

Recognition of Foreign Molecules by the Immune System

12

The immune system in vertebrates provides a defense mechanism against foreign parasites such as viruses and bacteria. Three main properties are essential to its operation: specific recognition of foreign molecules, which also involves discrimination between self and nonself; the ability to destroy the foreign parasite; and a memory mechanism that results in a more rapid response to a second infection by the same microorganism.

Foreign invaders are recognized through specific and tight binding of the proteins of the immune system to molecules specific to the foreign organisms. The sites on foreign molecules that are recognized by the immune system are called **antigenic determinants**, and they interact with two different classes of antigen receptors on the surface of the two major cell types of the immune system. **Antibodies**, which are also known as **immunoglobulins**, act as antigen receptors on the surface of B cells, which are stimulated by antigen binding to secrete antibodies into the bloodstream (Figure 12.1a). A related, though quite distinct molecule, the **T-cell receptor**, acts as the antigen re-

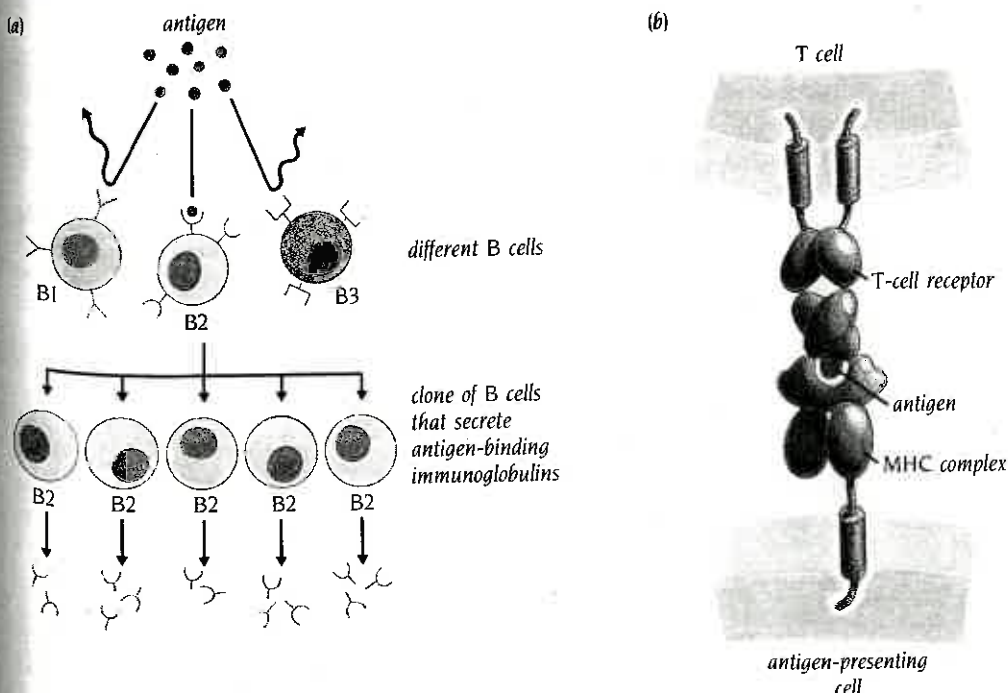


Figure 12.1 (a) The clonal selection theory. An antigen (red) activates those B cells that have immunoglobulin molecules on their surfaces that can recognize and bind the antigen. This binding triggers production of a clone of identical B cells that secrete soluble antigen-binding immunoglobulins into the blood stream. (b) T cells recognize foreign viral antigens (red), through a T-cell receptor (blue) that can bind degraded fragments of the antigen when they are associated with an MHC molecule (green). The polypeptide chains of both T-cell receptors and MHC molecules are folded into immunoglobulin-like domains, represented as ellipsoids in the diagram.

ceptor for T cells one of whose principal functions is to recognize and destroy virus-infected cells (Figure 12.1b).

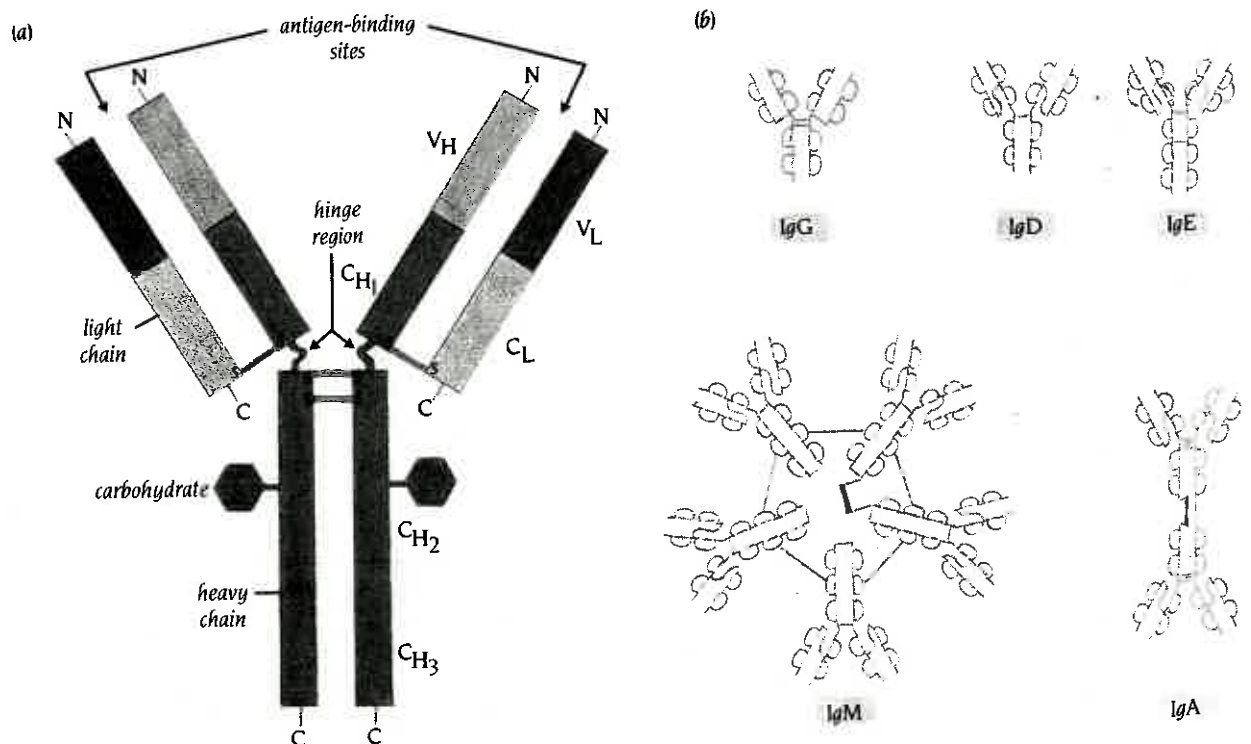
Antibody molecules at the surface of B cells recognize and bind intact soluble molecules or molecules on the surface of invading microorganisms. This stimulates the B cells first to multiply and to differentiate and then to synthesize and to secrete into the blood more copies of the antibody. The soluble antibody molecules bind to the cognate antigen and the complexes are recognized and removed by macrophages.

T cells, by contrast, directly destroy infected cells, and their receptors recognize antigenic determinants only when presented as part of a complex with a third important class of protein molecules, the molecules encoded by the major histocompatibility complex, or MHC, on the surface of cells. MHC molecules are so called because of the part they play in transplant rejection: each individual carries a selected set of MHC molecules on the surface of his or her cells, and the difference in the precise set of MHC molecules carried by the cells of different individuals forms the basis of discrimination between self and nonself. The physiological function of the MHC molecules, however, is to bind to degraded fragments of antigen generated inside infected cells and display them for recognition by T cells.

Thus T-cell receptors exist only on the surface of T cells, and the antigenic determinants they recognize are peptides derived from foreign proteins and complexed with MHC molecules on the surfaces of cells. Antibodies, by contrast, exist both as cell surface and as secreted soluble molecules and interact directly with intact antigens by forming very specific antigen-antibody complexes. Antibodies are the most structurally diverse of all known proteins, and any antigen presented to the immune system triggers the production of many different and specific antibodies.

In this chapter we will see that all three of the molecules directly involved in the specific recognition of antigen—immunoglobulin, the T-cell receptor, and the molecules of the MHC—belong to a family of proteins that seems to have evolved by duplication and diversification from a single ancestral domain. Because antibodies are naturally produced in very large quantities as soluble molecules, the crystal structure of immunoglobulin was solved long before either of the other two molecules could be approached, and hence the presumed ancestral domain structure is known as an immunoglobulin or immunoglobu-

Figure 12.2 (a) The immunoglobulin molecule, IgG, is built up from two copies each of two different polypeptide chains, heavy (H) and light (L). The L chain folds into two domains: V_L with variable sequence between different IgG molecules and C_L with constant sequence. The H chain folds into four domains: one variable V_H and three constant domains, C_{H1} , C_{H2} , and C_{H3} . Disulfide bridges connect the four chains. The antigen binding sites are at the ends of the variable domains. (b) Schematic polypeptide chain structure of different immunoglobulin molecules. IgM is a pentamer where the heavy chains have one variable domain and four constant domains. An additional polypeptide chain, the J chain, (black bar) is associated with the pentameric molecule. The J chain also links two units to form the dimeric IgA. IgG, IgD and IgE are monomeric immunoglobulin molecules.



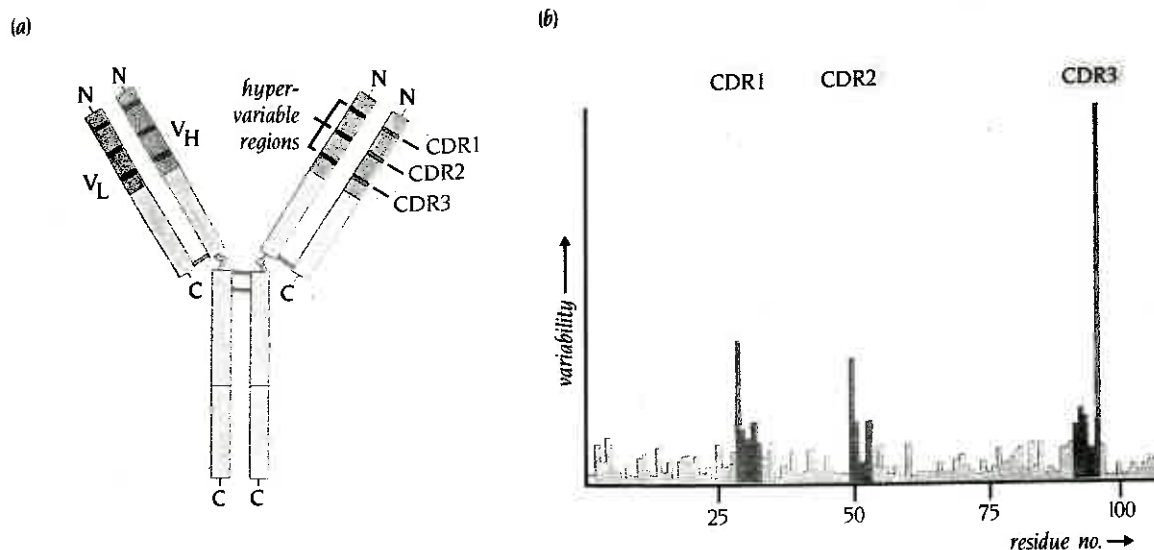


Figure 12.3 Certain regions within the 110 amino acid variable domains show a high degree of sequence variations. These regions determine the antigen specificity and are called hypervariable regions or complementarity determining regions, CDR (a). There are three such regions, CDR1–CDR3, in each variable domain. The sequences of a large number of variable domains in L chains have been compared, and the sequence variability is plotted as a function of residue number along the polypeptide chain in (b). The three large peaks in this diagram correspond to the three hypervariable regions in (a): CDR1–CDR3. [(b) Adapted from T.T. Wu and E.A. Kabat, *J. Exp. Med.* 132: 211, 1970.]

lin-like domain. The T-cell receptor has only just become available in soluble form for crystallization, but the crystal structure of an MHC molecule is now known. We examine the three-dimensional structures of antibodies and their bound antigens and discuss the relationship between the genetic basis for the diversification of antibodies and the structural basis for specific recognition. We also describe strategies for the design of novel antibodies including antibodies with enzymatic function. Finally, we examine the structure of an MHC molecule and show that the structural principles for recognition and binding of antigens to MHC are quite different from those of antibodies and that they reflect the special requirement for co-recognition of MHC and antigen by the T-cell receptor.

