4. MHC & Antigen Presentation

Background

- The MHC molecules are glycoproteins encoded in a large cluster of genes located on chromosome 6. They were first identified by their potent effect on the immune response to transplanted tissue (see later).
- For that reason, the gene complex was termed the ''major histocompatibility complex.'' MHC genes (called the H-2 complex in mice) were first recognized in 1937 as a barrier to transplantation in mice.
- In humans, these genes are often called human leukocyte antigens (HLA), as they were first discovered through antigenic differences between white blood cells from different individuals.
- MHC is the term for the region located on the short arm of chromosome 6p21.31 in humans and chromosome 17 in mice. In humans, it contains more than 200 genes (Figure 1).
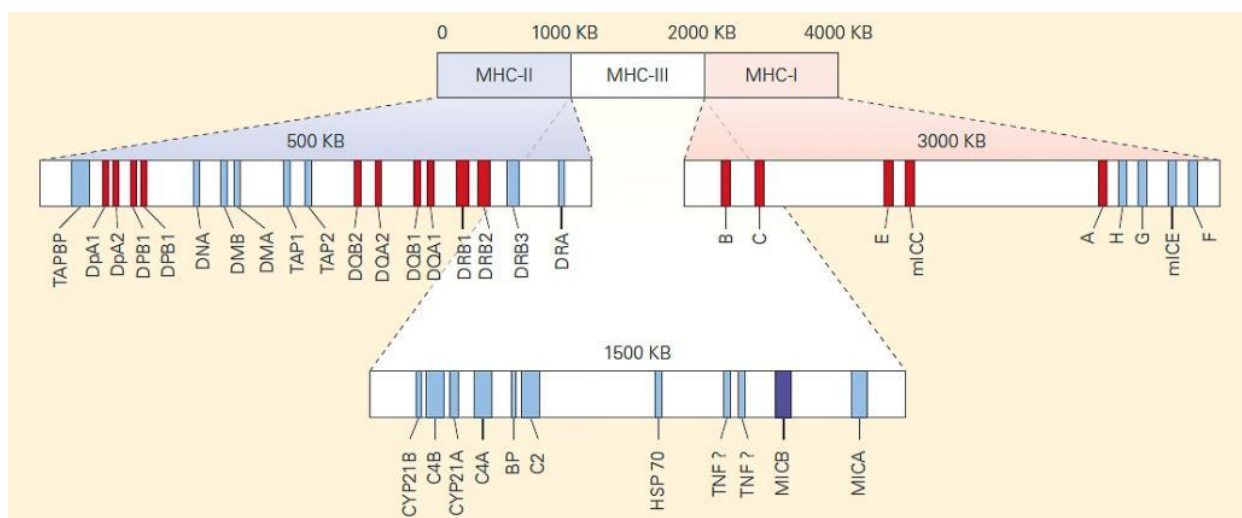

Figure 1. Genetic map of the MHC regions. This map has been simplified to demonstrate organizational themes within the MHC. There are more than 200 genes within these

regions. [Bellanti, JA (Ed). Immunology IV: Clinical Applications in Health and Disease. I Care Press, Bethesda, MD, 2012]

