MHC

BSc (H)

Sem - IV

ARUN JANA

The major histocompatibility complex is a collection of genes arrayed within a long continuous stretch of DNA on chromosome 6 in humans and on chromosome 17 in mice.

The MHC is referred to as the HLA complex in humans and as the H-2 complex in mice.

Although the arrangement of genes is somewhat different, in both cases the MHC genes are organized into regions encoding three classes of molecules

Class I MHC genes encode glycoproteins expressed on the surface of nearly all nucleated cells; the major function of the class I gene products is presentation of peptide antigens to Tc cells.

Class II MHC genes encode glycoproteins expressed primarily on antigen-presenting cells (macrophages, dendritic cells, and B cells), where they present processed antigenic peptides to TH cells.

Class III MHC genes encode, in addition to other products, various secreted proteins that have immune functions, including components of the complement system and molecules involved in inflammation.

Class I MHC molecules encoded by the K and D regions in mice and by the A, B, and C loci in humans were the first discovered, and they are expressed in the widest range of cell types. These are referred to as *classical class I molecules.*

Additional genes or groups of genes within the H-2 or HLA complexes also encode class I molecules; these genes are designated *nonclassical class I genes.*

The two chains of the *class II MHC molecules* are encoded by the IA and IE regions in mice and by the DP, DQ, and DR regions in humans. The terminology is somewhat confusing, since the D region in mice encodes class I MHC molecules, whereas the D region (DR, DQ, DP) in humans refers to genes encoding class II MHC molecules!

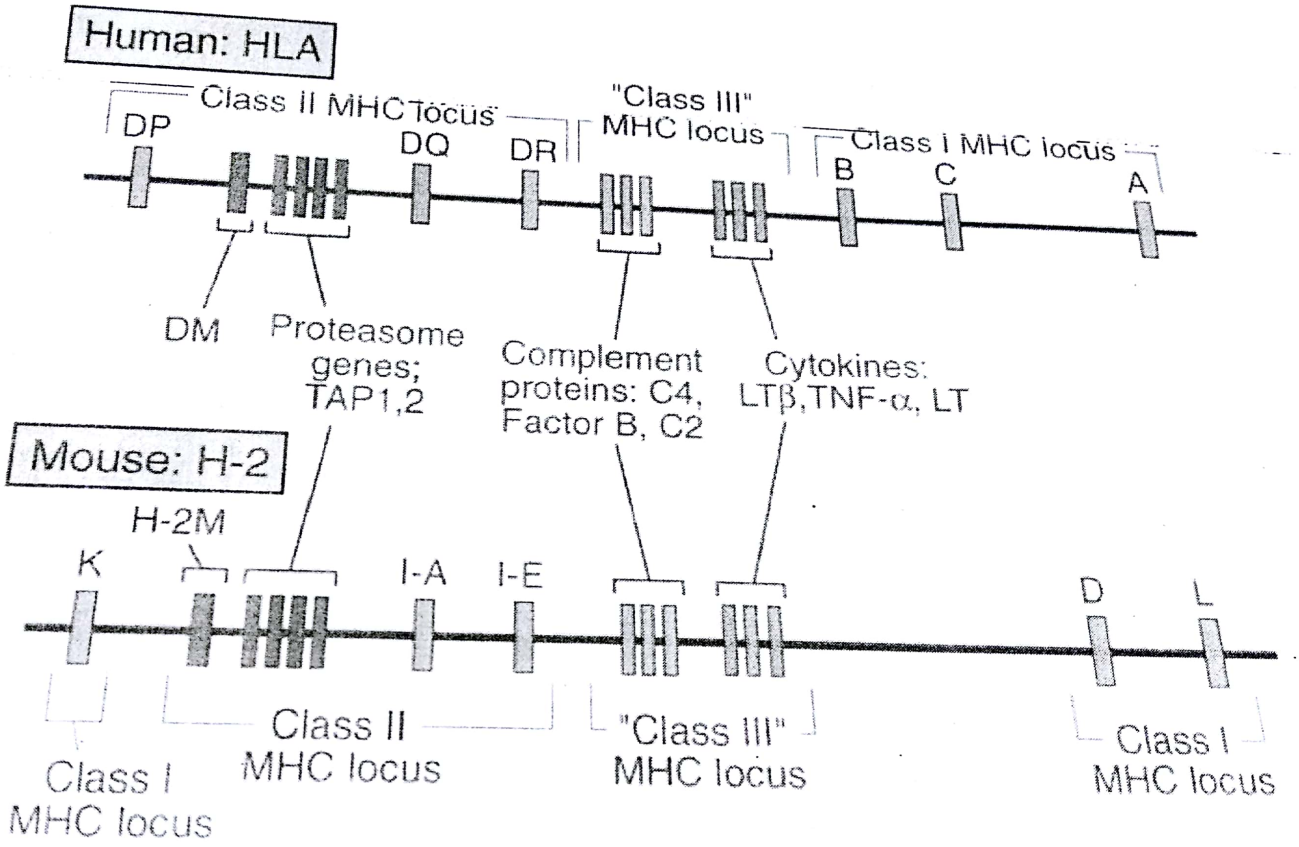
Class III MHC region, which is flanked by the class I and II regions, encodes molecules that are critical to immune function but have little in common with class I or II molecules. Class III products include the complement components C4, C2, BF, and inflammatory cytokines, including tumor necrosis factor (TNF) and heat-shock proteins.

