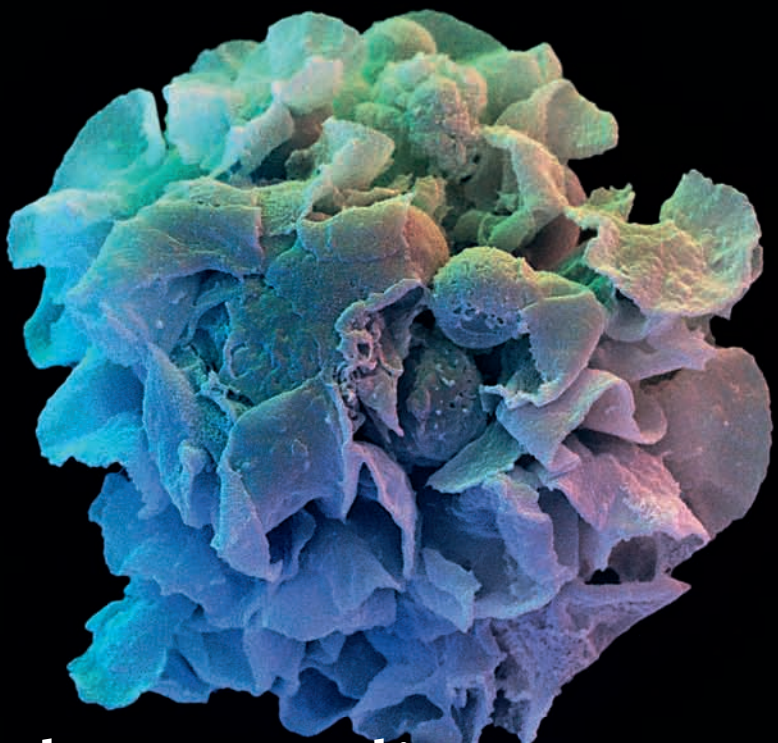
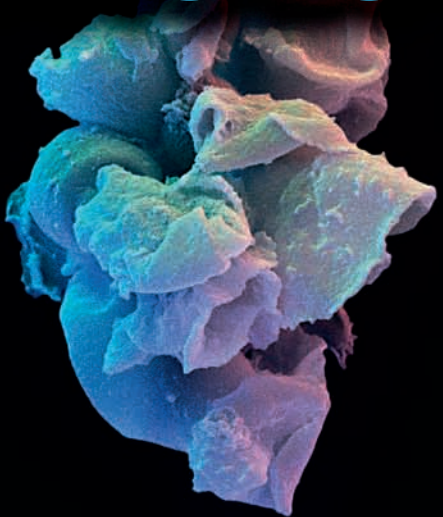
PETER WOOD



Understanding IMMUNOLOGY

THIRD EDITION



Understanding Immunology

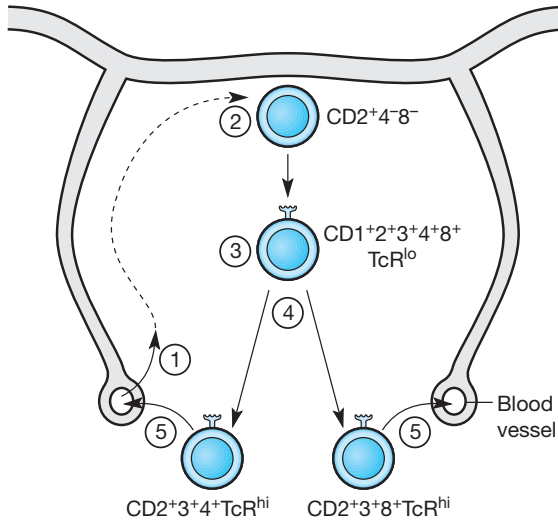


Figure 5.5 T cell development in the thymus. Cells enter the thymus from venules at the cortico-medullary junction (1) and travel to the subcapsular area of the outer cortex (2). These cells, which do not express CD4 or CD8, divide rapidly for a while and then stop dividing and migrate into the cortex (3), where they express both CD4 and CD8. As the cells traverse from the cortex to the medulla they stop expressing CD4 or CD8 (4) and leave the medulla as naïve T cells through blood vessels at the cortico-medullary junction (5).