- The principal function of the MHC is to present antigen to T cells to discriminate between self (our cells and tissues) and nonself (the invaders or modified self).
- Two main characteristics of the MHC make it difficult for pathogens to evade immune responses:
  - First, the MHC is polygenic. It contains several different MHC-I and MHC-II genes so that every individual possesses a set of MHC molecules with different ranges of peptide-binding specificities.
  - Second, the MHC is extremely polymorphic. The MHC genes display the greatest degree of polymorphism in the human genome. There are multiple variants of each gene within the population as a whole. The different variants that are inherited by an individual from a parent are known as alleles. The numbers of alleles recognized at the classical loci are presented in Table 10-1.
- Polymorphic sites are found predominantly in specific regions of the MHC-I and MHC-II molecules called domains.
- Although each HLA molecule shows slight differences in its amino acid sequence from one another, causing a slightly altered three-dimensional structure in the peptide-binding cleft, the basic structures of MHC-I and MHC-II molecules are very similar
- Figure 2 shows the similarity in structure between the two classes, although the manner in which the peptide is bound and presented in the binding cleft differs between Class I and II (explained below)
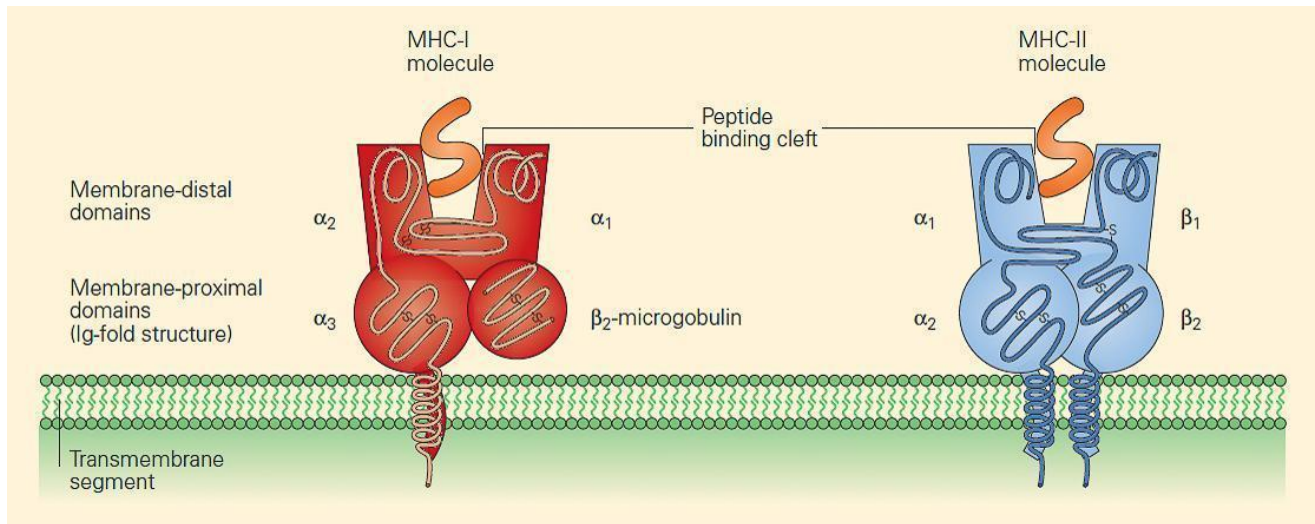
Figure 2. MHC-I and MHC-II molecules have a very similar structure. In each case, a cleft or groove is formed that cradles the pep-tide. The charge characteristics of the groove determine which peptides can be presented. [Bellanti, JA (Ed). Immunology IV: Clinical Applications in Health and Disease. I Care Press, Bethesda, MD, 2012]

- The charge characteristics of the groove determine which peptides can be presented. Since different antigenic peptides have different shapes and charge characteristics, it is important that the human population overall has a large array of different HLA molecules, each with different shaped peptide-binding areas (clefts) to cope with the multitude of self and nonself peptides presented.

MHC Structure and Function

- The MHC has three regions: MHC-I, MHC-II, and MHC-III (Figure 1).
- The classical HLA antigens encoded in each region include HLA-A, -B, and -C in the MHC-I region, and HLA-

DR, -DQ, and -DP in the MHC-II region.