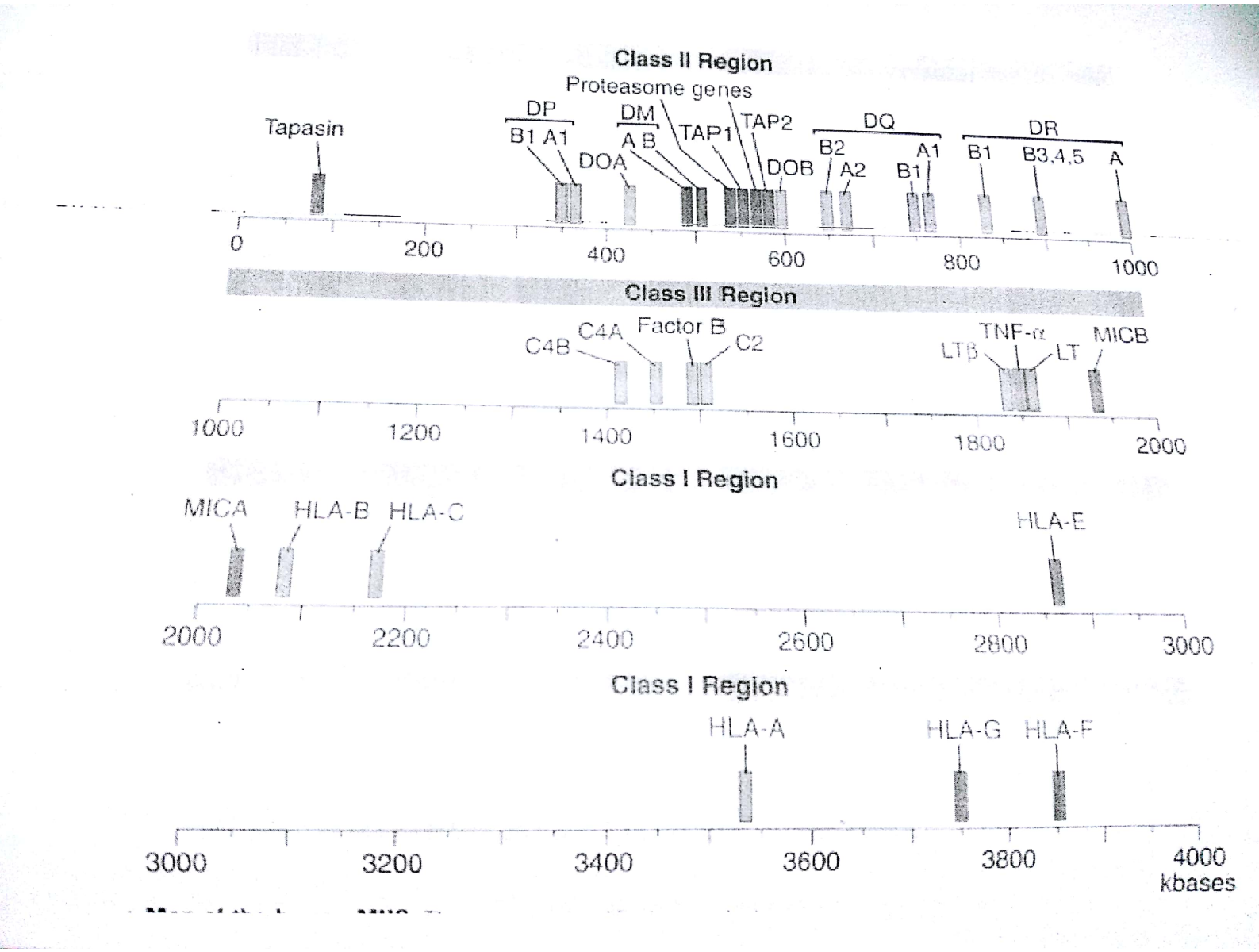
GENOMIC ORGANIZATION OF THE MHC

- In humans, the MHC is located on the short arm of chromosome 6, and β2-microglobulin is encoded by a gene on chromosome 15.
- The human MHC occupies a large segment of DNA,extending about 3500 kilobases (kb). (For comparison, a large human gene may extend up to 50 to 100kb, and the size of the entire genome of the bacterium *Escherichia coli* is approximately 4500kb.)
- In classical genetic terms, the MHC locus extends about 4 centimorgans, meaning that crossovers within the MHC occur with a frequency of about 4% at each meiosis.

Many of the proteins involved in the processing of protein antigens and the presentation of peptides to T cells are encoded by genes located within the MHC.

In other words, this genetic locus contains much of the information needed for the machinery of antigen presentation.





- *The class I genes, HLA-A, HLA-B, and HLA-C,* are in the most telomeric portion of the HLA locus, and the class II genes are the most centromeric in the HLA locus. Within the class II locus are genes that encode several proteins that play critical roles in antigen processing.

- One of these proteins, called the transporter associated with antigen processing (TAP), is a heterodimer that transports peptides from the cytosol into the endoplasmic reticulum, where the peptides can associate with newly synthesized class I molecules.

- The two subunits of the TAP dimer are encoded by two genes within the class II region. Other genes in this cluster encode subunits of a cytosolic protease complex, called the proteasome, that degrades cytosolic proteins into peptides that are subsequently presented by class I MHC molecules.

- Another pair of genes, called *HLA-DMA and HLA-DMB, encodes a nonpolymorphic heterodimeric* class II-like molecule, called HLA-DM (or H-2M in mice), that is involved in peptide binding to class II molecules.

- Between the class I and class II gene clusters are genes that code for several components of the complement system; for three structurally related cytokines, tumor necrosis factor (TNF), lymphotoxin-α (LT-α), and LT-β; and for some heat shock proteins.

