

Model for the Interaction of T-cell Receptors with Peptide/MHC Complexes

P.J. BJORKMAN* AND M.M. DAVIS†

*Division of Biology, California Institute of Technology, Pasadena, California 91125;

†Howard Hughes Medical Institute and Department of Microbiology and Immunology, Stanford University, Stanford, California 94305

The immune response against a viral infection is mediated by two different types of cells known as B and T lymphocytes. The receptor on the B cell is the well-characterized antibody molecule, which exists in a membrane-bound form and in a secreted form involved in the initiation of complement-mediated killing and the inactivation of viral particles by direct binding. The recognition molecule on T cells is the membrane-bound T-cell antigen receptor, which has specificity for a combination of foreign antigen with a molecule of the major histocompatibility complex (MHC), as first demonstrated by Zinkernagel and Doherty (1974). MHC proteins exist in two closely related forms called class I and class II MHC molecules, both of which are cell-surface glycoproteins that are highly polymorphic in the human population. In general, class II MHC molecules are involved in interactions with T-helper cells, which cooperate with B cells to make antibody. Class I MHC molecules are recognized by T-killer cells, or cytotoxic T lymphocytes, that lyse virally infected cells. In both cases, T-cell recognition of antigen together with a "self" MHC molecule is termed MHC "restricted" recognition (for review, see Davis 1985; Kronenberg et al. 1986).

It has now been established that MHC-restricted T-cell receptors (TCRs) recognize peptide fragments of antigens (presumably derived from intracellular processing) bound to an MHC molecule at what appears to be a single site. Two major lines of evidence point to the involvement of peptides in T-cell recognition: Peptide fragments of an antigen added to the outside of fixed class II MHC-bearing target cells can be recognized by T-helper cells specific for the appropriate combination of antigen and MHC type (Shimonkevitz et al. 1983). Subsequently, short synthetic peptides (8–30 residues long) were shown to bind to purified class II proteins (Babbitt et al. 1985). In some cases, the binding of a particular peptide by an MHC molecule correlated with the ability of an animal of that MHC type to mount an immune response against the antigen from which the peptide was derived (Buus et al. 1987), suggesting a reason for the observation that there is a correlation between immunological response to an antigen and certain specificities of histocompatibility molecules (McDevitt et al. 1972). It has been much more difficult to demonstrate peptide binding to purified class I MHC molecules, although class I and

class II MHC molecules are similar in domain organization, sequence, and presumably three-dimensional structure (Brown et al. 1988). However, virus-specific T-killer cells have been shown to lyse an uninfected target cell of appropriate class I specificity to which peptide fragments of a viral protein have been added (Townsend et al. 1986). By analogy with the work described for class II MHC molecules, it is assumed that the class I molecules on the target cell bind peptides, and that it is a peptide/MHC complex that is recognized by TCRs on the cytotoxic killer cells. Some T cells recognize foreign or "non-self" MHC molecules in the apparent absence of antigen, although in these cases, it is possible that the peptide-binding site is occupied by an endogenous peptide. The reactivity of T cells and antibodies against foreign MHC molecules leads to host rejection of transplanted tissue.

The discovery that MHC molecules bind antigenic peptides and present them to T cells has allowed the correlation of susceptibility to autoimmune disease with certain MHC alleles to be understood on a more molecular basis. Any given MHC molecule binds only a subset of peptides tested, and as discussed above, the ability to bind a peptide can determine the immune response to an antigen. Increased susceptibility to autoimmune diseases such as ankylosing spondylitis and insulin-dependent diabetes mellitus are found in individuals of certain MHC types, and such diseases arise when the body's immune system attacks its own proteins (for review, see Todd et al. 1988). It is thought that the MHC molecules correlated with autoimmune disease bind peptides from self proteins, leading to tissue damage by self-reactive T-cell clones.

One hopes that an increased understanding of the physical nature of antigenic peptide interactions with MHC molecules will allow the design of peptides to stimulate the immune response against a viral infection, or the design of high-affinity ligands that block the self-reactive recognition of MHC molecules involved in autoimmune diseases. Before this can be accomplished, we will need to understand not only the forces comprising the peptide/MHC complex, but also how TCRs bind to it. In this paper, we briefly discuss what the crystal structure of a human class I MHC molecule reveals about how histocompatibility molecules might interact with TCRs, and we present a hypothetical model for how MHC-restricted TCRs bind to the pep-

tide/MHC complex. The three-dimensional structure of a TCR has not yet been determined, but because of the similarity between TCRs and antibodies, we can use the known structure of an Fab to serve as a first-order TCR model structure, in order to make an educated guess about how T-cell recognition of the peptide/MHC complex occurs.

TCR Structure and Potential Diversity

TCRs are membrane-bound disulfide-linked heterodimers that resemble the Fab fragments of immunoglobulins in sequence and domain organization (for review, see Kronenberg et al. 1986). Each polypeptide chain contains a domain with sequence similarity to antibody variable (V) domains, followed by a domain similar to an antibody constant (C) domain. Following the constant-like domain in each chain, there is a hinge region containing a cysteine residue involved in the formation of an interchain disulfide bond, and then a hydrophobic membrane-spanning sequence and short cytoplasmic tail. Figure 1 shows a schematic representation of the primary sequence organization of the four TCR polypeptide chains that make up the two different types of receptor heterodimers (α/β and γ/δ). As in immunoglobulins, both the α/β and γ/δ TCRs are assembled by the relatively random joining of different coding segments to C-region genes (V and J join to C in the case of α and γ chains; V, D, and J join to C in the case of β and δ chains). Although variability in TCR V regions is less localized than in immunoglobulin V regions, hypervariability is found in the locations corresponding to the three classic immunoglobulin hypervariable regions or complementarity-determining regions (CDRs) (Patten et al. 1984; Barth et al. 1985), which are known in antibodies to form the principal points of contact with antigen (Wu and Kabat 1970; Kabat et al. 1987). In both antibodies and TCRs, the first and second CDR are encoded within the V gene segment itself, and the third CDR is formed by the junction of the V gene segment with D and J gene segments (in the case of immunoglobulin heavy chains

and TCR β and δ chains) or with J gene segments alone (immunoglobulin light chains and TCR α and γ).

