4 Antibodies

- **Circulating antibodies** recognize antigen in serum and tissue fluids.
- **There are five classes of antibody – IgG, IgA, IgM, IgD and IgE**.
- **Immunoglobulins have a basic unit of two light chains and two heavy chains in a light-heavy-heavy-light arrangement.** The heavy chains differ between classes. IgA and IgM occur as oligomers of the four-chain unit.
- **The chains are folded into discrete regions called domains.** There are two domains in the light chain and four or five in the heavy chain, depending on their class.
- **Hypervariable regions form the antigen-binding sites.** There are three such regions in the V domains of each light and heavy chain. The folding of the domains causes them to be clustered at the distal tips of the molecule, producing two antigen-binding sites for each four-chain unit.
- **All antibodies are bifunctional.** They exhibit one or more **effector functions** in addition to antigen binding. These biological activities (e.g. complement activation and cell binding) are localized to sites that are distant from the antigen-binding sites (mostly in the Fc region).
- **Receptors for immunoglobulins** are expressed by mononuclear cells, neutrophils, natural killer (NK) cells, eosinophils and mast cells. They interact with the Fc regions of different classes of immunoglobulins and promote activities such as phagocytosis, tumour cell killing and mast cell degranulation. Most of the Fcγ receptors are members of the **immunoglobulin superfamily** and have two or three extracellular immunoglobulin domains.
- **Antibodies are highly specific for the three-dimensional conformation of the epitope.**
- **Antibody affinity** is a measure of the strength of the bond between an antibody's combining site and a single epitope. The functional affinity or avidity of the interaction additionally depends on the number of binding sites on the antibody and their ability to react with multiple epitopes on the antigen.
- **The immune system is able to recognize and respond to many antigens** by generating great diversity in the antibodies produced by the B cells.
- **Immunoglobulin light chains are encoded by V and J gene segments**; heavy chains are also encoded by V and J gene segments with additional diversity provided by the D gene segment.
- **Diversity is achieved by the recombination of a limited number of V, D and J gene segments** to produce a vast number of variable domains.
- **Immunoglobulin heavy and light chains** undergo structural modifications called somatic mutation following antigen stimulation which fine tunes their affinity.
- **Recombination of V, D and J gene segments** of immunoglobulins is controlled, at least in part, by two recombination activating genes (RAG-1 and RAG-2).
- **In addition to simple combinations of V, D and J regions**, diversity in immunoglobulins depends upon N-region diversification, joining-site variation and multiple D regions.
- **Immunoglobulin class-switching involves recombination of VDJ genes** with various C region genes and differential RNA splicing.

The recognition of foreign antigen is the hallmark of the specific adaptive immune response. Two distinct types of molecules are involved in this process – the immunoglobulins and the T-cell antigen receptors (TCRs). Diversity and heterogeneity are characteristic features of these molecules. In both cases there is evidence of extensive gene rearrangements which generate immunoglobulins or TCRs capable of recognizing many different antigens. T-cell receptors are discussed in detail in Chapter 5.

The immunoglobulins are a group of glycoproteins present in the serum and tissue fluids of all mammals. Some are carried on the surface of B cells, where they act as receptors for specific antigens. Others (antibodies) are free in the blood or lymph. Contact between B cells and antigen is needed to cause the B cells to develop into antibody forming cells (AFCs), also called plasma cells, which secrete large amounts of antibody. ('Plasma cell' is the original histological term used to describe AFCs seen in blood and tissues.) The membrane-bound immunoglobulin on a precursor B cell has the same binding specificity as the antibody produced by the mature AFC (*Fig. 4.1*).

IMMUNOGLOBULINS – A FAMILY OF PROTEINS

Five distinct classes of immunoglobulin molecule are recognized in most higher mammals, namely IgG, IgA, IgM, IgD and IgE. They differ in size, charge, amino acid composition and carbohydrate content. In addition to the difference between classes, the immunoglobulins within each class are also very heterogeneous. Electrophoretically the immunoglobulins show a unique range of heterogeneity which extends from the γ to the α fractions of normal serum (*Fig. 4.2*).

Immunoglobulins are bifunctional molecules

Each immunoglobulin molecule is bifunctional. One region of the molecule is concerned with binding to antigen while a different region mediates so-called effector functions. Effector functions include binding of the immunoglobulin to host tissues, to various cells of the immune system, to some phagocytic cells, and to the first component (C1q) of the classical complement system.

Surface and secreted antibodies

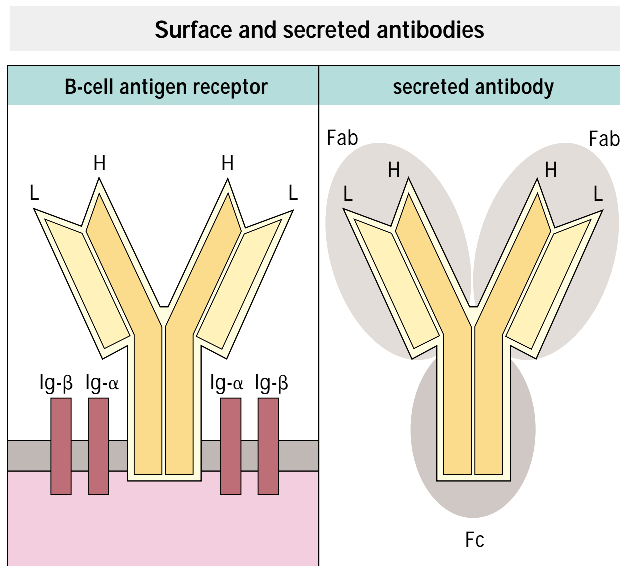


Fig. 4.1 The B-cell antigen receptor (left) consists of two identical heavy (H) chains and two identical light (L) chains. In addition, secondary components (Ig-α and Ig-β) are closely associated with the primary receptor and are thought to couple it to intracellular signalling pathways. Circulating antibodies (right) are structurally identical to the primary B-cell antigen receptors, except that they lack the transmembrane and intracytoplasmic sections. Many proteolytic enzymes cleave antibody molecules into three fragments – two identical Fab (antigen binding) fragments and one Fc (crystallizable) fragment.

Distribution of the major human immunoglobulins

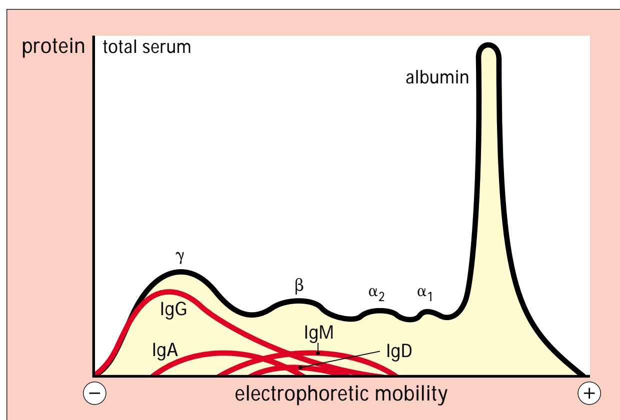


Fig. 4.2 Electrophoresis of human serum showing the distribution of the four major immunoglobulin classes. Serum proteins are separated according to their charges in an electric field, and classified as α_1, α_2, β and γ, depending on their mobility. (The IgE class has a similar mobility to IgD but cannot be represented quantitatively because of its low level in serum.) IgG exhibits the most charge heterogeneity, the other classes having a more restricted mobility in the β and fast γ regions.

Immunoglobulin class and subclass depends on the structure of the heavy chain

The basic structure of all immunoglobulin molecules is a unit consisting of two identical light polypeptide chains and two identical heavy polypeptide chains. These are linked together by disulphide bonds. The class and subclass of an immunoglobulin molecule are determined by its heavy chain type. Thus the four human IgG subclasses (IgG1, IgG2, IgG3 and IgG4) have heavy chains called γ1, γ2, γ3 and γ4 that differ slightly, although all are recognizably γ heavy chains.

The four subclasses of human IgG (IgG1–IgG4) occur in the approximate proportions of 66%, 23%, 7% and 4%, respectively. There are also known to be subclasses of human IgA (IgA1 and IgA2), but none have been described for IgM, IgD or IgE. This range of immunoglobulin class and subclass reflects isotypic variations in the immunoglobulin genes (see below). Immunoglobulin subclasses appear to have arisen late in evolution. Thus, the human IgG subclasses are very different from the four known subclasses of IgG that have been identified in the mouse. All immunoglobulins are glycoproteins, but the carbohydrate content ranges from 2–3% for IgG, to 12–14% for IgM, IgD and IgE. The physicochemical properties of the immunoglobulins are summarized in *Figure. 4.3*.

IgG – The major immunoglobulin in normal human serum, accounting for 70–75% of the total immunoglobulin pool, IgG consists of a single four-chain molecule with a sedimentation coefficient of 7S and a molecular weight of 146 000. However, IgG3 proteins are slightly larger than the other subclasses; due to the slightly heavier γ3 chain.

IgM – Accounts for approximately 10% of the immunoglobulin pool. The molecule is a pentamer of the basic four-chain structure. The individual heavy chains have a molecular weight of approximately 65 000 and the whole molecule has a molecular weight of 970 000.

IgA – Represents 15–20% of the human serum immunoglobulin pool. In humans more than 80% of IgA occurs as a monomer of the four-chain unit, but in most mammals the IgA in serum is mainly polymeric, occurring mostly as a dimer. IgA is the predominant immunoglobulin in seromucous secretions such as saliva, colostrum, milk, and tracheobronchial and genitourinary secretions. Secretory IgA (s-IgA), may be of either subclass (IgA1 or IgA2), exists mainly in the 11S, dimeric form and has a molecular weight of 385 000 due to its association with another protein, known as the secretory component.

IgD – Accounts for less than 1% of the total plasma immunoglobulin but is a major component of the surface membrane of many B cells.

IgE – Though scarce in serum, is found on the surface membrane of basophils and mast cells in all individuals.

Physicochemical properties of human immunoglobulin classes

property	immunoglobulin type									
	IgG1	IgG2	IgG3	IgG4	IgM	IgA1	IgA2	sIgA	IgD	IgE
heavy chain	γ_1	γ_2	γ_3	γ_4	μ	α_1	α_2	α_1/α_2	δ	ϵ
mean serum conc. (mg/ml)	9	3	1	0.5	1.5	3.0	0.5	0.05	0.03	0.00005
sedimentation constant	7s	7s	7s	7s	19s	7s	7s	11s	7s	8s
mol. wt ($\times 10^3$)	146	146	170	146	970	160	160	385	184	188
half-life (days)	21	20	7	21	10	6	6	?	3	2
% intravascular distribution	45	45	45	45	80	42	42	trace	75	50
carbohydrate (%)	2–3	2–3	2–3	2–3	12	7–11	7–11	7–11	9–14	12

Fig. 4.3 Each immunoglobulin class has a characteristic type of heavy chain. Thus IgG posesses γ chains; IgM, μ chains; IgA, α chains; IgD, δ chains; and IgE, ε chains. Variation in heavy chain structure within a class gives rise to immunoglobulin subclasses. For example, the human IgG pool consists of four subclasses reflecting four distinct types of heavy chain. The properties of the immunoglobulins vary between the different classes. Note that in secretions, IgA occurs in a dimeric form (s-IgA) in association with a protein chain termed the secretory component. The serum concentration of s-IgA is very low, whereas the level in intestinal secretions can be very high.

The basic structure of IgG1

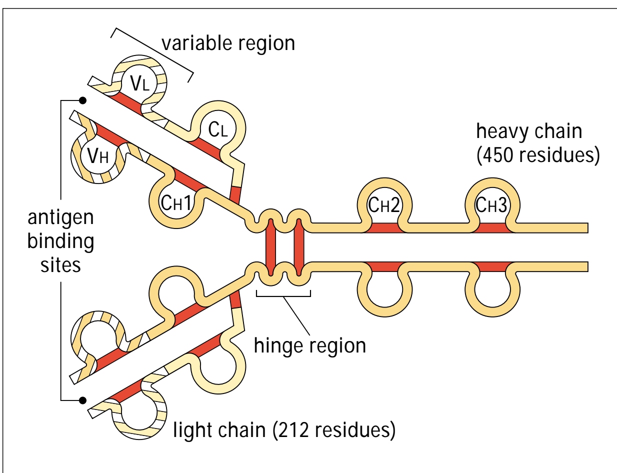


Fig. 4.4 The N-terminal end of IgG1 is characterized by sequence variability (V) in both the heavy and light chains, referred to as the VH and VL regions respectively. The rest of the molecule has a relatively constant (C) structure. The constant portion of the light chain is termed the CL region. The constant portion of the heavy chain is further divided into three structurally discrete regions: CH1, CH2 and CH3. These globular regions, which are stabilized by intrachain disulphide bonds, are referred to as 'domains'. The sites at which the antibody binds antigen are located in the variable domains. The hinge region is a segment of heavy chain between the CH1 and CH2 domains. Flexibility in this area permits the two antigen-binding sites to operate independently. There is close pairing of the domains except in the CH2 region (see *Fig. 4.6*). Carbohydrate moieties are attached to the CH2 domains.

The basic four-chain model for immunoglobulin molecules (*Fig. 4.4*) is based on two distinct types of polypeptide chain. The smaller (light) chain has a molecular weight of 25 000 and is common to all classes, whereas the larger (heavy) chain has a molecular weight of 50 000–77 000 and is structurally distinct for each class or subclass. The polypeptide chains are linked together by covalent and non-covalent forces.

All light chains have one variable and one constant region

The light chains of most vertebrates have been shown to exist in two distinct forms called kappa (κ) and lambda (λ). These are isotypes, being present in all individuals. Either of the light chain types may combine with any of the heavy chain types, but in any one immunoglobulin molecule both light chains and both heavy chains are of the same type.

Hilschmann, Craig and others in 1965 established that light chains consist of two distinct regions. The C-terminal half of the chain (approximately 107 amino acid residues) is constant except for certain allotypic and isotypic variations (see *Fig. 4.27*) and is called the CL (Constant : Light chain) region, whereas the N-terminal half of the chain shows much sequence variability and is known as the VL (Variable : Light chain) region.