- The genes within the MHC that encode these diverse proteins have been called class III MHC genes. Between HLA-C and HLA-A, and telomeric to HLA-A, are many genes that are called class I-like because they resemble class I genes but exhibit little or no polymorphism.

- Some of these encode proteins that are expressed in association with β2-microglobulin and are called class IB molecules, to distinguish them from the classical polymorphic class I molecules.

- Among the class IB molecules is HLA-G, which may playa role in antigen recognition by natural killer (NK) cells, and HLA-H, which appears to be involved in iron metabolism and has no known function in the immune system. Many of the class I-like sequences are pseudogenes.

- The functions of most of these class I-like genes and pseudo genes are not known. One role may be that during evolution, these DNA sequences serve as repositories of coding sequences that are used for generating polymorphic sequences in conventional class I and class II MHC molecules by the process of gene conversion.

- In this process, a portion of the sequence of one gene is replaced with a portion of another gene without a reciprocal recombination event.

- *Gene conversion* is a more efficient mechanism than point mutation for producing genetic variation without loss of function because several changes can be introduced at once, and amino acids necessary for maintaining protein structure can remain unchanged if identical amino acids at those positions are encoded by both of the genes involved in the conversion event.

- It is clear from population studies that the extraordinary polymorphism of MHC molecules has been generated by gene conversion and not by point mutations.

- The mouse MHC, located on chromosome 17, occupies about 2000 kb of DNA, and the genes are organized in an order slightly different from the human MHC gene.

- One of the mouse class I genes *(H-2K) is centromeric to* the class II region, but the other class I genes and the nonpolymorphic class IB genes are telomeric to the class II region.

- As in the human, β2-microglobulin is encoded not by the MHC but by a gene located on a separate chromosome (chromosome 2).

