Then the points of the cleft surface were assigned two-dimensional coordinates (s,t); s is simply z; t is essentially the distance along the curve, and specifically the location of the point in the trace relative to the base point, with negative t indicating a point that precedes the base point. The st coordinate system facilitates mapping of the electrostatic, hydrophobic, hydrogen bonding, concavity/convexity characteristics of the surface of the cleft.

To display the pockets of the cleft, we mapped the concavity/convexity of the floor of the cleft as follows. Choose a point (s_0,t_0) in the floor of the cleft. In the $(s_{-1} = s_0 - 1)$ trace, find the point with t-coordinate (t_{-1}) closest in XY coordinates to (s_0,t_0). In the $(s_{-2} = s_{-1} - 1)$ trace, find the point with t-coordinate (t_{-2}) closest in XY coordinates to (s_{-1},t_{-1}). Continue in this fashion, and obtain t-coordinates $t_{-5}, t_{-4}, \cdots, t_{+5}$. Align the 11 traces surrounding (s_0, t_0) so that t_i is at the same level as t_0, $i = -5,5$, and identify an 11 point-by-11 point patch of the surface surrounding (s_0, t_0). Fit by least squares a quadratic in (s,t) to the X-coordinates of the patch, and a quadratic in (s,t) to the Y-coordinates of the patch. The linear equation $z = s$ is the best fit (exact) to the Z-coordinates of the patch. Compute the Gaussian curvature of the surface defined by the quadratics. If the Gaussian curvature is greater than 0.0, (s_0,t_0) is a concave or a convex point. There is a simple second derivative test to distinguish concave from convex. We have found Do Carmo[15] to be a very good reference for differential geometry.

GRAPHICAL REPRESENTATIONS

Plastic Model of the Cleft

As described above, we consucted three-dimensional plastic models of the clefts in HLA-A2 and HLA-Aw68. These models are constructed at the scale 1 cm = 1 Å and are particularly useful for modeling the position of an antigenic fragment in the binding site. Shown in Color Plate 1 are pictures of (a) the whole model of HLA-A2 with the influenza matrix peptide 58–66 (amino acids GILGFVFTL)[16,17] fitting into the cleft, and (b) a picture of a portion of the model of HLA-A2 (pocket B) with the N-terminal glycine and isoleucine of that peptide. The amino terminal GLY of the peptide is placed deep in pocket A of the cleft and does not appear in Color Plate 1a, but can be seen in Color Plate 1b. The peptide is made of HGS Biochemistry Molecular Model parts that also scale to 1 cm = 1 Å. It is rather easy to manipulate the peptide within the cleft, and when an interesting position is found, to read the coordinates of the atoms of the peptide relative to the coordinate system established for the class I molecule. These coordinates can be read to within about 0.4-Å accuracy and readily adjust to acceptable bond lengths and angles with CHARMm.

As shown in Color Plate 1b, the model may be separated at any integer-valued Z-cross section to examine the class I atoms in that cross section and their proximity to peptide atoms. The cross section in the foreground of Color Plate 1b is $z = 10$ (cs10); pocket B runs approximately from cs8 to cs11. One can see the preponderance of carbon atoms surrounding the pocket, making it suitable to accept an hydrophobic side chain such as the isolucine shown in Color Plate 1a. Latron et al.[17] present experimental evidence that the peptide is oriented in the cleft with N-terminus to the left ($z \doteq 4$) and the C-terminus to the right ($z \doteq 28$). The B pocket is well located to accept the second residue side chain of the peptide.

Geometric Map of the Cleft Floor

"Pockets" in the cleft have been a major feature of the descriptions of the MHC binding regions.[7,10,14] Looking down into the plastic model one can easily see the cavities in the floor of the cleft that are referred to as pockets, and are believed to be the locations of the peptide amino- and carboxyl-end groups and certain peptide side chains. We show in Figure 1 a two-dimensional map of the floor of the HLA-

