7552dc33_921-950 1/12/02 9:28 AM Page 926 mac76 mac76:385_reb:

33.2 THE IMMUNOGLOBULIN FOLD CONSISTS OF A BETA-SANDWICH FRAMEWORK WITH HYPERVARIABLE LOOPS

An IgG molecule consists of a total of 12 immunoglobulin domains. These domains have many sequence features in common and adopt a common structure, the *immunoglobulin fold* (Figure 33.9). Remarkably, this same structural domain is found in many other proteins that play key roles in the immune system.

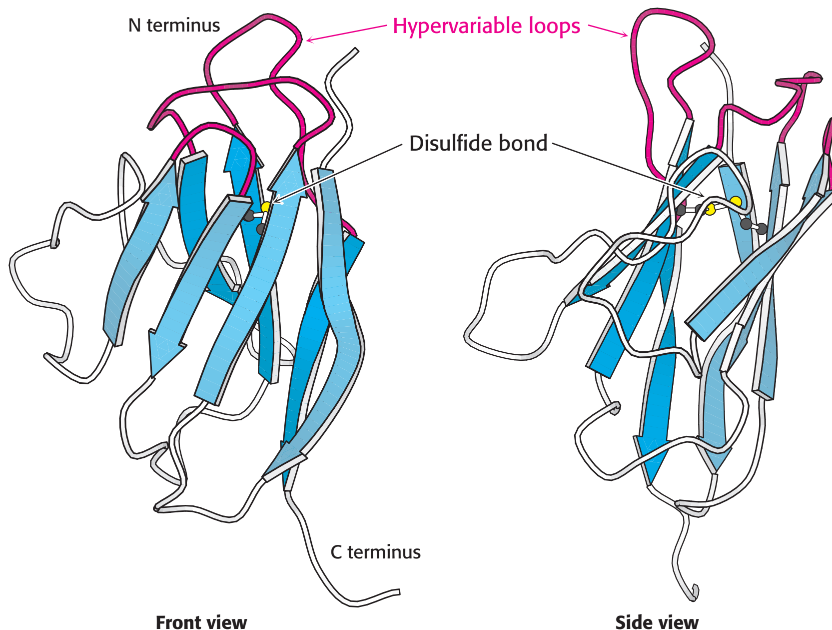

FIGURE 33.9 Immunoglobulin fold. An immunoglobulin domain consists of a pair of β-sheets linked by a disulfide bond and hydrophobic interactions. Three hypervariable loops lie at one end of the structure.

The immunoglobulin fold consists of a pair of β sheets, each built of antiparallel β strands, that surround a central hydrophobic core. A single disulfide bond bridges the two sheets. Two aspects of this structure are particularly important for its function. First, three loops present at one end of the structure form a potential binding surface. These loops contain the hypervariable sequences present in antibodies and in T-cell receptors (see Sections 33.3 and 33.5.2). Variation of the amino acid sequences of these loops provides the major mechanism for the generation of the vastly diverse set of antibodies and T-cell receptors expressed by the immune system. These loops are referred to as *hypervariable loops* or *complementarity-determining regions (CDRs)*. Second, the amino terminus and the carboxyl terminus are at opposite ends of the structure, which allows structural domains to be strung together to form chains, as in the L and H chains of antibodies. Such chains are present in several other key molecules in the immune system.

The immunoglobulin fold is one of the most prevalent domains encoded by the human genome— more than 750 genes encode proteins with at least one immunoglobulin fold recognizable at the level of amino acid sequence. Such domains are also common in other multicellular animals such as flies and nemotodes. However, from inspection of amino acid sequence alone, immunoglobulin-fold domains do not appear to be present in yeast or plants. However, structurally similar domains are present in these organisms, including the key photosynthetic electron-transport protein plastocyanin in plants (Section 19.3.2). Thus, the immunglobulin-fold family appears to have expanded greatly along evolutionary branches leading to animals—particularly, vertebrates.