- There are β2-microglobulin-associated proteins other than class I MHC molecules that may serve important functions in the immune system. These include the neonatal Fc receptor, and the CD1 molecules, which are involved in presenting lipid and other nonpeptide antigens to unusual populations of T cells. These proteins are homologous to the class 1 MHC α chain but are encoded outside the MHC, on different chromosomes.

Properties of MHC Molecules

ARUN JANA

All MHC molecules share certain structural characteristics that are critical for their role in peptide display and antigen recognition by T lymphocytes.

Each MHC molecule consists of an extracellular peptide-binding cleft, or groove, followed by immunoglobulin (Ig}-like domains and transmembrane and cytoplasmic domains.

Class I molecules are composed of one polypeptide chain encoded in the MHC and a second, non-MHC encoded chain, whereas class II molecules are made up of two MHC-encoded polypeptide chains. Despite this difference, the overall three-dimensional structures of class I and class II molecules are similar.

The polymorphic amino acid residues of MHC molecules are located in and adjacent to the peptide binding cleft.

This cleft is formed by the folding of the amino termini of the MI-IC-encoded proteins and is composed of paired a-helices resting on a floor made up of an eight-stranded ~-pleated sheet. The polymorphic residues, which are the amino acids that vary among different MHC alleles, are located in and around this cleft. This portion of the MHC molecule binds peptides for display to T cells, and the antigen receptors ofT cells interact with the displayed peptide and with the a-helices ofthe MHC molecules.

The nonpolymorphic Ig-like domains of MHC molecules contain binding sites for the T cell molecules CD4 and CD8.

CD4 and CD8 are expressed on distinct subpopulations of mature T lymphocytes and participate, together with antigen receptors, in the recognition of antigen; that is, CD4 and CD8 are T cell "coreceptors". CD4 binds selectively to class II MHC molecules. and CD8 binds to class I molecules.

This is why $CD4^+$ T cells recognize only peptides displayed by class II molecules, and $CD8^+$ T Cells recognize peptides presented by class I molecules. Most $CD4^+$ T cells function as helper cells, and most $CD8^+$ cells are CTLs.

Class I MHC Molecules

- Class I molecules consist of two noncovalently linked polypeptide chains: an MHC-encoded 44-47 kD **α chain** (or heavy chain) and a non-MHC-encoded 12 kD subunit called **β_2-microglobulin**.
- Each α-chain is oriented so that about three quarters of the complete polypeptide extends into the extracellular milieu.
- A short hydrophobic segment spans the cell membrane, and the carboxy-terminal residues are located in the cytoplasm. The amino-terminal (N-terminal) αl and α2 segments of the α chain, each approximately 90 residues long, interact to form a platform of an eight-stranded, anti parallel β-pleated sheet supporting two parallel strands of α helix. This forms the peptide-binding cleft of class I molecules.
- Its size is large enough (-25A x lOA x 11A) to bind peptides of 8 to 11 amino acids in a flexible, extended conformation.
- The ends of the class I peptide-binding cleft are closed so that larger peptides cannot be accommodated.
- Therefore, native globular proteins have to be "processed" to generate fragments that are small enough to bind to MHC molecules and to be recognized by T cells.