The polypeptide chains of antibodies are divided into domains

The basic structure of all immunoglobulin (Ig) molecules comprises two identical light chains and two identical heavy chains linked together by disulfide bonds (Figure 12.2a). There are two different classes, or isotypes, of light chains, λ and κ ; there is no known functional distinction between light-chain isotypes. Heavy chains, by contrast, have five different isotypes that divide the immunoglobulins into different functional classes, IgG, IgM, IgA, IgD, and IgE (Figure 12.2b), each with different effector properties in the elimination of antigen. Each class of heavy chains can combine with either of the two different classes of light chains.

Immunoglobulins of class G, IgG, which is the major type of immunoglobulin in normal human serum and which we will discuss in this chapter, are monomers of the basic unit. Each chain of an IgG molecule is divided into domains of about 110 amino acid residues. The **light chains** have two such domains, and the **heavy chains** have four.

The most remarkable feature of the antibody molecule is revealed by comparison of the amino acid sequences from many different immunoglobulin IgG molecules. This comparison shows that the amino-terminal domain of each polypeptide chain is highly variable whereas the remaining domains have constant sequences for each class. A light chain is thus built up from one amino-terminal variable domain (V_L) and one carboxy-terminal constant domain (C_L) and a heavy chain from one amino-terminal **variable domain** (V_H), followed by three **constant domains** (C_{H1} , C_{H2} , and C_{H3}).

The variable domains are not uniformly variable throughout their lengths: in particular, three small regions show much more variability than the rest of the chain; they are called hypervariable regions or **complementarity determining regions**, CDR1–CDR3 (Figure 12.3). They vary both in size and in

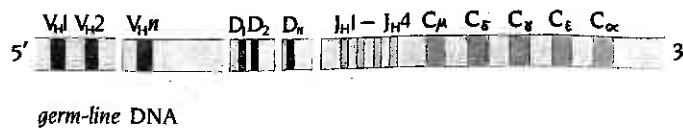


Figure 12.4 Variable domains in immunoglobulins (and T-cell receptors) are made by combinatorial joining of gene segments. Three such segments, V, D, and J, are joined to make the variable domain of a heavy chain. In the mouse the gene-segment pool for an H chain contains about 1,000 V segments, 12 D segments, 4 J segments, and an ordered cluster of C segments, each encoding a different class of H chains.

sequence between different immunoglobulins. These are the regions that determine the specificity of the antigen-antibody interactions. The remaining parts of the variable domains (except a fourth hypervariable region of unknown function in the heavy chain) have quite similar amino acid sequences. In fact, all variable domains show significant overall sequence homology and hence evolutionary relationships among each other. Similarly, the different constant domains—C_L, C_{H1}, C_{H2}, and C_{H3}—show significant (30–40%) sequence identity to each other and to constant domains from immunoglobulin chains of different classes. This strongly suggests that the genes that code for current constant domains arose by successive gene duplication of an ancestral antibody gene. In contrast, there is no significant sequence homology between constant and variable domains, even though, as we will see, there are striking structural similarities. Before describing the structures of the immunoglobulin domains, however, we must pause to outline the genetic mechanisms underlying the variability of immunoglobulin molecules.

Antibody diversity is generated by several different mechanisms

It has been estimated that a human body generates 50 million new antibody-producing B cells every day. The vast majority of these cells produce an antibody molecule that is unique to that cell and its progeny. The genetic information for this diversity is contained in about 1,000 small segments of DNA. The elucidation of the mechanisms for the generation of this **antibody diversity** is one of the truly outstanding achievements of modern biology.

The gene segments encoding the variable and constant regions of antibodies are clustered in three gene pools, one for the class G heavy chain and one for each of the two light-chain isotypes, on separate chromosomes. In the heavy-chain gene pool, the variable domain is encoded by three types of segments, V, D, and J (Figure 12.4). V codes for approximately the first 90 residues, D for the hypervariable region CDR3, and J for the remaining 15 residues of the variable domain.

There are about 1,000 different V segments, about 10 different D segments of variable lengths, and about 4 different J segments in the heavy-chain gene pool. The DNA for the variable domain of a new B cell is assembled by random joining of one of each of these segments into a single continuous exon (Figure 12.5). This process is called **combinatorial joining**, and there are about 40,000

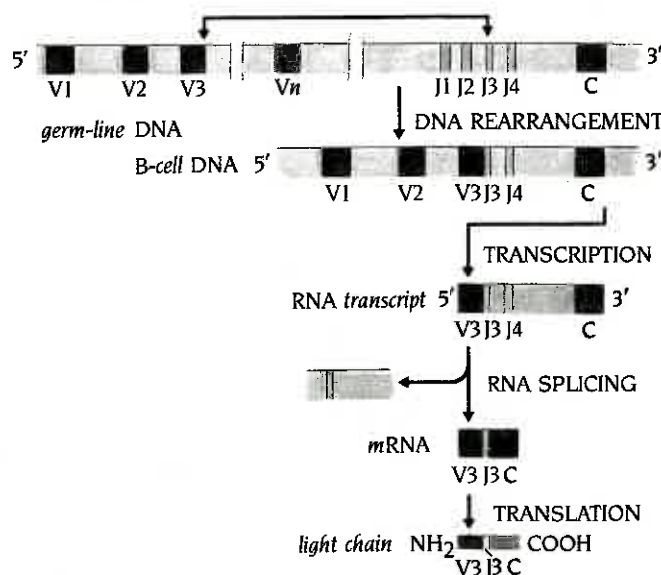


Figure 12.5 The V-J joining process involved in making a κ light chain in the mouse. In the germ-line DNA the cluster of 4 J gene segments is separated from the C gene segment by a short intron and from the about 250 V gene segments by long regions of DNA. During B-cell development the chosen V gene segment (V3) is moved to lie precisely next to the chosen J segment (J3). The “extra” J gene (J4) and the intron sequence are transcribed along with the V3, J3, and C segments and then removed by RNA splicing to generate mRNA molecules.

different ways it can occur. Furthermore, the joining of these segments is not precise. This creates additional diversity, called **junctional diversity**, where D recombines with V and J. In addition, extra nucleotides can be added during the joining procedure, giving still further diversity. The D segment encodes CD3, which is thus a focus of particular diversity in the heavy chain.

The newly created V-D-J exon becomes joined to one of the eight C segments that encode the constant region domains of the heavy chain as separate exons. Usually, a variable region is first expressed in an IgM molecule. Later, further genetic rearrangements will delete the mu (μ) exons encoding the IgM constant region, and the variable region will become joined to a second set of constant-region exons, most commonly those encoding IgG.

The genetic mechanisms underlying light-chain diversity are similar except that there are no D segments, and the diversity of CDR3 therefore depends largely upon junctional diversity generated during V-J joining.

By these genetic mechanisms it has been estimated that at least 90,000 different heavy chains and 3,000 different light chains can be produced. We will see that from structural considerations it is reasonable to assume that almost any heavy chain can combine with almost any light chain to produce a functional antibody. Since the antigen binding sites are built up from the CDR regions of both chains, this ability of light and heavy chains to combine and form functionally viable pairs generates considerable additional diversity. By this combinatorial association the number of different antibody molecules that a new B cell can choose from is increased to $3,000 \times 90,000 = 270$ million. The actual number of antibody molecules with different antigen binding sites is, however, several orders of magnitude higher because of yet one more mechanism for increasing diversity in mature B cells. In these cells point mutations can be introduced into the exons for the variable domains by somatic hypermutation.

All immunoglobulin domains have similar three-dimensional structure

Complete IgG molecules are difficult to crystallize, presumably because of the conformational flexibility in the hinge region between C_{H1} and C_{H2} of the heavy chains (Figure 12.2). When an IgG molecule is treated mildly with the proteolytic enzymes papain or pepsin, however, the heavy chains are cleaved in the hinge region to give two identical Fab fragments and one Fc fragment (Figure 12.6) that can be crystallized relatively easily. The Fab molecule comprises one complete light chain that is linked by a disulfide bridge to a fragment of the heavy chain consisting of V_H and C_{H1} . The Fc molecule comprises C_{H2} and C_{H3} from both heavy chains, which are also linked by disulfide bridges. All detailed high-resolution x-ray structural information on immunoglobulins have been obtained from structure determinations of Fab, Fc, or similar fragments including those of complexes with antigen. Until recently, because of the extreme

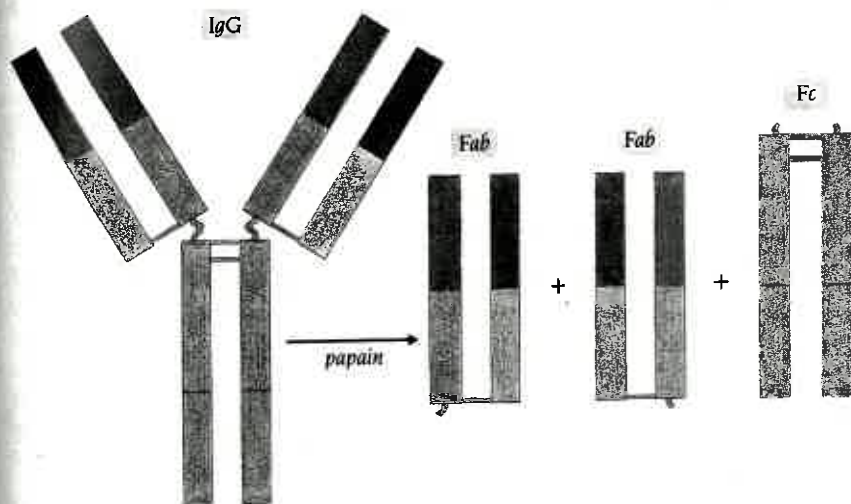


Figure 12.6 Enzymatic cleavage of immunoglobulin IgG. The enzyme papain splits the molecule in the hinge region yielding two Fab fragments and one Fc fragment.

diversity of serum antibodies, the only source of homogeneous immunoglobulins in quantities large enough for sequence and x-ray studies was a naturally occurring tumor, myeloma, which produces large quantities of a single immunoglobulin. Today monoclonal antibodies directed against almost any antigen can be routinely produced in large amounts using hybridoma techniques.

The first detailed description of an immunoglobulin domain was obtained before the advent of monoclonal antibody technology by Allen Edmundson and collaborators at Argonne National Laboratories, U.S.A. The high-resolution structure they determined in 1973 was that of a Bence-Jones protein, a tumor-derived molecule that is a dimer of two light chains. Roberto Poljak and collaborators, at Johns Hopkins Medical School, Baltimore, determined also in 1973 the first high-resolution structure of an intact immunoglobulin Fab fragment, and the same year David Davies and collaborators at the U.S. National Institutes of Health (NIH) demonstrated the binding of an antigenic determinant to the antigen binding site in a Fab fragment. Since then the structures of various other proteolytic fragments of immunoglobulin molecules have been solved to high resolution, and the structures of a few intact IgG molecules have been solved to low or medium resolution. All of the domains, both constant and variable, from both heavy and light chains, have proved to have a similar structure, which has become known as the immunoglobulin fold.