Although TCRs and immunoglobulins share a similar organization of diversity and mechanisms for its generation, a closer look at TCR diversity shows a striking concentration of sequence polymorphism in the CDR3-equivalent region as compared to this region in antibodies (discussed in more detail in Davis and Bjorkman 1988). In contrast, TCR diversity within the CDR1- and CDR2-equivalent regions is far less pronounced than in immunoglobulins. The primary cause of decreased TCR diversity in the first and second CDR regions is that there are far fewer TCR V gene segments than immunoglobulin V gene segments. Assuming that the V domains of TCRs pair to form a combining site (as is the case for antibody V regions), the combinatorial diversity resulting from a random pairing of TCR V domains shows even greater disparity from the amount of diversity resulting from antibody V region pairing. Estimates for the amounts of sequence diversity possible for immunoglobulins and TCRs are compared in Table 1. As a result of the many mechanisms generating diversity within the CDR3-equivalent or junctional region in TCRs, the potential diversity is estimated to be between 4 and 7 orders of magnitude higher in TCRs as compared to immunoglobulins. On the other hand, the amount of potential TCR nonjunctional diversity (within the CDR1- and CDR2-equivalent regions) is estimated to be between 10 and 1000 times less than that possible for antibodies.

To rationalize the striking concentration of TCR diversity within the junctional region and relative lack of diversity elsewhere, it should be relevant to consider what TCRs are known to do; namely, to recognize a large number of small molecules (i.e., peptides) embedded in physically larger and much less diverse MHC molecules. The simplest interpretation of the skewing of diversity in the TCR case toward the CDR3-equivalent region is that the amino acids in this region are mainly interacting with peptide determinants, and that residues within the much less diverse CDR1- and CDR2-equivalent regions are primarily involved in contacts to MHC determinants.

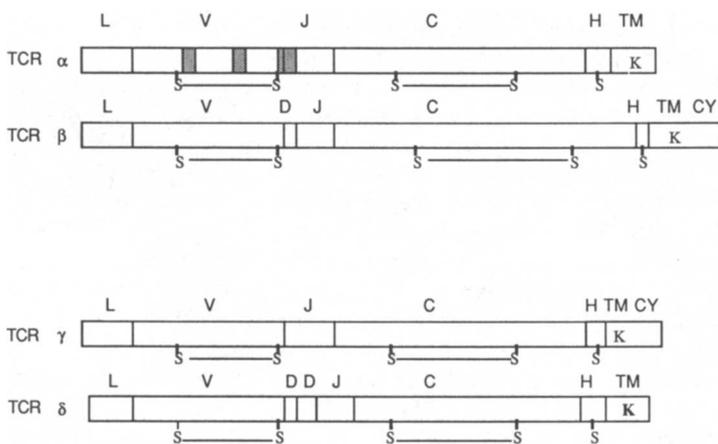


Figure 1. Schematic representation of primary structure of TCR polypeptides. Leader (L), variable (V), diversity (D), and joining (J) gene segment contributions are indicated for the variable domains. The constant portion of each TCR polypeptide is divided into the immunoglobulin-like constant (C) region, a hinge (H), transmembrane (TM) region, and cytoplasmic (CY) domain. The approximate locations of cysteine residues involved in disulfide bonds are indicated by an external "S" connected to another cysteine indicated by an external "S". The cysteine in the hinge region is presumed to form an interchain disulfide bond. The approximate location of CDR1-, CDR2-, and CDR3-equivalent regions are indicated by shading in the primary structure of the TCR α -chain.

Table 1. Sequence Diversity in T-cell Receptor and Immunoglobulin Genes

| | Immunoglobulin | | TCR α/β | | TCR γ/δ | |
|------------------------------|----------------|----------------|--------------------|----------------|---------------------|----------------------|
| | H | κ | α | β | γ | δ |
| Variable segments | 250-1000 | 250 | 100 | 25 | 7 | 10 |
| Diversity segments | 10 | 0 | 0 | 2 | 0 | 2 |
| Ds read in all frames | rarely | — | — | often | — | often |
| N-region addition | V-D, D-J | none | V-J | V-D, D-J | V-J | V-D1, D1-D2, D1-J |
| Joining segments | 4 | 4 | 50 | 12 | 2 | 2 |
| Variable region combinations | 62,500-250,000 | | | 2500 | | 70 |
| Junctional combinations | | $\sim 10^{11}$ | | $\sim 10^{15}$ | | $\sim 10^{18}$ |

Calculated potential amino acid sequence diversity in TCR and immunoglobulin genes without allowance for somatic mutation. The approximate number of V gene segments are listed for the four TCR polypeptides and contrasted with immunoglobulin heavy and light chains. CDR1 and CDR2 are encoded within the V gene segments. The pairing of random V regions generates the combinatorial diversity listed as "variable region combinations." Because there are fewer TCR V gene segments than immunoglobulin V gene segments, the combinatorial diversity is lower in TCRs than in immunoglobulins. Estimates for the number of unique sequences possible within the junctional region are contrasted for TCRs and immunoglobulins. Amino acids within CDR3 are encoded almost entirely within the D and/or J region gene segments. (The last few amino acids encoded by a TCR V gene segment can contribute to diversity within the TCR CDR3-equivalent region, but the effects of these residues on junctional diversity are not included in these calculations.) The mechanisms for generation of diversity within the junctional region that are used for this calculation include usage of different D and J gene segments, N region addition up to six nucleotides at each junction, variability in the 3' joining position in V and J gene segments, and translation of D regions in different reading frames. Numbers are corrected for out-of-frame joining codon redundancy and N-region mimicry of germ-line sequences as described by Elliott et al. (1988).

