Principles of MHC class I-mediated antigen presentation and T cell selection

H.-G. Ljunggren and C.J. Thorpe
Microbiology and Tumor Biology Center, Karolinska Institute, Stockholm, Sweden

Summary. Class I molecules of the major histocompatibility complex (MHC) are expressed on the cell surface of almost all nucleated mammalian cells. Their main function is to transport and present peptides, derived from intracellularly degraded proteins, to cytotoxic T cells (CTL). They are also directly involved in the process leading to maturation and selection of a functional CD8+ T cell repertoire. MHC class I molecules consist of a highly polymorphic membrane-spanning heavy chain of approximately 45 kD that is non-covalently associated with a light chain, β2-microglobulin (B2m). Class I molecules bind peptides, usually 8-11 amino acids in length. The majority of the class I-bound peptides are generated in the cytosol and are subsequently translocated into the lumen of the endoplasmic reticulum (ER) through the ATP-dependent transporter associated with antigen processing 1/2 (TAP1/2). Here, we provide an up-to-date review summarizing the most essential parts relating to MHC class I-mediated antigen processing, presentation and T cell selection. A particular emphasis is devoted to the structure of MHC class I molecules.

Key words: MHC class I, Antigen processing, Antigen presentation, T cell selection, Cytotoxic T lymphocyte

Structure of MHC class I molecules

MHC class I molecules are expressed on the cell surface of almost all nucleated mammalian cells. They consist of a highly polymorphic MHC-encoded membrane-spanning heavy chain of approximately 45 kD that is noncovalently associated with a light chain, β2-microglobulin (B2m) (Bjorkman et al., 1987; reviewed in Bjorkman and Parham, 1990). Class I molecules have four domains, three of which are formed by the class I heavy chain and one formed by B2m (Fig. 1). The α-3 domain of the heavy chain as well as B2m have a folded structure that closely resembles that of immunoglobulins. In contrast, the α-1 and α-2 domains of the heavy chain form two α-helices, topping a sheet of eight β-strands. This structure forms a cleft in which peptide antigens, normally 8-11 amino acids in length, can bind. The complex between peptide antigen and MHC constitutes the structural unit that is recognized by the T cell receptor (TCR; Fig. 2).

Peptide binding to MHC class I

The refined crystallographic structures of MHC class I molecules have revealed the detailed architecture of the peptide binding groove (reviewed by Madden, 1995). All class I structures analyzed to date have a closed peptide binding groove and conserve features that hold onto the peptide termini (see below). As a consequence, peptide binding by classical class I gene products usually requires free NH2 and COOH-termini. The bound peptides display a narrow size distribution encompassing 8-11 amino acids (reviewed by Rammensee et al., 1993). Peptides that bind to class I molecules are tightly bound primarily by virtue of contacts to the peptide’s amino acid side chains with the class I molecule. Pockets along the groove (designated A through F) may accommodate predominant amino acid side chains of the peptide, thereby anchoring the peptide onto the class I molecule (Madden, 1995). While the A and F pockets are fairly
Fig. 1. Three dimensional structure of the extracellular portion of an MHC class I molecule, represented here by the mouse allele H-2Kb. The molecule is comprised of a heavy chain (red) consisting of three domains, a single domain light chain (green) and a peptide (yellow). The peptide and the light chain, β2-microglobulin, are non-covalently attached to the heavy chain. These three units fold to form a compact structure which is easily visualized in panels a and b. A ribbon trace of the molecule is presented in panels c and d and this representation clearly shows the architecture of the molecule, with the «vice-like» peptide binding groove composed of two α-helices and a β-pleated sheet, clearly visible atop the two immunoglobulin-like domains. The side view presented in panel d demonstrates the slightly skewed symmetry of the molecule which may play a role in the recognition of the molecule by the T cell receptor (TCR). A slightly asymmetric molecule will ensure a greater number of productive engagements of the TCR.
well conserved, B through E have distinct sizes and character in different allelic variants of MHC class I molecules, thereby imposing different sequence constraints on the bound peptide. A consequence of this is that class I binding peptides contain allele specific sequence motifs, defined by the position and the identity of a couple of «anchoring» residues, one of which is the C-terminus (Rammensee et al., 1993).