The immunoglobulin fold is best described as two antiparallel β sheets packed tightly against each other

All constant domains are built up from seven β strands arranged so that four strands form one β sheet and three strands form a second sheet (Figure 12.7). The sheets are closely packed against each other and joined together by a disulfide bridge from β strand 6 in the three-strand sheet to β strand 2 in the four-strand sheet. The topology of the sheets is quite simple, as can be seen from the topological diagrams in Figure 12.7. Even though the actual structure consists of two separate β sheets, it is convenient to regard the topology in the form of a Greek key barrel.

The loops are short, and as a result, a majority of the residues of the domain are in the two β sheets. These framework residues have the same structure in all constant domains of all immunoglobulins that have been studied. In other words, the structure is the same in both heavy and light chains.

Most of the invariant residues of the constant domain, including the disulfide bridge, are found in the sheets of the framework region. These invariant residues have two important functions. One is to form and stabilize the framework by packing the β sheets through hydrophobic interactions to give a hydrophobic core between the sheets. Their second function involves interactions between constant domains of different chains to form a complete immunoglobulin molecule.

The remaining residues that are not part of the framework region form the loops between the β strands. These loops may vary in length and sequence between immunoglobulin chains of different classes but are constant within each class; the sequence of the loops is invariant. The functions of these loops are not known, but they are probably involved in the effector functions of antibodies. When an antibody-antigen complex has been formed, signals are mediated through the Fc region to different systems, such as phagocytic cells or the complement system, which will destroy the complex. These activities are triggered by specific interactions of ligands with the constant domains of immunoglobulins.

The overall structure of the variable domain is very similar to that of the constant domain but there are nine β strands instead of seven. The two additional β strands are inserted into the loop region that connects β strands 3 and 4 (red in Figure 12.8). Functionally, this part of the polypeptide chain is important since it contains the hypervariable region CDR2. The two extra β strands provide the framework that positions CDR2 close to the other two hypervariable regions in the domain structure (Figure 12.8).

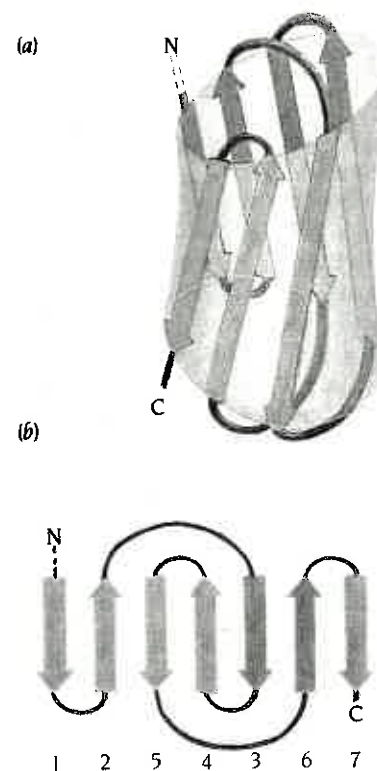


Figure 12.7 The constant domains of immunoglobulins are folded into a compressed antiparallel β barrel built up from one three-stranded β sheet packed against a four-stranded sheet (a). The topological diagrams (b) show the connected Greek key motifs of this fold.

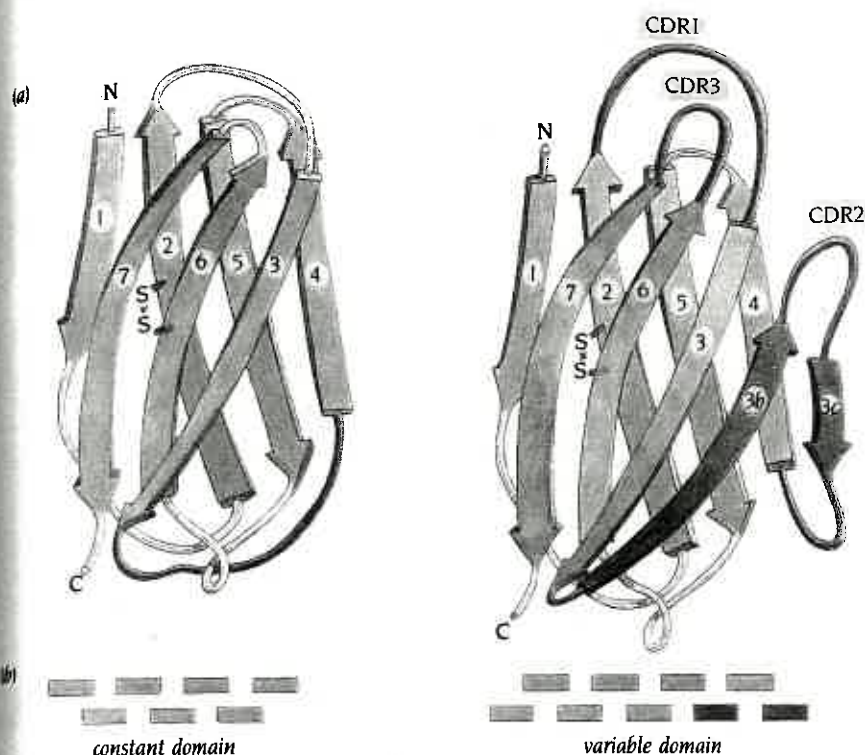


Figure 12.8 (a) Comparison of the structures of the constant and variable domains of immunoglobulins. β strands labeled 1–7 have the same topology and similar structures. There are two extra β strands, 3b and 3c (red) in the variable domain. The loop between these strands contain the hypervariable region CDR2. The remaining CDR regions are at the same end of the barrel in the loops connecting β strands 2 and 3 and strands 6 and 7. A disulfide bond bridges strand 2 in one sheet with strand 6 in the other sheet in both the constant and the variable domains. (b) The β strands viewed end on, illustrating that one β sheet has the same four β strands in the two domains.

The hypervariable regions are clustered in loop regions at one end of the variable domain

The specificity of immunoglobulins is determined by the sequence and size of the hypervariable regions in the variable domains. From amino acid sequence comparisons of several hundred different variable domains these hypervariable regions have been defined as amino acids 24–34, 50–56, and 89–97 in the light chains and 31–35, 50–65, and 95–102 in the heavy chains. In the three-dimensional structures these residues occur in the loop regions that connect β strands 2–3, 3b–3c, and 6–7. We can see in Figure 12.8 that these loop regions are clustered together in space, at one end of the β sheets. CDR2 and CDR3 are hairpin loops between β strands 3b–3c and 6–7 in the five-strand β sheet. The third hypervariable region CDR1 is a crossover connection from β strand 2 in the four-strand sheet to strand 3 in the five-strand sheet. All strands in the five-strand sheet thus contribute loops to the hypervariable regions, whereas only one strand from the four-strand sheet is involved. The additional loop at this end of the sheet, between β strands 4 and 5 in the four-strand sheet, does not have variable sequence and does not participate in antigen binding. It is thus apparent that loops from the five-strand β sheet contribute more residues to the complete antigen binding site than the four-strand sheet.

The variable domains of immunoglobulins are excellent examples of the important structural principle discussed in Chapter 4 in connection with α/β -barrel structures. Functional residues in protein molecules are frequently provided by loop regions that are attached to one end of a stable structural framework built up mainly by secondary structure elements. In α/β -barrel structures this framework is built from eight parallel β strands surrounded by eight helices, whereas in the variable domains of immunoglobulins it consists of nine antiparallel β strands arranged in two sheets packed against each other. The fact that there exists an almost unlimited number of sequence variations in the loop regions of antibodies demonstrates that the folding of the structural framework is quite insensitive to changes in the functional loop regions. Such loop regions are ideal candidates for attempts to design novel protein molecules with predetermined function by recombinant DNA techniques. It has already been shown by these techniques that stable domains are formed when CDR loops from one species are inserted into a framework from a different species.

However, in order to engineer a predetermined function, it is necessary to know how constant the β -sheet framework is and how accurately one can predict and model the conformations of CDR loops.

Comparison of the known structures of variable domains have shown that about 70 residues in the light chain and about 80 residues in the heavy chain form the β -sheet frameworks that have very similar structures. Therefore, all variable domains would be expected to have the same framework structures for homologous residues. Given the framework, the key question to all attempts to model variable domains of known sequence and unknown structure is then the following: Can the structure of CDR loops vary randomly, or are there certain preferred conformations that can be deduced from the lengths and sequences of these regions?

Cyrus Chothia at MRC, Cambridge, U.K., and Arthur Lesk at EMBL, Heidelberg, have attempted to answer this question by a careful analysis of the CDR-loop conformations in the known structures. For each observed conformation they identified those residues that were mainly responsible for maintaining the conformation, either through interaction with framework residues or through their ability to assume unusual conformations. Examples of the latter are Gly or Pro. For a preserved conformation these residues should be conserved or very conservatively substituted. Examination of the sequences of variable domains of unknown structure showed that many have CDR loops that are similar in size to those of one of the known structures and contain identical residues at the sites responsible for the observed conformation. These results have interesting implications for the molecular mechanisms involved in the generation of antibody diversity, in addition to their value in modeling and design of new antibodies. For five of the six hypervariable regions of most immunoglobulins (CDR3 of the heavy chain is an exception) there seems to be only a small repertoire of main-chain conformations, most of which are known from the set of immunoglobulin structures so far determined. Sequence variations within the hypervariable regions modulate the surface that these canonical structures present to the antigens. Sequence variations within both the framework and the hypervariable regions shift the canonical structures relative to each other by small but significant amounts.

We will give three examples from their analysis, which is further described in Chapter 16. The loop between β strands 3b and 3c in CDR2 of the light chain is three residues long in all known sequences. This loop has very similar conformation in the known structures, and 95% of the 250 known sequences of mouse and human light chains fit the sequence constraints of this canonical structure. There is thus a high probability that this loop has the same conformation in almost all immunoglobulins. Four different conformations of CDR2 from the heavy chain has been found in the 10 x-ray structures that were analyzed. For each of these conformations there are specific constraints at one or two positions in the amino acid sequence of this hypervariable region, which ranges from three to six residues. For 236 of the 300 known sequences of heavy chains the sequence of the CDR2 loop fits one of these constraints and would therefore be expected to have one of these four canonical CDR2-loop structures. The CDR3 loop of the heavy chain, however, varies in length from 6 to 14 residues in the known sequences and forms a long hairpin loop. For such large loops many conformations are possible. The one actually found will depend both on the size of the loop and on the packing against the rest of the protein. This loop, therefore, is difficult to predict and model correctly.