At this point, it would be useful to compare the three-dimensional structure of a TCR to the structure of HLA-A2 to see if there is any structural reason to hypothesize the alignment of CDR1 and 2 with MHC residues, and CDR3 with residues on a bound peptide. Although the crystal structure of a TCR is unknown, sequence data suggest that TCR V regions are folded into β -sandwich structures resembling immunoglobulin V regions (Chien et al. 1984; Hedrick et al. 1984; Yanagi et al. 1984). In antibodies, the hypervariable or CDR regions are located on loops that connect β -strands, and the V regions from the heavy and light chains (V_H and V_L) are paired such that CDR1, 2, and 3 from each domain are clustered at the ends of the Fab arms of the molecule, forming the antigen-binding site (for review, see Davies and Metzger 1983). In all of the three-dimensional structures of antibodies that are known, V_H and V_L domains have a conserved mode of interaction in which they are paired about an approximate twofold symmetry axis (i.e., one domain is related to the other by a 180° rotation). Thus, in different Fabs, the CDR loops are found in the same relative positions with respect to each other (Chothia et al. 1985). A study of immunoglobulin sequences and structures has identified conserved amino acids that are critical for maintaining the V_H - V_L contact surface, and most of these amino acids are found in homologous positions in TCR V region sequences (Novotny et al. 1986; Chothia et al. 1988). It is therefore likely that TCR V domains will not only fold into tertiary structures similar to antibody V regions, but that the chain pairing and resulting combining sites of TCRs will also be similar to those described for antibodies. This conclusion has been recently reached after a systematic study of immunoglobulin structures and a comparison of TCR and immunoglobulin sequences by Chothia et al. (1988). These authors conclude that TCR and anti-

body V domains will share a similar β -sheet framework, as well as a similar domain interface. Strongly conserved residues are also found in TCR β sequences at sites homologous to those that form the conserved contact between the framework of the V and C domains of the immunoglobulin heavy chain, implying that V_β and C_β will associate in the same way as V_H and C_H1 . We can therefore use what is known about antigen binding from crystal structures of Fab and of Fab/antigen complexes to serve as a guideline for understanding how TCRs bind to their ligand.

In antibodies, the first and second CDRs on V_L are separate from their counterparts on V_H , and the space between them is occupied by the CDR3 regions from each chain (Davies and Metzger 1983). In Figure 2A, the relative locations of the six CDR regions in an immunoglobulin combining site are shown (for our purposes to represent the approximate structure of a TCR combining site). In this view of the V_H - V_L pair, the pseudo twofold axis is perpendicular to the plane of the paper and located between the two CDR3 loops. Structural studies of antibodies complexed with protein antigens have shown that side chains from all six CDRs can contact the antigen, and that the entire antigen/antibody interface is a rather flat surface with protrusions and depressions in the antibody being complementary to the antigen surface (Amit et al. 1986; Colman et al. 1987; Sheriff et al. 1987). Assuming these details are also true for the combining site for a TCR, we now turn to a discussion of the structure of the TCR ligand, i.e., the peptide/MHC complex.

MHC Structure

The three-dimensional structure of HLA-A2, a human class I MHC molecule, has been determined (Bjorkman et al. 1987a,b), and it is likely that other

