

CONTENTS

Acknowledgments ix

1	Prologue: The Goals of This Book	1
2	Introduction to Basic Immunology	9
3	Development of the T Cell Repertoire	53
4	Characterization of the Fitness Landscape of Highly Mutable Viruses	87
5	Affinity Maturation of the Antibody Response	143
6	Intracellular Signaling in Lymphocytes	200
7	Epilogue: The Incompleteness of This Book	218

Index 225

2

Introduction to Basic Immunology

2.1. Introduction

The goal of this chapter is to introduce readers to the main components of the immune system and how they are thought to function, to prepare readers for the more detailed mechanistic exploration of some topics in the following chapters. More comprehensive descriptions of immunobiology can be found in standard immunology textbooks, such as the two listed in the suggested readings at the end of this chapter. We begin with a very brief elaboration of the history of the development of modern immunology mentioned in chapter 1.

2.2. A Very Brief History of Immunology

The history of the development of immunology until the eighteenth century can be traced by following the story of how humans battled smallpox. Smallpox probably started circulating in humans about 12,000 years ago and killed millions of people over many centuries. Our ancestors observed that people who recovered from smallpox could become caregivers to those afflicted with the disease without becoming sick again. This led to the practice of a form of vaccination against smallpox in India and China (since probably as early as 1500 AD). In China, material from smallpox sores was dried and powdered, and after a waiting period, the powder was inhaled nasally through a silver tube. In India, a similar procedure involved storing material from smallpox sores for a while and then mixing it with rice and rubbing it into punctures made in the skin of uninfected people. This procedure was called *variolation*, after the name of the causative virus, *variola*. The technique was brought to the West by an Englishwoman, Mary Wortley Montagu, and for a while it

was practiced by some, as it could protect against smallpox. However, variolation caused unpleasant smallpox-like symptoms, and it was a dangerous procedure that could result in the full-blown disease and even cause small epidemics. Thus, it was not widely adopted.

Edward Jenner developed the first effective and safe procedure to protect humans from smallpox infection. Jenner heard that dairy maids who were exposed to cowpox (which is usually harmless to humans) did not contract smallpox or respond to variolation. After researching this subject for some years, in May 1796, Jenner inoculated a boy named James Phipps with pus from a cowpox-infected milkmaid named Sarah Nelmes. Two months later, Jenner showed that Phipps was protected from smallpox, as he did not acquire any symptoms upon being variolated. Subsequently, Jenner carried out the procedure on others and showed that they also were protected from the disease, and his technique became widely practiced. Coordinated international vaccination drives ultimately eradicated smallpox from the planet, thus ending tremendous suffering. Jenner is thus rightly regarded as a heroic historic figure. More broadly, vaccination has protected humanity from diverse diseases and saved millions of lives.

When Jenner did his work, nothing was known about the immune system or that microorganisms cause disease. Although many people contributed to the science, Louis Pasteur and Robert Koch led efforts showing that infectious diseases were caused by microorganisms.

Observations such as mold growing on bread led to the popular idea that microorganisms were spontaneously generated. Although others had previously provided evidence against this notion, Pasteur conducted a clever experiment that clearly demonstrated that microorganisms were not spontaneously generated. He had previously described that heating liquid food products to high temperatures and sealing the container prevented spoiling. This is the basis for the process of pasteurization of beer and milk. Pasteur stored boiled water in two swan-necked flasks, one of which was tilted as shown in Fig. 2.1. Microorganisms grew in the water only in the tilted flask. In the upright flask, airborne germs that entered the neck were trapped on the glass surfaces and so could not access and colonize the water. This demonstration was the death knell of the spontaneous generation theory.

Koch worked on diverse bacteria, such as those that cause cholera, anthrax, and tuberculosis. He developed procedures to isolate and grow bacteria, which allowed them to be studied in detail. Based on his studies, Koch established

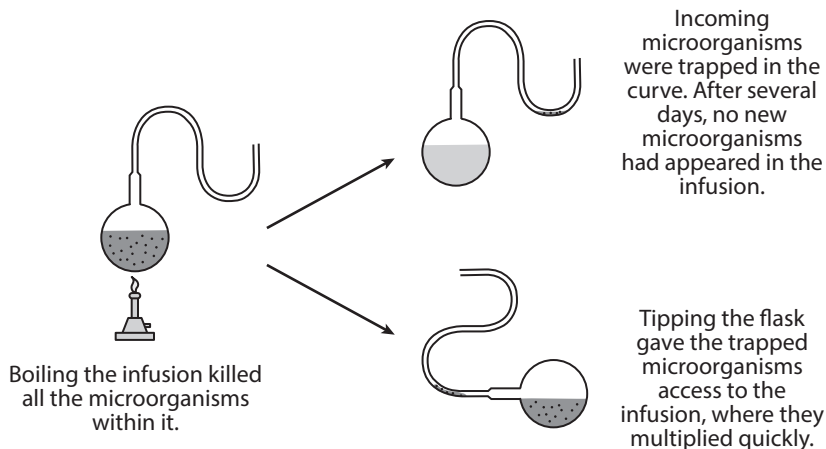


FIGURE 2.1. Schematic depiction of Pasteur's experiment disproving the idea that microorganisms are spontaneously generated.

the following postulates for determining that a particular pathogen is the causative agent of a disease:

1. The microorganism must be present in every instance of the disease.
2. The microorganism must be isolated from a human with the disease and grown in culture. Culturing bacteria involves providing nutrients that enable bacteria to multiply.
3. The microorganism grown in culture must cause the same disease upon injection in animal models.
4. Samples from the animal in which disease thus occurs must contain the same organism that was present in the original human with the disease.

Koch's postulates are still used to establish the causative agents of bacterial infections.

Koch and Pasteur were vigorous competitors, and both developed vaccines against diseases that mostly affected animals. A serendipitous observation led Pasteur to develop a new paradigm for vaccination that was different from Jenner's. In experimenting with chicken cholera, Pasteur inoculated chickens with the causative bacteria. When the chickens did not get sick, he found that the culture in which the bacteria were grown had spoiled, and so the bacteria were damaged. After culturing the bacteria anew, Pasteur inoculated the same chickens again, presumably to save resources. They still did not get sick! This

is how he discovered that inoculation with a killed or attenuated (thus harmless) form of the same pathogen could safely confer protection. This was a paradigm shift compared to Jenner's strategy, which required finding harmless pathogens related to the dangerous pathogen that caused severe human disease. Since Pasteur's discovery, many vaccines have been developed that are composed of weakened, killed, or altered variants of a pathogen that causes human disease. The annual influenza vaccine and polio vaccines are examples.

The success of vaccination strategies devised by Jenner, Pasteur, Koch, and others motivated scientists to seek the mechanisms underlying how vaccination works, which led to the development of the modern science of immunology. A major advance was made by Metchnikoff, who, in 1882, showed that certain cells could eat and digest (phagocytose) invading species as part of the immune response. In 1890 von Behring and Kitasato had a very significant insight into the mechanism by which vaccines might work. They showed that blood (or serum) from vaccinated or infected people contained a substance that specifically bound to the pathogen. They termed this substance *antibodies*. Antibodies were used as a therapy for diphtheria, which used to kill many children.

Since then, major advances have been made in understanding the fundamental aspects of immunology. It is now clear that antibodies and various other cell types (such as the ones discovered by Metchnikoff) play important roles in mediating an immune response. The following is an incomplete list of scientists who made major contributions to these advances: Landsteiner, Ehrlich, Pauling, Burnett, Tonegawa, Baltimore, Doherty, Zinkernagel, Unanue, Eisen, Benacerraf, Davis, Mak, Bjorkman, Wylie, Weiss, Steinman, Janeway, Hoffmann, Medzhitov, Boettler, Miller, and Nussenzweig. We will learn about some of their discoveries in later parts of this chapter and book.

2.3. Classes of Microbial Pathogens

Prior to learning how the immune system combats pathogens, let us briefly consider the different types of pathogens that infect humans and cause disease. Table 2.1 lists common pathogens and the routes they take to invade human bodies. Pathogens are typically classified into four categories: viruses, bacteria, fungi, and parasites. Examples of viruses include influenza and the human immune deficiency virus (HIV). The World Health Organization (WHO) estimates that 290,000 to 650,000 people die each year from influenza

TABLE 2.1. Examples of pathogens, the diseases they cause, and their portals of entry

Route of entry	Pathogen	Disease
Mucosal surfaces		
Airway	Influenza virus	Influenza
	<i>Neisseria meningitidis</i>	Meningococcal meningitis
Gastrointestinal tract	<i>Salmonella typhi</i>	Typhoid fever
	Rotavirus	Diarrhea
Reproductive tract	<i>Treponema pallidum</i>	Syphilis
External epithelia		
External surface	<i>Tinea pedis</i>	Athlete's foot
Wounds and abrasions	<i>Bacillus anthracis</i>	Anthrax
	<i>Clostridium tetani</i>	Tetanus
	<i>Pasteurella tularensis</i>	Tularemia
Insect bites	Flavivirus	Yellow fever
	<i>Borrelia burgdorferi</i>	Lyme disease
	<i>Plasmodium</i> spp.	Malaria

Source: Adapted from C. Janeway, P. Travers, M. Walport, and M. Shlomchik, *Immunobiology*, 5th edition, Garland (2001).

infections. HIV, the causative agent of acquired immune deficiency syndrome (AIDS), is estimated by the WHO to have killed well over 40 million people. While the immune system usually combats influenza successfully and clears the virus, HIV is not documented to have been definitively cleared by natural immune responses by anyone, although the immune system of rare individuals can control the virus to persist only at very low levels. There is no vaccine or cure for HIV, but new drugs are highly effective at prophylaxis, and lifelong use of antiretroviral drugs can control the virus to undetectable levels.

A bacterium that poses a major global health challenge is *Mycobacterium tuberculosis*, the causative agent of tuberculosis. The WHO estimates that tuberculosis killed 1.25 million people in 2023 alone. Antibiotics are administered to treat bacterial infections like tuberculosis, but strains resistant to such drugs have evolved and pose a major challenge. Currently, there is no effective vaccine against tuberculosis that provides long-term protection.

The immune system usually combats fungi (like *Candida*) very effectively, except when the immune system is compromised (e.g., in persons infected with HIV). Parasites live at the expense of the host they invade. Parasitic infections are common in tropical countries and are major causes of human suffering in, for

example, parts of Africa and Asia. A prominent example of a disease-causing parasite is *Plasmodium*, which is the causative agent of malaria. The WHO estimates that in each of the years 2022 and 2023 roughly 600,000 people died of malaria.

2.4. The Immune System of Higher Organisms Comprises Diverse Organs and Cells

The immune system of organisms that evolved after jawed vertebrate fish can roughly be divided into two inextricably linked parts called the innate and adaptive immune systems. The innate immune system comprises various types of cells and molecules that serve as first responders when a pathogen invades the host. These first responders perform two critical functions. Many types of innate immune cells can engulf pathogens and secrete substances that create an inhospitable environment for the pathogen. These actions of the innate immune system aim to control the pathogen, but they also make the host feel unwell (with fever, loss of appetite, chills, and the like). Despite the effects of innate immunity on pathogens, infection can be established.

The second key function of the innate immune system is to alert the cells of the adaptive immune system to the presence of the pathogen and its type, and to mediate the activation of adaptive immunity. The innate immune response subsequently begins to wane. The host starts feeling well in 7–10 days when the adaptive immune system mounts a bespoke response tailored to the specific infecting pathogen. A memory of this specific pathogen is imprinted in the adaptive immune system when the infection is cleared. This pathogen-specific immunological memory enables robust and rapid responses upon reinfection with the same pathogen, and this is the basis of vaccination.

In contrast, the innate immune response is not specific to a particular pathogen and, to our knowledge, does not exhibit immune memory of specific pathogens. However, recent evidence suggests that some cells of the innate immune system can maintain a hypersensitive state for some time after exposure to certain components common to many bacteria. This phenomenon can offer some protection against different pathogens in the same class for some time.

Figure 2.2 depicts the cell types that make up the innate and adaptive immune systems, all of which originate in the bone marrow. Hematopoietic stem cells differentiate into platelets, red blood cells, and the precursors of immune cells. Cells of the innate and adaptive immune systems originate from the common myeloid progenitor and the common lymphoid progenitor, respectively.

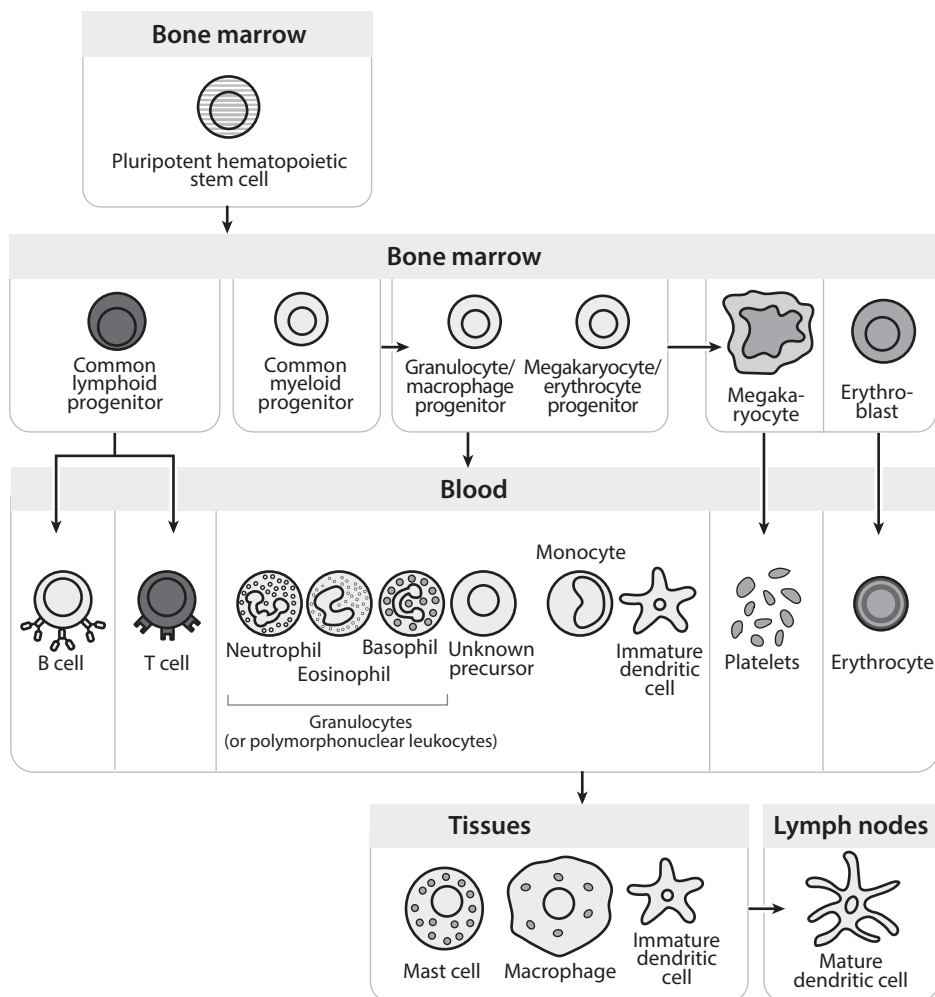


FIGURE 2.2. The various cell types of the immune system differentiate from pluripotent hematopoietic stem cells that originate in bone marrow. The common lymphoid progenitor also differentiates into NK cells (not shown). Redrawn from C.A. Janeway, P. Travers, M. Walport, and M. Shlomchik, *Immunobiology*, 5th edition, Garland Publishing (2001).