The antigen binding site is formed by close association of the hypervariable regions from both heavy and light chains

Fab fragments contain one arm of the IgG molecule with an intact antigen binding site. In these molecules, as in intact IgG, the light chain is associated with the heavy chain in such a way that C_L associates with C_{H1} and V_L with V_H (Figure 12.9). These associations are very tight and extensive. The segments of the light and heavy polypeptide chains that join V_L to C_L and V_H to C_{H1} have

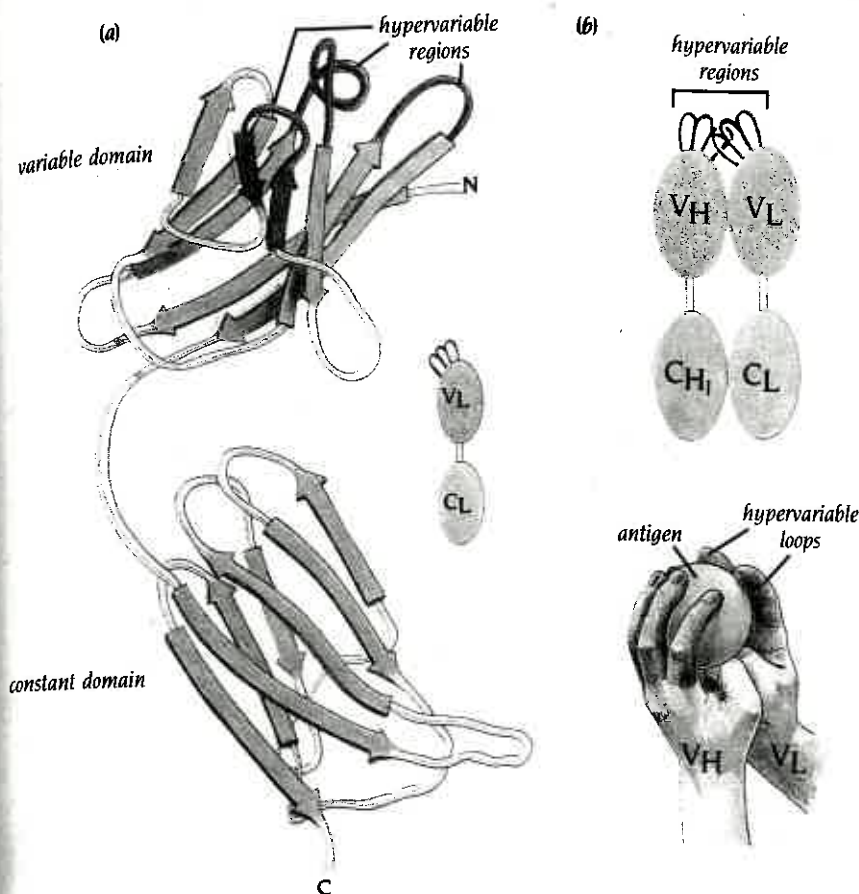


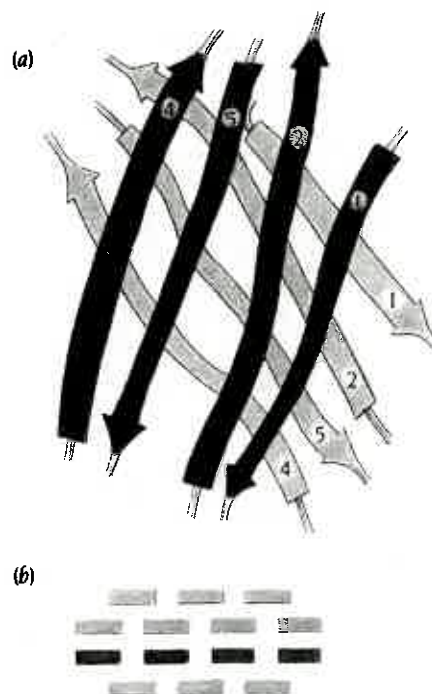
Figure 12.9 (a) The variable and constant domains in the light chain of immunoglobulins are folded into two separate globular units. In both domains the four-stranded β sheet is blue and the other sheet is green. The hypervariable CDR regions are at one end of this elongated molecule. (b) In the Fab fragment as well as in the intact immunoglobulin molecules the domains associate pairwise so that V_H interacts with V_L and C_{H1} with C_L. By this interaction the CDR regions of both variable domains are brought close to each other and together form the antigen binding site.

few contacts. As a result, the Fab fragment consists of two globular regions, one formed by the two constant domains and the other by the two variable domains. Together, they make an elongated molecule. The hypervariable regions of both the light and heavy chains are close together at one end of this elongated molecule. We will now examine more closely how these domains associate.

The constant domains associate by interactions between the four-stranded β sheets in both C_L and C_{H1}. The sheets are closely packed with mainly hydrophobic residues in the interface. The directions of the β strands in the two domains are almost at right angles to each other (Figure 12.10). This mode of packing β sheets against each other occurs frequently within the core of single domains that are built up from two β sheets, for example, in retinol-binding protein discussed in Chapter 5, but they are rather unusual for domain-domain associations. The effect of this packing is to bury the central regions of the four-strand β sheets away from the solvent. Each four-strand β sheet has the three-strand β sheet of the same domain on one side and the four-strand β sheet from the second domain on its other side (Figure 12.10). This places rather stringent constraints on the sequence of these regions and most of the residues that occur in the interface are, in fact, conserved. In addition, residues that form the hydrophobic core within each domain are in general conserved.

The variable domains associate in a very different manner. It is obvious from Figure 12.11 that if they were associated in the same way as the constant domains, via the four-strand β sheets, the CDR loops, which are linked mainly to the five-strand β sheet, would be too far apart on the outside of each domain to contribute jointly to the antigen binding site. Thus in the variable domains the five-strand β sheets form the domain-domain interaction area. Furthermore, the relative orientation of the β strands in the two domains is closer to parallel than in the constant domains (Figure 12.11), and the curvature of the five-strand β sheets is such that they do not pack tightly against each other. Instead, each sheet forms half a barrel. When the two sheets associate, four of the five

Figure 12.10 Schematic diagrams of the packing of the four-stranded β sheets of the constant domains C_{H1} and C_L in an Fab fragment of IgG. The sheets are viewed perpendicular to the β strands in (a) and end-on in (b), where the four-stranded β sheet is blue.



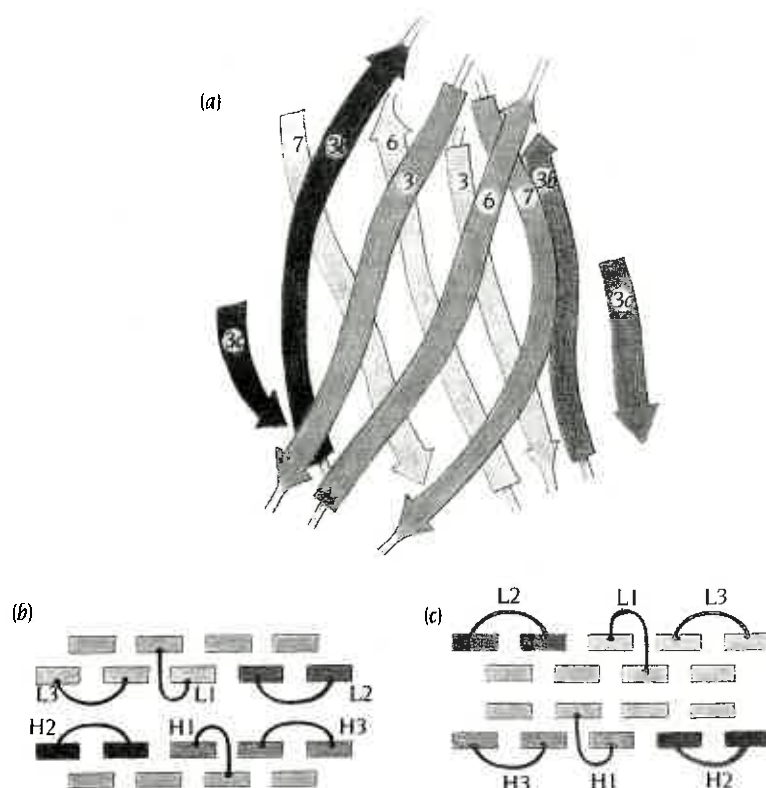


Figure 12.11 The two variable domains V_H and V_L are packed against each other so that the six hypervariable regions are close to each other. (a) Schematic diagram of the packing of the five-stranded β sheets of V_H and V_L in IgG. Only four of the β strands are involved in packing the variable domains against each other. Strand 3c is not involved. (b) Schematic diagram of the strands viewed end-on in the variable domain. Hypervariable loop regions are red. (c) Diagram illustrating hypothetical packing of the variable domains through their four-stranded sheets. The hypervariable regions would be far apart.

strands from both domains complete a barrel structure of eight antiparallel β strands (Figure 12.12).

The sheets associate in such a way that CDR loops from both sheets are at the same end of the barrel. These six CDR loops form the complete antigen binding site, which extends over a large surface outside the top of the barrel (Figure 12.13).

The antigen binding site binds haptens in crevices and protein antigens through large flat surfaces

The six CDR loops from the two chains at the rim of the eight-strand barrel provide an ideal arrangement for generating antigen binding sites of different shapes depending on the size and sequence of the loops. They can create flat extended binding surfaces for protein antigens or specific deep binding cavities for small hapten molecules.

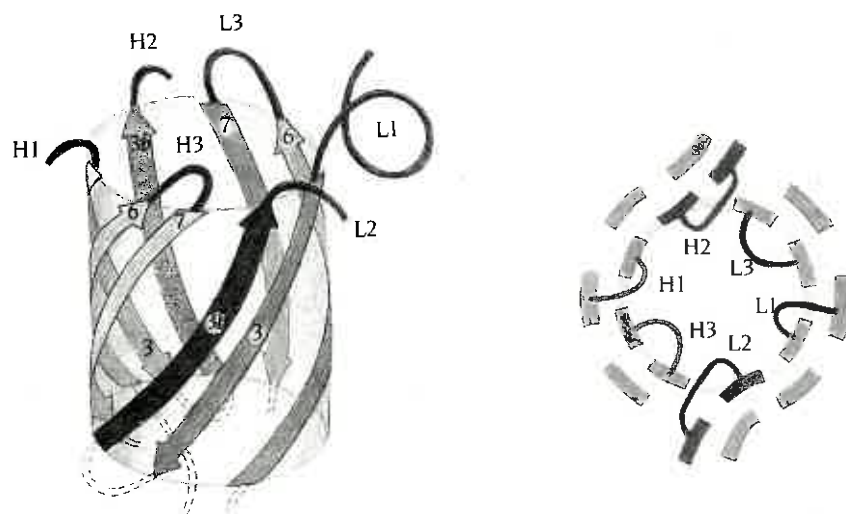


Figure 12.12 Schematic diagram of the barrel arrangement of four β strands from each of the variable domains in Fab. The six hypervariable regions, CDR1–CDR3 from the light chain (L1–L3) and from the heavy chain (H1–H3), are at one end of this barrel. (From J. Novotny et al., *J. Biol. Chem.* 258: 14435, 1983.)

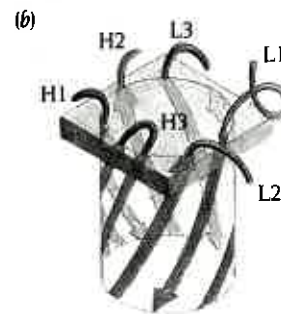
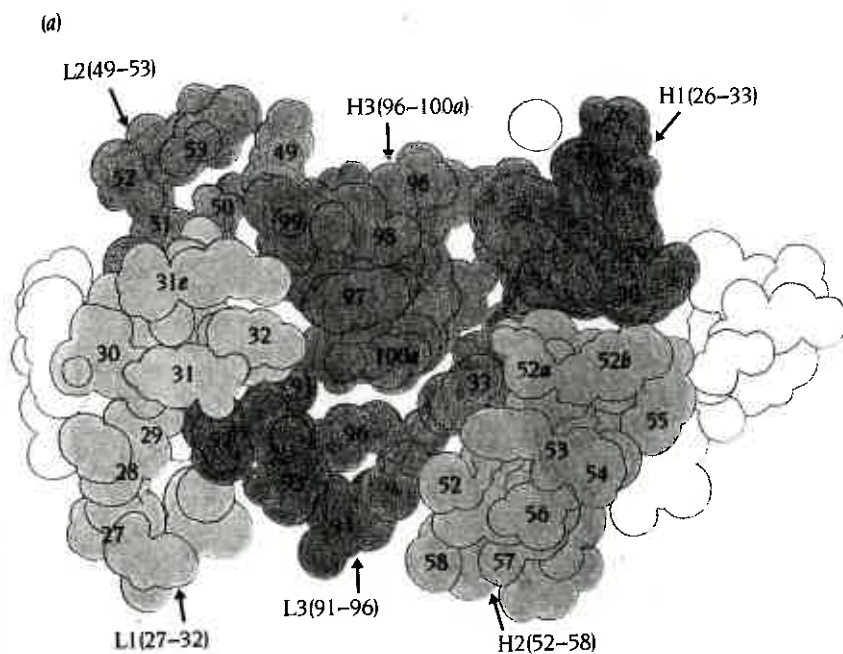


Figure 12.13 Drawing of a space-filling model (a) of the hypervariable regions of an Fab fragment. The superpositions of five sections are shown, cut through a model as shown in (b). It is clearly seen that all six hypervariable regions (L1-L3, H1-H3) contribute to the surface shown here. (From C. Chothia and A. Lesk, *J. Mol. Biol.* 196: 914, 1987.)