- The MHC-III region includes several genes involved in the complement cascade (C4A, C4B, C2, and FB) (see section 6, Complement), the TNF-a and TNF-b (LTa) genes, the CYP21 gene that encodes an enzyme in steroid metabolism, the HSP70 gene that encodes a chaperone, and many other genes of unknown immunological function.
- In general, when we refer to MHC, we are referring to either MHC-I or MHC-II molecules.
- Shown in Figure 3 is a schematic representation of the chromosomal locations and genetic loci responsible for MHC-I and MHC-II synthesis.
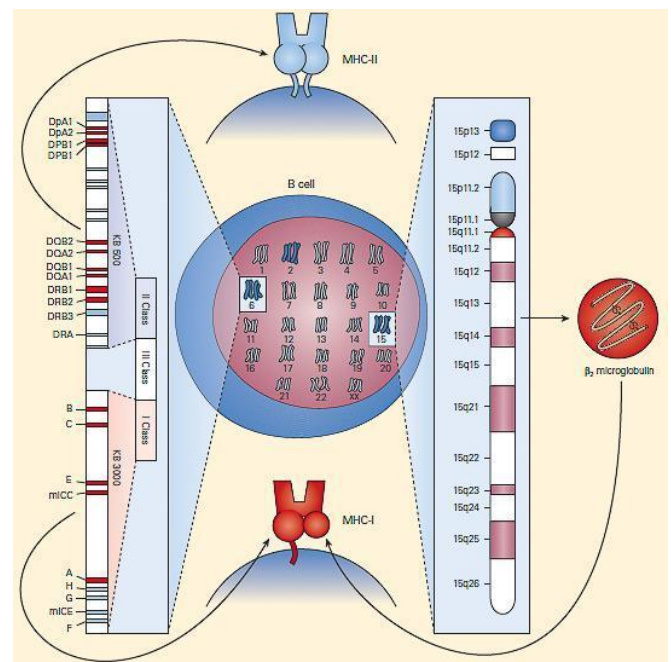

Figure 3. Schematic representation of the chromosomal location and genetic loci responsible for MHC-I and MHC-II synthesis. [Bellanti, JA (Ed). Immunology IV: Clinical Applications in Health and Disease. I Care Press, Bethesda, MD, 2012]

- **MHC-I molecules** consist of two polypeptide chains, a larger a chain encoded on chromosome 6 in the MHC region and a smaller b2 microglobulin encoded on chromosome 15 (Figures 2 and 3).
- The class I a chains consist of a single polypeptide composed of three extracellular domains named a_1, a_2, and a_3, a transmembrane region that anchors it in the plasma membrane, and a short intracytoplasmic tail (Figure 2).
- The b_2 microglobulin consists of a single non-polymorphic molecule noncovalently bound to the alpha chain and is encoded on chromosome 15 (Figure 2 and Figure 3). The a_1 and a_2 domains fold together into a single structure consisting of two segmented a helices lying on a sheet of eight antiparallel b strands.
- The folding of the a_1 and a_2 domains creates a long cleft or groove that is the site at which peptide antigens bind to the MHC-I molecule and are presented to the CD8 lymphocyte.
- **MHC-II molecules** consist of two polypeptide chains, a and b, both encoded in the MHC-II region on chromosome 6 and noncovalently linked to one another (Figure 2 and Figure 3).
- The a and b chains each consist of two extracellular domains referred to as a_1 and a_2 and b_1 and b_2, respectively, and, similar to the MHC-I a chain, the a and b chains of the MHC-II molecule also consist of a transmembrane segment and a cytoplas-mic tail (Figure 2).
- The extracellular membrane-proximal a_2 and b_2 domains are homologous to immunoglobulin-constant domains.
- The crystallographic structure of the MHC-II molecule shows that it is folded very much like the MHC-I molecule (Figure 4).
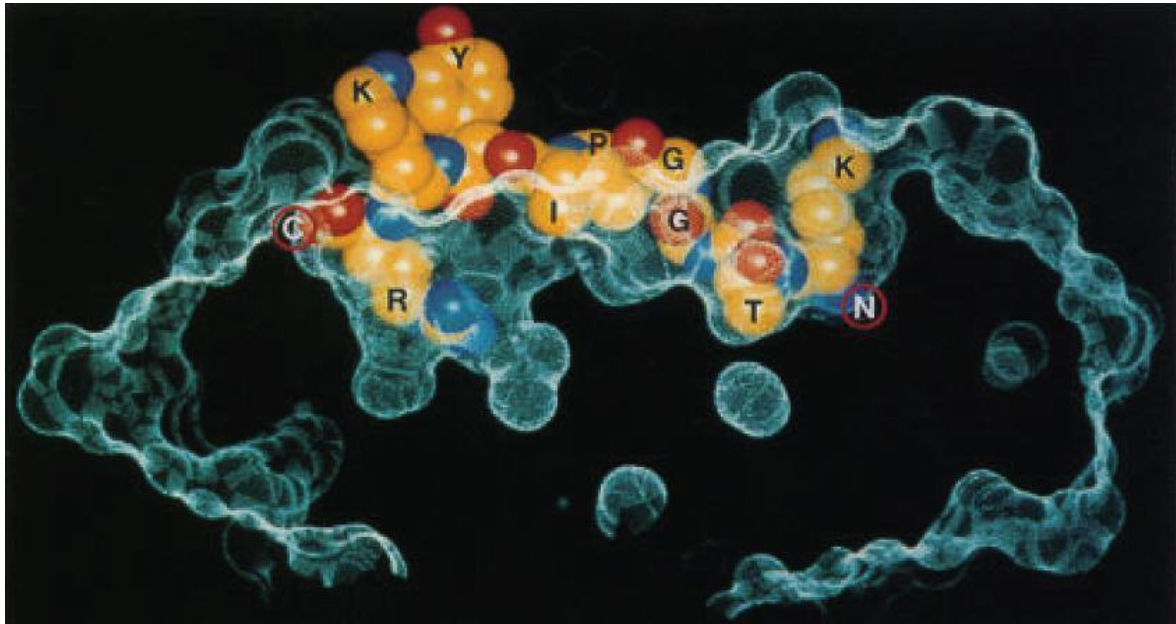
Figure 4. Structure of an HLA-DQ molecule. An influenza virus nucleoprotein peptide (KTGGPIYKR) bound to HLA-A*6801, shows insertion of Thr (T) and Arg (R) buried in specificity pockets of the HLA molecule. (Reproduced with permission from Guo HC, Madden DR, Silver ML, et al. Comparison of the P2 specificity pocket in three human histocompatibility antigens: HLA-A*6801, HLA-A*0201, and HLA-B*2705. Proc Natl Acad Sci USA. 1993;90:8053–7.) [Bellanti, JA (Ed). Immunology IV: Clinical Applications in Health and Disease. I Care Press, Bethesda, MD, 2012]