The N- and C-termini of the peptides are almost always located in identical orientations. The termini are rigidly fixed in this position by a conserved network of hydrogen bonding ligands and water molecules. In the majority of cases studied, the peptides are prevented from extending from the cleft by large «walls» consisting of conserved residues (Fig. 3). A comparison of different peptides (Fig. 4) clearly shows that whilst the N- and C-termini of peptides bind in a similar manner, regardless of which allele they are bound to, they deviate dramatically in the center of the cleft. For example, the H-2Kb- and H-2Db-bound peptides sequester a central anchor in the cleft, whereas the peptides bound to the HLA-A2 molecule bulge out of the cleft in the center. Furthermore, the H-2Kb- and H-

Fig. 2. Predicted structure of the T cell receptor (TCR):MHC:peptide superassembly. The TCR sequences bear a striking resemblance to those of Fab fragments upon which the model is based. It is widely believed that the most diverse regions, the CDR3 regions, produced by recombination of V (D), and J segments are those which primarily recognize the peptide antigen bound in the jaws of the MHC molecule. The less diverse CDR1 and CDR2 regions are presumed to recognize the less diverse, but nevertheless polymorphic a1 and a2 helices of the presenting MHC molecule.

Fig. 3. Top view of the MHC class I peptide binding site demonstrating the integral role of the peptide in forming the structure. In essence the peptide forms the core of a zip, holding the two helices in position.
Peptides binding to MHC molecules conform approximately to structural patterns that are partly dependent on the cleft architecture of the allele to which the peptide is bound. This pattern is, to a certain, but not to an exclusive extent, dependent on the peptide length. Panel a shows peptides bound to the mouse molecule, H-2Kb, and panel b portrays peptide bound to the mouse molecule, H-2Db, and panel c represents peptides bound to the human molecule HLA-A2.

Fig. 4. The shape of the MHC-bound peptide.
2D\textsuperscript{b}-bound peptides differ in the location of their bulge. The H-2K\textsuperscript{b} molecule bulges before the central anchor, the H-2D\textsuperscript{b} molecule after the anchor. This feature is strongly correlated with the presence of a large hydrophobic ridge in the cleft of the H-2D\textsuperscript{b} molecule which the peptide must circumnavigate to bury both its central and C-terminal anchor. Thus, crystallographic analysis demonstrate a clear correlation between bound peptide shape and MHC encoded cleft architecture.

**MHC class I-mediated antigen presentation**

Degradation of endogenous proteins in the cells results in the generation of short peptides that associate with class I molecules of the MHC. Class I/peptide complexes are then presented at the cell surface for recognition by cytotoxic T cells (CTL). The molecular events underlying this process are referred to as antigen processing and presentation (reviewed in Yewdell and Bennink, 1992). The basic events in this process are summarized below and schematically illustrated in Figure 5.

Much of the available information about class I-restricted antigen processing and presentation has been identified by the study of mutant cell lines, such as the murine RMA-S and the human.\textsuperscript{174} (and its derivative T2) and .\textsuperscript{134} cell lines. These cell lines, when infected with virus, fail to be recognized by antigen-specific CTL (Townsend et al., 1989; Hosken and Bevan, 1990). They accumulate «empty» class I heavy chain/\(\beta\)2m complexes in the ER, most of which fail to leave the ER. Consequently, these cells express very few MHC class I molecules at the cell surface. Despite being low in numbers, these class I molecules readily bind synthetic class I-binding peptides when added to the culture medium. Addition of such presentable peptides restores recognition of these mutant cells by CTL and, at least in part, reconstitutes class I cell surface expression (Townsend et al., 1989; Ljunggren et al., 1990). Since the stability of «empty» class I molecules is enhanced at low temperature, culture of these cells at reduced temperature can also partially restore cell surface expression (Ljunggren et al., 1990).

The defect in these mutant cell lines has been

---

**Fig. 5.** Schematic summary of MHC class I-mediated antigen presentation. Endogenous proteins, including proteins synthesized after for example, a virus infection, are degraded to peptide fragments by proteolytic machineries such as proteasomes. Peptides are then translocated over the ER membranes by the TAP1/2 transporters. In the ER they assemble with class I molecules of the MHC, and these complexes are then transported via the Golgi apparatus to the cell surface where they can be recognized by CD8\textsuperscript{+} CTL.
Specificity of TAP for a particular C-terminal amino acid, whereas mouse binding specificity of MHC class I proteins. Further efficient transport requires free N- and C-terminal residues, whereas less is known about the ER lumen. Formal proof for the transporter function of TAP1/2 has more recently come from in vitro peptide translocation assays with microsomes from wild-type and TAP1 mutant mice (Sheperd et al., 1993), or with permeabilized wild-type and T2 mutant cells (Androlewicz et al., 1993; Neefjes et al., 1993).