IgG has a 'typical' antibody structure

The IgG molecule may be thought of as a 'typical' antibody (Fig. 4.4). It has two intrachain disulphide bonds in the light chain – one in the variable region and one in the constant region (*Fig. 4.5*). There are four such bonds in the heavy (γ) chain, which is twice the length of the light chain. Each disulphide bond encloses a peptide loop of 60–70 amino acid residues; if the amino acid sequences of these loops are compared a striking degree of homology is revealed. Essentially this means that each immunoglobulin peptide chain is composed of a series of globular regions with very similar secondary and tertiary structure (folding). This is shown for the light chain in *Figure 4.5*.

The peptide loops enclosed by the disulphide bonds represent the central portion of a 'domain' of about 110 amino acid residues. In both the heavy and the light chains

Basic folding in the light chain

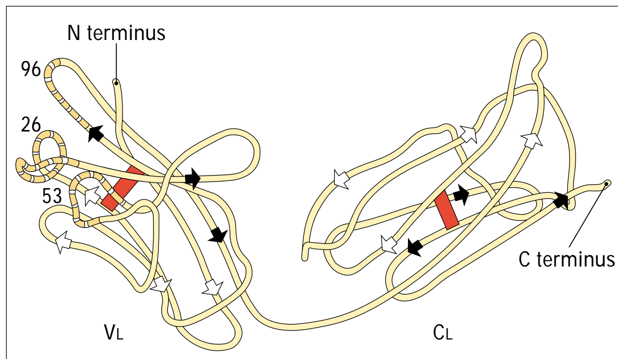


Fig. 4.5 The immunoglobulin domains in the light chain share a basic folding pattern with several straight segments of polypeptide chain lying parallel to the long axis of the domain. Light chains have two domains – one constant and one variable. Within each domain, the polypeptide chain is arranged in two layers, running in opposite directions, with many hydrophobic amino acid side-chains between the layers. One of the layers has four segments (arrowed white), the other has three (arrowed black); both are linked by a single disulphide bond (red). Folding of the VL domains causes the hypervariable regions to become exposed in three separate but closely disposed loops. One numbered residue from each hypervariable region is identified.

the first of these domains corresponds to the variable region, VH and VL respectively. In the heavy chain of IgG, IgA and IgD there are three further domains, which make up the constant part of the chain, CH1, CH2 and CH3. In both μ and ε chains there is an additional domain immediately after CH1 (*Fig. 4.6*). Thus, the C-terminal domains of IgM and IgE heavy chains (referred to as Cμ4 and Cε4) are homologous to the CH3 domain of IgG (Cγ3).

X-ray crystallography has provided structural data on complete IgG molecules, making it possible to construct both α-carbon backbone and computer-generated atomic models for this class of immunoglobulin (*Fig. 4.7*). These show the Y-shaped and T-shaped structures that have also been visualized by electron microscopy.

Homologous domains of the light and heavy chains are paired in the Fab region (indicated in *Fig. 4.6*). The Cγ3 domains of the γ heavy chains are also paired, but the Cγ2 domains are separated by carbohydrate moieties. Despite the structural similarities between domains there are striking differences at the level of domain interaction. For example, the variable domains associate with each other through their three-segment layers, whereas the constant domains associate through their four-segment layers. (See *Fig. 4.5* for an explanation of the layers in light chain domains.)

IgG – With human IgG, the four subclasses differ only slightly in their amino acid sequences. Most of the differences are clustered in the hinge region and give rise to differing patterns of interchain disulphide bonds between the four proteins. The most striking structural difference is the elongated hinge region of IgG3, which accounts for its higher molecular weight and possibly for some of its enhanced biological activity (*Fig. 4.6(2)*).

IgM – Human IgM is usually found as a pentamer of the basic four-chain unit (*Fig. 4.6(3)*). The μ chains of IgM differ from γ chains in amino acid sequence and have an extra constant region domain. The subunits of the pentamer are linked by disulphide bonds between the Cμ3 domains, and possibly by disulphide bonds between the C-terminal 18-residue peptide tailpieces. The complete molecule consists of a densely packed central region with radiating arms, as seen in electron micrographs.

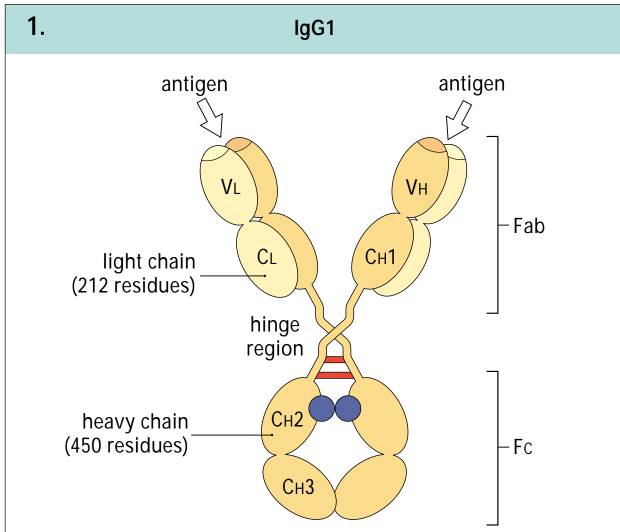
Photographs of IgM antibodies binding to bacterial flagella show molecules adopting a ‘crab-like’ configuration (*Fig. 4.8*). This suggests that flexion readily occurs between the Cμ2 and Cμ3 domains, although note that this region is not structurally homologous to the IgG hinge. The dislocation resulting in the ‘crab-like’ configuration appears to be related to the activation of complement by IgM.

Two other features characterize the IgM molecule: an abundance of oligosaccharide units associated with the μ chain, and an additional peptide chain, the J (joining) chain, thought to assist the process of polymerization prior to secretion by the AFC. The J chain is an Ig-like domain of 137 amino acid residues. One J chain is incorporated into the IgM structure by disulphide bonding to the 18-residue peptide tailpiece of the separate monomers. Binding is to the penultimate cysteine residues of the tailpieces. If J chains are not freely available, there is evidence that hexameric IgM becomes the preferred form.

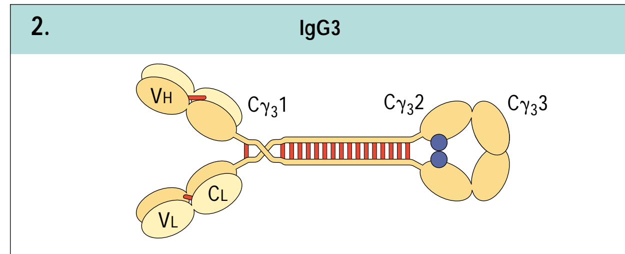
Fig. 4.6 In each of the structures shown the carbohydrate side-chains are shown in blue. Inter heavy chain disulphide bonds are shown in red but interchain bonds between H and L chains are omitted.
(**1**) A model of IgG1 indicating the globular domains of heavy (H) and light (L) chains. Note the apposition of the CH3 domains and the separation of the CH2 domains. The carbohydrate units lie between the CH2 domains.
(**2**) Polypeptide chain structure of human IgG3. Note the elongated hinge region.
(**3**) IgM heavy chains have five domains with disulphide bonds cross-linking adjacent Cμ3 and Cμ4 domains. The possible location of the J chain is shown. IgM does not have extended hinge regions, but flexion can occur about the Cμ2 domains.
(**4**) The secretory component of s-IgA is probably wound around the dimer and attached by two disulphide bonds to the Cα2 domain of one IgA monomer. The J chain is required to join the two subunits.
(**5**) This diagram of IgD shows the domain structure and a characteristically large number of oligosaccharide units. Note also the presence of a hinge region and short octapeptide tailpieces.
(**6**) IgE can be cleaved by enzymes to give the fragments $F(ab')_2$, Fc and Fc′. Note the absence of a hinge region.

Structural characteristics of various human immunoglobulins

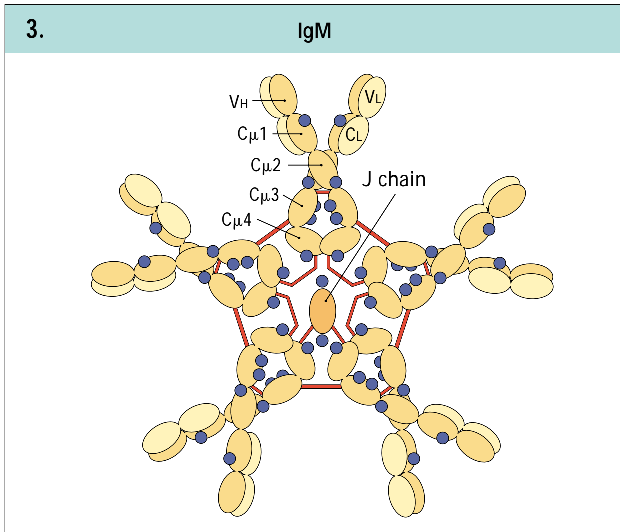
1. IgG1



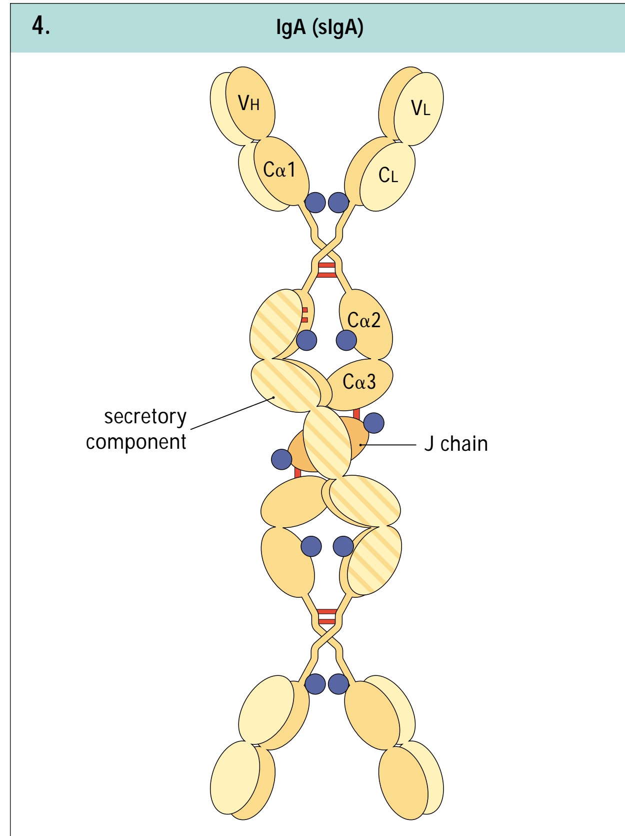
2. IgG3



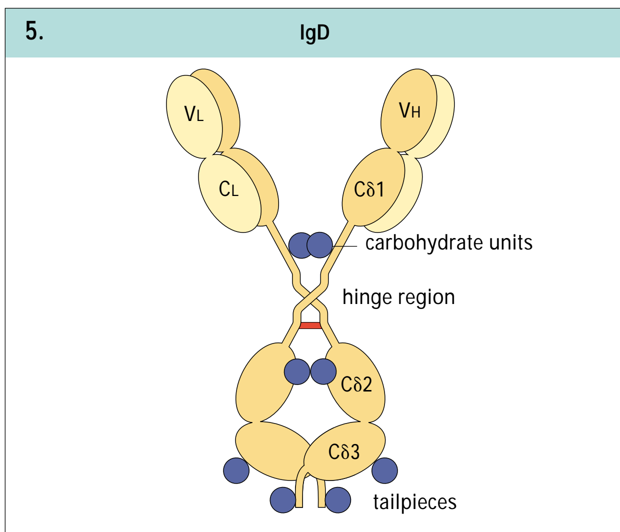
3. IgM



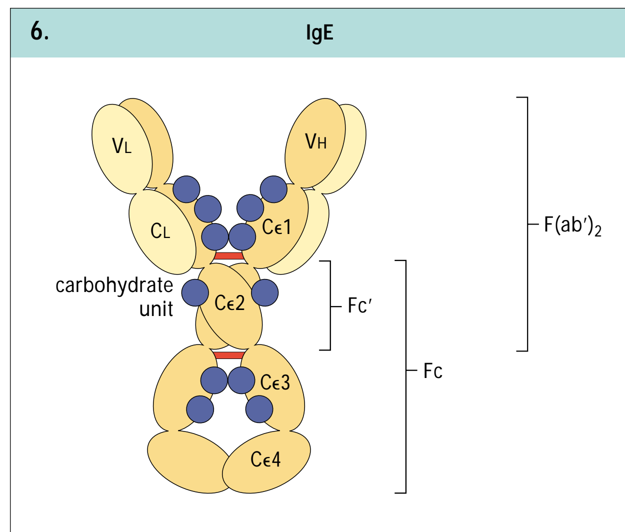
4. IgA (sIgA)



5. IgD



6. IgE



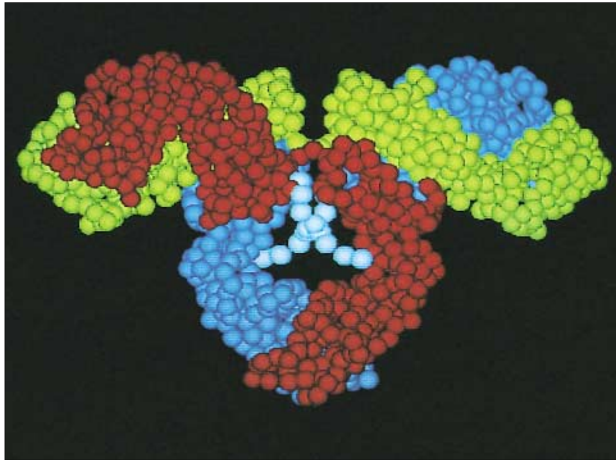


Fig. 4.7 Computer-generated model of the hinge-deleted human IgG1 protein Dob. Such proteins lack the flexibility characteristic of normal IgG molecules. Their rigidity permits structural determinations at a higher resolution. One heavy chain is shown in blue and one in red, with two light chains being depicted in green. Carbohydrate bound to the Fc portion of the molecule is shown in turquoise. The model suggests that interactions between the Cγ2 domains are weak whereas those between Cγ3 domains are strong. (The structure of this immunoglobulin was determined by David R. Davies *et al. Proc Natl Acad Sci USA* 1977; **74**; the computer graphics were generated using the system developed by Richard J. Feldmann at the National Institutes of Health.)