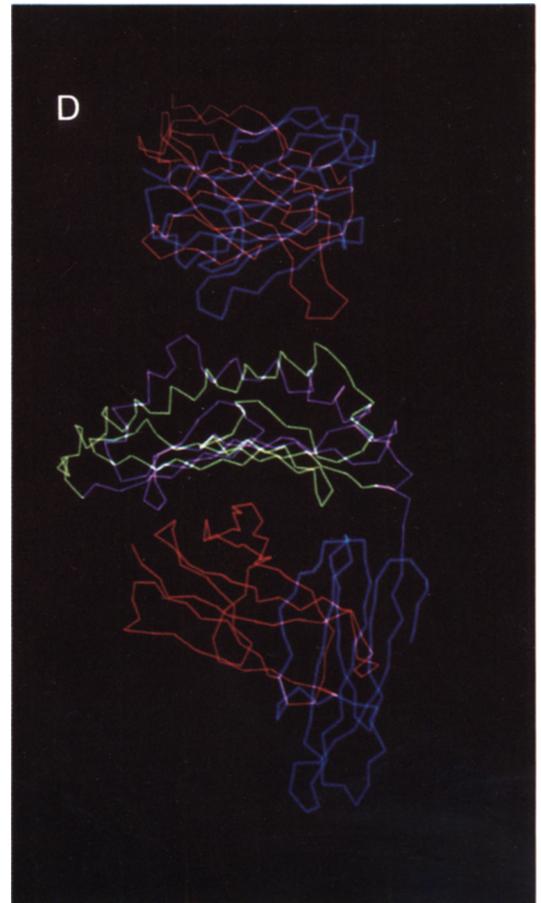
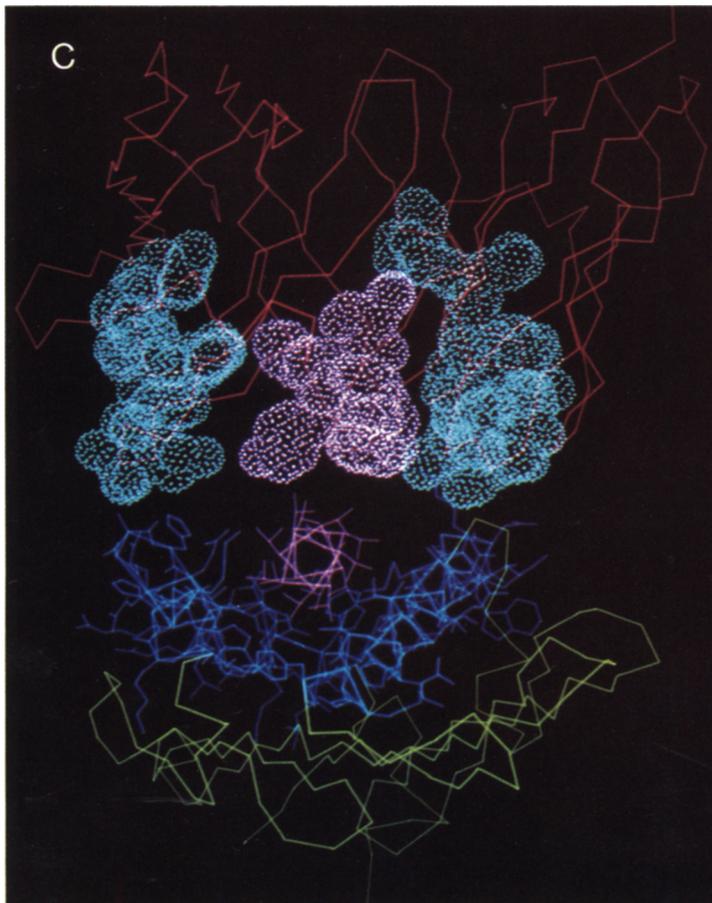
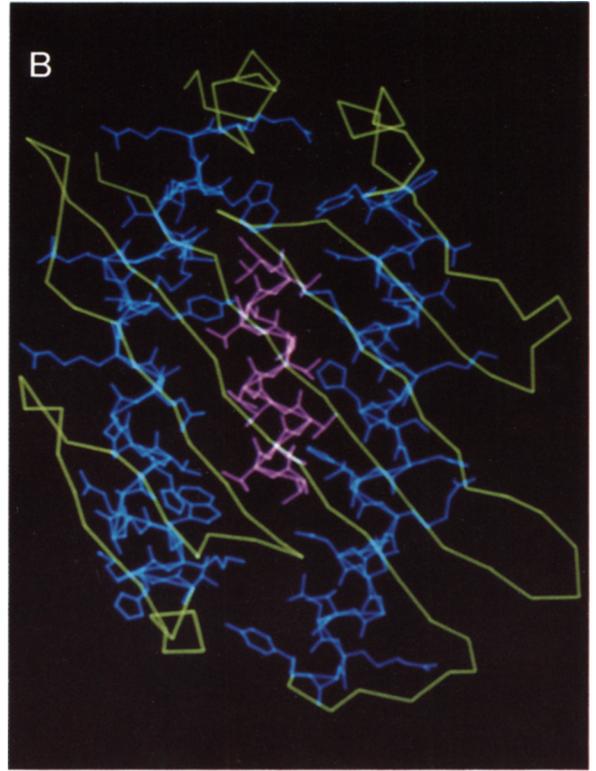
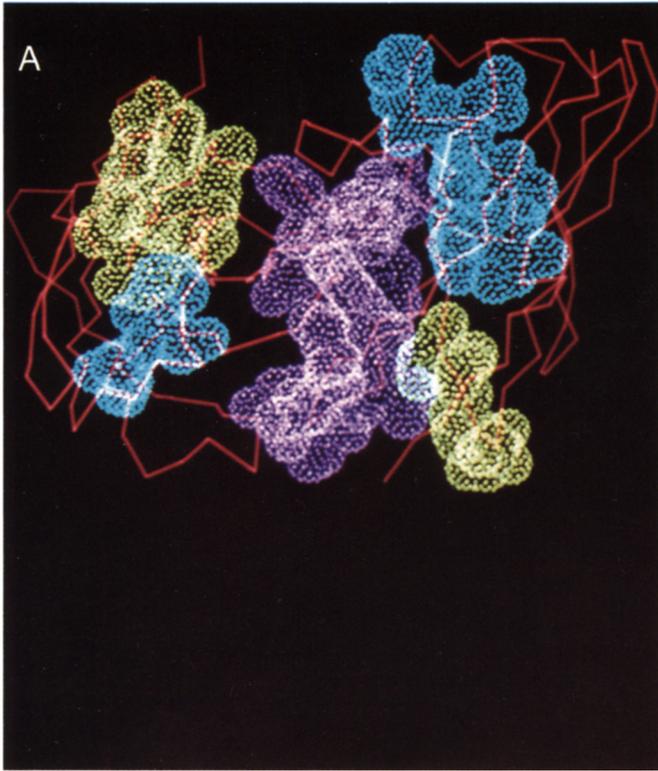


Figure 2. (See facing page for legend.)

class I as well as class II MHC molecules (Brown et al. 1988) will fold into similar tertiary structures. The HLA-A2 molecule is composed of two structural motifs: The membrane proximal α_3 and β_2 microglobulin domains are folded like immunoglobulin C domains, and the membrane-distal α_1 and α_2 domains consist of a platform of eight β -strands topped by two α -helices. The α_1 and α_2 domains have similar tertiary structures, each consisting of four β -strands and a long helical region. As is also the case for immunoglobulin and presumably TCR V domains, the α_1 and α_2 domains are related by an approximate twofold symmetry axis when they are paired in the HLA structure. Since these are the domains that show most of the polymorphism between different alleles, one would expect that the peptide-binding site and TCR recognition regions would be located on this part of the molecule. Several lines of evidence suggest the peptide-binding site on MHC molecules is the deep groove that separates the two α -helices on its top surface, and that this surface is therefore the part of the MHC molecule that the TCR recognizes: (1) The groove is located on the top of the molecule, and is thus in a likely position to interact with a molecule on the surface of another cell. (2) The groove is ~ 25 Å long and ~ 10 Å wide and deep, dimensions consistent with the expectation that MHC molecules bind a processed (i.e., peptide) form of an antigen. (3) Many of the residues that form the sides and the bottom of the site are highly polymorphic or have been identified to be critical for T-cell recognition of class I molecules, fitting with the expectation that MHC polymorphisms affect the peptide binding and T-cell reactivity. (4) The crystal structure of HLA-A2 shows that the site appears to be occupied by a molecule or mixture of molecules that evidently copurified and cocrystallized with HLA-A2. Although other interpretations are possible, it seems most likely

that these molecules represent a heterogeneous mixture of endogenous peptides perhaps added during synthesis of HLA. If the peptide-binding site is occupied, the top surface of the peptide/MHC complex is relatively flat, as can be shown by model-building short peptides into the HLA-A2 structure. (In fact, in the original electron density maps of HLA-A2, the top surface of the molecule appears fairly level because of the presence of the unknown occupant[s] in this site.) A view of the top surface of HLA-A2 combined with a 12-residue peptide model-built (as an α -helix) into the presumed peptide-binding site is shown in Figure 2B (to scale with respect to the V regions combining site, Fig. 2A). The HLA-A2 structure is discussed in more detail by Bjorkman et al. (1987a,b).