33.3 ANTIBODIES BIND SPECIFIC MOLECULES THROUGH THEIR HYPERVARIABLE LOOPS

For each class of antibody, the amino-terminal immunoglobin domains of the L and H chains (the variable domains, designated V_L and V_H) come together at the ends of the arms extending from the structure. The positions of the complementarity-determining regions are striking. These hypervariable sequences, present in three loops of each domain, come together so that all six loops form a single surface at the end of each arm (Figure 33.10). Because virtually any V_L can pair with any V_H, *a very large number of different binding sites can be constructed by their combinatorial association*.

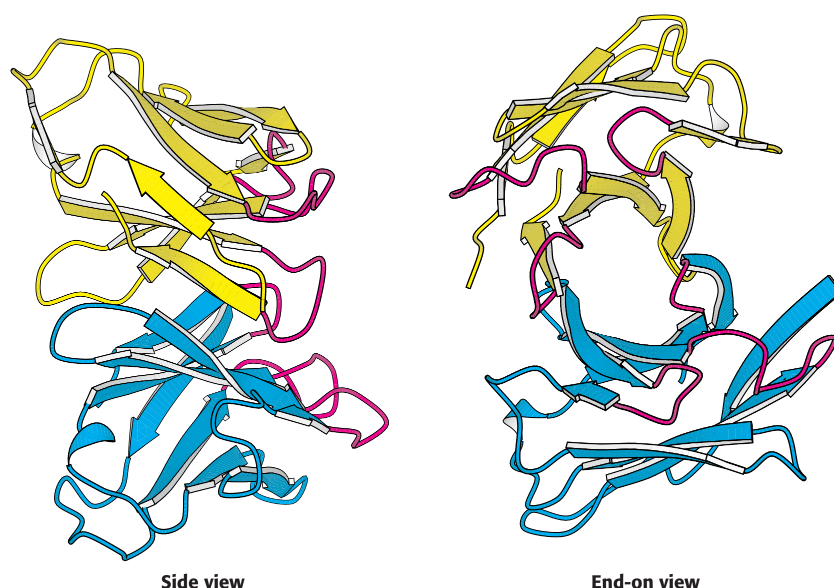

FIGURE 33.10 Variable domains. Two views of the variable domains of the L chain (yellow) and the H chain (blue); the complementarity-determining regions (CDRs) are shown in red. The six CDRs come together to form a binding surface. The specificity of the surface is determined by the sequences and structures of the CDRs.

33.3.1 X-Ray Analyses Have Revealed How Antibodies Bind Antigens

The results of x-ray crystallographic studies of many large and small antigens bound to F_{ab} molecules have been sources of much insight into the structural basis of antibody specificity. The binding of antigens to antibodies is governed by the same principles that govern the binding of substrates to enzymes. The apposition of complementary shapes results in numerous contacts between amino acids at the binding surfaces of both molecules. Numerous hydrogen bonds, electrostatic interactions, and van der Waals interactions, reinforced by hydrophobic interactions, combine to give specific and strong binding.

A few aspects of antibody binding merit specific attention, inasmuch as they relate directly to the structure of immunoglobulins. The binding site on the antibody has been found to incorporate some or all of the CDRs in the variable domains of the antibody. Small molecules (e.g., octapeptides) are likely to make contact with fewer CDRs, with perhaps 15 residues of the antibody participating in the binding interaction. Macromolecules often make more extensive contact, interacting with all six CDRs and 20 or more residues of the antibody. Small molecules often bind in a cleft of the antigen-binding region. Macromolecules such as globular proteins tend to interact across larger, fairly flat apposed surfaces bearing complementary protrusions and depressions.

A well-studied case of small-molecule binding is seen in an example of phosphorylcholine bound to F_{ab}. Crystallographic analysis revealed

7552dc33_921-950 1/12/02 9:28 AM Page 928 mac76 mac76:385_reb:

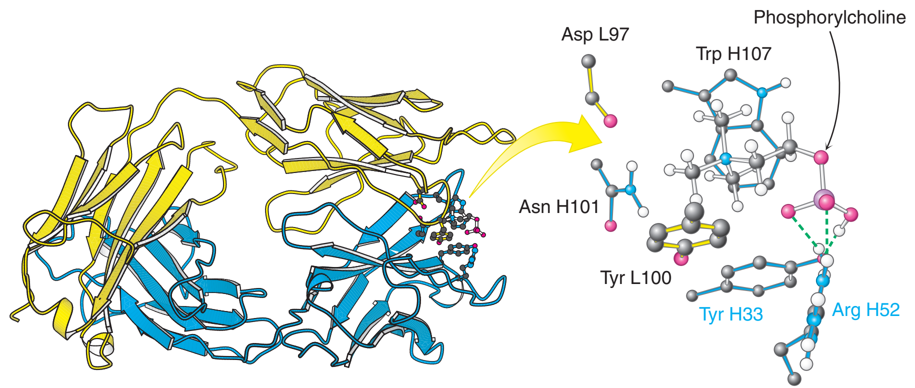

FIGURE 33.11 Binding of a small antigen. The structure of a complex between an F_{ab} fragment of an antibody and its target—in this case, phosphorylcholine. Residues from the antibody interact with phosphorylcholine through hydrogen bonding and electrostatic and van der Waals interactions.