Though it is clear that TAP transporters provide the ER lumen with a majority of the peptides that subsequently bind to MHC class I molecules, some peptides may reach the class I molecules via other pathways (reviewed by Heemels and Ploegh, 1993). One source of such peptides is that derived from signal sequences, which have been released from proteins translocated into the ER via the signal sequence recognition particle (Henderson et al., 1992; Wei and Cresswell, 1992). Other sources may be peptides that accumulate in the ER as a result of degradation of ER resident proteins or peptides that have been translocated over the ER membrane by as yet unknown transporter complexes. In addition, other pathways of TAP-independent antigen presentation may exist. The presentation of Sendai virus antigens in T2 cells may represent one such pathway (Zhou et al., 1993; Liu et al., 1995). The presentation of this antigen is not only independent of the TAP transporter, but also insensitive to the drug Brefeldin A, which blocks egress of proteins from the ER to the cell surface.

Specificity of peptide transport and antigen processing

The specificity of the TAP transporter has been characterized with respect to length, charge, hydrophobicity and sequence of peptide substrates as well as interallelic and interspecies differences (reviewed by Römisch, 1994). A peptide length of at least eight to nine amino acids is generally needed for transport to occur whereas less is known about the absolute limits with regard to maximum peptide length. Efficient peptide transport requires free N- and C-termini. The preferred N-terminal amino acids are polar, small or charged. Peptides with phenylalanine and glutamic acid in the second to last position are the least efficiently transported. Human TAPs prefer substrates with hydrophobic or charged amino acid residues at the C-terminus, whereas mouse TAPs selectively transport peptides with hydrophobic C-terminal amino acids. The specificity of TAP for a particular C-terminal amino acid in the peptide substrate generally agrees with the known binding specificity of MHC class I proteins. Furthermore, it is now generally accepted that TAP-dependent transport of peptides requires ATP hydrolysis (Androlewicz et al., 1993; Neefjes et al., 1993; Sheperd et al., 1993; reviewed by Römisch, 1994).

Detailed knowledge about the proteolytic machinery that is responsible for the generation of class I MHC-binding peptides is still limited. Recent studies suggest a role for large proteolytic enzyme complexes, termed proteasomes, in this process (reviewed by Goldberg and Rock, 1992; Driscoll, 1994; Rubin and Finley, 1995). Proteasomes consist of 13-15 distinct subunits and contain multiple active sites that catalyze peptide bond cleavage on the carboxyl side of hydrophobic, basic and acidic amino acid residues. Two subunits of the proteasome, termed low molecular weight protein (LMP) 2 and 7, have been found to be of particular interest since they are encoded within the MHC class II region, in close proximity to the TAP1 and 2 genes. LMP2 and 7 associate with the proteasome and are legitimate members of the proteasome gene family since they exhibit sequence similarities with other subunits. They are inducible by interferon-α and exhibit amino acid sequence polymorphisms, features shared with the MHC class I molecules and the TAPs. Their precise role in antigen processing is unknown, but incorporation of LMP2 and 7 into proteasomes alters the cleavage specificity of the complexes in a way that is predicted to be beneficial for antigen processing. This idea has been supported by studies with cell lines and mice that harbor mutations in the LMPs (Howard and Seelig, 1993; Fehling et al., 1994; Van Kaer et al., 1994).

The exact nature of the peptides that are generated by proteasomes is unknown. It has been proposed that the proteasome degrades peptides with an extended N-terminus, which is trimmed after transport to the ER (Rammensee et al., 1993). Rather little is known about how peptides, after generation in the cytosol, reach the TAP transporters and how they reach, after translocation over the ER membrane, the class I molecules.

T cell selection

The immune system needs to acquire the capacity to recognize a spectrum of potential pathogens, none of which it has encountered during its development. Recent studies have given a remarkable insight into these events, even though they are as yet far from completely understood at a molecular level.

T cell precursors are derived from the bone marrow and mature in the thymus. During this maturation T cells go through a complex, and yet not completely understood, selection process where useless and harmful T cells are deleted while useful T cells are spared and released to a life in the periphery. Two types of cellular selection in the thymus are central to the repertoire determination of T cells (Janeway, 1994). Firstly, T cells acquire the capacity to recognize antigens in the context of MHC molecules through a process referred to as positive selection (reviewed in von Boehmer, 1994; Jameson et al., 1995). This process is critically...
dependent on interactions of the TCR with MHC molecules expressed by the selecting cells of the thymus. Engagement of the TCR with MHC class I glycoproteins induces immature T cells to differentiate into CD48+ cytotoxic T cells, whereas TCR engagement with class II molecules induces differentiation to the CD4+8+ helper T cell lineage. Secondly, through negative selection, potentially autoreactive T cells are eliminated from the mature T cell repertoire (reviewed by Nossal, 1994).