A Model for TCR Recognition of the Peptide/MHC Complex

When the (hypothetical) TCR combining site is compared to the top surface of the peptide/MHC complex that is its ligand, it is interesting to note that the α -helices that make up the sides of the MHC peptide-binding site are separated by about the same distance as separates the CDR1 and 2 regions of one V domain from another (18 Å). Thus, the relatively flat surface of a peptide/MHC complex could interact with the combining site of an immunoglobulin-like TCR such that the limited diversity in the CDR1- and CDR2-equivalent regions on TCR V_α and V_β contacts the side chains of the MHC α -helices, leaving the centrally located and very diverse CDR3-equivalent regions to interact with the peptide. Using immunoglobulin V regions as a model for a TCR and HLA-A2 with a hypothetical peptide as the MHC molecule, this type of interaction is shown in Figure 2C. A more extensive discussion of this model for TCR recognition of peptide/MHC com-

Figure 2. Representation of the structures of (A) the immunoglobulin antigen-binding site, (B) the peptide-binding site of an MHC molecule, and (C) the alignment of CDRs in a hypothetical TCR over a peptide/MHC complex. (A) View from the top (the direction of antigen) of an immunoglobulin combining site (Fab J539; Suh et al. 1986) with the three CDRs on each chain highlighted with van der Waals surfaces (CDR1: yellow, CDR2: blue, CDR3: pink, carbon- α backbone of V_H and V_L shown in red). Similarities between TCRs and immunoglobulins suggest that TCR combining sites will preserve the same general features, having CDR1 and CDR2 from one domain separate from their counterparts on the partner domain and the space between them occupied by CDR3 from each domain (see text for details). (B) Top surface of an MHC molecule with the side chains of the amino acids located in two helices on either side of the peptide-binding site highlighted in blue. A hypothetical peptide (a 12-mer polyvaline) has been model-built into the peptide-binding site as an α -helix (shown in pink). Carbon- α backbone of HLA-A2 (Bjorkman et al. 1987a,b) is shown in green. Note that the distance between the MHC α -helices is approximately the same as the distance the first and second CDRs on one V domain are separated from their counterparts on the partner V domain (see part A). (C) Model for TCR interaction with a peptide/MHC complex. The molecules in this figure are rotated $\sim 90^\circ$ with respect to their orientation in parts A and B on this figure. The antibody V domains (top of figure) here represent a TCR bound to a peptide/MHC complex (bottom of figure). The carbon- α backbone of the V domains (red) is shown with the amino acids within CDR1 and CDR2 highlighted with blue van der Waals surfaces, and amino acids within CDR3 highlighted in pink van der Waals surfaces. The MHC carbon- α backbone is shown in green with side chains located on the α -helices highlighted in blue. A (hypothetical) peptide model built into the binding site as an α -helix is shown in pink. The relatively flat surface of the V region combining site is complementary to the peptide/MHC complex, with the first and second CDR from each V domain "fitting" over an MHC α -helix, leaving the centrally located CDR3 regions aligned over the peptide (see text for details). (D) Side view of model for TCR interacting with an MHC molecule. The V domains and MHC molecule have been rotated to show that there is sufficient space on the top surface of the MHC molecule for TCR V regions (here depicted as immunoglobulin V regions) to bind in several different registers along the MHC α -helices. The carbon- α backbone of V_H and V_L (here representing TCR V_α and V_β) are shown in red and blue (top of figure). At the bottom of the figure are the four domains of a class I MHC molecule (α_1 : green, α_2 : pink, α_3 : blue, β_2 -microglobulin: red).

plexes has been published (Davis and Bjorkman 1988), and a similar model has been suggested by Claverie and co-workers (Claverie et al. 1989).

This model for TCR and MHC/peptide interactions is consistent with some recent work correlating TCR α/β sequences with known antigen-MHC specificities. For example, in some cytochrome-*c*-specific T-cell clones, changes in the junctional region alter the specificity for peptide without altering MHC specificity (Fink et al. 1986; Winoto et al. 1986). Also, some of the TCRs from these clones show a selection for certain amino acids within the CDR3-equivalent region, suggesting that junctional residues are important for peptide recognition (Hedrick et al. 1988). This idea was tested by changing one of the conserved junctional residues by site-directed mutagenesis, with the result that the mutated TCR displayed a different fine specificity for antigen (Engel and Hedrick 1988).

Even within the confines of our model for TCR interaction with peptide/MHC complexes, however, one would not always expect a direct correlation of TCR junctional residues with specificity for a particular peptide, for the simple reason that changes in MHC residues could alter the conformation or orientation of a bound peptide, thus requiring a compensatory change in the TCR residues contacting the peptide. Also, we would not expect that the V_α and V_β gene segments used by T cells restricted to the same MHC molecule would always be the same, because the surface of a peptide/MHC complex is large enough to allow a TCR (assuming an immunoglobulin-like binding site) to bind in different registers along the MHC α -helices. In Figure 2D, the V-regions/MHC complex has been rotated by 90° so that the MHC α -helices are approximately in the plane of the paper to demonstrate that there is room on the top surface of the MHC molecule for a TCR molecule to bind in several different regions. In addition, because antibody (and presumably TCR) V regions pair with approximate dyad symmetry and because the MHC α_1 - α_2 domains are approximately dyad symmetric, the interaction of CDR1 and CDR2 residues with MHC determinants and CDR3 with peptide determinants can be accomplished in either of two orientations related to each other by 180° . (In other words, the interaction depicted in Figure 2C would look the same if the V regions dimer were rotated by 180° about its pseudo-dyad axis, which is vertical in this figure).