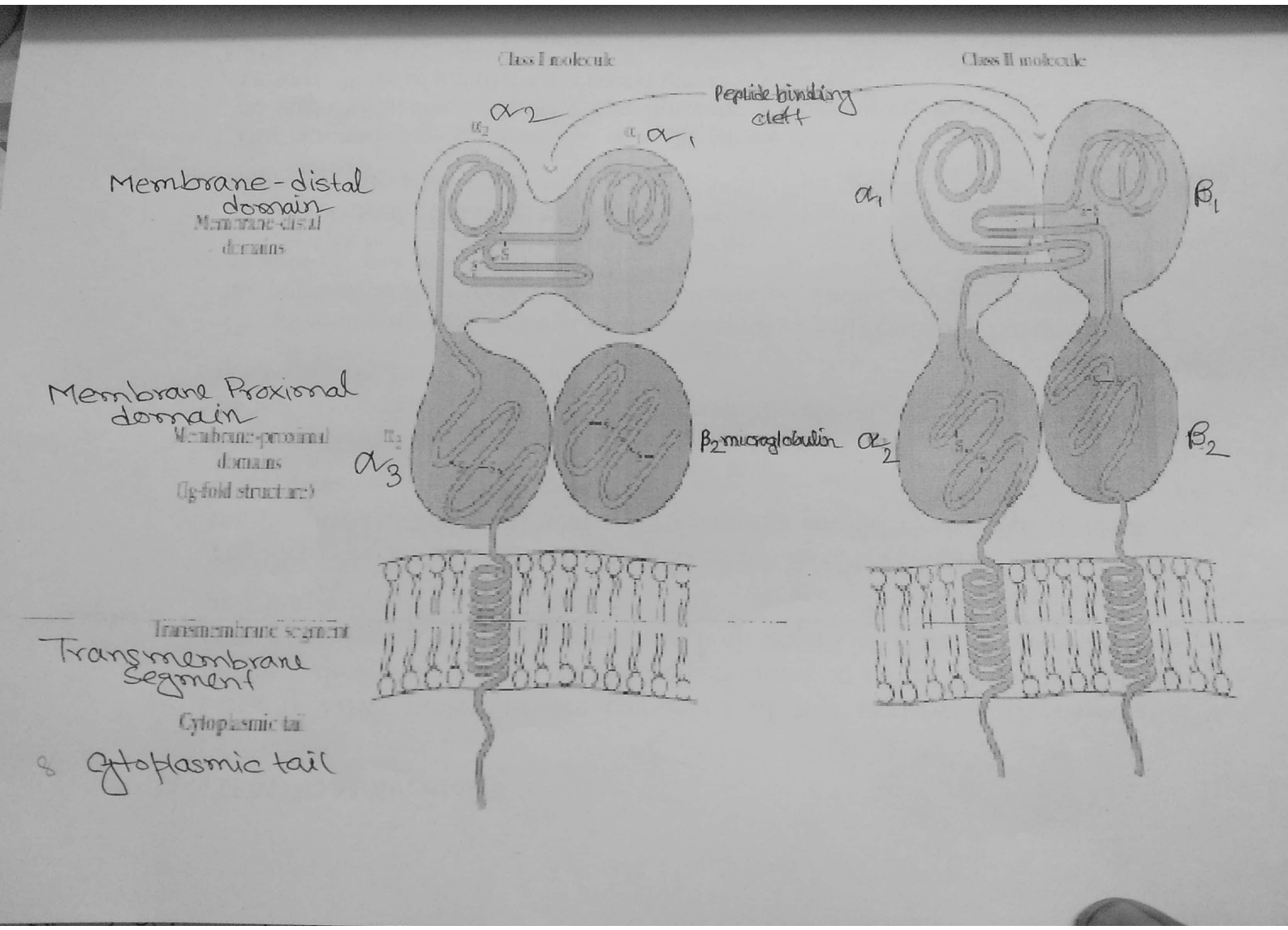
- The polymorphic residues of class I molecules are confined to the α1 and α2 domains, where they contribute to variations among different class-I alleles in peptide binding and T cell recognition.
- The α3 segment of the a chain folds into an Ig domain whose amino acid sequence is conserved among all class I molecules. This segment contains the binding site for CD8.
- At the carboxy-terminal end of the α3 segment is a stretch of approximately 25 hydrophobic amino acids that traverses the lipid bilayer of the plasma membrane.
- Immediately following this are approximately 30 residues located in the cytoplasm, included in which is a cluster of basic amino acids that interact with phospholipid head groups of the inner leaflet of the lipid bilayer and anchor the MHC molecule in the plasma membrane.
- The light chain of class I molecules, which is encoded by a gene outside the MHC, is called β_2-microglobulin for its electrophoretic mobility (β_2), size (micro), and solubility (globulin).
- β_2-microglobulin interacts noncovalently with the α3 domain of the α chain. Like the α3 segment, β_2-microglobulin is structurally homologous to an Ig domain and is invariant among all class I molecules.

Class II MHC Molecules

- Class II MHC molecules are composed of two noncovalently associated polypeptide chains, a 32 to 34 kD α chain and a 29 to 32 kD β chain. Unlike class I molecules, the genes encoding both chains of class II molecules are polymorphic.
- The amino-terminal αl and βl segments of the class II chains interact to form the peptide-binding cleft, which is structurally similar to the cleft of class I molecules.
- Four strands of the floor of the cleft and one of the α-helical walls are formed by the αl segment, and the other four strands of the floor and the second wall are formed by the β1 segment.
- The polymorphic residues are located in the αl and βl segments, in and around the peptide-binding cleft, as in class I molecules . In human class II molecules, most of the polymorphism is in the β chain.
- In class II molecules, the ends of the peptide-binding cleft are open, so that peptides of 30 residues or more can fit.
- The α2 and β2 segments of class II molecules, like class I α3 and β2-microglobulin, are folded into Ig domains and are nonpolymorphic among various alleles of a particular class II gene.

- The β2 segment of class II molecules contains the binding site for CD4, similar to the binding site for CD8 in the α3 segment of the class I heavy chain.
- In general, α chains of one class II MHC locus most often pair with β chains of the same locus and less commonly with β chains of other loci.
- The carboxy-terminal ends of the α2 and β2 segments continue into short connecting regions followed by approximately 25-amino acid stretches of hydrophobic transmembrane residues.
- In both chains, the transmembrane regions end with clusters of basic amino acid residues, followed by short, hydrophilic cytoplasmic tails.

The fully assembled class II molecule is a heterotrimer consisting of an α chain, a β chain, and a bound antigenic peptide, and stable expression of class II molecules on cell surfaces requires the presence of all three components of the heterotrimer.



Feature	Class I MHC	Class II MHC
Polypeptide chains	α (44–47 kD) β_2-Microglobulin (12 kD)	α (32–34 kD) β (29–32 kD)
Locations of polymorphic residues	α1 and α2 domains	α1 and β1 domains
Binding site for T cell coreceptor	α3 region binds CD8	β2 region binds CD4
Size of peptide-binding cleft	Accommodates peptides of 8-11 residues	Accommodates peptides of 10-30 residues or more
Nomenclature		
Human	HLA-A, HLA-B, HLA-C	HLA-DR, HLA-DQ, HLA-DP
Mouse	H-2K, H-2D, H-2L	I-A, I-E

Abbreviations: HLA, human leukocyte antigen; MHC, major histocompatibility complex

Characteristics of Peptide-MHC Interactions

MHC molecules show a broad specificity for peptide binding, in contrast to the fine specificity of antigen recognition of the antigen receptors of T lymphocytes.

There are several important features of the interactions of MHC molecules and antigenic peptides.

Each class I or class II MHC molecule has a single peptide-binding cleft that binds one peptide at a time, but each MHC molecule can bind many different pep tides.

The peptides that bind to MHC molecules share structural features that promote this interaction.

One of these features is the size of the peptide-class I molecules can accommodate peptides that are 8 to 11 residues long, and class II molecules bind peptides that may be 10 to 30 residues long or longer, the optimal length being 12 to 16 residues. In addition, peptides that bind to a particular allelic form of an MHC molecule contain amino acid residues that allow complementary interactions between the peptide and that allelic MHC molecule. The residues of a peptide that bind to MHC molecules are distinct from those that are recognized by T cells.

The association of antigenic pep tides and MHC molecules is a saturable interaction with a very slow off-rate.

In a cell, several chaperones and enzymes facilitate the binding of pep tides to MHC molecules . Once formed, most peptide-MHC complexes are stable, and kinetic dissociation constants are indicative of long half-lives that range from hours to many days.

This extraordinarily slow off-rate of peptide dissociation from MHC molecules allowspeptide-MHC complexes to persist long enough on the surfaces of antigen-presenting cells that the antigen can be found by T cells. This feature of antigen display enables the few T cells that are specific for the antigen to locate the antigen as the cells circulate through tissues. and thus to generate effective immune responses against the antigen.

The MHC molecules of an individual do not discriminate between foreign pep tides (e.g., those derived from microbial proteins) and pep tides derived from the proteins of that individual (self antigens).

Structural Basis of Peptide Binding to MHC Molecules

The binding of peptides to MHC molecules is a noncovalent interaction mediated by residues both in the peptides and in the clefts of the MHC molecules.

- Protein antigens are proteolytically cleaved in antigen presenting cells to generate the peptides that will be bound and displayed by MHC molecules . These peptides bind to the clefts of MHC molecules in an extended conformation. Once bound. the peptides and their associated water molecules fill the clefts, making extensive contacts with the amino acid residues that form the β-strands of the floor and the α-helices of the walls of the cleft.

- In most MHC molecules, the β-strands in the floor of the cleft contain "pockets."

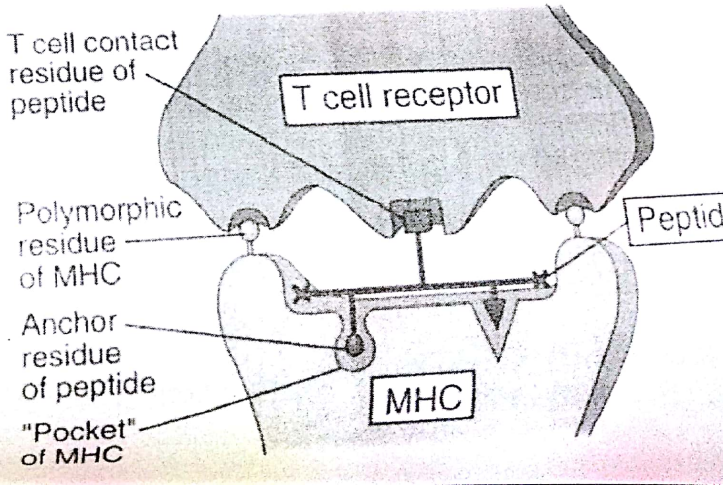
- The amino acid residues of a peptide may contain side chains that fit into these pockets and bind to complementary amino acids in the MHC molecule, often through hydrophobic interactions.

Such residues of the peptide are called anchor residues because they contribute most of the favorable interactions of the binding (i.e., they anchor the peptide in the cleft of the MHC molecule).

- The anchor residues of peptides may be located in the middle or at the ends of the peptide. Each MHC-binding peptide usually contains only one or two anchor residues, and this presumably allows greater variability in the other residues of the Peptide, that are recognized by specific T cells.
- Not all peptides use anchor residues to bind to MHC molecules, especially to class II molecules.
- Specific interactions of peptides with the α-helical sides of the MHC cleft also contribute to peptide binding by forming hydrogen bonds or charge interactions (salt bridges).
- Class I-binding peptides usually contain hydrophobic or basic amino acids at their carboxyl termini that also contribute to the interaction.
- Because many of the residues in and around the peptide-binding cleft of MHC molecules are polymorphic (i.e. they differ among various MHC alleles), different alleles favor the binding of different peptides.
- This is the structural basis of the function of MHC genes as "immune response genes"; only animals that express MHC alleles that can bind a particular peptide and display it to T cells can respond to that peptide.

The antigen receptors of T cells recognize both the antigenic peptide and the MHC molecules, with the peptide being responsible for the fine specificity of antigen recognition and the MHC residues accounting for the MHC restriction of the T cells.

- A portion of the bound peptide is exposed from the open top of the cleft of the MHC molecule. and the amino acid side chains of this portion of the peptide are recognized by the antigen receptors of specific T cells.
- The same T cell receptor also interacts with polymorphic residues of the a-helices of the MHC molecule itself.
- Predictably, variations in either the peptide antigen or the peptide-binding cleft of the MHC molecule will alter presentation of that peptide or its recognition by T cells.
- In fact, one can enhance the immunogenicity of a peptide by incorporating into it a residue that strengthens its binding to commonly inherited MHC molecules in a population.



CLASS I MHC–PEPTIDE INTERACTION

Class I MHC molecules bind peptides and present them to $CD8^+$ T cells. In general, these peptides are derived from endogenous intracellular proteins that are digested in the cytosol.

The peptides are then transported from the cytosol into the cisternae of the endoplasmic reticulum, where they interact with class I MHC molecules. This process, known as the cytosolic or endogenous processing pathway.

Each type of class I MHC molecule (K, D, and L in mice or A, B, and C in humans) binds a unique set of peptides. In addition, each allelic variant of a class I MHC molecule (e.g., H-2Kk and H-2Kd) also binds a distinct set of peptides.

Because a single nucleated cell expresses about 10^5 copies of each class I molecule, many different peptides will be expressed simultaneously on the surface of a nucleated cell by class I MHC molecules.

The bound peptides isolated from different class I molecules have been found to have two distinguishing features: they are eight to ten amino acids in length, most commonly nine, and they contain specific amino acid residues that appear to be essential for binding to a particular MHC molecule.