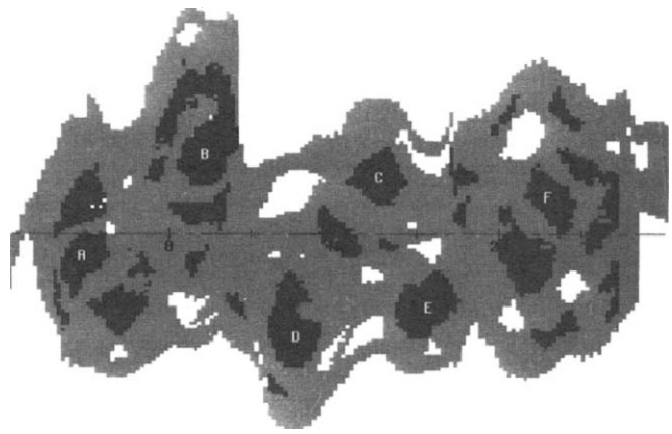

Figure 1. Geometry of the HLA-A2 cleft floor is illustrated, with black showing regions of concavity of the protein (associated with pockets). White shows regions of convexity (bumps on the protein surface). The letters A–F mark the locations of the pockets of the cleft as labeled in Saper et al.[10] The numbers mark Z-coordinates.

A2 cleft that shows the concave and convex regions of the floor. We define concave and convex from the perspective of the protein. A point on the water accessible surface is a concave point if intervals connecting nearby surface points lie outside the protein (thus, on the water accessible side of the surface); convex means that intervals connecting nearby surface points lie inside the protein (on the protein side of the surface). The pockets of the cleft are generally associated with concave regions of the surface, and appear in Figure 1 as black areas. The white areas are convex regions and appear as small bumps on the protein surface protruding into the cleft. Gray marks either relatively flat or saddle points of the surface. Marked on the map in Figure 1 are the locations of the pockets A–F distinguished by Saper et al.[10] The maps are most accurate near the horizontal center line, where the vertical traces are aligned, and are limited to the floor of the cleft below $y \leq 3$ to avoid further distortion at the upper and lower edges. The concave regions are also rather subtly displayed in Color Plates 3 and 4, by using black characters to mark atoms associated with concave regions, white characters otherwise. One can see groups of black characters associated with the various pockets A–F in both figures.

Comparison of the HLA-A2 and HLA-Aw68 Clefts

Pictures of cross sections perpendicular to the main axis of the cleft are particularly useful for comparing differences among class I MHC binding clefts. "Pocket C"[10] of HLA-A2 is a minor indentation compared to pocket C of HLA-Aw68, as illustrated in the two cross sections, cs19, of Color Plates 2a and b. The tan regions in the pictures show the cleft points (points that lie within a 1.4-Å sphere that does not intersect the protein). The intermediate tan color marks what we identify as pocket region, determined as follows. The lighter tan region at the top of the cleft shows the points that not only lie in a 1.4-Å sphere not intersecting the protein, but also lie within a 3.0-Å sphere that does not intersect the protein, and are considered to be cleft points that are not within pockets.[14] The remaining region includes some shallow sections along the water accessible surface that are too shallow to be considered pockets and are marked with the darkest tan color. We distinguish pocket points to be those grid points that lie within a 0.6-Å sphere that lies totally within the remaining region, and have marked them with an intermediate tan color. The length 0.6 Å was chosen heuristically.

As described in Saper et al.,[10] pocket C in HLA-A2 "is located on the inner wall of the α_1 domain α-helix above NE2 of HIS 74." This can be seen as the intermediate tan region in the lower right portion of the cleft in Color Plate 2a. The sections clearly show that four of the 11 side chains differing between the two clefts (70, 74, 97, and 114), are involved in the formation of this pocket. Atoms from all four side chains are seen in both cross sections. The much deeper pocket C in HLA-Aw68 strongly suggests a side chain difference in the specificity of antigenic peptides that bind to HLA-Aw68 as opposed to HLA-A2. Because the pockets are located in cs19, and assuming an extended conformation of a nonamer peptide with orientation as suggested by Latron et al.,[17] the side chain difference would occur in the fifth or sixth residue of the nonamer.

The beginning of the E pocket of HLA-A2 appears on the lower left of the cleft in Color Plate 2a. The corresponding identation in HLA-Aw68 is so slight that our algorithm does not record pocket E. As other class I structures are determined, the identity and location of pockets may not be so discretely and uniquely defined as is suggested by the references to and comparisons of pockets A–F in the two molecules HLA-A2 and HLA-Aw68. Particularly, the pockets near the middle of the cleft seem likely to change. One can see in Color Plate 2 that pocket C of HLA-A2 is not sufficient to accommodate a side-chain; the center of pocket C of HLA-Aw68 is actually forward a bit, to cs20. Nevertheless, the language is useful for initial comparisons, and because amino acids at the ends of the clefts are relatively well conserved between different class I molecules, and form pockets thought to accommodate the peptide amino and carboxyl terminal groups, the pockets at the ends of the cleft would be expected to also be conserved between the different class I molecules.