Since peptides are an integral part of the MHC molecules, it is likely that the MHC molecules which are recognized by the TCRs of immature thymocytes during both positive and negative selection are occupied by self-peptides. This raises two central questions (Janeway, 1994) about the fate of a thymocyte during T cell selection. (i) What is the role of self-peptides in T cell self-peptide selection? This raises two central questions (Janeway, 1994) about the fate of a thymocyte during T cell selection. (i) What is the role of self-peptides in T cell positive selection? i.e. Does the type of peptide bound with MHC molecules expressed by selecting cells contribute to the specificity of T cell selection? (ii) What determines whether a thymocyte undergoes positive versus negative selection?

The observation that synthetic class I binding peptides can partially restore expression of MHC class I molecules on the surface of TAP1 deficient cells, and partially also B2m deficient cells, opened the possibility to use fetal thymus organ cultures (FTOC) to directly address the questions imposed in the previous section (Ashton-Rickardt et al., 1993; Hogquist et al., 1993). FTOC assays allow one to manipulate class I expression in the fetal thymus and to eliminate the enormous complexity caused by the presence of self peptides that are normally bound with class I molecules. For simplicity, we will here focus on studies with TAP1-/— mice (Van Kaer et al., 1992), and in the end compare these studies with similar studies performed with B2m-/— mice.

Single peptides, when added to the organ cultures from TAP1 mutant thymii were able to reconstitute surface expression of class I molecules on thymic epithelial cells, but only a subset of these peptides was able to induce positive selection of CD8+ T lymphocytes. Furthermore, more complex mixtures of peptides were very effective at inducing T cell positive selection while being rather inefficient in stabilizing class I cell surface expression. These data indicate that peptide does not simply serve to stabilize the class I structure during T cell positive selection but that it contributes to the specificity of this process. Since the epithelial cells in the thymus express MHC molecules that are bound with more than one peptide it is conceivable than in vivo every thymocyte is selected on more than one peptide. The overall avidity of this interaction might therefore determine whether the thymocyte will be positively selected or not. In order to address the factors that determine the outcome of positive and negative selection, TAP1 mutant mice were mated with a strain of mice transgenic for a TCR derived from a lymphocytic choriomeningitis virus (LCMV) peptide-specific H-2Dβ-restricted T cell clone (Ashton-Rickardt et al., 1994). Positive selection of the transgenic T cells was impaired in these mice. In FTOC from these mice, transgenic T cells were positively selected when low concentrations of nascent LCMV peptide were added to the culture, whereas at higher concentrations of this peptide negative selection occurred. Increased peptide concentration of this peptide negative selection occurred. Increased peptide concentration serves to stabilize more H-2Dβ molecules at the cell surface, thereby increasing the amount of ligand that is available for interaction with the TCR. It was therefore concluded that the overall avidity between the TCRs of the thymocytes and the MHC/peptide complexes of the selecting cells in the thymus determines the fate of a thymocyte during T cell selection. This has led to the formulation of an avidity model for T cell selection (Ashton-Rickardt et al., 1994).

Studies performed in B2m deficient mice have similarly demonstrated a critical role for peptide in T cell selection (Hogquist et al., 1993, 1994; Sebzda et al., 1994). In contrast to studies performed by Tonegawa and colleagues in the TAP1/-/ model and Ohashi and colleagues in the B2m-/— model, Bevan and colleagues (also using B2m deficient mice in their studies) have failed to detect any positive selection by strong agonist/ligands. These studies have strongly supported the notion that selection on peptides that act as antagonists or weak agonist/antagonists for mature T cells drives positive selection (Hogquist et al., 1995). Such peptide ligands normally differ from the nominal antigen in one or two residues. The molecular basis for these differences, and the natural ligands for positive selection of T cells, are currently not understood at the molecular level.

Conclusion

The major pathway for class I-mediated antigen processing and presentation has been elucidated during the last five years. The crystal structures of the first MHC class I molecules have likewise provided a structural basis for antigen presentation. The scientific community is currently eagerly awaiting a complete crystal structure of the T cell receptor, and the T cell receptor complexed with an MHC molecule. In a similar way, the elucidation of the molecular basis for T cell selection, including intracellular signaling events in T cells upon interaction with antigen, still awaits complete resolution. It is most likely that insights into these events will be reached during the coming five years. In addition, MHC class I molecules have recently been shown to be of importance in relation to NK cell development and target cell recognition (Ljunggren and Kärre, 1990; Kärre, 1995), and much interest is currently directed towards this area of research as well.

Acknowledgements. We thank the members of our groups for stimulating discussions. Hans-Gustaf Ljunggren is supported by the Swedish Medical Research Council and the Swedish Cancer Foundation.
MHC class I structure and function

References