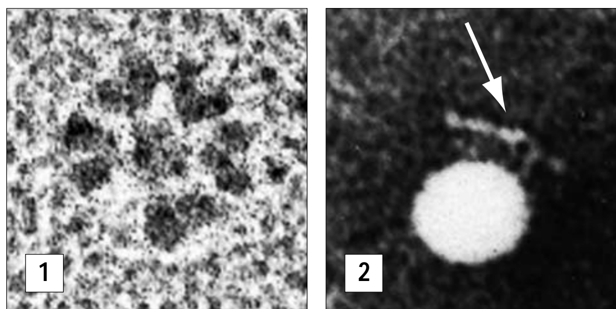


Fig. 4.8 Electron micrographs of IgM molecules. (**1**) In free solution, deer IgM adopts the characteristic star-shaped configuration. ×195 000. (Courtesy of Drs E. Holm Nielson P. Storgaard and Prof S-E. Svehag.) (**2**) Rabbit IgM antibody (arrowed) in 'crab-like' configuration with partly visible central ring structure bound to a poliovirus virion ×190 000. (Courtesy of Dr B. Chesebro and Professor S-E Svehag.)

IgA – The 472 amino acid residues of the α chain are arranged in four domains: VH, Cα1, Cα2 and Cα3. A feature shared with IgM is an additional C-terminal 18-residue peptide with a penultimate cysteine residue, which is able to bind covalently to a J chain to form dimers. Electron micrographs of IgA dimers show double Y-shaped structures, suggesting that the monomeric subunits are

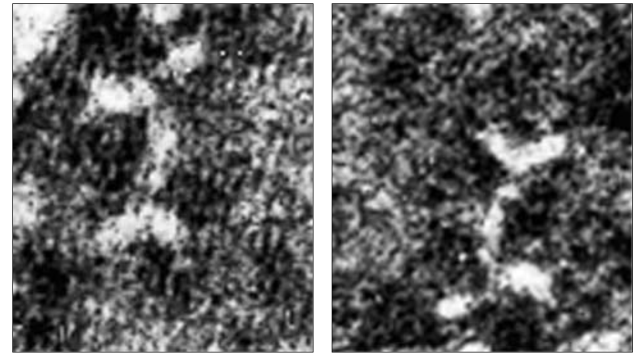


Fig. 4.9 Electron micrographs of human dimeric IgA molecules. The double Y-shaped appearance suggests that the monomeric subunits are linked end to end through the C-terminal Cα3 domain ×250 000. (Courtesy of Professor S-E. Svehag.)

linked end-to-end through the C-terminal Cα3 regions (*Fig. 4.9*).

Secretory IgA (s-IgA) exists mainly in the form of a molecule sedimenting at 11S (mol. wt 380 000). The complete molecule is made up of two units of IgA, one secretory component (mol. wt 70 000) and one J chain (mol. wt 15 000) (*Fig. 4.6(4)*). It is not clear how the various peptide chains are linked together. In contrast to the J chain, secretory component is not synthesized by plasma cells but by epithelial cells. IgA held in dimer configuration by a J chain, and secreted by submucosal plasma cells, actively binds secretory component as it traverses epithelial cell layers. Bound secretory component facilitates the transport of s-IgA into secretions, as well as protecting it from proteolytic attack.

IgD – Less than 1% of the total immunoglobulin in serum is IgD. This protein is more susceptible to proteolysis than IgG1, IgG2, IgA or IgM, and also has a tendency to undergo spontaneous proteolysis. There appears to be a single disulphide bond between the δ chains and a large amount of carbohydrate distributed in multiple oligosaccharide units (*Fig. 4.6(5)*).

IgE – The structure of IgE is shown in *Figure 4.6(6)*. The higher molecular weight of the ε chain (72 500) is explained by the larger number of amino acid residues (approximately 550) distributed over five domains (VH, Cε1, Cε2, Cε3 and Cε4).

Immunoglobulins are the prototypes of the immunoglobulin superfamily of molecules

The basic domain structure first identified in antibodies also occurs in a number of other molecules, including many cell surface molecules involved in immune reactions. Examples of such molecules are the adhesion molecules ICAM-1 and VCAM-1 (see Chapter 3), the T-cell antigen receptor and MHC molecules (see Chapter 5), as well as several receptors for antibodies, encountered later in this chapter. Such molecules are said to belong to the immunoglobulin supergene family (IgSF). The principle elements

of the domain are two β-pleated sheets arranged opposite each other and usually stabilized by one or more disulphide bonds. This is sometimes referred to as a β-barrel. The domain structure must have developed early in evolution and has been used since as the basis for different molecules as vertebrates have radiated and developed.

INTERACTION OF ANTIBODIES WITH ANTIGENS

Antibodies form multiple non-covalent bonds with antigen

X-ray crystallography studies of antibody V domains show that the hypervariable regions are clustered at the end of the Fab arms; particular residues in these regions interact specifically with antigen. The framework residues do not usually form bonds with the antigen. However, they are essential for producing the folding of the V domains and maintaining the integrity of the binding site. The binding of antigen to antibody involves the formation of multiple non-covalent bonds between the antigen and amino acids of the binding site. Considered individually, the attractive forces (hydrogen and electrostatic bonds, Van der Waals and hydrophobic forces) are weak by comparison with covalent bonds. However, the large number of interactions results in a large total binding energy.

The conformations of target antigen and binding site are complementary

The strength of a non-covalent bond is critically dependent on the distance (d) between the interacting groups. The force is proportional to $1/d^2$ for electrostatic forces, and to $1/d^7$ for Van der Waals forces. Thus the interacting groups must be close (in molecular terms) before these forces become significant (*Fig. 4.10*). In order for an antigenic determinant (epitope) and an antibody-combining site (paratope) to combine (*Fig. 4.11*), there must be suitable atomic groupings on opposing parts of the antigen and antibody, and the shape of the combining site must fit the epitope, so that several non-covalent bonds can form simultaneously. If the antigen and the combining site are complementary in this way, there will be sufficient binding energy to resist thermodynamic disruption of the bond. However, if electron clouds of the antigen and antibody overlap, steric repulsive forces come into play which are inversely proportional to the 12th power of the distance between the clouds: $F \propto 1/d^{12}$. These forces have a vital role in determining the specificity of the antibody molecule for a particular antigen, and its ability to discriminate between antigens, as any variation from the ideal complementary shape will cause a decrease in the total binding energy through increased repulsive forces and decreased attractive forces (*Fig. 4.11*). This is not to say that the antigen-binding site is completely inflexible; when antigen binds to antibody individual amino acid residues may become slightly displaced from their position in the free state. This is referred to as 'induced fit', but it will only occur when the energy gain in the overall antigen–antibody bond offsets that needed to induce the fit.

An examination of the interaction between lysozyme and the Fab of an antibody to lysozyme has shown that the antigen epitope and the binding site have complementary surfaces. These surfaces extend beyond the hypervariable regions. In total, 17 amino acid residues on the antibody contact 16 residues on the lysozyme molecule (*Fig. 4.12*). All the hypervariable regions contribute to the antibody-

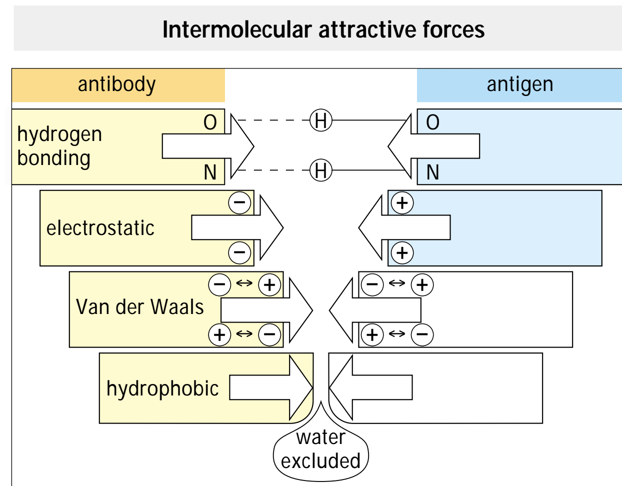


Fig. 4.10 The forces binding antigen to antibody require the close approach of the interacting groups. Hydrogen bonding results from the formation of hydrogen bridges between appropriate atoms. Electrostatic forces derive from the attraction of oppositely charged groups located on two protein side-chains. Van der Waals bonds are generated by the interaction between electron clouds (here represented as induced oscillating dipoles). Hydrophobic bonds (which may contribute up to half the total strength of the antigen–antibody bond) rely on the association of non-polar, hydrophobic groups so that contact with water molecules is minimized. The distance between the interacting groups that gives optimum binding depends on the type of bond.

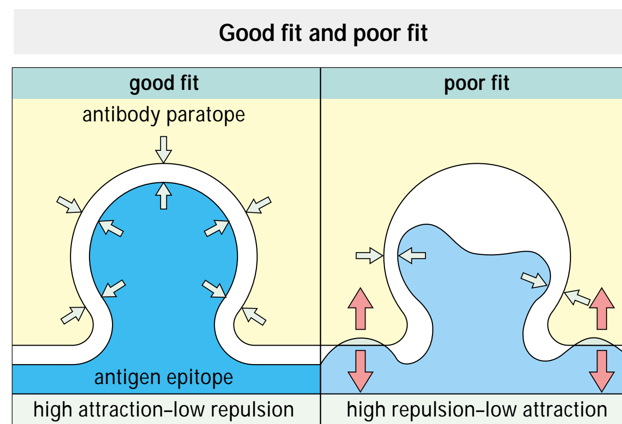


Fig. 4.11 A good fit between the antigenic determinant and the binding site of the antibody will create ample opportunities for intermolecular attractive forces to be created and few opportunities for repulsive forces to operate. Conversely, when there is a poor fit, the reverse is true. When electron clouds overlap, high repulsive forces are generated which override any small forces of attraction.

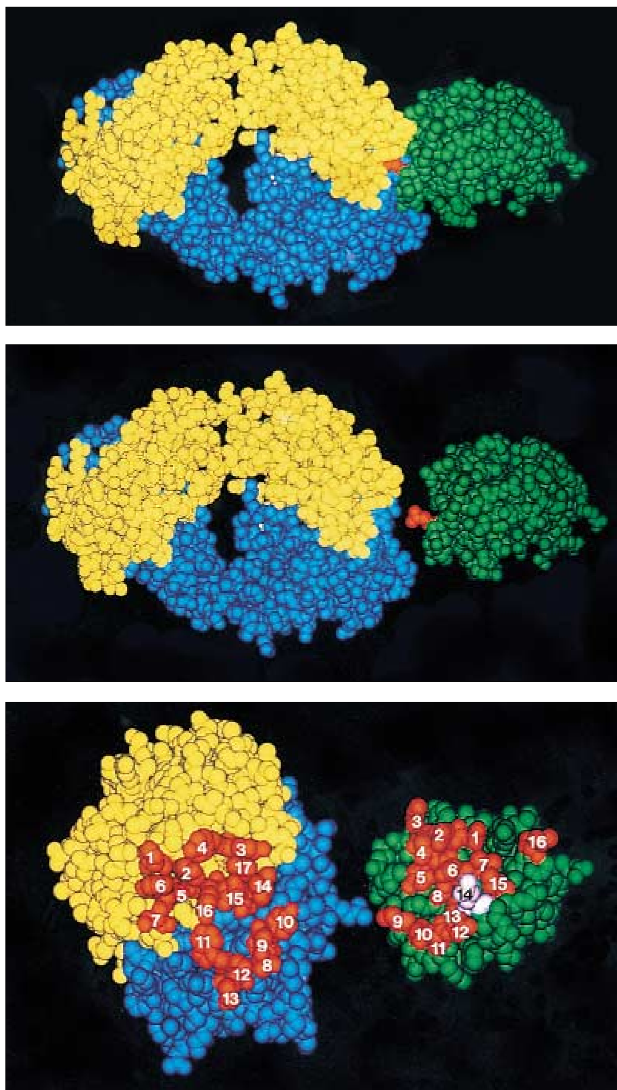


Fig. 4.12 The Fab–lysozyme complex. Upper: Lysozyme (green) binds to the hypervariable regions of the heavy (blue) and light (yellow) chains of the Fab fragment of antibody D1.3. Centre: The separated complex with Glu 121 visible (red). This residue fits into the centre of the cleft between the heavy and light chains. Lower: The same molecules rotated 90° to show the contact residues which contribute to the antigen–antibody bond. (Courtesy of Dr R. J. Poljak, from Science 1986;**233**:747–53, with permission.)

binding site, although the third hypervariable region, formed by the V–D–J junction in the heavy chain gene, lying at the centre of the combining site, appears to be most important. This may also be related to the greater variability generated by recombination of the V, D and J segments.