Haptens, a special class of antigens, are small molecules that induce specific antibody production when they are attached to a protein that acts as a carrier. Phosphorylcholine is one such. It has been widely used in the investigation of immune responses. The specific binding of this hapten to an Fab fragment has been studied crystallographically by David Davies and co-workers at NIH, U.S.A. The binding cavity is about 15 Å wide at the mouth and 12 Å deep. It is lined by residues from all CDR loops except the short CDR2 from the light chain. A number of side chains from these loops bind phosphorylcholine (Figure 12.14) in a manner that is very similar to that of a substrate or inhibitor binding to an enzyme.

In enzyme-catalyzed reactions substrates in their transition state bind very tightly to the enzyme and are positioned close to polar or charged side chains that participate in catalyzing the enzymatic reaction. It has recently been

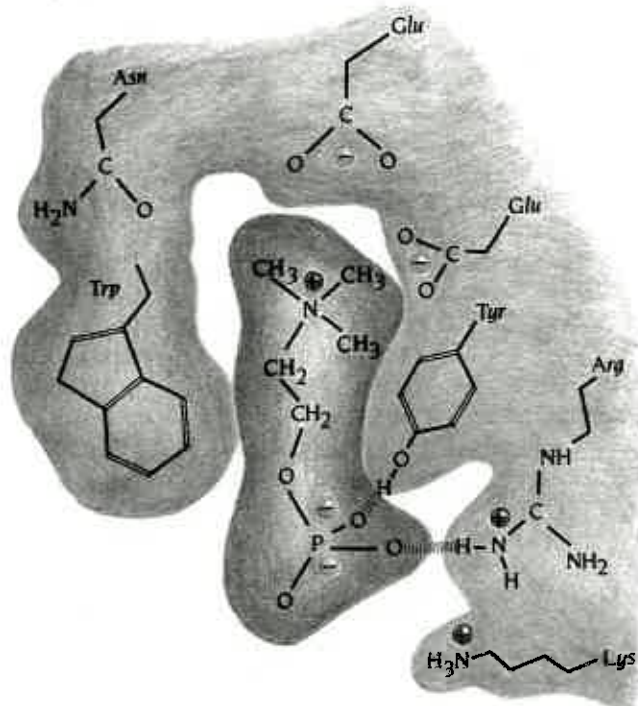


Figure 12.14 Schematic representation of the specific interactions between phosphorylcholine (brown) and the protein side groups (green) in Fab. The binding cavity is in a cleft between the light and the heavy chains. Choline binds in the interior while the phosphate group is toward the surface. (Adapted from E.A. Padlan et al., *Immunochemistry* 13: 945, 1976.)

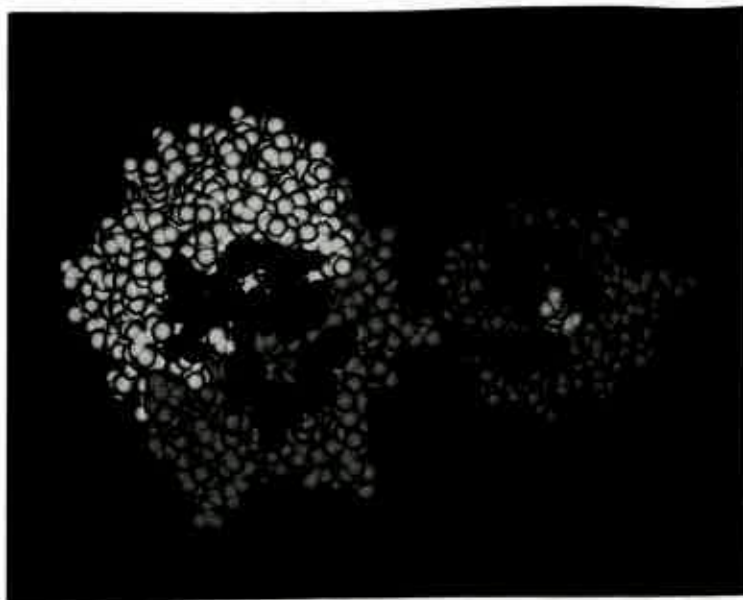
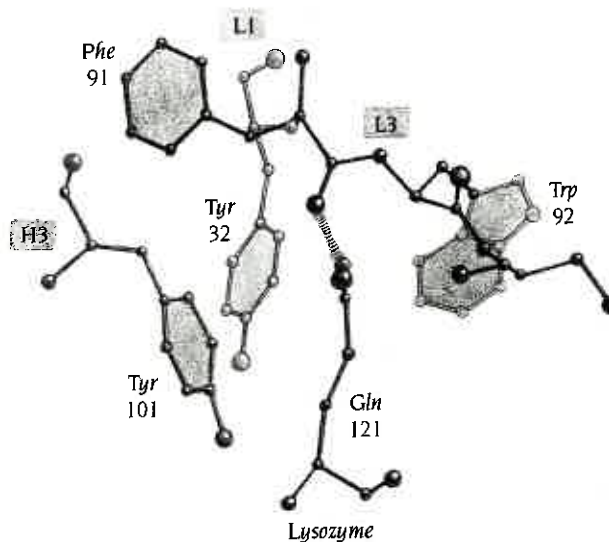
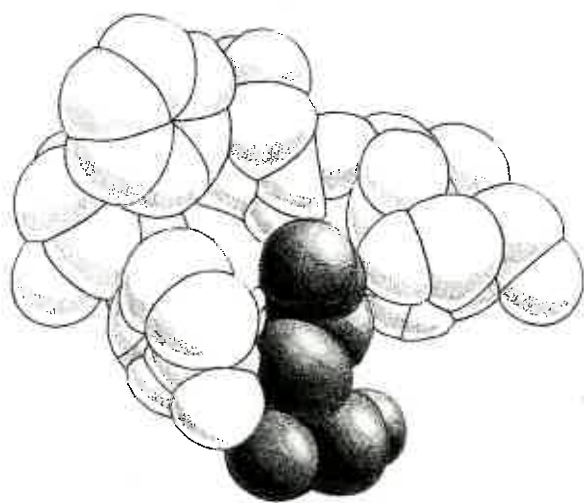


Figure 12.15 Space-filling representation of a complex between lysozyme (green) and the Fab fragment of a monoclonal antilysozyme (blue and yellow). The Fab fragment and the antigen (lysozyme) have been separated in this diagram, and their combining surfaces are viewed end-on. Atoms that are in contact in the complex are colored red both in Fab and lysozyme, except Gln 121 in lysozyme, which is violet. The diagram illustrates the large size of the interaction surfaces. (After A.G. Amit et al., *Science* 233: 749, 1986; courtesy of R. Poljak.)

possible to induce production of antibodies with enzymatic activity by choosing haptens that are analogues of the transition state of the substrate for a chemical reaction. When a transition-state analogue is used as a hapten, it induces an array of different antibodies that bind the hapten specifically. Some of these antibodies happen to have charged or polar side chains suitably placed to promote catalytic reactions of a corresponding substrate. These antibodies can be isolated as monoclonals and improved upon by recombinant DNA techniques. This approach seems to be one promising way to convert antibodies into enzymes with completely novel catalytic functions not previously present in living organisms.

The analysis of hapten binding to antibodies was possible before monoclonal antibodies became available because some myeloma proteins bind haptens. Thus hapten binding was studied as early as 1974. Once monoclonal antibodies with any chosen specificity could be produced, it became possible in principle to study the interactions of antibodies with protein antigens. However, the path from a cell line to a high-resolution three-dimensional antigen-antibody structure is long, and not until 1986 was the first such investigation reported by Roberto Poljak and co-workers at the Pasteur Institute, France. They chose a lysozyme-Fab complex because the structure of the enzyme lysozyme was known and its antigenic properties had been extensively studied.

The shape of the interaction area between lysozyme and the CDR loops of the antibody is quite different from the hapten-binding crevice. The interaction extends over a large area with maximum dimensions of about $20 \times 30 \text{ \AA}$ (Figure 12.15). It is an irregular rather flat surface with small protuberances and depressions that are complementary in the antigen and the antibody. Residues from all six CDR loops contribute to the antibody surface and residues from two stretches of the polypeptide chain of lysozyme, 18–27 and 116–129, form the antigen surface. There are 17 antibody residues that make close contact with 16 lysozyme residues. These residues pack against each other with a density similar to that found in the interior of protein molecules. Exchange of some of these residues would sterically interfere with packing or abolish bonding contacts and therefore cannot be tolerated. For example, Gln 121 in lysozyme protrudes from the lysozyme surface and fits snugly into a cleft between the CDR3 loops of V_H and V_L (Figure 12.16). On the other hand, there are some imperfections in the form of holes in the interaction area. Appropriate changes in the CDR loops could fill these holes and produce an even better fit. This is the sort of adjustment that can presumably occur through somatic mutations in the antibody genes and would account for the observation that a second immunization often gives rise to higher-affinity antibodies.



Two of the 17 antibody residues that make contact with the lysozyme antigen are from framework regions. These are residues bordering CDR in the sequence, indicating that the functional distinction between framework and CDR residues is not absolute in the border between the two regions. But at the physical center of the interaction area is CDR3 of the heavy chain, which makes a proportionally greater contribution to the formation of this area than any of the other CDR loops. Four residues in the heavy-chain CDR3 participate in interactions with lysozyme, and they are all encoded in the D segment, the special significance of which now becomes clear. CDR3 of the heavy chain is the principal focus of the somatic mechanisms for generating antibody diversity and is central to the specific recognition of protein antigens. It is also the loop that eludes the predictive formula of Chothia and Lesk.

The surface of lysozyme that interacts with the antibody is built up by residues from two different regions of the polypeptide chain and is critically dependent on an intact three-dimensional structure of the protein. Yet immunization with isolated peptides from a protein molecule can induce antibodies that also bind to the intact protein. The peptides that induce these antibodies must presumably adopt the same conformation as in the intact protein; otherwise the known structures of protein-antibody complexes tell us nothing about the mechanism of this recognition.

The structure of lysozyme in the complex is the same as that in crystals of free lysozyme. No conformational changes of the main chain are seen even in the regions that bind to the antibody. The binding area in lysozyme, the antigenic determinant, is in regions exposed to solvent, but these regions are not particularly flexible. The structure of the Fab fragment complexed to lysozyme is very similar to the structure of other Fab fragments not bound to their corresponding antigen. This implies that no gross conformational changes occur upon antigen binding. This is only partly true for a second antigen-antibody complex, the structure of which was determined by Peter Colman and collaborators, CSIRO, Melbourne, Australia, where the antigen was the influenza virus protein neuraminidase, whose structure is described in Chapter 5. In this complex there is a difference within the interaction area between V_H and V_L such that the CDR loops have been displaced about 3 Å compared with other Fab structures. It is not known if this difference is due to antibody variability or to changes induced by antigen binding. In addition, the structure of the antigen, neuraminidase, has been somewhat distorted so that it is no longer catalytically active. On the whole, however, it seems that the classical "lock-and-key" description is an adequate first-order simplification to describe the interaction of antibodies with protein antigens. There is no evidence so far of major conformational changes induced by binding of antigen to antibody, but small local structural changes may occur to increase the complementarity between the interacting surfaces of antigen and antibody.