It is likely that an exact correlation between TCR hypervariable region sequences and MHC and peptide specificities will never be possible, due to the spatial proximity of CDR residues in the combining site. Also, it is a bit of an oversimplification to assume that TCR interactions with ligand will exclusively involve residues in the CDR loops, because crystal structures of Fabs complexed to protein antigens have shown that framework residues adjacent to the CDR loops can contact the antigen (Amit et al. 1986; Sheriff et al. 1987). However, the structural complementarity between the MHC peptide-binding site and the (hypothetical) TCR

combining site depicted in Figure 2C may provide an explanation for the original suggestion by Jerne (1971) that T-cell antigen receptors and MHC molecules coevolved to have some affinity for each other and gives us some idea of what MHC-restricted TCR recognition means at the molecular level. This model also provides a rationale for the decreased number of TCR V gene segments as compared to immunoglobulin V gene segments, because the variability within the portion of the protein they encode is primarily predicted to interact with a limited number of MHC molecules within any given individual (on the order of 6 class I and 20 class II MHC molecules in a heterozygous individual). Finally, the model may serve as a useful framework for the design of experiments to alter TCR residues and test the effect on MHC or peptide recognition.

Thymic Selection

In a way that is not fully understood, the T cells that leave the thymus have receptors that can recognize foreign antigen plus self-MHC, but not self-MHC alone (or self-MHC plus a self-peptide) when they are in the periphery (for review, see Marrack and Kappler 1987). It is difficult to understand how this selection can occur, since the huge number of foreign peptides that the TCR must recognize are not present in the thymus. The selection of receptors is thought to operate on two levels: There is a "positive" selection for TCRs that recognize antigen in the context of their own MHC, and a "negative" selection against self-reactive receptors to ensure that individuals are tolerant to their own proteins. It has been suggested that all TCRs with some affinity for self-MHC are selected, but only thymocytes with a low affinity for self are allowed to mature because the process of tolerance deletes cells with TCRs that have a moderate to high affinity for self-MHC (Marrack and Kappler 1987). The TCRs with low affinity for self-MHC are assumed to have an increased affinity for some combination of self-MHC plus foreign peptide that they will encounter in the periphery.

It is difficult to understand, however, how a positively selected TCR could bind tightly enough to distinguish residues on a self-MHC molecule from a non-self molecule, yet apparently disregard interactions with any self-peptides that are present on MHC molecules in the thymus during selection. As can be seen by examination of the HLA-A2 structure, a TCR binding to the top surface of the peptide/MHC complex would, by necessity, interact with a bound peptide. In fact, because of the physical proximity of side chains contributed by the MHC molecule and side chains from the peptide, almost any shape of molecule recognizing the top of the MHC molecule could not help but interact with parts of a bound peptide. As a way around the dilemma of the absence of foreign peptides during positive selection, it has been postulated that the foreign peptides to which the TCR responds in the periphery resemble self-peptides that are bound to

MHC molecules in the thymus. As has been suggested by Marrack and Kappler (1987), the MHC molecules on the thymus cortical epithelial cells might bind a different spectrum of peptides than is found elsewhere in the animal. In this scenario, these peptides would not be found on bone-marrow-derived cells, hence the process of tolerance would not delete cells with TCRs specific for self-MHC plus a thymic cortical epithelial-specific peptide. These cells would then go on to join the peripheral T-cell pool, where they would respond to foreign peptides that resemble the peptide upon which they were selected. A variant of this idea suggests that spontaneous mutation of self-proteins could generate enough variants to self-peptides to allow TCRs to be selected to recognize a self-MHC plus a mutated self-peptide that resembles a foreign peptide to be encountered later in the periphery (Kourilsky and Claverie 1989). However, it is hard to imagine that enough variability in self-peptides would be generated to account for a T-cell response against a universe of foreign antigens.

Another way to reconcile the existence of positive selection with the absence of foreign antigens in the thymus would be to assume that the MHC molecules in the thymic cortical epithelium have empty peptide-binding sites, which is not unreasonable from a purely structural point of view. Since the residues lining the site in HLA-A2 are not exclusively hydrophobic (Bjorkman et al. 1987b), one could imagine that the peptide-binding site could be filled with water molecules, resulting in an empty MHC molecule with a structure similar to that of HLA-A2 (in which the site was found to be occupied by an unknown molecule or mixture of molecules). An empty class II MHC molecule is also a structural possibility, since an examination of the class II MHC sequences shows that the residues predicted to be in the peptide-binding site are also not particularly hydrophobic. If TCRs are positively selected by binding to MHC molecules with empty peptide-binding sites, they could recognize the MHC α -helices using residues within their CDR1 and CDR2 loops, leaving the selection of residues within the CDR3 junctional region to be almost completely unconstrained. Following a selection of TCRs with affinity for self-MHC, those TCRs with an affinity for self-MHC plus a self-peptide could be eliminated by whatever mechanism controls negative selection, leaving a number of TCRs capable of recognizing a self-MHC plus foreign peptide complex.

The idea that positive selection occurs on empty MHC molecules assumes that at least some MHC molecules in the thymic cortical epithelium are capable of reaching the cell surface without a bound peptide. However, recent results from Townsend's laboratory suggest that the intracellular assembly and cell-surface expression of class I MHC molecules are dependent on peptide binding (Townsend et al., this volume). It is still possible, though, that some small percentage of MHC molecules manage to reach the cell surface of the thymic cortical epithelium without a bound peptide.

TCRs that are positively selected by binding to these empty molecules could survive a subsequent negative selection step, whereas the majority of TCRs that are positively selected on MHC molecules occupied with a self-peptide would be eliminated during negative selection.

TCR Recognition of Superantigens

Recently, several self-antigens in combination with murine class II MHC molecules have been shown to form ligands that are recognized by virtually all T cells whose receptors bear certain V_{β} elements, without apparent regard to V_{α} , J_{α} , D_{β} , or J_{β} gene usage (Kappler et al. 1987, 1988; McDonald et al. 1988; Pullen et al. 1988). The best-studied such antigen is encoded within the Mls locus, and mice combining particular Mls and MHC haplotypes are found to have deleted T cells bearing certain V_{β} TCRs during the establishment of self-tolerance. Assuming that Mls products are processed and presented as peptides bound to MHC molecules, one interpretation of the correlation of TCR V_{β} sequences with reactivity to Mls is that specific germline V_{β} -encoded residues are in direct contact with antigen. Since this interpretation is in direct conflict with our model for TCR interaction with the peptide/MHC complex, we will discuss an alternative explanation for these data originally suggested by Janeway and colleagues (Janeway et al. 1989).