The common myeloid progenitor differentiates into the granulocyte/macrophage progenitor, from which the innate immune cells are derived (granulocytes, monocytes, dendritic cells). Granulocytes are so called because granules are observed inside them. Monocytes differentiate in tissues into cells called *macrophages*. *Mast cells* are also a part of the innate immune system, and

they are derived from an unknown progenitor. *Erythrocytes* are red blood cells, and *platelets* play an important role in blood clotting.

The common lymphoid progenitor differentiates into T lymphocytes (T cells) and B lymphocytes (B cells), and these are the key players that mount pathogen-specific adaptive immune responses. The common lymphoid progenitor also differentiates into cells called *natural killer* (NK) cells, which are considered to be a part of the innate immune system that play an important role in immune defense. We will not consider NK cells further.

In this book, we will largely focus on the adaptive immune system and how it responds to and overcomes most pathogen infections. The goal is to illustrate how the convergence of ideas from physics, biology, and medicine can lead to advances in immunology that can potentially be harnessed to prevent and cure diseases. However, the adaptive immune system's function is dependent on the innate immune system, and so we begin with a short description of the functions of its cells.

2.5. Overview of the Innate Immune System

The innate immune system can detect the presence of pathogens, as well as the type of infection (e.g., bacterial, viral, parasitic), and then signal the adaptive immune system to mount appropriate responses. Cells of the innate immune system express several types of receptors on their outer membrane and in intracellular endosomal compartments that recognize specific types of pathogen components. For example, some receptors can bind to specific types of sugars and other constituents typically expressed on the surface of bacteria and not in higher organisms. Others can detect intracellular viral DNA and RNA. The innate immune system also comprises molecules that can detect chemicals produced by a functioning intracellular pathogen. The sensing of conserved elements of pathogens, together with signatures of their functions, informs the innate immune system about the nature of invading pathogens.

Many types of cells and molecules mediate the functions of innate immunity. For example, eosinophils and basophils play an important role in controlling parasites. Mast cells reside near small blood vessels and are implicated in allergic reactions. We will focus on cells called phagocytes because they are very important for both innate and adaptive immunity. *Phagocytic cells* were discovered by Metchnikoff. They engulf and “digest” microorganisms in intracellular vesicles, and this process is called phagocytosis (derived from *phagein*,

the ancient Greek word for “to eat”). Neutrophils, macrophages, and dendritic cells (DCs) are important examples of phagocytic cells.

Neutrophils express receptors on their surface that can bind to common substances on pathogen surfaces (such as carbohydrates and lipids). They can then engulf the pathogen and produce hydrogen peroxide, superoxide anion, and other substances toxic to bacteria. Neutrophils die shortly after phagocytosis, and dead neutrophils are a major component of pus. Neutrophils are among the first lines of immune defense. Individuals with neutrophil deficiency suffer from uncontrolled bacterial infections and live for only a short time without therapeutic interventions.

Macrophages and DCs play key roles in activating adaptive immunity. A hierarchical set of blood proteins called *complement* is normally inactive. The first protein in the hierarchy is activated by interactions with a pathogen or with antibodies bound to a pathogen. The activated component of the complement then activates the next protein in the hierarchy, and so on, thus setting up a chain reaction of activation of many complement proteins that are able to bind to pathogens. Phagocytic cells have receptors on their surface that can bind to complement bound to pathogens, which promotes their activation and concomitant pathogen engulfment. Macrophages internalize pathogens in phagosomes (Fig 2.3), and the actions of enzymes therein can break down the proteins of many pathogens. Lysosomes, which are bodies containing antimicrobial agents within macrophages, can fuse with phagosomes, and the antimicrobial agents can damage bacterial pathogens.

Macrophages and DCs express a family of evolutionarily ancient receptors called *toll-like receptors* (TLRs). The importance of TLRs for mediating an immune response was first discovered in fruit flies, and different types of TLRs and other types of receptors that can detect pathogens have now been identified. These receptors “recognize” different types of pathogen components. For example, some TLRs recognize viral DNA or RNA, while others recognize bacterial cell-wall components (e.g., lipopolysaccharide). Thus, receptors like TLRs inform the immune system of the type of infection (viral, bacterial, etc.). DCs are long-lived cells that reside in tissues in immature form. When their TLRs bind to pathogen components, the DCs engulf them, and the DCs’ phenotype changes into a mature form that does not engulf pathogens. Mature DCs migrate from tissues to organs, where they encounter cells of the adaptive immune system and help activate the adaptive immune response. TLRs are also expressed in endosomes, and they detect the presence of nucleic acids.

Macrophages engulf and digest bacteria to which they bind

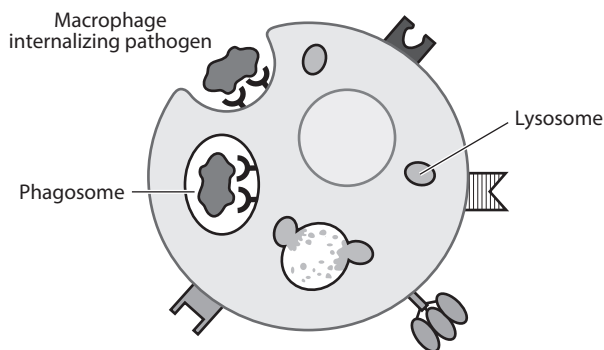


FIGURE 2.3. Macrophages are an important class of phagocytes. Phagosomes are bodies inside macrophages wherein enzymes break down pathogens. Lysosomes are bodies in macrophages that contain antimicrobial agents. Redrawn from C.A. Janeway, P. Travers, M. Walport, and M. Shlomchik, *Immunobiology*, 5th edition, Garland Publishing (2001).

Regulatory mechanisms have evolved that allow endosomal TLRs to differentiate between self and foreign nucleic acids.

TLR binding to pathogen components or their complexes with blood proteins initiates a cascade of biochemical reactions in macrophages and DCs resulting in transcription of new genes that cause them to secrete molecules called cytokines and chemokines. These chemicals are important for many functions, including the type of immune response that results. For example, certain kinds of cytokines, as well as complement components, cause inflammation, which is critical for the immune response for many reasons. Interestingly, this issue was heatedly debated in Metchnikoff's times, when many believed inflammation was part of the pathogenesis of disease-causing microbes, rather than part of the immune response.

Inflammation is characterized by heat, redness, swelling, and pain. The action of cytokines results in the dilation of local blood vessels and increased blood flow, which manifest as heat and redness. Inflammation induces blood vessel walls to express molecules that can bind to cells of the immune system, promoting their adhesion to the blood vessel walls. Inflammation also causes the endothelial cells making up the walls of blood vessels to separate from one another,

thus increasing the permeability of the blood vessel walls. Enhanced adhesion of immune cells to the more permeable walls results in migration of these cells from blood vessels into tissues. The enhanced permeability of blood vessel walls also results in leakage of blood components into tissues, causing swelling. Cytokines also activate certain types of neurons associated with experiencing pain.

Cytokines have many additional functions that regulate immune responses. One of these functions is to act on pathogens (effector functions). For example, the cytokine $\text{TNF}\alpha$ induces localized clotting of blood, thereby isolating the infected area. Other cytokines raise the body temperature (fever), which serves to lower pathogen activity. A family of cytokines called *interferons* (IFNs) has potent antiviral activity—for example, blocking viral replication. Cytokines and chemokines also attract neutrophils and macrophages to the site of an infection. In addition to these types of effector functions, cytokines secreted by innate immune cells also regulate the adaptive immune response. Depending upon the kind of pathogen detected, different kinds of cytokines are secreted, stimulating particular types of adaptive immune responses (discussed later in this chapter).

All immune responses result in a cost to the host, with different types of responses associated with particular costs. For example, the inflammatory response makes one feel sick, and certain types of responses cause host tissue damage. If a pathogen or its products are detected deeper in the host, more severe immune responses with a higher cost to the host arise. For example, a pathogen detected in skin may result in localized inflammation, whereas the same pathogen detected in blood causes a severe systemic response. It is thought that the immune system balances the cost to the host due to the pathogen against the cost due to damage caused by the immune response (immunopathology). Aberrant regulation of these processes can be very dangerous for the host. For example, uncontrolled secretion of cytokines, called *cytokine storms*, are such events often accompanying severe infections and can lead to death.

2.6. The Adaptive Immune System

Pathogens evolve strategies to evade innate immune responses and, despite these responses, may succeed in establishing infection. As noted earlier, the innate immune system also mediates the activation of cells of the adaptive immune system, which then results in a response specific for a particular pathogen (e.g., a particular strain of influenza virus or a specific coronavirus).

An obvious question then arises: How can the adaptive immune system mount pathogen-specific responses against an enormous number of existing and future pathogens? In the 1950s, an Australian immunologist, Macfarlane Burnet, proposed an answer to this question. Since the late nineteenth century, it was known that blood from animals and humans infected or immunized with a specific pathogen contained substances, called *antibodies*, that could bind specifically to the pathogen or the toxins it produced. Burnet proposed that lymphocytes were the source of antibodies, arguing that each lymphocyte displayed a few (maybe even one) antibodies. When an individual is exposed to a pathogen, the lymphocytes displaying antibodies specific for this pathogen (specific clones) are activated, and they proliferate and carry out functions that damage this pathogen and eliminate it.

Burnet's idea, which was not based on hard experimental data, is called the *clonal selection theory*. While reality is more complex, Burnet was in essence correct, and his work had an enormous impact on immunology. Today we know lymphocytes orchestrate pathogen-specific adaptive immune responses, there are two kinds of lymphocytes, and precursors to both types are first produced in the bone marrow. The lymphocytes that then develop and mature in the organ called the thymus (T) are called *T lymphocytes* (T cells), and those that develop in the bone marrow (B) are called *B lymphocytes* (B cells). B cells produce antibodies, and T cells have many functions, which include killing infected cells.

Humans have many billion T cells and a similar number of B cells. T cells and B cells express receptors on their surface called *T cell receptors* (TCRs) and *B cell receptors* (BCRs), respectively. A huge diversity of TCRs and BCRs is generated, and most, but not all, T cells and B cells express a single type of receptor. Thus, with high probability, the TCR (BCR) expressed by a T cell (B cell) is distinct from that expressed by another T cell (B cell). A given TCR or BCR can bind to a specific molecular component of a particular pathogen (the receptor's *ligand*), and TCRs and BCRs bind to very different types of pathogen-derived ligands. *Receptor specificity* means that a TCR or BCR that binds well to a particular ligand will likely not bind well to another ligand that is even somewhat different. Most T cells exhibit distinct specificities, but some can exhibit the same specificity, meaning they can bind to the same ligand. Each T cell can also exhibit some cross-reactivity, in that it can bind to a few ligands. These statements are also true for B cells. Thus, because the repertoire of B cells and T cells in every individual can generate a huge diversity of TCRs and BCRs, each person's immune system has the potential to mount pathogen-specific responses against diverse pathogens.

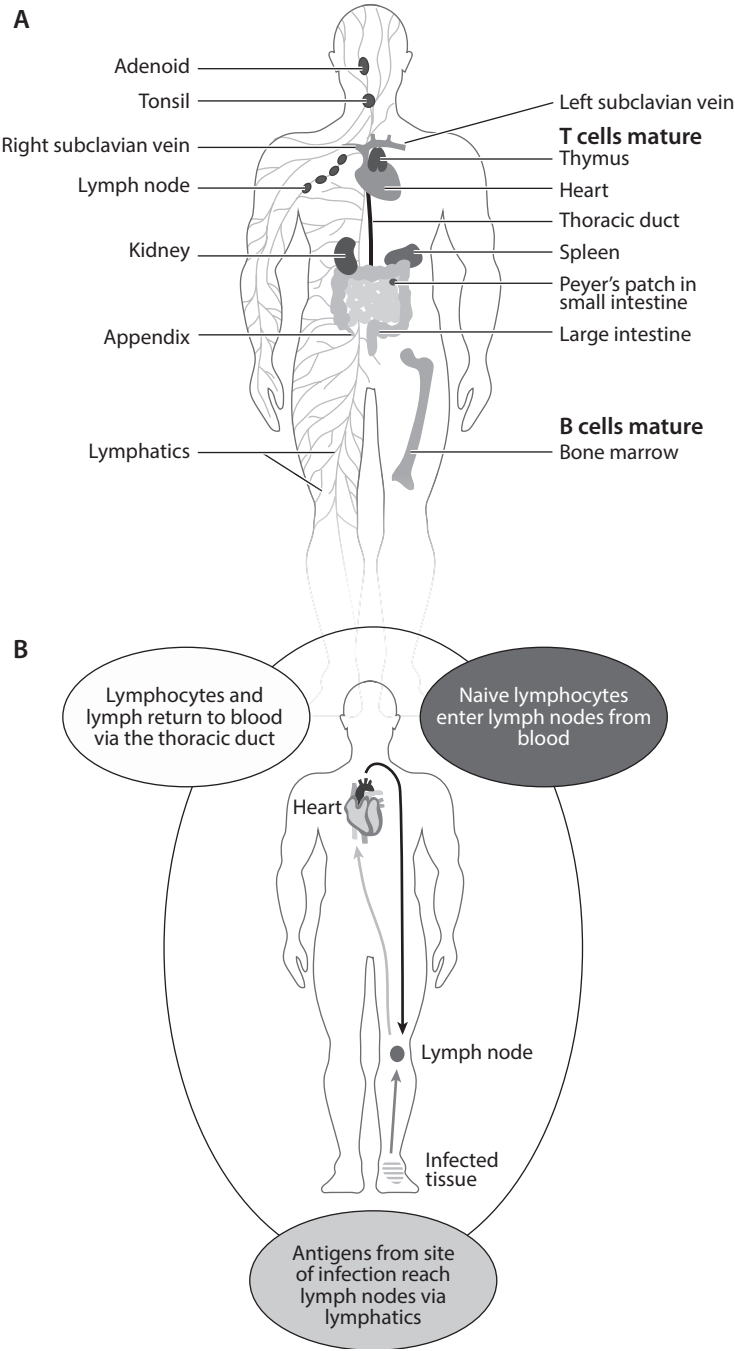
We next discuss how lymphocytes encounter pathogens, how the diversity of receptors is generated, the structures and ligands of TCRs and BCRs, and the functions of T cells and B cells.

2.6.1. T cells and B cells First Encounter Pathogens in Lymphoid Tissues

Humans and other organisms with an adaptive immune system have a number of lymphoid organs, the primary ones of which are the thymus and the bone marrow. As discussed in the previous section, T cells and B cells are produced in the bone marrow. B cells mature in the bone marrow, and T cells mature in the thymus (discussed in detail in chapter 3). After maturation, lymphocytes enter the bloodstream. The other lymphoid organs, called peripheral or secondary lymphoid organs, are distributed throughout the body (Fig. 2.4A). These organs include the spleen, ubiquitously distributed lymph nodes, and other lymphoid tissues. Lymphocytes first encounter pathogens in the peripheral lymphoid organs.

A set of vessels, called the *lymphatic vessels*, connect tissues to peripheral lymphoid organs (Fig. 2.4B). For example, the so-called afferent lymphatic vessels drain extracellular fluid from tissues to the closest lymph node. Upon infection this fluid, called *lymph*, contains pathogens as well as DCs and macrophages that may have detected and phagocytosed pathogens. Lymphocytes circulating in blood enter lymphoid organs by squeezing through blood vessels in lymph nodes. Multiphoton microscopy techniques that image lymph nodes in anesthetized mice vividly show the dynamics of this process. A snapshot from such a video (Fig. 2.5) shows T cells labeled with fluorescent markers squeezing through the walls of a large blood vessel called the high endothelial venule to enter a lymph node.

The lymph nodes are highly organized structures (Fig. 2.6). B cells are largely present in structures called follicles. T cells are present in structures around the follicles. T cells and B cells encounter pathogens and innate immune cells that can activate lymphocytes in lymph nodes and other peripheral lymphoid organs. After activation and proliferation, lymphocytes and their products (e.g., antibodies), as well as lymph, leave a lymph node by so-called efferent lymphatic vessels and return to the blood by the thoracic duct (Fig. 2.4B). Activated effector lymphocytes and antibodies then migrate into tissues by squeezing through blood vessel walls and attack pathogens at the sites of infection. The principal effector functions of lymphocytes are described later in this chapter.



2.6.2. The Structures of the BCR and TCR

Figure 2.7A shows an X-ray crystal structure of a BCR. Each receptor is composed of two identical light chains (of molecular weight ~25 kDa each) and two identical heavy chains (of molecular weight ~50 kDa each). As shown in the schematic (Fig. 2.7B), the two heavy chains are linked to each other by disulfide bonds, and each heavy chain is associated with a light chain. There are two types of light chains, called λ and κ . Both light chains of a particular BCR are of the same type, either λ or κ . There do not appear to be any functional differences between these two types of light chains, but it is interesting that the ratio of λ and κ chains varies among species. Each heavy and light chain has a so-called constant region and a variable region. The constant regions exhibit very limited diversity from one BCR to another. The different forms of heavy-chain constant regions and their functions will be discussed later in this chapter. Most of the diversity among BCRs is focused on the variable regions of the heavy and the light chains. The variable regions of the two chains come together to form a BCR's antigen binding site. Since each BCR consists of two identical heavy and light chains, each BCR has two identical binding sites that can bind to a pathogen's molecular components. Different BCRs are likely to have different variable regions and so are likely to exhibit different binding specificities. Each short arm of the Y-shaped BCR is called a *fragment for*

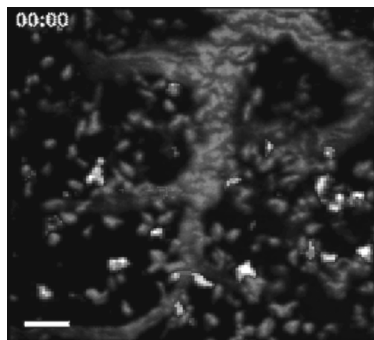


FIGURE 2.5. T cells labeled with a fluorescent marker are shown squeezing out of a large blood vessel (high endothelial venule) to enter a lymph node. The cells shown in white are those that have internalized an antigen (e.g., a DC). Image from the laboratory of U. von Andrian. Reproduced with permission.

FIGURE 2.4. (A) The primary lymphoid organs (bone marrow and thymus), secondary lymphoid organs, and lymphatic vessels. (B) The flow of immune cells and antigens through lymphatic vessels and blood. Redrawn from C.A. Janeway, P. Travers, M. Walport, and M. Shlomchik, *Immunobiology*, 5th edition, Garland Publishing (2001).

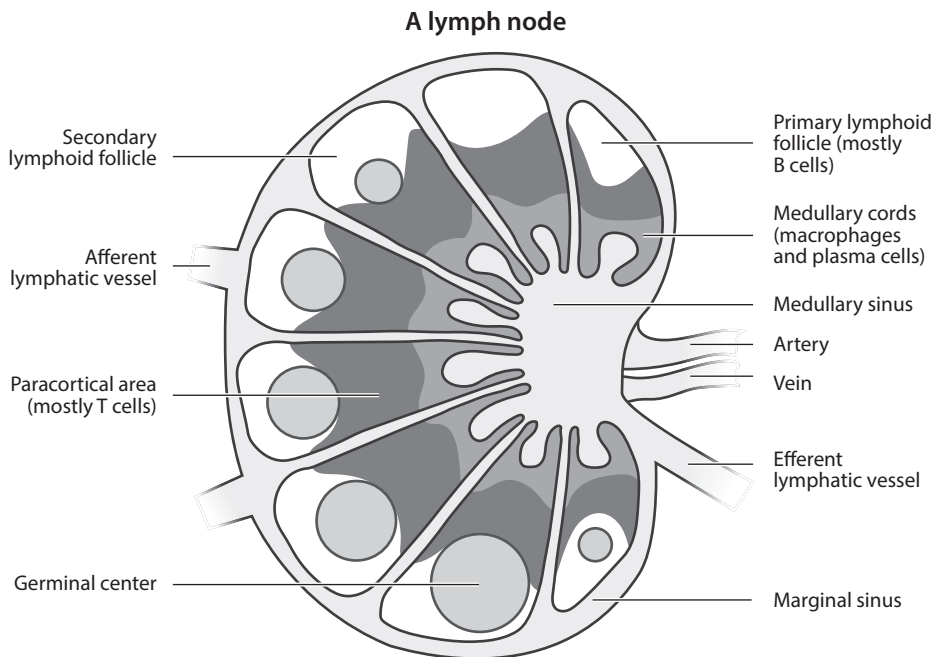


FIGURE 2.6. Schematic representation of the organization of different types of immune cells in a lymph node. Fluid from lymphatic vessels enters the lymph node via the afferent lymphatic vessel and leaves by the efferent lymphatic vessel. The organization of the lymph node helps promote interactions between T helper cells and B cells. The germinal centers discussed in section 2.6.8 are also depicted. Redrawn from C.A. Janeway, P. Travers, M. Walport, and M. Shlomchik, *Immunobiology*, 5th edition, Garland Publishing (2001). Copyright © 2001 by Garland Publishing.

antigen binding (Fab) and is often used in experiments to determine the binding properties and antigen specificities of a BCR.

When a BCR is bound to a ligand, the part of the BCR protruding into the cytoplasm of the B cell on which the BCR is expressed associates with other cytoplasmic molecules to initiate a cascade of biochemical reactions that ultimately can lead to new gene transcription programs and activation of the B cell. Activated B cells can mount potent pathogen-specific responses, by mechanisms discussed later in this chapter and in more detail in chapter 5.

The TCR is also composed of two chains of different types linked by disulfide bonds (Figs. 2.8A and B). The most common type of TCR is one with two types of chains, called α and β . A less common type of TCR is made up of chains called

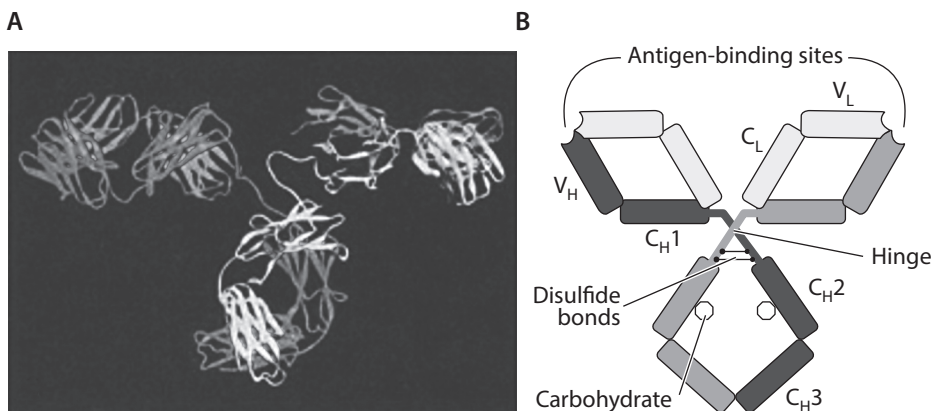


FIGURE 2.7. (A) A crystal structure of a BCR. The two identical heavy chains are shown in white and dark gray for clarity, and the light chains are shown in lighter gray. (B) Schematic representation of the BCR showing the disulfide bonds that link the two chains. Both heavy and light chains have variable and constant regions, labeled V_H , V_L , C_H , and C_L , respectively. Reproduced from C.A. Janeway, P. Travers, M. Walport, and M. Shlomchik, *Immunobiology*, 5th edition, Garland Publishing (2001).

γ and δ ; the function of these $\gamma\delta$ T cells is not completely understood. Each chain of an $\alpha\beta$ TCR has a constant region and a variable region. The variable regions of the α and β chains come together to form the binding site of the TCR. The binding sites of different TCRs are usually distinct, but the short tail of the TCR protruding into the cytoplasm of the T cell on which the TCR is expressed does not vary from one TCR to another. In chapter 6 we will discuss how this tail on a TCR bound to a ligand associates with other cytoplasmic molecules to initiate a cascade of biochemical reactions that can ultimately lead to new gene transcription programs and T cell activation. The effector functions of activated T cells help mount an adaptive immune response.

2.6.3. How a Great Diversity of Lymphocyte Receptors Is Generated

According to current estimates, roughly 10^7 different TCR β chain amino acid sequences and a similar or higher diversity of BCR sequences are generated in each individual. As there are roughly 20,000 protein-coding genes in humans, if receptor variants were encoded in the human genome as separate genes, with each lymphocyte expressing one possible variant, the number of genes

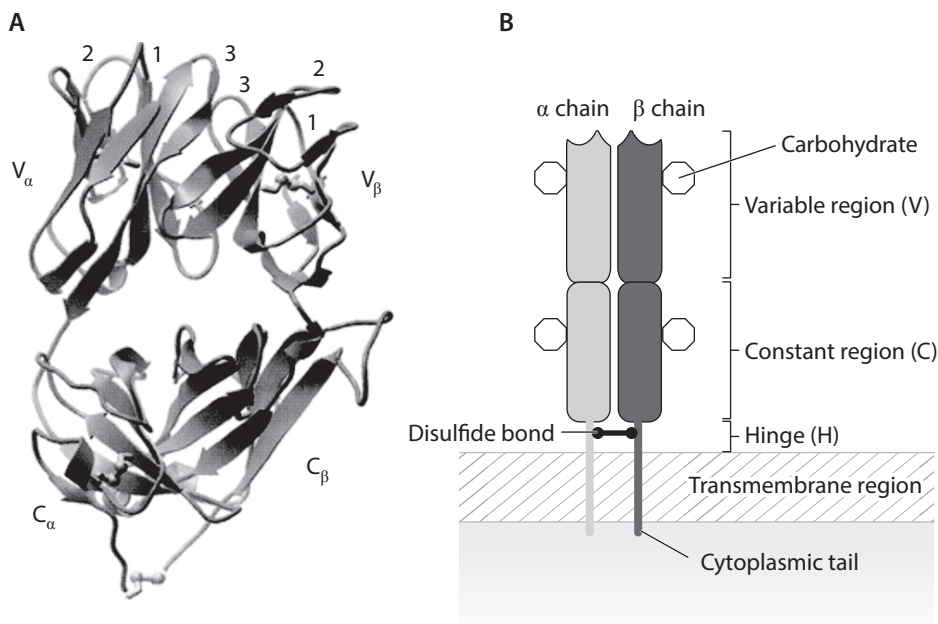


FIGURE 2.8. (A) The structure of an $\alpha\beta$ TCR. The α and β chains have constant and variable regions. The variable regions of the two chains come together to form the antigen binding region of the TCR. (B) Schematic representation of the structure of the TCR showing the disulfide bond that links the α and β chains. The cytoplasmic tail of the TCR is important for intracellular biochemical reactions that can lead to T cell activation upon TCR binding to an antigen. Reproduced from C.A. Janeway, P. Travers, M. Walport, and M. Shlomchik, *Immunobiology*, 5th edition, Garland Publishing (2001).

encoding for lymphocyte receptors would greatly exceed the total number of human coding genes. Susumu Tonegawa's elegant experiments in 1976 revealed the mechanism underlying the generation of the observed enormous diversity of BCR sequences. The mechanism for generating diverse TCR sequences is very similar.

The genes that code for a BCR are inherited from one's parents not as a single gene but as gene segments. The variable regions of light chains are encoded in two segments: the light chain variable region, V_L and the light chain joining region, J_L , which is so called because it joins the variable region to the constant region of the BCR gene. Similarly, the variable region of the heavy chain is made up of three segments: the V_H region, the J_H region, and the D_H

region (diversity region). There are many variants of each type of variable gene segment. During the development of each B cell, one variant of each type of gene segment is stochastically selected and connected by the process of DNA recombination to result in a single gene for a heavy or a light chain. The single gene resulting from this process of recombination codes for a particular variable region of a heavy chain or a light chain sequence. The entire process is called VDJ recombination. The variable regions of the heavy and the light chains are joined to the gene segments that code for their corresponding constant regions. The translated heavy and light chain proteins, whose gene segments recombine independently, then pair to form a particular BCR. The process of DNA recombination is irreversible, so if a B cell divides, the progeny will express the same BCR as the parent. In circumstances to be described later, the progeny can also express a mutated form of the receptor.

Noncoding regions adjacent to the different types of gene segments ensure that a V_L region can combine only with a J_L region (and not a different type of V_L region), D_H regions combine with J_H regions, and V_H regions are joined to the D_H - J_H complex. Some enzymes mediating recombination are lymphocyte specific, but others are not.

The diversity of BCRs that can be generated this way in a single human is enormous. As noted earlier, there are two types of light chains, κ and λ . For κ chains, there are 40 types of V_L regions and 5 types of J_L regions, so there are 200 different possible combinations of V_L and J_L genes that could code for a κ light chain. For λ chains, there are ~ 30 types of V_L regions and 4 types of J_L regions, resulting in 120 possible genes for λ light chains. Thus, there are 320 possible variants of light chains. There are roughly 40 possible V_H gene segments, 27 types of D_H segments, and 6 types of J_H segments. Therefore, there are roughly 6500 types of heavy chains. This simple calculation suggests that $6500 \times 320 \sim 2 \times 10^6$ distinct BCR genes could be assembled. This diversity is significantly augmented by *junctional diversity*, that is, the stochastic addition or deletion of some nucleotides at the junctions during recombination of the gene segments.

The actual diversity of the BCR repertoire available to a human is different from that obtained from the simple estimates just discussed, for several reasons. For example, although most combinations of heavy and light chains fold into functional BCRs, some do not. Also, the frequency of recombination of the different types of variable gene segments is not uniform, as assumed for making simple estimates.

The availability of high-throughput sequencing data has made it possible to create statistical models to quantitate the process of VDJ recombination

and thus characterize the statistical properties of T cell and B cell repertoires. We will not describe these models in detail, but the suggested readings at the end of this chapter will lead the reader to these elegant efforts to learn the statistical rules of VDJ recombination.

The basic approach is to write a model for the probability of observing a sequence that is the product of VDJ recombination, considering the possibility of different recombination scenarios resulting in the same sequence and the effects of sequencing errors. The probabilistic model for sequence generation contains parameters that are inferred by maximizing the likelihood of the observed sequence data. The TCR and BCR sequences in blood samples taken from organisms (including humans) have been used as data for inferring these maximum likelihood probabilistic models for VDJ recombination and characteristics of receptor repertoires.

The general scheme employed to generate a diversity of TCRs is very similar to that described for B cells. The α chains are encoded by diverse V and J genes, and the β chains are encoded by V, D, and J genes. But there are some differences. For example, the constant regions of the α and β chains of TCRs exhibit essentially no diversity. There is only one type of C_α region, and there are two types of C_β regions, with no functional differences between them. The reason is that the constant regions of T cells are important only for transducing biochemical signals inside the cell when a TCR is productively bound to its ligand and do not mediate T cell effector functions. The constant regions of BCRs have more diversity in helping carry out different B cell effector functions, as we shall soon describe.

2.6.4. *What Features of Pathogens Do BCRs Bind (Recognize)?*

BCRs can bind to diverse pathogenic components, including proteins, parts of proteins, and parts of sugar molecules or proteins on the surfaces of bacteria and viruses. The part of the pathogenic component that a BCR binds to is called its *epitope*. BCRs bind to their epitopes via the antigen binding site created by the variable regions of their heavy and light chains. A BCR's epitope can comprise amino acids that are not contiguous along the sequence of the pathogenic protein, sugar, or other component, owing to the nonlinear shapes of pathogenic proteins or sugars and the BCR antigen binding site. Antibody-antigen binding is mediated by noncovalent interactions like hydrogen bonds and electrostatic and van der Waals interactions. Three-dimensional shape

complementarity between the BCR's antigen binding site and the epitope favors stronger noncovalent interactions.

B cells principally interact with pathogens present in blood, lymph nodes, and extracellular spaces in tissues. They can also interact with intracellular pathogens, such as viruses, because such pathogens traverse extracellular spaces to infect new host cells. If a BCR binds to a surface protein of a pathogen sufficiently strongly, intracellular biochemical reactions can result in B cell activation. Activated B cells undergo processes described later in this chapter and chapter 5 that ultimately lead to the secretion of modified and soluble forms of the BCR, called antibodies. Antibody binding to pathogens in blood and extracellular spaces results in functions mediated by the innate immune system that can eliminate the pathogen (discussed later). It is important to note that some antibodies bind to epitopes of the pathogen's proteins that prevents the pathogen from infecting new cells. For example, an antibody can bind to an epitope on the surface protein of a virus that includes the region that binds to host cell receptors to infect cells. Such antibodies are called *neutralizing antibodies* and are very potent, because they inhibit infection and act through other effector functions. Antibodies can also act on infected cells via various effector functions, for example, by binding to virus particles budding out of the infected cell.

2.6.5. *What Features of Intracellular Pathogens Do TCRs Recognize?*

The immune system mounts pathogen-specific responses against infected cells principally through the action of T cells. Molecular markers of intracellular pathogens in phagocytic cells and other cell types infected by pathogens are displayed on the surface of these cells, appropriately called *antigen presenting cells* (APCs). Through their TCR, T cells recognize the presence of the pathogen-derived molecular markers displayed by APCs.

Viruses replicate by synthesizing their proteins in host cells. For example, most infecting RNA viruses make many copies of their RNA genome in the cytoplasm of infected cells, and the host cell's ribosome then translates these RNA molecules into proteins. Retroviruses, like HIV, integrate their genome into the host cell's DNA and hijack the host transcriptional and translational apparatus to produce viral proteins, which are then assembled into new virus particles in the cytosol and on the host cell's membrane. Some bacteria also replicate in the cytosol. Proteins in the cytosol can be cleaved into short

peptides by the host cell's proteasomes, which are large protein complexes that usually degrade used or misfolded host proteins. The pathogen-derived peptides generated in the cytosol by proteasomes are the molecular markers of pathogens ultimately displayed on the surface of APCs.

Proteins or peptides that are eventually transported to the cell surface are assembled in a subcellular structure called the *endoplasmic reticulum* (ER). Specialized proteins transport peptides generated in the cytosol across intracellular membranes into the ER. The *major histocompatibility complex* (MHC) genes in vertebrates encode information for two classes of proteins called MHC class I and MHC class II molecules. MHC class I molecules are present in excess in the ER, which allows them to rapidly bind to some of the peptides transported from the cytosol to the ER. Subsequently, the peptide-MHC complexes (pMHC) are transported to the cell surface and act as molecular flags of infection.

Cytotoxic T lymphocytes (CTL) are a kind of T cell with TCRs that can potentially recognize (i.e., bind sufficiently strongly to) peptides bound to MHC class I molecules. In addition to the TCR, the CTLs display a molecule called CD8 on their surface. CD8 can bind to MHC class I molecules, and the TCR binds to the pMHC complex (Fig. 2.9). An enzyme called Lck that can initiate downstream biochemical reactions necessary for T cell activation is linked to the cytoplasmic tail of CD8. Other molecules also bind across the T cell–APC junction to mediate T cell activation. In chapter 6 we will discuss the biochemical signals triggered by TCR–pMHC interactions that result in activation. In chapter 3 we will describe the conditions necessary for activated T cells to proliferate and mediate effector functions. The effector functions of CTLs will be described later in this chapter.

Many bacteria and some parasites are taken up by phagocytic cells in cell membrane-bound cytosolic vesicles called *endosomes*. B cells also internalize pathogens in endosomes upon their BCR binding to them sufficiently strongly—a process necessary for potent antibody production (discussed in a later section). Pathogens that live and replicate in vesicles and endosomes are not accessible to proteasomes. Proteases and peptidases in endosomes and vesicles can degrade proteins to generate short peptides, which are then delivered to the cell surface by binding to MHC class II proteins. MHC class II proteins are also present in the ER, but they are prevented from binding peptides in this compartment because a protein called *MHC class II-associated invariant chain* binds to MHC class II and partially blocks the peptide-binding groove in the molecule. MHC class II-associated invariant chain also plays a key role in targeting MHC class II molecules to acidic vesicles and endosomes. In these

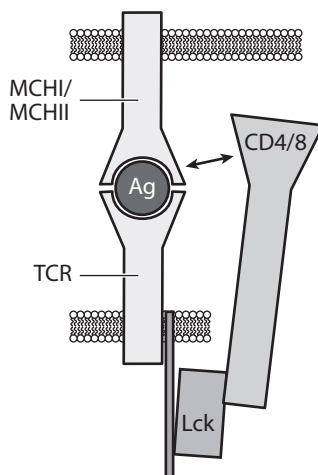


FIGURE 2.9. The TCR on a CD8 T cell (CTL) can interact with an antigenic peptide (Ag) bound to MHC class I molecules. CTLs also express a molecule called CD8 that can bind to MHC class I molecules. An enzyme called Lck bound to the cytoplasmic tail of CD8 is important for initiating intracellular biochemical reactions upon TCR binding to antigen. The TCR on T helper cells can bind to Ag bound to MHC class II molecules. In T helper cells, a molecule called CD4 binds to MHC class II; CD4 is also associated with Lck.

acidic compartments, the invariant chain is cleaved, allowing the binding of peptides therein to MHC class II molecules. Specialized intracellular vesicles transport the pMHC class II complexes thus generated to the cell surface.

T helper cells are another type of T cell. They can bind to peptide-MHC class II molecules displayed on the surface of phagocytic cells and B cells that have ingested pathogen. T helper cells express a molecule called CD4 on their surface that can bind to MHC class II molecules, and the TCR binds to the pMHC complex (Fig. 2.9). The same enzyme bound to CD8 is also bound to CD4, and it can initiate downstream biochemical reactions necessary for T cell activation. There are many varieties of T helper cells, which mediate different effector functions, as will be described later in this chapter.

Almost all types of cells express MHC class I proteins, perhaps because all cells can potentially be infected by viruses, which replicate in the cytosol. Red blood cells do not express MHC Class I. The parasite that causes malaria can live in red blood cells protected from the action of CTLs. MHC class II molecules are largely expressed on B cells and cells presenting antigen to T cells (DCs, macrophages, etc.).

2.6.6. *There Is a Great Diversity of MHC Molecules in a Population, Exhibiting Different Peptide Binding Characteristics*

In humans the MHC genes are located in chromosome 6 and are called *human leukocyte antigen* (HLA) genes. Three loci code for HLA class I molecules, and three code for HLA class II molecules. The three major HLA class I genes are called HLA-A, HLA-B, and HLA-C, and the three class II genes are called HLA-DR, HLA-DP, and HLA-DQ. Another type of HLA molecule (HLA-E) is also expressed on most cells, but usually at low levels. The HLA genes are among the most polymorphic genes in humans, with thousands of variant alleles in the human population. Each individual can express up to two alleles (one from each parent) at each locus. Thus, an individual can typically express three to six types of HLA class I molecules and three to six types of HLA class II molecules. Different HLA molecules tend to bind to different peptides, thereby enabling an individual to bind to peptides derived from diverse pathogens. Because of the enormous diversity of HLA alleles in the population, two individuals who are not identical twins are highly unlikely to express an identical set of HLA molecules, although some may be common. To emphasize the importance of the polymorphism of HLA genes for protecting the human population from diverse pathogens, consider the scenario in which a single type of HLA molecule was expressed in all humans. If a pathogen evolved whose peptides did not bind to this HLA molecule, no one in the population would be able to mount a T cell response to this pathogen, thus potentially wiping out the population.

To generalize for vertebrates, we will return to referring to the molecules that present peptides to T cells as MHC molecules. MHC class I molecules consist of two chains that form four domains (Fig. 2.10A). One chain, called the α chain, has a molecular weight of 43 kDa and comprises three domains (α_1 , α_2 , and α_3). The other chain, called β_2 -microglobulin, has a molecular weight of 12 kDa and is noncovalently associated with the α chain. Only the α chain spans the cell membrane. The α_1 and α_2 domains of MHC class I molecules come together to form the peptide-binding groove. The free amino and carboxy terminal ends of peptides bind to the MHC groove by hydrogen bonding and electrostatic interactions to invariant sites found at both ends of MHC class I molecules (Fig. 2.10B). The peptides sit in elongated forms in the groove; peptides that are 8–11 amino acids long can bind to MHC class I molecules, with 9-mers being most common. Two amino acids of the peptide, called *anchor residues*, are buried in the MHC class I molecule groove to further facilitate peptide binding, with diverse peptides able to bind to a given

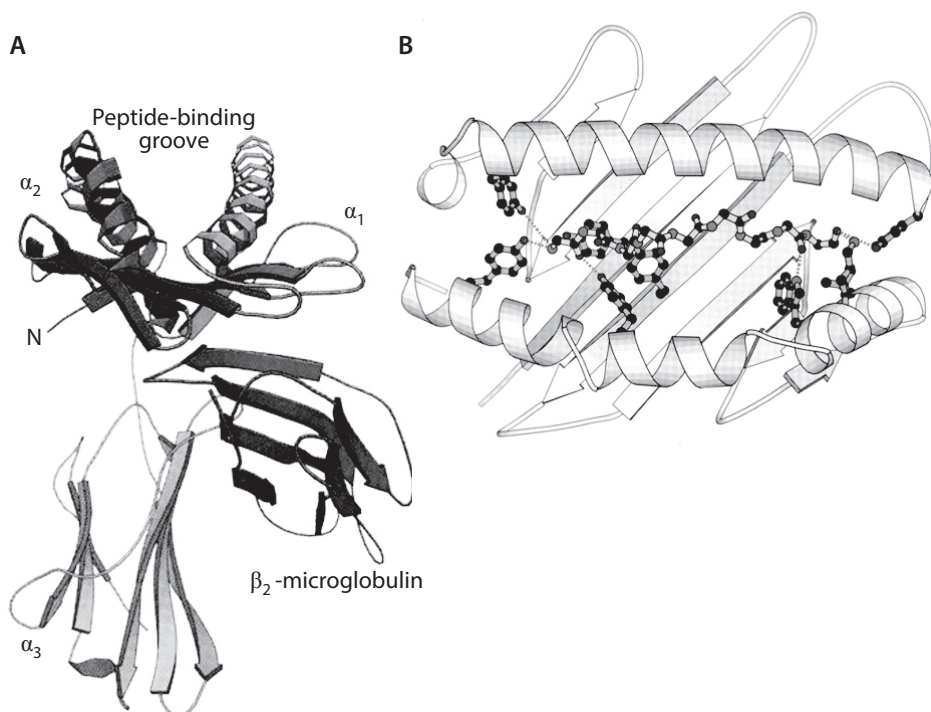


FIGURE 2.10. (A) A schematic representation of an MHC class I molecule showing the α_1 , α_2 , and α_3 domains and the β_2 microglobulin. (B) A bound peptide is shown in the MHC molecule's peptide binding groove. Reproduced from C.A. Janeway, P. Travers, M. Walport, and M. Shlomchik, *Immunobiology*, 5th edition, Garland Publishing (2001). Copyright © 2001 by Garland Publishing. Used by permission of W.W. Norton & Company, Inc.

MHC class I molecule. Different MHC class I alleles prefer different types and distinct positions of anchor residues. The major sites of polymorphism in MHC molecules lie in the peptide-binding groove, thus enabling them to bind to different sequences of peptides. Machine learning-based models trained on experimental data have been developed to predict the binding of different peptides to distinct MHC molecules. They are reasonably accurate in predicting whether a peptide will bind to a given MHC class I molecule.

MHC class II proteins are made up of two chains (α and β), both of which span the membrane (Fig. 2.11A). There are four domains: α_1 , α_2 , β_1 , and β_2 . The ends of the peptide-binding groove in MHC class II molecules are more open compared with those of MHC class I molecules, and the ends of the

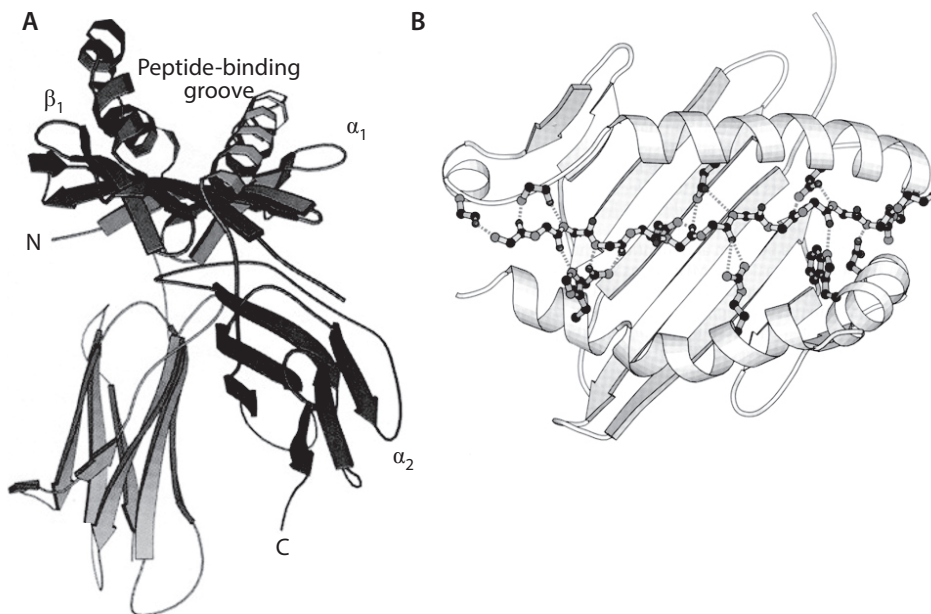


FIGURE 2.11. (A) A schematic representation of an MHC class II molecule showing the different α and β domains. (B) A bound peptide is shown in the MHC class II molecule's peptide binding groove, which is more open than the MHC class I molecule's groove. Reproduced from C.A. Janeway, P. Travers, M. Walport, and M. Shlomchik, *Immunobiology*, 5th edition, Garland Publishing (2001). Copyright © 2001 by Garland Publishing. Used by permission of W.W. Norton & Company, Inc.

peptide hang out of the groove (Fig. 2.11B). Thus, MHC class II molecules bind to peptides that are longer and of more variable length (~13–17 amino acids). Anchor residues are buried in the groove and are essential for peptide binding. Because of the open ends and diversity of peptide lengths that can bind to MHC class II proteins, it is harder to predict which peptides are likely to bind to MHC class II molecules using machine learning–based inference (compared with predictions for MHC class I molecules), but these models are becoming increasingly more accurate.

Crystal structures of TCRs interacting with pMHC molecules show that the TCR is oriented diagonally over the peptide and peptide-binding groove (Fig. 2.12). In MHC class I molecules, the α chain of the TCR sits over the α_2 domain of the MHC and the amino terminal end of the peptide; the β chain of the TCR lies over the α_1 domain and the carboxy terminal of the peptide.

The most variable regions of the TCR are the complementary determining region 3 (CDR) loops. CDR3 loops of the TCR are aligned on top of the central exposed amino acids of the peptide. Structures of pMHC class II–TCR complexes reveal similar features.

2.6.7. T Cell Activation and Effector Functions

Phagocytes, such as DCs and macrophages, internalize pathogens and display pathogen-derived pMHC molecules on their surface. Sufficiently long-lived binding of a T cell's TCR to a pMHC molecule results in intracellular biochemical reactions that can lead to new gene transcription programs and T cell activation. In chapter 6 we will discuss these biochemical reactions that enable T cell activation and why sufficiently long-lived TCR–pMHC binding is required. Host proteins (e.g., misfolded ones) are also processed into short peptides and displayed on the surface as pMHC molecules. Considered together, the topics discussed in chapters 3 and 6 will provide insights into how T cells discriminate between self- and pathogen-derived pMHC molecules and how the biochemical network in T cells enables T cells to respond to just a few pathogen-derived pMHC molecules in a background of many self-pMHC molecules. Activated T cells can proliferate and mediate different effector functions.

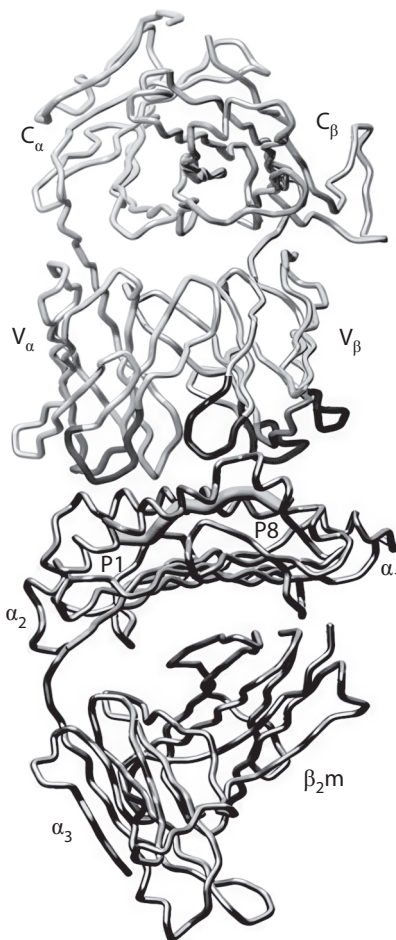


FIGURE 2.12. A crystal structure of a TCR binding to a peptide bound to an MHC class I molecule. The text describes more details of the structure. Reproduced from C.A. Janeway, P. Travers, M. Walport, and M. Shlomchik, *Immunobiology*, 5th edition, Garland Publishing (2001). Copyright © 2001 by Garland Publishing. Used by permission of W.W. Norton & Company, Inc.

A CD8 T cell expressing a TCR that binds sufficiently strongly to a pMHC molecule displayed on an APC can become activated and proliferate. If an activated CD8 T cell interacts with another infected cell displaying the same pMHC molecule that activated it, it can secrete various products that can kill the infected cell, which is why CD8 T cells are also called cytotoxic T lymphocytes (CTLs) or killer T cells. CTLs mediate an important effector function of the immune system by clearing infected cells.

As noted earlier, phagocytes, B cells, and others express MHC class II proteins, and T helper cells can be activated by sufficiently long-lived interactions between their TCRs and pMHC class II molecules. Activated T helper cells can proliferate and differentiate into a variety of effector cell types and coordinate an immune response in different ways. *Regulatory T cells* (T_{regs}) are one type of T helper cell, and as we will discuss in chapter 3, they play an important role in suppressing autoimmune responses. Activated T helper cells that are not T_{regs} differentiate into two broad classes of effector cells called Th1 and Th2 cells. Innate immune cells, especially phagocytes, secrete specific cytokines based on the kind of pathogen detected, and the cytokine environment drives the response of the adaptive immune system. Cytokine IL-12 promotes differentiation to Th1 cells, and IL-4 is necessary for generating Th2 cells. Various other factors, such as the types of hormones in the environment and antigen dose, also determine whether activated T helper cells differentiate into Th1 or Th2 cells. This differentiation occurs 48–72 hours after activation, so there is ample time for innate immunity to sense the type of danger and instruct the adaptive immune system appropriately.

The two broad categories of immune responses to infection with a microbial pathogen are called type I and type II and are principally orchestrated by Th1 and Th2 cells, respectively. While these two types of cells carry out different functions and secrete various cytokines, the clearest difference between them is that Th1 cells secrete interferon-gamma ($\text{IFN}\gamma$), and Th2 cells secrete IL-4. $\text{IFN}\gamma$ stimulates upregulation of MHC class I and II molecules, intracellular killing of pathogens, and secretion of cytokines that cause inflammation. Such a proinflammatory response accompanied by CTL activation and killing of infected cells is the hallmark of a type I immune response. Type II responses directed by Th2 cells are characterized by production of large amounts of antibodies (see next section). Responses to parasites and allergens are type II responses. It is important to note that type I responses do not actively suppress type II responses, and so antibodies are produced during a type I response. Conversely, type II responses actively suppress type I responses.

2.6.8. B Cell Activation and Antibody Production

As noted earlier, if a B cell's BCR binds sufficiently strongly to its epitope, intracellular biochemical reactions inside the B cell contribute to its activation. A BCR bound to its pathogenic epitope can also internalize the pathogen, because the B cell membrane, and its associated cytoskeleton, exerts a force that pulls the BCR toward the membrane. If the BCR binds sufficiently strongly to the pathogenic epitope, the pathogen is internalized. Otherwise, the BCR-pathogen epitope bond breaks. (The process of antigen internalization is discussed later in this section.) Internalized antigen is then processed, and the resulting pMHC class II complexes are displayed on the B cell surface. B cells with BCR that bind more strongly to the antigen are likely to internalize more antigen and display more pMHC molecules on their surface.

T helper cells that have been independently activated by APCs can then interact with B cells displaying antigen-derived pMHC molecules. The probability of interactions between these B cells and T cells activated by the same pathogen is enhanced by the organization of the lymph node (Fig. 2.6). The B cells migrate through the T cell zone in lymph nodes to increase the encounter probability. Interactions between the TCR on T cells and pMHC molecules displayed on B cells with internalized antigen can result in B cell activation, and B cells displaying larger numbers of pMHC molecules are more likely to have productive interactions with T cells. Activation also requires other molecules on the T cell surface and depends on the cytokines secreted by the T cells.

A key interaction is that between a protein called CD40 expressed on the B cell surface and its ligand on the T cell surface.

The BCR has a hydrophobic constant region anchoring it to the B cell membrane. Activated B cells proliferate, with some differentiating into a form called plasma cells, and secrete a form of the BCR with a hydrophilic constant region. This soluble form of the BCR is an antibody. As mentioned in the previous chapter, antibodies are what von Behring and Kitasato found in the serum of vaccinated or infected individuals, and they have important effector functions.

The first phase of antibody production just described is consistent with Burnet's proposal that there are many B cell clones with different antibodies, and those binding to a particular infecting pathogen sufficiently strongly expand upon encounter. The discovery of VDJ recombination explained how a diverse repertoire of B cell clones is created in a single individual.

But an important puzzle remained. The antibodies produced during the first stage of their production have a relatively low affinity for the pathogen. Well

Simplest haptens

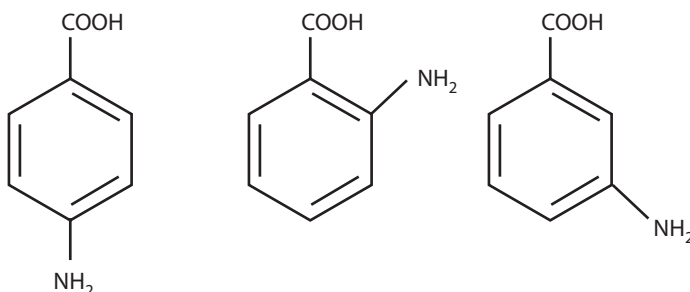


FIGURE 2.13. The haptens shown are isomers; immunization with one of them results in antibodies specific for the immunizing hapten and not for others.

before the discovery of VDJ recombination, it was observed that the antigen affinity and potency of antibodies increased with time after immunization. This was definitively demonstrated by Eisen and coworkers, who immunized rabbits with small-molecule antigens called *haptens* (Fig. 2.13). Antibodies specific for the immunizing hapten, and not the others, developed, and the affinity of the antibodies increased with time after immunization. It was subsequently shown that the increase in affinity with time is accompanied by mutations (so-called somatic mutations) in the BCR sequences initially activated by the immunizing antigen. These mutations and the increasing affinity of antibodies for the pathogen with time after infection or vaccination are the result of a Darwinian evolutionary process called *affinity maturation*. Studies in mice revealed the mechanisms underlying affinity maturation, which we now describe.

After the first phase of the antibody response described earlier, activated B cells form and enter structures in lymph nodes (Fig 2.14) called *germinal centers* (GCs), where affinity maturation occurs. GCs contain many types of cells in addition to B cells, notably phagocytes called *tingible-body macrophages*.

The GC is divided into two zones (Fig. 2.14), labeled the *dark zone* (DZ), because it appears dark in images obtained by light microscopy, and the *light zone* (LZ), because it has a lighter appearance. The DZ contains B cells, and the LZ contains a network of cells called *follicular dendritic cells* (FDCs), B cells, and T cells (mostly CD4 T cells).

Affinity maturation has been studied extensively over the years, and recent studies using multiphoton microscopy to image GCs in live animals or explanted lymph nodes kept “alive” in culture have clarified several aspects of the

(continued...)

INDEX

Note: Page numbers followed by *f* refer to figures.

- ACE method. *See* Adaptive Cluster Expansion (ACE) method
- acquired immunodeficiency syndrome. *See* AIDS
- activated B cells, 24, 29; and antibody production, 37–43; differentiation of, into plasma cells, 37; entering germinal centers (GCs), 38, 39, 39*f*
- activated CD4 T cells: differentiation, 83–84
- activated CD8 T cell, 36
- activated T cells, 30, 35, 53, 78–82; differentiation of, 51*f*; effector functions of, 25, 35–36, 221; proliferation of, 49–50, 80–81; as sensed by regulatory T cells, 84. *See also* T cell activation
- activated T helper cells, 36
- activation-induced cytidine deaminase (AID), 39, 40, 43, 144, 145, 175–76
- activation threshold, 67, 78
- Adaptive Cluster Expansion (ACE) method, 101–6
- adaptive immune system, 14, 19–47; activation of, by innate immune system, 16, 17, 19, 49, 88, 221; cell types of, 14–16; and cytokine regulation, 19, 36; degraded by HIV, 90; macrophages and DCs activation in, 17–18; and pathogen-specific immunological memory, 14, 47, 50; and pathogen-specific responses, 14, 20–21; schematic summary of, 49–51, 51*f*
- adenovirus vectors, 89, 134
- adjuvants, 88
- afferent lymphatic vessels, 21
- affinity maturation (AM), 37–47, 143–96; B cells and, 38, 48, 49, 51, 157, 160–61, 175, 187; GC B cells and, 38–39, 42, 43–45, 144; mutation of BCRs during, 185; in response to a single antigen, 144–49; simple model of, upon immunization with multiple variant antigens, 154–73; simulations of, with a shape-space model, 182–89; stochastic agent-based simulation method of, 173–82
- AIDS, 13, 90, 123
- allergens: responses to, 36, 45–47
- $\alpha\beta$ TCR, 24–25, 26*f*
- amino acids: and HIV polyproteins (Potts model), 99, 101, 106–8; hydrophobicity of, 70, 72, 73, 174; at P6 and P7 sites, 72–73; for YAc62 and B3KS06, 72. *See also* peptide amino acids
- amino acid strings, and interaction free energy between BCR paratope and antigen's epitope, 174
- amino acid variation (entropy), 96–97
- anchor residues, 32, 33, 34
- ancient ideas about disease, 2–3
- antibiotics, 13
- antibodies, 12, 20, 37; affinity (free energy of binding) of, for an antigen, 155, 196; affinity and potency of, for antigens, 38,

- antibodies (*continued*)
 - 42–43, 155–68, 173–89, 196; binding to epitopes, 29; binding to pathogens, 45–47, 143; binding to surface proteins, 47, 143; breadth of coverage of, 155; breadth state of, 155–68, 156*f*, 187, 188–89; classes of, 42, 43; and class switching, 45; clearing pathogens in concert with phagocytes, 45–46, 46*f*; and effector functions, 29, 45–47, 143; and interaction free energy with antigen/epitopes, 174–75, 182–83, 184*f*; key functions of, 143; “net” of, targeting different epitopes, 195–96; neutralization capability of, 155; relative abundance of different classes of, in blood, 43; strain-specific, 151; and type II responses, 36. *See also* bnAbs
- antibody-bound bacteria, and opsonization, 45
- antibody-bound pathogens, interaction of, with complement, 45
- antibody evolution, 149–96; computer simulations of more complex models of, 173–89; and simple model of AM upon immunization with multiple variant antigens, 154–73; simulations of, with a shape-space model, 182–89
- antibody production: and B cell activation, 37–45; first phase of, 37–38, 43; generation of, by class switching, 43
- antibody response, 47; affinity maturation of, 37–47, 143–96
- antigenic epitopes: interactions of, with BCR’s paratope (“string model”), 174; and shape space model, 183
- antigenic peptides, 58
- antigen internalization, 37, 40–41, 42
- antigen presenting cells. *See* APCs
- antigen specificity of T cells, thymic selection influencing, 58–78
- APCs, 29–30, 37, 40, 82–83, 200, 204, 215
- apoptosis (programmed death), 40, 42, 73, 83
- artificial intelligence, 4, 5, 6, 196
- attenuated vaccines, 12, 87–88
- autoimmune diseases, 2, 3, 76, 78
- autoimmune regulator (AIRE) gene, 55–56
- autoimmune responses: T cell repertoire
 - avoidance of, 78, 80, 84; why T cells and B cells suppress, 47–48
- autoimmunity, 77, 78, 85
- autoreactive T cells, 78
- B6 mice, 74, 75, 76, 77
- bacteria, 11, 12, 30, 45, 46*f*
- bacterial toxins, antibodies binding to, 45
- basophils, 15*f*
- B cell activation, 24, 29; and antibody production, 37–45
- B cell receptors. *See* BCRs
- B cells, 16, 20, 38, 48, 49, 51, 157, 160–61, 175, 187; and affinity maturation, 38–45, 144–49, 154–89. *See also* germline B cells
- Bcl-2 and Bcl-6, 40
- BCRs, 20; antigen binding sites of, 23, 25*f*, 28; constant and variable regions within, 23, 25*f*, 28; disulfide bonds in, 25*f*; diversity of, 27–28; heavy and light chains within, 23, 25, 26–27; hydrophobic constant region of, 37; and interaction free energy with paratopes, 174, 183; and probability of AID-mediated mutations at specific sites, 145; structure of, 23–25, 25*f*; and VDJ recombination, 27–28, 37–38, 49
- Bim $-/-$ mice and Bim $+/+$ mice studies, 73–74, 75
- binding free energies: and BCR–antigen interactions, 174, 183; and TCR binding to peptide-MHC molecule, 64
- biomolecular condensate, 216
- blood: flow of immune cells through, 22*f*; relative abundance of different classes of antibodies in, 43
- blood clotting, 19
- blood vessel dilation, 18
- blood vessel walls, enhanced permeability in inflammation, 18–19

- B lymphocytes. *See* B cells
- bnAbs, 149–52, 152*f*, 153*f*; mechanistic models for evolution of, 154–89; PGT121, 191, 192–93; VRCO1, 191, 192
- Boltzmann distribution, 100
- Boltzmann entropy formula, 135
- bone marrow, 14, 15, 20, 21, 48, 54, 55
- broadly neutralizing antibodies. *See* bnAbs
- Burnet, Macfarlane, 20, 37
- cancer immunotherapy, 2, 220–21
- cancer vaccines, 134
- canonical ensemble, 135, 137, 139, 140
- CD3 chains (TCRs), 201
- CD4 coreceptor, 92, 93, 150, 201
- CD8 coreceptor, 31, 55, 201
- CD4 T cells, 31, 31*f*, 55, 73, 92, 149; binding to MHC molecule, 31, 201; HIV infecting and killing, 89–90, 94–95; HIV's spike binding to, 122, 150, 152, 176
- CD8 T cells, 30, 31*f*, 36, 55, 73
- CD28 molecule, 83
- CD-40 (protein), 37, 41, 43
- CD80 molecule, 83
- CDR3 region, 40, 63, 72, 85; and hydrophobic residues, 75*f*, 77*f*
- channel capacity, 215
- chemokines, 18, 19
- chicken cholera inoculation, 11–12
- class switching, 43, 45
- clonal selection theory, 20
- cognate pMHC molecules, 200
- common lymphoid progenitor, 14, 15, 16
- common myeloid progenitor, 14, 15–16
- compensatory mutations, 96
- complement, 17, 18, 45
- complement activation, 45, 46*f*
- cortex, 55
- cost to the host of immune responses, 19
- COVID-19 pandemic, 1, 2, 7, 44, 222; and epidemiological models, 222; and vaccine development, 88–89, 219
- cross-entropy (ACE method), 102–6
- cross-reactive antibodies from sequential immunization with variant antigens, 190–93
- CTLA4, 221
- cytokines: actions of, 18–19; and activated T cell proliferation, 80–81; and adaptive immune response, 19; effector functions of, 19; and regulation of T cell proliferation and differentiation, 84; and Th1 or Th2 differentiation, 36; T helper cell secretion of, 43; uncontrolled secretion of, 19
- cytokine storms, 19
- cytosol, 29–30
- cytotoxic T lymphocytes (CTLs), 30, 36, 49; and CD8 coreceptor, 31; effector function of, 36; vaccine-induced responses of, 149
- dark zone (DZ), 38, 39, 39*f*, 40, 144, 145
- Darwinian evolution of B cells, 38, 49, 154–89
- DEC205 antibody–OVA construct, 41–42
- dendritic cells (DCs), 17–18, 83
- dengue, 89
- direct coupling analysis (DCA), 122
- disease: and microorganisms, 3, 10; pathogens as causative agent of, 11
- DNA viruses, 92
- double-negative thymocytes, 55
- double-positive thymocytes, 55
- downregulation of CD4 and HLA proteins, 93
- effector cells (Th1 and Th2), 36
- effector functions, 19, 29; and activated T cell, 25, 35–36, 221; of antibodies, 45–47, 143
- effector T cells, 85
- efferent lymphatic vessels, 21
- elite controllers, 118, 123–24, 125
- endoplasmic reticulum (ER), 30
- endosomal TLR, and differentiation between self and foreign nucleic acids, 18
- endosomes, 17, 30

- entropy: cross-entropy (ACE method), 102–6; and information, 137–40; intensive (ACE method), 102–6; and Potts model, 108, 109; Shannon's, 138–40; in statistical mechanics, 99–100, 135–37, 140
- ENV polyprotein of HIV, 120, 122
- eosinophils, 16
- epitope masking, 45, 195
- epitopes, 45, 124–33, 154, 173; immuno-dominant, 194–95; “net” of antibodies targeting different, 195–96; represented by modeling the antigen as a string of N spins, 173–74
- erythrocytes, 16
- escape mutations, 114, 118, 125–28, 130, 132
- escape time, 125–26, 127, 132, 133
- evolutionary dynamics, 45, 149–93
- evolution of antibodies that can target diverse mutant strains of a pathogen, 149–93
- experimental evidence for quorum testing by T cells, 82–85
- extracellular signal-regulated kinases (ERKs), 203
- extra germinal center (EGC) processes, 43; affinity-dependent selection and expansion, 43, 44f; differentiation into plasma cells, 43–44; upon reexposure to antigen, 44
- extreme value statistics: analysis of thymic selection effects in terms of, 62–69; primer on, 59–62
- Fc receptors, 45
- fetal thymic organ culture (FTOC), and positive and negative selection thresholds, 56, 67
- fever, 19
- fitness landscape, 87–134, 140; of HIV mutants, 109–19; of HIV polyproteins, 98–119; and prevalence landscape of HIV polyproteins, 98–109; and tests against clinical data, 119–20, 123–33; and tests against in vitro measurements of HIV mutant strains, 119–20, 121–23
- follicles, 21
- follicular dendritic cells (FDCs), 38, 40, 42, 44–45, 144–46, 160, 194; and GC processes, 194; interacting with antigens, 41f, 145–46, 160, 189, variant antigens displaying homogeneously on, 160, 190; variant antigens distributed heterogeneously on, 160–61
- fragment for antigen binding (Fab), 24
- fungi, 12, 13
- Gag polyprotein of HIV, 106, 107, 120, 121, 128, 130
- $\gamma\delta$ TCRs, 25
- GCs, 38, 39f; cell types in, 38; dark and light zones of, 38, 39f, 144, 145; formation of, 39; memory cells entering, 193–94; Perelson's model of, 144–47, 149; response of, to reexposure to antigen, 44–45, 44f; secondary, 44–45, 194–95; evolutionary dynamics in 154–89
- germinal centers. *See* GCs
- germline B cells, 145, 147, 157, 163, 175, 189, 191, 193–95; activated with variants of full HIV spike, 192–93; evolving into bnAbs targeting HIV's CD4 binding site, 154–55, 157; GC seeded with, 175–82
- germline BCR sequences, 218
- Gillespie algorithm, 161, 162, 211
- gp120 monomer, 190, 191f, 192f
- granulocytes, 15
- Growth factor receptor-bound protein 2 (Grb2), 203, 213
- Gumbell distribution, 62, 63
- haemagglutinin (HA) protein (influenza viral spike), 150, 153f; binding to sialic acid (RBS), 150, 151; bnAbs targeting, 195
- Hamiltonian, 114, 115–16, 117, 119

- Hamming distance, 112
- haptens, 38
- heat, redness, and pain in inflammation, 18–19
- heavy chain BCRs, 23; constant and variable regions of, 23, 25*f*; variable region encoding of, 26–28
- Helmholtz free energy, 137
- hepatitis C virus (HCV), 76, 77, 88, 89
- higher organisms, immune system cells and organs of, 14–16
- highly mutable viruses, fitness landscape of, 87–140
- HIV, 3, 13, 29, 76, 77, 89, 219; and activated germline B cells, 154–55; antibody response and virus population coevolution in, 196; biology of, 91–96; bnAbs against, 149–51, 152, 154, 189, 191–93; CD4 binding site of, 122, 150, 152, 154, 173, 176, 190; drugs for prophylaxis against, 134; as enveloped retrovirus, 92, 93*f*; evading immune responses, 89, 96; fitness landscape of, 97–119, 149; gene coding of, 92, 93; genome of, 95; and gp120 and gp41 proteins, 92, 93; infecting and killing human CD4 T cells, 89–90, 94–95; infections and mortality rates of, 91; life cycle of, 92–93, 94*f*; mutational vulnerabilities of, 89, 96–134, 140; mutations of, during life cycle, 95–96; replication rates of, 96; reverse transcription mutations in, 95–96, 131; structure of, 92, 93*f*; and temporal evolution, 93–95, 95*f*, 151; vaccine for, 92, 151, 153; viral spike of, 92, 93, 122, 149–50, 152*f*, 153, 154, 173, 176, 190, 192–93
- HIV mutant strains: connection between fitness and prevalence of, 109–19; diversity of, 89, 90*f*; least biased model for prevalence of, 99–100, 135, 140; and tests against clinical data, 123–33; and tests against in vitro measurements of fitness, 120–23
- HIV polypeptides, 92, 111, 120, 134; and Adaptive Cluster Expansion (ACE) method, 101–6; and amino acids (Potts model), 99, 101, 106–8; inferring fitness landscapes of, 98–119; inferring the prevalence landscape of, 98–109; Los Alamos National Laboratory database of sequences of, 106; sequence data for, 98–104, 108, 112, 119
- HLA alleles, 76, 77–78
- HLA class I alleles, 76
- HLA class I genes, 32
- HLA class II genes, 32
- HLA class I molecules, 32, 77, 220
- HLA class II molecules, 32, 144
- HLA genes, 32, 78, 117, 124
- HLA molecules, peptides binding to, 32, 77, 220
- HLA proteins, downregulation of, 93
- Hopfield's kinetic proofreading model, 203–11
- host-derived pMHC molecules, 47–48
- “hot spots,” 58, 66, 67, 71
- human immunodeficiency virus. *See* HIV
- human leukocyte antigens. *See* HLA *entries*
- human papillomavirus, 88
- humoral immune response, 194, 195, 196, 200, 219; modeling, 154
- humoral memory, 47
- Huseby, E. S., 58, 61, 64, 72–73
- hydrophobic amino acids, 71–73, 74–76, 77
- hydrophobicity, 70, 72, 73, 174
- ICAM-1, 83
- IgA antibodies, 43, 47
- IgD antibodies, 42, 43
- IgE antibodies, 43, 45–46
- IgG antibodies, 43, 44–45, 47
- IgG-antigen immune complexes, 44
- IgM antibodies, 42, 43, 45
- IL-2, 80; consumed by regulatory T cells, 80–81; secretion of, by T cells, 83–84
- IL-4, 36
- IL-12, 36

- ul style="list-style-type: none; padding-left: 0;">
- immature T cell development, 48, 54, 54*f*, 55
- immune cells: flowing through lymphatic vessels and blood, 22*f*; organization of, in lymph nodes, 21, 24*f*
- immune complexes, 40
- immune defense, 16, 17
- immune memory responses, 89
- immune pressure, 114–15, 117, 129–30
- immune receptor tyrosine-based activation motifs (ITAMs), 201; phosphorylation of, 201, 206, 208–9
- immune repertoire sequencing data, 27–28, 218–19
- immune system, 1, 3; early understanding of, 2–3; misregulation of, 2; modulation of, 2; and organs and cells of higher organisms, 14–16; summary of, 49–51. *See also* adaptive immune system; innate immune system
- immunity: early observations of, 3; origin of the word, 3
- immunization: and antigen affinity and potency of antibodies, 38; computer simulations of complex models of, 173–89; with haptens, 38; prime and boost, 159, 165, 169, 171–72, 180; sequential, with cross-reactive antibodies, 190–93; and “shape space” model, 182–89; and simple model of affinity maturation, 154–73; and “string model,” 173–82; computer simulations of, with variant antigens, 178–82. *See also* sequential immunization
- immunodeficiency, 90
- immunodominance hierarchies, 195
- immunodominance patterns, 219
- immunodominant epitopes, 194–95
- immunogenicity, 220
- immunogens, 90–91, 134, 149, 154; administration of first and second, 179–82; boost, 159, 172, 180; and Kullback-Leibler divergence, 159, 160–61; long-term memory of past exposure to, 196; prime, 159, 165, 169, 172, 180
- immunoglobulins, 42
- immunological memory, pathogen-specific, 1–2, 14, 47, 50
- immunological synapse, 215
- immunology: brief history of, 9–12; importance of, for human health, 1–2; integrated experimental, clinical, and theoretical/computational approaches to, 5–7; mechanistic principles of, 4–7; as relatively young science, 2–3; scientists contributing to advances in, 12
- immunophysics, 221, 222
- immunotherapy, 2, 220–21
- inactivated vaccines, 12, 87–88
- infectious diseases, 3, 10, 87–88
- inflammation/inflammatory response, 18–19, 36
- influenza, 1, 12–13; and bnAbs, 149–50, 151, 152*f*, 195; as RNA virus, 92; vaccines for, 3, 12, 89, 118, 151, 153
- influenza strains, 89, 196, 219; memory B cells and T cells and, 47; stem-specific response to, 151
- influenza virus: diversity of, 90*f*; hemagglutinin (HA) protein on viral spike of, 150, 153, 195; and neuraminidase, 150; receptor binding site of, 173; viral spike of, 150–51, 153*f*, 154
- information theory: and entropy, 137–40; and statistical mechanics, 99, 135–40
- innate immune responses, 19, 88; as critical for activating adaptive immune system, 16, 17, 19, 49, 88, 221; as not pathogen specific, 14
- innate immune system, 14, 16–19, 221; and antibody binding, 29; cell types in, 14–16, 36; functions of, 14; and other physiological systems, 221; and paucity of work using physical models, 221; and receptor binding, 16; schematic summary of, 49, 50*f*; sensing conserved elements of pathogens, 16

- integrase, 92, 93
- intensive entropy (ACE method), 102–6
- interaction free energies, 62–65, 69–70;
 - BCR/antibody and antigen, 182–83, 184*f*, 185–86; and BCR/antibody interactions with antigenic epitopes, 174–75; between strings of amino acids, 174
- interferon-gamma (IFN γ), 36, 125
- interferons (IFNs), 19
- intergenic epistasis, 98
- intracellular pathogens: antibodies binding to, 45; antibodies' memory of past exposure to, 196; mutant strains of, 149–89; TCRs' recognition of, 29–31
- intracellular signaling in lymphocytes, 200–216
- Ising model/Ising magnet, 99, 100, 109–10, 111, 113–15
- isomorphism, 109–10

- Jenner, Edward, 3, 10, 12, 87
- J** matrix, 122–23, 126–27
- junctional diversity, 27

- κ light chain BCRs, 23, 27
- killed vaccines, 12, 87–88
- killer T cells, 36
- kinetic proofreading, 204
- kinetic proofreading model, 203–11, 215;
 - and complex regulation of T cell signaling, 211–16; and equilibrium considerations/constraints, 204–10, 212
- Koch, Robert, 3, 10, 87; and vaccine development, 11, 12
- Koch's postulates, 10–11
- Kullback-Leibler divergence (KLD), 158–59, 179, 187, 189; and evolutionary trajectories of GC B cells, 166*f*, 167; and number of bnAbs/GC after first immunization, 164–65; optimal value of, 164, 173; for sequential rounds of immunization, 159–61

- λ light chain BCRs, 23, 27
- Lck enzyme, 31, 201, 206, 210, 214
- “least biased” model for prevalence of HIV strains, 99–100, 135, 140
- ligands, 20
- light chain BCRs, 23; constant and variable regions of, 23, 25*f*; variable region encoding of, 26
- light zone (LZ), 38, 39*f*, 40, 144, 145
- light zone (LZ) B cells, 40, 42, 178
- linker for activation of T cells (LAT), 201–3, 213, 216
- Listeria monocytogenes*, 69
- Los Alamos National Laboratory database of HIV polyproteins and proteins, 106
- lymph, 21
- lymphatic vessels, 21, 22*f*; flow of immune cells through, 22*f*
- lymph nodes, 21, 22; organization of immune cells in, 21, 24*f*, 37; structure of, 21, 24*f*
- lymphocyte receptors, generation of diversity of, 25–28
- lymphocytes, 21; effector functions of, 21; and first encounter with pathogens, 21; intracellular signaling in, 200–216; and pathogen-specific adaptive immune responses, 20; types of, 20
- lysis of pathogens, 45, 46*f*
- lysosomes, 17

- machine learning approaches, 4, 5, 6, 33, 34, 77, 196, 219, 220, 222
- macrophages, 15, 17, 19; activating adaptive immunity, 17; expressing toll-like receptors, 17–18; ingestion by, 46*f*; internalizing pathogens, 17, 18
- major histocompatibility complex. *See* MHC entries
- malaria, 14, 31, 89
- MAPK pathways, 203
- mast cells, 15, 16
- mature T cell repertoire, 55
- medulla, 55

- membrane-proximal T cell signaling pathway, 201–3, 213–16
- memory B cells, 39, 42, 43, 44, 49, 143, 178, 186, 194; affinity-dependent selection and expansion of, 43, 44*f*; entering GCs upon exposure to another variant antigen, 194–95; higher-affinity, 194; providing long-term protection from reinfection, 47
- memory T cells, 47, 83
- Metchnikoff, Élie, 12, 16, 18
- MHC class II–associated invariant chain, 30–31
- MHC class I molecules, 30; α chain (with 3 domains) of, 32; cell expression of, 31; β_2 -microglobulin of, 32, 33; CD8 binding to, 30, 201; chains and domains of, 32, 33*f*; interacting with TCR, 34–35; and peptide-binding groove, 32, 33*f*; predicting peptide binding, 33
- MHC class II molecules, 30, 144; expression of, 36; binding to CD4, 31, 201; chains and domains of, 33, 34–35, 35*f*; interacting with TCR, 35; and peptide-binding groove, 33, 34; predicting peptide binding with, 34; targeted to acidic vesicles and endosomes, 30–31
- MHC genes, 32
- MHC molecules: diversity of, exhibiting different binding characteristics, 31–32; sites of polymorphism in, 33
- microbes, 2, 3
- microbial pathogens, 12–14, 87–88
- microcanonical ensemble, 135–37, 138, 139
- microorganisms: and disease, 3, 10; as not spontaneously generated, 10, 11
- microscopy, 3
- microstates and entropy, 135–37, 139–40
- mitogen-activated protein kinase (MAPK) pathways, 203
- Miyazawa-Jernigan potential, 70
- monocytes, 15
- Montagu, Mary Wortley, 9
- Monte Carlo simulations, 101, 106, 107, 124, 134, 161
- mouse T cell repertoires, statistical properties of, 84–85
- mRNA vaccines, 88–89, 134
- multiple variant antigens: inducing bnAbs, 151; practical questions about, 151–53; simple AM model with, 154–73
- mutant strains of pathogens, antibody evolution targeting, 149–89. *See also* HIV mutant strains
- mutational distance, 180–81, 183, 185; and bnAb titers, 179, 180*f*; and bnAb titers after second immunization, 180, 182*f*; and distribution of breadths of antibodies, 179, 181*f*; and GC survival fraction, 179, 181*f*; and median breadth of antibodies produced upon immunization with a cocktail of variant antigens, 187, 188*f*
- mutational fidelity, 116, 117, 118–19, 131
- mutational vulnerabilities of HIV, 89, 96–119, 140; and tests against clinical data, 123–33; and tests against in vitro measurements of fitness, 121–23
- mutations: affinity-affecting, and B cells, 176; in BCR sequences, 38, 39–40, 145, 175–79; in HIV life cycle, 95–96; and probability of observing a mutational trajectory of a B cell, 146–48; to pathogenic peptides in T cell maturation, 67
- Mycobacterium tuberculosis*, 13
- naïve B cells, 42, 43, 44, 48, 159, 194
- natural killer (NK) cells, 15, 16
- Nef, 93
- negative selection, 55–58, 84; and binding free energy, 65; and characteristics of T cell repertoire, 58; and hydrophobic amino acids, 72–73, 74–75; and interaction free energies, 62, 64
- negative selection thresholds, 56–57
- neoantigens, 220–21