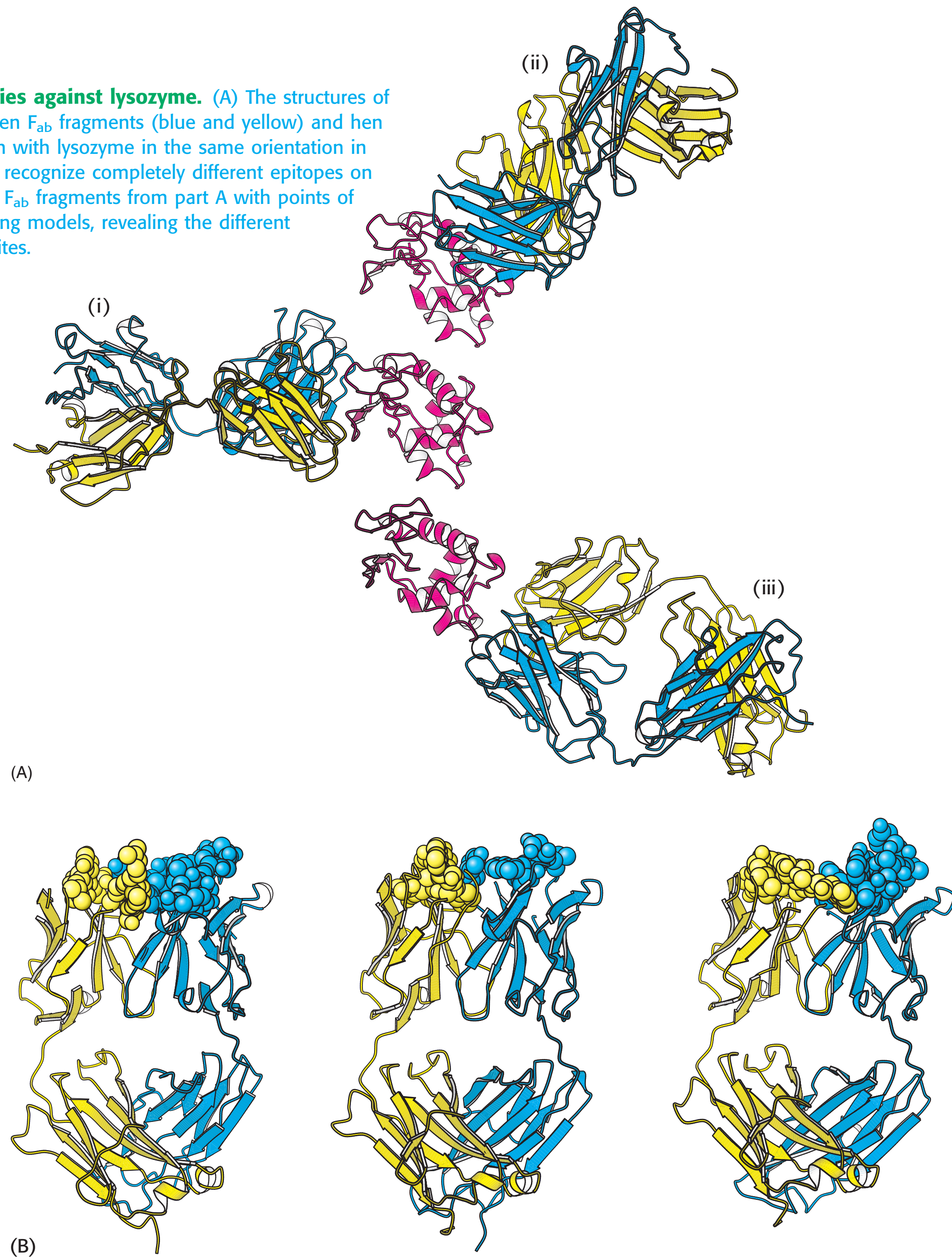

FIGURE 33.12 Antibodies against lysozyme. (A) The structures of three complexes (i, ii, iii) between F_{ab} fragments (blue and yellow) and hen egg-white lysozyme (red) shown with lysozyme in the same orientation in each case. The three antibodies recognize completely different epitopes on the lysozyme molecule. (B) The F_{ab} fragments from part A with points of contact highlighted as space-filling models, revealing the different shapes of the antigen-binding sites.

7552dc33_921-950 1/12/02 9:28 AM Page 929 mac76 mac76:385_reb:

phosphorylcholine bound to a cavity lined by residues from five CDRs—two from the L chain and three from the H chain (Figure 33.11). The positively charged trimethylammonium group of phosphorylcholine is buried inside the wedge-shaped cavity, where it interacts electrostatically with two negatively charged glutamate residues. The negatively charged phosphate group of phosphorylcholine binds to the positively charged guanidinium group of an arginine residue at the mouth of the crevice and to a nearby lysine residue. The phosphate group is also hydrogen bonded to the hydroxyl group of a tyrosine residue and to the guanidinium group of the arginine side chain. Numerous van der Waals interactions, such as those made by a tryptophan side chain, also stabilize this complex.

The binding of phosphorylcholine does not significantly change the structure of the antibody, yet induced fit plays a role in the formation of many antibody–antigen complexes. A malleable binding site can accommodate many more kinds of ligands than can a rigid one. Thus, induced fit increases the repertoire of antibody specificities.

33.3.2 Large Antigens Bind Antibodies with Numerous Interactions

How do large antigens interact with antibodies? A large collection of antibodies raised against hen egg-white lysozyme has been structurally characterized in great detail (Figure 33.12). Each different antibody binds to a distinct surface of lysozyme. Let us examine the interactions present in one of these complexes in detail. This antibody binds two polypeptide segments that are widely separated in the primary structure, residues 18 to 27 and 116 to 129 (Figure 33.13).

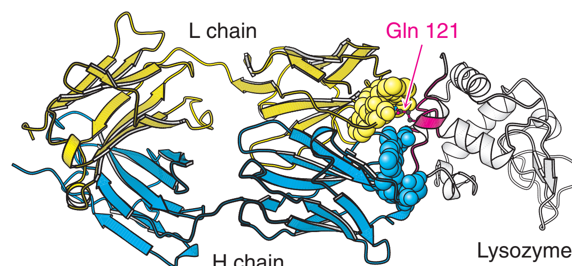

FIGURE 33.13 Antibody–protein interactions. The structure of a complex between an F_{ab} fragment and lysozyme reveals that the binding surfaces are complementary in shape over a large area. A single residue of lysozyme, glutamine 121, penetrates more deeply into the antibody combining site.

All six CDRs of the antibody make contact with this epitope. The region of contact is quite extensive (about 30 × 20 Å). The apposed surfaces are rather flat. The only exception is the side chain of glutamine 121 of lysozyme, which penetrates deeply into the antibody binding site, where it forms a hydrogen bond with a main-chain carbonyl oxygen atom and is surrounded by three aromatic side chains. The formation of 12 hydrogen bonds and numerous van der Waals interactions contributes to the high affinity (K_d = 20 nM) of this antibody–antigen interaction. Examination of the F_{ab} molecule without bound protein reveals that the structures of the V_L and V_H domains change little on binding, although they slide 1 Å apart to allow more intimate contact with lysozyme.