Chemical Maps of the Cleft Floor

The electrostatic, hydrophobic, and hydrogen bonding profiles of the cleft determine the binding characteristics of the class I molecule. Shown in Color Plate 3 is a map of the floor of the HLA-A2 cleft showing chemical characteristics of the exposed atoms. Each grid point of the cleft was associated with the atom of the MHC with surface nearest the grid point, and the physical/chemical characteristic of that atom was assigned to the grid point. We began with the atomic solvation parameter types used by Eisenberg and McLaughlin[18] and refined by Wessen and Eisenberg,[19] yielding five categories: carbon, uncharged oxygen or nitrogen, sulfur, charged oxygen, and charged nitrogen. Wessen and Eisenberg suggest that for the shared charge pairs in GLU, ASP, and ARG, the atom with largest exposed surface area (number of associated grid points) should be shown as charged. We use this procedure with the exception that salt bridges involving exposed atoms should first determine charged atoms. The uncharged side chain oxygens and nitrogens are further distinguished as hydrogen bond donor/acceptor according to Ippolito, et al.[20] The uncharged nitrogens are marked as donors in all side chains except HIS, in which the nitrogens are marked as acceptor/donor. The uncharged oxygens are marked as acceptors in all side chains except SER, THR, and HIS, in which the hydroxyl oxygens are shown as acceptor/donors. The rare appearance of sulfur is not distinguished as to hydrogen bond type.

The effect of first marking the atoms in salt bridges as charged can be seen in ARG 97, where NH2 is clearly interacting with OD1 of ASP 77. NH2 of ARG 97 has smaller exposed surface area than does NH1, and OD1 has smaller exposed surface area than does OD2 of ASP 77. Nevertheless, ARG 97 NH2 and ASP 77 OD1 are well positioned to form a salt bridge (distance between the centers is 3.110 Å), and are marked as the charged atoms of the respective side chains.

In Color Plate 3 the hydrophobic designation follows Wesson and Eisenberg, with five categories, carbon (color

code C), uncharged oxygen and nitrogen (O/N), charged oxygen (O^-), charged nitrogen (N^+), and sulfur (S). The uncharged oxygens and nitrogens are further distinguished as to hydrogen donor type. The hydroxyl oxygens (acceptor/donor, O HB A/D) are distinguished from the remaining uncharged oxygens (acceptors, O HB A); the histidine nitrogens (acceptor/donor, N HB A/D) are distinguished from the other uncharged side chain nitrogens (donors, N HB D). Symbols are shown for all of the non-carbon atoms and for carbon atoms for which there are at least 50 surface grid points. The pockets of the cleft are indicated by black lettering of the atoms located near concave regions of the surface (see Figure 1).

Another feature of the surface of the cleft is the temperature factor of the surface atoms. Shown in Color Plate 4 is a map of those temperature factors for the HLA-A2 cleft floor, color coded according to the scale shown. It can be seen that atoms in one region shown in red, from ARG 97, HIS 114, and TYR 116, have high thermal factors, in the range from 32 to 40. Saper et al.[10] suggest that the high thermal factors may be associated with flexibility of the relevant side chains required to accommodate different bound peptide orientations. A comparable picture of HLA-Aw68 shows no such high temperature cluster.

DISCUSSION

The plastic model we constructed is very accurate, stable and robust; the conformation of the cleft is well defined and clear. The model can be opened at any 1-Å Z-interval to observe the side chains and atoms forming the cleft at that point. A disadvantage of the model is that it does not allow movement of side chains. Each atom is positioned exactly as in the original Protein Data Bank file, but only in that position. However, the exact definition of the cleft, including the clear identification of the pockets, has proven to be very valuable in understanding the positioning of the antigen in the cleft, and the positioning of antigen side chains into the cleft pockets. We believe it to be a valuable alternative and supplement to the ball-and-stick and space-filling models that have been more generally used. An earlier construction of the MHC molecule with ball-and-stick plastic was very helpful, but because of the inaccuracy in location of atoms it did not show the pockets in the cleft, and it lacks the stability and robustness of our more recent model.