- The major differences between the two MHC class molecules lie at the ends of their peptide-binding clefts, which are more open in MHC-II molecules compared with MHC-I molecules. The MHC-II molecule cleft is made up of a noncovalent association between the a_1 and b_1 domains and that binds the peptide through multiple van der Waals forces and hydrogen bonds (Figure 5).
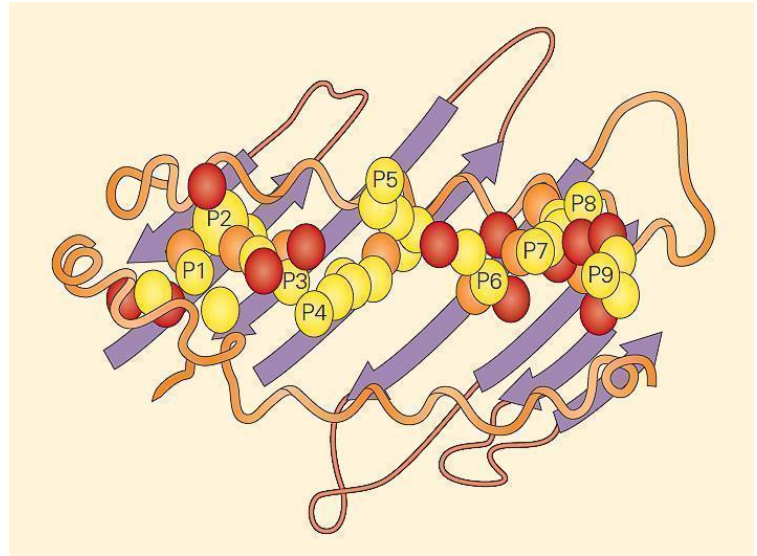
Figure 5. An example of a peptide held within an MHC-II groove. The fit of the peptide within the groove is very specific. The MHC-II molecule cleft is made up of a noncovalent association between the a1 and b1 domains that bind the peptide through multiple van der Waals forces and hydrogen bonds. The a1 and b1 domains are shown lying on a sheet of eight antiparallel b strands. The folding of the a1 and b1 domains creates a long cleft or groove that is the site at which peptide antigens bind to the MHC-II molecule and are presented to the CD4 lymphocyte. [Bellanti, JA (Ed). Immunology IV: Clinical Applications in Health and Disease. I Care Press, Bethesda, MD, 2012]

- The main consequence of this difference is that the ends of a peptide bound to an MHC-I molecule are buried within the molecule whereas the ends of peptides bound to MHC-II molecules are not.