33.4 DIVERSITY IS GENERATED BY GENE REARRANGEMENTS

A mammal such as a mouse or a human being can synthesize large amounts of specific antibody against virtually any foreign determinant within a matter of days of being exposed to it. We have seen that antibody specificity is determined by the amino acid sequences of the variable regions of both light

7552dc33_921-950 1/12/02 9:28 AM Page 930 mac76 mac76:385_reb:

and heavy chains, which brings us to the key question: How are different variable-region sequences generated?

The discovery of distinct variable and constant regions in the L and H chains raised the possibility that the genes that encode immunoglobulins have an unusual architecture that facilitates the generation of a diverse set of polypeptide products. In 1965, William Dreyer and Claude Bennett proposed that multiple *V (variable) genes* are separate from a single *C (constant) gene* in embryonic (germ-line) DNA. According to their model, one of these V genes becomes joined to the C gene in the course of differentiation of the antibody-producing cell. A critical test of this novel hypothesis had to await the isolation of pure immunoglobulin mRNA and the development of techniques for analyzing mammalian genomes. Twenty years later, Susumu Tonegawa found that V and C genes are indeed far apart in embryonic DNA but are closely associated in the DNA of antibody-producing cells. Thus, immunoglobulin genes are rearranged in the differentiation of lymphocytes.

33.4.1 J (Joining) Genes and D (Diversity) Genes Increase Antibody Diversity

Sequencing studies carried out by Susumu Tonegawa, Philip Leder, and Leroy Hood revealed that V genes in embryonic cells do not encode the entire variable region of L and H chains. Consider, for example, the region that encodes the κ light-chain family. A tandem array of 40 segments, each of which encodes approximately the first 97 residues of the variable domain of the L chain, is present on human chromosome 2 (Figure 33.14).

FIGURE 33.14 The κ light-chain locus. This part of human chromosome 2 includes an array of 40 segments that encode the variable (V) region (approximately residues 1–97) of the light chain, an array of 5 segments that encode the joining (J) region (residues 98–110), and a single region that encodes the constant (C) region.

However, the variable region of the L chain extends to residue 110. Where is the DNA that encodes the last 13 residues of the V region? For L chains in undifferentiated cells, this stretch of DNA is located in an unexpected place: near the C gene. It is called the *J gene* because it *j*oins the V and C genes in a differentiated cell. In fact, a tandem array of five J genes is located near the C gene in embryonic cells. In the differentiation of an antibody-producing cell, a V gene becomes spliced to a J gene to form a complete gene for the variable region (Figure 33.15). RNA splicing generates

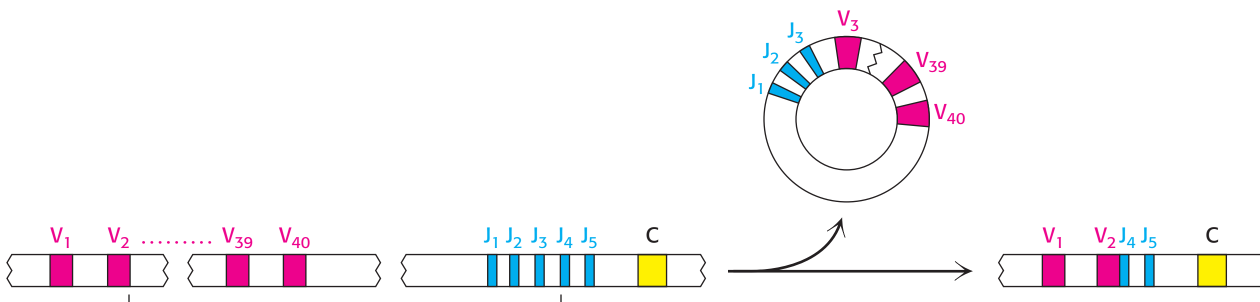

FIGURE 33.15 VJ recombination. A single V gene (in this case, V_2) is linked to a J gene (here, J_4) to form an intact VJ region. The intervening DNA is released in a circular form. Because the V and J regions are selected at random and the joint between them is not always in exactly the same place, many VJ combinations can be generated by this process.