The materials required for the model are two 4′ by 8′ sheets of 0.080″ (≒ 2-mm) thick GP CAST acrylic, 5′ of 3/16″ diameter acrylic rod, solvent for acrylic, 32 pictures, and 64 magnets; it requires about 12 hours of machine time and 20 hours for two people to set up and assemble. Although not inexpensive, our construction is not so expensive as the earlier model, and we believe it is a useful addition to the tools for understanding molecular structure.

The cross sections we display on the computer can be scaled, 1 cm = 1 Å on the screen, and are programmed to advance through the molecule in either the ± Z-direction. This allows the user to determine which atoms are surrounding a point on the surface of the cleft or a pocket region. We have also added antigenic peptides in the cleft, and displayed them in the computer cross sections, where the geometrical and chemical fit can be examined. With smaller scale, we have simultaneously displayed cross sections of as many as four different MHC structures, some with peptide in the cleft, thus making easy and informative comparisons.

The methods we describe here for displaying the cleft in the MHC class I molecule are applicable to a broad range of molecular problems, particularly problems of drug design, more general protein–protein interactions, and protein–DNA complexes.

ACKNOWLEDGEMENTS

The authors appreciate very much the assistance of Charles Stanley and Gerry Sheppard in setting up the milling machine and advice on running it. The authors also gladly acknowledge the beneficial discussions with Dr. Temple Smith. This work was supported by Grant AI 30535 from the National Institute of Allergy and Infectious Disease.

REFERENCES

1 Rotzsche, O. and Falk, K. Naturally-occurring peptide antigens derived from the MHC class-I-restricted processing pathway. *Immunol. Today* 1991, **12,** 447–455
2 Germain, R.N. The second class story. *Nature (London)* 1991, **353,** 605
3 Jardetzky, T.S., Lane, W.S., Robinson, R.A., Madden, D.R., and Wiley, D.C. Identification of self peptides bound to purified HLA-B27. *Nature (London)* 1991, **353,** 326
4 Rudensky, A.Y., Preston-Hulbert, P., Soon-Cheol, H., Barlow, A., and Janeway, C.A. Jr. Sequence analysis of peptides bound to MHC class II molecules. *Nature (London)* 1991, **353,** 622
5 Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature (London)* 1987, **329,** 506–511
6 Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature (London)* 1987, **329,** 512–518
7 Garrett, T.P.J., Saper, M.A., Bjorkman, P.J., Strominger, J.L., and Wiley, D.C. Specificity pockets for the side chains of peptide antigens in HLA-Aw68. *Nature (London)* 1989, **342,** 692
8 Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer E.F. Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., and Tasumi, M. The protein Data Bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* 1977, **112,** 535–542
9 Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F., and Weng, J. Protein Data Bank in *Crystallographic Databases—Information Content, Software Systems, Scientific Applications*. (F.H. Allen, G. Bergerhoff, and R. Sievers, Eds.) Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester, 1987, 107–132
10 Saper, M.A., Bjorkman, P.J., and Wiley, D.C. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6-Å resolution. *J. Mol. Biol.* 1991, **219,** 277–319

11 Richmond, T.J. Solvent accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J. Mol. Biol.* 1984, **178,** 63–89
12 Brooks, C.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* 1983, **4,** 187–217
13 Polygen Corporation, Waltham, MA 02254
14 Madden, D.R., Gorga, J.C., Strominger, J.L., and Wiley, D.C. The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. *Nature (London)* 1991, **353,** 321–325
15 Do Carmo, M.P. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, Englewood Cliffs, 1976
16 Gotch, F., McMichael, A., and Rothbard, J. Recognition of influenza matrix protein by HLA-A2-restricted cytotoxic T lymphocytes. *J. Exp. Med.* 1988, **168,** 2045–2057
17 Latron, F., Moots, R., Rothbard, J.B., Garrett, T.P.J., Strominger, J.L., and McMichael, A. Positioning of a peptide in the cleft of HLA-A2 by complementing amino acid changes. *Proc. Nat. Acad. Sci. USA* 1991, **88,** 11325–11329
18 Eisenberg, D. and McLaughlin, A.D. Solvation energy in protein folding and binding. *Nature (London)* 1986, **319,** 199–203
19 Wesson, L., and Eisenberg, D. Atomic solvation parameters applied to molecular dynamics of proteins in solution. *Protein Sci.* 1992, **1,** 227–235
20 Ippolito, J.A., Alexander, R.S., and Christianson, D.W. Hydrogen bond stereochemistry in protein structure and function. *J. Mol. Biol.* 1990, **215,** 457–471