- This difference allows more flexibility in the length and types of peptides that MHC-II molecules can bind. Peptides that bind a specific class II molecule will share the same middle anchor residues but may vary in length and sequence of other residues.

Expression of MHC Molecules MHC-I

- MHC- I proteins are expressed on all nucleated cells, in contrast to MHC-II molecules, which are restricted to antigen-presenting cells (APCs)
- Lymphocytes, macrophages, dendritic cells, Langherans cells, and some endothelial cells are the predominant cells that express MHC-II.
- Nonnucleated cells such as mammalian red blood cells express little or no MHC-I and thus, pathogens within red blood cells can go undetected by cytotoxic T cells, e.g., malaria.
- The expression of both MHC-I and MHC-II molecules is regulated by cytokines.
- Interferon-g (INF-y) increases the expression of MHC-I or MHC-II molecules and can induce the expression of MHC-II molecules on certain cell types that do not normally express them. This may be very important both in normal immunologic function and in autoimmunity.
- The level of MHC molecule expression plays an important role in T cell activation and therefore differences in levels of expression are significant.
- Shown in Table 1 is a comparison of the principal differences between MHC-I and MHC-II molecules.

Table 1. Features of MHC-I and MHC-II molecules

Feature	MHC-I	MHC-II
Polypeptide chains	A single a chain (44–47 kD) noncovalently linked to the b2-microglobulin chain (12 kD)	A single a chain (32–34 kD) non-covalently linked to a single b chain (29–32 kD)
Distribution	All nucleated cells	Antigen-presenting cells
Composition of antigen-binding clefts	α1 and α12 domains	α1 and β1 domains
Binding site for T cell co-receptor	CD8 binds to the α3 region	CD4 binds to the β2 region
Size of peptide-binding cleft	Accommodates peptides of 8–11 residues	Accommodates peptides of 13–25 residues or more
Nomenclature in the human	HLA-A, HLA-B, HLA-C	HLA-DR, HLA-DQ, HLA-DP

Antigen Presentation

- T cells recognize foreign antigens in the form of short peptides that have been processed and dis-played on the cell surface bound to MHC-I or MHC-II molecules (Figure 5).
- Antigens are often categorized according to whether they are derived from (1) viruses, intracellular bacteria, or protozoan parasites (endogenous pathogens); or (2) exogenous pathogens that replicate outside of the cell.
- Intracellular antigens are presented to T cells by any nucleated cell because MHC-I expression is ubiquitous.
- In contrast, exogenous antigens are taken up by professional APCs, which process the antigens and

present them in the context of MHC-II. An important function of a professional APC, e.g., dendritic cell (DC), is to deliver a second signal (costimulation) to the T cell to alert it to the presence of infection.