Figure 12.16 Detailed views of the environment of Gln 121 in the lysozyme-antilysozyme complex. Gln 121 in lysozyme is colored green both in the space filling representation to the left and in the ball and stick model to the right. This side chain of the antigen fits into a hole between CDR3 regions of both the heavy (Tyr 101) and the light (Trp 92) chains as well as CDR1 from the light chain (Tyr 32). (After A.G. Amit et al., *Science* 233: 751, 1986.)

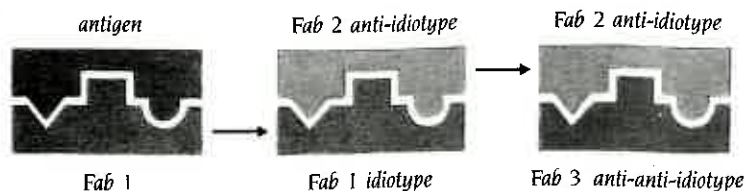


Figure 12.17 Schematic diagram illustrating the idiotype-anti-idiotype reaction.

The structure of an idiotype-anti-idiotype complex has been determined

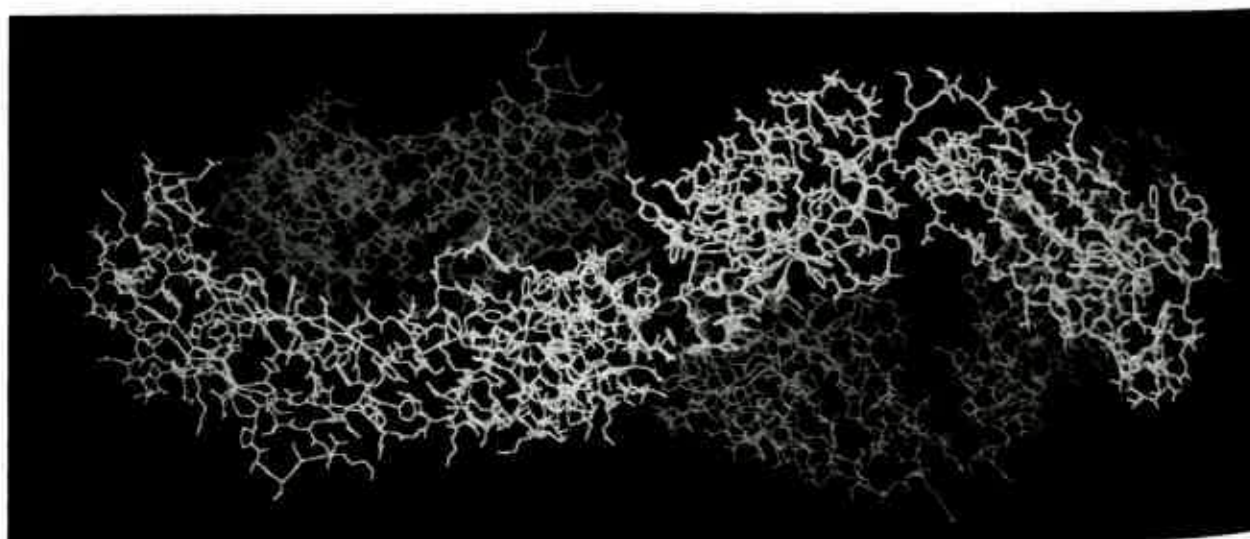
The characterization of the interaction between antigen and antibody as a flat surface provides a satisfactory structural explanation of idiotype-anti-idiotype reactions in immunology. An idiotope is an antigenic determinant in the variable domains of an antibody. An idiotype in antibody 1 can induce the production of antibody 2 whose CDR loops bind to the CDR loops of antibody 1. This is called an **idiotype-anti-idiotype** reaction (Figure 12.17). Networks of anti-idiotype and anti-anti-idiotype interactions have been proposed as a mechanism for regulating immune responses, but extensive interactions of this sort were hard to envisage on the basis of the cleftlike antigen binding site suggested by the phosphorylcholine complex. However, an interaction area in the form of a flat surface can easily accommodate idiotype-anti-idiotype reactions.

It is also easily imaginable that the anti-idiotype of antibody 2 can structurally mimic the antigenic determinant that induced antibody 1, an idea that has provided the basis for an innovative strategy for isolating receptor molecules by raising antibodies against their ligands and then raising anti-idiotypic antibodies against those antibodies. If the anti-ligand antibody mimics the binding site of the receptor, then an anti-idiotypic antibody may recognize the original binding site in the receptor molecule.

This attractive hypothesis has only partly been validated by a recent structure determination by the group of Roberto Poljak of an idiotype-anti-idiotype complex. Poljak has managed to crystallize the complex between the Fab fragment shown in Figure 12.15 and the Fab fragment of one of its anti-idiotype immunoglobulins. Since the first Fab fragment was derived from a monoclonal antibody against lysozyme, it was expected that its anti-idiotype would mimic lysozyme and show similar interactions with the idiotype. This is not the case.

The structure of the complex shows that the CDR regions of the idiotype interacts with the CDR regions of the anti-idiotype as predicted (Figure 12.18). The details of these interactions are, however, not the same as in the idiotype-lysozyme complex. The anti-idiotype Fab fragment is oriented in the complex in such a way that its L chain occupies the same interaction area with the

Figure 12.18 The three-dimensional structure of a complex between Fab fragments from an idiotype immunoglobulin and its anti-idiotype. The idiotype Fab fragment (L chain, yellow; H chain, orange) is from a monoclonal antibody against lysozyme. The anti-idiotype Fab fragment (L chain, light blue; H chain, dark blue) is from a monoclonal antibody against the idiotype. The interaction between these two Fab fragments takes place largely through the hypervariable regions of both fragments. (Courtesy of R. Poljak.)



idiotype as lysozyme does in its complex. The interactions between the idiotype and the H chain of the anti-idiotype have no counterpart in the lysozyme complex. The amino acid residues from the anti-idiotype L chain that are involved in this interaction are different from those of lysozyme in its interaction with the idiotype. Details of this interaction are not available at the time of writing this book, since the structure is not refined, but it is quite clear that the binding interactions are different in the lysozyme complex compared to the anti-idiotype complex.

An IgG molecule has several degrees of conformational flexibility

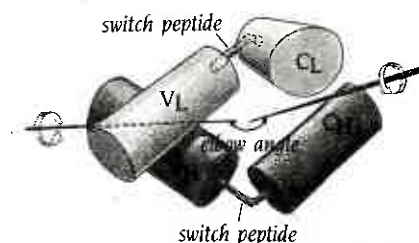
Early physicochemical and electron microscopy studies suggested that the Fab arms were flexible and could move relative to each other and to the Fc stem of intact IgG molecules. The region of the heavy chains that connects C_{H1} with C_{H2} and therefore Fab with Fc is called the **hinge region**. The amino acid sequence of this region includes the cysteine residues that form S-S bridges between heavy chains and has a high proportion of proline residues that frequently are found in flexible regions of proteins. The flexibility of the hinge region is important for the function of immunoglobulins, but it makes immunoglobulins difficult to crystallize. Therefore, it has been expedient to use antibodies from which the hinge has been deleted.

Medium-resolution x-ray studies of an IgG molecule lacking its hinge region, and therefore with Fc rigidly attached to Fab, has provided the picture of an IgG molecule shaped like a T with the Fab arms at right angles to the Fc stem; this structure, however, may not be representative for a functionally intact IgG molecule that includes the hinge region. The C_{H2} and C_{H3} domains of each of the heavy chains associate in $C_{H2}:C_{H2}$ and $C_{H3}:C_{H3}$ pairs. The C_{H3} pair associate in the same way as the $C_{H1}:C_L$ pair discussed previously. In contrast, there is no protein-protein interaction between the C_{H2} domains: instead, a carbohydrate chain is attached to each C_{H2} domain in the interface region and forms a weak bridge between them.

In crystals of IgG molecules with intact hinge regions only the Fab moiety and the beginning of the hinge region are visible: the C-terminal residues of the hinge region and the Fc parts of the molecules are disordered. From a careful study of one such molecule and a proteolytic fragment containing two joined Fab arms, Robert Huber and co-workers in Munich, Germany, have deduced that there are at least four different orientations of the Fc region in the crystal. The crystal thus contains a random mixture of at least four different conformations of the same IgG molecule. Furthermore, the orientation of the two Fab arms is such that the intact IgG molecule must resemble a Y more than a T. This shows that the hinge region is a flexible link that allows the Fab arms to move relative both to each other and to Fc. This flexibility has important functional consequences. It means that the two antigen binding sites at the tips of the Fab arms can move relative to each other to bind antigenic determinants separated by different distances and in different orientations. It also provides flexibility to the Fc region relative to the Fab that facilitates the effector functions of IgG.

There is another flexible region in the IgG molecule that is functionally more obscure. The region of the polypeptide chain that links the variable domain to the adjacent constant domain, the so-called **switch peptide**, can have different conformations. As a result the variable and constant regions of Fab are oriented differently with respect to each other. This orientation can be defined by the angle between the local twofold axes that relates V_H to V_L and C_{H1} to C_L , respectively (Figure 12.19). This "elbow" angle is different in different Fab structures and varies from 132° to 172° . Even the same Fab molecule can have different "elbow" angles in different crystal forms. It has been suggested that antigen binding might trigger a change in the "elbow" angle and thereby transmit a signal through C_{H1} to the Fc effector region. There is, however, no experimental evidence to substantiate this idea. In contrast, no correlation has

Figure 12.19 Schematic diagram illustrating the "elbow" angle between the variable and constant domains in immunoglobulins.



been observed between the size of the "elbow" angle and antigen binding. These results strongly suggest that the observed spread of "elbow" angles reflects an inherent flexibility in Fab with no specific biological function.

The structure of a human MHC molecule has provided insights into the molecular mechanism of T-cell activation

Few protein crystal structures have had such a direct impact as that of the human class I **MHC antigen**, human lymphocyte antigen A2 (HLA-A2), which was determined by Pamela Björkman in the group of Don Wiley at Harvard University. MHC molecules are crucial to the activation of T cells, which play a major role in immune defense. **T-cell activation** is initiated when the antigen receptors on the surface of these cells bind foreign antigens in combination with MHC molecules on the surface of other cells. T-cells are also activated by foreign MHC molecules on the surface of cells from a genetically different individual. An individual's own MHC molecules do not normally activate that individual's T cells because all T cells bearing potentially self-reactive receptors are eliminated during ontogeny.

There have been major questions in understanding these observations at the molecular level, primarily the following: How can a very limited number of MHC molecules recognize and bind an almost unlimited number of antigens and present them to T-cell receptors in such a way that they discriminate between self and nonself? Do the T cells see antigen and MHC molecules, which are firmly attached to the surface of cells, as a single structure or as two separate entities? These questions were immediately answered by the structure of one HLA molecule, in combination with many years of detailed molecular genetic analysis of T-cell-MHC interactions. Furthermore, the structure allows us now to ask very specific questions about the immune system and answer them by site-directed mutagenesis of MHC molecules.