Antibody affinity indicates the strength of a single antigen–antibody bond

The strength of the bond between an antigen and an antibody is known as the antibody affinity. It is the sum of the attractive and repulsive forces described above (*Fig. 4.13*). Interaction of the antibody-combining site with antigen

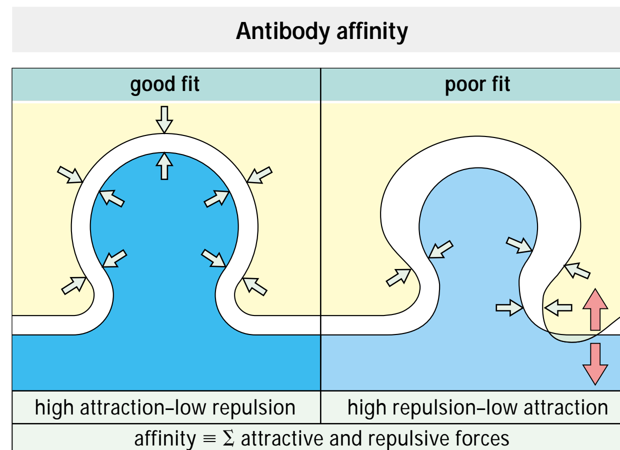


Fig. 4.13 The affinity with which antibody binds antigen is the sum of the attractive and repulsive forces between them. A high-affinity antibody implies a good fit and, conversely, a low-affinity antibody implies a poor fit.

can be investigated thermodynamically. To measure the affinity of a single combining site, it is necessary to use a monovalent antigen, or even a single isolated antigenic determinant (a hapten). Because the non-covalent bonds between antibody and epitope are dissociable, the overall combination of an antibody and antigen must be reversible; thus the Law of Mass Action can be applied to the reaction and the equilibrium constant, K, can be determined. This is the affinity constant (*Fig. 4.14*).

Antibody avidity indicates the overall strength of interaction between antibody and antigen

Because each antibody unit of four polypeptide chains has two antigen-binding sites, antibodies are potentially multivalent in their reaction with antigen. In addition, antigen can also be monovalent (e.g. haptens) or multivalent (e.g. microorganisms). The strength with which a multivalent antibody binds a multivalent antigen is termed avidity, to differentiate it from the affinity of a single antigenic determinant for an individual combining site. The avidity of

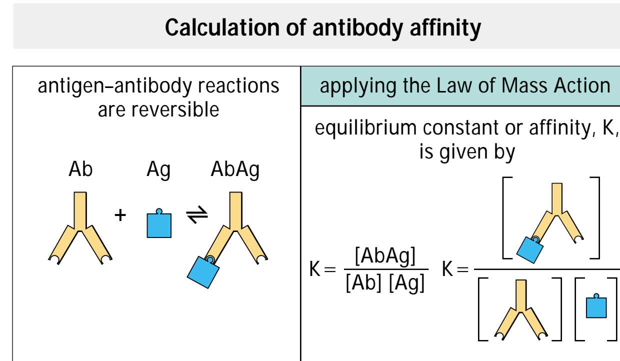
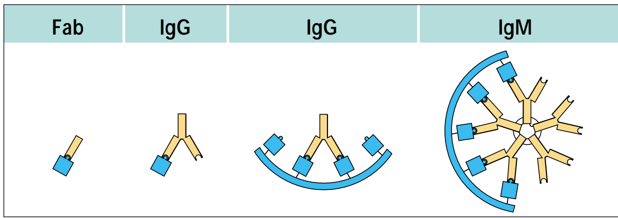


Fig. 4.14 All antigen–antibody reactions are reversible. The Law of Mass Action can therefore be applied, and the antibody affinity (given by the equilibrium constant, K) can be calculated. (Square brackets refer to the concentrations of the reactants.)

Affinity and avidity



antibody	Fab	IgG	IgG	IgM
effective antibody valence	1	1	2	up to 10
antigen valence	1	1	n	n
equilibrium constant (L/M)	10^4	10^4	10^7	10^{11}
advantage of multivalence	–	–	10^3-fold	10^7-fold
definition of binding	affinity	affinity	avidity	avidity
	intrinsic affinity		functional affinity	

Fig. 4.15 Multivalent binding between antibody and antigen (avidity or functional affinity) results in a considerable increase in stability as measured by the equilibrium constant, compared with simple monovalent binding (affinity or intrinsic affinity, here arbitrarily assigned a value of 10^4 L/M^{-1}). This is sometimes referred to as the 'bonus effect' of multivalency. Thus there may be a 10^3-fold increase in the binding energy of IgG when both valencies (combining sites) are utilized and a 10^7-fold increase when IgM binds antigen in a multivalent manner.

an antibody for its antigen is dependent on the affinities of the individual combining sites for the determinants on the antigen. It is greater than the sum of these affinities if both antibody-binding sites can combine with the antigen. This is because all the antigen–antibody bonds must be broken simultaneously before the antigen and antibody dissociate (*Fig. 4.15*). In normal physiological situations, avidity is likely to be more relevant than affinity, as naturally occurring antigens are multivalent. However, the precise measurement of hapten–antibody affinity is more likely to give an insight into the immunochemical nature of the antigen–antibody reaction.

Kinetics of antibody–antigen reactions

Measurements of antibody affinity relate to equilibrium conditions. Affinity indicates the tendency of the antibodies to form stable complexes with the antigen. However, for many biological activities of antibodies, it is possible that the kinetics of the reaction may also be significant. Kinetics measures the forward rate (on-rate) constant $K_{1,2}$ (mol^{-1}s^{-1}) and the reverse rate (off-rate) constant $K_{2,1}$ (s^{-1}). At equilibrium, the ratio of the two constants gives the equilibrium constant, or affinity, of the antibody. It has been claimed that differences in affinity are primarily the result of differences in off-rates, but more recently it has been shown that affinity can also be influenced by differences in on-rates.

It has been suggested that B-cell selection and stimulation during a maturing antibody response depend upon both selection for the ability of antibodies to bind to antigens rapidly (kinetic selection) and selection for the ability to bind antigens tightly (thermodynamic selection).

Antibody specificity and affinity

Antigen–antibody reactions can show a high level of specificity. For example, antibodies to measles virus will bind to the measles virus and confer immunity to this disease, but will not combine with, or protect against, an unrelated virus such as polio. The specificity of an antiserum is equal to the sum of the actions of every antibody in that antiserum. The antibody population may contain many antigen-binding sites, each reacting with a different epitope, or even with different parts of the same epitope (*Fig. 4.16*). However, when some of the epitopes of an antigen, A, are shared by another antigen, B, then a proportion of the antibodies directed to A will also react with B. This phenomenon is termed cross-reactivity.

Antibodies recognize the overall conformation of antigens

Clearly, antibodies recognize the overall shape of an epitope rather than particular chemical residues (*Fig. 4.17*). Antibodies are capable of expressing remarkable specificity, and are able to distinguish between small differences in the primary amino acid sequence of protein antigens, in addition to differences in charge, optical configuration and steric conformation. One consequence of this specificity is that many antibodies will bind only to native antigens, or to fragments of antigens that retain sufficient tertiary structure to permit the multiple interactions required for bond formation (*Fig. 4.18*).

When considering antibodies which bind to protein antigens, we can distinguish some which interact with epitopes consisting of a single contiguous stretch of amino acids (a continuous epitope) from others which bind to epitopes formed from separated segments of the polypeptide chain (discontinuous epitopes). Antibodies which bind to discontinuous epitopes often do not bind to denatured antigens e.g. in Western blots.

Specificity can also create a problem when one wishes to produce antibodies for immunological assays. It is often

Specificity, cross-reactivity and non-reactivity

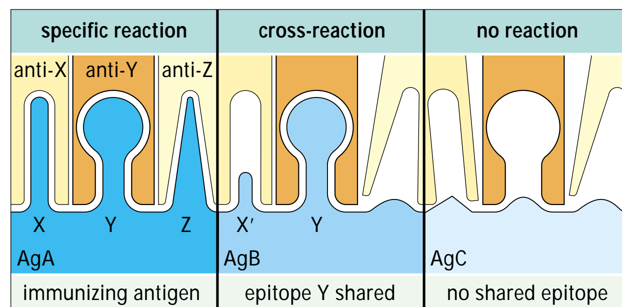


Fig. 4.16 Antiserum specificity results from the action of a population of individual antibody molecules (anti-X, anti-Y, anti-Z) directed against different epitopes (X, Y, Z) on the same or different antigen molecules. Antigen A (AgA) and antigen B (AgB) have epitope Y in common. Antiserum raised against AgA (anti-XYZ) not only reacts specifically with AgA, but cross-reacts with AgB (through recognition of epitopes Y and X′). The antiserum gives no reaction with AgC because there are no shared epitopes.

Specificity and cross-reactivity

radical (R)	sulphonate	arsonate	carboxylate
	tetrahedral	tetrahedral	planar
ortho	+ +	–	–
meta	+ + +	+	±
para	±	–	–

Fig. 4.17 Antiserum raised to the meta isomer of aminobenzene sulphonate (the immunizing hapten), is mixed with ortho and para isomers of aminobenzene sulphonate, and also with the three isomers (ortho, meta, para) of two different but related antigens: aminobenzene arsonate and aminobenzene carboxylate. The antiserum reacts specifically with the sulphonate group in the meta position, but will cross-react (although more weakly) with sulphonate in the ortho position. Further, weaker cross-reactions are possible when the antiserum is reacted with either the arsonate group or the carboxylate group in the meta, but not in the ortho or para position. Arsonate is larger than sulphonate and has an extra hydrogen atom, while carboxylate is the smallest. These data suggest that an antigen's configuration is as important as the individual chemical groupings that it contains.

Configurational specificity

antiserum	antigen		
	lysozyme	isolated 'loop' peptide	reduced 'loop'
anti-lysozyme	+ +	+	–
anti-'loop' peptide	+	+ +	–

Fig. 4.18 The lysozyme molecule possesses an intrachain bond (red) which produces a loop in the peptide chain. Antisera raised against whole lysozyme (anti-lysozyme) and the isolated loop (anti-'loop' peptide), are able to distinguish between the two. Neither antiserum reacts with the isolated loop in its linear, reduced form. This demonstrates the importance of tertiary structure in determining antibody specificity.

easier to synthesize short polypeptide antigens of known primary structure than it is to purify sufficient amounts of the native antigen for immunization. However, antibodies to the synthetic polypeptides often do not bind well or predictably to the antigen in its native form.

FUNCTIONS OF ANTIBODIES

The primary function of an antibody is to bind antigen. In a few cases this has a direct effect, for example by neutralizing bacterial toxin, or by preventing viral attachment to host cells. In general, however, the interaction of antibody and antigen is without significance unless secondary 'effector' functions come into play (*Fig. 4.19*).

Selected effector functions of human immunoglobulins

Function		immunoglobulin									
		IgG1	IgG2	IgG3	IgG4	IgM	IgA1	IgA2	sIgA	IgD	IgE
complement fixation (classical pathway)		++	+	+++	–	+++	–	–	–	–	–
placental transfer		+	+	+	+	–	–	–	–	–	–
Binding to cell surface receptors on:											
mononuclear cells	FcγRI	++	–	+++	++	–	–	–	–	–	–
	FcγRIIa	+	(+)	++	–	–	–	–	–	–	–
	FcγRIIIa	+	–	+	–	–	–	–	–	–	–
	FcμR	–	–	–	–	+	–	–	–	–	–
	FcεRII	–	–	–	–	–	–	–	–	–	++
	FcαR	–	–	–	–	–	++	++	++	–	–
neutrophils	FcγRIIa	+	–	+	–	–	–	–	–	–	–
	FcγRIIIb	+	–	+	–	–	–	–	–	–	–
	FcαR	–	–	–	–	–	++	++	++	–	–
mast cells/ basophils	FcεRI	–	–	–	–	–	–	–	–	–	+++

Fig. 4.19 These effector functions are associated with different parts of the Fc region. Placental transfer of IgG in man and intestinal transport in rodents are mediated by an MHC class-I-like receptor molecule (see *Fig. 4.24*). A complex family of receptor molecules able to bind immunoglobulin continues to be discovered (selected examples are listed here). FcμR is expressed by activated B cells but not by T cells or monocytes. FcεRII is also expressed on eosinophils, platelets, T cells and B cells.

IgG class – IgG is the most important class of immunoglobulin in secondary immune responses and, unlike IgM, is distributed evenly between the intravascular and extravascular pools.

A major effector mechanism of the human IgG1 and IgG3 subclasses is the activation of the classical pathway of complement. The latter is a complex group of serum proteins involved in the elimination of pathogens and the mediation of inflammation (see Chapter 3). The IgG2 subclass is less effective at complement activation and IgG4 appears to be inactive.

In humans IgG molecules of all subclasses cross the placenta and confer a high degree of passive immunity to the newborn (*Fig. 4.20*). In some species, e.g. the pig, maternal immunoglobulin is only transferred to the offspring postnatally. In such cases there is a selective transport of such IgG across the gastrointestinal tract via a specific receptor. The IgG subclasses also interact with a complex array of Fc receptors expressed on various cells, as summarized in *Figure 4.19* and discussed further in the section below.

IgM class – IgM is the predominant antibody in primary immune responses. The protein is largely confined to the intravascular pool and is frequently associated with the immune response to antigenically complex, blood-borne infectious organisms.

Once bound to its target, IgM is a potent activator of the classical pathway complement. Unlike IgG-mediated activation where two antibody molecules in close apposition are required, a single molecule of bound IgM is able to initiate the cascade because adjacent Fc regions are intrinsic to the structure.

IgA class – Although there are significant levels of IgA in human serum it is generally accepted that the secretory form of the protein is, in a functional sense, the most important. Secretory IgA is assembled during an active transport process as locally produced dimeric IgA passes across mucosal epithelium (*Fig. 4.21*).

In human serum IgA1 is the predominant subclass (approximately 90% of total IgA) and in many secretions such as nasal secretions, tears, saliva and milk IgA1 will account for 70–95% of total IgA. However, in the colon IgA2 predominates (approximately 60% of the total IgA). It is of interest that many microorganisms in the upper respiratory tract have adapted to their environment by releasing proteases that cleave IgA1.

IgD class – The precise biological function of this class of immunoglobulin remains unclear although it may play a role in antigen-triggered lymphocyte differentiation.

IgE class – Despite its low serum concentration, the IgE class is characterized by its ability to bind avidly to

Immunoglobulins in the serum of the fetus and newborn child

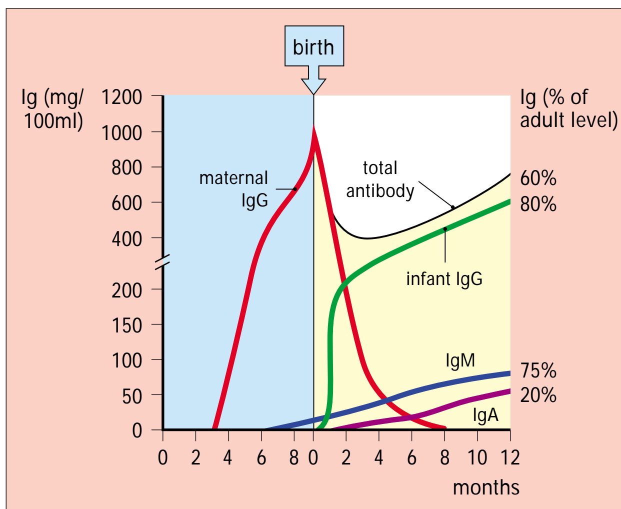


Fig. 4.20 IgG in the fetus and newborn infant is derived solely from the mother. This maternal IgG has disappeared by the age of 9 months, by which time the infant is synthesizing its own IgG. The neonate produces its own IgM and IgA; these classes cannot cross the placenta. By the age of 12 months, the infant produces 80% of its adult level of IgG, 75% of its adult IgM level and 20% of its adult IgA level.

Transport of IgA across the mucosal epithelium

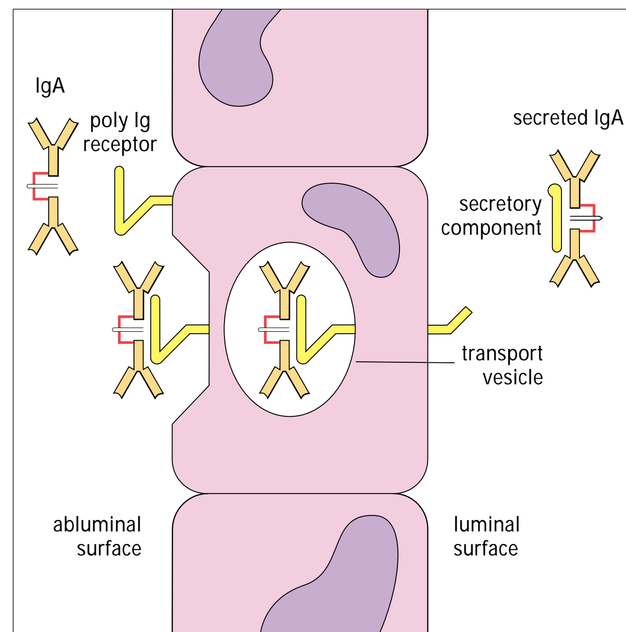


Fig. 4.21 IgA dimers secreted into the intestinal lamina propria by plasma cells bind to poly-Ig receptors on the internal (abluminal) surface of the epithelial cells. The s-IgA–receptor complex is then endocytosed and transported across the cell while still bound to the membrane of transport vesicles. These vesicles fuse with the plasma membrane at the luminal surface, releasing IgA dimers with bound secretory component derived from cleavage of the receptor. The dimeric IgA is protected from proteolytic enzymes in the lumen by the presence of this secretory component.

circulating basophils and tissue mast cells through the high affinity FcεRI receptor (see next section). It also sensitizes cells on mucosal surfaces such as the conjunctival, nasal and bronchial mucosae. This class of immunoglobulin may have evolved to provide immunity against helminthic parasites but in developed countries it is now more commonly associated with allergic diseases such as asthma and hay fever.

Fc RECEPTORS

There are three types of cell surface receptor for IgG

IgG receptors mediate several effector functions and have overlapping biological activities, which are triggered by cross-linking with the appropriate immunoglobulin. The major activities are phagocytosis, antibody-dependent cell-mediated cytotoxicity (ADCC), mediator release and enhancement of antigen presentation.

Three groups of human IgG receptor are now recognized on cell surfaces: FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16). They are all characterized by extracellular domains showing significant homology with immunoglobulin V regions (*Fig. 4.22*), i.e. they belong to the immunoglobulin superfamily, as does FcαR, a receptor specific for IgA molecules.

FcγRI (CD64) in man binds monomeric IgG with high affinity (10^8–10^9 M^{-1}) and has a more restricted distribution than the other receptors. Primarily, it is expressed on all cells of the mononuclear phagocyte lineage, and is involved in the phagocytosis of immune complexes (see Chapter 9).

FcγRII (CD32) is broadly distributed on cells and is frequently the only receptor to be expressed. It binds only complexed or polymeric IgG with a low intrinsic affinity ($<10^7$ M^{-1}). On B cells, it has a particular function of moderating cell activation when the levels of specific antibody are high (see *Fig. 11.5*).

FcγRIII (CD16) is extensively glycosylated and is expressed as a molecule with a range of molecular weights (50 000–80 000). FcγRIIIa is expressed on macrophages, NK cells and some T cells and interacts with complexed as well as monomeric IgG (intrinsic affinity 3×10^7 M^{-1}). The GPI-linked FcγRIIIb is selectively expressed on granulocytes and has a low affinity for IgG ($<10^7$ M^{-1}).

The three Fcγ receptors generate 12 different isoforms and genetic polymorphism has been described for both FcγRII and FcγRIII. In addition to this intrinsic heterogeneity there is evidence that the receptors are expressed on cell surfaces as complexes in association with other chains. Two such chains have been identified to date in unrelated receptor complexes:

- FcγRI is associated with disulphide-linked dimers of the γ chain also seen in FcεRI.
- FcγRIIIa can associate with the same dimers of γ chains, or with dimers of ζ chains from the TCR complex, or with γ–ζ heterodimers.

In addition to preventing degradation of the FcγRIIIa complex in the endoplasmic reticulum, these associated chains appear to be essential for signal transduction. In the case of the GPI-anchored FcγRIIIb there appears to be no requirement for either γ or ζ chains (*Fig. 4.22*).

Selected phagocyte receptors interacting with IgG

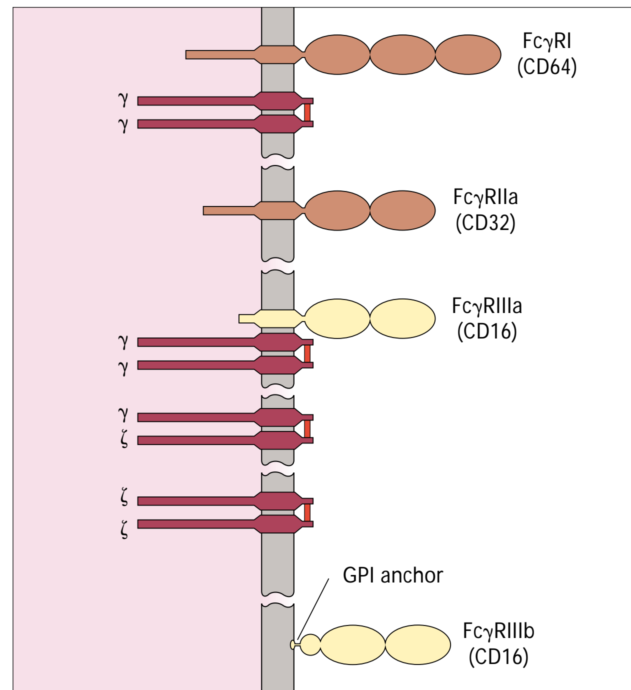


Fig. 4.22 The human Fcγ receptor structures shown are those for FcγRI (expressed by monocytes), FcγRIIa (expressed by monocytes and neutrophils), FcγRIIIa (expressed by monocytes and attached as a normal transmembrane protein) and FcγRIIIb (expressed by neutrophils and attached by a phosphatidyl inositol glycan [GPI] membrane anchor). Each receptor belongs to the immunoglobulin superfamily and expresses two or three extracellular immunoglobulin-like domains. Several of the receptors are now known to exist as complexes with various disulphide-linked subunits. FcγRI and FcγRIIIa both associate with dimers of the γ chain originally described as part of the high-affinity FcεRI complex (see *Fig 4.23*). FcγRIIIa has also been shown to associate with dimers of the ζ chain found in the TCR–CD3 complex. In the case of FcγRIIIa these subunits can associate as either homodimers (γ–γ or ζ–ζ) or as heterodimers (γ–ζ). They appear to be essential for surface expression and signal transduction. In FcγRI interactions, the receptor appears to bind a structural motif centred around Leu 235 in the CH2 domain, present in IgG1, IgG3 and IgG4.

Two distinct Fcε receptors bind to IgE

Two different receptors for IgE on cells are now known (*Fig. 4.23*). The high affinity receptor (FcεRI) is found on mast cells and basophils and is the 'classical' IgE receptor. This receptor is part of the immunoglobulin supergene family and quite distinct from the low affinity Fc receptor for IgE (FcεRII) found on leucocytes and lymphocytes. The low-affinity receptor has not evolved from the immunoglobulin superfamily, but has substantial homology with several animal lectins such as mannose-binding lectin (MBL).

IgE Fc receptors

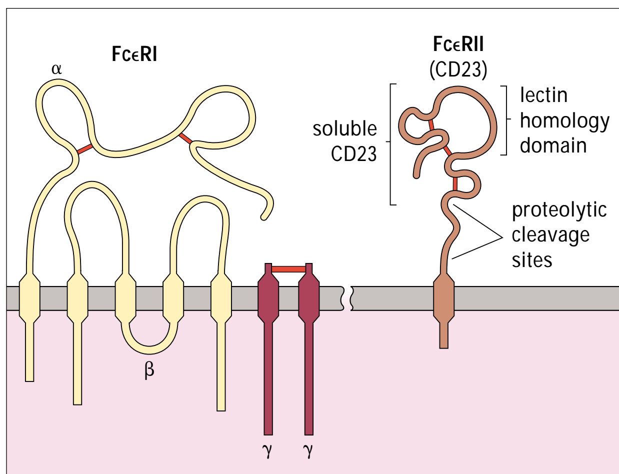


Fig. 4.23 The model for FcεRI proposes a tetramer consisting of one α chain with two disulphide-linked immunoglobulin-like loops. The β chain has two extracellular portions near two γ chains which are linked by disulphide bonds (red). The α chain is crucial for IgE binding. The model for FcεRII is hypothetical, and is based on sequence data and the homology with animal lectins. Proteolytic cleavage can release several types of IgE-binding factors, including the 25-kDa soluble CD23 molecule, which contains the lectin domain. This cleavage is inhibited by IgE, accounting for the apparent increase of FcεRII expression on lymphocytes cultured in the presence of IgE.

FcεRI is the high-affinity IgE receptor

FcεRI has a tetrameric structure (*Fig. 4.23*). The α chain (45 kDa) is glycosylated and exposed on the cell surface. Antibodies against the α chain can block IgE binding to the receptor and trigger histamine release from rat basophil leukaemia cells. The carbohydrate probably protects the α chain from serum protease activity, as it does with many other cell-surface proteins. It is unlikely that the carbohydrate on the α chain plays a role in IgE binding and IgE-mediated histamine release. The single β chain (33 kDa) and the two disulphide-linked γ chains (9 kDa) are essential components of the $\alpha\beta\gamma_2$ receptor unit. They are required for receptor expression on the cell surface, and may have a role in signal transduction.

The receptor interacts with the distal portion of the IgE heavy chain, that is, regions of the Cε2 and/or Cε3 domains. The interaction is highly specific and the binding constant for IgE is very high (approximately 10^{10} M^{-1}). Neither the interaction of monovalent IgE with the receptor complex, nor the binding of substrate to a single IgE, appear to activate mast cells or basophils, since no histamine release occurs. It is the cross-linking of several surface-bound IgEs, by antigen or by other molecules, that stimulates degranulation.

The carbohydrate associated with IgE itself does not seem to be of importance in its interaction with FcεRI. Its role seems to be in the secretion of IgE from B cells. The high-affinity receptor was thought to be limited to mast cells and basophils, but some data suggest that receptors may also be found on Langerhan's cells and stem cells.

FcεRII is the low-affinity IgE receptor

The human lymphocyte FcεRII or CD23 antigen (45 kDa) shows the characteristics of a membrane-bound molecule, i.e. a transmembrane domain, but it is unusual in that it lies 'upside-down' in the cell membrane, the C-terminus being extracellular (*Fig. 4.23*).

Two forms of the human FcεRII have now been identified, cloned and sequenced. They differ only in the N-terminal cytoplasmic region, the extracellular domains being identical. The FcεRIIa is normally expressed on B cells, whereas expression of FcεRIIb is inducible on T cells, B cells, monocytes and eosinophils by the cytokine IL-4. Expression of FcεRIIb is often increased on B cells and monocytes of individuals with eczema, and on lymphocytes of hayfever sufferers.

Many of the sites of effector functions in antibodies have been identified

In contrast to the rapid progress made in localizing the antigen-binding sites of antibodies, the precise structural locations of most effector functions have proved to be elusive. Enzymic subfragments and peptide inhibition studies provided provisional data, but further progress was slow until the technique of site-directed mutagenesis was introduced. This allowed researchers to selectively alter amino acids at different positions in the known peptide sequence, and thus to assess the importance of specific residues for particular functions. An investigation of complement activation by IgG was one of the first uses of this technique. Earlier studies had already suggested that the C1q subcomponent of C1 interacted with the Cγ2 domain of IgG. Site-directed mutagenesis was used to localize the binding site for C1q to three side chains in the Cγ2 domain, Glu 318, Lys 320 and Lys 322. This IgG sequence motif appears to be the common feature in interactions between C1q and IgG molecules.