The molecular nature of the Mls product is not known, but it has been suggested to be many things, including a unique peptide bound to an MHC molecule (Kappler et al. 1988). However, the T-cell response to Mls differs in several ways from a response to conventional antigens that are recognized as peptides embedded in MHC molecules: The response is very strong, stimulating about 20% of all T cells, and is not strictly MHC restricted (Janeway et al. 1989). Recently, the T-cell response to certain bacterial enterotoxins has been found to closely resemble the response to Mls, both in its strength and in the correlation of certain V_{β} TCR chains with T-cell reactivity, and the Mls product and the enterotoxins are now being called "superantigens" (White et al. 1989). In the case of T-cell recognition of the bacterial enterotoxins, antigen processing (i.e., degradation into peptides) is not required, and furthermore, the intact protein has been shown to bind to purified class II MHC molecules (Fraser 1989). It is therefore likely that the T-cell response to the superantigens may represent a special mode of recognition in which the intact antigen has some intrinsic affinity for TCR V_{β} subunits. The model proposed by Janeway and colleagues (Janeway et al. 1989) suggests that a superantigen (the Mls product or an enterotoxin) binds as an intact protein to the sides of the class II MHC molecule and the V_{β} TCR subunit, effectively forming a bridge between the T cell and antigen-presenting cell. The interaction between the TCR and peptide/MHC complex is thus undisturbed by the superantigen and could be as we have proposed (Davis and Bjorkman

1988 and this paper). The MHC peptide-binding site could even be occupied by an endogenous peptide, since the superantigen is proposed to bind in a region distinct from this site. Recent data from the laboratory of Mathis and Benoist support the contention that superantigens do not interact with MHC molecules like conventional peptides, in that mutations within the peptide-binding site of a class II molecule that abrogated peptide presentation had no effect on the recognition of the enterotoxin SEB (Dellabona et al., this volume). The idea that superantigens are presented to T cells in a fundamentally different way from processed peptides bound in the MHC peptide-binding site is attractive, because it is difficult to envision how the V_{β} portion of the TCR could exclusively contact a peptide bound in the binding site of an MHC molecule as we understand the architecture of this site from the crystal structure of HLA-A2 (Bjorkman et al. 1987a,b).