Recognition of antigen is different in MHC compared to immunoglobulins

The class I MHC molecules are plasma membrane proteins expressed in all cells and which are composed of two polypeptide chains: a heavy chain that spans the membrane bilayer and a noncovalently attached light chain. The light chain is invariant, but the class I MHC heavy chain is the most genetically **polymorphic protein** known. Each human individual has three class I genes, and each of those genes has a very large number of alleles so that any given individual is likely to have six different class I molecules out of an estimated total of about 100 in the human population. Exactly how this polymorphism contributes to the recognition of antigen by T cells is a major question to which a general answer emerged from the HLA-A2 structure.

The light chain of the class I heterodimer, which is called β_2 microglobulin (β_2m) has amino acid sequence homology to the immunoglobulin constant domains and essentially the same three-dimensional structure. The extracellular portion of the heavy chain is divided into three domains, $\alpha 1$, $\alpha 2$, and $\alpha 3$, each about 90 amino acids long. The carboxy-terminal domain $\alpha 3$ also shows sequence and structural homology to immunoglobulin constant domains, whereas $\alpha 1$ and $\alpha 2$ have quite different structures. The extracellular portion of HLA-A2 composed of $\alpha 1$, $\alpha 2$, $\alpha 3$, and β_2m was removed from the cells by proteolytic digestion and crystallized for high-resolution x-ray analysis.

The structure is divided into two globular regions (Figure 12.20). One region is built up from the two immunoglobulin-like domains $\alpha 3$ and β_2m . Since $\alpha 3$ is attached to the membrane in the intact HLA molecule, this region is closest to the cell surface. These domains are paired through their four-strand β sheets but in detail quite differently from the pairing of the constant domains $C_{H1}:C_L$ and $C_{H3}:C_{H3}$ in IgG.

The second region, which presumably is facing away from the surface and which contains the antigen binding site, is composed of domains $\alpha 1$ and $\alpha 2$.

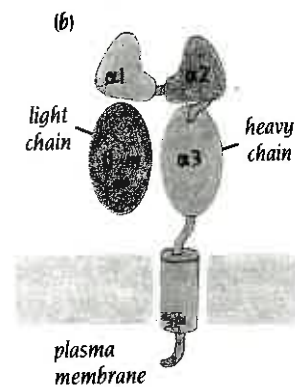
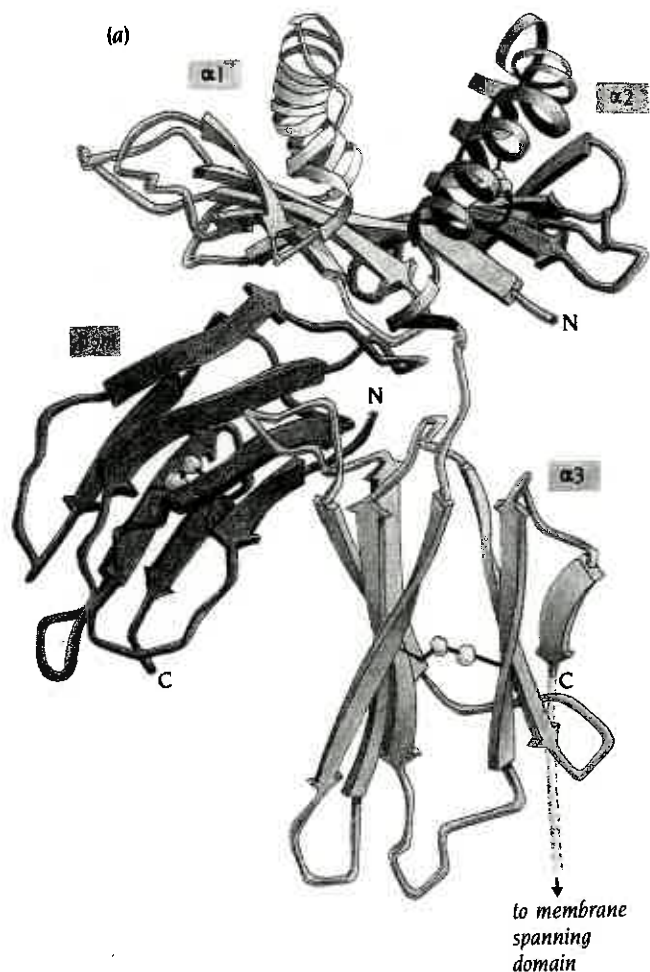


Figure 12.20 (a) Schematic representation of the path of the polypeptide chain in the structure of HLA-A2. Disulfide bonds are indicated as two connected spheres. The molecule is shown with the membrane proximal immunoglobulin-like domains ($\alpha 3$, $\beta_2 m$) at the bottom and the polymorphic $\alpha 1$ and $\alpha 2$ domains at the top. (b) The domain arrangement in HLA-A2. (Adapted from P.J. Björkman et al., *Nature* 329: 506, 1987.)

Each domain has a very similar structure, which is quite simple (Figure 12.21). Starting from the N terminus the chain forms four up-and-down antiparallel β strands called a "W" followed by a helical region across the β strands on one side of the β sheet. The two domains associate by their "W" regions in such a way that they are hydrogen bonded to each other in an antiparallel fashion. By this association the structure of the complete region has a "floor" of a continuous eight-strand antiparallel β sheet (Figure 12.21). This floor sits on top of the immunoglobulin-like domains $\alpha 3$ and $\beta_2 m$. The two helical regions are above the floor almost parallel to each other and separated by a large distance, about 18 Å from center to center. A large crevice that faces the solution is thus formed

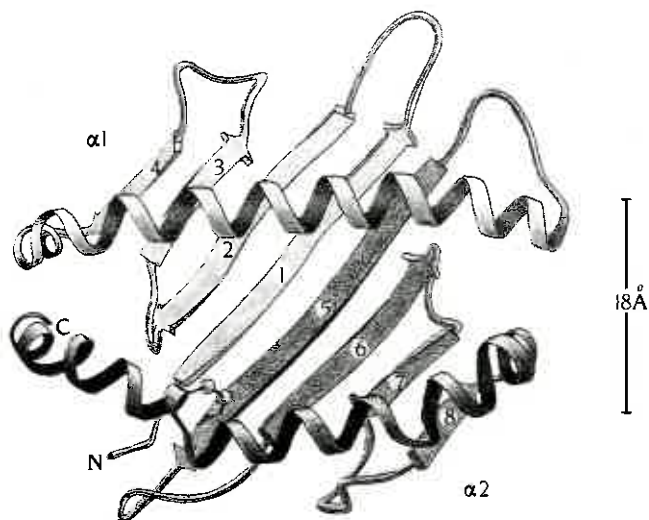


Figure 12.21 Schematic representation of the top surface of HLA-A2. The $\alpha 1$ (yellow) and $\alpha 2$ (green) domains are viewed from the top of the molecule, showing the empty antigen binding site as well as the surface that is presumably contacted by a T-cell receptor. (Adapted from P.J. Björkman et al., *Nature* 329: 512, 1987.)

with the floor forming its bottom and the helices its sides. This crevice is the antigen binding site. In the actual structure that was determined this crevice was occupied by an unknown antigen, probably a short peptide.

It is immediately obvious from this structure that the HLA molecule and the antigen bind to the T-cell receptor as one entity. The antigen is presented in the form of a short peptide that binds in the crevice and exposes some of its side chains to the receptor. The dimensions of the crevice, 25 Å long, 10 Å wide, and 11 Å deep, allow the binding of peptides of 8 residues if they are in an extended form and of about 20 residues if they are folded into an α helix.

The constraints of this binding pocket are wholly consistent with immunological studies showing that T cells sensitized to cells bearing a specific class I MHC molecule and infected with influenza virus will recognize cells bearing the same MHC incubated with short peptides from influenza virus proteins. It is believed that intracellular pathogens such as viruses betray their presence when fragments of viral proteins escape from the infected cells and become associated with the MHC molecules on the surface. In other cases, extracellular pathogens or foreign proteins may be internalized by cells, degraded in their lysosomes, and returned to the cell surface for presentation to T cells in association with MHC. In either case, the HLA structure reveals the nature of the antigen-MHC complex, and the effect of MHC polymorphism on their function can be more clearly understood.

The polymorphisms are due to a large number of point mutations at specific positions in the amino acid sequences of the $\alpha 1$ and $\alpha 2$ domains. Insight into the structural basis for this polymorphism was provided in 1989 by Don Wiley's group. They determined the structure of a second MHC molecule, HLA-Aw68, which is related to HLA-A2. The principal differences between the two structures are in the size and precise location of 13 amino acids (Figure 12.22) that have been mutated. Five of these differences (orange in Figure 12.22) are located on the central β strands of the "floor." They all point up between the two helical regions and would be buried if the binding site were occupied. Another five mutations have occurred (red in Figure 12.22) on the sides of the helices facing the binding site and are thus also in a position to bind foreign antigens. Although each of these individual amino acid differences in the binding cleft causes only local structural changes, confined to the actual side-chain difference, the 10 substitutions taken together substantially transform the shape and charge of parts of the cleft but leave other parts unchanged (Figure 12.23). Several pockets that extend from the groove beneath the α helices change drastically and presumably cause major changes in specificity for side chains of the peptide antigen.

One amino acid difference between HLA-A2 and HLA-Aw68 is at residue 62 (blue in Figure 12.22), whose side chain is at the top face of the helix in $\alpha 1$ and

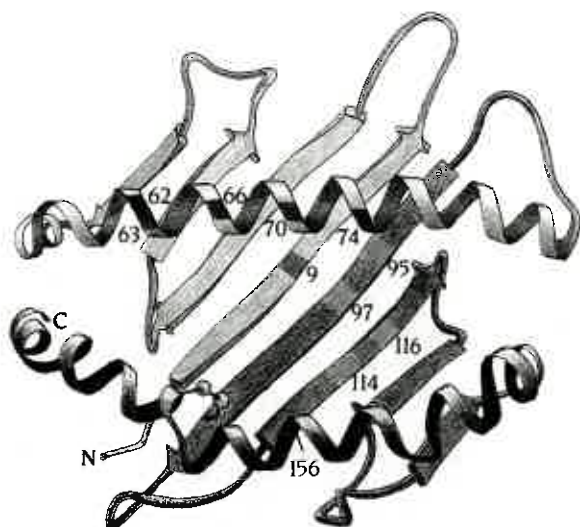


Figure 12.22 Schematic diagram illustrating the location of amino acid differences between two HLA-A molecules that are responsible for their polymorphic difference. Residues that point into the antigenic binding site are orange and red. Residue 62 (blue) may be involved in interactions with the T-cell receptor. (Adapted from P.J. Björkman et al., *Nature* 329: 506, 1987.)

points toward the solvent. This residue is a strong candidate for making direct contacts with the T-cell receptor and may possibly be involved in the discrimination between self and foreign HLA. Thus an immune response involving T cells depends on whether, first, a peptide derived from the pathogen can be successfully bound in the MHC antigen pocket of any of an individual's six class I molecules and, second, whether any of his or her T-cell receptors can recognize the resulting combination of antigen with the remaining exposed polymorphic residues.

We now have the following picture of T-cell activation by HLA. The T-cell receptor, which is composed of four immunoglobulin-like domains, probably has a binding surface similar to the antibody binding surface. This T-cell-receptor surface recognizes a large surface of the HLA molecule. The HLA surface is composed of a pattern of side chains from the top of the two helical regions in HLA as well as side chains from the bound peptide antigen between the helices. The pattern of side chains on this surface that is given by self-HLA and foreign antigens is recognized by the T-cell receptor, which binds the complex with consequent activation of T cells. T cells bearing receptors that bind self-HLA and self-antigen never mature but are destroyed inside the thymus.

Conclusion

IgG antibody molecules are composed of two light chains and two heavy chains joined together to form a T- or Y-shaped molecule. Each light chain has one variable domain and one constant domain, while each heavy chain has one variable and three constant domains. All domains have a similar three-dimensional structure known as the immunoglobulin fold. The stem of the molecule is formed by association between two constant domains from each of the heavy chains. Each arm is formed by association between a light chain and the amino-terminal half of a heavy chain. The hinge region between the stem and the arms is flexible and provides possibilities for different relative orientations of the arms and stem.

The constant domain has a stable framework structure composed of two antiparallel β sheets comprising seven β strands, four in one sheet and three in the other. The variable domains have a similar framework structure but comprising nine β strands, five in one sheet and four in the other. The three hypervariable regions are in loops at one end of the five-strand β sheet. The variable domains associate at the end of the arms through their five-strand β sheets in such a way that the hypervariable loop regions from both domains are close together at the top of a barrel of β strands from both the light and the heavy chains.

This arrangement of the hypervariable loop regions can form antigen binding sites of different shapes depending on the size and sequence of the loops. Haptens induce antibodies with binding cavities, and protein antigens induce antibodies with extended binding surfaces, which contain grooves or pockets for protruding antigen residues.

The lysozyme-antilysozyme interaction area is a large irregular complementary surface where antibody residues from all six hypervariable regions make close contacts with lysozyme residues from two different regions of the polypeptide chain. No gross conformational changes occur in antibody or antigen when they bind together. The third hypervariable region of the heavy chain is in a central position of the interaction area and makes a large number of contacts with the antigen. This region corresponds to the D gene segment, which, through junctional diversity and addition of nucleotides during the recombination events, has the largest sequence variation of the hypervariable regions.

HLA molecules of class I, which bind foreign antigens and activate T-cell receptors, are composed of one light chain, β_2m , and one transmembrane heavy chain. The extracellular portion of the heavy chain is divided into three domains $\alpha 1$, $\alpha 2$, and $\alpha 3$. The light chain and $\alpha 3$, both of which have the immunoglobulin fold, associate to form one region of the HLA molecule close to

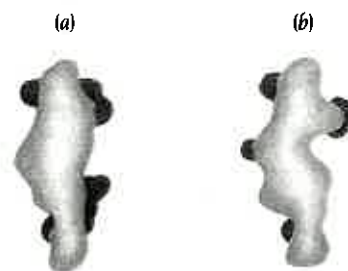


Figure 12.23 Comparison of the antigen binding site in HLA-A2 (a) and HLA-Aw68 (b). The blue regions illustrate the shapes of the cleft between the two α helices, and the red regions are pockets beneath the helices that might bind side chains from the antigen peptides. (Adapted from T.P.J. Garret et al., *Nature* 342: 692, 1989.)

the cell surface. Domains $\alpha 2$ and $\alpha 3$, which have the same unique fold, associate to form a second region, which contains the antigen binding site. This site has a floor of antiparallel β strands with two α helices on top of the floor separated by 18 Å. Antigens in the form of peptides 8–20 residues long bind in the crevice between these helices. Residues from the β strands on the floor and from the two helices on the sides contribute to antigen binding. HLA and antigen thus present to the T-cell receptors a large binding surface composed of side chains from the peptide antigen flanked by side chains from the two helices of HLA.

Selected readings

General

- Alzari, P.N.; Lascombe, M.-B.; Poljak, R.J. Three-dimensional structure of antibodies. *Annu. Rev. Immunol.* 6: 555–580, 1988.
- Amzel, L.M.; Poljak, R.J. Three-dimensional structure of immunoglobulins. *Annu. Rev. Biochem.* 48: 961–997, 1976.
- Colman, P.M. Structure of antibody-antigen complexes: implications for immune recognition. *Adv. Immunol.* 43: 99–132, 1988.
- Davies, D.R.; Metzger, H. Structural basis of antibody function. *Annu. Rev. Immunol.* 1: 87–117, 1983.
- Davies, D.R.; Padlan, E.A. Antibody-antigen complexes. *Annu. Rev. Biochem.* 59: 439–473, 1990.
- Davies, D.R.; Sheriff, S.; Padlan, E.A. Antibody-antigen complexes. *J. Biol. Chem.* 263: 10541–10544, 1988.
- Flavell, R.A., et al. Molecular biology of the H-2 histocompatibility complex. *Science* 233: 437–443, 1986.
- Kennedy, R.C.; Melnick, J.L.; Dreesman, G.R. Anti-idiotypes and immunity. *Sci. Am.* 255(1): 40–63, 1986.
- Leder, P. The genetics of antibody diversity. *Sci. Am.* 247(5): 72–83, 1982.
- Lerner, R.A. Synthetic vaccines. *Sci. Am.* 248(2): 48–62, 1983.
- Lerner, R.A.; Tramontano, A. Catalytic antibodies. *Sci. Am.* 258(3): 42–53, 1988.
- Mariuzza, R.A.; Phillips, S.E.V.; Poljak, R.J. The structural basis of antigen-antibody recognition. *Annu. Rev. Biophys. Biophys. Chem.* 16: 139–159, 1987.
- Marquart, M.; Deisenhofer, J. Three-dimensional structure of antibodies. *Immunol. Today* 3: 160–166, 1982.
- Marrack, P.; Kappler, J. The T cell and its receptor. *Sci. Am.* 254(2): 28–37, 1986.
- Rees, A.R.; de la Paz, P. Investigating antibody specificity using computer graphics and protein engineering. *Trends Biochem. Sci.* 11: 144–148, 1986.
- Riechmann, L., et al. Reshaping human antibodies for therapy. *Nature* 332: 323–327, 1988.
- Shokat, K.M.; Schultz, P.G. Catalytic antibodies. *Annu. Rev. Immunol.* 8: 335–364, 1990.
- Tonegawa, S. Somatic generation of antibody diversity. *Nature* 302: 575–581, 1983.

Specific structures

- Amit, A.G., et al. Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. *Science* 233: 747–753, 1986.
- Bentley, G.A., et al. Three-dimensional structure of an idiotope-anti-idiotope complex. *Nature* 348: 254–257, 1990.
- Bjorkman, P.J., et al. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 329: 506–512, 1987.
- Bjorkman, P.J., et al. The foreign antigen binding site and T-cell recognition regions of class I histocompatibility antigens. *Nature* 329: 512–518, 1987.
- Brown, J.H., et al. A hypothetical model of the foreign antigen binding site of Class II histocompatibility molecules. *Nature* 332: 845–850, 1988.
- Brucoleri, R.E.; Haber, E.; Novotny, J. Structure of antibody hypervariable loops reproduced by a conformational search algorithm. *Nature* 335: 564–568, 1988.
- Chothia, C.; Boswell, D.R.; Lesk, A.M. The outline structure of the T-cell $\alpha\beta$ receptor. *EMBO J.* 7: 3745–3755, 1988.
- Chothia, C., et al. Domain association in immunoglobulin molecules. The packing of variable domains. *J. Mol. Biol.* 186: 651–663, 1985.
- Chothia, C., et al. Conformations of immunoglobulin hypervariable regions. *Nature* 343: 877–883, 1989.
- Colman, P.M., et al. Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. *Nature* 326: 358–363, 1987.
- Cygler, M., et al. Crystallization and structure determination of an autoimmune antipoly (dT) immunoglobulin Fab fragment at 3.0 Å resolution. *J. Biol. Chem.* 262: 643–648, 1987.
- Deisenhofer, J. Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9 and 2.8 Å resolution. *Biochemistry* 20: 2361–2369, 1981.
- Ely, K.R.; Herron, J.N.; Edmundson, A.B. Three-dimensional structure of a hybrid light chain dimer: protein engineering of a binding cavity. *Mol. Immunol.* 27: 101–114, 1990.
- Ely, K.R., et al. Three-dimensional structure of a light chain dimer crystallized in water. Conformational

- flexibility of a molecule in two crystal forms. *J. Mol. Biol.* 210: 601-615, 1989.
- Epp, O., et al. Crystal and molecular structure of a dimer composed of the variable portions of the Bence-Jones protein. *Eur. J. Biochem.* 45: 513-520, 1974.
- Furey, W., Jr., et al. Structure of a novel Bence-Jones protein (Rhe) fragment at 1.6 Å resolution. *J. Mol. Biol.* 167: 661-692, 1983.
- Garret, T.P.J., et al. Specificity pockets for the side chains of peptide antigens in HLA-Aw68. *Nature* 342: 692-696, 1989.
- Herron, J.N., et al. Three-dimensional structure of a fluorescein-fab complex crystallized in 2-methyl-2,4-pentandiol. *Prot.: Struct., Funct., Gen.* 5: 271-280, 1989.
- Huber, R., et al. Crystallographic structure studies of an IgG molecule and an Fc fragment. *Nature* 264: 415-420, 1976.
- Janda, K.D.; Benkovic, S.J.; Lerner, R.A. Catalytic antibodies with lipase activity and R or S substrate selectivity. *Science* 244: 437-440, 1989.
- Lascombe, M.-B., et al. Three-dimensional structure of Fab R 19.9, a monoclonal murine antibody specific for the p-azobenzenearsonate group. *Proc. Natl. Acad. Sci. USA* 86: 607-611, 1989.
- Lesk, A.M.; Chothia, C. Elbow motion in the immunoglobulins involves a molecular ball-and-socket joint. *Nature* 335: 188-190, 1988.
- Marquart, M., et al. Crystallographic refinement and atomic models of the intact immunoglobulin molecule Kol and its antigen-binding fragment at 3.0 and 1.9 Å resolution. *J. Mol. Biol.* 141: 369-391, 1980.
- Novotny, J., et al. Molecular anatomy of the antibody binding site. *J. Biol. Chem.* 258: 14433-14437, 1983.
- Padlan, E.A., et al. Structure of an antibody-antigen complex. Crystal structure of the HyHEL-10 Fab-lysozyme complex. *Proc. Natl. Acad. Sci. USA* 86: 5938-5942, 1989.
- Poljak, R.J., et al. Three-dimensional structure of the Fab fragment of a human immunoglobulin at 2.8 Å resolution. *Proc. Natl. Acad. Sci. USA* 70: 3305-3310, 1973.
- Pollack, S.J.; Jacobs, J.W.; Schultz, P.G. Selective chemical catalysis by an antibody. *Science* 234: 1570-1573, 1986.
- Rajan, S.S., et al. Three-dimensional structure of the Mcg IgG1 immunoglobulin. *Mol. Immunol.* 20: 787-799, 1983.
- Satow, Y., et al. Phosphorylcholine-binding immunoglobulin Fab McPC603. An x-ray diffraction study at 2.7 Å. *J. Mol. Biol.* 190: 593-604, 1986.
- Schiffer, M., et al. Structure of a λ-type Bence-Jones protein at 3.5 Å resolution. *Biochemistry* 12: 4620-4631, 1973.
- Segal, D.M., et al. The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin Fab and the nature of the antigen binding site. *Proc. Natl. Acad. Sci. USA* 71: 4298-4302, 1974.
- Sheriff, S., et al. Three-dimensional structure of an antibody-antigen complex. *Proc. Natl. Acad. Sci. USA* 84: 8075-8079, 1987.
- Shokat, K.M., et al. A new strategy for the generation of catalytic antibodies. *Nature* 338: 269-271, 1989.
- Silverton, E.W.; Navia, M.A.; Davies, D.R. Three-dimensional structure of an intact human immunoglobulin. *Proc. Natl. Acad. Sci. USA* 74: 5140-5144, 1977.
- Stanfield, R.L., et al. Crystal structures of an antibody to a peptide and its complex with peptide antigen at 2.8 Å. *Science* 248: 712-719, 1990.
- Suh, S.W., et al. The galactan-binding immunoglobulin Fab J539: an x-ray diffraction study at 2.6 Å resolution. *Prot.: Struct., Funct., Gen.* 1: 74-80, 1986.
- Tramontano, A.; Janda, K.D.; Lerner, R.A. Catalytic antibodies. *Science* 234: 1566-1570, 1986.