5.4 During their development lymphocytes must generate huge numbers of Ig and TCR receptors with different antigen specificities

One vital aspect of B and T cell differentiation is that the cells must generate antigen receptors, Ig or TCR, for huge numbers of different antigens. As described above, estimates are that in excess of 10^{11} different antigen specificities are required and these must be generated from a much smaller number of genes. So how is this achieved? Although the genes coding Ig H and L chains and TCR α and β chains are different, the B and T cells use the same molecular mechanisms to convert a limited number of genes into a huge variety of protein products. This is achieved through the specialised structure of the Ig and TCR genes and a unique molecular mechanism of Ig and TCR gene rearrangement.

5.4.1 Ig and TCR genes have special structures

Antibody molecules consist of two heavy chains and two light chains joined together by disulphide bonds (see Chapter 3). The TCR consists of an α-chain and a β-chain (see Chapter 4). The locations of the Ig and TCR genes are shown in Table 5.1. Each Ig or TCR protein chain has a variable (V) region and a constant (C) region. Like all other proteins, Ig and TCRs

are coded for by genes that are transcribed and translated into the final protein. However there are unique features about the Ig and TCR genes that enable large numbers of different proteins to be generated from a limited number of genes. These features apply to all Ig and TCR genes but can best be explained by describing the human Ig κ-light chain gene, which has the simplest structure.

Table 5.1 Location of Ig and TCR genes

Gene	Chromosome
IgH	14
κ	2
λ	22
TCRα	14
TCRβ	7

[a] The TCRδ genes are located between those coding for Vα and J.

Gene and protein structure of the human κ-light chain

At the protein level the κ-light chain consists of a variable region domain and a constant region domain as described in Chapter 3. The genes coding for the κ-light chain consist of exons, which are sequences of DNA coding for the protein, and introns, which are parts of the gene that do not code for the protein. In this respect the κ-light chain gene is similar to most other eukaryotic genes. Figure 5.6 shows the structure of a typical gene, that for α-globin, as well as the κ-light chain gene. Superficially they appear to be similar in that they both contain introns and exons. The α-globin gene has been chosen for its simplicity; it contains three exons and two introns. Other genes can contain dozens of introns and exons. The mRNA transcripts for α-globin are shown in Figure 5.6 and it can be seen that all of the exons are used to code for the final protein sequence. This is typical for most genes although sometimes the primary RNA can be processed in slightly different ways – so-called differential processing – to give different forms of the protein. However, this differential splicing results in only a few different forms of the protein.

If we now look at the κ-light chain gene, two important features are seen that distinguish this gene from typical genes. At the protein level the κ-chain consists of a variable and a constant region. A single exon codes for the constant region of the κ-chain; this is often referred to as the '*C*κ gene' although strictly speaking it is a gene segment. The variable region is coded by two exons. Most of the variable region is coded for by a single exon called the κ variable gene, or Vκ (Figure 5.6). Part of the CDR3 variable region is also coded for by the Vk gene but the part closest to the constant κ region is coded for by a separate gene segment called the J gene, or Jκ. The J stands for 'joining' because the J segment joins the V and C segments. Each of the 76 Vκ genes has a different nucleotide sequence, especially in the CDRs; similarly, the 5 Jκ genes have different nucleotide sequences.