REFERENCES

- Amit, A.G., R.A. Mariuzza, S.E.V. Phillips, and R.J. Poljak. 1986. Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. *Science* **233**: 747.
- Barth, R.K., B.S. Kim, B.S. Lan, N.C. Lan, T. Hunkapiller, N. Sobeck, A. Winoto, H. Gershenfeld, C. Okada, D. Hansburg, I.L. Weissman, and L. Hood. 1985. The murine T cell receptor uses a limited repertoire of expressed V_{β} gene segments. *Nature* **316**: 517.
- Babbitt, B.P., P.M. Allen, G. Matsueda, E. Haber, and E.R. Unanue. 1985. Binding of immunogenic peptides to immunoglobulin histocompatibility molecules. *Nature* **324**: 317.
- Bjorkman, P.J., M.A. Saper, B. Samraoui, W.S. Bennett, J.L. Strominger, and D.C. Wiley. 1987a. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* **329**: 506.
- . 1987b. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* **329**: 512.
- Brown, J.H., T. Jardetzky, M.A. Saper, B. Samraoui, P.J. Bjorkman, and D.C. Wiley. 1988. A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules. *Nature* **332**: 845.
- Buus, S., A. Sette, S.M. Colon, C. Miles, and H.M. Grey. 1987. The relation between major histocompatibility complex (MHC) restriction and the capacity of Ia to bind immunogenic peptides. *Science* **235**: 1353.
- Chien, Y.-H., D. Becker, T. Lindsten, M. Okamura, D. Cohen, and M.M. Davis. 1984. A third type of murine T cell receptor gene. *Nature* **312**: 31.
- Chothia, C., D.R. Boswell, and A.M. Lesk. 1988. The outline structure of the T-cell $\alpha\beta$ receptor. *EMBO J.* **7**: 3745.
- Chothia, C., J. Novotny, R. Brucoleri, and M. Karplus. 1985. Domain association in immunoglobulin molecules. The packing of variable domains. *J. Mol. Biol.* **186**: 651.
- Claverie, J.-M., A. Prochnicka-Chalufour, and L. Bougueret. 1989. Implications of a Fab-like structure for the T-cell receptor. *Immunol. Today* **10**: 10.
- Colman, P.M., W.G. Laver, J.N. Varghese, A.T. Baker, P.A. Tulloch, G.M. Air, and R.G. Webster. 1987. Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. *Nature* **326**: 358.
- Davies, D.R. and H. Metzger. 1983. Structural basis of antibody function. *Annu. Rev. Immunol.* **1**: 87.
- Davis, M.M. 1985. Molecular genetics of the T cell receptor β chain. *Annu. Rev. Immunol.* **3**: 537.
- Davis, M.M. and P.J. Bjorkman. 1988. T-cell antigen receptor genes and T-cell recognition. *Nature* **334**: 395.
- Elliott, J.R., E.P. Rock, M.M. Davis, and Y.-H. Chien. 1988. The adult T cell receptor δ chain is diverse and distinct from that of fetal thymocytes. *Nature* **331**: 627.
- Engel, I. and S.M. Hedrick. 1988. Site-directed mutations in the VDJ junctional region of a T cell receptor β chain cause changes in antigenic peptide recognition. *Cell* **54**: 473.
- Fink, P.J., L.A. Matis, D.L. McElligott, M. Bookman, and S.M. Hedrick. 1986. Correlations between T-cell specificity and the structure of the antigen receptor. *Nature* **321**: 219.
- Fraser, J.D. 1989. High affinity binding of staphylococcal enterotoxins A and B to HLA-DR. *Nature* **339**: 221.
- Hedrick, S.M., E.A. Nielsen, J. Kavaler, D.I. Cohen, and M.M. Davis. 1984. Sequence relationships between putative T cell receptor polypeptides and immunoglobulins. *Nature* **308**: 153.
- Hedrick, S.M., I. Engel, D.L. McElligott, P.J. Fink, M.-L. Hsu, D. Hansburg, and L.A. Matis. 1988. Selection of amino acid sequences in the beta chain of the T cell antigen receptor. *Science* **239**: 1541.
- Janeway, C.A., J. Yagi, P.J. Conrad, M.E. Katz, B. Jones, S. Vroegop, and S. Buxser. 1989. T-cell responses to MIs and to bacterial proteins that mimic its behavior. *Immunol. Rev.* **107**: 61.
- Jerne, N.K. 1971. The somatic generation of immune recognition. *Eur. J. Immunol.* **1**: 1.
- Kabat, E.A., T.T. Wu, M. Reid-Miller, H.M. Perry, and K.S. Gottesman. 1987. *Sequences of proteins of immunological interest*, 4th edition. Public Health Service, National Institutes of Health, Washington, D.C.
- Kappler, J.W., N. Roehm, and P. Marrack. 1987. T cell tolerance by clonal elimination in the thymus. *Cell* **49**: 273.
- Kappler, J.W., U. Staerz, J. White, and P. Marrack. 1988. Self-tolerance eliminates T cells specific for MIs-modified products of the major histocompatibility complex. *Nature* **332**: 35.
- Kourilsky, P. and J.-M. Claverie. 1989. MHC restriction, alloreactivity, and thymic education: A common link? *Cell* **56**: 327.
- Kronenberg, M., G. Siu, L.E. Hood, and N. Shastri. 1986. The molecular genetics of the T-cell antigen receptor and T-cell antigen recognition. *Annu. Rev. Immunol.* **4**: 529.
- MacDonald, H.R., R. Schneider, R.K. Lees, R.C. Howe, H. Acha-Orbea, H. Festenstein, R.M. Zinkernagel, and H. Hengartner. 1988. T cell receptor V_{β} use predicts reactivity and tolerance to MIs-encoded antigens. *Nature* **332**: 40.
- Marrack, P. and J. Kappler. 1987. The T cell receptor. *Science* **238**: 1073.
- McDevitt, H.O., B.D. Deak, D.G. Shreffler, J. Klein, J.H. Stimpfling, and G.D. Snell. 1972. Genetic control of the immune response. Mapping of the Ir-1 locus. *J. Exp. Med.* **128**: 1.
- Novotny, J., S. Tonegawa, H. Saito, D.M. Kranz, and H.N. Eisen. 1986. Secondary, tertiary, and quaternary structure of the T-cell specific immunoglobulin-like polypeptide chains. *Proc. Natl. Acad. Sci.* **83**: 742.
- Patten, P., T. Yokota, J. Rothbard, Y.-H. Chien, K.-I. Arai, and M.M. Davis. 1984. Structure, expression and divergence of T cell receptor β -chain variable regions. *Nature* **312**: 40.
- Pullen, A.M., P. Marrack, and J.W. Kappler. 1988. The T-cell repertoire is heavily influenced by tolerance to polymorphic self-antigens. *Nature* **335**: 796.
- Sheriff, S., E.W. Silverton, E.A. Padlan, G.H. Cohen, S.J. Smith-Gill, B.C. Finzel, and D.R. Davies. 1987. Three-dimensional structure of an antibody-antigen complex. *Proc. Natl. Acad. Sci.* **84**: 8075.
- Shimonkevitz, R., J.W. Kappler, P. Marrack, and H.M. Grey. 1983. Antigen recognition by H-2 restricted T cells. Cell-free antigen processing. *J. Exp. Med.* **158**: 303.

- Todd, J.A., H. Acha-Orbea, J.I. Bell, N. Chao, Z. Fronek, C.O. Jacob, M. McDermott, A.A. Sinha, L. Timmerman, L. Steinman, and H.O. McDevitt. 1988. A molecular basis for MHC class II-associated autoimmunity. *Science* **240**: 1003.
- Suh, S.W., T.N. Bhat, M.A. Navia, G.H. Cohen, D.N. Rao, S. Rudikoff, and D.R. Davies. 1986. The galactan-binding immunoglobulin Fab J539: An X-ray diffraction study at 2.6 Å resolution. *Proteins* **1**: 74.
- Townsend, A.R.M., J. Rothbard, F.M. Gotch, G. Behador, D. Wraith, and A.J. McMichael. 1986. The epitopes of influenza nucleoprotein recognized by cytotoxic T lymphocytes can be defined with short synthetic peptides. *Cell* **44**: 959.
- White, J., A. Herman, A.M. Pullen, R. Kubo, J.W. Kappler, and P. Marrack. 1989. The V β -specific superantigen staphylococcal enterotoxin B: Stimulation of mature T cells and clonal deletion in neonatal mice. *Cell* **56**: 27.
- Winoto, A., J.L. Urban, N.C. Lan, J. Goverman, L. Hood, and D. Hansburg. 1986. Predominant use of a V α gene segment in mouse T-cell receptors for cytochrome c. *Nature* **324**: 679.
- Wu, T.T. and E.A. Kabat. 1970. An analysis of the sequences of the variable regions of Bence-Jones and myeloma light chains and the implications for antibody complementarity. *J. Exp. Med.* **132**: 211.
- Yanagai, Y., Y. Yoshikai, K. Leggett, S.P. Clark, I. Aleksander, and T.W. Mak. 1984. A human T cell-specific cDNA clone encodes a protein having extensive homology to immunoglobulin chains. *Nature* **308**: 145.
- Zinkernagel, R.M. and P.C. Doherty. 1974. Restriction of in vitro T-cell mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semi-allogeneic system. *Nature* **248**: 701.