Statistical Mechanical Concepts in Immunology

Arup K. Chakraborty[1,2,4,5] and Andrej Košmrlj[3,5]

Departments of [1]Chemical Engineering, [2]Chemistry, [3]Physics, and [4]Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

[5]Ragon Institute of MGH, MIT, and Harvard, Cambridge, Massachusetts 02129; email: arupc@mit.edu

Annu. Rev. Phys. Chem. 2010. 61:283–303

The *Annual Review of Physical Chemistry* is online at physchem.annualreviews.org

This article's doi: 10.1146/annurev.physchem.59.032607.093537

0066-426X/10/0505-0283$20.00

Key Words

T cell, B cell, antigen recognition, theory and computation, lymphocyte development

Abstract

Higher organisms, such as humans, have an adaptive immune system that usually enables them to successfully combat diverse (and evolving) microbial pathogens. The adaptive immune system is not preprogrammed to respond to prescribed pathogens. Yet it mounts pathogen-specific responses against diverse microbes and establishes memory of past infections (the basis of vaccination). Although major advances have been made in understanding pertinent molecular and cellular phenomena, the mechanistic principles that govern many aspects of an immune response are not known. We illustrate how complementary approaches from the physical and life sciences can help confront this challenge. Specifically, we describe work that brings together statistical mechanics and cell biology to shed light on how key molecular/cellular components of the adaptive immune system are selected to enable pathogen-specific responses. We hope these examples encourage physical chemists to work at this crossroad of disciplines where fundamental discoveries with implications for human health might be made.

INTRODUCTION

The immune system of an organism combats invading pathogens, thereby protecting the host from disease. Jawed vertebrates, such as humans, have an adaptive immune system that enables them to mount pathogen-specific immune responses (1). The importance of the adaptive immune response for human health is highlighted by the opportunistic infections that afflict individuals with compromised adaptive immune systems [e.g., those who have progressed to AIDS after infection with the human immunodeficiency virus (HIV)]. The flexible adaptive immune system can also go awry, and many diseases (e.g., multiple sclerosis and type I diabetes) are the consequence of the adaptive immune system failing to discriminate between markers of self and nonself. The suffering caused by autoimmune diseases and the need to combat diverse infectious agents have motivated a great deal of experimental research aimed toward understanding how the adaptive immune system is regulated. These efforts have led to some spectacular discoveries (2–10). Yet an understanding of the principles that govern the emergence of an immune or autoimmune response has proven to be elusive. The practical impact of this missing knowledge is highlighted by the inability to design vaccines against many scourges on the planet (e.g., HIV).

An important barrier for the development of mechanistic principles that describe adaptive immunity is that the pertinent processes involve cooperative dynamic events with many participating components that must act collectively for an immune or autoimmune response to emerge. Moreover, these processes span a spectrum of timescales and length scales that range from interactions between molecules on cells to phenomena that affect the entire organism; feedback loops between processes on different spatiotemporal scales are also important. It is often hard to intuit underlying principles from experimental observations because of the complexity of these hierarchically organized collective processes. The importance of stochastic effects further confounds intuition.

Statistical mechanics can relate observations to the underlying microscopic stochastic events that occur in a complex interacting system. Statistical mechanical theory, associated computations, and complementary experiments have therefore helped uncover mechanisms underlying complex physical and chemical phenomena. In this review, we describe recent work that brings together statistical mechanics and cell biology to uncover new concepts in immunology. This type of interdisciplinary research is beginning to shed light on some basic questions in biology with implications for human health. Our goal is to illustrate the challenges and excitement at this crossroad of the physical and life sciences, with a view toward attracting more physical chemists to this area of inquiry.

Immunology is a vast field with a wealth of interesting problems that can benefit from complementary experimental and theoretical research. Space limitations allow us to focus on essentially one topic. To properly define this topic, we begin with a brief exposition of basic immunology.

BASIC IMMUNOLOGY

Higher organisms are constantly exposed to infectious microbial pathogens. Yet the development of infectious diseases is relatively rare. This is because diverse types of cells that compose the innate immune system are efficient in preventing pathogenic microorganisms from establishing an infection. The components of the innate immune system respond to common features of diverse pathogens, but are not specific for individual pathogens. Some bacteria and many viruses can evade or overcome the innate mechanisms of host defense. The adaptive immune system mounts pathogen-specific immune responses against such invading microorganisms. Adaptive immunity also establishes memory of past infections, thereby conferring the ability to mount rapid immune responses to pathogens encountered previously.

Two Arms of the Adaptive Immune System

The adaptive immune system has two arms, called cellular and humoral immunity. T lymphocytes (T cells) and B lymphocytes (B cells) are the key regulators of cellular and humoral immunity, respectively. T cells and B cells express immunoglobulin proteins on their surfaces, which are called T cell receptors (TCRs) and B cell receptors (BCRs), respectively. The genes encoding these receptors are inherited as gene segments that stochastically recombine during the synthesis of T cells and B cells in the bone marrow. Each gene assembled in a given lymphocyte is thus distinct, enabling the generation of a great diversity of T cells and B cells expressing different receptors. Different lymphocytes can potentially respond to specific pathogens as distinct receptors can potentially recognize (i.e., bind sufficiently strongly to) molecular signatures of specific foreign invaders. Thus, the adaptive immune system can mount pathogen-specific responses against diverse infectious microbes.

TCR: T cell receptor

BCR: B cell receptor

APC: antigen-presenting cell

The diverse lymphocytes bearing different TCRs and BCRs that are generated in the bone marrow by the stochastic recombination of gene segments do not all become part of an organism's army of T cells and B cells that battle pathogens. Rather, T cells and B cells undergo development processes that allow only a small fraction of the generated cells to become part of an organism's repertoire of lymphocytes. T cells develop in an organ called the thymus (the T stands for thymus). B cells develop in the bone marrow (the B stands for bone marrow) and also, upon activation, in lymphoid organs. The primary focus of this review is on studies that are enabling an understanding of how these developmental processes shape the T and B cell repertoire such that adaptive immunity exhibits both remarkable pathogen specificity and the ability to combat myriad pathogens. In this context, we also briefly touch on the molecular machinery that translates receptor binding to pathogenic markers into cellular responses. To consider these issues, we note a few more biological facts.

What is Recognition, Where Does it Occur, and What do T Cells and B Cells Recognize?

Cells of the innate immune system (e.g., dendritic cells, macrophages) engulf pathogens (also called antigens) present in different parts of an organism's body. These cells are called antigen-presenting cells (APCs) because they express molecular signatures of the ingested antigens on their surface. Extracellular fluid from tissues, which contains pathogens or APCs harboring pathogens, drains into lymphoid organs (e.g., lymph nodes, spleen) via the lymphatic vessels. The lymphatic vessels also enable lymphocytes to circulate among the blood, lymphoid organs, and tissues. In lymphoid organs, lymphocytes can interact with pathogen-bearing APCs and pathogens and recognize them as foreign. (We define the term recognize precisely below.)

If a lymphocyte recognizes pathogens in a lymph node, a series of intracellular biochemical reactions occurs (called signaling) that results in gene transcription programs that cause the lymphocyte to become activated; i.e., it begins to proliferate and acquire the ability to carry out functions that can mediate an immune response. Activated lymphocytes thus generated, bearing receptors specific for the infecting pathogen, then leave the lymph node and enter the blood via lymphatic vessels. When they encounter the same pathogen's molecular markers in the blood or tissues, they can carry out effector functions to eliminate the infection. For example, certain kinds of activated T cells can kill cells infected by the pathogen, thereby killing the pathogen as well.

The BCRs and TCRs expressed on B cells and T cells can bind to species that are called ligands (**Figure 1**). B cells protect against pathogens in blood or extracellular spaces. The ligands of the BCR include proteins, fragments of proteins, and molecules on the surface of viruses or bacteria.

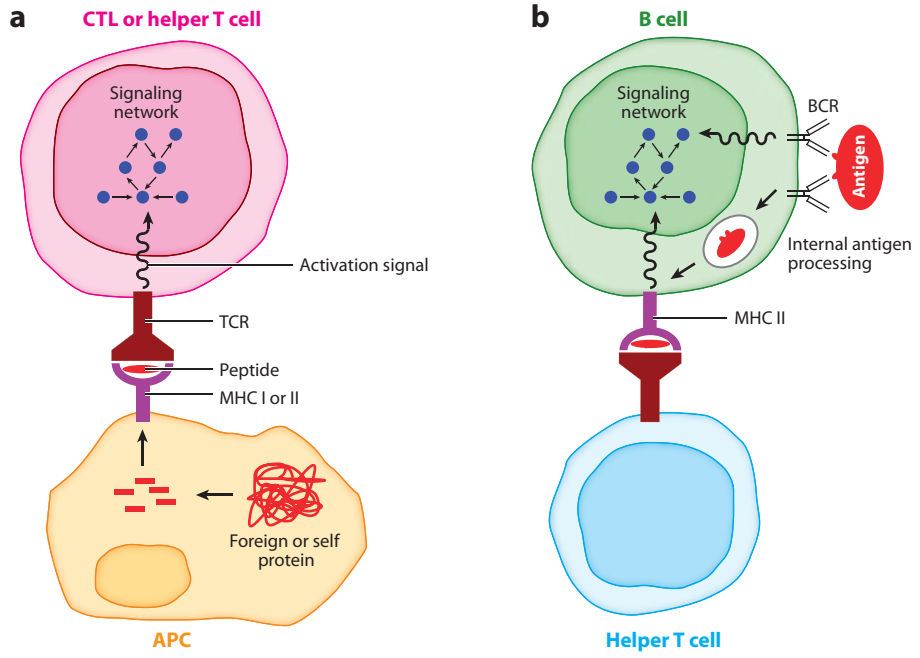


Figure 1

Lymphocyte recognition of signatures of pathogens. (*a*) T cells. Antigen-presenting cells (APCs) engulf pathogens and process their proteins into short peptides, which are bound to major histocompatibility complex (MHC) proteins and presented on the surface. T cell receptors (TCRs) bind to peptide MHCs, and sufficiently strong binding enables intracellular signaling and gene transcription, leading to T cell activation. APCs also present self-peptides derived from self-proteins, but typically T cells are not activated by them. CTL, cytotoxic T lymphocyte. (*b*) B cells. B cell receptors (BCRs) bind directly to antigens and their products, and sufficiently strong binding results in productive intracellular signaling and internalization of the antigen, bound to BCR. The antigen is processed by the B cell, which then presents the corresponding peptide MHCs on its surface. Recognition of these peptide MHCs by an activated T helper cell's TCR is usually necessary for B cell activation.

T cells evolved to combat intracellular pathogens. Proteins synthesized by intracellular pathogens are cut up into short peptide fragments by enzymes in cells harboring the pathogen. These peptide fragments can potentially bind to protein products of the host's major histocompatibility complex (MHC) genes. There are two kinds of MHC proteins, called MHC class I and MHC class II. Typically, a human will have up to six types of MHC class I proteins, and up to six types of MHC class II proteins. Pathogen-derived peptides (p) bound to MHC proteins are ultimately expressed on the surface of APCs and infected cells. These pMHCs are the TCR ligands.

MHC: major histocompatibility complex

pMHC: peptide major histocompatibility complex

CTL: cytotoxic T lymphocyte

When we say that a T cell recognizes a particular pathogen-derived pMHC, what we mean is that its TCR binds to it sufficiently strongly, which allows productive intracellular signaling and activation to occur. T cells activated by peptides presented by MHC class II proteins proliferate and differentiate into many cell types called T helper cells (for reasons noted below). T cells activated by peptides presented by MHC class I molecules become cytotoxic T lymphocytes (CTLs). When activated CTLs encounter cells in tissues that express the pMHC molecules that originally activated them, they can kill these cells by secreting various chemicals.

Similarly, when we say that a B cell recognizes a ligand, we mean that its BCR binds sufficiently strongly to it, which results in signaling and also causes the internalization of the pathogen.

pMHCs, with the peptide derived from the internalized pathogen, are presented on the B cell surface. Signaling induced by the binding of these class II pMHCs with a T helper cell's TCR (activated by the same pMHC) augments BCR signaling to activate B cells. Activated B cells proliferate and differentiate into plasma cells that secrete a soluble form of its BCR. These soluble immunoglobulins are called antibodies. Antibodies act on pathogens in extracellular spaces and in blood in a variety of different ways to help clear infections. Both T and B cells also differentiate into memory cells that mount rapid immune responses upon reinfection with the same pathogen.

MOLECULAR MECHANISM THAT ENABLES THE LYMPHOCYTE SIGNALING NETWORK TO BE ON OR OFF

The lymphocyte signaling network does not exhibit a continuous increase in response as the stimulus (e.g., TCR-pMHC binding strength) is progressively increased. Rather, it only responds strongly above a threshold stimulus level. This feature is necessary for sharp discrimination between recognized and unrecognized ligands. Readers not interested in molecular details may skip this section and take this result as a fact in the following sections that are the focus of this review.

Signaling through BCRs and TCRs is mediated by biochemical networks that are similar, but not identical. We briefly discuss one example of signaling through BCRs and TCRs that describes a phenomenon that is important when considering T cell development in the thymus. It also exemplifies how statistical mechanical computations and cell biology were brought together to elucidate the molecular machinery that causes these signaling networks to discriminate sharply between ligands they recognize and those they do not. A fuller review of such studies of lymphocyte signaling can be found elsewhere (11).

When cells are stimulated by receptor-ligand binding, many biochemical reactions that are part of a network can occur. These reactions modify proteins inside the cell. The modified proteins are referred to as activated signaling molecules. The number of activated signaling molecules is not zero in unstimulated cells, as a basal level of signaling is maintained (**Figure 2*a***). When cells are weakly stimulated (for example, by binding to very few ligands or ligands that bind weakly to receptors), basal levels of signaling are maintained. Recently, it was shown that when lymphocytes are stimulated by increasing doses or strengths of receptor-ligand binding, the population of cells does not respond by continuously increasing the number of active downstream signaling molecules (12, 13). Rather, beyond a threshold stimulus level, the population of cells splits into two subpopulations, one that turns on a large number of active signaling molecules and another that maintains basal signaling levels (**Figure 2*a***). Thus, the signaling network is said to exhibit a digital or on-off response. This is important as it allows lymphocytes to become activated only when a particular ligand is present in a sufficient amount, a feature with implications for specific recognition and robustness to noise.

Upon binding of pMHC molecules to the TCR, a kinase called Lck can bind to the cytoplasmic domains of the receptor and phosphorylate a number of residues (called ITAMs) therein (**Figure 2*b***) (14). Doubly phosphorylated ITAMs can then bind a protein called ZAP70, which is then phosphorylated by Lck (7, 15). Phosphorylated ZAP70 can then phosphorylate a number of tyrosine residues on a large protein called LAT (16). This enables the assembly of a complex containing many different proteins. One component of this complex is a protein called PLCγ, which via a series of steps can activate proteins called RasGRP (17–19). An important node in the lymphocyte signaling network (and others) is the activation of proteins called Ras from an inactive GDP-bound state to the active GTP-bound state (20). Active Ras can initiate diverse downstream signaling pathways that ultimately activate transcription factors that activate lymphocytes. RasGRP proteins are enzymes that can activate Ras proteins (21). But, in lymphocytes, there is

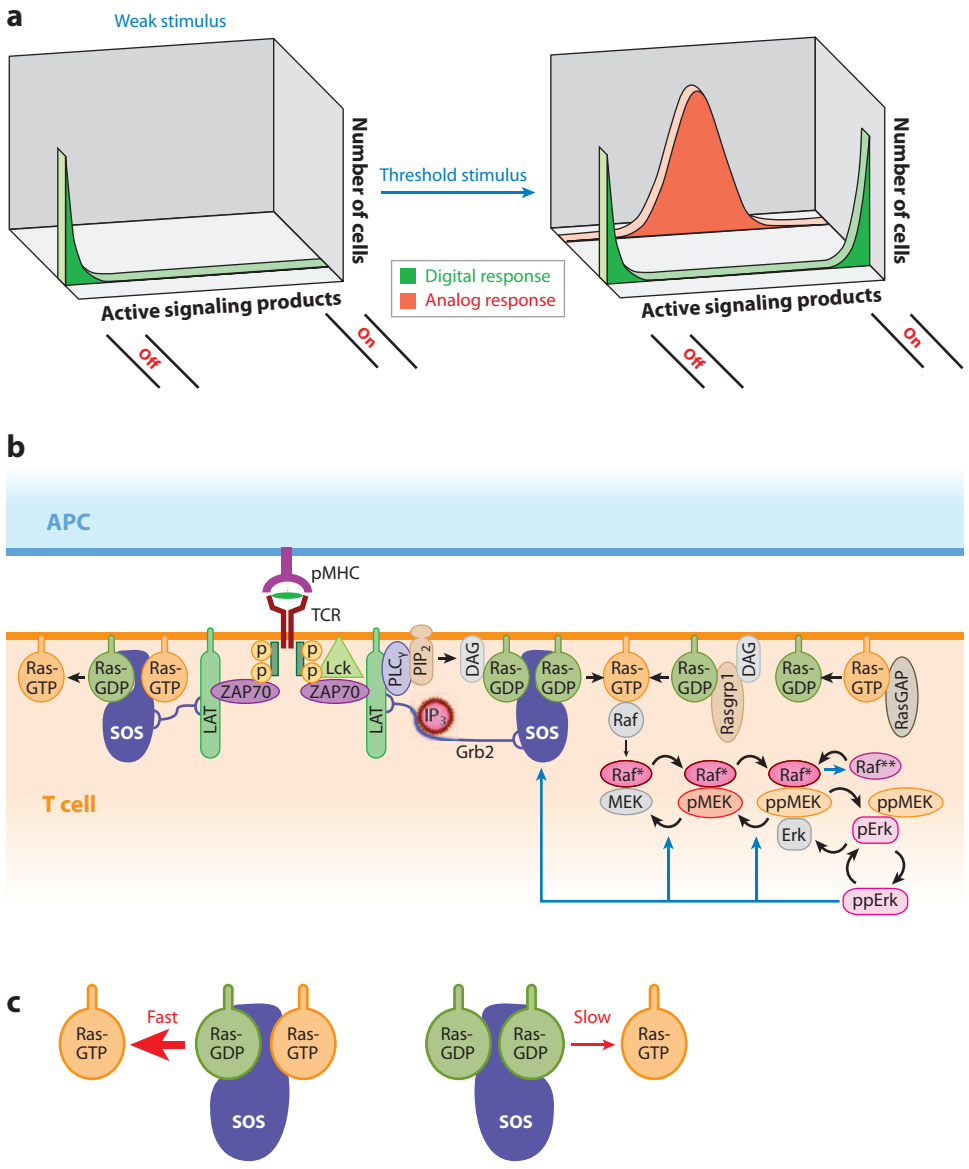


Figure 2

Digital signaling in lymphocytes. (*a*) Histograms showing the number of cells with a particular amount of an activated downstream signaling protein. For weak stimulus, a basal level of signaling characteristic of unstimulated cells is maintained (the off state). When stimulus exceeds a threshold value, some cells turn on a large amount of downstream signaling (the on state), whereas others still exhibit basal signaling. A continuous (or analog) response is shown to contrast with such a digital response. Digital responses enable sharp discrimination between the ligand type and quantity that are recognized and those that are not. (*b*) Schematic representation of the membrane-proximal T cell signaling network. (*c*) Ras activation by the enzyme SOS is subject to positive feedback regulation because the catalytic rate of conversion is much faster if the product of the catalysis (active GTP-bound Ras) is bound to an allosteric site (see 24, 25).

another protein assembled in the LAT signaling complex that can also activate Ras. This protein is called SOS (22), and crystallographic and biochemical experiments show that Ras activation by SOS is subject to positive feedback regulation (23, 24). In addition to an enzymatic site where Ras-GDP binds and is activated, SOS has another site where either GDP- or GTP-associated Ras can bind (24, 25) (**Figure 2*b*,*c***). If Ras-GTP (the product of the catalysis at the enzymatic site) is bound to this distal site, the catalytic rate of conversion increases roughly 75 times (24, 25).

Deterministic ordinary differential equations describing these processes suggested that this positive feedback loop leads to bistability; i.e., for intermediate levels of stimulation, more than one steady state is stable (13). T cells are extremely sensitive and can be activated by small numbers of ligands (26–28), and cytoplasmic signaling proteins can be present in small numbers. Therefore, intrinsic stochastic fluctuations can be important. These cell-to-cell stochastic variations in the numbers of activated signaling molecules can be obtained by solving the master equations describing the network of biochemical reactions (29). Several computer codes that can numerically solve the master equations using the Gillespie algorithm (30, 31) and Green's function methods (32) have recently been published (33–38). Two such codes (34, 38) allow the specification of the signaling network in a high-level format without the need for coding, and one of them can simulate systems that are spatially inhomogeneous extremely fast (38). The spatial organization of signaling components can be important in mediating outcome, and has recently been elucidated by a combination of computational and experimental methods rooted in physical chemistry and cell biology (39–47).

Stochastic simulations (13) of the biochemical network shown in **Figure 2*b*** predicted that Ras activation in lymphocytes is digital; i.e., above a threshold stimulus, two populations of cells should exist with greatly differing amounts of active Ras proteins (i.e., a digital response as illustrated in **Figure 2*a*** was predicted). This is a stochastic manifestation of the bistability predicted by the deterministic calculations. Many of the parameters involved in the stochastic simulations are unknown. Therefore, prior to carrying out experimental tests, it is important to establish the robustness of the qualitative phenomenon to changes in unknown parameters. The development of fast algorithms that can carry out exhaustive parameter sensitivity analyses for stochastic simulations and the proper sampling of nonequilibrium states remain a challenge. However, deterministic analyses can often be used as a guide. For example, aided by such analyses and stochastic computations, Das et al. (13) showed that the predicted digital signaling was robust as long as a few features of the network were true. This type of robustness is quite common in biological networks (48).

Experimental studies (13) showed that signaling in lymphocytes was indeed digital upon carefully changing the stimulus level. However, owing to technical reasons, only a downstream product of Ras activation could be assayed in single-cell experiments that could interrogate whether signaling is digital. Therefore, one could argue that the digital signaling observed in the experiments resulted from other feedback loops not related to SOS-mediated Ras activation. Computational models can be used to design genetic and biochemical experiments that can discriminate between such possibilities. The work reported by Das et al. (13) is an example of complementary physical and biological studies that established that the molecular origin of digital signaling in lymphocytes is feedback regulation of Ras activation. The results also showed that this feedback loop results in hysteresis, which confers lymphocytes with short-term memory of past encounters with antigen.

SELECTION OF AN ANTIGEN-SPECIFIC, YET DEGENERATE, T CELL REPERTOIRE DURING DEVELOPMENT

TCR recognition of pMHC molecules is both highly specific and degenerate. It is specific because if a TCR recognizes a pMHC molecule, most point mutations of the peptide's amino acids abrogate

recognition (49, 50). However, a given TCR can also recognize diverse peptides (51–54). This specificity-degeneracy conundrum is made vivid by dividing the world of peptides into classes, with the members of each class having sequences that are closely related. For example, peptides within a class could differ by just point mutations. A TCR can discriminate quite well between peptides within a class of closely related peptide sequences (as point mutants of the peptides it recognizes are not recognized with high probability). But, at the same time, TCRs are not so good at distinguishing between peptides with different sequences as a given TCR can recognize many peptides.

How do TCRs recognize pMHC molecules in this specific, yet degenerate, fashion? Addressing this fundamental question in biology could help in understanding the immune response to pathogens and its aberrant regulation (as in autoimmunity). Recent work suggests that processes that occur during T cell development in the thymus select TCR sequences that can simultaneously exhibit specificity and degeneracy. The mechanism for specificity of TCR-pMHC interactions suggested by these studies is somewhat distinct from Fisher's (55) lock-and-key metaphor for specificity in biology.

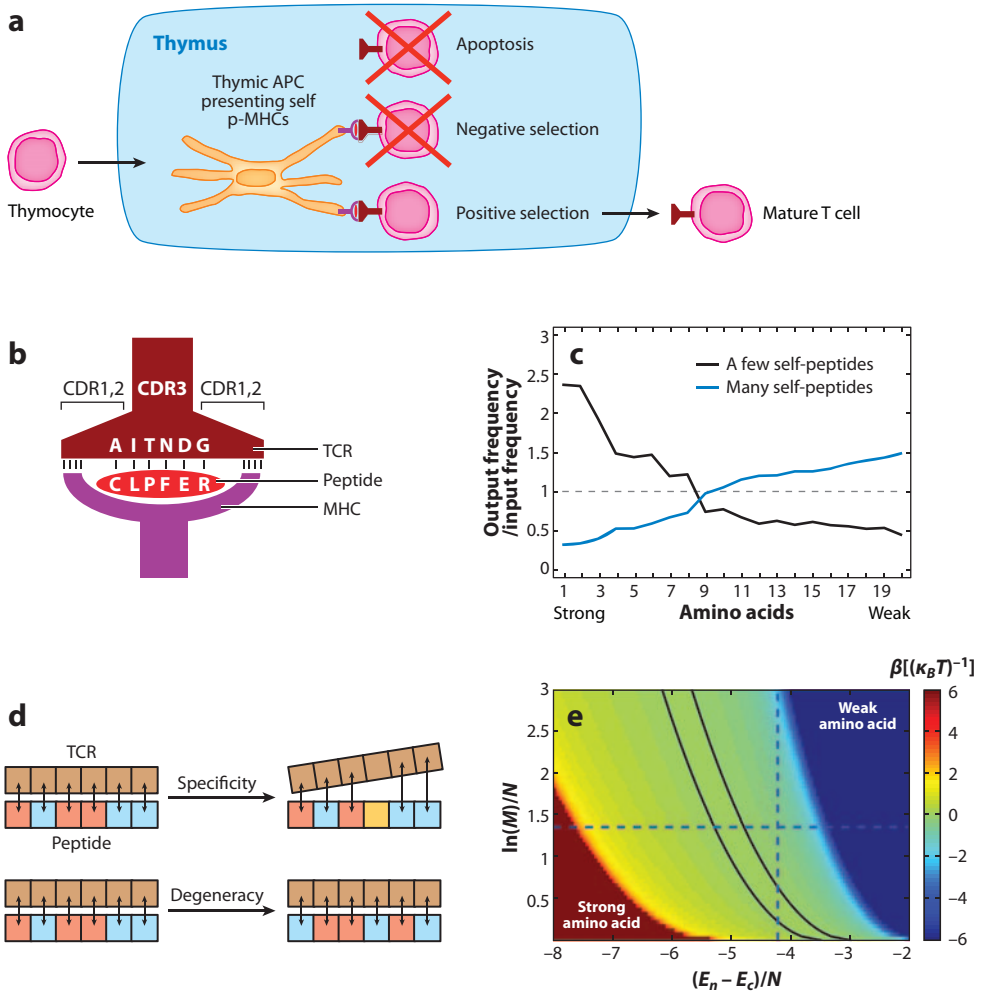
After synthesis, baby T cells (thymocytes) go to the thymus, an organ located behind the sternum (**Figure 3*a***) (10, 56–59). Cells in the thymus display pMHC molecules on their surface with the peptides derived from diverse parts of the host proteome. For a T cell to exit the thymus and become part of the host's repertoire of T cells, it must pass the following two tests: (*a*) It must not be negatively selected. That is, its TCR must not bind to any self-pMHC molecule with a binding free energy that exceeds a threshold (denoted E_n below). (*b*) It must bind at least one self-pMHC molecule with a binding free energy that exceeds another threshold (denoted E_p below). The negative selection threshold is sharply defined, whereas positive selection occurs over a range of binding strengths (a few k_BT) (60, 61). Statistical mechanical models (62, 63)

Figure 3

Thymic selection of T cells, and its consequences for the antigen-recognition properties of the T cell repertoire. (*a*) Immature T cells (thymocytes) develop in the thymus. Thymocytes migrate through the thymus and interact with diverse self peptide major histocompatibility complexes (self-pMHCs) presented on the surface of thymic antigen presenting cells (APCs). A T cell's receptor (TCR) must bind to at least one of these self-pMHCs weakly to exit the thymus and become a part of the individual's T cell repertoire (positive selection). A T cell with a TCR that binds to any self-pMHC with an affinity that exceeds a sharply defined threshold dies in the thymus (negative selection). (*b*) Schematic representation of the interface between TCR and pMHCs. The region of the TCR contacting the peptide is highly variable and is modeled by string of amino acids of length N. The peptide is also treated similarly. The binding free energy between the TCR and the entire pMHC is computed as described in the text. (*c*) The ordinate is the ratio of the frequency of occurrence of an amino acid in the peptide contact residues of selected TCRs and the preselection frequency. TCRs selected against many types of self-peptides in the thymus have peptide contact residues that are enriched in amino acids that interact weakly with other amino acids. In contrast, the peptide contact residues of TCRs selected against only a few self-peptides comprise more strongly interacting amino acids. (*d*) A mechanism for the puzzle of how TCR recognition of pathogen-derived peptides is both specific and degenerate emerges from statistical mechanical theory and is illustrated in the schematic. Peptide amino acids shown in red, or blue, are not identical; a given color represents that the amino acid is among the stronger complements of the major amino acid contact on the TCR. Sufficiently strong interactions required for recognition are mediated by several moderate interactions. (*e*) Representation of the dependency of the parameter β, a measure of amino acid composition of selected TCRs, on the number of self-peptides ln M/N and the threshold for negative selection E_n with $(E_p - E_n)/N = 0.5\ k_BT$. The region between the black lines corresponds to $\beta = 0$, to the right (left) of which negative (positive) selection is dominant, and weak (strong) amino acids are selected. The blue dashed lines indicate the relevant parameter values for thymic selection in mouse. This figure is adapted from figure 2 in Reference 62.

that complement experiments have recently taken steps toward the goal of answering how thymic selection processes may tune the T cell repertoire. Huseby et al. (50) discovered that T cells that develop in genetically altered mice that express only one type of self-pMHC in the thymus can be stimulated, with relatively high probability, by point mutants of pathogen-derived pMHCs that they recognize. In contrast, for T cells that develop in mice with many types of pMHC ($\sim 10^3$–10^4) in the thymus, most such point mutations abrogate recognition (50).

To consider this difference in recognition properties due to the number of self-pMHCs present in the thymus during T cell development, we present the following simple model (62, 63). The pMHCs are divided into two parts, as shown in **Figure 3*b***. TCRs comprise three loops. The CDR1 and CDR2 loops largely make contact with MHC amino acids, whereas the CDR3 loop contains almost all the peptide contact residues of the TCR. Therefore, the TCR is also divided into two parts (**Figure 3*b***). To assess which kinds of T cells survive thymic selection and their properties vis-à-vis recognition of pathogen-derived pMHCs, one needs a way to estimate the free energies of TCR-pMHC binding. Huseby et al. (49, 50) used inbred mice in their experiments, so the MHC proteins were all the same. Although the CDR1 and CDR2 loops vary from one TCR



to another, they are not hypervariable. Therefore, Košmrlj et al. (62, 63) represented the TCR-MHC interactions using a continuous variable (E_c), which varies from one TCR to another. The CDR3 loops are hypervariable (1) and so are the self-peptides; thus the peptide contact residues and the amino acids of the peptides were represented explicitly. Various models of the interactions between these amino acids can be envisaged. The simplest is one in which each amino acid on the peptide has a major contact residue on the TCR, and other interactions are ignored. The details of such simplifications and the potential function describing these interactions do not seem to affect qualitative results (e.g., see 62, 63, and below).

Previous studies (64–66) that represented TCR-pMHC interactions in a manner similar to **Figure 3*b*** (string models) did not have an explicit treatment of amino acids or consider the variation of the character of the T cell repertoire with the number of self-peptides in the thymus. Therefore, it was difficult to obtain mechanistic insights with experimental consequences. Košmrlj et al. (62, 63) used numerical calculations and analytical statistical mechanical theory to suggest a mechanism for how a specific, yet degenerate T cell repertoire is designed in the thymus using the model depicted in **Figure 3*b***. Their results are consistent with the reports by Huseby et al. (49, 50), and some theoretical predictions have been tested positively by experiments (63). Their conclusions also suggest new avenues for theoretical and experimental research.

An important concept emerging from the theoretical studies is that the major constraint on the kinds of sequences of TCR peptide contact residues that can emerge from the thymus is the requirement of avoiding negative selection against diverse self-pMHCs. This is because the same quenched sequence of peptide contact residues must avoid interacting strongly with diverse peptide sequences. Positive selection does not present this frustrating effect as once a TCR is positively selected by a self-pMHC, the other peptides are only relevant for negative selection.

Košmrlj et al. (62, 63) show that, with high probability, the constraints imposed by negative selection result in the elimination of TCRs with peptide contact residues that interact strongly with other amino acids or have flexible side chains (**Figure 3*c***). This is because such TCRs are likely to interact strongly with at least one of the diverse sequences of self-peptides presented in the thymus. This argument can be formalized in a number of ways, perhaps most elegantly by casting thymic selection as an extreme value distribution problem (67). A T cell bearing a particular TCR will survive only if its most attractive interaction with the diverse pMHCs presented in the thymus lies between the positive and negative selection thresholds, E_p and E_n, respectively. By treating the self-peptide sequences as independent random strings of amino acids [bioinformatic data (68, 69) suggest that this is a reasonable approximation], Košmrlj et al. (62) showed that for a TCR with a sequence of peptide contact residues, $\vec{t}$, the distribution of the strongest interaction with M self-pMHCs is sharply peaked around

$$E_0(\vec{t}) = E_c + \sum_{i=1}^{N} \varepsilon(t_i) - \sqrt{(2 \ln M) \sum_{i=1}^{N} \nu(t_i)}, \qquad (1)$$

where N is the number of TCR peptide contact residues and $\varepsilon(a)$ and $\nu(a)$ are the average and variance of the free energy of interaction of amino acid a with all others, respectively. E_0 must lie between E_n and E_p for the T cell bearing this TCR to develop into a mature T cell. From this condition and Equation 1, it is evident that increasing the number M of self-pMHCs increases the constraints presented by the requirement of avoiding negative selection. To counterbalance this pressure, for large M, TCRs are enriched with weakly interacting amino acids. A similar effect can be obtained with amino acids with smaller variance of interactions, but this effect is less pronounced because of the square root. This result is independent of the form of the statistical

potential between contacting amino acids. Different potentials only change the identities of weak and strong amino acids.

Numerical simulations of the model in **Figure 3*b*** (63) were the first to report this result (**Figure 3*c***) suggested by physical arguments and later formalized by extreme value distribution analysis (62). This prediction is supported by the analysis of available crystal structures of TCR-pMHCs (63). However, given the small number of published crystal structures, more data are required to further test this idea. If these results continue to be tested positively, taken together with experiments, they provide a mechanistic explanation for the origin of the long-standing specificity/degeneracy puzzle.

Because selected TCR sequences are predominantly composed of peptide contact residues with weakly interacting amino acids or those without flexible side chains, they recognize pathogen-derived pMHCs, with peptides that are composed of amino acids, which are among the strongest complements of the TCR peptide contact residues. This results in sufficiently strong binding (required for recognition) via many moderate interactions, a result consistent with recent experiments (49). The mechanistic origin of specificity (**Figure 3*d***) can then be summarized as follows: (*a*) Most point mutations to the amino acids of the antigenic peptide will result in weakening interactions with the TCR because peptides recognized by weakly interacting TCR peptide contact residues are enriched in amino acids that bind strongly to the TCR peptide contact residues (63). (*b*) Weakening a contact has a significant effect on the total binding free energy because each contact makes a significant contribution (49, 63). (*c*) Strong nonlinearities in the T cell signaling network, leading to T cells being either on or off (60, 61), imply that even moderate changes in binding free energy can abrogate recognition.

This mechanism for TCR-pMHC specificity is distinct from Fischer's (55) lock-and-key metaphor. Interactions between the TCR and the MHC dock the TCR over its ligand in essentially the same orientation (70, 71)—this may be analogous to shape complementarity, but it is not ligand specific. The peptide contact residues of the TCR then scan the peptide to assess if there is a sufficient number of moderate interactions to mediate recognition (**Figure 3*d***). An appropriate metaphor may be that the TCR peptide contact residues scan a bar code (peptide). The statistical aspects of TCR-pMHC recognition also make degeneracy or cross-reactivity to peptides with different sequences the flip side of the coin. Although point mutations can abrogate recognition with high probability, making a number of changes to the peptide sequence such that a sufficient number of moderate interactions is still obtained will allow recognition by the same TCR (**Figure 3*d***). This may also be why two peptides with different sequences and conformations in the MHC groove can be recognized by the same TCR.

For T cells selected against only one type of peptide in the thymus, the frustrating effects of the negative-selection constraint are not present. Thus, it is predicted that TCRs with strongly interacting or flexible peptide contact residues will emerge from the thymus with high probability (**Figure 3*c***). These amino acids will dominate the interaction free energy with recognized pMHCs; i.e., recognition will be mediated by few important contacts (49, 63), rather than many moderate interactions. T cells with such TCRs are cross-reactive to most point mutants of peptides that they recognize because only mutations that affect the few important contacts will abrogate recognition.

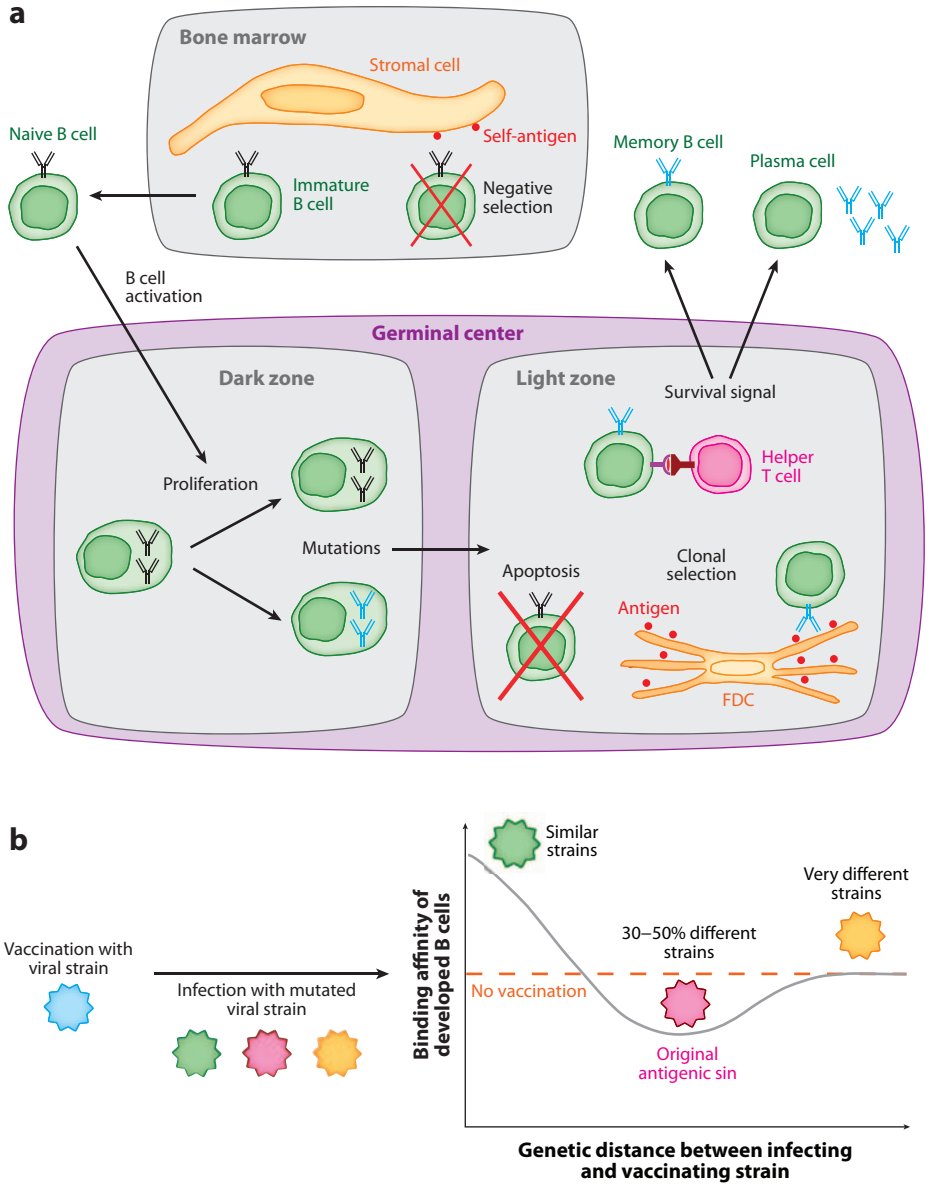
The probabilities with which amino acids with different characteristics are chosen for the peptide contact residues of the TCRs in the T cell repertoire depend on conditions (e.g., the number of peptides present in the thymus). This dependency can be formalized by using statistical mechanical methods that apply in the limit of very long peptides (62); remarkably, the results seem to hold even for short peptides. The thymic selection condition,

$$E_n < E_0(\vec{t}) < E_p, \tag{2}$$

can be interpreted as a microcanonical ensemble of TCR sequences $\vec{t}$ that are acceptable if the value of the Hamiltonian, $E_0(\vec{t})$, falls in the interval (E_n, E_p). In the limit of long peptides, the probability for TCR selection is governed by the Boltzmann weight

$$p(\vec{t}) \propto \left(\prod_{i=1}^{N} f_{t_i}\right) \exp[-\beta E_0(\vec{t})]. \quad (3)$$

Here f_a are the natural frequencies of the different amino acids prior to selection, whereas the effect of thymic selection is captured in the parameter β, which is determined by the condition that the average energy falls in the interval (E_n, E_p). The complication presented by the square



root term in Equation 1 is easily dealt with by Hamiltonian minimization (72) and introducing an effective Hamiltonian,

$$H_0(\vec{t}) = E_c + \sum_i [\varepsilon(t_i) - \gamma\nu(t_i)] - \ln M/(2\gamma). \quad (4)$$

This corresponds to Boltzmann weights

$$p(\vec{t}) \propto \prod_i \{f_{t_i} \exp[-\beta(\varepsilon(t_i) - \gamma\nu(t_i))]\}, \quad (5)$$

for which thermodynamic quantities are easily computed. γ is determined by minimizing the effective Hamiltonian $H_0(\vec{t})$ with respect to γ. β is determined by constraining the average energy to the range (E_n, E_p), while maximizing entropy. For $\beta > 0$, positive selection is dominant and stronger amino acids are selected, whereas for $\beta < 0$, negative selection is dominant and weaker amino acids are selected. The resulting phase diagram is shown in **Figure 3*e***.

The complementary theoretical (62, 63) and experimental studies (49, 50) summarized above have shed light on the specificity/degeneracy puzzle and mechanisms underlying how the antigen-recognition properties of the T cell repertoire are shaped during development. In addition to further examining these ideas, we suggest an important future direction of research. Unlike the inbred mice used in experimental models, the outbred human population has a vast diversity of MHC genes. How do particular MHC genes affect the characteristics of the T cell repertoire that an individual possesses? Answering this question might shed light on why individuals with certain MHC genes are more prone to autoimmunity or more likely to be able to control certain viral infections.

B CELL DEVELOPMENT UPON ANTIGEN RECOGNITION

Antibodies secreted by B cells were discovered much earlier than T cells. Most of the early understanding of adaptive immunity emerged from thinking about antibodies, and the first major concept in adaptive immunity, Burnet's clonal selection theory (73, 74), was tested in their context. This concept posits that a diversity of cells, each with its own antigen specificity, is synthesized, and upon interaction with infectious agents, the pathogen-specific cells (clones) proliferate. These ideas have proven to be largely correct. Remarkably, Burnet's ideas precede Tonegawa's (75) discovery of stochastic recombination of immunoglobulin genes that produce combinatorial diversity of BCRs or an understanding of B cell developmental processes.

Like T cells in the thymus, B cells undergo negative selection in the bone marrow (**Figure 4*a***); B cells expressing BCRs that recognize self-antigens are eliminated. Activation of a

Figure 4

B cell development and the original antigenic sin. (*a*) Immature B cells are synthesized in the bone marrow where they are exposed to self-antigens on the surface of stromal cells. B cells that bind too strongly to self-antigens are removed by negative selection. The rest of the B cells are released in the bloodstream as short-lived naïve B cells. Upon antigen recognition, activated naïve B cells migrate to the germinal center, where they develop better binding affinity for antigen (affinity maturation). First they enter the dark zone, where they proliferate and their receptors undergo somatic hypermutations in the variable region. Next they enter the light zone, where they die by apoptosis unless they receive two survival signals. The first signal comes from binding to antigens presented on the surface of follicular dendritic cells (FDCs). B cells with higher affinity for antigens have an advantage (clonal selection) when competing for low amounts of antigen presented on FDCs. The second signal is provided by an activated T helper cell, as shown in **Figure 1*b***. B cells that leave the germinal centers differentiate into memory cells and plasma cells. (*b*) The effect of vaccination with viral strain on subsequent infection with mutated viral strain. Depending on the genetic distance between the vaccinating and infecting strains, vaccination can be advantageous or deleterious.

B cell that emerges from the bone marrow requires its BCR to bind with moderate affinity to a ligand. These similarities notwithstanding, the development and activation of T cells and B cells occur quite differently.

GC: germinal center

BCRs recognize diverse structurally distinct ligands, unlike TCRs, which can only recognize pathogen-derived peptides presented in complex with a defined MHC protein. Naïve B cells cannot be activated just by their BCRs binding sufficiently strongly to ligands (**Figure 1*b***). The need for cooperation with T helper cells in order to be activated should provide B cells with an additional level of protection against erroneous autoimmune responses compared with T cells. Once activated, B cells undergo a round of developmental processes in the lymphoid organs wherein mutations to the BCR occur, and the mutants that bind more strongly to the infecting pathogen are selected. Activated T cells do not undergo further developmental processes after activation. Thus, it appears that characteristics that confer T cells with pathogen specificity and self-tolerance must be tuned in the thymus once, whereas BCRs on B cells can evolve these attributes after they exit the bone marrow. The implications of these differences on how the TCR and BCR repertoires are shaped, and how this affects their recognition of antigens, compose a largely unexplored area that would benefit from interdisciplinary studies.

The processes that have been described extensively by mathematical models (76–88) occur after a naïve B cell recognizes a pathogen via its BCR. These processes (**Figure 4*a***), collectively termed affinity maturation, involve a period of rapid proliferation, mutation in variable regions of BCR, selection of strongly binding mutants (clones), and differentiation into memory and plasma cells. In a few weeks, affinity maturation results in antibodies that bind to pathogens strongly and eliminate them by diverse processes.

Activated naïve B cells migrate to dynamically created areas within lymph nodes called germinal centers (GCs) (**Figure 4*a***). B cells first enter a so-called dark zone where they rapidly proliferate and their receptors undergo mutations in the genes encoding the variable regions of the BCR and their flanking regions (89, 90). These mutations are called somatic hypermutations because the rate of mutation is extraordinarily high (89). Rapid proliferation and somatic hypermutations produce some cells with receptors of higher binding affinity for antigen, but also many cells with receptors of unchanged or lower affinity. Clonal selection occurs in the light zone where the receptors on B cells compete for the limited amount of antigen presented on follicular dendritic cells (91, 92). Cells that bind to antigens will receive survival signals, whereas others die. Cells with high-affinity receptors are more likely to be successful during this competition. B cell survival is also predicated on the binding of T helper cells (93–95).

During GC processes, the affinity of antibodies for the infecting antigen are increased by at least one or two orders of magnitude (1) via a large number (5 to 10, or sometimes higher) of mutations (78, 89). Soon after high-affinity mutants appear, they take over GCs and are thus oligoclonal (90, 96, 97). Another important feature of in vivo GC reactions is an all-or-none behavior (96, 98). GCs either have no high-affinity cells at all or are dominated by them.

Although there have been many theoretical studies of these processes (76–88), the most significant early analyses are those due to Perelson and coworkers (76–78, 86, 88). For example, Kepler & Perelson (77) developed a deterministic model of B cell proliferation, mutation, and competition for antigen. B cells were divided into a discrete number of classes based on the affinity of their BCRs for the infecting pathogen. Events occurring in the GC were described with first-order differential equations that represented clonal selection (proliferation and death determined by affinity in a proportional manner) and transitions between affinity classes via a time-dependent mutation rate.

The main findings from this model can be summarized as follows. Using the Pontryagin maximum principle (99), they determined that the optimal schedule of mutations during B cell

development, which maximizes the total binding affinity for antigen, is a periodic function with mutations turning on and off. This is consistent with the fact that proliferation and high mutation rates are observed in the dark zone, and clonal selection occurs in the light zone. This periodicity suggests that some B cells may re-enter the GC after maturation, an idea referred to as the recycling hypothesis. Experimental data supporting the recycling hypothesis are still missing. However, Perelson and colleagues (78) analyzed the effect of only one pass through the GC and found that this does not produce the observed number of high-affinity B cells. Perhaps, this is why more recent studies also employ the recycling hypothesis.

Deterministic models without explicit treatment of protein sequences, exemplified by the studies of Perelson and coworkers, led to significant new insights. Yet they cannot describe the interplay between stochastic effects and the selection of certain sequences during the GC reaction. Thus, they cannot describe oligoclonality (most developed cells are descendants of very few cells) and the all-or-none behavior noted above. Two recent classes of studies aim to describe these features and their origins.

Shakhnovich and coworkers (100) developed a stochastic mesoscopic model of the humoral response, in which population dynamics of the main components (B cells, host cells, and pathogens) are determined by stability and interactions of the relevant proteins. Following their earlier work (101) on the statistical mechanics of protein folding, they modeled proteins as sequences of 27 amino acids folded into compact structures on a lattice. In this model, the stability of the native structure of the protein and the interaction strength (affinity) between two proteins are calculated in terms of Boltzmann probabilities. Sequences of pathogen-derived proteins as well as BCRs and antibodies are allowed to mutate. Only stable proteins are functional. The processes of B cell activation, differentiation, and the production of antibodies were treated by making the dynamics of the immunoglobulin proteins depend on their interaction with pathogen-derived and host cell proteins. The clonal selection of B cells in the GC was incorporated by having a faster replication rate and slower death rate for those cells that bind the pathogen-derived proteins better. For a pathogen to infect a cell, its proteins need to be stable, and bound antibodies prevent pathogens from infecting the host cell. Once the free pathogen infects a host cell, it starts replicating. Once a threshold number of pathogens is produced, it kills the host cell and free pathogens are released. Host cells also divide.

Compared with earlier work (76–88), the main new feature of this model is a qualitative genotype-phenotype relationship, because protein characteristics are determined from coarse-grained sequences of their genomes. This allows the recapitulation of the observed oligoclonality of the antibody response and the all-or-none behavior of GC reactions. The number of progenitor B cells (those at the time of infection) that seeded the population of B cells at the end of the immune response depends on the death rate of B cells in the model. The optimal death rate that results in monoclonal B cells and does not impair the healing probability of the host is $\sim$2.5 per day, which is consistent with experimental observations (102), but it does depend on parameters in the model. By computing the sequence entropy of the developed antibodies, Shakhnovich and coworkers (100) show that a monoclonal population is still quite diverse, as there are many amino acid sequences on protein surfaces that do not impair binding affinity. The distribution of mature affinity shows a bimodal (all-or-none-like) distribution, which was captured because stochastic effects were included.

Another class of models incorporates sequence information in descriptions of B cell development using a variant of ideas from spin-glass physics that were first explicated in the context of biological evolution by Kauffman and colleagues (82, 83) as the NK model. Deem and coworkers (103–107) have used NK models to study biophysical problems, including B cell development and host-pathogen dynamics. The total free energy of the BCR-antigen complex is a sum of random

free energies of interactions within a BCR subdomain, between BCR subdomains, and between BCRs and antigen. GC events are represented by several rounds of BCR mutations, and clonal selection of BCRs with antigen affinities in the top 20%.

This model was used (104) to describe a well-known phenomenon, called the original antigenic sin (108, 109). Because of high mutation rates of some viruses, the infecting viral strain is different from the strain used in a vaccine (e.g., in a flu shot). Using the NK model, a relationship between the binding affinity of the mature antibody and the infecting viral strain was calculated as a function of the genetic distance between the infecting and vaccinating strains (**Figure 4*b***). If the genetic distance between these strains is short, vaccination promotes the development of better antibodies. If the strains are quite different, vaccination is obviously irrelevant. For intermediate differences in genotypes of the strains, the developed antibodies have a lower binding affinity for the pathogen compared to if the individual had not been vaccinated. The free energy landscape of NK models is rugged. Deem and coworkers (104) describe original antigenic sin as trapping in a locally deep minimum in this landscape, which would not occur without prior vaccination. Physically, this is because of two effects. First, memory cells specific for the vaccinating strain are present in larger numbers than naïve B cells specific for the infecting strain (1). However, the memory cells bind to the infecting strain more weakly if there is a sufficiently large genetic distance between them. Therefore, existing memory cells dominate the response, which results in weaker binding mature antibodies. Furthermore, because the memory cells were the product of affinity maturation, they acquired many mutations to bind strongly to the vaccinating strain. These mutations may need to be undone to bind strongly to an infecting strain that is sufficiently different. Naïve B cells may not require the undoing of mutations.

NK models have been used to consider T cell vaccination strategies against cancer (107). In agreement with previous experiments (110), calculations suggest that it is better to vaccinate with different strains in different lymph nodes compared to vaccinating with the same strains in one lymph node. The former protocol produces memory cells with broader coverage because there is no competition for different strains.

The studies noted above, which include mesoscopic representations of protein sequences, have recapitulated well-understood phenomena and have provided new insights into the phenotype-genotype relationships during B cell development and host-pathogen dynamics. However, new experimental consequences have not been explicated. Therefore, new concepts in immunology emerging from these studies are not yet evident.

SUMMARY AND FUTURE DIRECTIONS

Understanding how an adaptive immune response emerges and how it is misregulated presents an interesting challenge in basic science. New discoveries in this regard will have important implications for human health. Uncovering the principles that determine molecular and cellular phenomena pertinent to adaptive immunity will greatly benefit by bringing together approaches from the physical and life sciences. Restricted by our scientific expertise, we focus on the interface of statistical mechanics and cell biology in this review, but experimental techniques drawn from physical chemistry can contribute greatly.

Because of space limitations, above we discuss only one topic to illustrate the crossroad of immunology and statistical mechanics, e.g., the development of the T and B cell repertoires. We present a mechanism for how T cell recognition of antigenic peptides is both specific and degenerate that has emerged from recent work. These studies suggest that the origin of specificity in T cell recognition of antigen may be distinct from the lock-and-key metaphor used to describe specificity of enzyme-substrate interactions. We also describe studies on B cell development, which

extend previous more phenomenological models to provide insights into the genotype-phenotype relationships during the evolution of activated B cells (and antibodies). Some of these ideas, if pursued further, may have implications for vaccination strategies.

Many problems in immunology at the molecular, cellular, and organism level remain unsolved and present future research opportunities. However, a key challenge just beginning to be addressed is the following. Much progress (theoretical and experimental) has been made to describe host-pathogen dynamics in humans (especially in the context of HIV). Basic molecular and cellular immunology has focused largely on experimental models that are inbred mice or cell lines. This is because these models can be readily manipulated to test ideas (such as predictions emerging from statistical mechanical theory and computation). A fundamental challenge is to bridge the gap between basic molecular discoveries and the immune response of individual humans with specific genes (e.g., people have diverse MHC genes). Much is to be gained by making the two meet, and understanding the effects of fluctuations in host genetics is an interesting statistical mechanics problem.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

A.K.C. is grateful to Profs. Mark Davis, Michael Dustin, Herman Eisen, Eric Huseby, Jeroen Roose, Andrey Shaw, and Arthur Weiss for educating him about basic immunology over the past decade. He is also grateful to the National Institutes of Health for generous research support. A.K. and A.K.C. both greatly value ongoing and past collaborations with Prof. Mehran Kardar.

LITERATURE CITED

1. Janeway C. 2005. *Immunobiology: The Immune System in Health and Disease*. New York: Garland Sci. 823 pp.
2. Allison JP, McIntyre BW, Bloch D. 1982. Tumor-specific antigen of murine T-lymphoma defined with monoclonal antibody. *J. Immunol.* 129:2293–300
3. Dialynas DP, Wilde DB, Marrack P, Pierres A, Wall KA, et al. 1983. Characterization of the murine antigenic determinant, designated L3T4a, recognized by monoclonal antibody GK1.5: Expression of L3T4a by functional T cell clones appears to correlate primarily with class II MHC antigen reactivity. *Immunol. Rev.* 74:29–56
4. Hedrick SM, Cohen DI, Nielsen EA, Davis MM. 1984. Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. *Nature* 308:149–53
5. Yanagi Y, Yoshikai Y, Leggett K, Clark SP, Aleksander I, Mak TW. 1984. A human T cell-specific cDNA clone encodes a protein having extensive homology to immunoglobulin chains. *Nature* 308:145–49
6. Unanue ER. 1984. Antigen-presenting function of the macrophage. *Annu. Rev. Immunol.* 2:395–428
7. Chan AC, Iwashima M, Turck CW, Weiss A. 1992. ZAP-70: a 70 kd protein-tyrosine kinase that associates with the TCR zeta chain. *Cell* 71:649–62
8. Irving BA, Weiss A. 1991. The cytoplasmic domain of the T cell receptor zeta chain is sufficient to couple to receptor-associated signal transduction pathways. *Cell* 64:891–901
9. Gallegos AM, Bevan MJ. 2006. Central tolerance: good but imperfect. *Immunol. Rev.* 209:290–96
10. Hogquist KA, Baldwin TA, Jameson SC. 2005. Central tolerance: learning self-control in the thymus. *Nat. Rev. Immunol.* 5:772–82

11. Chakraborty AK, Das J. 2010. Pairing computation with experimentation: a powerful coupling for understanding T cell signalling. *Nat. Rev. Immunol.* 10:59–71
12. Altan-Bonnet G, Germain RN. 2005. Modeling T cell antigen discrimination based on feedback control of digital ERK responses. *PLoS Biol.* 3:e356
13. Das J, Ho M, Zikherman J, Govern C, Yang M, et al. 2009. Digital signaling and hysteresis characterize Ras activation in lymphoid cells. *Cell* 136:337–51
14. Iwashima M, Irving BA, van Oers NS, Chan AC, Weiss A. 1994. Sequential interactions of the TCR with two distinct cytoplasmic tyrosine kinases. *Science* 263:1136–39
15. Neumeister EN, Zhu Y, Richard S, Terhorst C, Chan AC, Shaw AS. 1995. Binding of ZAP-70 to phosphorylated T-cell receptor zeta and eta enhances its autophosphorylation and generates specific binding sites for SH2 domain-containing proteins. *Mol. Cell Biol.* 15:3171–78
16. Zhang W, Sloan-Lancaster J, Kitchen J, Trible RP, Samelson LE. 1998. LAT: the ZAP-70 tyrosine kinase substrate that links T cell receptor to cellular activation. *Cell* 92:83–92
17. Houtman JC, Higashimoto Y, Dimasi N, