Chapter 3/4/5:

Link: NCBI/literature/bookshelf/

Immunobio 5

Chapter 3. **Antigen** **Recognition** by B-cell and **T-cell** **Receptors**

* The structure of a typical antibody molecule
  + 3-1. IgG antibodies consist of four polypeptide chains
  + 3-2. Immunoglobulin heavy and light chains are composed of constant and variable regions
  + 3-3. The antibody molecule can readily be cleaved into functionally distinct fragments
  + 3-4. The immunoglobulin molecule is flexible, especially at the hinge region
  + 3-5. The domains of an immunoglobulin molecule have similar structures
  + Summary
* The interaction of the antibody molecule with specific **antigen**
  + 3-6. Localized regions of hypervariable sequence form the antigenbinding site
  + 3-7. Antibodies bind **antigens** via contacts with amino acids in CDRs, but the details of binding depend upon the size and shape of the **antigen**
  + 3-8. Antibodies bind to conformational shapes on the surfaces of **antigens**
  + 3-9. **Antigen**-antibody interactions involve a variety of forces
  + Summary
* **Antigen** **recognition** by T cells
  + 3-10. The **antigen** receptor on T cells is very similar to a Fab fragment of immunoglobulin
  + 3-11. A **T-cell** receptor recognizes **antigen** in the form of a complex of a foreign peptide bound to an MHC molecule
  + 3-12. T cells with different functions are distinguished by CD4 and CD8 cell-surface proteins and recognize peptides bound to different classes of MHC molecule
  + 3-13. The two classes of MHC molecule are expressed differentially on cells
  + 3-14. The two classes of MHC molecule have distinct subunit structures but similar three-dimensional structures
  + 3-15. Peptides are stably bound to MHC molecules, and also serve to stabilize the MHC molecule on the cell surface
  + 3-16. MHC class I molecules bind short peptides of 8–10 amino acids by both ends
  + 3-17. The length of the peptides bound by MHC class II molecules is not constrained
  + 3-18. The crystal structures of several MHC:peptide:**T-cell** receptor complexes all show the same **T-cell** receptor orientation over the MHC:peptide complex
  + 3-19. A distinct subset of T cells bears an alternative receptor made up of γ and δ chains

**(A) The structure of a typical antibody molecule**

**Antibodies** are the secreted form of the B-cell receptor. An [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) is identical to the B-cell receptor of the cell that secretes it except for a small portion of the C-terminus of the heavy-chain [constant region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2727/). In the case of the B-cell receptor the C-terminus is a hydrophobic membrane-anchoring sequence, and in the case of **antibody** it is a hydrophilic sequence that allows secretion. Since they are soluble, and secreted in large quantities, **antibodies** are easily obtainable and easily studied. For this reason, most of what we know about the B-cell receptor comes from the study of **antibodies**.

**Antibody** molecules are roughly Y-shaped molecules consisting of three equal-sized portions, loosely connected by a flexible tether. Three schematic representations of [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **structure**, which has been determined by X-ray crystallography, are shown in [Fig. 3.1](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A324/?report=objectonly" \t "object). The aim of this part of the chapter is to explain how this **structure** is formed and how it allows **antibody** molecules to carry out their dual tasks—binding on the one hand to a wide variety of antigens, and on the other hand to a limited number of effector molecules and cells. As we will see, each of these tasks is carried out by separable parts of the **molecule**. The two arms of the Y end in regions that vary between different **antibody** molecules, the V regions. These are involved in [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) binding, whereas the stem of the Y, or the [C region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2649/), is far less variable and is the part that interacts with effector cells and molecules.

**[Figure 3.1](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A324/?report=objectonly" \t "object)**

**Structure** of an **antibody** **molecule**. Panel a illustrates a ribbon diagram based on the X-ray crystallographic **structure** of an IgG **antibody**, showing the course of the backbones of the polypeptide chains. Three globular regions form a Y. The two antigen-binding [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A324/?report=objectonly" \t "object)

All **antibodies** are constructed in the same way from paired heavy and light polypeptide chains, and the generic term immunoglobulin is used for all such proteins. Within this general category, however, five different classes of [**immunoglobulins**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/)—[IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/), [IgD](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2906/), [IgG](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2908/), [IgA](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2905/), and [IgE](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2907/)—can be distinguished by their C regions, which will be described more fully in Chapter 4. More subtle differences confined to the V region account for the [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/) of [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) binding. We will use the IgG [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule** as an example to describe the general structural features of **immunoglobulins**.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27144/)

**3-1. IgG antibodies consist of four polypeptide chains**

[IgG](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2908/) **antibodies** are large molecules, having a molecular weight of approximately 150 kDa, composed of two different kinds of polypeptide chain. One, of approximately 50 kDa, is termed the **heavy** or **H chain**, and the other, of 25 kDa, is termed the **light** or [L chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2985/) ([Fig. 3.2](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A326/?report=objectonly" \t "object)). Each IgG **molecule** consists of two heavy chains and two light chains. The two heavy chains are linked to each other by disulfide bonds and each [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/) is linked to a [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/) by a disulfide bond. In any given immunoglobulin **molecule**, the two heavy chains and the two light chains are identical, giving an [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule** two identical [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-binding sites (see [Fig. 3.1](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A324/?report=objectonly" \t "object)), and thus the ability to bind simultaneously to two identical structures.

**[Figure 3.2](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A326/?report=objectonly" \t "object)**

Immunoglobulin molecules are composed of two types of protein chain: heavy chains and light chains. Each immunoglobulin **molecule** is made up of two heavy chains (green) and two light chains (yellow) joined by disulfide bonds so that each heavy chain is [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A326/?report=objectonly" \t "object)

Two types of [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/), termed **lambda** (λ) and **kappa** (κ), are found in **antibodies**. A given immunoglobulin either has κ chains or λ chains, never one of each. No functional difference has been found between **antibodies** having λ or κ light chains, and either type of light chain may be found in **antibodies** of any of the five major classes. The ratio of the two types of light chain varies from species to species. In mice, the average κ to λ ratio is 20:1, whereas in humans it is 2:1 and in cattle it is 1:20. The reason for this variation is unknown. Distortions of this ratio can sometimes be used to detect the abnormal proliferation of a [clone](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2704/) of B cells. These would all express the identical light chain, and thus an excess of λ light chains in a person might indicate the presence of a B-cell tumor producing λ chains.

By contrast, the class, and thus the effector function, of an [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/), is defined by the **structure** of its [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/). There are five main **heavy-chain classes** or [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/), some of which have several subtypes, and these determine the functional activity of an **antibody** **molecule**. The five major classes of immunoglobulin are immunoglobulin M ([IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/)), immunoglobulin D ([IgD](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2906/)), immunoglobulin G ([IgG](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2908/)), immunoglobulin A ([IgA](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2905/)), and immunoglobulin E ([IgE](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2907/)). Their heavy chains are denoted by the corresponding lower-case Greek letter (μ, δ, γ, α, and ε, respectively). IgG is by far the most abundant immunoglobulin and has several subclasses (IgG1, 2, 3, and 4 in humans). Their distinctive functional properties are conferred by the carboxy-terminal part of the heavy chain, where it is not associated with the [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/). We will describe the **structure** and functions of the different heavy-chain isotypes in Chapter 4. The general structural features of all the isotypes are similar and we will consider IgG, the most abundant isotype in plasma, as a **typical** **antibody** **molecule**.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27144/)

**3-2. Immunoglobulin heavy and light chains are composed of constant and variable regions**

The amino acid sequences of many immunoglobulin heavy and light chains have been determined and reveal two important features of [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) molecules. First, each chain consists of a series of similar, although not identical, sequences, each about 110 amino acids long. Each of these repeats corresponds to a discrete, compactly folded region of protein **structure** known as a protein domain. The [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/) is made up of two such [immunoglobulin domains](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2935/), whereas the [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/) of the [IgG](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2908/) **antibody** contains four (see [Fig. 3.1a](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A324/?report=objectonly" \t "object)). This suggests that the immunoglobulin chains have evolved by repeated duplication of an ancestral gene corresponding to a single domain.

The second important feature revealed by comparisons of amino acid sequences is that the amino-terminal sequences of both the heavy and light chains vary greatly between different **antibodies**. The [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/) in sequence is limited to approximately the first 110 amino acids, corresponding to the first domain, whereas the remaining domains are constant between immunoglobulin chains of the same isotype. The amino-terminal variable or V domains of the heavy and light chains (VH and VL, respectively) together make up the V region of the [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) and confer on it the ability to bind specific [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/), while the constant domains ([C domains](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2648/)) of the heavy and light chains (CH and CL, respectively) make up the [C region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2649/) (see [Fig. 3.1b, c](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A324/?report=objectonly" \t "object)). The multiple heavy-chain C domains are numbered from the amino-terminal end to the carboxy terminus, for example CH1, CH2, and so on.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27144/)

**3-3. The antibody molecule can readily be cleaved into functionally distinct fragments**

The protein domains described above associate to form larger globular domains. Thus, when fully folded and assembled, an [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule** comprises three equal-sized globular portions joined by a flexible stretch of polypeptide chain known as the [hinge region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2874/) (see [Fig. 3.1b](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A324/?report=objectonly" \t "object)). Each arm of this Y-shaped **structure** is formed by the association of a [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/) with the amino-terminal half of a [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/), whereas the trunk of the Y is formed by the pairing of the carboxy-terminal halves of the two heavy chains. The association of the heavy and light chains is such that the VH and VL domains are paired, as are the CH1 and CL domains. The CH3 domains pair with each other but the CH2 domains do not interact; carbohydrate side chains attached to the CH2 domains lie between the two heavy chains. The two [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-binding sites are formed by the paired VH and VL domains at the ends of the two arms of the Y (see [Fig. 3.1b](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A324/?report=objectonly" \t "object)).

Proteolytic enzymes (proteases) that cleave polypeptide sequences have been used to dissect the **structure** of [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) molecules and to determine which parts of the **molecule** are responsible for its various functions. Limited digestion with the protease papain cleaves **antibody** molecules into three fragments ([Fig. 3.3](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A329/?report=objectonly" \t "object)). Two fragments are identical and contain the [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-binding activity. These are termed the [Fab fragments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2810/), for **F**ragment **a**ntigen **b**inding. The Fab fragments correspond to the two identical arms of the **antibody** **molecule**, which contain the complete light chains paired with the VH and CH1 domains of the heavy chains. The other fragment contains no antigen-binding activity but was originally observed to crystallize readily, and for this reason was named the [Fc fragment](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2815/), for **F**ragment **c**rystallizable. This fragment corresponds to the paired CH2 and CH3 domains and is the part of the **antibody** **molecule** that interacts with effector molecules and cells. The functional differences between heavy-chain [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/) lie mainly in the Fc fragment.

**[Figure 3.3](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A329/?report=objectonly" \t "object)**

The Y-shaped immunoglobulin **molecule** can be dissected by partial digestion with proteases. Papain cleaves the immunoglobulin **molecule** into three pieces, two Fab fragments and one Fc fragment (upper panels). The Fab fragment contains the V regions and [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A329/?report=objectonly" \t "object)

The protein fragments obtained after proteolysis are determined by where the protease cuts the [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule** in relation to the disulfide bonds that link the two heavy chains. These lie in the [hinge region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2874/) between the CH1 and CH2 domains and, as illustrated in [Fig. 3.3](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A329/?report=objectonly" \t "object), papain cleaves the **antibody** **molecule** on the amino-terminal side of the disulfide bonds. This releases the two arms of the **antibody** as separate [Fab fragments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2810/), whereas in the [Fc fragment](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2815/) the carboxy-terminal halves of the heavy chains remain linked.

Another protease, pepsin, cuts in the same general region of the [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule** as papain but on the carboxy-terminal side of the disulfide bonds (see [Fig. 3.3](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A329/?report=objectonly" \t "object)). This produces a fragment, the **F(ab′)**2 **fragment**, in which the two [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-binding arms of the **antibody** **molecule** remain linked. In this case the remaining part of the [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/) is cut into several small fragments. The F(ab′)2 fragment has exactly the same antigen-binding characteristics as the original **antibody** but is unable to interact with any effector **molecule**. It is thus of potential value in therapeutic applications of **antibodies** as well as in research into the functional role of the Fc portion.

Genetic engineering techniques also now permit the construction of many different [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/)-related molecules. One important type is a truncated Fab comprising only the [V domain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3342/) of a [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/) linked by a stretch of synthetic peptide to a V domain of a [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/). This is called [single-chain Fv](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3247/), named from **F**ragment **v**ariable. Fv molecules may become valuable therapeutic agents because of their small size, which allows them to penetrate tissues readily. They can be coupled to protein toxins to yield immunotoxins with potential application, for example, in tumor therapy in the case of a Fv specific for a tumor [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) (see Chapter 14).

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27144/)

**3-4. The immunoglobulin molecule is flexible, especially at the hinge region**

The [hinge region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2874/) that links the Fc and Fab portions of the [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule** is in reality a flexible tether, allowing independent movement of the two Fab arms, rather than a rigid hinge. This has been demonstrated by electron microscopy of **antibodies** bound to haptens. These are small molecules of various sorts, typically about the size of a tyrosine side chain. They can be recognized by **antibody** but are only able to stimulate production of antihapten **antibodies** when linked to a larger protein **carrier**(see [Appendix I, Section A-1](https://www.ncbi.nlm.nih.gov/books/n/imm/A2395/#A2399)). An [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) made of two identical hapten molecules joined by a short flexible region can link two or more anti-hapten **antibodies**, forming dimers, trimers, tetramers, and so on, which can be seen by electron microscopy ([Fig. 3.4](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A331/?report=objectonly" \t "object)). The shapes formed by these complexes demonstrate that **antibody** molecules are flexible at the hinge region. Some flexibility is also found at the junction between the V and [C domains](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2648/), allowing bending and rotation of the [V domain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3342/) relative to the C domain. For example, in the **antibody** **molecule** shown in [Fig. 3.1a](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A324/?report=objectonly" \t "object), not only are the two hinge regions clearly bent differently, but the angle between the V and C domains in each of the two Fab arms is also different. This range of motion has led to the junction between the V and C domains being referred to as a ‘molecular balland-socket joint.’ Flexibility at both the hinge and V-C junction enables the binding of both arms of an **antibody** **molecule** to sites that are various distances apart, for example, sites on bacterial cell-wall polysaccharides. Flexibility at the hinge also enables the **antibodies** to interact with the **antibody**-binding proteins that mediate immune effector mechanisms.

**[Figure 3.4](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A331/?report=objectonly" \t "object)**

**Antibody** arms are joined by a flexible hinge. An antigen consisting of two hapten molecules (red balls in diagrams) that can cross-link two antigen-binding sites is used to create antigen:**antibody** complexes, which can be seen in the electron micrograph. [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A331/?report=objectonly" \t "object)

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27144/)

**3-5. The domains of an immunoglobulin molecule have similar structures**

As we saw in [Section 3-2](https://www.ncbi.nlm.nih.gov/books/NBK27144/#A327), immunoglobulin heavy and light chains are composed of a series of discrete protein domains. These protein domains all have a similar folded **structure**. Within this basic three-dimensional **structure**, there are distinct differences between V and [C domains](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2648/). The structural similarities and differences can be seen in the diagram of a [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/) in [Fig. 3.5](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A333/?report=objectonly" \t "object). Each domain is constructed from two β sheets, which are elements of protein **structure** made up of strands of the polypeptide chain (β strands) packed together; the sheets are linked by a disulfide bridge and together form a roughly barrel-shaped **structure**, known as a [β barrel](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2617/). The distinctive folded **structure** of the immunoglobulin protein domain is known as the **immunoglobulin fold**.

**[Figure 3.5](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A333/?report=objectonly" \t "object)**

The **structure** of immuno-globulin constant and variable domains. The upper panels show schematically the folding pattern of the constant (C) and variable (V) domains of an immunoglobulin light chain. Each domain is a barrel-shaped **structure** in which strands [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A333/?report=objectonly" \t "object)

Both the essential similarity of V and [C domains](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2648/) and the critical difference between them are most clearly seen in the bottom panels of [Fig. 3.5](https://www.ncbi.nlm.nih.gov/books/NBK27144/figure/A333/?report=objectonly" \t "object), where the cylindrical domains are opened out to reveal how the polypeptide chain folds to create each of the β sheets and how it forms flexible loops as it changes direction. The main difference between the V and C domains is that the [V domain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3342/) is larger, with an extra loop. We will see in [Section 3-6](https://www.ncbi.nlm.nih.gov/books/n/imm/A335/#A336) that the flexible loops of the V domains form the [antigen-binding site](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2582/) of the immunoglobulin **molecule**.

Many of the amino acids that are common to C and V domains of immuno-globulin chains lie in the core of the immunoglobulin fold and are critical to its stability. For that reason, other proteins having sequences similar to those of [**immunoglobulins**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) are believed to form domains of similar **structure**, and in many cases this has been demonstrated by crystallography. These **immunoglobulin-like domains** are present in many other proteins of the [immune system](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2921/), and in proteins involved in cell-cell recognition in the nervous system and other tissues. Together with the **immunoglobulins** and the T-cell receptors, they make up the extensive [immunoglobulin superfamily](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2940/).

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27144/)

**Summary**

The [IgG](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2908/) [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule** is made up of four polypeptide chains, comprising two identical light chains and two identical heavy chains, and can be thought of as forming a flexible Y-shaped **structure**. Each of the four chains has a variable (V) region at its amino terminus, which contributes to the [antigen-binding site](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2582/), and a constant (C) region, which determines the isotype. The isotype of the [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/) determines the functional properties of the **antibody**. The light chains are bound to the heavy chains by many noncovalent interactions and by disulfide bonds, and the V regions of the heavy and light chains pair in each arm of the Y to generate two identical antigen-binding sites, which lie at the tips of the arms of the Y. The possession of two antigen-binding sites allows **antibody** molecules to cross-link antigens and to bind them much more stably. The trunk of the Y, or [Fc fragment](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2815/), is composed of the carboxy-terminal domains of the heavy chains. Joining the arms of the Y to the trunk are the flexible hinge regions. The Fc fragment and hinge regions differ in **antibodies** of different [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/), thus determining their functional properties. However, the overall organization of the domains is similar in all isotypes.

# (B) The interaction of the antibody molecule with specific antigen

We have described the structure of the [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule** and how the V regions of the heavy and light chains fold and pair to form the [**antigen**-binding site](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2582/). In this part of the chapter we will look at the **antigen**-binding site in more detail. We will discuss the different ways in which **antigens** can bind to **antibody** and address the question of how variation in the sequences of the **antibody** V domains determines the [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/) for **antigen**.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27160/)

## 3-6. Localized regions of hypervariable sequence form the antigenbinding site

The V regions of any given [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule** differ from those of every other. Sequence [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/) is not, however, distributed evenly throughout the V regions but is concentrated in certain segments of the V region. The distribution of variable amino acids can be seen clearly in what is termed a variability plot ([Fig. 3.6](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A337/?report=objectonly" \t "object)), in which the amino acid sequences of many different **antibody** V regions are compared. Three segments of particular variability can be identified in both the VH and VL domains. They are designated [**hypervariable**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2894/)**regions** and are denoted HV1, HV2, and HV3. In the light chains these are roughly from residues 28 to 35, from 49 to 59, and from 92 to 103, respectively. The most variable part of the domain is in the HV3 region. The regions between the hypervariable regions, which comprise the rest of the [V domain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3342/), show less variability and are termed the [framework regions](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2825/). There are four such regions in each V domain, designated FR1, FR2, FR3, and FR4.

#### [Figure 3.6](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A337/?report=objectonly" \t "object)

There are discrete regions of hypervariability in V domains. A variability plot derived from comparison of the amino acid sequences of several dozen heavy-chain and light-chain V domains. At each amino acid position the degree of variability is the ratio [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A337/?report=objectonly" \t "object)

The [framework regions](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2825/) form the β sheets that provide the structural framework of the domain, whereas the [hypervariable](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2894/) sequences correspond to three loops at the outer edge of the [β barrel](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2617/), which are juxtaposed in the folded domain ([Fig. 3.7](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A338/?report=objectonly" \t "object)). Thus, not only is sequence diversity concentrated in particular parts of the [V domain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3342/) but it is localized to a particular region on the surface of the **molecule**. When the VH and VL domains are paired in the [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule**, the hypervariable loops from each domain are brought together, creating a single hypervariable site at the tip of each arm of the **molecule**. This is the binding site for [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/), the [**antigen**-binding site](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2582/) or **antibody** combining site. The three hypervariable loops determine **antigen** [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/) by forming a surface complementary to the **antigen**, and are more commonly termed the **complementarity-determining regions**, or CDRs (**CDR1**, **CDR2**, and **CDR3**). Because CDRs from both VH and VL domains contribute to the **antigen**-binding site, it is the combination of the heavy and the [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/), and not either alone, that determines the final **antigen** specificity. Thus, one way in which the [immune system](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2921/) is able to generate **antibodies** of different specificities is by generating different combinations of heavy- and light-chain V regions. This means of producing [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/) is known as [combinatorial diversity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2714/); we will encounter a second form of combinatorial diversity when we consider in Chapter 4 how the genes encoding the heavy- and light-chain V regions are created from smaller segments of DNA.