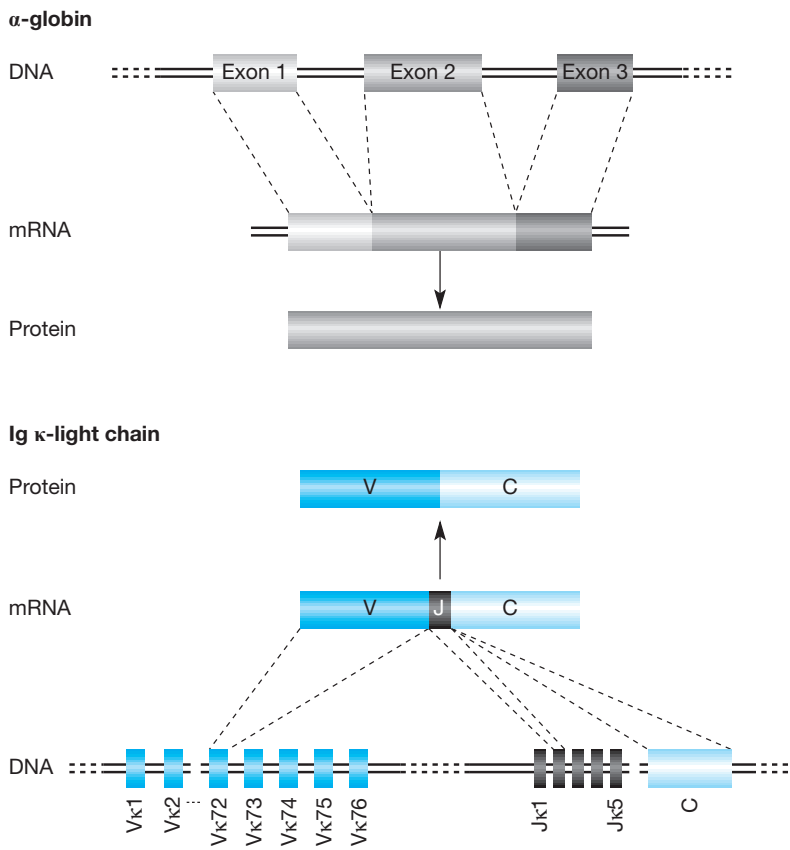


Figure 5.6 Structure of genes coding for α-globin and κ-light chain. All three exons of the α-globin gene are translated into RNA, which is processed and translated into protein. By contrast, only one of 76 Vκ genes and one of five Jκ genes are transcribed into RNA, along with the Cκ gene.

5.4.2 Each developing B cell chooses to use only one of its V and one of its J kappa genes

Although the developing B cell has a large number of V genes (or exons) and, to a lesser extent, J genes, **each B cell uses only one of the V genes and one of the J genes**. As shown in Figure 5.6, the DNA of the gene in a B cell has one of its 76 V genes joined to one of its five J genes and it is from this VJ DNA sequence that RNA will be transcribed. This 'choosing' to utilise only one out of the many exons occurs because of a specialised molecular process called **gene rearrangement** that occurs during the development of the B cell (Figure 5.7). This gene rearrangement involves the DNA encoding the κ-chain folding so that one of the V-region genes is positioned next to one of the J-region genes. The DNA is then cut so