- **Endogenous antigens,** including misfolded proteins and pathogen-derived peptides, are processed by the proteasome (Figure 6A).
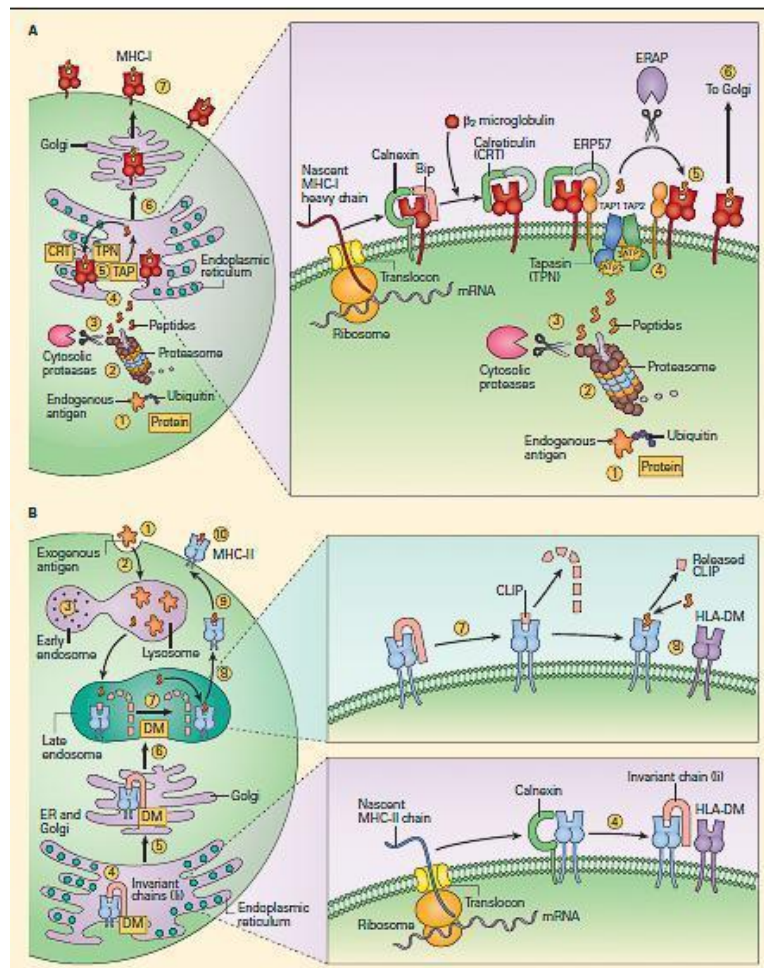

Figure 6. Peptide loading of MHC-I and MHC-II molecules. Panel A: shows the synthesis and peptide loading of MHC-I through the endogenous pathway. Endogenous proteins (e.g., a self-protein or a viral protein) synthesized in the cytoplasm are modified initially by ubiquitin (1), following which they are processed by the proteasomes (2). After trimming

by cytosolic proteases (3), the peptides enter the endoplasmic reticulum via the TAP 1 and TAP 2 transporters (4). The MHC-I alpha chain, which is initially formed as a linear peptide in the ER, is then folded with the help of several chaperones (calnexin, calreticulin [CRT]). Binding immunoglobulin protein (BiP) and endoplasmic reticulum protein 57 (ERP57), during which the b2 microglobulin is added to the alpha chain, complete the synthesis of the complete MHC-I molecule (right inset in panel A). The complex is held together by tapasin (TPN), which facilitates transfer of the peptide to the antigen-binding cleft (5). The peptide-loaded MHC-I complex is then transferred to the Golgi (6) and then transported to the surface of the cell (7). Panel B: Shows the uptake of protein and peptide loading of MHC-II through the exogenous pathway. Exogenous proteins are taken up (1) and processed in the early endosomal compartment (2) and cleaved into peptides by cathepsins and other acid proteases (3). MHC-II molecules are formed in the endoplasmic reticulum with the help of the chaperone calnexin (4) and are held ready by the invariant chain (li); the complex is later fused with the HLA-DM (DM) (right lower inset in panel B). After passage of the li-loaded MHC-II-DM

complex through the Golgi (5) into the late endosomes (6), the invariant chain is cleaved by acid proteases, leaving a residual peptide referred to as the class II-associated invariant chain peptide (CLIP) (7) in the MHC-II cleft (right upper inset in panel B). The HLA-DM facilitates the insertion of the peptide in the MHC-II cleft replacing CLIP (8). The MHC molecule loaded with peptide is transported (9) and expressed on the cell surface (10). [Bellanti, JA (Ed). Immunology IV: Clinical Applications in Health and Disease. I Care Press, Bethesda, MD, 2012]

- This complex of proteases typically generates peptides of four to twenty amino acids with a hydrophobic carboxy terminus. After trimming of the peptide by cytosolic proteases, the antigenic peptides are translocated to the endoplasmic reticulum by the transporters associated with antigen processing (TAP1 and TAP2 molecule).
- Meanwhile, a new MHC-I molecule is being synthesized in the endoplasmic reticulum.
- As it folds, it is bound by calnexin, which is then replaced by calreticulin and b_2
- The new MHC-I molecule associates with the MHC-I peptide loading complex. Tapasin physically links the MHC-I molecules and the TAP transporters. As the peptide enters from the cytosol, the cleft of the class I molecule receives it and the peptide-bound MHC-I molecule dissociates from the peptide-loading complex and is recruited to the cell surface.