- neuraminidase, 150
- neutralization, 45, 46f, 155
- neutralizing antibodies, 29, 47, 149
- neutrophils, 17, 19
- NOD MHC mice, and CD4 T cells
 - hydrophobicity, 74, 75, 77
- noncoding regions, 27
- nucleocapsid, 93
- opsonization, 45, 46f
- OT-1 TCR, and OVA peptide, 56, 57
- pain, 19
- parasites/parasitic infections, 12, 13–14, 16, 30, 36
- Pasteur, Louis, 3, 10; experiments by, 10, 11;
 - and vaccine development, 11–12, 87–88
- pathogen-derived immunogenic peptides, 220
- pathogen-derived peptides, 30; and
 - probability of T cell activation, 80–82
- pathogen-derived pMHC molecules, 200, 203; T cell discrimination between
 - self- and, 203–11; T cell response to, 35, 53
- pathogenic peptides, 30, 53, 69, 82;
 - expressed on APCs, 82; interaction free energy of mature TCR with, 69–71; predicting T cell recognition, 66–67; T cell recognition of, 70–71, 71f; effect of thymic selection on T cell specificity for, 58, 64, 65
- pathogens, 1; and adaptive immune system, 14, 20–21; and antibodies, 45–47; BCRs binding to, 28–29; as causative agents of a disease, 11; classes of, 12–14; cleared by action of antibodies working with phagocytes, 45, 46f; diseases caused by and portals of entry for, 13; in lymphoid tissues, 21–22; vaccine-countering features of, 89; and innate immune system, 14; intracellular (*see* intracellular pathogens)
- pathogen-specific immunological memory, 1–2, 14, 47, 50
- pathogen's proteins: DNA encoding, 89;
 - targets in, 90; as vaccine material, 88, 89
- PD1, 221
- peptide amino acids, 62, 63–64, 67, 69
- peptide-binding groove, 32, 33, 57
- peptide epitopes, 124, 125, 127–28, 220
- peptide-MHC molecules (pMHC), 30, 51, 54, 144; antigen-derived, 177; class II complexes, 37; host-derived, 47–48; pathogen-derived, 35; potency of, in activating T cells, 56–57; and T cell recognition, 58, 68; TCR binding to/interactions with, 30, 31f, 34–35, 65–69; and T helper cells, 31, 41–42, 49–50. *See also* self-pMHC molecules
- peptide potency, 56–57
- peptides: binding to HLA molecules, 32, 77, 220; and immunogenicity, 220; randomly chosen, and probability of T cell activation, 78–82. *See also* epitopes; pathogen-derived peptides; pathogenic peptides; self-peptides
- PGT121 bnAb, 191–93
- phagocytes, 16–18, 35, 36; complement receptors on, 45; Fc receptors on, 45; interaction of antigen-pathogen complex with, 45; working with antibodies to clear pathogens, 45–46, 46f
- phagocytosis, 16–17
- phagosomes, 17, 18
- Phipps, James, 10
- phospholipase C gamma (PLC- γ), 201–2
- phosphorylation, and TCR complex, 201, 206, 208–9, 210, 213
- plasma cells, 37, 39, 42, 43–44, 49, 178, 186
- Plasmodium*, 14
- platelets, 16
- pluripotent hematopoietic stem cells, 14, 15
- Pol coding, 93
- polio vaccines, 12, 88
- polyproteins: HIV, 92
- positive selection, 55–58, 72
- postagrarian societies, 1
- Potts model, 99, 101, 106–8

- prevalence connection to fitness of HIV
 - proteins, 109–19
- prevalence landscape of HIV polyproteins, 98–109; and Adaptive Cluster Expansion (ACE) method, 101–6; and Potts model, 99, 101, 106–8
- primary lymphoid organs, 21, 22*f*
- prime immunogens, 159, 165, 169–71, 180
- probability of T cell maturation, 62–63
- proinflammatory response, 36
- protease, 92, 93
- proteasomes, 30
- Protein–Protein Interactions Thermo-dynamic (PINT) database, 148
- protein–protein interface, and probability distribution of the change in binding free energy, 148
- proteomes, 54, 56, 69, 77, 96, 117–18, 134; entropy of variation, 96–97; HIV, 76, 96
- proviruses, 92
- pus, 17
- quorum sensing by T cells, 78–82; experimental evidence, 82–85
- Raf kinase, 203
- Ras activation, 202*f*, 203, 213–14, 213*f*
- Ras-GDP, 203, 213
- Ras-GTP, 203, 213, 214*f*
- Ras guanine nucleotide-releasing proteins (RasGRPs), 203
- receptor binding site (RBS), 150, 173
- receptors, 16, 17; toll-like, 17–18. *See also* BCRs; TCRs
- receptor specificity, 20
- recombinant DNA technology, 88
- recombinant proteins, 88
- red blood cells, 31
- redness, heat and swelling, 18–19
- regulatory T cells (T_{regs}), 36, 85; consuming IL-2, 80–81, 84; density of, 84; suppressing T cell responses, 80–81
- reinfection, long-term protection from, 47
- retroviruses, 29, 92, 93
- reverse transcriptase, 93, 95
- reverse transcription mutations: HIV, 95–96, 131
- RNA viruses, 29, 92
- Sabin, Albert, 88
- Salk, Jonas, 88
- SARS-CoV-2: new strains of, 196; as RNA virus, 92; spike proteins of, 194; vaccines for, 44, 88–89, 134
- secondary GCs, 44; subdominant responses promoted in, 194–95
- secondary lymphoid organs, 21, 22*f*
- self-peptides, and effects on thymic development of the T cell repertoire, 62–77
- self-pMHC molecules: half-life of TCR bond with, 200, 203, 208, 211; and probability of avoiding autoimmune response, 80; T cell discrimination between pathogen-derived pMHC and, 203–11; and T cell recognition of pathogenic peptides, 71; and T cell specificity, 58; and TCR binding, 35, 53, 56–58, 65–69, 200, 203–11
- self-reactivity, 73
- self-tolerance: of T cell repertoire, 53, 78, 80, 82; of T cells and B cells, 47–48
- sequential immunization, 165, 195; cross-reactive antibodies production, 190–93; for optimizing bnAB production, 189; prime-boost, 171; with variant antigens, 172, 182, 190–93
- Shannon's entropy, 138–40
- shape space model (BCR-antigen interactions), 182–89; interaction free energy and, 182–83, 184*f*, 185; for studying a cocktail of variant antigens, 185–89
- single-positive T cells, 55
- smallpox, 2, 92; global eradication of, 2; vaccination for, 2, 3, 9; variolation to treat, 9–10, 87

- somatic hypermutation, 39–40
- SOS (Son of Sevenless) enzyme, 203, 213;
 - bistability due to positive feedback regulation of, 213, 214f; and Ras activation, 203, 213–14; structure of, 213f
- SOSIP crystal structure, 122–23
- spleen, 21
- spontaneous generation theory, 10, 11
- subdominant responses, 194–95
- subunit vaccines, 88
- swelling, 18, 19

- T cell activation, 203; probability of, by a randomly chosen peptide, 78–82; and quorum needed for effective response, 82–85; sensitivity of, to amino acid mutations in a stimulatory peptide, 61. *See also* activated T cells
- T cell-APC junction, 30
- T cell immunogenicity, 220
- T cell immunotherapy, 220
- T cell maturation: affinity threshold for, 56; and extreme value statistics, 62–69; and high peptide antigen specificity, 66; and hot spots, 58, 66, 67; and point mutations to pathogenic peptides, 67; and positive and negative selection threshold, 56–58, 62, 64, 66; probability of, 62–69; and interaction free energies, 56, 62–64
- T cell-mediated autoimmune diseases, 76, 78
- T cell receptors. *See* TCRs
- T cell repertoire, 53–85; and analysis of thymic selection effects in terms of extreme value statistics, 62–69; and basic biology of thymic development, 55–58; becoming peptide and MHC specific, 60, 61; collective effects enabling relative self-tolerance and pathogen sensitivity of, 78–85; in engineered mice, 84–85; and how thymic selection influences antigen specificity of T cells, 58–78; and probability of avoidance of autoimmune responses of every self-peptide–MHC molecule, 78, 80; and quorum sensing by T cells, 78–85
- T cells, 17, 20; and antigenic peptides, 58, 64, 65; antigen specificity of, 58–78; clustering around APCs, 83–84; density of, 83, 84; development processes of, 48; and discrimination between self- and pathogen-derived pMHC molecules, 203–11; first encountering pathogens in lymphoid tissues, 21–22; and interaction with B cells activated by the same pathogen, 37; and kinetic proofreading model, 203–11; quorum sensing by, 78–85; and recognition of specific peptide–MHC molecules, 68–69; as self-tolerant, 47–48; single-positive, 55; squeezing out of a large blood vessel, 21, 23f; in structures around follicles, 21; and suppression of autoimmune responses, 48; *See also* activated T cells
- T cell signaling, 200–201; biochemical steps of, 201–3, 202f, 213–14; complex regulation of, 211–16; and kinetic proofreading model, 203–11; membrane-proximal pathway for, 201–3, 213–16; and positive and negative feedback loops, 212–15
- T cell-targeted peptide epitope, 124
- TCR α chains: constant and variable regions of, 25, 28; encoding of, 28; rearrangement of, 55
- TCR β chains: and amino acids at P6 and P7 sites, 72–73; amino acid sequences of, 25; constant and variable regions of, 25, 28; encoding of, 28; gene rearrangement in, 55
- TCR-pMHC interactions, 30, 31f, 35, 201; and biochemical steps after engagement, 201–3, 202f, 213–14; and immunological synapse, 215

- TCRs, 20; antigen binding sites of, 25; binding to cognate pMHC, 200; chains of, linked by disulfide bonds, 24–25, 26*f*; cytoplasmic tail of, 201; discriminating between self- and pathogen-derived pMHC molecules, 203–11; diversity of, 28; gene encoding in, 28; interacting with pMHC complex, 30, 34–35; and interaction free energies with pMHCs, “string model” of, 70; and interaction free energy with MHC, 63, 64, 65; and kinetic proofreading model, 203–11; recognizing intracellular pathogen features, 29–31; and sequence data, 218–19; structure of, 24–25; and VDJ recombination, 28, 49, 50–51, 57
- TCR-self-pMHC interactions, 53, 56–58, 65–69, 200; and half-life of ligand binding, 200, 203, 208, 211; and signaling pathway, 201, 202*f*
- T follicular helper cells, 144, 175, 177
- Th1 and Th2 cells, 36; and cytokine differentiation and secretion, 36; and immune responses, 36
- T helper cells, 36, 177–78; and antibody production, 43; and APCs displaying pMHC molecules, 41; and B cells displaying antigen-derived pMHC molecules, 37; and class switching, 43; in the GC, 144; and GC B cells, 40–42, 144; and peptide-MHC class II molecules, 31
- thymic selection: and the antigen specificity of T cells, 58–78; and extreme value statistics, 62–69; and T cell specificity for pathogenic peptides, 58, 64, 65
- thymic stroma, 55
- thymocytes, 55; and average number of types of peptides encountered in the thymus, 64–65; development of, in FTOC, 56–57; differentiation and development of, 55; double-negative, 55; double-positive, 55; interaction of, with self-pMHC molecules, 56; maturation of, 56–58, 62–69; studies of, in FTOC, 56–57; and TCR and affinity thresholds for positive and negative selection, 56–57; and thymic stromal cells, 55
- thymus, 20, 21, 54–55, 54*f*; development of, 55–58; immature T cell development in, 48, 54, 54*f*; positive and negative selection in, 55–58
- tingible-body macrophages, 38, 42
- T lymphocytes. *See* T cells
- TNF α , 19
- toll-like receptors (TLRs), 17–18, 88
- transgenic mice experiments: and positive and negative selection thresholds, 56–57; and thymocytes expressing YAe62 or B3K506 TCR, 72–73
- tuberculosis, 13, 89
- type 1 diabetes in humans, susceptibility to, 76
- type I immune response, 36
- type II immune response, 36
- vaccination, 2–3; against pathogens that have evolved sophisticated strategies to evade human immune responses, 90–91; fundamental challenges of, 154; history of, 87–89; with multiple variants of viral spikes inducing bnAbs, 151; and selection forces to efficiently evolve bnAbs, 154; with single variant of a viral spike, 151; smallpox, 2, 3, 9, 10; strategies for, 10, 11–12
- vaccine-induced immune responses: HIV evading, 89, 90; and influenza, 118
- vaccine-induced memory response, 89
- vaccines, 1–2, 3, 50; antibodies’ role in, 12; challenges in developing, 89–91; design principles for, 90–91; eliciting neutralizing antibodies, 47, 149; killed/inactivated or attenuated, 11–12, 87–88; mRNA, 88–89; subunit, 88
- variant antigens: affinity of a BCR/antibody for a panel of, 155–68, 173–89; and

- cross-reactive antibodies, 190–93; and mean breadth of antibodies, 188–89, 188*f*; optimal cocktail of, 187, 188*f*; selection force for binding, 189; for sequential immunization, 172, 182, 190–93; and shape space model, 185–89. *See also* multiple variant antigens
- variolation, 9–10, 87
- VDJ recombination: and BCRs, 27–28, 37–38, 49; probabilistic model for, 28; and TCRs, 28, 49, 50–51, 53
- viral DNA and RNA, 16, 17, 92, 95
- viral evolution, dynamics, 128–29, 131, 132–33
- viral genomes, recombination, 129
- viral infection, and adaptive immune system, 49–51, 51*f*
- viral proteome, high and low entropy of variation of, 96–97
- viral spike: of HIV, 92, 93, 122, 149–50, 152*f*; of influenza, 150–51, 153*f*; vaccination with a single variant of, 151
- viruses, 12, 45; classes of, 92; transmission of, 92. *See also* HIV; influenza virus
- VRCO1, binding to gp120, 191, 192*f*
- World Health Assembly, and smallpox eradication, 2
- World Health Organization, and deaths from pathogens, 12, 13, 14
- Wright-Fisher dynamics, 128–29, 131
- ZAP-70 gene, 201, 210, 214
- zeta chains (TCRs), 201