Nonameric peptides bind to class I molecules with a 100- to 1000-fold higher affinity than do peptides that are either longer or shorter, suggesting that this peptide length is most compatible with the closed-ended peptide-binding cleft in class I molecules.

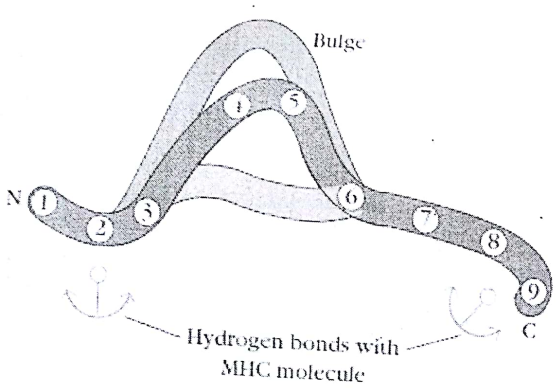
The ability of an individual class I MHC molecule to bind to a diverse spectrum of peptides is due to the presence of the same or similar amino acid residues at several defined positions along the peptides . Because these amino acid residues anchor the peptide into the groove of the MHC molecule, they are called anchor residues.

The side chains of the anchor residues in the peptide are complementary with surface features of the binding cleft of the class I MHC molecule. The amino acid residues lining the binding sites vary among different class I allelic variants and determine the identity of the anchor residues that can interact with the molecule.

All peptides that bind to class I molecules contain a carboxyl-terminal anchor. These anchors are generally hydrophobic residues (e.g., leucine, isoleucine), although a few charged amino acids have been reported.
Besides the anchor residue found at the carboxyl terminus, another anchor is often found at the second or second and third positions at the amino-terminal end of the peptide.
In general, any peptide of correct length that contains the same or similar anchor residues will bind to the same class I MHC molecule.

CLASS II MHC–PEPTIDE INTERACTION

- Class II MHC molecules bind peptides and present these peptides to CD4 T cells.
- In general, these peptides are derived from exogenous proteins (either self or nonself), which are degraded within the endocytic processing pathway.
- Most of the peptides associated with class II MHC molecules are derived from membrane bound proteins or proteins associated with the vesicles of the endocytic processing pathway.
- The membrane-bound proteins presumably are internalized by phagocytosis or by receptor-mediated endocytosis and enter the endocytic processing pathway.
- Peptides generally contain 13–18 amino acid residues, somewhat longer than the nonameric peptides that most commonly bind to class I molecules.
- The peptide-binding cleft in class II molecules is open at both ends , allowing longer peptides to extend beyond the ends.
- Peptides bound to class II MHC molecules maintain a roughly constant elevation on the floor of the binding cleft, another feature that distinguishes peptide binding to class I and class II molecules.
- Central core of 13 amino acids determines the ability of a peptide to bind class II. unlike class I–binding peptides, they lack conserved anchor residues.



- Hydrogen bonds between the backbone of the peptide and the class II molecule are distributed throughout the binding site rather than being clustered predominantly at the ends of the site as for class I–bound peptides.
- Peptides that bind to class II MHC molecules contain an internal sequence comprising 7–10 amino acids that provide the major contact points.
- Generally, this sequence has an aromatic or hydrophobic residue at the amino terminus and three additional hydrophobic residues in the middle portion and carboxyl-terminal end of the peptide.
- In addition, over 30% of the peptides eluted from class II molecules contain a proline residue at position 2 and another cluster of prolines at the carboxyl-terminal end.

TABLE 7-2 Peptide binding by class I and class II MHC molecules

	Class I molecules	Class II molecules
Peptide-binding domain	$\alpha 1/\alpha 2$	$\alpha 1/\beta 1$
Nature of peptide-binding cleft	Closed at both ends	Open at both ends
General size of bound peptides	8–10 amino acids	13–18 amino acids
Peptide motifs involved in binding to MHC molecule	Anchor residues at both ends of peptide; generally hydrophobic carboxyl-terminal anchor	Anchor residues distributed along the length of the peptide
Nature of bound peptide	Extended structure in which both ends interact with MHC cleft but middle arches up away from MHC molecule	Extended structure that is held at a constant elevation above the floor of MHC cleft