- This complex machinery has several quality-control steps, such that MHC-I molecules that fail to assemble properly are degraded. Ultimately, the peptide presented on MHC-I molecule will stimulate a CD8 T cell response (Figure 7)
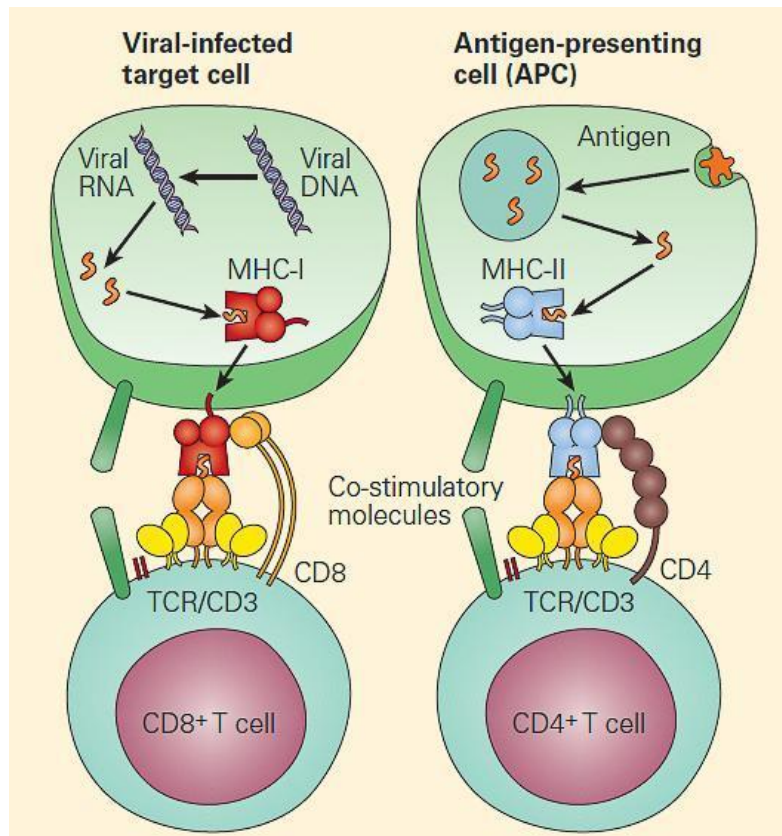

Figure 7. Endogenous antigens are generally presented to CD8þ T cells (left panel), and exogenous antigens are generally presented to CD4þ T cells (right panel). [Bellanti, JA (Ed). Immunology IV: Clinical Applications in Health and Disease. I Care Press, Bethesda, MD, 2012]

- **Exogenous antigens** are processed quite differently (Figure 6B).
- Bacterial proteins are cleaved by proteases, cathepsins,

- and metalloproteases in the acid environment of the endocytic pathway.
- Meanwhile, MHC-II molecules assemble in the endoplasmic reticulum with another molecule called the invariant chain (li). The newly synthesized molecules transit the endoplasmic reticulum and the Golgi apparatus.
- After passage of the li-loaded MHC-II-DM complex through the Golgi into the late endosomes, the invariant chain is cleaved by acid proteases, leaving a residual peptide referred to as the CLass II-associated Invariant chain Peptide (CLIP) in the MHC-II cleft.
- CLIP occludes the MHC-II cleft and prevents peptides from loading until the molecule is in the lysosomal or late endosomal compartment containing the peptides.
- At that point, HLA-DM molecules remove CLIP from the cleft and stabilize the molecule while the peptide is loaded into the cleft.
- HLA-DO molecules can also facilitate this process in some settings.
- The fully assembled and loaded MHC-II molecule is recruited to the surface and serves to stimulate predominantly CD4 positive T cells (Figure 6B and Figure 7).

Watch Video of Antigen Presentation by MHC II:

/var/www/html/wp-content/uploads/2015/08/AAPC_TCell_Final1.flv
Reproduced with permission from Bellanti, JA (Ed). Immunology IV: Clinical Applications in Health and Disease. I Care Press, Bethesda, MD, 2012.