In the case of IgM, complement activation seems to involve a different mechanism. Free circulating IgM in the star-shaped configuration is clearly incapable of activating complement, whereas IgM bound to antigen is a potent activator. Feinstein and colleagues suggested that the process of IgM binding to a polymeric or latticed antigen dislocates the $F(ab')_2$ units out of their original plane and leads to the so-called crab-like configuration visualized by electron microscopy (*Fig. 4.8*). These conformational changes would unveil a ring of C1q binding sites that are hidden in the star-shaped configuration of IgM by the close juxtaposition of the subunits. Candidate residues are His 430, Asp/Gly 432 and Pro 436, which occupy a structural location in the Cμ3 domain that is analogous to the proposed C1q binding site in the Cγ2 domain of IgG.

IgG molecules interact with a wide range of cellular Fc receptors. Site-directed mutagenesis studies suggest that the high-affinity FcγRI receptor of monocytes interacts with a motif centred around a leucine residue at position 235 of the IgG heavy chain, between the Cγ2 domain and the hinge region.

The interactions between maternal IgG and the MHC class-I-like FcRn expressed on the intestinal epithelium of the neonatal rat have now been studied at high resolution (*Fig. 4.24*) and are believed to mimic closely the binding

The intestinal Fc receptor FcRn

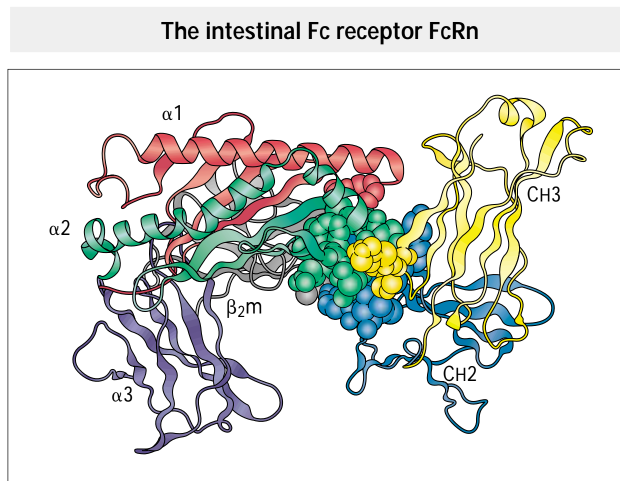


Fig. 4.24 Principal interactions between neonatal rat intestinal FcRn and the Fc of maternal IgG (derived from milk) are illustrated by ribbon diagrams of FcRn (domains α1, α2, α3 and β_2m are shown in red, light green, purple and grey respectively) and of Fc (CH2 and CH3 domains are shown in blue and yellow). The main contact residues of the FcRn (α1 domain, 90; α2 , 113–119 and 131–135; β_2m 1–4 and 86) are depicted as space-filling structures. (Reproduced with permission from Ravetch and Margulies. New tricks for old molecules. *Nature* 1994;**372**:323–4.)

Interaction of IgE with IgE receptors

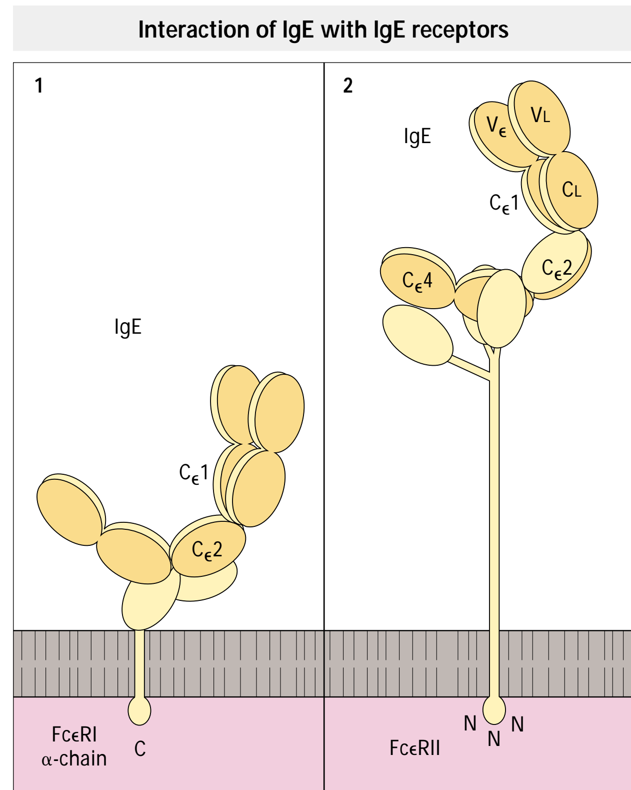


Fig. 4.25 Possible models for IgE interactions with (**1**) the FcεRI receptor and (**2**) the FcεRII receptor. The model shown in (**1**) envisages the α chain of the FcεRI receptor oriented at an angle to the membrane and interacting through the membrane-proximal Ig-like domain with the convex surface of an IgE molecule at the Cε2 and Cε3 interface. Such a model explains why IgE can, apparently, bind only a single FcεRI α-chain. The model shown in (**2**) illustrates the interaction of the Cε3 domains of the IgE molecule with two of the three lectin domains of the membrane-bound FcεRII receptor.

of the human placental counterpart hFcRn with maternal IgG. Fc contacts FcRn mainly in the junction between the CH2 and CH3 domains, overlapping the Fc binding site for Protein A. It is probable that three or four histidine residues on the Fc interface are critical. Titration of histidines could account for the binding of IgG to FcRn at pH 6.5 (the pH of ingested milk) and its release at pH 7.5 (the pH of blood).

The interaction between Protein A of *Staphylococcus aureus* and the Fc region of IgG has also been mapped in some detail. The data suggest a binding site spanning the Cγ2–Cγ3 junction in the Fc region.

A genetic engineering approach has been used to study the sites on the IgE molecule that interact with mast cells through the FcεRI receptor, or with B cells through the FcεRII receptor. Recombinant peptides containing ε-chain sequences were synthesized and then used to inhibit IgE–receptor interactions. In the case of FcεRI interactions, a 76-residue peptide spanning the Cε2–Cε3 junction appears to be critical. By contrast, the FcεRII site appears to recognize a motif involving residues in the Cε3 domains of both ε chains. Possible models for the interaction of IgE with the FcεRI and FcεRII receptors are illustrated in *Figure 4.25*.

GENERATION OF ANTIBODY DIVERSITY

Antibodies are remarkably diverse and provide enough different combining sites to recognize the millions of antigenic shapes in the environment. Each class of antibody also has a characteristic effector region so that, for instance, IgE can bind to Fc receptors on mast cells, whereas IgG can bind to phagocytes. It has been estimated that an individual produces more different forms of antibody than all the other proteins of the body put together. In fact, we produce more types of antibody than there are genes in our genome. How can all this diversity be generated? Ideas about the formation of antibodies have changed considerably over the years, but it is perhaps surprising how close Ehrlich came with his side-chain hypothesis at the beginning of the twentieth century (*Fig. 4.26*). His idea of antigen-induced selection is close to our present view of clonal selection, except that he placed several different receptors on the same cell.

Theories of antibody formation

After Ehrlich the situation became more complicated. The problem was that many new organic compounds were being synthesized and Landsteiner was showing that the immune system could react with the production of specific antibody for each new compound. It was not thought possible that the immune system could, by natural selection, have maintained genes for all these antibodies directed at

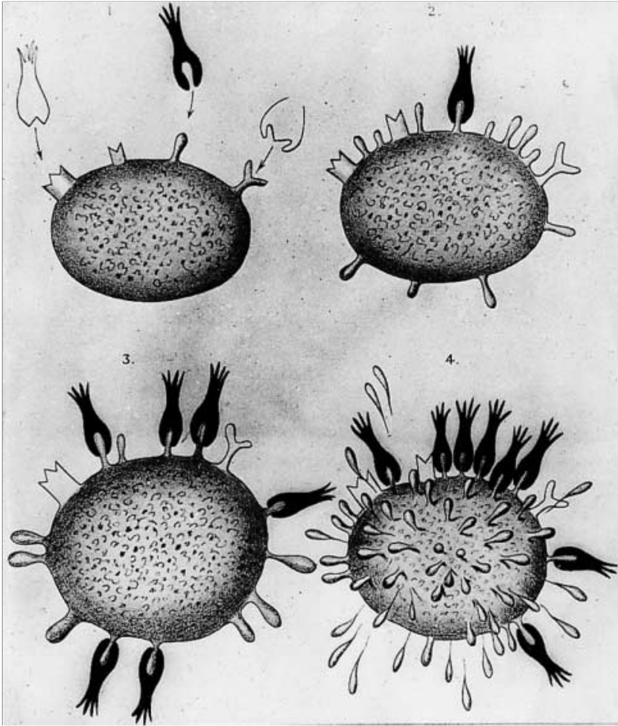


Fig. 4.26 Ehrlich's side-chain theory. Ehrlich proposed that the combination of antigen with a preformed B-cell receptor (now known to be antibody) triggered the cell to produce and secrete more of those receptors. Although the diagram indicates that he thought a single cell could produce antibodies to bind more than one type of antigen, it is evident that he anticipated both the clonal selection theory and the idea that the immune system could generate receptors before contact with antigen.

novel, artificial compounds. This led to the development of the instructive hypothesis, which proposed a flexible antibody molecule that is acted on by antigen to form a complementary binding site. With the spectacular progress in molecular biology in the 1950s and 1960s the instructive hypothesis became untenable, as it became clear that the mechanism for the proposed 'instruction' simply did not exist. The circle turned and selective theories came back into favour, with both Jerne and Burnet putting forward the idea of clonal selection. Each lymphocyte produces one type of immunoglobulin only, and the antigen selects and stimulates cells carrying that immunoglobulin type.

This still left the problem of antibody diversity. One solution was to postulate the existence of a separate gene for each antibody specificity. This immediately presented a problem. Looking at the structure of a light chain, half the chain is variable in amino acid sequence but the other half is constant; similarly with heavy chains, a quarter of the chain is variable while the rest is constant. How, if there are so many genes, is it possible to maintain this constancy of sequence in the constant regions? Dreyer and Bennett proposed a solution to this problem by suggesting that the constant and variable portions of the chains are coded for by separate genes, with one or only a few genes coding for the constant (C) region and many genes coding for the variable (V) region. At this point, the theory still needed to account for the multiple variable regions. A solution to this aspect of the diversity problem was suggested by the idea of somatic mutation, which proposes that relatively few germ line genes give rise to many mutated genes during the lifetime of the individual. It was also suggested that a number of gene segments could recombine to give a complete V gene. During cutting and joining of the DNA, extra nucleotides may be inserted at the cut ends to give further variability. This is called N-region diversity because the sequence is non-germ-line encoded. Rather than mutation, a panel of pseudogenes can also be used to alter the sequences within the variable region by a process of gene conversion. This gave five possible solutions to the problem of generating diversity:

- Multiple V-region genes in the germ line.
- Somatic recombination between elements forming a V-region gene.
- Gene conversion.
- Nucleotide addition.
- Somatic mutation.

It is now known that mammals may use all five mechanisms to generate diversity. Interestingly, sharks rely on having a large number of antibody genes, and do not use somatic recombination, while chickens have small numbers of antibody genes that undergo a very high level of gene conversion (see *Figs 13.20, 13.21*)

Immunoglobulins show isotypic, allotypic and idiotypic variation

Immunoglobulins are composed of heavy and light chains, the light chains being either kappa (κ) or lambda (λ) type. Because virtually any light chain can combine with any heavy chain, the number of possible antigen-binding sites is the product of the number of heavy and light chains. Part of the variability in immunoglobulin structure is derived from the interaction of these separate polypeptide chains. For example, if there are 10^4 different light chains each capable of binding with any of 10^4 different heavy chains, then 10^8 different antibody specificities are theoretically possible. Separate diversification mechanisms exist for each of the chains, as they are encoded on separate chromosomes. For example, in humans the genes encoding kappa and lambda light chains are found on chromosomes 2 and 22 respectively whereas the heavy chain gene locus is found on chromosome 14.

Polymorphic forms of immunoglobulins derive from variation in many parts of the molecule (*Fig. 4.27*).

Isotypic variation

The genes for isotypic variants are present in all healthy members of a species. For example, the genes for $\gamma 1$, $\gamma 2$, $\gamma 3$, $\gamma 4$, μ, $\alpha 1$, $\alpha 2$, δ, ε, κ and λ chains are all present in the human genome, and are therefore isotypes.

Allotypic variation

This refers to genetic variation between individuals within a species, involving different alleles at a given locus. For example, the variant of IgG3 called G3m(b^0) is characterized by a phenylalanine at position 436 of the $\gamma 3$ heavy chain. It is not found in all people and is therefore an

Variability of immunoglobulin structure

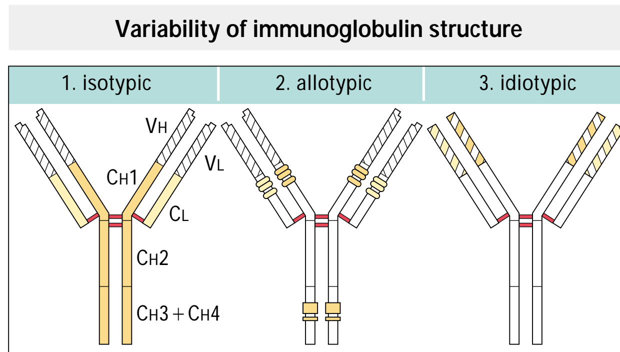


Fig. 4.27 All immunoglobulins have the basic four-chain structure. There are three types of immunoglobulin variability:
(1) Isotypic variation is present in the germ line of all members of a species, producing the heavy (μ, δ, γ, ε, α) and light chains (Igκ, λ) and the V-region frameworks (subgroups).
(2) Allotypic variation is intraspecies allelic variability.
(3) Idiotypic variation refers to the diversity at the antigen-binding site (paratope) and in particular relates to the hypervariable segments.

allotype. Allotypes occur mostly as variants of heavy chain constant regions.

Idiotypic variation

Variation in the variable domain, particularly in the highly variable segments known as hypervariable regions, produces idiotypes. These determine the binding specificity of the antigen-binding site. Idiotypes are usually specific for individual B-cell clones (private idiotypes), but are sometimes shared between different B-cell clones (public, cross-reacting or recurrent idiotypes).

Hypervariable sequences in the antigen-binding site allow antibodies to bind a range of antigens

Within the variable regions of both heavy and light chains, some short polypeptide segments show exceptional variability. Termed hypervariable regions, these segments are located near amino acid positions 30, 50 and 95 (*Fig. 4.28*). Because they create the antigen-binding site, hypervariable regions are sometimes referred to as complementarity determining regions (CDRs). The intervening peptide segments are called framework regions (FRs). In both light and heavy chain V regions there are three CDRs (CDR1–CDR3) and four FRs (FR1–FR4).

The variable regions of the light and heavy chains are folded in such a way that the regions of hypervariability are brought together to create the surface structure that binds antigen. These regions are, in the main, associated with bends in the peptide chain (see *Fig. 4.5*).

IMMUNOGLOBULIN GENE RECOMBINATION

Light chain genes recombine V and J segments to make a gene for the VL domain

With the advent of recombinant DNA techniques, it became possible to attempt analysis of the genes encoding

Amino acid variability in the variable regions of immunoglobulins

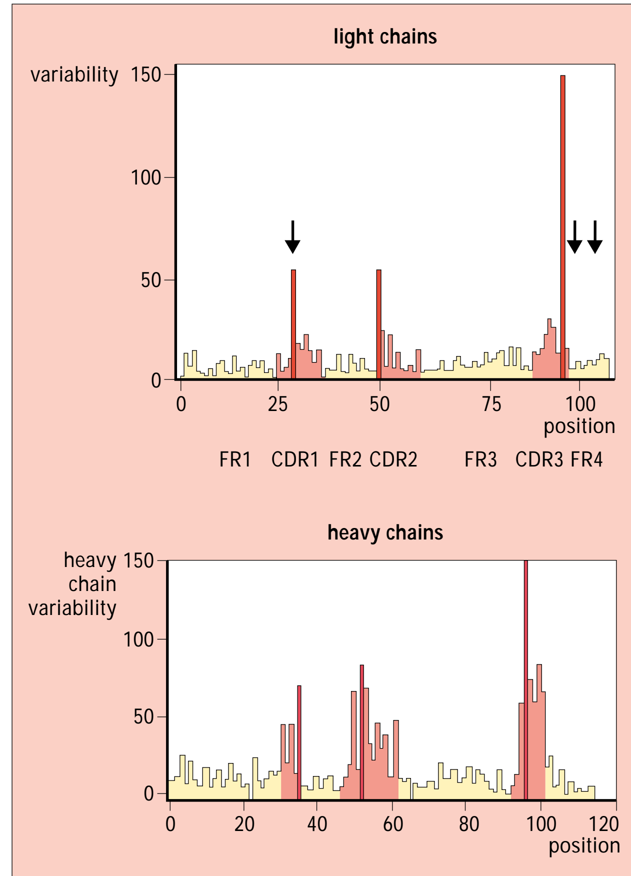


Fig. 4.28 Variability is calculated by comparing the sequences of many individual chains and, for any position, is equal to the ratio of the number of different amino acids found at that position, to the frequency of the most common amino acid. The areas of greatest variability, of which there are three in the VL domain, are termed the hypervariable regions. In some sequences studied, extra amino acids have been found, but these are excluded here to enhance comparison; their positions are indicated by arrows. The areas shaded pink denote regions of hypervariability (CDR), and the most hypervariable positions are shaded red. The four framework regions (FR) are shown in yellow. (Courtesy of Professor E. A. Kabat.)

antibodies. In accord with the Dreyer and Bennett hypothesis, it was found that two separate segments of DNA code for the constant and variable regions of light chains. In cells not producing antibody, these gene segments are far apart on the chromosomes, but in antibody-forming cells they are brought closer together. However, even in a fully differentiated B cell, the two gene segments do not join directly, but remain about 1500 base pairs apart. Between the V and C segments, and joined onto the V segment in rearranged chromosomes, is a short section of DNA known as the J (joining) segment. (Not to be confused with the J chain present in IgM and dimeric IgA.)

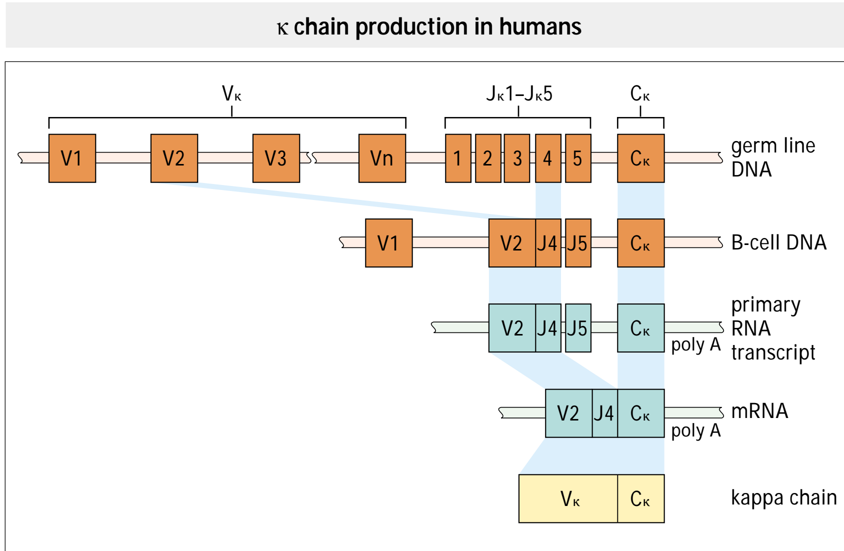


Fig. 4.29 During differentiation of the pre-B cell one of several Vκ genes on the germ line DNA (V1–Vn) is recombined and apposed to a Jκ segment (Jκ1–Jκ5). The B cell transcribes a segment of DNA into a primary RNA transcript that contains a long intervening sequence of additional J segments and introns. This transcript is processed into mRNA by splicing the exons together, and is translated by ribosomes into kappa (κ) chains; B-cell DNA is coloured light brown; RNA is coloured green; and immunoglobulin peptides are coloured yellow. The rearrangement illustrated is only one of the many possible recombinations.

Structure of the light chain systems

The Vκ-gene segment codes for the V region of the antibody light chain, up to and including amino acid 95; the Jκ-gene segment codes for the rest of the V region (*Fig. 4.29*). There is only one constant-region gene but in humans, of 76 Vκ genes, 16 have minor defects and 25 are pseudogenes leaving 35 potentially functional genes. During differentiation of lymphoid cells, there is rearrangement of the DNA such that one of the V genes is joined to one of five J genes. The number of possible κ-chain variable regions that can be produced in this way is approximately 150 (30×5). Each V segment is preceded by a leader or signal sequence, a short hydrophobic segment responsible for targeting the chain to the endoplasmic reticulum. The leader sequence is cleaved in the endoplasmic reticulum and the antibody molecule is then processed through the intracellular secretory pathway.

The λ-gene locus in humans contains a set of V genes, but each of seven C genes is accompanied by just one J gene. Despite the difference, the overall process of V–J recombination is similar to that of the κ genes (*Fig. 4.30*).

Following rearrangement of the V and J genes, there is still an intron (a non-coding intervening sequence) between the recombined VJ gene and the gene for the C region. This whole stretch of DNA (from the leader to the end of the C gene, including introns) is then transcribed into heterogeneous nuclear RNA (hnRNA), that is, unprocessed mRNA. A process of RNA splicing then removes the introns, leaving mRNA that can be translated into protein.

Heavy chain genes recombine V, D and J gene segments to make a gene for the VH domain

The heavy chain variable region is also encoded by V- and J-segment genes. Additional diversity is provided by a third gene segment, the D (diversity)-segment gene (*Fig. 4.31*). The D segment is highly variable both in the number of codons and in the sequence of base pairs. In antibodies

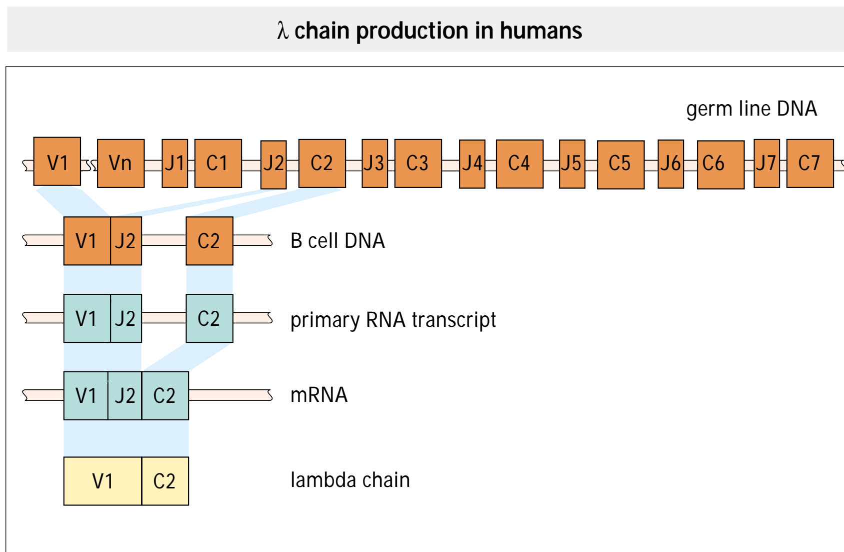


Fig. 4.30 During B-cell differentiation, one of the germ line Vλ genes recombines with a J-segment to form a V–J combination. The rearranged gene is transcribed into a primary RNA transcript complete with introns (non-coding segments occurring between the genes), exons (which code for protein) and a poly-A tail. This is spliced to form messenger RNA (mRNA) with loss of the introns, and then translated into protein.

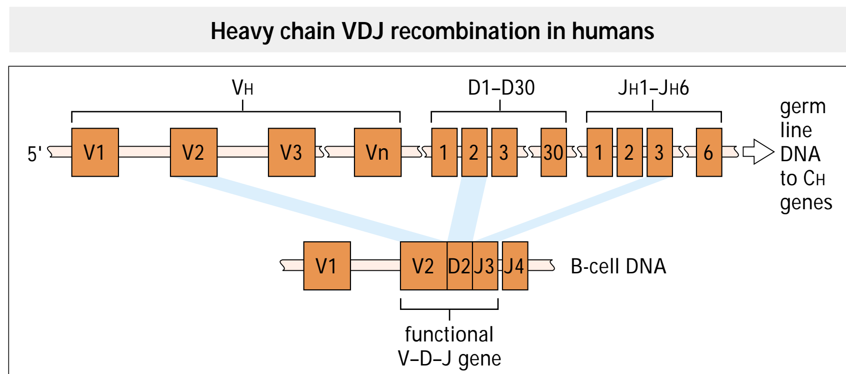


Fig. 4.31 The heavy chain gene loci combine three segments to produce the exon (V–D–J gene) which codes for the V_H domain. Of some 80 V genes, about 50 are functional, and the others are pseudogenes. The V gene recombines with one of 30 D segments and one of six J segments, to produce a functional V–D–J gene in the B cell.

binding dextran this section comprises two amino acids; in those binding phosphorylcholine up to eight amino acids are inserted; in anti-levan antibodies this section is completely missing. More than one D segment may join to form an enlarged D region. The D region may be read in three possible reading frames without generating stop codons, so adding to diversity. So far, 30 germ line D segments have been identified. Eighty-seven V_H segments are found on chromosome 14, of which at least 32 are pseudogenes. The recombination junction of V, D and J segments in the heavy chain is largely responsible for variability in the complementarity-determining region, CDR-3, which forms an essential part of the antigen-binding site. In some systems, such as the family of anti-dextran antibodies, the differences between antibodies are nearly all situated in this region.

V regions are rearranged and expressed in a programmed manner during early fetal life

When animals are immunized with selected antigens during fetal life or soon after birth, the ability to respond to each antigen develops in a precise order, suggesting a programmed pattern of development. In humans the V region nearest to the J regions is utilized first, and it is interesting that V6-1, the nearest V gene, is a single conserved sequence. In all primate species examined, only a single copy is present, and no sequence variation occurs within a species. Between humans and other primates, only 2% of the nucleotides vary.

This fetal repertoire is overrepresented in autoantibodies, indicating that autoimmunity might, in part, be the result of dysregulation of these early sequences. There is similar overrepresentation of particular V segments in tumours of early B cells, 20 V segments being present in 85% of chronic lymphocytic leukaemias.

Recombination sequences flanking the V, D and J genes direct joining of the gene segments

The recombination of gene segments is a key feature of the generation of a functional gene for both light and heavy chain variable regions. The precise mechanism by which this recombination occurs is now becoming clear, and specific base sequences that act as recombination signal sequences (RSS) have been identified. A signal sequence is found downstream (3′) of V- and D-segment genes. It consists of a heptamer CACAGTG or its analogue, followed by a spacer of unconserved sequence (12 or 23 bases), and then a nonamer ACAAAAACC or its analogue. Immediately upstream (5′) of all germ line D and J segments is a corresponding signal sequence of a nonamer and then a heptamer, again separated by an unconserved sequence (12 or 23 bases). The heptameric and nonameric sequences following a V_L, V_H or D segment are complementary to those preceding the J_L, D or J_H segments with which they recombine. The 12 and 23 base spacers correspond to either one or two turns of the DNA helix (*Fig. 4.32*).

The recombination process is controlled at least in part by two recombination-activating genes (RAG-1 and RAG-2). Mice whose RAG-1 and RAG-2 genes have been 'knocked out' lack mature T and B cells, because of a failure to produce the T-cell receptors and immunoglobulins respectively. Interleukin-7 (IL-7), a cytokine produced by bone-marrow stromal cells, has been shown to influence levels of RAG-1 and RAG-2 expression. Studies with synthetic DNA substrates containing heptamer and nonamer sequences, together with 12b and 23b spacers and purified RAG-1 and RAG-2, have established that these proteins are able to mediate the formation of a synaptic complex between two RSS. Initially, RAG-1 binds to the nonamer sequence, then RAG-2, which cannot bind

Recombination sequences in immunoglobulin genes

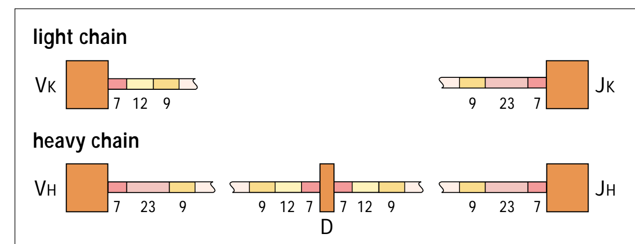


Fig 4.32 The recombination sequences in the light chain genes (top) and heavy chain genes (bottom) consist of heptamers (7), 12 or 23 unconserved bases and nonamers (9). The sequences of heptamers and nonamers are complementary and the nonamers act as signals for the recombination activating genes to form a synapsis between the adjoining exons. Similar recombination sequences are present in the T-cell receptor V, D and J gene segments (see Chapter 5).

to DNA on its own, binds to the RAG-1–DNA complex, particularly next to a 12b spacer. Synapsis then occurs between 12b and 23b spacers. It is possible that this is mediated between the RAG-1–RAG-2 complex on the nonamer–12b spacer sequence and a RAG-1 complex on the nonamer–23b sequence. After binding, cleavage is directed by the heptamer sequences at the borders between the V–(D–)J coding regions and the heptamers.

Cleavage is initiated by introducing a nick in the area bordering the 5′ end of the signal heptamer and the coding region. This is then converted, by the RAG proteins, into a double-strand break, generating a hairpin coding end. *In vitro*, hairpin formation is clearly seen at the cut coding ends. This also occurs *in vivo*, but is less obvious, as the hairpins are often further modified before joining.

The place at which V and J segment genes join may vary slightly

Slight variations in the positions at which recombination takes place generate additional diversity. For example, the 95th residue of the κ light chain is usually encoded by the last codon of the V-segment gene; the 96th is frequently encoded by the first Jκ triplet. Sometimes, however, the 96th amino acid is encoded by a composite triplet formed by the second and third bases, or third base alone, of the first Jκ triplet, with the other bases of the triplet coming from the intron 3′ to the V-segment gene (*Fig. 4.33*). This will lead to variations in amino acid sequence. Obviously, to produce a functional light chain the correct reading frame must be preserved; if the gene segments join out of phase, non-functional antibodies are produced.

In the case of the heavy chain, similar imprecision in joining occurs between the D and JH segment genes and can extend over as many as 10 nucleotides (*Fig. 4.34*). Furthermore, a few nucleotides may be inserted between D and JH and between VH and D by means of the enzyme, terminal deoxynucleotidyl transferase, without the need for a template. The addition of these N-nucleotides is called N-region diversity. In mice, terminal deoxynucleotidyl transferase activity increases with age, giving rise to long N segments in adult animals. Thus the recombinational variability of the D region can be so great that no recognizable D-gene segment remains.

Light chain diversity created by variable recombination

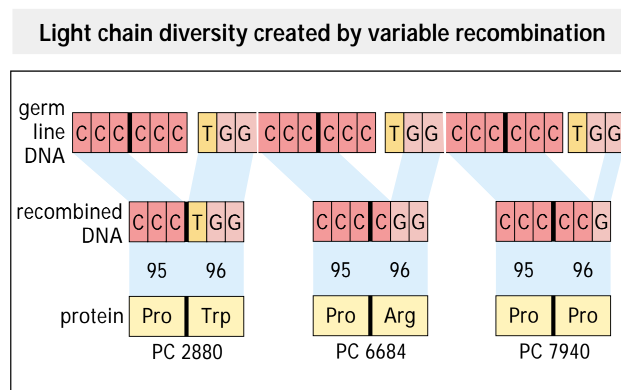


Fig. 4.33 The same Vκ21 and J1 sequences of the germ line create three different amino acid sequences in the proteins PC2880, PC6684 and PC7940 by variable recombination. PC2880 has proline and tryptophan at positions 95 and 96, caused by recombination at the end of the CCC codon. Recombination one base down produces proline and arginine in PC6684. Recombination two bases down from the end of Vκ21 produces proline and proline in PC7940.

Severe combined immune deficiency (SCID) mice do not generate functional T or B cells because of a defect in V–(D)–J recombination.

SOMATIC MUTATION

Immunoglobulin heavy and light chain genes undergo structural modifications after antigen stimulation

The idea that somatic mutations could occur during the lifetime of an individual and thus increase the diversity of antibodies has been strongly argued for many years. Most sequences in the murine λ1 light chain system are identical, with a few variations in the complementarity-determining regions giving eight sequences in all. However, only one Vλ1 gene segment has been found per haploid genome, and this corresponds to the main shared prototype sequence; thus, all variant sequences must be generated by

Heavy chain diversity created by variable recombination and N region diversity

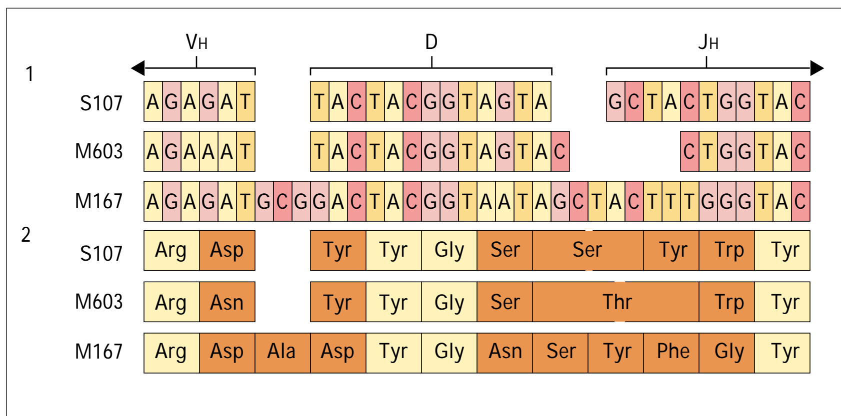


Fig. 4.34 The DNA sequence (**1**) and amino acid sequence (**2**) of three heavy chains of anti-phosphorylcholine are shown. Variable recombination between the germ line, V, D and J regions and N region insertion causes variation (orange) in amino acid sequences. In some cases (e.g. M167) there appear to be additional inserted codons. However, these are in multiples of three, and do not alter the overall reading frame.

somatic mutations produced by single base changes. Somatic mutants have also been identified in κ light chains and in heavy chains.

There is some evidence that the region of DNA encoding the variable region may be particularly susceptible to mutation. For example, examination of the nucleotide sequences of two anti-phosphorylcholine antibodies shows them to have numerous mutations from the germ line sequence. These mutations are found in both introns and exons of the V region, but not in adjoining sequences, implying that the whole V region is particularly mutable (*Fig. 4.35*). Somatic mutation occurs in germinal centres, and cells which have produced a higher-affinity antibody are selected for survival (see Chapter 8). This process is dependent on both T cells and germinal centres. Athymic mice lack T cells and germinal centres and there is no affinity maturation.

Antibody diversity thus arises at several levels. First there are the multiple V genes recombining with J and D segments. The imprecision with which recombination occurs achieves further variation. Note that the structures of the first and second hypervariable regions are coded for entirely by germ line genes. CDR-2 (HV2) already has wide diversity encoded in the genome, CDR-1 is diversified by somatic mutation, and variability in CDR-3 is largely the result of recombination. As virtually any light chain may pair with any heavy chain the combinatorial binding of heavy and light chains amplifies the diversity enormously (*Fig. 4.36*). Somatic hypermutation probably contributes less than 5% of total sequence variability, but up to 90% of B cells express VH genes which have undergone somatic mutation.

Mutations in the DNA of two heavy chain genes

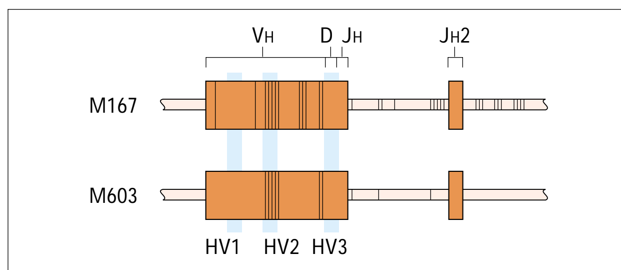


Fig. 4.35 DNA of two IgG antibodies to phosphorylcholine are illustrated. Both antibodies share the T15 idiotype. Black lines indicate positions where the sequence has mutated from the germ line sequence. There are large numbers of mutations in both the introns and exons of both genes, but particularly in the second hypervariable region (HV2). By comparison, there are no mutations in the regions encoding the constant genes, which implies that the mutational mechanism is highly localized.

Seven mechanisms for the generation of antibody diversity
1. multiple germ line V genes
2. V–J and V–D–J recombinations
3. N-nucleotide addition
4. gene conversion
5. recombinational inaccuracies
6. somatic point mutation
7. assorted heavy and light chains

Fig. 4.36 Since each mechanism can occur with any of the others, the potential for increased diversity multiplies at each step of immunoglobulin production.

CRITICAL THINKING • The specificity of antibodies (Explanations on pp. 453–454)

The human rhinovirus HRV14 is formed from four different polypeptides: one of them (VP4) is associated with viral RNA in the core of the virus, while the other three polypeptides (VP1–VP3) make up the shell of the virus – the capsid.

4.1 When virus is propagated in the presence of neutralizing anti-viral antiserum it is found that mutated forms of the virus develop. Mutations are detected in VP1, VP2 or VP3, but never in VP4. Why should this be so?

The most effective neutralizing antibodies are directed against the protein VP1 – this is termed an immunodominant antigen. Two different monoclonal antibodies against VP1 were developed and used to induce mutated forms of the virus. When the sequences of the mutated variants were compared with the original virus, it was found that only certain amino acid residues became mutated (see table below).

Antibody	Amino acid number	Residue in wild type	Observed mutations
VP1-a	91	Glu	Ala, Asp, Gly, His, Asn, Val, Tyr
	95	Asp	Gly, Lys
VP1-b	83	Gln	His
	85	Lys	Asn
	138	Glu	Asp, Gly
	139	Ser	Pro

4.2 What can you tell about the epitopes which are recognized by the two different monoclonal antibodies?

4.3 When the binding of the antibody VP1-a is measured against the different mutant viruses, it is found that it binds with high affinity to the variant with Glycine (Gly) at position 138, with low affinity to the variant with Gly at position 95 and it does not bind to the variant with Lysine (Lys) at position 95.

How can you explain these observations?

DISCUSSION POINT How could you use this information to design a vaccine against a common cold virus?

FURTHER READING

Blackwell TK, Alt FW. Mechanism and developmental program of immunoglobulin gene rearrangement in mammals. *Annu Rev Genet* 1989;**23**:605–36.

Burton DR. Antibody: the flexible adaptor molecule. *Trends Biochem Sci* 1990;**15**:64–9.

Conrad DH. The low affinity receptor for IgE. *Annu Rev Immunol* 1990;**8**:623–45.

Davies AC, Schulman MJ. IgM – molecular requirements for its assembly and function. *Immunol Today* 1989;**10**:118–22, 127–8.

Davies DR, Metzger H. Structural basis of antibody function. *Annu Rev Immunol* 1984;**1**:87–117.

Feinstein A, Richardson N, Taussig MJ. Immunoglobulin flexibility in complement activation. *Immunol Today* 1986;**7**:169–73.

Harriman W, Volk H, Defanoux N, *et al.* Immunoglobulin class switch recombinations. *Annu Rev Immunol* 1993;**11**:361–84.

Honjo T, Alt FW. *Immunoglobulin Genes*. London: Academic,1995.

Hunkapiller T, Hood L. Diversity of the immunoglobulin gene superfamily. *Adv Immunol* 1990;**44**:1–63.

Kerr MA. The structure and function of human IgA. *Biochem J* 1990;**271**:285–96.

Kilian M, Russell MW. Function of mucosal immunoglobulins. In: Ogra P, Mestecky J, Lamm ME, *et al.* (eds). *Handbook of Mucosal Immunology*. San Diego: Academic, 1994:127–37.

Lewis SM. The mechanism of V(D)J joining: lessons from molecular, immunological and comparative analyses. *Adv Immunol* 1994;**56**:27–150.

Möller G. (ed). Immunoglobulin D: structure, synthesis, membrane representation and function. *Immunol Rev* 1977;**37**.

Möller G. (ed). Immunoglobulin E. *Immunol Rev* 1978;**41**.

Möller G. (ed). Fc receptors. *Immunol Rev* 1992;**125**.

Möller G. (ed). The B-cell antigen receptor complex. *Immunol Rev* 1993;**132**.

Nisonoff A. *Introduction to Molecular Immunology*. 2nd edn. Baltimore: Sinauer Associates, 1984:326.

Pascual V, Capra JD. Human immunoglobulin heavy-chain variable region genes: organization, polymorphism and expression. *Adv Immunol* 1991;**49**:1–74.

Ravetch JV, Kinet J-P. Fc receptors. *Annu Rev Immunol* 1991;**9**:457–92.

Ravetch JV, Margulies DH. New tricks for old molecules. *Nature* 1994;**372**:323–4.

Shakib F. (ed). *The Human IgG Subclasses. Molecular analysis of structure, function, and regulation*. Oxford: Pergamon, 1990.

Van de Winkel JGJ, Capel PJA. Human IgG Fc receptor heterogeneity: molecular aspects and clinical implications. *Immunol Today* 1993;**14**:215–21.

INTERNET REFERENCES

The ImMunoGeneTics database IMGT was originally established in 1989 and the IMGT/LIGM-DB is a comprehensive database of Ig and TcR sequences from human and other vertebrates
http://www.ebi.ac.uk/imgt

V base contains the germline variable region sequences of more than 1000 published sequences
http://www.mrc-cpe.cam.ac.uk/imt-doc/public

The Kabat database contains sequences of proteins of immunological interest
http://immuno.bme.nwu.edu