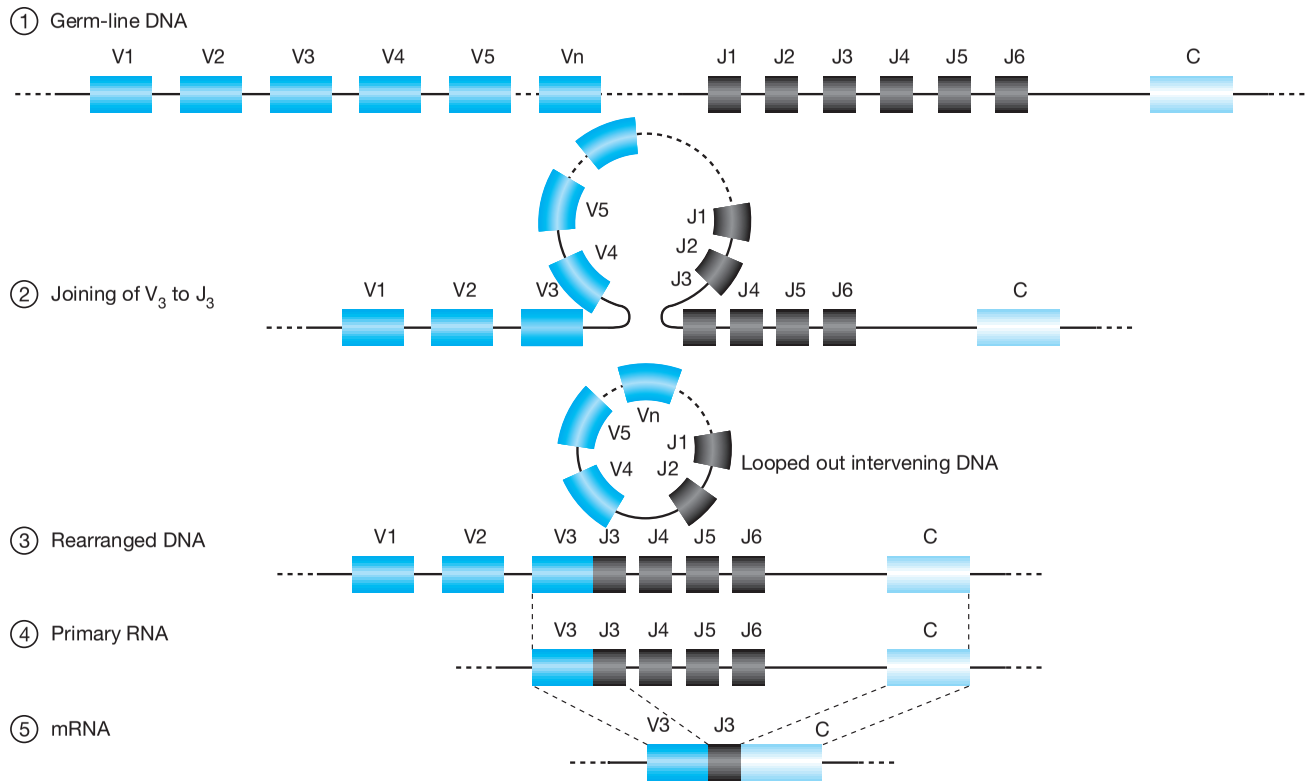


Figure 5.7 Ig and TcR gene rearrangement. Each T cell or B cell can use one V gene and one J gene from the germ-line DNA ①. In this example V3 is joining to J3 during the gene rearrangement ②. The DNA between V3 and J3 is looped out and excised, leaving V3 joined to J3 ③. All the DNA from V3 to C in the rearranged gene is transcribed to give the primary transcript ④ and this is processed to give the mRNA ⑤, which is translated into protein.

During their development lymphocytes must generate huge numbers of Ig and TCR receptors with different antigen specificities