#### [Figure 3.7](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A338/?report=objectonly" \t "object)

The hypervariable regions lie in discrete loops of the folded structure. When the hypervariable regions (CDRs) are positioned on the structure of a V domain it can be seen that they lie in loops that are brought together in the folded structure. In the [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A338/?report=objectonly" \t "object)

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27160/)

## 3-7. Antibodies bind antigens via contacts with amino acids in CDRs, but the details of binding depend upon the size and shape of the antigen

In early investigations of [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) binding to **antibodies**, the only available sources of large quantities of a single type of [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule** were tumors of **antibody**-secreting cells. The **antigen** specificities of the tumor-derived **antibodies** were unknown, so many compounds had to be screened to identify ligands that could be used to study **antigen** binding. In general, the substances found to bind to these **antibodies** were haptens (see [Section 3-4](https://www.ncbi.nlm.nih.gov/books/n/imm/A323/#A330)) such as phosphorylcholine or vitamin K1. Structural analysis of complexes of **antibodies** with their hapten ligands provided the first direct evidence that the [hypervariable](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2894/) regions form the [**antigen**-binding site](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2582/), and demonstrated the structural basis of [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/) for the hapten. Subsequently, with the discovery of methods of generating monoclonal **antibodies** (see [Appendix I, Section A-12](https://www.ncbi.nlm.nih.gov/books/n/imm/A2395/#A2422)), it became possible to make large amounts of pure **antibodies** **specific** for many different **antigens**. This has provided a more general picture of how **antibodies** interact with their **antigens**, confirming and extending the view of **antibody**-**antigen** interactions derived from the study of haptens.

The surface of the [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) **molecule** formed by the juxtaposition of the CDRs of the heavy and light chains creates the site to which an [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) binds. Clearly, as the amino acid sequences of the CDRs are different in different **antibodies**, so are the shapes of the surfaces created by these CDRs. As a general principle, **antibodies** bind ligands whose surfaces are complementary to that of the **antibody**. A small **antigen**, such as a hapten or a short peptide, generally binds in a pocket or groove lying between the heavy- and light-chain V domains ([Fig. 3.8](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A340/?report=objectonly" \t "object), left and center panels). Other **antigens**, such as a protein **molecule**, can be of the same size as, or larger than, the **antibody** **molecule** itself, and cannot fit into a groove or pocket. In these cases, the interface between the two molecules is often an extended surface that involves all of the CDRs and, in some cases, part of the framework region of the **antibody** ([Fig. 3.8](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A340/?report=objectonly" \t "object), right panel). This surface need not be concave but can be flat, undulating, or even convex.

#### [Figure 3.8](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A340/?report=objectonly" \t "object)

**Antigens** can bind in pockets or grooves, or on extended surfaces in the binding sites of **antibodies**. The panels in the top row show schematic representations of the different types of binding site in a Fab fragment of an **antibody**: left, pocket; center, [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A340/?report=objectonly" \t "object)

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27160/)

## 3-8. Antibodies bind to conformational shapes on the surfaces of antigens

The biological function of **antibodies** is to bind to pathogens and their products, and to facilitate their removal from the body. An [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) generally recognizes only a small region on the surface of a large **molecule** such as a polysaccharide or protein. The structure recognized by an **antibody** is called an [antigenic determinant](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2584/) or [epitope](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2800/). Some of the most important pathogens have polysaccharide coats, and **antibodies** that recognize epitopes formed by the sugar subunits of these molecules are essential in providing immune protection from such pathogens. In many cases, however, the **antigens** that provoke an [immune response](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2918/) are proteins. For example, protective **antibodies** against viruses recognize viral coat proteins. In such cases, the structures recognized by the **antibody** are located on the surface of the protein. Such sites are likely to be composed of amino acids from different parts of the polypeptide chain that have been brought together by protein folding. Antigenic determinants of this kind are known as **conformational** or discontinuous epitopes because the structure recognized is composed of segments of the protein that are discontinuous in the amino acid sequence of the [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) but are brought together in the three-dimensional structure. In contrast, an epitope composed of a single segment of polypeptide chain is termed a **continuous** or linear epitope. Although most **antibodies** raised against intact, fully folded proteins recognize discontinuous epitopes, some will bind peptide fragments of the protein. Conversely, **antibodies** raised against peptides of a protein or against synthetic peptides corresponding to part of its sequence are occasionally found to bind to the natural folded protein. This makes it possible, in some cases, to use synthetic peptides in vaccines that aim at raising **antibodies** against a pathogen protein.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27160/)

## 3-9. Antigen-antibody interactions involve a variety of forces

The **interaction** between an [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) and its [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) can be disrupted by high salt concentrations, extremes of pH, detergents, and sometimes by competition with high concentrations of the pure [epitope](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2800/) itself. The binding is therefore a reversible noncovalent **interaction**. The forces, or bonds, involved in these noncovalent interactions are outlined in [Fig. 3.9](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A343/?report=objectonly" \t "object).

#### [Figure 3.9](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A343/?report=objectonly" \t "object)

The noncovalent forces that hold together the **antigen**:**antibody** complex. Partial charges found in electric dipoles are shown as δ+ or δ-. Electrostatic forces diminish as the inverse square of the distance separating the charges, whereas [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A343/?report=objectonly" \t "object)

Electrostatic interactions occur between charged amino acid side chains, as in salt bridges. Interactions also occur between electric dipoles, as in hydrogen bonds, or can involve short-range van der Waals forces. High salt concentrations and extremes of pH disrupt [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-[**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) binding by weakening electrostatic interactions and/or hydrogen bonds. This principle is employed in the purification of **antigens** using affinity columns of immobilized **antibodies**, and vice versa for **antibody** purification (see [Appendix I, Section A-5](https://www.ncbi.nlm.nih.gov/books/n/imm/A2395/#A2406)). Hydrophobic interactions occur when two hydrophobic surfaces come together to exclude water. The strength of a hydrophobic **interaction** is proportional to the surface area that is hidden from water. For some **antigens**, hydrophobic interactions probably account for most of the binding energy. In some cases, water molecules are trapped in pockets in the interface between **antigen** and **antibody**. These trapped water molecules may also contribute to binding, especially between polar amino acid residues.

The contribution of each of these forces to the overall **interaction** depends on the particular [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) and [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) involved. A striking difference between **antibody** interactions with protein **antigens** and most other natural protein-protein interactions is that **antibodies** possess many aromatic amino acids in their **antigen**-binding sites. These amino acids participate mainly in van der Waals and hydrophobic interactions, and sometimes in hydrogen bonds. In general, the hydrophobic and van der Waals forces operate over very short ranges and serve to pull together two surfaces that are complementary in shape: hills on one surface must fit into valleys on the other for good binding to occur. In contrast, electrostatic interactions between charged side chains, and hydrogen bonds bridging oxygen and/or nitrogen atoms, accommodate **specific** features or reactive groups while strengthening the **interaction** overall.

For example, in the complex of hen egg-white lysozyme with the [**antibody**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) D1.3 ([Fig. 3.10](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A344/?report=objectonly" \t "object)), strong hydrogen bonds are formed between the **antibody** and a particular glutamine in the lysozyme **molecule** that protrudes between the VH and VL domains. Lysozymes from partridge and turkey have another amino acid in place of the glutamine and do not bind to the **antibody**. In the high-affinity complex of hen egg-white lysozyme with another **antibody**, HyHel5 (see [Fig. 3.8c](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A340/?report=objectonly" \t "object)), two salt bridges between two basic arginines on the surface of the lysozyme interact with two glutamic acids, one each from the VH CDR1 and CDR2 loops. Again, lysozymes that lack one of the two arginine residues show a 1000-fold decrease in affinity. Although overall surface complementarity must play an important part in [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-**antibody** inter-actions, **specific** electrostatic and hydrogen-bonding interactions appear to determine **antibody** affinity. In most **antibodies** that have been studied at this level of detail, only a few residues make a major contribution to the binding energy. Genetic engineering by site-directed mutagenesis can further tailor an **antibody**'s binding to its complementary [epitope](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2800/).

#### [Figure 3.10](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A344/?report=objectonly" \t "object)

The complex of lysozyme with the **antibody** D1.3. The **interaction** of the Fab fragment of D1.3 with hen egg-white lysozyme is shown, with the lysozyme in blue, the heavy chain in purple and the light chain in yellow. A glutamine residue of lysozyme, shown [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27160/figure/A344/?report=objectonly" \t "object)

# (C) Antigen recognition by T cells

In contrast to the [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/), which interact with pathogens and their toxic products in the extracellular spaces of the body, [**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) only recognize foreign **antigens** that are displayed on the surfaces of the body's own cells. These **antigens** can derive from pathogens such as viruses or intracellular [bacteria](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2622/), which replicate within cells, or from pathogens or their products that cells have internalized by endocytosis from the extracellular fluid.

[**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) can detect the presence of an intracellular pathogen because infected cells display on their surface peptide fragments derived from the pathogen's proteins. These foreign peptides are delivered to the cell surface by specialized host-cell glycoproteins. These are encoded in a large cluster of genes that were first identified by their powerful effects on the [immune response](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2918/) to transplanted tissues. For that reason, the gene complex was called the [**major histocompatibility complex**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3043/)**(MHC)**, and the peptide-binding glycoproteins are still known as [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/). The **recognition** of [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) as a small peptide fragment bound to an MHC molecule and displayed at the cell surface is one of the most distinctive features of **T cells**, and will be the focus of this part of the chapter. How peptide fragments of **antigen** become complexed with MHC molecules will be considered in Chapter 5.

In this part of the chapter we will describe the structure and properties of the [T-cell **antigen** receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3285/), [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/), or TCR for short. As might be expected from their function as highly variable **antigen**-**recognition** structures, T-cell receptors are closely related to [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) molecules in the structure of their genes. There are, however, important differences between T-cell receptors and [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) that reflect the special features of **antigen** **recognition** by the T-cell receptor, and its lack of effector functions.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27098/)

## 3-10. The antigen receptor on T cells is very similar to a Fab fragment of immunoglobulin

T-cell receptors were first identified using monoclonal antibodies that bound only one [cloned T-cell line](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2705/) but not others and that could specifically inhibit [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) **recognition** by that [clone](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2704/) of [**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/), or specifically activate them (see [Appendix I, Section A-19](https://www.ncbi.nlm.nih.gov/books/n/imm/A2395/#A2436)). These [clonotypic](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2706/) antibodies were then used to show that each T cell bears about 30,000 **antigen**-receptor molecules on its surface, each receptor consisting of two different polypeptide chains, termed the [**T-cell receptor**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/)**α** (**TCRα**) and β (**TCRβ**) **chains**, linked by a disulfide bond. These α:β heterodimers are very similar in structure to the Fab fragment of an immunoglobulin molecule ([Fig. 3.11](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A348/?report=objectonly" \t "object)), and they account for **antigen** **recognition** by most **T cells**. A minority of **T cells** bear an alternative, but structurally similar, receptor made up of a different pair of polypeptide chains designated γ and δ. γ:δ T-cell receptors appear to have different **antigen**-**recognition** properties from the α:β T-cell receptors, and the function of γ:δ **T cells** in immune responses is not yet entirely clear. In the rest of this chapter, we shall use the term T-cell receptor to mean the α:β receptor, except where specified otherwise. Both types of T-cell receptor differ from the membrane-bound immunoglobulin that serves as the B-cell receptor: a T-cell receptor has only one [**antigen**-binding site](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2582/), whereas a B-cell receptor has two, and T-cell receptors are never secreted, whereas immunoglobulin can be secreted as [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/).

#### [Figure 3.11](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A348/?report=objectonly" \t "object)

The T-cell receptor resembles a membrane-bound Fab fragment. The Fab fragment of antibody molecules is a disulfide-linked heterodimer, each chain of which contains one immunoglobulin C domain and one V domain; the juxtaposition of the V domains forms the [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A348/?report=objectonly" \t "object)

Our initial insights into the structure and function of the [α:β T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2530/) came from studies of cloned cDNA encoding the receptor chains. The amino acid sequences predicted from T-cell receptor cDNAs show clearly that both chains of the T-cell receptor have an amino-terminal variable (V) region with homology to an immunoglobulin [V domain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3342/), a constant (C) region with homology to an immunoglobulin C domain, and a short [hinge region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2874/) containing a cysteine residue that forms the interchain disulfide bond ([Fig. 3.12](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A349/?report=objectonly" \t "object)). Each chain spans the lipid bilayer by a hydrophobic transmembrane domain, and ends in a short cytoplasmic tail. These close similarities of T-cell receptor chains to the heavy and light immunoglobulin chains first enabled prediction of the structural resemblance of the T-cell receptor heterodimer to a Fab fragment of immunoglobulin.

#### [Figure 3.12](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A349/?report=objectonly" \t "object)

Structure of the T-cell receptor. The T-cell receptor heterodimer is composed of two trans-membrane glycoprotein chains, α and β. The extracellular portion of each chain consists of two domains, resembling immunoglobulin V and C domains, [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A349/?report=objectonly" \t "object)

Recently, the three-dimensional structure of the [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) has been determined. The structure is indeed similar to that of an [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) Fab fragment, as was suspected from earlier studies on the genes that encoded it. The T-cell receptor chains fold in much the same way as those of a Fab fragment ([Fig. 3.13a](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A350/?report=objectonly" \t "object)), although the final structure appears a little shorter and wider. There are, however, some distinct differences between T-cell receptors and [Fab fragments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2810/). The most striking difference is in the Cα domain, where the fold is unlike that of any other immunoglobulin-like domain. The half of the domain that is juxtaposed with the Cβ domain forms a [β sheet](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2619/) similar to that found in other immunoglobulin-like domains, but the other half of the domain is formed of loosely packed strands and a short segment of α helix ([Fig. 3.13b](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A350/?report=objectonly" \t "object)). The intramolecular disulfide bond, which in immunoglobulin-like domains normally joins two β strands, in a Cα domain joins a β strand to this segment of α helix.

#### [Figure 3.13](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A350/?report=objectonly" \t "object)

The crystal structure of an α:β T-cell receptor resolved at 2.5 Å. In panels a and b the α chain is shown in pink and the β chain in blue. Disulfide bonds are shown in green. In panel a, the T-cell receptor is viewed [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A350/?report=objectonly" \t "object)

There are also differences in the way in which the domains interact. The interface between the V and [C domains](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2648/) of both [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) chains is more extensive than in antibodies, which may make the hinge joint between the domains less flexible. And the interaction between the Cα and Cβ domains is distinctive in being assisted by carbohydrate, with a sugar group from the Cα domain making a number of hydrogen bonds to the Cβ domain (see [Fig. 3.13b](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A350/?report=objectonly" \t "object)). Finally, a comparison of the variable binding sites shows that, although the complementarity-determining region ([CDR](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2681/)) loops align fairly closely with those of [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) molecules, there is some displacement relative to those of the antibody molecule ([Fig. 3.13c](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A350/?report=objectonly" \t "object)). This displacement is particularly marked in the Vα CDR2 loop, which is oriented at roughly right angles to the equivalent loop in antibody V domains, as a result of a shift in the β strand that anchors one end of the loop from one face of the domain to the other. A strand displacement also causes a change in the orientation of the Vβ CDR2 loop in two of the seven Vβ domains whose structures are known. As yet, the crystallographic structures of only seven T-cell receptors have been solved to this level of resolution, so it remains to be seen to what degree all T-cell receptors share these features, and whether there is more [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/) to be discovered.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27098/)

## 3-11. A T-cell receptor recognizes antigen in the form of a complex of a foreign peptide bound to an MHC molecule

**Antigen** **recognition** by T-cell receptors clearly differs from **recognition** by B-cell receptors and antibodies. **Antigen** **recognition** by B cells involves direct binding of immunoglobulin to the intact [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) and, as discussed in [Section 3-8](https://www.ncbi.nlm.nih.gov/books/n/imm/A335/#A341), antibodies typically bind to the surface of protein **antigens**, contacting amino acids that are discontinuous in the primary structure but are brought together in the folded protein. [**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/), on the other hand, were found to respond to short contiguous amino acid sequences in proteins. These sequences were often buried within the native structure of the protein and thus could not be recognized directly by T-cell receptors unless some unfolding of the protein **antigen** and its ‘processing’ into peptide fragments had occurred ([Fig. 3.14](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A352/?report=objectonly" \t "object)).

#### [Figure 3.14](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A352/?report=objectonly" \t "object)

Differences in the **recognition** of hen egg-white lysozyme by immunoglobulins and T-cell receptors. Antibodies can be shown by X-ray crystallography to bind epitopes on the surface of proteins, as shown in panel a, where the epitopes for three antibodies [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A352/?report=objectonly" \t "object)

The nature of the [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) recognized by [**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) became clear with the realization that the peptides that stimulate **T cells** are recognized only when bound to an MHC molecule. These cell-surface glycoproteins are encoded by genes within the [major histocompatibility complex](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3043/) (MHC). The ligand recognized by the T cell is thus a complex of peptide and MHC molecule. The evidence for involvement of the MHC in T-cell **recognition** of **antigen** was at first indirect, but it has recently been proved conclusively by stimulating **T cells** with purified peptide:MHC complexes. The [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) interacts with this ligand by making contacts with both the MHC molecule and the **antigen** peptide.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27098/)

## 3-12. T cells with different functions are distinguished by CD4 and CD8 cell-surface proteins and recognize peptides bound to different classes of MHC molecule

[**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) fall into two major classes that have different effector functions. The two classes are distinguished by the expression of the cell-surface proteins [CD4](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2668/) and [CD8](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2671/). These two types of T cell differ in the class of MHC molecule they recognize. There are two classes of MHC molecule—**MHC class I** and **MHC class II**—which differ in their structure and expression pattern on tissues of the body (see [Section 3-13](https://www.ncbi.nlm.nih.gov/books/NBK27098/#A358)). CD4 and CD8 were known as markers for different functional sets of **T cells** for some time before it became clear that they play an important part in the direct **recognition** of MHC class II and MHC class I molecules, respectively. CD4 binds to the MHC class II molecule and CD8 to the MHC class I molecule. During [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) **recognition**, depending on the type of T-cell, CD4 or CD8 molecules associate on the T-cell surface with the [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) and bind to invariant sites on the MHC portion of the composite MHC:peptide ligand. This binding is required for the T cell to make an effective response, and so CD4 and CD8 are called co-receptors.

[CD4](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2668/) is a single-chain molecule composed of four immunoglobulin-like domains ([Fig. 3.15](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A354/?report=objectonly" \t "object)). The first two domains (D1 and D2) of the CD4 molecule are packed tightly together to form a rigid rod some 60 Å long, which is joined by a flexible hinge to a similar rod formed by the third and fourth domains (D3 and D4). CD4 binds MHC class II molecules through a region that is located mainly on a lateral face of the first domain, D1. Because CD4 binds to a site on the β2 domain of the MHC class II molecule that is well away from the site where the [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) binds ([Fig. 3.16a](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A355/?report=objectonly" \t "object)), the CD4 molecule and the T-cell receptor can bind the same peptide:MHC class II complex. CD4 interacts strongly with a cytoplasmic [tyrosine kinase](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3340/) called [Lck](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2992/), and can deliver this tyrosine kinase into close proximity with the signaling components of the T-cell receptor complex. This results in enhancement of the signal that is generated when the T-cell receptor binds its peptide:MHC class II ligand, as we will discuss further in Chapter 6. When CD4 and the T-cell receptor can simultaneously bind to the same MHC class II:peptide complex, the sensitivity of a T cell to [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) presented by MHC class II molecules is markedly increased; the T-cell in this case requires 100-fold less **antigen** for activation.

#### [Figure 3.15](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A354/?report=objectonly" \t "object)

The outline structures of the CD4 and CD8 co-receptor molecules. . The CD4 molecule contains four immunoglobulin-like domains, as shown in diagrammatic form in panel a, and as a ribbon diagram of the structure in panel b. The amino-terminal domain, D [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A354/?report=objectonly" \t "object)

#### [Figure 3.16](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A355/?report=objectonly" \t "object)

The binding sites for CD4 and CD8 on MHC class II and class I molecules lie in the immunoglobulin-like domains. The binding sites for CD4 and CD8 on the MHC class II and class I molecules, respectively, lie in the immunoglobulin-like domains nearest to [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A355/?report=objectonly" \t "object)

[CD4](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2668/) binding to an MHC class II molecule on its own is weak, and it is not clear whether such binding would be able to transmit a signal to the interior of the T cell. As shown in [Fig. 3.17](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A356/?report=objectonly" \t "object), CD4 can form homodimers through a site in the D4 domain, which leaves the MHC-binding site free to interact with an MHC class II molecule*.* Thus, the CD4 dimer could cross-link two MHC class II molecules and thus the two T-cell receptors bound to them. Whether the dimerization of CD4 is important in its [co-receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2733/) function is not known at present.

#### [Figure 3.17](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A356/?report=objectonly" \t "object)

CD4 is capable of forming dimers. The structure of the extra-cellular domains of the CD4 molecule has been determined by X-ray crystallography. Two molecules of CD4 can interact with each other through their D4 domains, forming homodimers. The site that [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A356/?report=objectonly" \t "object)

Although [CD4](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2668/) and [CD8](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2671/) both function as co-receptors, their structures are quite distinct. The CD8 molecule is a disulfide-linked heterodimer consisting of an α and a β chain, each containing a single immunoglobulin-like domain linked to the membrane by a segment of extended polypeptide chain (see [Fig. 3.15](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A354/?report=objectonly" \t "object)). This segment is extensively glycosylated, which is thought to be important in maintaining this polypeptide in an extended conformation and protecting it from cleavage by proteases. CD8α chains can also form homo-dimers, although these are not found when the CD8β chains are present.

[CD8](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2671/) binds weakly to an invariant site in the α3 domain of an MHC class I molecule ([Fig. 3.16b](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A355/?report=objectonly" \t "object)), which is equivalent to the site in MHC class II molecules to which [CD4](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2668/) binds. Although only the interaction of the CD8α homodimer with MHC class I is so far known in detail, it is clear from this that the MHC class I binding site of the CD8 α:β heterodimer will be formed by the interaction of the CD8α and β chains. In addition, CD8 (most probably through the α chain) interacts with residues in the base of the α2 domain of the MHC class I molecule. Binding in this way, CD8 leaves the upper surface of the MHC class I molecule exposed and free to interact simultaneously with a [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/), as shown in [Fig. 3.18](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A357/?report=objectonly" \t "object). Like CD4, CD8 also binds [Lck](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2992/) through the cytoplasmic tail of the α chain and brings it into close proximity with the T-cell receptor. And as with CD4, the presence of CD8 increases the sensitivity of [**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) to [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) presented by MHC class I molecules by about 100-fold. Thus, CD4 and CD8 have similar functions and bind to the same approximate location in MHC class I and MHC class II molecules even though the structures of the two [co-receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2733/) proteins are only distantly related.

#### [Figure 3.18](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A357/?report=objectonly" \t "object)

CD8 binds to a site on MHC class I molecules distant from that to which the T-cell receptor binds. The relative positions of the T-cell receptor and CD8 molecules bound to the same MHC class I molecule can be seen in this hypothetical reconstruction of [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A357/?report=objectonly" \t "object)

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27098/)

## 3-13. The two classes of MHC molecule are expressed differentially on cells

MHC class I and MHC class II molecules have a distinct distribution among cells that reflects the different effector functions of the [**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) that recognize them ([Fig. 3.19](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A359/?report=objectonly" \t "object)). MHC class I molecules present peptides from pathogens, commonly viruses, to [CD8](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2671/) [cytotoxic **T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2755/), which are specialized to kill any cell that they specifically recognize. As viruses can infect any nucleated cell, almost all such cells express MHC class I molecules, although the level of constitutive expression varies from one cell type to the next. For example, cells of the [immune system](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2921/) express abundant MHC class I on their surface, whereas liver cells (hepatocytes) express relatively low levels (see [Fig. 3.19](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A359/?report=objectonly" \t "object)). Nonnucleated cells, such as mammalian red blood cells, express little or no MHC class I, and thus the interior of red blood cells is a site in which an infection can go undetected by cytotoxic **T cells**. As red blood cells cannot support viral replication, this is of no great consequence for viral infection, but it may be the absence of MHC class I that allows the *Plasmodium* species that cause malaria to live in this privileged site.

#### [Figure 3.19](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A359/?report=objectonly" \t "object)

The expression of MHC molecules differs between tissues. MHC class I molecules are expressed on all nucleated cells, although they are most highly expressed in hematopoietic cells. MHC class II molecules are normally expressed only by a subset of hematopoietic cells [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A359/?report=objectonly" \t "object)

In contrast, the main function of the [CD4 **T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2669/) that recognize MHC class II molecules is to activate other effector cells of the [immune system](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2921/). Thus MHC class II molecules are normally found on B [lymphocytes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3021/), dendritic cells, and macrophages—cells that participate in immune responses—but not on other tissue cells (see [Fig. 3.19](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A359/?report=objectonly" \t "object)). When CD4 **T cells** recognize peptides bound to MHC class II molecules on B cells, they stimulate the B cells to produce [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) Likewise, CD4 **T cells** recognizing peptides bound to MHC class II molecules on macrophages activate these cells to destroy the pathogens in their vesicles We shall see in Chapter 8 that MHC class II molecules are also expressed on specialized [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-presenting cells in lymphoid tissues where naive **T cells** encounter **antigen** and are first activated. The expression of both MHC class I and MHC class II molecules is regulated by cytokines, in particular interferons, released in the course of immune responses. Interferon-γ ([IFN](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2902/)-γ), for example, increases the expression of MHC class I and MHC class II molecules, and can induce the expression of MHC class II molecules on certain cell types that do not normally express them. [Interferons](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2964/) also enhance the antigenpresenting function of MHC class I molecules by inducing the expression of key components of the intracellular machinery that enables peptides to be loaded onto the [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/).

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27098/)

## 3-14. The two classes of MHC molecule have distinct subunit structures but similar three-dimensional structures

The two classes of MHC molecule differ from each other in their structure and also have different distributions on the cells of the body. Their different structures enable the two classes of [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) to serve distinct functions in [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) presentation, binding peptides from different intracellular sites and activating different subsets of [**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/), as we will see in Chapter 5. Despite their differences in subunit structure, however, MHC class I and class II molecules are closely related in overall structure. In both classes, the two paired protein domains nearest the membrane resemble [immunoglobulin domains](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2935/), whereas the two domains distal to the membrane fold together to create a long cleft, or groove, which is the site at which a peptide binds. Purified peptide:MHC class I and peptide:MHC class II complexes have been characterized structurally, allowing us to describe in detail both the MHC molecules themselves and the way in which they bind peptides.

MHC class I structure is outlined in [Fig. 3.20](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A361/?report=objectonly" \t "object). MHC class I molecules consist of two polypeptide chains, a larger α chain encoded in the MHC genetic locus, and a smaller noncovalently associated chain, β2**-microglobulin**, which is not polymorphic and is not encoded in the MHC locus. Only the class I α chain spans the membrane. The complete molecule has four domains, three formed from the MHC-encoded α chain, and one contributed by [β2-microglobulin.](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2620/) The α3 domain and β2-microglobulin have a folded structure that closely resembles that of an immunoglobulin domain. The most remarkable feature of MHC class I molecules is the structure of the folded α1 and α2 domains. These two domains form the walls of a cleft on the surface of the molecule; this is the site of peptide binding. They also are sites of polymorphisms that determine T-cell [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) **recognition** (see Chapter 5).

#### [Figure 3.20](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A361/?report=objectonly" \t "object)

The structure of an MHC class I molecule determined by X-ray crystallography. Panel a shows a computer graphic representation of a human MHC class I molecule, HLA-A2, which has been cleaved from the cell surface by the enzyme papain. The surface of the [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A361/?report=objectonly" \t "object)

An MHC class II molecule consists of a noncovalent complex of two chains, α and β, both of which span the membrane ([Fig. 3.21](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A362/?report=objectonly" \t "object)). The MHC class II α and β chains are both encoded within the MHC. The crystallographic structure of the MHC class II molecule shows that it is folded very much like the MHC class I molecule. The major differences lie at the ends of the peptide-binding cleft, which are more open in MHC class II molecules compared with MHC class I molecules. The main consequence of this is that the ends of a peptide bound to an MHC class I molecule are substantially buried within the molecule, whereas the ends of peptides bound to MHC class II molecules are not. Again, the sites of major polymorphism are located in the peptide-binding cleft, which in the case of an MHC class II molecule are formed by the α1, and β1 domains (see Chapter 5).

#### [Figure 3.21](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A362/?report=objectonly" \t "object)

MHC class II molecules resemble MHC class I molecules in overall structure. The MHC class II molecule is composed of two trans-membrane glycoprotein chains, α (34 kDa) and β (29 kDa), as shown schematically in panel d. Each chain has two [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A362/?report=objectonly" \t "object)

In both MHC class I and class II molecules, bound peptides are sandwiched between the two α-helical segments of the MHC molecule ([Fig. 3.22](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A363/?report=objectonly" \t "object)). The [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) interacts with this compound ligand, making contacts with both the MHC molecule and with the peptide fragment of [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/).

#### [Figure 3.22](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A363/?report=objectonly" \t "object)

MHC molecules bind peptides tightly within the cleft. When MHC molecules are crystallized with a single synthetic peptide **antigen**, the details of peptide binding are revealed. In MHC class I molecules (panels a and c) the peptide is bound in an elongated conformation [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A363/?report=objectonly" \t "object)

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27098/)

## 3-15. Peptides are stably bound to MHC molecules, and also serve to stabilize the MHC molecule on the cell surface

An individual can be infected by a wide variety of different pathogens the proteins of which will not generally have peptide sequences in common. If [**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) are to be alerted to all possible infections, then the [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) on each cell (both class I and class II) must be able to bind stably to many different peptides. This behavior is quite distinct from that of other peptide-binding receptors, such as those for peptide hormones, which usually bind only a single type of peptide. The crystal structures of peptide:MHC complexes have helped to show how a single binding site can bind peptides with high affinity while retaining the ability to bind a wide variety of different peptides.

An important feature of the binding of peptides to [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) is that the peptide is bound as an integral part of the MHC molecule's structure, and MHC molecules are unstable when peptides are not bound. The stability of peptide binding is important, because otherwise, peptide exchanges occurring at the cell surface would prevent peptide:MHC complexes from being reliable indicators of infection or of uptake of specific [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/). As a result of this stability, when MHC molecules are purified from cells, their bound peptides co-purify with them, and this has enabled the peptides bound by specific MHC molecules to be analyzed. The peptides are released from the MHC molecules by denaturing the complex in acid, and can then be purified and sequenced. Pure synthetic peptides can also be incorporated into previously empty MHC molecules and the structure of the complex determined, revealing details of the contacts between the MHC molecule and the peptide. From the sequences of peptides bound to specific MHC molecules, combined with structural analysis of the peptide:MHC complex, a detailed picture of the binding interactions has been built up. We will first discuss the peptide-binding properties of MHC class I molecules.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27098/)

## 3-16. MHC class I molecules bind short peptides of 8–10 amino acids by both ends

The binding of a peptide in the peptide-binding cleft of an MHC class I molecule is stabilized at both ends by contacts between atoms in the free amino and carboxy termini of the peptide and invariant sites that are found at each end of the cleft of all MHC class I molecules ([Fig. 3.23](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A366/?report=objectonly" \t "object)). These contacts are thought to be the main stabilizing contacts for peptide:MHC class I complexes because synthetic peptide analogues lacking terminal amino and carboxyl groups fail to bind stably to MHC class I molecules. Other residues in the peptide serve as additional anchors. Peptides that bind to MHC class I molecules are usually 8–10 amino acids long. The peptide lies in an elongated conformation along the groove; variations in peptide length appear to be accommodated, in most cases, by a kinking in the peptide backbone. However, two examples of MHC class I molecules where the peptide is able to extend out of the groove at the carboxy terminus suggest that some length variation may also be accommodated in this way.

#### [Figure 3.23](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A366/?report=objectonly" \t "object)

Peptides are bound to MHC class I molecules by their ends. MHC class I molecules interact with the back-bone of a bound peptide (shown in yellow) through a series of hydrogen bonds and ionic interactions (shown as dotted blue lines) at each end of the peptide. [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A366/?report=objectonly" \t "object)

These interactions give all MHC class I molecules their broad peptide-binding [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/). In addition, [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) are highly polymorphic. There are hundreds of different versions, or alleles, of the MHC class I genes in the human population as a whole, and each individual carries only a small selection of them. The main differences between the allelic MHC variants are found at certain sites in the peptide-binding cleft, resulting in different amino acids in key peptide interaction sites in the different MHC variants. The consequence of this is that the different MHC variants preferentially bind different peptides. The peptides that can bind to a given MHC variant have the same or very similar amino acid residues at two or three particular positions along the peptide sequence. The amino acid side chains at these positions insert into pockets in the MHC molecule that are lined by the polymorphic amino acids. Because the binding of these side chains anchors the peptide to the MHC molecule, the peptide residues involved have been called [anchor residues](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2572/). Both the position and identity of these anchor residues can vary, depending on the particular MHC class I variant that is binding the peptide. However, most peptides that bind to MHC class I molecules have a hydrophobic (or sometimes basic) anchor residue at the carboxy terminus ([Fig. 3.24](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A367/?report=objectonly" \t "object)). Changing an anchor residue can prevent the peptide from binding and, conversely, most synthetic peptides of suitable length that contain these anchor residues will bind the appropriate MHC class I molecule, in most cases irrespective of the amino acids at other positions in the peptide. These features of peptide binding enable an individual MHC class I molecule to bind a wide variety of different peptides, yet allow different MHC class I allelic variants to bind different sets of peptides.

#### [Figure 3.24](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A367/?report=objectonly" \t "object)

Peptides bind to MHC molecules through structurally related anchor residues. Peptides eluted from two different MHC class I molecules are shown. The anchor residues (green) differ for peptides that bind different alleles of MHC class I molecules but are [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A367/?report=objectonly" \t "object)

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27098/)

## 3-17. The length of the peptides bound by MHC class II molecules is not constrained

Peptide binding to MHC class II molecules has also been analyzed by elution of bound peptides and by X-ray crystallography, and differs in several ways from peptide binding to MHC class I molecules. Peptides that bind to MHC class II molecules are at least 13 amino acids long and can be much longer. The clusters of conserved residues that bind the two ends of a peptide in MHC class I molecules are not found in MHC class II molecules, and the ends of the peptide are not bound. Instead, the peptide lies in an extended conformation along the MHC class II peptide-binding groove. It is held in this groove both by peptide side chains that protrude into shallow and deep pockets lined by polymorphic residues, and by interactions between the peptide backbone and side chains of conserved amino acids that line the peptide-binding cleft in all MHC class II molecules ([Fig. 3.25](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A369/?report=objectonly" \t "object)). Although there are fewer crystal structures of MHC class II-bound peptides than of MHC class I, the available data show that amino acid side chains at residues 1, 4, 6, and 9 of an MHC class II-bound peptide can be held in these binding pockets.

#### [Figure 3.25](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A369/?report=objectonly" \t "object)

Peptides bind to MHC class II molecules by interactions along the length of the binding groove. A peptide (yellow; shown as the peptide backbone only, with the amino terminus to the left and the carboxy terminus to the right), is bound by an MHC class [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A369/?report=objectonly" \t "object)

The binding pockets of MHC class II molecules are more permissive in their accommodation of different amino acid side chains than are those of the MHC class I molecule, making it more difficult to define [anchor residues](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2572/) and predict which peptides will be able to bind particular MHC class II molecules ([Fig. 3.26](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A370/?report=objectonly" \t "object)). Nevertheless, by comparing the sequences of known binding peptides, it is usually possible to detect a pattern of permissive amino acids for each of the different alleles of MHC class II molecules, and to model how the amino acids of this peptide [sequence motif](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3234/) will interact with the amino acids that make up the peptide-binding cleft in the MHC class II molecule. Because the peptide is bound by its backbone and allowed to emerge from both ends of the binding groove there is, in principle, no upper limit to the length of peptides that could bind to MHC class II molecules. However, it appears that longer peptides bound to MHC class II molecules are trimmed by peptidases to a length of 13–17 amino acids in most cases. Like MHC class I molecules, MHC class II molecules that lack bound peptide are unstable, but the critical stabilizing interactions that the peptide makes with the MHC class II molecule are not yet known.

#### [Figure 3.26](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A370/?report=objectonly" \t "object)

Peptides that bind MHC class II molecules are variable in length and their anchor residues lie at various distances from the ends of the peptide. The sequences of a set of peptides that bind to the mouse MHC class II Ak allele are shown in the upper panel. [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A370/?report=objectonly" \t "object)

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27098/)

## 3-18. The crystal structures of several MHC:peptide:T-cell receptor complexes all show the same T-cell receptor orientation over the MHC:peptide complex

At the time that the first X-ray crystallographic structure of a [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) was published, a structure of the same T-cell receptor bound to a peptide:MHC class I ligand was also produced. This structure ([Fig. 3.27](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A372/?report=objectonly" \t "object)), which had been forecast by site-directed mutagenesis of the MHC class I molecule, showed the T-cell receptor aligned diagonally over the peptide and the peptide-binding groove, with the T-cell receptor α chain lying over the α2 domain and the amino-terminal end of the bound peptide, the T-cell receptor β chain lying over the α1 domain and the carboxy-terminal end of the peptide, with the CDR3 loops of both T-cell receptor α and T-cell receptor β meeting over the central amino acids of the peptide. The T-cell receptor is threaded through a valley between the two high points on the two surrounding α helices that form the walls of the peptide-binding cleft.

#### [Figure 3.27](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A372/?report=objectonly" \t "object)

The T-cell receptor binds to the MHC:peptide complex. Panel a: the T-cell receptor binds to the top of the MHC:peptide complex, straddling, in the case of the class I molecule shown here, both the α1 and α2 domain helices. The CDRs of the [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A372/?report=objectonly" \t "object)

Analysis of other MHC class I:peptide:[T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) complexes and of the single example so far of an MHC class II:peptide:T-cell receptor complex ([Fig. 3.28](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A373/?report=objectonly" \t "object)) shows that all of them have a very similar orientation, particularly for the Vα domain, although some [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/) does occur in the location and orientation of the Vβ domain. In this orientation, the Vα domain makes contact primarily with the amino terminus of the bound peptide, whereas the Vβ domain contacts primarily the carboxy terminus of the bound peptide. Both chains also interact with the α helices of the MHC class I molecule (see [Fig. 3.27](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A372/?report=objectonly" \t "object)). The T-cell receptor contacts are not symmetrically distributed over the MHC molecule, so whereas the Vα CDR1 and CDR2 loops are in close contact with the helices of the MHC:peptide complex around the amino terminus of the bound peptide, the β-chain CDR1 and CDR2 loops, which interact with the complex at the carboxy terminus of the bound peptide, have variable contributions to the binding. This suggests that the Vα contacts are responsible for the conserved orientation of the T-cell receptor on the MHC:peptide complex.

#### [Figure 3.28](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A373/?report=objectonly" \t "object)

The T-cell receptor interacts with MHC class I and MHC class II molecules in a similar fashion. The structure of a T-cell receptor binding to an MHC class II molecule has been determined, and shows the T-cell receptor binding to an equivalent site, and [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27098/figure/A373/?report=objectonly" \t "object)

Comparison of the three-dimensional structure of the [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) to that of the same T-cell receptor complexed to its MHC-peptide ligand could address the question of whether the T-cell receptor, like some other receptors, undergoes a conformational change, or ‘induced fit,’ in its three-dimensional structure when it binds its specific ligand. To date, there is no certain answer, owing to the limitations of the available structures. For one thing, all the T-cell receptors analyzed to date have either been bound to ligands that do not produce activation, or are bound to ligands that can activate them, but the comparable unliganded receptor structures are not available. Also, the crystals of the T-cell receptor are all formed at 0 °C or below, which locks the receptor into a single conformation. However, subtly different peptides can have strikingly different effects when the same T cell recognizes either of the two peptides complexed with MHC. This could be due to differences in how T-cell receptor conformation is altered by binding the two related yet different ligands. Recent evidence also suggests that the temperature at which the T-cell receptor binds to a particular peptide:MHC complex makes a large difference in the extent of T-cell receptor aggregation; protein conformation is affected by temperature, and so these differences may well result from a conformational change.

From an examination of these structures it is hard to predict whether the main binding energy is contributed by [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) contacts with the bound peptide, or by T-cell receptor contacts with the MHC molecule. It is known that alterations as simple as changing a leucine to isoleucine in the peptide are sufficient to alter the T-cell response from strong killing to absolutely no response at all. Studies show that mutations of single residues in the presenting [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) can have the same effect. Thus, the [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/) of T-cell **recognition** involves both the peptide and its presenting MHC molecule. This dual specificity underlies the MHC restriction of T-cell responses, a phenomenom that was observed long before the peptidebinding properties of MHC molecules were known. We will recount the story of how MHC restriction was discovered when we return to the issue of how MHC polymorphism affects [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) **recognition** by [**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) in Chapter 5. Another consequence of this dual specificity is a need for T-cell receptors to be able to interact appropriately with the **antigen**-presenting surface of MHC molecules. It appears that there is some inherent specificity for MHC molecules encoded in the T-cell receptor genes, as well as selection during T-cell development for a repertoire of receptors able to interact appropriately with the particular MHC molecules present in that individual. We will be discussing the evidence for this in Chapter 7.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27098/)

## 3-19. A distinct subset of T cells bears an alternative receptor made up of γ and δ chains

During the search for the gene for the [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) α chain, another T-cell receptorlike gene was unexpectedly discovered. This gene was named T-cell receptor γ, and its discovery led to a search for further T-cell receptor genes. Another receptor chain was identified using [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) to the predicted sequence of the γ chain and was called the δ chain. It was soon discovered that a minority population of [**T cells**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) bore a distinct type of T-cell receptor made up of γ:δ heterodimers rather than α:β heterodimers. The development of these cells is described in [Sections 7-13](https://www.ncbi.nlm.nih.gov/books/n/imm/A820/#A832) and [7-14](https://www.ncbi.nlm.nih.gov/books/n/imm/A820/#A834).

To date, there is no crystallographic structure of a [γ:δ T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2831/), although it is expected to be similar in shape to α:β T-cell receptors. γ:δ T-cell receptors may be specialized to bind certain kinds of ligands, including heat-shock proteins and nonpeptide ligands such as mycobacterial lipid **antigens**. It seems likely that γ:δ T-cell receptors are not restricted by the ‘classical’ MHC class I and class II molecules. They may bind the free [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/), much as [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) do, and/or they may bind to peptides or other **antigens** presented by nonclassical MHC-like molecules. These are proteins that resemble MHC class I molecules but are relatively nonpolymorphic.

Chapter 4. The Generation of Lymphocyte Antigen Receptors

* The generation of diversity in immunoglobulins
  + 4-1. **Immunoglobulin** genes are rearranged in antibody-producing cells
  + 4-2. The DNA sequence encoding a complete V region is generated by the somatic recombination of separate gene segments
  + 4-3. There are multiple different V-region gene segments
  + 4-4. Rearrangement of V, D, and J gene segments is guided by flanking DNA sequences
  + 4-5. The reaction that recombines V, D, and J gene segments involves both lymphocyte-specific and ubiquitous DNA-modifying enzymes
  + 4-6. The diversity of the **immunoglobulin** repertoire is generated by four main processes
  + 4-7. The multiple inherited gene segments are used in different combinations
  + 4-8. Variable addition and subtraction of nucleotides at the junctions between gene segments contributes to diversity in the third hypervariable region
  + 4-9. Rearranged V genes are further diversified by somatic hypermutation
  + 4-10. In some species most **immunoglobulin** gene diversification occurs after gene rearrangement
  + Summary
* T-cell receptor gene rearrangement
  + 4-11. The T-cell receptor loci comprise sets of gene segments and are rearranged by the same enzymes as the **immunoglobulin** loci
  + 4-12. T-cell receptors concentrate diversity in the third hypervariable region
  + 4-13. γ:δ T-cell receptors are also generated by gene rearrangement
  + 4-14. Somatic hypermutation does not generate diversity in T-cell receptors
  + Summary
* **Structural** **variation** in **immunoglobulin** **constant** **regions**
  + 4-15. The **immunoglobulin** heavy-chain isotypes are distinguished by the structure of their **constant** **regions**
  + 4-16. The same VH exon can associate with different CH genes in the course of an immune response
  + 4-17. Transmembrane and secreted forms of **immunoglobulin** are generated from alternative heavy-chain transcripts
  + 4-18. Antibody C **regions** confer functional specialization
  + 4-19. IgM and IgA can form polymers
  + 4-20. Various differences between immunoglobulins can be detected by antibodies

# Chapter 4The Generation of Lymphocyte Antigen Receptors

**Lymphocyte** [**antigen** **receptors**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2590/), in the form of [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) on B cells and T-cell **receptors** on [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/), are the means by which [lymphocytes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3021/) sense the presence of **antigens** in their environment. The **receptors** produced by each **lymphocyte** have a unique **antigen** [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/), which is determined by the structure of their [**antigen**-binding site](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2582/), as described in Chapter 3. Because each person possesses billions of lymphocytes, these cells collectively provide the individual with the ability to respond to a great variety of **antigens**. The wide range of **antigen** specificities in the **antigen** receptor repertoire is due to variation in the amino acid sequence at the **antigen**-binding site, which is made up from the variable (V) regions of the receptor protein chains. In each chain the V region is linked to an invariant constant (C) region, which provides effector or signaling functions.

Given the importance of a diverse repertoire of **lymphocyte** **receptors** in the defense against infection, it is not surprising that a complex and elegant genetic mechanism has evolved for generating these highly variable proteins. Each receptor chain variant cannot be encoded in full in the genome, as this would require more genes for [**antigen** **receptors**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2590/) than there are genes in the entire genome. Instead, we will see that the V regions of the receptor chains are encoded in several pieces—so-called [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/). These are assembled in the developing **lymphocyte** by somatic DNA recombination to form a complete V-region sequence, a mechanism known generally as **gene**[**rearrangement**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/). Each type of gene segment is present in multiple copies in the germline genome. The selection of a gene segment of each type during gene rearrangement occurs at random, and the large number of possible different combinations accounts for much of the diversity of the receptor repertoire.

In the first two parts of this chapter we will describe the gene [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) mechanism that generates the V regions of immunoglobulin and [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) genes. The basic mechanism is common to both B cells and [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/), and involves many if not all of the same enzymes. We will describe the details of the enzymology of this recombination process, the evolution of which was probably critical to the evolution of the vertebrate adaptive [immune system](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2921/).

In B cells, but not [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/), the rearranged V region undergoes additional modification, known as somatic hypermutation. This does not occur until after B cells encounter and become activated by [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/). In these cells, the V regions of the assembled immunoglobulin genes undergo a high rate of point mutation that creates additional diversity within the expanding [clone](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2704/) of B cells responding to **antigen**.

In the third part of the chapter we consider the limited, but functionally important, diversity of immunoglobulin C regions. The C regions of T-cell **receptors** do not show such diversity as they function only as part of a membrane-bound [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) receptor. Their role is to anchor and support the V regions at the cell surface as well as linking the binding of **antigen** by the V regions to the receptor-associated intracellular signaling complex. The C regions of [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) also serve these functions but in addition the C regions of the [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/) are responsible for the effector functions of the secreted immunoglobulins, or antibodies, made by activated B cells. These CH regions come in several different versions, or [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/), each of which has a different effector function. In B cells that have become activated by **antigen**, the heavy-chain V region can become associated with a different CH region by a further [somatic recombination](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3257/) event, in the process known as [isotype switching](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2973/). This enables the different heavy-chain C regions, each with a different function, to be represented among antibodies of the same **antigen** [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/).

# The generation of diversity in immunoglobulins

Virtually any substance can elicit an [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) response. Furthermore, the response even to a simple [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) bearing a single [antigenic determinant](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2584/) is diverse, comprising many different antibody molecules each with a unique affinity, or binding strength, for the antigen and a subtly different [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/). The total number of antibody specificities available to an individual is known as the [antibody repertoire](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2577/), or [immunoglobulin repertoire](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2939/), and in humans is at least 1011, perhaps many more. The number of antibody specificities present at any one time is, however, limited by the total number of B cells in an individual, as well as by each individual's encounters with antigens.

Before it was possible to examine the immunoglobulin genes directly, there were two main hypotheses for the origin of this **diversity**. The [germline theory](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2844/) held that there is a separate gene for each different immunoglobulin chain and that the [antibody repertoire](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2577/) is largely inherited. By contrast, **somatic diversification theories** proposed that the observed repertoire is generated from a limited number of inherited V-region sequences that undergo alteration within B cells during the individual's lifetime. Cloning of the immunoglobulin genes revealed that the antibody repertoire is, in fact, generated by DNA rearrangements during B-cell development. As we will see in this part of the chapter, a DNA sequence encoding a V region is assembled at each locus by selection from a relatively small group of inherited [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/). **Diversity** is further enhanced by the process of somatic hypermutation in mature activated B cells. Thus the [somatic diversification theory](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3254/) was essentially correct, although the concept of multiple germline genes embodied in the germline theory also proved true.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27140/)

## 4-1. Immunoglobulin genes are rearranged in antibody-producing cells

In nonlymphoid cells, the [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) encoding the greater part of the V region of an immunoglobulin chain are some considerable distance away from the sequence encoding the [C region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2649/). In mature B [lymphocytes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3021/), however, the assembled V-region sequence lies much nearer the C region, as a consequence of gene [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/). Rearrangement within the immunoglobulin genes was originally discovered 25 years ago, when it first became possible to study the organization of the immunoglobulin genes in both B cells and nonlymphoid cells using restriction enzyme analysis and Southern blotting. In this procedure, chromosomal DNA is first cut with a restriction enzyme, and the DNA fragments containing particular V- and C-region sequences are identified by hybridization with radiolabeled DNA probes specific for the relevant DNA sequences. In germline DNA, from nonlymphoid cells, the V- and C-region sequences identified by the probes are on separate DNA fragments. However, in DNA from an [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/)-producing [B cell](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2618/) these V- and C-region sequences are on the same DNA fragment, showing that a rearrangement of the DNA has occurred. A typical experiment using human DNA is shown in [Fig. 4.1](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A438/?report=objectonly" \t "object).

#### [Figure 4.1](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A438/?report=objectonly" \t "object)

Immunoglobulin genes are rearranged in B cells. The two photographs on the left (germline DNA) show a Southern blot of a restriction enzyme digest of DNA from nonlymphoid cells from a normal person. The locations of immunoglobulin DNA sequences are identified [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A438/?report=objectonly" \t "object)

This simple experiment showed that segments of genomic DNA within the immunoglobulin genes are rearranged in cells of the B-lymphocyte lineage, but not in other cells. This process of [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) is known as [somatic recombination](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3257/), to distinguish it from the meiotic recombination that takes place during the production of gametes.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27140/)

## 4-2. The DNA sequence encoding a complete V region is generated by the somatic recombination of separate gene segments

The V region, or [V domain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3342/), of an immunoglobulin heavy or [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/) is encoded by more than one gene segment. For the light chain, the V domain is encoded by two separate DNA segments. The first segment encodes the first 95–101 amino acids of the light chain and is termed a **V gene segment** because it encodes most of the V domain. The second segment encodes the remainder of the V domain (up to 13 amino acids) and is termed a **joining** or **J gene segment**.

The rearrangements that lead to the production of a complete immunoglobulin light-chain gene are shown in [Fig. 4.2](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A440/?report=objectonly" \t "object) (center panel). The joining of a V and a J gene segment creates a continuous exon that encodes the whole of the light-chain V region. In the unrearranged DNA, the [V gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3343/) are located relatively far away from the [C region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2649/). The [J gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2976/) are located close to the C region, however, and joining of a V segment to a J gene segment also brings the V gene close to a C-region sequence. The J gene segment of the rearranged V region is separated from a C-region sequence only by an intron. In the experiment shown in [Fig. 4.1](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A438/?report=objectonly" \t "object), the germline DNA fragment identified by the ‘V-region probe’ contains the V gene segment, and that identified by the ‘C-region probe’ actually contains both the J gene segment and the C-region sequence. To make a complete immunoglobulin light-chain messenger RNA, the V-region exon is joined to the C-region sequence by RNA splicing after transcription (see [Fig. 4.2](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A440/?report=objectonly" \t "object)).

#### [Figure 4.2](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A440/?report=objectonly" \t "object)

V-region genes are constructed from gene segments. Light-chain V-region genes are constructed from two segments (center panel). A variable (V) and a joining (J) gene segment in the genomic DNA are joined to form a complete light-chain V-region exon. Immunoglobulin [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A440/?report=objectonly" \t "object)

A heavy-chain V region is encoded in three [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/). In addition to the V and [J gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2976/) (denoted VH and JH to distinguish them from the light-chain VL and JL), there is a third gene segment called the **diversity** or **D**H**gene segment**, which lies between the VH and JH gene segments. The process of recombination that generates a complete heavy-chain V region is shown in [Fig. 4.2](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A440/?report=objectonly" \t "object) (right panel), and occurs in two separate stages. In the first, a DH gene segment is joined to a JH gene segment; then a VH gene segment rearranges to DJH to make a complete VH-region exon. As with the light-chain genes, RNA splicing joins the assembled V-region sequence to the neighboring C-region gene.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27140/)

## 4-3. There are multiple different V-region gene segments

For simplicity, we have so far discussed the formation of a complete immunoglobulin V-region sequence as though there were only a single copy of each gene segment. In fact, there are multiple copies of all of the [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) in germline DNA. It is the random selection of just one gene segment of each type to assemble a V region that makes possible the great **diversity** of V regions among [**immunoglobulins**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/). The numbers of functional gene segments of each type in the human genome, as determined by gene cloning and sequencing, are shown in [Fig. 4.3](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A442/?report=objectonly" \t "object). Not all the gene segments discovered are functional, as a proportion have accumulated mutations that prevent them from encoding a functional protein. These are termed ‘pseudogenes.’ Because there are many V, D, and [J gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2976/) in germline DNA, no single one is essential. This reduces the evolutionary pressure on each gene segment to remain intact, and has resulted in a relatively large number of pseudogenes. Since some of these pseudogenes can undergo [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) just like a normal functional gene segment, a significant proportion of rearrangements will incorporate a pseudogene and thus be nonfunctional.

#### [Figure 4.3](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A442/?report=objectonly" \t "object)

The numbers of functional gene segments for the V regions of human heavy and light chains. These numbers are derived from exhaustive cloning and sequencing of DNA from one individual and exclude all pseudogenes (mutated and nonfunctional versions of a [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A442/?report=objectonly" \t "object)

The immunoglobulin [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) are organized into three clusters or genetic loci—the κ, λ, and heavy-chain loci. These are on different chromosomes and each is organized slightly differently, as shown in [Fig. 4.4](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A443/?report=objectonly" \t "object) for humans. At the λ light-chain locus, located on chromosome 22, a cluster of Vλ gene segments is followed by four sets of Jλ gene segments each linked to a single Cλ gene. In the κ light-chain locus, on chromosome 2, the cluster of Vκ gene segments is followed by a cluster of Jκ gene segments, and then by a single Cκ gene. The organization of the heavy-chain locus, on chromosome 14, resembles that of the κ locus, with separate clusters of VH, DH, and JH gene segments and of CH genes. The heavy-chain locus differs in one important way: instead of a single C-region, it contains a series of C regions arrayed one after the other, each of which corresponds to a different isotype. Generally, a cell expresses only one at a time, beginning with [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/). The expression of other [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/), such as [IgG](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2908/), can occur through [isotype switching](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2973/), as will be described in [Section 4-16](https://www.ncbi.nlm.nih.gov/books/n/imm/A469/#A473).

#### [Figure 4.4](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A443/?report=objectonly" \t "object)

The germline organization of the immunoglobulin heavy- and light-chain loci in the human genome. The genetic locus for the λ light chain (chromosome 22) has about 30 functional Vλ gene segments and four pairs of functional Jλ gene [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A443/?report=objectonly" \t "object)

The human [V gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3343/) can be grouped into families in which each member shares at least 80% DNA sequence identity with all others in the **family**. Both the heavy-chain and κ-chain V gene segments can be subdivided into seven such families, whereas there are eight families of Vλ gene segments. The families can be grouped into clans, made up of families that are more similar to each other than to families in other clans. Human VH gene segments fall into three such clans. All of the VH gene segments identified from amphibians, reptiles, and mammals also fall into the same three clans, suggesting that these clans existed in a common ancestor of these modern animal groups.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27140/)

## 4-4. Rearrangement of V, D, and J gene segments is guided by flanking DNA sequences

A system is required to ensure that DNA rearrangements take place at the correct locations relative to the V, D, or J gene segment coding regions. In addition, joins must be regulated such that a V gene segment joins to a D or J and not to another V. DNA rearrangements are in fact guided by conserved noncoding DNA sequences that are found adjacent to the points at which recombination takes place. These sequences consist of a conserved block of seven nucleotides—the **heptamer** 5′CACAGTG3′—which is always contiguous with the coding sequence, followed by a nonconserved region known as the spacer, which is either 12 or 23 nucleotides long. This is followed by a second conserved block of nine nucleotides—the **nonamer** 5′ACAAAAACC3′ ([Fig. 4.5](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A445/?report=objectonly" \t "object)). The spacer varies in sequence but its conserved length corresponds to one or two turns of the DNA double helix. This brings the heptamer and nonamer sequences to the same side of the DNA helix, where they can be bound by the complex of proteins that catalyzes recombination. The heptamer-spacer-nonamer is called a **recombination signal sequence** (**RSS**).

#### [Figure 4.5](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A445/?report=objectonly" \t "object)

Conserved heptamer and nonamer sequences flank the gene segments encoding the V regions of heavy (H) and light (λ and κ) chains. The spacer (white) between the heptamer (orange) and nonamer (purple) sequences is always either approximately [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A445/?report=objectonly" \t "object)

Recombination only occurs between [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) located on the same chromosome. It generally follows the rule that only a gene segment flanked by a RSS with a 12-base pair (bp) spacer can be joined to one flanked by a 23 bp spacer RSS. This is known as the [12/23 rule](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3191/). Thus, for the [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/), a DH gene segment can be joined to a JH gene segment and a VH gene segment to a DH gene segment, but VH gene segments cannot be joined to JH gene segments directly, as both VH and JH gene segments are flanked by 23 bp spacers and the DH gene segments have 12 bp spacers on both sides (see [Fig. 4.5](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A445/?report=objectonly" \t "object)).

It is now apparent, however, that, even though it violates the [12/23 rule](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3191/), direct joining of one D gene segment to another can occur in most species. In humans, D-D fusion is found in approximately 5% of antibodies and is the major mechanism accounting for the unusually long CDR3 loops found in some heavy chains. By creating extra-long CDR3s and unusual amino acid combinations, these D-D fusions add further to the **diversity** of the [antibody repertoire](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2577/).

The mechanism of DNA [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) is similar for the heavy- and light-chain loci, although only one joining event is needed to generate a light-chain gene whereas two are needed to generate a complete heavy-chain gene. The commonest mode of rearrangement ([Fig. 4.6](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A446/?report=objectonly" \t "object), left panels) involves the looping-out and deletion of the DNA between two [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/). This occurs when the coding sequences of the two gene segments are in the same orientation in the DNA. A second mode of recombination can occur between two gene segments that have opposite transcriptional orientations. This mode of recombination is less common, although such rearrangements account for about half of all Vκ to Jκ joins; the transcriptional orientation of half of the human Vκ gene segments is opposite to that of the Jκ gene segments. The mechanism of recombination is essentially the same, but the DNA that lies between the two gene segments meets a different fate ([Fig. 4.6](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A446/?report=objectonly" \t "object), right panels). When the RSSs in such cases are brought together and recombination takes place, the intervening DNA is not lost from the chromosome but is retained in an inverted orientation.

#### [Figure 4.6](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A446/?report=objectonly" \t "object)

V-region gene segments are joined by recombination. In every V-region recombination event, the signals flanking the gene segments are brought together to allow recombination to take place. For simplicity, the recombination of a light-chain gene is illustrated; [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A446/?report=objectonly" \t "object)

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27140/)

## 4-5. The reaction that recombines V, D, and J gene segments involves both lymphocyte-specific and ubiquitous DNA-modifying enzymes

The molecular mechanism of V-region DNA [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/), or [**V(D)J recombination**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3354/), is illustrated in [Fig. 4.7](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A448/?report=objectonly" \t "object). The 12 bp spaced and 23 bp spaced RSSs are brought together by interactions between proteins that specifically recognize the length of spacer and thus enforce the [12/23 rule](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3191/) for recombination. The DNA molecule is then broken in two places and rejoined in a different configuration. The ends of the heptamer sequences are joined precisely in a head-to-head fashion to form a [signal joint](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3244/) in a circular piece of extrachromosomal DNA, which is lost from the genome when the cell divides. The V and [J gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2976/), which remain on the chromosome, join to form what is called the [coding joint](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2710/). This junction is imprecise, and consequently generates much additional [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/) in the V-region sequence.

#### [Figure 4.7](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A448/?report=objectonly" \t "object)

Enzymatic steps in the rearrangement of immunoglobulin gene segments. Rearrangement begins with the binding of RAG-1, RAG-2, and high mobility group (HMG) proteins (not shown). These RAG-1:RAG-2 complexes (domes, colored green or purple for clarity although [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A448/?report=objectonly" \t "object)

The complex of enzymes that act in concert to effect somatic [V(D)J recombination](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3354/) is termed the **V**(**D**)**J recombinase**. The products of the two genes [RAG-1](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3196/) and ***RAG-2*** (**recombination-activating genes**) comprise the lymphoid-specific components of the recombinase. This pair of genes is only expressed in developing [lymphocytes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3021/) while they are engaged in assembling their [antigen receptors](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2590/), as is described in more detail in Chapter 7. They are essential for V(D)J recombination. Indeed, these genes, when expressed together, are sufficient to confer on nonlymphoid cells such as fibroblasts the capacity to rearrange exogenous segments of DNA that contain appropriate RSSs; this is how *RAG-1* and *RAG-2* were initially discovered.

Although the RAG proteins are required for [V(D)J recombination](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3354/), they are not the only enzymes in the recombinase. The remaining enzymes are ubiquitously expressed DNA-modifying proteins that are involved in double-stranded DNA repair, DNA bending, or the modification of the ends of the broken DNA strands. They include the enzyme DNA ligase IV, the enzyme DNA-dependent protein kinase (DNA-PK), and Ku, a well-known autoantigen, which is a heterodimer (Ku 70:Ku 80) that associates tightly with DNA-PK.

[V(D)J recombination](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3354/) is a multistep enzymatic process in which the first reaction is an endonucleolytic cleavage requiring the coordinated activity of both RAG proteins. Initially, two RAG protein complexes, each containing [RAG-1](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3196/), RAG-2, and high-mobility group proteins, recognize and align the two RSSs that are guiding the join (see [Fig. 4.7](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A448/?report=objectonly" \t "object)). RAG-1 is thought to specifically recognize the nonamer of the RSS. At this stage, the [12/23 rule](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3191/) is established through mechanisms that are still poorly understood. The endonuclease activity of the RAG protein complexes then makes two single-strand DNA breaks at sites just 5′ of each bound RSS, leaving a free 3′-OH group at the end of each coding segment. This 3′-OH group then hydrolyzes the phosphodiester bond on the other strand, sealing the end of the double-stranded DNA to create a DNA ‘hairpin’ out of the gene segment coding region. This process simultaneously creates a flush double-stranded break at the ends of the two heptamer signal sequences. The DNA ends do not float apart, however, but are held tightly in a complex by the RAG proteins and other associated DNA repair enzymes until the join is completed. The two RSSs are precisely joined to form the [signal joint](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3244/). Coding joint formation is more complex. First, the DNA hairpin is nicked open by a single-stranded break, again by the RAG proteins. The nicking can happen at various points along the hairpin, which leads to sequence [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/) in the eventual joint. The DNA repair enzymes in the complex then modify the opened hairpins by removing nucleotides (by exonuclease activity) and by randomly adding nucleotides (by [terminal deoxynucleotidyl transferase](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3293/), TdT). It is not known if addition and deletion of nucleotides at the ends of coding regions occurs simultaneously or in a defined order. Finally, ligases such as DNA ligase IV join the processed ends together to generate a continuous double-stranded DNA, thus reconstituting a chromosome that includes the rearranged gene. This enzymatic process seems to create **diversity** in the joint between [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/), while ensuring that the RSS ends are ligated without modification, and that unintended genetic damage such as a chromosomal break is avoided.

The recombination mechanism controlled by the RAG proteins shares many interesting features with the mechanism by which retroviral integrases catalyze the insertion of retroviral DNA into the genome, and also with the transposition mechanism used by transposons (mobile genetic elements that encode their own transposase, allowing them to excise and reinsert themselves in the genome). Even the structure of the *RAG* genes themselves, which lie close together in the chromosome and lack the usual mammalian introns, is reminiscent of a transposon. Indeed, it has recently been shown that the RAG complex can act as a transposase *in vitro*. These features have provoked speculation that the RAG complex originated as a transposase whose function was adapted by vertebrates to allow V gene segment recombination, thus leading to the advent of the vertebrate adaptive [immune system](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2921/). Consistent with this idea, no genes homologous to the *RAG* genes have been found in nonvertebrates.

The *in vivo* roles of the enzymes involved in [V(D)J recombination](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3354/) have been established through natural or artificially induced mutations. Mice in which either of the *RAG* genes is knocked out suffer a complete block in lymphocyte development at the gene [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) stage. Mice lacking TdT do not add extra nucleotides to the joints between [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/). A mutation that was discovered some time ago results in mice that make only trivial amounts of [**immunoglobulins**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) or T-cell receptors. Such mice suffer from a **s**evere **c**ombined **i**mmune **d**eficiency—hence the name ***scid*** for this mutation. These mice have subsequently been found to have a mutation in the enzyme DNA-PK that prevents the efficient rejoining of DNA at gene segment junctions. Mutations of other proteins that are involved in DNA joining also give the *scid* phenotype. (Omenn Syndrome, in *Case Studies in*[*Immunology*](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2945/), see Preface for details)

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27140/)

## 4-6. The diversity of the immunoglobulin repertoire is generated by four main processes

Antibody **diversity** is generated in four main ways. Two of these are consequences of the recombination process just discussed (see [Sections 4-4](https://www.ncbi.nlm.nih.gov/books/NBK27140/#A444) and [4-5](https://www.ncbi.nlm.nih.gov/books/NBK27140/#A447)) which creates complete immunoglobulin V-region exons during early B-cell development. The third is due to the different possible combinations of a heavy and a [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/) in the complete immunoglobulin molecule. The fourth is a mutational process that occurs in mature B cells, acting only on rearranged DNA encoding the V regions.

The gene [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) that combines two or three [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) to form a complete V-region exon generates **diversity** in two ways. First, there are multiple different copies of each type of gene segment, and different combinations of gene segments can be used in different rearrangement events. This [combinatorial **diversity**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2714/) is responsible for a substantial part of the **diversity** of the heavy- and light-chain V regions. Second, junctional **diversity** is introduced at the joints between the different gene segments as a result of addition and subtraction of nucleotides by the recombination process. A third source of **diversity** is also combinatorial, arising from the many possible different combinations of heavy- and light-chain V regions that pair to form the [antigen-binding site](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2582/) in the immunoglobulin molecule. The two means of generating combinatorial **diversity** alone could give rise, in theory, to approximately 3.5 × 106 different [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) molecules (see [Section 4-7](https://www.ncbi.nlm.nih.gov/books/NBK27140/#A450)). Coupled with junctional **diversity**, it is estimated that as many as 1011 different receptors could make up the repertoire of receptors expressed by naive B cells. Finally, somatic hypermutation introduces point mutations into the rearranged V-region genes of activated B cells, creating further **diversity** that can be [selected](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3228/) for enhanced binding to antigen. We will discuss these mechanisms at greater length in the following sections.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27140/)

## 4-7. The multiple inherited gene segments are used in different combinations

There are multiple copies of the V, D, and [J gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2976/), each of which is capable of contributing to an immunoglobulin V region. Many different V regions can therefore be made by selecting different combinations of these segments. For human κ light chains, there are approximately 40 functional Vκ gene segments and five Jκ gene segments, and thus potentially 200 different Vκ regions. For λ light chains there are approximately 30 functional Vλ gene segments and four Jλ gene segments, yielding 120 possible Vλ regions. So, in all, 320 different light chains can be made as a result of combining different light-chain gene segments. For the heavy chains of humans, there are 65 functional VH gene segments, approximately 27 DH gene segments, and 6 JH gene segments, and thus around 11,000 different possible VH regions (65 × 27 × 6 ≈ 11,000). During B-cell development, [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) at the heavy-chain gene locus to produce any one of the possible heavy chains is followed by several rounds of cell division before light-chain gene rearrangement takes place. The particular combination of gene segments used to produce a [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/) does not appear to restrict the choice of gene segments that can be recombined to assemble a light-chain [variable region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3350/). Thus, in theory any one possible heavy chain can be produced together with any one possible [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/) in a single [B cell](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2618/). As both the heavy- and the light-chain V regions contribute to [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/), each of the 320 different light chains could be combined with each of the approximately 11,000 heavy chains to give around 3.5 × 106 different antibody specificities. This theoretical estimate of [combinatorial **diversity**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2714/) is based on the number of germline [V gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3343/) contributing to functional antibodies (see [Fig. 4.3](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A442/?report=objectonly" \t "object)); the total number of V gene segments is larger, but the additional gene segments are pseudogenes and do not appear in expressed immunoglobulin molecules.

In practice, [combinatorial **diversity**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2714/) is likely to be less than one might expect from the theoretical calculations above. One reason for this is that not all [V gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3343/) are used at the same frequency; some are common in antibodies, while others are found only rarely. It is also clear that not every [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/) can pair with every [light chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3006/); certain combinations of VH and VL regions result in failure to assemble a stable immunoglobulin molecule. Cells that have heavy and light chains that cannot pair may continue to undergo light-chain gene [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) until a suitable light chain is produced, or may be eliminated, but in both cases a heavy- and light-chain combination that does not pair is lost from the repertoire. Nevertheless, it is thought that most heavy and light chains can pair with each other, and that this type of combinatorial **diversity** has a major role in the formation of an [immunoglobulin repertoire](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2939/) with a wide range of specificities. In addition, two further processes add greatly to repertoire **diversity**—imprecise joining of V, D, and [J gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2976/) and somatic hypermutation.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27140/)

## 4-8. Variable addition and subtraction of nucleotides at the junctions between gene segments contributes to diversity in the third hypervariable region

Of the three [hypervariable](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2894/) loops in the protein chains of [**immunoglobulins**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/), two are encoded within the V gene segment DNA. The third (HV3 or CDR3, see [Fig. 3.6](https://www.ncbi.nlm.nih.gov/books/n/imm/A335/figure/A337/?report=objectonly" \t "object)) falls at the joint between the V gene segment and the J gene segment, and in the [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/) is partially encoded by the D gene segment. In both heavy and light chains, the **diversity** of CDR3 is significantly increased by the addition and deletion of nucleotides at two steps in the formation of the junctions between [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/). The added nucleotides are known as [P-nucleotides](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3150/) and [N-nucleotides](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3108/) and their addition is illustrated in [Fig. 4.8](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A452/?report=objectonly" \t "object).

#### [Figure 4.8](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A452/?report=objectonly" \t "object)

The introduction of P- and N-nucleotides at the joints between gene segments during immunoglobulin gene rearrangement. The process is illustrated for a DH to JH rearrangement; however, the same steps occur in VH to DH and in VL to JL rearrangements. After [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A452/?report=objectonly" \t "object)

[P-nucleotides](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3150/) are so called because they make up palindromic sequences added to the ends of the [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/). After the formation of the DNA hairpins as described in [Section 4-5](https://www.ncbi.nlm.nih.gov/books/NBK27140/#A447), the RAG protein complex catalyzes a single-stranded cleavage at a random point within the coding sequence but near the original point at which the hairpin was first formed. When this cleavage occurs at a different point from the initial break, a single-stranded tail is formed from a few nucleotides of the coding sequence plus the complementary nucleotides from the other DNA strand (see [Fig. 4.8](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A452/?report=objectonly" \t "object)). In most light-chain gene rearrangements, DNA repair enzymes then fill in complementary nucleotides on the single-stranded tails which would leave short palindromic sequences at the joint, if the ends are rejoined without any further exonuclease activity (see below). In heavy-chain gene rearrangements and in some human light-chain genes, however, [N-nucleotides](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3108/) are first added by a quite different mechanism.

[N-nucleotides](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3108/) are so called because they are nontemplate-encoded. They are added by the enzyme [terminal deoxynucleotidyl transferase](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3293/) (TdT) to single-stranded ends of the coding DNA after hairpin cleavage. After the addition of up to 20 nucleotides by this enzyme, the two single-stranded stretches at the ends of the [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) form base pairs over a short region. Repair enzymes then trim off any nonmatching bases, synthesize complementary bases to fill in the remaining single-stranded DNA, and ligate it to the [P-nucleotides](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3150/) (see [Fig. 4.8](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A452/?report=objectonly" \t "object)). N-nucleotides are found especially in the V-D and D-J junctions of the assembled heavy-chain gene; they are less common in light-chain genes because TdT is expressed for only a short period in B-cell development, during the assembly of the heavy-chain gene, which occurs before that of the light-chain gene.

Nucleotides can also be deleted at gene segment junctions. This is accomplished by as yet unidentified exonucleases. Thus, the length of heavy-chain CDR3 can be even shorter than the smallest D segment. In some instances it is difficult, if not impossible, to recognize the D segment that contributed to CDR3 formation because of the excision of most of its nucleotides. Deletions may also erase the traces of P-nucleotide palindromes introduced at the time of hairpin opening. For this reason, many completed V(D)J joins do not show obvious evidence of [P-nucleotides](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3150/).

As the total number of nucleotides added by these processes is random, the added nucleotides often disrupt the reading frame of the coding sequence beyond the joint. Such frameshifts will lead to a nonfunctional protein, and DNA rearrangements leading to such disruptions are known as [nonproductive rearrangements](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3110/). As roughly two in every three rearrangements will be nonproductive, many B cells never succeed in producing functional immunoglobulin molecules, and junctional **diversity** is therefore achieved only at the expense of considerable wastage. We will discuss this further in Chapter 7.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27140/)

## 4-9. Rearranged V genes are further diversified by somatic hypermutation

The mechanisms for generating **diversity** described so far all take place during the [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) of [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) in the initial development of B cells in the central lymphoid organs. There is an additional mechanism that generates **diversity** throughout the V region and that operates on B cells in peripheral lymphoid organs after functional immunoglobulin genes have been assembled. This process, known as [somatic hypermutation,](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3256/) introduces point mutations into the V regions of the rearranged heavy- and light-chain genes at a very high rate, giving rise to mutant B-cell receptors on the surface of the B cells ([Fig. 4.9](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A454/?report=objectonly" \t "object)). Some of the mutant immunoglobulin molecules bind [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) better than the original B-cell receptors, and B cells expressing them are preferentially [selected](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3228/) to mature into [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/)-secreting cells. This gives rise to a phenomenon called affinity maturation of the antibody population, which we will discuss in more detail in Chapters 9 and 10.

#### [Figure 4.9](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A454/?report=objectonly" \t "object)

Somatic hypermutation introduces variation into the rearranged immunoglobulin variable region that is subject to negative and positive selection to yield improved antigen binding. In some circumstances it is possible to follow the process of somatic hypermutation [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A454/?report=objectonly" \t "object)

Somatic hypermutation occurs when B cells respond to [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) along with signals from activated [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/). The immunoglobulin C-region gene, and other genes expressed in the [B cell](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2618/), are not affected, whereas the rearranged VH and VL genes are mutated even if they are [nonproductive rearrangements](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3110/) and are not expressed. The pattern of nucleotide base changes in nonproductive V-region genes illustrates the result of somatic hypermutation without selection for enhanced binding to antigen. The base changes are distributed throughout the V region, but not completely randomly: there are certain ‘hotspots’ of mutation that indicate a preference for characteristic short motifs of four to five nucleotides, and perhaps also certain ill-defined secondary structural features. The pattern of base changes in the V regions of expressed immunoglobulin genes is different. Mutations that alter amino acid sequences in the conserved [framework regions](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2825/) will tend to disrupt basic [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) structure and are [selected](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3228/) against. In contrast, the result of selection for enhanced binding to antigen is that base changes that alter amino acid sequences, and thus protein structure, tend to be clustered in the CDRs, whereas silent mutations that preserve amino acid sequence and do not alter protein structure are scattered throughout the V region.

The mechanism of somatic hypermutation is poorly defined, but there have been several new discoveries that shed some light. It is known that mutation requires the presence of [enhancers](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2796/), DNA sequences that enhance the trans-cription of immunoglobulin genes in B cells, as well as a transcriptional promoter. The promoter, and the sequences that are the target of mutation need not derive from immunoglobulin V genes, however. The **generation** of new mutations in V regions in mutating B cells has recently been shown to be accompanied by double-stranded breaks in the DNA which are thought to then be repaired in an error-prone way. In addition, it has recently been discovered that deficiency in an RNA editing enzyme called Activation Induced Cytidine Deaminase, blocks the accumulation of somatic hypermutations. The mechanism by which this enzyme contributes to hypermutation is unknown. Interestingly, deficiency of this enzyme also abrogates the [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) of C-region genes that underlies the immunoglobulin class switching seen in activated B cells (see [Section 4-16](https://www.ncbi.nlm.nih.gov/books/n/imm/A469/#A473)).

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27140/)

## 4-10. In some species most immunoglobulin gene diversification occurs after gene rearrangement

As we have seen in the preceding sections, a proportion of the immunoglobulin **diversity** in an adult human derives from the existence of a variety of germline [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/), and a proportion from somatic alterations acquired during the lifetime of the individual. This particular combination of heritable and acquired components of **diversity** operates in several mammalian immune systems, including those of humans and mice. Other species achieve a mix of inherited and acquired **diversity** by different means. Overall, it would appear that there is strong selective pressure to generate sufficient **diversity** in the [immune system](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2921/) to protect the organism from common pathogens, and several different mechanisms have evolved toward this end.

In birds, rabbits, cows, pigs, sheep, and horses there is little or no [germline **diversity**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2843/) in the V, D, and [J gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2976/) that are rearranged to form the genes for the initial B-cell receptors, and the rearranged V-region sequences are identical or similar in most immature B cells. These B cells then migrate to specialized microenvironments, the best known of which is the [bursa of Fabricius](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2647/) in chickens. Here, B cells proliferate rapidly, and their rearranged immunoglobulin genes undergo further diversification. In birds and rabbits this occurs by a process that includes [gene conversion](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2834/), in which an upstream V segment pseudogene exchanges short sequences with the expressed rearranged V-region gene ([Fig. 4.10](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A456/?report=objectonly" \t "object)). In sheep and cows, diversification is the result of [somatic hypermutation,](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3256/) which occurs in an organ known as the ileal Peyer's patch. Somatic hypermutation probably also contributes to immunoglobulin diversification in birds and rabbits.

#### [Figure 4.10](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A456/?report=objectonly" \t "object)

The diversification of chicken **immunoglobulins** occurs through gene conversion. In chickens, all B cells express the same surface immunoglobulin (slg) initially; there is only one active V, D, and J gene segment for the chicken heavy-chain gene and one [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27140/figure/A456/?report=objectonly" \t "object)

# T-cell receptor gene rearrangement

The mechanism by which B-cell [antigen receptors](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2590/) are generated is such a powerful means of creating diversity that it is not surprising that the antigen receptors of [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) bear structural resemblances to [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) and are generated by the same mechanism. In this part of the chapter we describe the organization of the [**T-cell** **receptor**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) loci and the generation of the **genes** for the individual **T-cell** **receptor** chains.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27145/)

## 4-11. The T-cell receptor loci comprise sets of gene segments and are rearranged by the same enzymes as the immunoglobulin loci

Like immunoglobulin heavy and light chains, [**T-cell** **receptor**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) α and β chains each consist of a variable (V) amino-terminal region and a constant (C) region (see [Section 3-10](https://www.ncbi.nlm.nih.gov/books/n/imm/A346/#A347)). The organization of the TCRα and TCRβ loci is shown in [Fig. 4.11](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A460/?report=objectonly" \t "object). The organization of the [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) is broadly homologous to that of the immunoglobulin gene segments (see [Sections 4-2](https://www.ncbi.nlm.nih.gov/books/n/imm/A436/#A439) and [4-3](https://www.ncbi.nlm.nih.gov/books/n/imm/A436/#A441)). The TCRα locus, like those for the immunoglobulin light chains, contains V and [J gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2976/) (Vα and Jα). The TCRβ locus, like that for the immunoglobulin heavy-chain, contains [D gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2757/) in addition to Vβ and Jβ gene segments.

#### [Figure 4.11](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A460/?report=objectonly" \t "object)

The germline organization of the human **T-cell** **receptor** α and β loci. The arrangement of the gene segments resembles that at the immunoglobulin loci, with separate variable (V), diversity (D), joining (J) gene segments, and constant (C) [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A460/?report=objectonly" \t "object)

The [**T-cell** **receptor**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) rearrange during **T-cell** development to form complete V-domain exons ([Fig. 4.12](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A461/?report=objectonly" \t "object)). **T-cell** **receptor** gene [**rearrangement**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) takes place in the [thymus](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3304/); the order and regulation of the rearrangements will be dealt with in detail in Chapter 7. Essentially, however, the mechanics of gene **rearrangement** are similar for B and [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/). The **T-cell** **receptor** gene segments are flanked by heptamer and nonamer recombination signal sequences (RSSs) that are homologous to those flanking immunoglobulin gene segments (see [Section 4-4](https://www.ncbi.nlm.nih.gov/books/n/imm/A436/#A444) and [Fig. 4.5](https://www.ncbi.nlm.nih.gov/books/n/imm/A436/figure/A445/?report=objectonly" \t "object)) and are recognized by the same enzymes. All known defects in **genes** that control [V(D)J recombination](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3354/) affect T cells and B cells equally, and animals with these genetic defects lack functional [lymphocytes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3021/) altogether (see [Section 4-5](https://www.ncbi.nlm.nih.gov/books/n/imm/A436/#A447)). A further shared feature of immunoglobulin and **T-cell** **receptor** gene **rearrangement** is the presence of P- and [N-nucleotides](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3108/) in the junctions between the V, D, and [J gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2976/) of the rearranged TCRβ gene. In T cells, P- and N-nucleotides are also added between the V and J gene segments of all rearranged TCRα **genes**, whereas only about half the V-J joints in immunoglobulin light-chain **genes** are modified by N-nucleotide addition and these are often left without any [P-nucleotides](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3150/) as well (see [Section 4-8](https://www.ncbi.nlm.nih.gov/books/n/imm/A436/#A451) and [Fig. 4.13](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A462/?report=objectonly" \t "object)).

#### [Figure 4.12](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A461/?report=objectonly" \t "object)

**T-cell** **receptor** α- and β-chain gene **rearrangement** and expression. The TCRα- and β-chain **genes** are composed of discrete segments that are joined by somatic recombination during development of the T cell. Functional α- [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A461/?report=objectonly" \t "object)

#### [Figure 4.13](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A462/?report=objectonly" \t "object)

The numbers of human **T-cell** **receptor** gene segments and the sources of **T-cell** **receptor** diversity compared with those of immunoglobulins. Note that only about half of human κ chains contain N-nucleotides. Somatic hypermutation as a source of diversity [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A462/?report=objectonly" \t "object)

The main differences between the immunoglobulin **genes** and those encoding **T-cell** receptors reflect the fact that all the effector functions of B cells depend upon secreted antibodies whose different heavy-chain C-region [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/) trigger distinct effector mechanisms. The effector functions of [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/), in contrast, depend upon cell-cell contact and are not mediated directly by the [**T-cell** **receptor**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/), which serves only for [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) recognition. Thus, the C regions of the TCRα and TCRβ loci are much simpler than those of the immunoglobulin heavy-chain locus. There is only one Cα gene and, although there are two Cβ **genes**, they are very closely homologous and there is no known functional distinction between their products. The **T-cell** **receptor** C-region **genes** encode only transmembrane polypeptides.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27145/)

## 4-12. T-cell receptors concentrate diversity in the third hypervariable region

The extent and pattern of [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/) in **T-cell** receptors and [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) reflect the distinct nature of their ligands. Whereas the [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-binding sites of immunoglobulins must conform to the surfaces of an almost infinite variety of different antigens, and thus come in a wide variety of shapes and chemical properties, the ligand for the [**T-cell** **receptor**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) is always a peptide bound to an MHC molecule. The antigen-recognition sites of **T-cell** receptors would therefore be predicted to have a less variable shape, with most of the variability focused on the bound antigenic peptide occupying the center of the surface in contact with the **receptor**.

In spite of differences in the sites of [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/), the three-dimensional structure of the [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-recognition site of a [**T-cell** **receptor**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) looks much like that of an [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) molecule (see [Sections 3-11](https://www.ncbi.nlm.nih.gov/books/n/imm/A346/#A351) and [3-7](https://www.ncbi.nlm.nih.gov/books/n/imm/A335/#A339), respectively). In an antibody, the center of the [antigen-binding site](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2582/) is formed by the CDR3s of the heavy and light chains. The structurally equivalent third [hypervariable](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2894/) loops (CDR3s) of the **T-cell** **receptor** α and β chains, to which the D and [J gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2976/) contribute, also form the center of the antigen-binding site of a **T-cell** **receptor**; the periphery of the site consists of the equivalent of the CDR1 and CDR2 loops, which are encoded within the germline [V gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3343/) for the α and β chains.

[**T-cell** **receptor**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) loci have roughly the same number of [V gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3343/) as do the immunoglobulin loci, but only B cells diversify rearranged V-region **genes** by somatic hypermutation. Thus, diversity in the CDR1 and CDR2 loops that comprise the periphery of the [antigen-binding site](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2582/) will be far greater among [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) molecules than among **T-cell** receptors. This is in keeping with the fact that the CDR1 and CDR2 loops of a **T-cell** **receptor** will mainly contact the relatively less variable MHC component of the ligand rather than the highly variable peptide component ([Fig. 4.14](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A464/?report=objectonly" \t "object)).

#### [Figure 4.14](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A464/?report=objectonly" \t "object)

The most variable parts of the **T-cell** **receptor** interact with the peptide bound to an MHC molecule. The positions of the CDR loops of a **T-cell** **receptor** are shown as the colored tubes in this figure superimposing onto the MHC:peptide complex. The CDR1 loops [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A464/?report=objectonly" \t "object)

The structural diversity of **T-cell** receptors is mainly attributable to combinatorial and junctional diversity generated during the process of gene [**rearrangement**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/). It can be seen from [Fig. 4.13](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A462/?report=objectonly" \t "object) that the [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/) in [**T-cell** **receptor**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) chains is focused on the junctional region encoded by V, D, and [J gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2976/) and modified by P- and [N-nucleotides](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3108/). The TCRα locus contains many more J gene segments than either of the immunoglobulin light-chain loci: in humans, 61 Jα gene segments are distributed over about 80 kb of DNA, whereas immunoglobulin light-chain loci have only five J gene segments at most (see [Fig. 4.13](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A462/?report=objectonly" \t "object)). Because the TCRα locus has so many J gene segments, the variability generated in this region is even greater for **T-cell** receptors than for [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/). This region encodes the CDR3 loops in immunoglobulins and **T-cell** receptors that form the center of the [antigen-binding site](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2582/). Thus, the center of the **T-cell** **receptor** will be highly variable, whereas the periphery will be subject to relatively little variation.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27145/)

## 4-13. γ:δ T-cell receptors are also generated by gene rearrangement

A minority of [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) bear **T-cell** receptors composed of γ and δ chains (see [Section 3-19](https://www.ncbi.nlm.nih.gov/books/n/imm/A346/#A374)). The organization of the TCRγ and TCRδ loci ([Fig. 4.15](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A466/?report=objectonly" \t "object)) resembles that of the TCRα and TCRβ loci, although there are important differences. The cluster of [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) encoding the δ chain is found entirely within the TCRα locus, between the Vα and the Jα gene segments. Because all Vα gene segments are oriented such that [**rearrangement**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) will delete the intervening DNA, any **rearrangement** at the α locus results in the loss of the δ locus. There are substantially fewer [V gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3343/) at the TCRγ and TCRδ loci than at either the TCRα or TCRβ loci or at any of the immunoglobulin loci. Increased junctional [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/) in the δ chains may compensate for the small number of V gene segments and has the effect of focusing almost all of the variability in the γ:δ **receptor** in the junctional region. As we have seen, the amino acids encoded by the junctional regions lie at the center of the [**T-cell** **receptor**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) binding site.

#### [Figure 4.15](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A466/?report=objectonly" \t "object)

The organization of the **T-cell** **receptor** γ- and δ-chain loci in humans. The TCRγ and TCRδ loci, like the TCRα and TCRβ loci, have discrete V, D, and J gene segments, and C **genes**. Uniquely, the locus encoding the [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27145/figure/A466/?report=objectonly" \t "object)

[T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) bearing γ:δ receptors are a distinct lineage of T cells whose functions are at present unknown. The ligands for these receptors are also largely unknown (see [Section 3-19](https://www.ncbi.nlm.nih.gov/books/n/imm/A346/#A374)). Some γ:δ **T-cell** receptors appear to be able to recognize [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) directly, much as antibodies do, without the requirement for presentation by an MHC molecule or processing of the antigen. Detailed analysis of the rearranged V regions of γ:δ **T-cell** receptors shows that they resemble the V regions of [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) molecules more than they resemble the V regions of α:β **T-cell** receptors.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27145/)

## 4-14. Somatic hypermutation does not generate diversity in T-cell receptors

When we discussed the generation of [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) diversity in [Section 4-9](https://www.ncbi.nlm.nih.gov/books/n/imm/A436/#A453), we saw that somatic hypermutation increases the diversity of all three complementarity-determining regions of both immunoglobulin chains. Somatic hypermutation does not occur in [**T-cell** **receptor**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) **genes**, so that [variability](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3347/) of the CDR1 and CDR2 regions is limited to that of the germline [V gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3343/). All the diversity in **T-cell** receptors is generated during [**rearrangement**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/) and is consequently focused on the CDR3 regions.

Why **T-cell** and B-cell receptors differ in their abilities to undergo somatic hypermutation is not clear, but several explanations can be suggested on the basis of the functional differences between T and B cells. Because the central role of [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) is to stimulate both humoral and cellular immune responses, it is crucially important that T cells do not react with self proteins. T cells that recognize [self antigens](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3230/) are rigorously purged during development (see Chapter 7) and the absence of somatic hypermutation helps to ensure that somatic mutants recognizing self proteins do not arise later in the course of immune responses. This constraint does not apply with the same force to B-cell receptors, as B cells usually require **T-cell** help to secrete antibodies. A [B cell](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2618/) whose **receptor** mutates to become self reactive would, under normal circumstances, fail to make [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) for lack of self-reactive T cells to provide this help (see Chapter 9).

A further argument is that [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) already interact with a self component, namely the MHC molecule that makes up the major part of the ligand for the **receptor**, and thus might be unusually prone to developing self-recognition capability through somatic hypermutation. In this case, the converse argument can also be made: because **T-cell** receptors must be able to recognize self [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) as part of their ligand, it is important to avoid somatic mutation that might result in the loss of recognition and the consequent loss of any ability to respond. However, the strongest argument for this difference between [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) and **T-cell** receptors is the simple one that somatic hypermutation is an adaptive specialization for B cells alone, because they must make very high-affinity antibodies to [capture](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2662/) toxin molecules in the extracellular fluids. We will see in Chapter 10 that they do this through somatic hypermutation followed by selection for [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) binding.

# Structural variation in immunoglobulin constant regions

So far we have focused on the **structural** **variation** inherent in the assembly of the V **regions** of the [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) molecule and [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/). We have seen how this **variation** creates a diverse repertoire of [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-specificities, and we have also considered how these variable **regions** are attached to **constant** **regions** in the monovalent heterodimeric T-cell receptor, and the Y-shaped four-chain structure of the divalent **immunoglobulin** molecule. However, we have discussed only the general **structural** features of the **immunoglobulin** [C region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2649/) as illustrated by [IgG](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2908/), the most abundant type of antibody in plasma (see [Section 3-1](https://www.ncbi.nlm.nih.gov/books/n/imm/A323/#A325)). Immunoglobulins can be made in several different forms, or [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/), and we now consider how this **structural** **variation** is generated by linking different heavy-chain **constant** **regions** to the same [variable region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3350/). The CH **regions**, which determine the class or isotype of the antibody and thus its effector functions, are encoded in separate genes located downstream of the V genes at the heavy-chain locus. Initially only the first of these genes, the Cμ gene, is expressed in conjunction with an assembled V gene. However, during the course of an antibody response activated B cells can switch to the expression of a different downstream CH gene by a process of [somatic recombination](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3257/) known as [isotype switching](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2973/). In this part of the chapter we consider the **structural** features that distinguish the CH **regions** of antibodies of the five major isotypes and confer on them their specialized functional properties as well as the mechanism of isotype switching. We also look at how alternative mRNA splicing allows the production of both membrane-bound and secreted forms of each isotype, and the simultaneous production of surface [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/) and [IgD](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2906/) in mature but naive B cells.

The use of [isotype switching](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2973/) and alternative mRNA splicing to generate **structural** and functional **variation** is unique to the **immunoglobulin** heavy-chain locus and does not occur in [T-cell receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3289/) genes. This reflects the fact that [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) act as soluble molecules that must both bind [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) and recruit a variety of other effector cells and molecules to deal with it appropriately, whereas the T-cell receptor functions only as a membrane-bound receptor to activate an appropriate cellular [immune response](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2918/).

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27106/)

## 4-15. The immunoglobulin heavy-chain isotypes are distinguished by the structure of their constant regions

The five main [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/) of **immunoglobulin** are [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/), [IgD](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2906/), [IgG](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2908/), [IgE](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2907/), and [IgA](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2905/). In humans, IgG antibodies can be further subdivided into four subclasses (IgG1, IgG2, IgG3, and IgG4), whereas IgA antibodies are found as two subclasses (IgA1 and IgA2). The IgG isotypes in humans are named in order of their abundance in serum, with IgG1 being the most abundant. The heavy chains that define these isotypes are designated by the lower-case Greek letters μ, δ, γ, ε, and α, as shown in [Fig. 4.16](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A471/?report=objectonly" \t "object), which also lists the major physical properties of the different human isotypes. IgM forms pentamers in serum, which accounts for its high molecular weight. Secreted IgA can occur as either a monomer or as a dimer.

#### [Figure 4.16](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A471/?report=objectonly" \t "object)

The properties of the human **immunoglobulin** isotypes. IgM is so called because of its size: although monomeric IgM is only 190 kDa, it normally forms pentamers, known as macroglobulin (hence the M), of very large molecular weight (see Fig. 4.23). IgA dimerizes [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A471/?report=objectonly" \t "object)

Sequence differences between **immunoglobulin** heavy chains cause the various [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/) to differ in several characteristic respects. These include the number and location of interchain disulfide bonds, the number of attached oligosaccharide moieties, the number of [C domains](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2648/), and the length of the [hinge region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2874/) ([Fig. 4.17](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A472/?report=objectonly" \t "object)). [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/) and [IgE](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2907/) heavy chains contain an extra C domain that replaces the hinge region found in γ, δ, and α chains. The absence of the hinge region does not imply that IgM and IgE molecules lack flexibility; electron micrographs of IgM molecules binding to ligands show that the Fab arms can bend relative to the Fc portion. However, such a difference in structure may have functional consequences that are not yet characterized. Different isotypes and subtypes also differ in their ability to engage various effector functions, as will be described in [Section 4-18](https://www.ncbi.nlm.nih.gov/books/NBK27106/#A479).

#### [Figure 4.17](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A472/?report=objectonly" \t "object)

The **structural** organization of the main human **immunoglobulin** isotype monomers. Both IgM and IgE lack a hinge region but each contains an extra heavy-chain domain. Note the differences in the numbers and locations of the disulfide bonds (black lines) linking [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A472/?report=objectonly" \t "object)

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27106/)

## 4-16. The same VH exon can associate with different CH genes in the course of an immune response

The V-region exons expressed by any given [B cell](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2618/) are determined during its early differentiation in the [bone marrow](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2640/) and, although they may subsequently be modified by [somatic hypermutation,](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3256/) no further [V(D)J recombination](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3354/) occurs. All the progeny of that B cell will therefore express the same assembled V genes. By contrast, several different C-region genes can be expressed in the B cell's progeny as the cells mature and proliferate in the course of an [immune response](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2918/). Every B cell begins by expressing [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/) as its B-cell receptor, and the first [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) produced in an immune response is always IgM. Later in the immune response, however, the same assembled V region may be expressed in [IgG](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2908/), [IgA](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2905/), or [IgE](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2907/) antibodies. This change is known as [isotype switching](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2973/). It is stimulated in the course of an immune response by external signals such as cytokines released by [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) or mitogenic signals delivered by pathogens, as we will discuss further in Chapter 9. Here we are concerned with the molecular basis of the isotype switch.

The **immunoglobulin** CH genes form a large cluster spanning about 200 kb to the 3′ side of the JH [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) ([Fig. 4.18](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A474/?report=objectonly" \t "object)). Each CH gene is split into several exons (not shown in the figure), each corresponding to an individual **immunoglobulin** domain in the folded [C region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2649/). The gene encoding the μ C region lies closest to the JH gene segments, and therefore closest to the assembled V-region exon after DNA [rearrangement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3199/). A complete μ heavy-chain transcript is produced from the newly rearranged gene. Any JH gene segments remaining between the assembled V gene and the Cμ gene are removed during RNA processing to generate the mature mRNA. μ heavy chains are therefore the first to be expressed and [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/) is the first **immunoglobulin** isotype to be expressed during B-cell development.

#### [Figure 4.18](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A474/?report=objectonly" \t "object)

The organization of the **immunoglobulin** heavy-chain C-region genes in mice and humans (not to scale). In humans, the cluster shows evidence of evolutionary duplication of a unit consisting of two γ genes, an ε gene and an α gene. [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A474/?report=objectonly" \t "object)

Immediately 3′ to the μ gene lies the δ gene, which encodes the [C region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2649/) of the [IgD](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2906/) [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/). IgD is coexpressed with [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/) on the surface of almost all mature B cells, although this isotype is secreted in only small amounts and its function is unknown. Indeed, mice lacking the Cδ exons seem to have essentially normal immune systems. B cells expressing IgM and IgD have not undergone [isotype switching](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2973/), which, as we will see, entails an irreversible change in the DNA. Instead, these cells produce a long primary transcript that is differentially cleaved and spliced to yield one of two distinct mRNA molecules. In one of these, the VDJ exon is linked to the Cμ exons to encode a μ heavy chain, and in the other the VDJ exon is linked to the Cδ exons to encode a δ heavy chain ([Fig. 4.19](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A475/?report=objectonly" \t "object)). The differential processing of the long mRNA transcript is developmentally regulated, with immature B cells making mostly the μ transcript and mature B cells making mostly the δ form along with some of the μ transcript.

#### [Figure 4.19](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A475/?report=objectonly" \t "object)

Co-expression of IgD and IgM is regulated by RNA processing. In mature B cells, transcription initiated at the VH promoter extends through both Cμ and Cδ exons. This long primary transcript is then processed by cleavage and polyadenylation [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A475/?report=objectonly" \t "object)

Switching to other [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/) occurs only after B cells have been stimulated by [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/). It occurs through a specialized nonhomologous DNA recombination mechanism guided by stretches of repetitive DNA known as switch **regions**. Switch **regions** lie in the intron between the JH [gene segments](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2836/) and the Cμ gene, and at equivalent sites upstream of the genes for each of the other heavy-chain isotypes, with the exception of the δ gene ([Fig. 4.20](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A476/?report=objectonly" \t "object), top panel). The μ [switch region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3272/) (Sμ) consists of about 150 repeats of the sequence [(GAGCT)*n* (GGGGGT)], where *n* is usually 3 but can be as many as 7. The sequences of the other switch **regions** (Sγ, Sα, and Sε) differ in detail but all contain repeats of the GAGCT and GGGGGT sequences. Exactly how these repetitive sequences promote switch recombination is unclear because the enzyme(s) that promote switch recombination have not been identified; however, it is thought that the repetitive sequences might promote short stretches of homologous alignment that in turn promote DNA recombination.

#### [Figure 4.20](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A476/?report=objectonly" \t "object)

Isotype switching involves recombination between specific switch signals. Repetitive DNA sequences that guide isotype switching are found upstream of each of the **immunoglobulin** C-region genes, with the exception of the δ gene. The figure illustrates [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A476/?report=objectonly" \t "object)

When a [B cell](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2618/) switches from coexpression of [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/) and [IgD](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2906/) to expression of an [IgG](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2908/) subtype, DNA recombination occurs between Sμ and the Sγ of that IgG subtype. The Cμ and Cδ coding **regions** are deleted, and γ heavy-chain transcripts are made from the recombined gene. [Figure 4.20](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A476/?report=objectonly" \t "object) (left panels) illustrates switching to γ3 in the mouse. Some of the progeny of this IgG-producing cell may subsequently undergo a further switching event to produce a different isotype, for example [IgA](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2905/), as shown in the bottom panel of [Fig. 4.20](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A476/?report=objectonly" \t "object). Alternatively, as shown in the right panels of [Fig. 4.20](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A476/?report=objectonly" \t "object), the switch recombination may occur between Sμ and one of the switch **regions** downstream of the Cγ genes so that the cell switches from IgM to IgA or [IgE](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2907/) (illustrated for IgA only). All switch recombination events produce genes that can encode a functional protein because the switch sequences lie in introns and therefore cannot cause frame shift mutations.

The enzymes that carry out [isotype switching](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2973/) have not been clearly defined. However, we do know that DNA repair enzymes are involved since switching is markedly reduced in Ku protein knockouts; Ku proteins are also essential for the rejoining of DNA during V(D)J joining (see [Section 4-5](https://www.ncbi.nlm.nih.gov/books/n/imm/A436/#A447)). Recently, it was discovered that deficiency in Activation Induced Cytidine Deaminase completely blocks isotype switching. As mentioned in [Section 4-9](https://www.ncbi.nlm.nih.gov/books/n/imm/A436/#A453), this deficiency also blocks somatic hypermutation. Activation Induced Cytidine Deaminise is thought to be an RNA editing enzyme and how it works to enable both hypermutation and switching is unknown at present. A deficiency in this enzyme in humans has now been associated with a form of immuno-deficiency known as Hyper [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/) type 2 syndrome, which is characterized by an absence of [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) other than IgM. A failure of [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) to activate isotype switching leads to a similar syndrome now classified as Hyper IgM type 1 (see [Section 11-9](https://www.ncbi.nlm.nih.gov/books/n/imm/A1494/#A1503)). (Hyper IgM Immunodeficiency, in *Case Studies in*[*Immunology*](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2945/), see Preface for details)

Isotype switch recombination is unlike [V(D)J recombination](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3354/) in several ways. First, all isotype switch recombination is productive; second, it uses different recombination signal sequences and enzymes; third, it happens after [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) stimulation and not during B-cell development in the [bone marrow](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2640/); and fourth, the switching process is not random but is regulated by external signals such as those provided by [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/), as will be discussed in Chapter 9.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27106/)

## 4-17. Transmembrane and secreted forms of immunoglobulin are generated from alternative heavy-chain transcripts

Immunoglobulins of all heavy-chain [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/) can be produced either in secreted form or as a membrane-bound receptor. All B cells initially express the transmembrane form of [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/); after [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) stimulation, some of their progeny differentiate into plasma cells producing the secreted form of IgM, whereas others undergo [isotype switching](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2973/) to express transmembrane [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) of a different isotype before switching to the production of secreted [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) of the new isotype.

The membrane forms of all [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/) are monomers comprised of two heavy and two light chains: [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/) and [IgA](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2905/) polymerize only when they are secreted. In its membrane-bound form the **immunoglobulin** [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/) has a hydrophobic transmembrane domain of about 25 amino acid residues at the carboxy terminus, which anchors it to the surface of the [B lymphocyte](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2637/). This transmembrane domain is absent from the secreted form, whose carboxy terminus is a hydrophilic secretory tail. The two different carboxy termini of the transmembrane and secreted forms of **immunoglobulin** heavy chains are encoded in separate exons and production of the two forms is achieved by alternative RNA processing ([Fig. 4.21](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A478/?report=objectonly" \t "object)). The last two exons of each CH gene contain the sequences encoding the secreted and the transmembrane **regions** respectively; if the primary transcript is cleaved and polyadenylated at a site downstream of these exons, the sequence encoding the carboxy terminus of the secreted form is removed by splicing and the cell-surface form of **immunoglobulin** is produced. Alternatively, if the primary transcript is cleaved at the polyadenylation site located before the last two exons, only the secreted molecule can be produced. This differential RNA processing is illustrated for Cμ in [Fig. 4.21](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A478/?report=objectonly" \t "object), but occurs in the same way for all isotypes.

#### [Figure 4.21](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A478/?report=objectonly" \t "object)

Transmembrane and secreted forms of immunoglobulins are derived from the same heavy-chain sequence by alternative RNA processing. Each heavy-chain C gene has two exons (membrane-coding (MC) yellow) that encode the transmembrane region and cytoplasmic [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A478/?report=objectonly" \t "object)

Although the production of membrane-bound and secreted versions of the [heavy chain](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2862/) is achieved by similar mechanisms to those that allow the co-expression of surface [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/) and [IgD](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2906/) (see [Fig. 4.19](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A475/?report=objectonly" \t "object)), these two instances of alternative RNA processing act at different stages in the life of the [B cell](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2618/), and on different primary transcripts. B cells make a long heavy-chain transcript that can be processed to give either transmembrane IgM or IgD before they are stimulated by [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/). A B cell that is activated ceases to coexpress IgD with IgM, either because μ and δ sequences have been removed as a consequence of an isotype switch or, in IgM-secreting plasma cells, because transcription from the VH promoter no longer extends through the Cδ exons. In activated B cells that differentiate to become [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/)-secreting plasma cells, much of the transcript is spliced to the secreted rather than transmembrane form of whichever isotype the B cell happens to express.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27106/)

## 4-18. Antibody C regions confer functional specialization

The secreted antibodies protect the body in a variety of ways, as we briefly outline here and discuss further in Chapter 9. In some cases it is enough for the [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) simply to bind [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/). For instance, by binding tightly to a toxin or virus, an antibody can prevent it from recognizing its receptor on a host cell. The V **regions** on their own are sufficient for this. The [C region](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2649/) is essential, however, for recruiting the help of other cells and molecules to destroy and dispose of pathogens to which the antibody has bound, and it confers functionally distinct properties on each of the various [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/).

The C **regions** of antibodies have three main effector functions. First, the Fc portions of different [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/) are recognized by specialized receptors expressed by immune effector cells. The Fc portions of IgG1 and IgG3 antibodies are recognized by [Fc receptors](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2816/) present on the surface of phagocytic cells such as macrophages and neutrophils, which can thereby bind and engulf pathogens coated with antibodies of these isotypes. The Fc portion of [IgE](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2907/) binds to a high-affinity [Fcε receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2817/) on mast cells, basophils, and activated eosinophils, enabling these cells to respond to the binding of specific [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) by releasing inflammatory mediators. Second, the Fc portions of antigen: [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) complexes can bind to [complement](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2720/) (see [Fig. 1.24](https://www.ncbi.nlm.nih.gov/books/n/imm/A75/figure/A78/?report=objectonly" \t "object)) and initiate the complement cascade, which helps to recruit and activate phagocytes, can aid the engulfment of microbes by phagocytes, and can also directly destroy pathogens. Third, the Fc portion can deliver antibodies to places they would not reach without active transport. These include the mucus secretions, tears, and milk ([IgA](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2905/)), and the fetal blood circulation by transfer from the pregnant mother ([IgG](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2908/)). In both cases, the Fc portion engages a specific receptor that leads to the active transport of the **immunoglobulin** through cells to reach different body compartments.

The role of the Fc portion in these effector functions can be demonstrated by studying enzymatically treated [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) that have had one or other domain of the Fc cleaved off (see [Section 3-3](https://www.ncbi.nlm.nih.gov/books/n/imm/A323/#A328)) or, more recently, by genetic engineering, which permits detailed mapping of the exact amino acid residues within the Fc that are needed for particular functions. Many kinds of microorganism seem to have responded to the destructive potential of the Fc portion by manufacturing proteins that either bind to it or proteolytically cleave it, and so prevent the Fc region from working. Examples of these are [Protein A](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3180/) and Protein G made by *Staphylococcus* species ([Fig. 4.22](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A480/?report=objectonly" \t "object)), and Protein D of *Haemophilus* species. Researchers can exploit these proteins to help to map the Fc and as immunological reagents (see [Appendix I, Section A-10](https://www.ncbi.nlm.nih.gov/books/n/imm/A2395/#A2419)). Not all **immunoglobulin** [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/) or subtypes have the same capacity to engage each of the effector functions. The differential capabilities of each isotype are summarized in [Fig. 4.16](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A471/?report=objectonly" \t "object). For example, IgG1 and IgG3 have higher affinity for the most common type of Fc receptor.

#### [Figure 4.22](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A480/?report=objectonly" \t "object)

Protein A of *Staphylococcus aureus* bound to a fragment of the Fc region of IgG. A fragment of the Fc portion of a single IgG heavy chain is complexed with a fragment of the **immunoglobulin**-binding Protein A from *Staphylococcus aureus*. The Fc fragment has [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A480/?report=objectonly" \t "object)

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27106/)

## 4-19. IgM and IgA can form polymers

Although all **immunoglobulin** molecules are constructed from a basic unit of two heavy and two light chains, both [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/) and [IgA](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2905/) can form multimers ([Fig. 4.23](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A482/?report=objectonly" \t "object)). IgM and IgA C **regions** contain a ‘tailpiece’ of 18 amino acids that contains a cysteine residue essential for polymerization. An additional separate 15 kDa polypeptide chain called the J chain promotes polymerization by linking to the cysteines of the tailpiece, which is found only in the secreted forms of the μ and α chains. (This J chain should not be confused with the J gene segment; see [Section 4-2](https://www.ncbi.nlm.nih.gov/books/n/imm/A436/#A439).) In the case of IgA, polymerization is required for transport through epithelia, as we discuss in Chapter 9. IgM molecules are found as pentamers, and occasionally hexamers (without J chain), in plasma, whereas IgA in mucous secretions, but not in plasma, is mainly found as a dimer (see [Fig. 4.23](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A482/?report=objectonly" \t "object)).

#### [Figure 4.23](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A482/?report=objectonly" \t "object)

The IgM and IgA molecules can form multimers. IgM and IgA are usually synthesized as multimers in association with an additional polypeptide chain, the J chain. In pentameric IgM, the monomers are cross-linked by disulfide bonds to each other and to the [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A482/?report=objectonly" \t "object)

The polymerization of **immunoglobulin** molecules is thought to be important in the binding of [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) to repetitive epitopes. The dissociation rate of an individual [epitope](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2800/) from an individual antibody binding site influences the strength of binding, or affinity, of that site: the lower the dissociation rate, the higher the affinity (see [Appendix I, Section A-9](https://www.ncbi.nlm.nih.gov/books/n/imm/A2395/#A2416)). An antibody molecule has two or more identical [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-binding sites, and if it attaches to two or more repeating epitopes on a single target antigen, it will only dissociate when all sites dissociate. The dissociation rate of the whole antibody from the whole antigen will therefore be much slower than the rate for the individual binding sites, giving a greater effective total binding strength, or avidity. This consideration is particularly relevant for pentameric [IgM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2909/), which has ten antigen-binding sites. IgM antibodies frequently recognize repetitive epitopes such as those expressed by bacterial cell-wall polysaccharides, but the binding of individual sites is often of low affinity because IgM is made early in immune responses, before somatic hypermutation and affinity maturation. Multisite binding makes up for this, dramatically improving the overall functional binding strength.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27106/)

## 4-20. Various differences between immunoglobulins can be detected by antibodies

When an **immunoglobulin** is used as an [antigen](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/), it will be treated like any other foreign protein and will elicit an [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/) response. [Anti-**immunoglobulin** antibodies](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2592/) can be made that recognize the amino acids that characterize the isotype of the injected antibody. Such anti-isotypic antibodies recognize all [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) of the same isotype in all members of the species from which the injected antibody came.

It is also possible to raise antibodies that recognize differences in immuno-globulins from members of the same species that are due to the presence of multiple alleles of the individual C genes in the population (genetic polymorphism). Such allelic variants are called allotypes. In contrast to anti-isotypic antibodies, anti-allotypic antibodies will recognize immuno-globulin of a particular isotype only in some members of a species. Finally, as individual antibodies differ in their V **regions**, one can raise antibodies against unique sequence variants, which are called idiotypes.

A schematic picture of the differences between idiotypes, allotypes, and [isotypes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2974/) is given in [Fig. 4.24](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A484/?report=objectonly" \t "object). Historically, the main features of [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/) were defined by using isotypic and allotypic genetic markers identified by antisera raised in different species or in genetically distinct members of the same species (see [Appendix I, Section A-10](https://www.ncbi.nlm.nih.gov/books/n/imm/A2395/#A2419)). The independent segregation of allotypic and isotypic markers revealed the existence of separate heavy-chain, κ, and λ genes.

#### [Figure 4.24](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A484/?report=objectonly" \t "object)

Different types of **variation** between immunoglobulins. Differences between **constant** **regions** due to usage of different C-region genes are called

isotypes; differences due to different alleles of the same C gene are called allotypes; differences due to particular [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27106/figure/A484/?report=objectonly" \t "object)

Chapter 5. **Antigen** Presentation to T Lymphocytes

* The **generation** of **T-cell** receptor **ligands**
  + 5-1. The MHC class I and class II molecules deliver peptides to the cell surface from two distinct intracellular compartments
  + 5-2. Peptides that bind to MHC class I molecules are actively transported from the cytosol to the endoplasmic reticulum
  + 5-3. Peptides for transport into the endoplasmic reticulum are generated in the cytosol
  + 5-4. Newly synthesized MHC class I molecules are retained in the endoplasmic reticulum until they bind peptide
  + 5-5. Peptides presented by MHC class II molecules are generated in acidified endocytic vesicles
  + 5-6. The invariant chain directs newly synthesized MHC class II molecules to acidified intracellular vesicles
  + 5-7. A specialized MHC class II-like molecule catalyzes loading of MHC class II molecules with peptides
  + 5-8. Stable binding of peptides by MHC molecules provides effective **antigen** presentation at the cell surface

# The generation of T-cell receptor ligands

The protective function of [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) depends on their ability to recognize cells that are harboring pathogens or that have internalized pathogens or their products. T cells do this by recognizing peptide fragments of pathogen-derived proteins in the form of complexes of peptides and [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) on the cell surface. Because the **generation** of peptides from an intact [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) involves modification of the native protein, it is commonly referred to as **antigen** processing, whereas the display of the peptide at the cell surface by the MHC molecule is referred to as **antigen** presentation. We have already described the structure of MHC molecules and seen how they bind peptide antigens in a cleft on their outer surface (see [Sections 3-15](https://www.ncbi.nlm.nih.gov/books/n/imm/A346/#A364) to [3-18](https://www.ncbi.nlm.nih.gov/books/n/imm/A346/#A371)). In this chapter we will look at how peptides are generated from pathogens present in the cytosol or in the vesicular compartment of cells and loaded onto MHC class I and MHC class II molecules, respectively, at different sites inside the cell. Both classes of MHC molecule must combine with a peptide before they can be stably expressed at the cell surface. Peptide-binding completes the folding and assembly of newly synthesized MHC class I molecules in the endoplasmic reticulum, while MHC class II molecules are prevented from binding peptides in the endoplasmic reticulum and are instead escorted to an endosomal compartment where loading with vesicular peptides occurs.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27137/)

## 5-1. The MHC class I and class II molecules deliver peptides to the cell surface from two distinct intracellular compartments

Infectious agents can replicate in either of two distinct intracellular compartments ([Fig. 5.1](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A556/?report=objectonly" \t "object)). [Viruses](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3359/) and certain [bacteria](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2622/) replicate in the cytosol or in the contiguous nuclear compartment ([Fig. 5.2](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A557/?report=objectonly" \t "object), left panel), whereas many pathogenic bacteria and some eukaryotic parasites replicate in the [endosomes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2794/) and lysosomes that form part of the vesicular system ([Fig. 5.2](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A557/?report=objectonly" \t "object), center panel). The [immune system](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2921/) has different strategies for eliminating pathogens from these two sites. Cells infected with viruses or with bacteria that live in the cytosol are eliminated by [cytotoxic T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2755/); as mentioned in Chapter 3, these T cells are distinguished by the [co-receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2733/) molecule [CD8](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2671/). The function of [CD8 T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2672/) is to kill infected cells; this is an important means of eliminating sources of new viral particles and obligate cytosolic bacteria, thus freeing the host of infection.

#### [Figure 5.1](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A556/?report=objectonly" \t "object)

There are two major intracellular compartments, separated by membranes. The first is the cytosol, which also communicates with the nucleus via the nuclear pores in the nuclear membrane. The second is the vesicular system, which comprises the endoplasmic [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A556/?report=objectonly" \t "object)

#### [Figure 5.2](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A557/?report=objectonly" \t "object)

Pathogens and their products can be found in either the cytosolic or the vesicular compartment of cells. Left panel: all viruses and some bacteria replicate in the cytosolic compartment. Their antigens are presented by MHC class I molecules to CD8 T cells. [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A557/?report=objectonly" \t "object)

Pathogens and their products in the vesicular compartments of cells are detected by a different class of T cell, distinguished by the [co-receptor](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2733/) molecule [CD4](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2668/) (see Chapter 3). [CD4 T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2669/) are specialized to activate other cells and fall into two functional classes: [TH1 cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3295/) (sometimes known as inflammatory T cells), whose main function is to activate macrophages to kill the intravesicular pathogens they harbor; and [TH2 cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3296/), or helper T cells, which activate B cells to make [antibody](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2575/). Microbial antigens may enter the vesicular compartment in either of two ways. Some [bacteria](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2622/), including the mycobacteria that cause tuberculosis and leprosy, invade macrophages and flourish in intracellular vesicles. Other bacteria proliferate outside cells, where they cause pathology by secreting toxins and other proteins. These bacteria and their toxic products can be internalized by phagocytosis, endocytosis, or [macropinocytosis](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3040/) into the intracellular vesicles of cells that then present [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) to T cells. These include the dendritic cells that specialize in initiating T cell responses (see [Section 1-6](https://www.ncbi.nlm.nih.gov/books/n/imm/A53/#A56)), macrophages that specialize in taking up particulate material (see [Section 2-3](https://www.ncbi.nlm.nih.gov/books/n/imm/A150/#A156)), and B cells that efficiently internalize specific **antigen** by receptor-mediated endocytosis of the **antigen** bound to their [surface immunoglobulin](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3270/) ([Fig. 5.2](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A557/?report=objectonly" \t "object), right panel).

To produce an appropriate response to infectious microorganisms, [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) need to be able to detect the presence of intracellular pathogens and to distinguish between foreign material coming from the cytosolic and vesicular compartments. This is achieved through the use of the different classes of MHC molecule. MHC class I molecules deliver peptides originating in the cytosol to the cell surface, where they are recognized by [CD8 T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2672/). MHC class II molecules deliver peptides originating in the vesicular system to the cell surface, where they are recognized by [CD4 T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2669/). As we saw in [Section 3-12](https://www.ncbi.nlm.nih.gov/books/n/imm/A346/#A353), CD8 and CD4 bind MHC class I and MHC class II molecules, respectively, and so help to ensure that the appropriate type of T cell is activated in response to a given pathogen.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27137/)

## 5-2. Peptides that bind to MHC class I molecules are actively transported from the cytosol to the endoplasmic reticulum

The [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) fragments that bind to MHC class I molecules are typically derived from viruses that take over the cell's biosynthetic mechanisms to make their own proteins. All proteins are made in the cytosol. The polypeptide chains of proteins destined for the cell surface, which include both classes of MHC molecule, are translocated during synthesis into the lumen of the endoplasmic reticulum. Here the chains must fold correctly, and assemble with each other if necessary, before the complete protein can be transported to the cell surface. Thus the peptide-binding site of the MHC class I molecule is formed in the lumen of the endoplasmic reticulum and is never exposed to the cytosol. This raised the question—how are peptides derived from viral proteins in the cytosol able to bind to MHC class I molecules for delivery to the cell surface?

The first clues came from mutant cells with a defect in [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) presentation by MHC class I molecules. Although both chains of MHC class I molecules are synthesized normally in these cells, the MHC class I proteins are present at abnormally low levels on the cell surface. This defect can be corrected by the addition of synthetic peptides to the medium bathing the cells, suggesting both that the mutation affects the supply of peptides to MHC class I molecules and that peptide is required for their maintenance at the cell surface. This was the first indication that [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) are unstable in the absence of bound peptide.

Analysis of the DNA of the mutant cells showed that two genes encoding members of the ATP-binding cassette, or ABC, **family** of proteins are mutant or absent in these cells. ABC proteins mediate ATP-dependent transport of ions, sugars, amino acids, and peptides across membranes in many types of cells, including [bacteria](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2622/). The two ABC proteins missing in the mutant cells are normally associated with the endoplasmic reticulum membrane. Transfection of the mutant cells with both genes restores presentation of peptides by the cell's MHC class I molecules. These proteins are now called **Transporters associated with Antigen Processing-1** and **-2** ([**TAP1**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3281/) and **TAP2**). The two TAP proteins form a heterodimer ([Fig. 5.3](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A559/?report=objectonly" \t "object)) and mutations in either TAP gene can prevent [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) presentation by MHC class I molecules. The genes *TAP1* and *TAP2* map within the MHC itself (see [Section 5-9](https://www.ncbi.nlm.nih.gov/books/n/imm/A574/#A575)), and are inducible by interferons, which are produced in response to virus infection.

#### [Figure 5.3](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A559/?report=objectonly" \t "object)

TAP1 and TAP2 form a peptide transporter in the endoplasmic reticulum membrane. All transporters of the ATP-binding cassette (ABC) **family** are composed of four domains: two hydrophobic transmembrane domains that have multiple transmembrane regions, and [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A559/?report=objectonly" \t "object)

In *in vitro* assays using normal cell fractions, microsomal vesicles that mimic the endoplasmic reticulum will internalize peptides, which then bind to MHC class I molecules already present in the microsome lumen. [Vesicles](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3356/) from [TAP1](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3281/) or TAP2 mutant cells do not transport peptides. Peptide transport into the normal microsomes requires ATP hydrolysis, proving that the TAP1:TAP2 complex is an ATP-dependent peptide transporter. Such experiments have also shown that the TAP complex has some [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/) for the peptides it will transport. It prefers peptides of eight or more amino acids with hydrophobic or basic residues at the carboxy terminus—the exact features of peptides that bind MHC class I molecules. The TAP transporter provides the answer to the question of how viral peptides gain access to the lumen of the ER in order to bind to MHC class I molecules, but leaves open the question of how these peptides are generated.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27137/)

## 5-3. Peptides for transport into the endoplasmic reticulum are generated in the cytosol

Proteins in cells are continually being degraded and replaced with newly synthesized proteins. Much cytosolic protein degradation is carried out by a large, multicatalytic protease complex called the [proteasome](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3177/) ([Fig. 5.4](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A561/?report=objectonly" \t "object)). The proteasome is a large cylindrical complex of some 28 subunits, arranged in four stacked rings, each of seven subunits, and it has a hollow core lined by the active sites of the proteolytic subunits of the proteasome. Proteins to be degraded are introduced into the core of the proteasome and are there broken down into short peptides.

#### [Figure 5.4](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A561/?report=objectonly" \t "object)

The structure of the proteasome. Proteasomes are found throughout the eukaryotes and the archaebacteria, and their structure and function are highly conserved. The structure shown here is from an archaebacterium, as the detailed structure of a mammalian [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A561/?report=objectonly" \t "object)

Various lines of evidence implicate the [proteasome](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3177/) in the production of peptide **ligands** for MHC class I molecules. For example, the proteasome takes part in the ubiquitin-dependent degradation pathway for cytosolic proteins, and experimentally tagging proteins with ubiquitin also results in more efficient presentation of their peptides by MHC class I molecules. Moreover, inhibitors of the proteolytic activity of the proteasome also inhibit [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) presentation by MHC class I molecules. Whether the proteasome is the only cytosolic protease capable of generating peptides for transport into the endoplasmic reticulum is not known.

Two subunits of the [proteasome](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3177/), called LMP2 and LMP7, are encoded within the MHC near the [*TAP1*](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3281/) and *TAP2* genes. Along with MHC class I and TAP molecules, their expression is induced by interferons, which are produced in response to viral infections. LMP2 and LMP7 substitute for two constitutively expressed subunits of the proteasome. A third subunit, MECL-1, which is not encoded within the MHC, is also induced by interferons and also displaces a constitutive proteasome subunit. These three inducible subunits and their constitutive counterparts are thought to be the active proteases of the proteasome. The replacement of the constitutive components by their interferon-inducible counterparts seems to change the [specificity](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3259/) of the proteasome: in interferon-treated cells, there is increased cleavage of polypeptides after hydrophobic and basic residues, and reduced cleavage after acidic residues. This produces peptides with carboxy-terminal residues that are preferred [anchor residues](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2572/) for binding to most MHC class I molecules and are also the preferred structures for transport by TAP.

MHC class I molecules also present peptides derived from membrane and secreted proteins, for example, the glycoproteins of viral envelopes. Membrane and secreted proteins are normally translocated into the lumen of the endoplasmic reticulum during their biosynthesis. Yet the peptides bound by MHC class I molecules bear evidence that such proteins are degraded in the cytosol. Asparagine-linked carbohydrate moieties commonly present on membrane-bound or secreted proteins can be removed in the cytosol by an enzyme reaction that changes the asparagine residue into aspartic acid, and this diagnostic sequence change can be seen in some peptides presented by MHC class I molecules. It now appears that endoplasmic reticulum proteins can be returned to the cytosol by the same translocation system that transported them into the endoplasmic reticulum in the first place. This newly discovered mechanism, known as retrograde translocation, may be the normal mechanism by which proteins in the endoplasmic reticulum are turned over, and by which misfolded proteins in the endoplasmic reticulum are removed and degraded. Once in the cytosol, the polypeptides are degraded by the [proteasome](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3177/). The resulting peptides can then be transported back into the lumen of the endoplasmic reticulum via TAP and loaded onto MHC class I molecules.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27137/)

## 5-4. Newly synthesized MHC class I molecules are retained in the endoplasmic reticulum until they bind peptide

Binding of peptide is an important step in the assembly of stable MHC class I molecules. When the supply of peptides into the endoplasmic reticulum is disrupted, as in the *TAP* mutant cells, newly synthesized MHC class I molecules are held in the endoplasmic reticulum in a partially folded state. This explains why cells with mutations in [*TAP1*](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3281/) or *TAP2* fail to express MHC class I molecules at the cell surface. The folding and assembly of a complete MHC class I molecule (see [Fig. 3.20](https://www.ncbi.nlm.nih.gov/books/n/imm/A346/figure/A361/?report=objectonly" \t "object)) depends on the association of the MHC class I α chain first with β2-microglobulin and then with peptide, and this process involves a number of accessory proteins with a chaperone-like function. Only after peptide has bound is the MHC class I molecule released from the endoplasmic reticulum and allowed to reach the cell surface. (MHC Class I Deficiency, in *Case Studies in*[*Immunology*](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2945/), see Preface for details)

In humans, newly synthesized MHC class I α chains that enter the endoplasmic reticulum membranes bind to a chaperone protein, [calnexin](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2660/), which retains the MHC class I molecule in a partially folded state in the endoplasmic reticulum. Calnexin also associates with partially folded **T-cell** **receptors**, [immunoglobulins](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2897/), and MHC class II molecules, and so has a central role in the assembly of many immunological proteins. When β2-microglobulin binds to the α chain, the partially folded α:β2-microglobulin heterodimer dissociates from calnexin and now binds to a complex of proteins, one of which—calreticulin—is similar to calnexin and probably carries out a similar chaperone function. A second component of the complex is the TAPassociated protein tapasin, also encoded by a gene that lies within the MHC. [Tapasin](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3282/) forms a bridge between MHC class I molecules and [TAP1](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3281/) and TAP2, allowing the partially folded α:β2-microglobulin heterodimer to await the transport of a suitable peptide from the cytosol. A third component of this complex is the chaperone molecule [Erp57](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2805/), a protein disulfide isomerase that may have a role in breaking and reforming the disulfide bond in the MHC class I α2 domain during peptide loading. Erp57 and calreticulin bind to a number of glycoproteins during their assembly in the endoplasmic reticulum and seem to be part of the cell's general quality-control mechanism. Finally, the binding of a peptide to the partially folded heterodimer releases it from the complex of TAP:tapasin:calreticulin:Erp57. The fully folded MHC class I molecule and its bound peptide are now able to leave the endoplasmic reticulum and are transported to the cell surface ([Fig. 5.5](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A563/?report=objectonly" \t "object)).

#### [Figure 5.5](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A563/?report=objectonly" \t "object)

MHC class I molecules do not leave the endoplasmic reticulum unless they bind peptides. Newly synthesized MHC class I α chains assemble in the endoplasmic reticulum with a membrane-bound protein, calnexin. When this complex binds β2-microglobulin [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A563/?report=objectonly" \t "object)

Most of the peptides transported by TAP will not bind to the [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) in that cell and are rapidly cleared out of the endoplasmic reticulum; there is evidence that they are transported back into the cytosol by an ATP-dependent transport mechanism distinct from that of TAP. It is not yet clear whether the TAP:tapasin complex directly loads peptides onto MHC class I molecules or whether binding to the TAP complex merely allows the MHC class I molecule to scan the transported peptides before they diffuse through the lumen of the endoplasmic reticulum and are transported back into the cytosol.

In cells with mutant TAP genes, the MHC class I molecules in the endoplasmic reticulum are unstable and are eventually translocated back into the cytosol, where they are degraded. Thus, the MHC class I molecule must bind a peptide to complete its folding and be transported onward from the endoplasmic reticulum. In uninfected cells, peptides derived from self proteins fill the peptide-binding cleft of the mature MHC class I molecules and are carried to the cell surface. In normal cells, MHC class I molecules are retained in the endoplasmic reticulum for some time, which suggests that they are present in excess of peptide. This is very important for the function of MHC class I molecules because they must be immediately available to transport viral peptides to the cell surface if the cell becomes infected. When a cell is infected by a virus, the presence of excess MHC class I molecules in the endoplasmic reticulum allows the rapid appearance of pathogen-derived peptides at the cell surface.

Because the presentation of viral peptides by MHC class I molecules signals [CD8 T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2672/) to kill the infected cell, some viruses have evolved ways of evading recognition by preventing the appearance of peptide:MHC class I complexes at the cell surface. The herpes simplex virus prevents the transport of viral peptides into the endoplasmic reticulum by producing a protein that binds to and inhibits TAP. Adenoviruses, on the other hand, encode a protein that binds to MHC class I molecules and retains them in the endoplasmic reticulum. Cytomegalovirus accelerates the retrograde translocation of MHC class I molecules back into the cytosol of the cell, where they are degraded. The advantage to a virus of blocking the recognition of infected cells is so great that it would not be surprising if other steps in the formation of MHC:peptide complexes, for example, the association of the MHC class I:chaperone complex with TAP, were found to be inhibited by some viruses.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27137/)

## 5-5. Peptides presented by MHC class II molecules are generated in acidified endocytic vesicles

Several classes of pathogen, including the protozoan parasite *Leishmania* and the mycobacteria that cause leprosy and tuberculosis, replicate inside intracellular vesicles in macrophages. Because they reside in membrane-bounded vesicles, the proteins of these pathogens are not accessible to proteasomes in the cytosol. Instead, after activation of the macrophage, proteins in vesicles are degraded by proteases within the vesicles into peptide fragments that bind to MHC class II molecules. In this way they are delivered to the cell surface where they can be recognized by [CD4 T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2669/). Extracellular pathogens and proteins that are internalized into endocytic vesicles are also processed by this pathway and their peptides are presented to CD4 T cells ([Fig. 5.6](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A565/?report=objectonly" \t "object)).

#### [Figure 5.6](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A565/?report=objectonly" \t "object)

Peptides that bind to MHC class II molecules are generated in acidified endocytic vesicles. In the case illustrated here, extracellular foreign antigens, such as bacteria or bacterial antigens, have been taken up by an **antigen**-presenting cell such as [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A565/?report=objectonly" \t "object)

Most of what we know about protein processing in the endocytic pathway has come from experiments in which simple proteins are fed to macrophages and are taken up by endocytosis; in this way the processing of added [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) can be quantified. Proteins that bind to [surface immunoglobulin](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3270/) on B cells and are internalized by receptor-mediated endocytosis are processed by the same pathway. Proteins that enter cells through endocytosis become enclosed in [endosomes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2794/), which become increasingly acidic as they progress into the interior of the cell, eventually fusing with lysosomes. The endosomes and lysosomes contain proteases, known as acid proteases, which are activated at low pH and eventually degrade the protein antigens contained in the vesicles. Larger particulate material internalized by phagocytosis or [macropinocytosis](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3040/) can also be handled by this pathway of **antigen** processing.

Drugs, such as chloroquine, that raise the pH of [endosomes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2794/), making them less acid, inhibit the presentation of antigens that enter the cell in this way, suggesting that acid proteases are responsible for the processing of internalized [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/). Among these acid proteases are the cysteine proteases cathepsins B, D, S, and L, the last of which is the most active enzyme in this **family**. [**Antigen** processing](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2589/) can be mimicked to some extent by digestion of proteins with these enzymes *in vitro* at acid pH. Cathepsins S and L may be the predominant proteases involved in the processing of vesicular antigens; mice that lack cathepsin B or cathepsin D show normal **antigen** processing, whereas mice with no cathepsin S cannot process **antigen**.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27137/)

## 5-6. The invariant chain directs newly synthesized MHC class II molecules to acidified intracellular vesicles

The function of MHC class II molecules is to bind peptides generated in the intracellular vesicles of macrophages, [immature dendritic cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2912/), B cells, and other [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-presenting cells and to present these peptides to [CD4 T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2669/). However, the biosynthetic pathway for MHC class II molecules, like that of other cell-surface glycoproteins, starts with their translocation into the endoplasmic reticulum, and they must therefore be prevented from binding prematurely to peptides transported into the endoplasmic reticulum lumen or to the cell's own newly synthesized polypeptides. As the endoplasmic reticulum is richly endowed with unfolded and partially folded polypeptide chains, a general mechanism is needed to prevent their binding in the open-ended MHC class II peptide-binding groove.

Binding is prevented by the assembly of newly synthesized MHC class II molecules with a protein known as the MHC class II-associated **invariant chain** (**Ii**). The invariant chain forms trimers, with each subunit binding noncovalently to an MHC class II α:β heterodimer ([Fig. 5.7](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A567/?report=objectonly" \t "object)). Ii binds to the MHC class II molecule with part of its polypeptide chain lying within the peptide-binding groove, thus blocking the groove and preventing the binding of either peptides or partially folded proteins. While this complex is being assembled in the endoplasmic reticulum, its component parts are associated with [calnexin](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2660/). Only when assembly is completed to produce a nine-chain complex is the complex released from calnexin for transport out of the endoplasmic reticulum. When it is part of the nine-chain complex, the MHC class II molecule cannot bind peptides or unfolded proteins, so that peptides present in the endoplasmic reticulum are not usually presented by MHC class II molecules. There is evidence that in the absence of invariant chains many MHC class II molecules are retained in the endoplasmic reticulum as complexes with misfolded proteins.

#### [Figure 5.7](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A567/?report=objectonly" \t "object)

The invariant chain is cleaved to leave a peptide fragment, CLIP, bound to the MHC class II molecule. A model of the trimeric invariant chain bound to MHC class II α:β heterodimers is shown on the left. The CLIP portion is shown in red, the [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A567/?report=objectonly" \t "object)

The invariant chain has a second function, which is to target delivery of the MHC class II molecules to a low-pH endosomal compartment where peptide loading can occur. The complex of MHC class II α:β heterodimers with invariant chain is retained for 2–4 hours in this compartment. During this time, the invariant chain is cleaved by acid proteases such as cathepsin S in several steps, as shown in [Fig. 5.7](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A567/?report=objectonly" \t "object). The initial cleavage events generate a truncated form of the invariant chain that remains bound to the MHC class II molecule and retains it within the proteolytic compartment. A subsequent cleavage releases the MHC class II molecule from the membrane-associated fragment of Ii, leaving a short fragment of Ii, called **CLIP** (for **class II-associated invariant-chain peptide**) still bound to the MHC class II molecule. MHC class II molecules associated with CLIP still cannot bind other peptides. CLIP must either dissociate or be displaced to allow a peptide to bind to the MHC molecule and enable the complex to be delivered to the cell surface. Cathepsin S cleaves Ii in most class II-positive cells, including antigenpresenting cells, whereas cathepsin L appears to substitute for cathepsin S in thymic cortical epithelial cells.

The endosomal compartment in which invariant chain is cleaved and MHC class II molecules encounter peptide is not yet clearly defined. Most newly synthesized MHC class II molecules are brought toward the cell surface in vesicles, which at some point fuse with incoming [endosomes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2794/). However, there is also evidence that some MHC class II:Ii complexes are first transported to the cell surface and then re-internalized into endosomes. In either case, MHC class II:Ii complexes enter the endosomal pathway and there encounter and bind pathogen-derived peptides. Immunoelectron-microscopy using antibodies tagged with gold particles to localize Ii and MHC class II molecules within cells suggests that Ii is cleaved and peptides bind to MHC class II in a particular endosomal compartment called the [MIIC](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3073/) (**MHC class** ***II*compartment**), late in the endosomal pathway ([Fig. 5.8](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A568/?report=objectonly" \t "object)).

#### [Figure 5.8](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A568/?report=objectonly" \t "object)

MHC class II molecules are loaded with peptide in a specialized intracellular compartment. MHC class II molecules are transported from the Golgi apparatus (labeled G in this electron micrograph of an ultrathin section of a B cell) to the cell surface [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A568/?report=objectonly" \t "object)

As with MHC class I molecules, MHC class II molecules in uninfected cells bind peptides derived from self proteins. MHC class II molecules that do not bind peptide after dissociation from the invariant chain are unstable; in the acidic pH of the endosome they aggregate and are rapidly degraded. It is therefore not surprising that peptides derived from MHC class II molecules form a large proportion of the self peptides presented by MHC class II molecules in normal uninfected cells. This suggests that, as for MHC class I molecules, MHC class II molecules are produced in excess. Thus, when a cell is infected by pathogens that proliferate in intracellular vesicles, when a phagocyte engulfs a pathogen, or when pathogen-derived proteins are internalized by B cells, the peptides generated from the pathogen proteins find plentiful empty MHC class II molecules to bind.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27137/)

## 5-7. A specialized MHC class II-like molecule catalyzes loading of MHC class II molecules with peptides

An unsuspected component of the vesicular [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-processing pathway has been revealed by analysis of mutant human B-cell lines with a defect in **antigen** presentation. MHC class II molecules in these cell lines assemble correctly with the invariant chain and seem to follow the normal vesicular route. However, they fail to bind peptides derived from internalized proteins and often arrive at the cell surface with the CLIP peptide still bound.

The defect in these mutant cells lies in an MHC class II-like molecule called [HLA-DM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2880/) in humans (H-2M in mice). The HLA-DM genes are found near the TAP and LMP genes in the class II region of the MHC (see [Fig. 5.10](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A570/?report=objectonly" \t "object)); they encode an α chain and a β chain that closely resemble those of other MHC class II molecules. The HLA-DM molecule is not expressed at the cell surface, however, but is found predominantly in the [MIIC](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3073/) compartment and does not appear to require peptide for stabilization. HLA-DM binds to and stabilizes empty MHC class II molecules that would otherwise aggregate; in addition, it catalyzes both the release of the CLIP fragment from MHC class II:CLIP complexes and the binding of other peptides to the empty MHC class II molecule ([Fig. 5.9](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A571/?report=objectonly" \t "object)).

#### [Figure 5.10](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A570/?report=objectonly" \t "object)

The genetic organization of the major histocompatibility complex (MHC) in human and mouse. The organization of the principal MHC genes is shown for both humans (where the MHC is called HLA and is on chromo-some 6) and mice (in which the MHC is called [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A570/?report=objectonly" \t "object)

#### [Figure 5.9](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A571/?report=objectonly" \t "object)

HLA-DM facilitates the loading of antigenic peptides onto class II molecules. The invariant chain binds to newly synthesized MHC class II molecules and blocks the binding of peptides and unfolded proteins in the endoplasmic reticulum and during the transport [(more...)](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A571/?report=objectonly" \t "object)

[HLA-DM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2880/) also catalyzes the release of unstably bound peptides from MHC class II molecules. In the presence of a mixture of peptides capable of binding to MHC class II molecules, as occurs in the [MIIC](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3073/), HLA-DM will continuously bind and rebind to peptide:MHC class II complexes, removing weakly bound peptides and allowing other peptides to replace them. Antigens presented by MHC class II molecules may have to persist on the surface of antigenpresenting cells for some days before encountering [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) able to recognize them. The ability of HLA-DM to remove unstably bound peptides, sometimes called ‘peptide editing,’ ensures that the peptide:MHC class II complexes displayed on the surface of the [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-presenting cell will survive long enough to stimulate the appropriate [CD4](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2668/) cells.

A second atypical MHC class II molecule, called [HLA](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2879/)-DO (H-2O in mice), is produced in thymic epithelial cells and B cells. This molecule is a heterodimer of the HLA-DNα chain and the HLA-DOβ chain (see [Fig. 5.10](https://www.ncbi.nlm.nih.gov/books/NBK27137/figure/A570/?report=objectonly" \t "object)). HLA-DO is not present at the cell surface, being found only in intracellular vesicles, and it does not appear to bind peptides. Instead, it acts as a negative regulator of [HLA-DM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2880/), binding to it and inhibiting both the HLA-DM-catalyzed release of CLIP from, and the binding of other peptides to, MHC class II molecules. Expression of the HLA-DOβ chain is not increased by interferon-γ ([IFN](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2902/)-γ), whereas the expression of HLA-DM is. Thus, during inflammatory responses, in which IFN-γ is produced by [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) and [NK cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3106/) (see [Section 2-25](https://www.ncbi.nlm.nih.gov/books/n/imm/A203/#A219)), the increased expression of HLA-DM is able to overcome the inhibitory effects of HLA-DO. Why the [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/)-presenting ability of thymic epithelial cells and of B cells should be regulated in this way is not known; in thymic epithelial cells the purpose may be to select developing [CD4 T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2669/) by using a repertoire of self peptides different from those to which they will be exposed as mature T cells.

The role of the [HLA-DM](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2880/) molecule in facilitating the binding of peptides to MHC class II molecules parallels that of the TAP molecules in facilitating peptide binding to MHC class I molecules. Thus it seems likely that specialized mechanisms of delivering peptides have coevolved with the [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) themselves. It is also likely that pathogens have evolved strategies to inhibit loading of peptides onto MHC class II molecules, much as viruses have found ways of subverting [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) processing and presentation through the MHC class I molecules.

[Go to:](https://www.ncbi.nlm.nih.gov/books/NBK27137/)

## 5-8. Stable binding of peptides by MHC molecules provides effective antigen presentation at the cell surface

For [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) to perform their essential function of signaling intracellular infection, the peptide:MHC complex must be stable at the cell surface. If the complex were to dissociate too readily, the pathogen in the infected cell could escape detection. In addition, MHC molecules on uninfected cells could pick up peptides released by MHC molecules on infected cells and falsely signal to [cytotoxic T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2755/) that a healthy cell is infected, triggering its unwarranted destruction. The tight binding of peptides by MHC molecules makes both of these undesirable outcomes unlikely.

The persistence of a peptide:MHC complex on a cell can be measured by its ability to stimulate [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/), while the fate of the [MHC molecules](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3065/) themselves can be directly followed by specific staining. In this way it can be shown that specific peptide:MHC complexes on living cells are lost from the surface and re-internalized as part of natural protein turnover at the same rate as the MHC molecules themselves, indicating that peptide binding is essentially irreversible. This stable binding enables even rare peptides to be transported efficiently to the cell surface by MHC molecules, and allows long-term display of these complexes on the surface of the infected cell. This fulfills the first of the requirements for effective [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) presentation.

The second requirement is that if a peptide should dissociate from a cell-surface MHC molecule, peptides from the surrounding extracellular fluid would not be able to bind to the empty peptide-binding site. In fact, removal of the peptide from a purified MHC class I molecule has been shown to require denaturation of the protein. When peptide dissociates from an MHC class I molecule at the surface of a living cell, the molecule changes conformation, the β2-microglobulin dissociates, and the α chain is internalized and rapidly degraded. Thus, most empty MHC class I molecules are quickly lost from the cell surface.

At neutral pH, empty MHC class II molecules are more stable than empty MHC class I molecules, yet empty MHC class II molecules are also removed from the cell surface. Empty MHC class II molecules aggregate readily, and internalization of such aggregates may account for the loss of empty MHC class II molecules from the surface of cells. Moreover, peptide loss from MHC class II molecules is most likely when the molecules transit through acidified [endosomes](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2794/) as part of the normal process of cell membrane recycling. At acidic pH, MHC class II molecules are able to bind peptides that are present in the vesicles, but those that fail to do so are rapidly degraded.

Thus, both MHC class I and class II molecules are effectively prevented from acquiring peptides from the surrounding extracellular fluid. This ensures that [T cells](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A3278/) act selectively on infected cells or on cells specialized for [**antigen**](https://www.ncbi.nlm.nih.gov/books/n/imm/A2528/def-item/A2579/) uptake and display, while sparing surrounding healthy cells.