Nomenclature of HLA

- The HLA nomenclature has developed historically from the original serological designations. Polymorphisms in proteins were originally defined by antibody reaction patterns. Modern definitions utilize DNA sequences to define alleles. The current nomenclature was recommended during the Tenth International Histocompatibility Workshop in 1987, with minor modifications added in 1990.
- Each chromosome is found twice (diploid) in each individual, and therefore a normal tissue type of an individual will involve twelve HLA antigens (three HLA class I loci [A, B, and C] from each parent and three class II loci [DR, DQ, and DP] from each parent).
- HLA-DM and HLA-DO are not highly polymorphic and are not typed. These twelve antigens are inherited co-dominantly.
- The MHC phenotype of a person describes which alleles the person carries without reference to inheritance. For example, someone might be typed as HLA-A1, -A3; B7, B8; Cw2, Cw4; DR15, DR4, DQ3, DQ6, DP4, DP4.
- A haplotype is the set of HLA antigens inherited from one parent. For example, the mother of the person whose HLA type is given above might have HLA-A3, -A69; B7, B45; Cw4, Cw9; DR15, DR17, DQ6, DQ2, DP2, DP4. Therefore, the A3, B7, Cw4, DR15, DQ6, and DP4 were all passed on from the mother to the child above. This group of antigens is a haplotype.
- Despite the enormous number of alleles at each expressed locus, the number of haplotypes observed in the population is much smaller than theoretical expectations. This is because certain alleles tend to occur together on the same haplotype rather than segregating randomly. This is called linkage disequilibrium.

Linkage Disequilibrium

- Linkage disequilibrium is a genetic phenomenon in which two alleles are found together with a higher frequency than normally expected.
- It is the non-random association between alleles at different loci. For example, if 16 percent of the population has a particular HLA-A antigen (A1) and 10 percent of the population has a particular HLA-B antigen (B8), the chance of finding A1 genetically linked to B8 on the same chromosome is given by the product of their gene frequencies (16 percent **x** 10 percent = 6 percent).
- In practice, this does not always occur. Certain combinations of A and B specificities occur more frequently than would be expected if their association were random. The combination of A1 and B8 is found at a frequency of 8.8 percent in human populations compared to an expected frequency of 1.6 percent. Such paired specificities are said to be in linkage disequilibrium.
- In Caucasians, the HLA-A1, B8, DR3 (DRB1*0301), DQ2 (DQB1*0201) haplotype is highly conserved in the population.
- At HLA class II, this phenomenon is so pronounced that the presence of specific HLA-DR alleles can be used to predict the HLA-DQ allele with a high degree of accuracy before testing. The HLA alleles are ordered on chromosome 6 as DP-DQ-DR-B-Cw-A.
- Those alleles that are physically closest to each other usually have the highest linkage disequilibrium. It is possible that certain haplotypes may be advantageous in some immunological sense, so that they have a positive selective advantage.

Rules that dictate the nomenclature of HLA:

- The prefix HLA precedes all antigens or alleles.
- A capital letter indicates a specific locus (A, B, C, or D). All genes in the region D are prefixed by the letter D followed by a second letter indicating the subregion of D (DR, DQ, DP, DM, or DO).
- Loci coding for the specific class II peptide chains are next identified (A1, A2, B1, and B2). Greek letters are used for protein designations, whereas Latin capital letters are used for gene/allele designations, i.e., DRp1 versus DRB1.
- Specific alleles are designated by an ''*'' followed by a two-digit number indicating the most closely associated serologic specificity, followed by a two-digit number that defines the unique allele. For example, the serologically defined HLA-A2 specificity actually comprises seventy-seven distinct variant alleles. These alleles are now referred to as HLA-A*02:01 through *02:99.
- Some alleles have a third two-digit number (HLA-B*35:01:01 and B*35:01:02) which indicates that the two variants differ by a silent nucleotide substitution, but not in amino acid sequence 6.

Quiz

[Now test your knowledge with these questions!](#)

Related Talk

/var/www/html/wp-content/uploads/2016/06/Reto-MHC-presentation.m4a

Reto Guler, University of Cape Town — TB Host Directed Therapy and MHC/Antigen Presentation