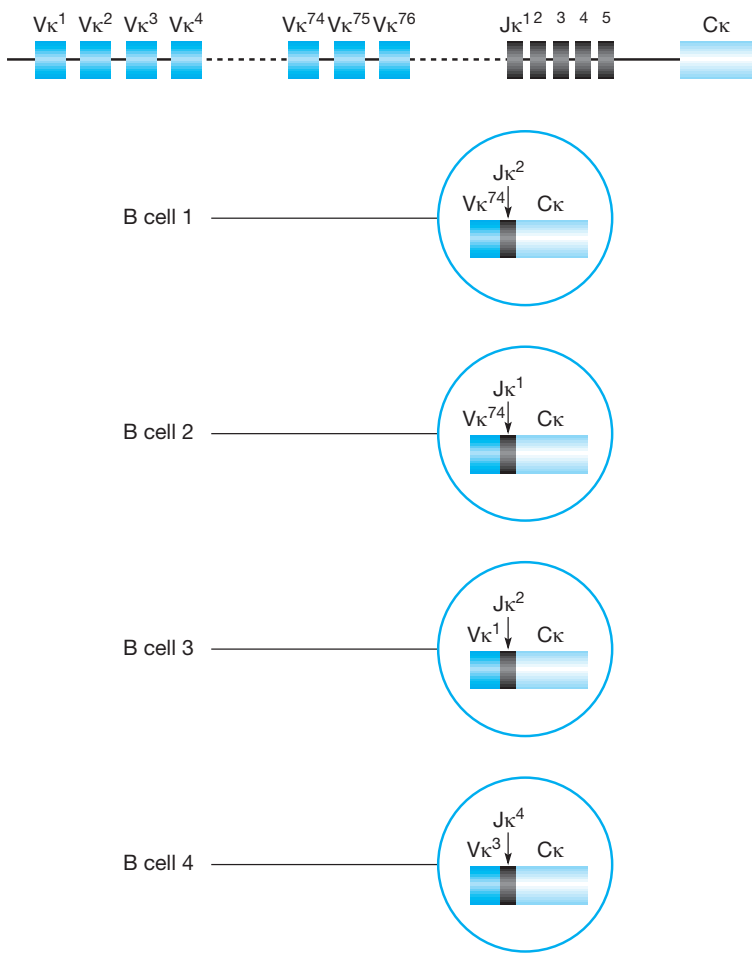


Figure 5.8 Different B cells utilise different Ig gene segments. Using the Ig κ-light chain as an example, there are 76 genes coding for Vκ and five genes coding for Jκ. Each B cell chooses at random one functional Vκ gene and one Jκ gene; the actual genes chosen will differ for each B cell. Therefore B cell 1 has chosen the Vκ74 and Jκ2. B cell 2 has chosen the same V, Vκ74, but a different J, Jκ1, and it can be seen that all four B cells have a different combination of V and J regions. Each V and J region has a unique nucleotide sequence and therefore each combination of V and J gives rise to a V-region protein with a slightly different amino acid sequence, and therefore different antigen specificity, from any other combination of V and J.

that the V gene can be joined to the J gene. The gene is now said to be rearranged and the rearranged DNA can be transcribed and translated into the κ-chain protein. Although each individual B cell uses only one V and one J segment, different B cells will 'choose' different V and J segments at random. Therefore the population of B cells in an individual will utilise all the different V and J segments (Figure 5.8).

The mechanism by which both the J gene and V genes code for the third complementarity-determining region (CDR3) of the variable region is important in increasing the variety of antigen specificities. We saw in Chapter 3 how Ig chains contain three regions that differ extensively in amino acid sequence in different antibodies. These regions contribute to the antigen-binding site. CDR1 and CDR2 are coded for exclusively by the V gene segment but will differ in different B cells because they will select different V gene segments. CDR3 is coded for partially by the V gene and partially by the J gene (Figure 5.9); therefore both the V gene and the J gene that a B cell selects will influence the antigen specificity of the Ig made by the B cell. This feature is very important in increasing the variety of amino acid sequences in CDR3.

Since there are 76 Vk genes and 5 Jk genes, the number of combinations V and J is 76 × 5 = 380. Therefore this mixing and matching gives 380 combinations (i.e. different DNA sequences have from 81 genes (76V + 5J). You can see already how the use of different gene segments increases variability.

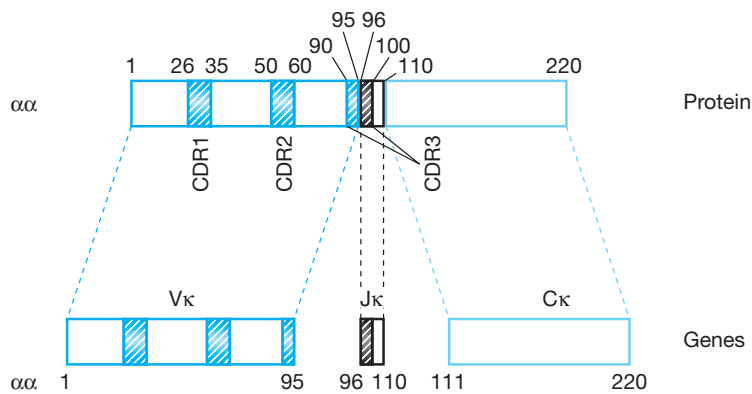


Figure 5.9 Relationship between Ig κ gene and protein. The Vκ-gene codes for the first 95 amino acids of V region of the protein. The Jκ gene codes for the final 15 amino acids of the V region of the protein. The CDR1 and 2 are coded for totally by the Vκ gene but the CDR3 is coded for partially by the Vκ gene and partially by the Jκ gene. Therefore different combinations of V and J will give different nucleotide sequences coding for CDR3 and therefore different amino acid sequences at CDR3. Because CDR1, 2 and 3 fold in the antibody molecule to form the antigen-binding site, the different amino acid sequences at CDR3, formed by different combinations of Vκ and Jκ genes, will have different antigen specificities.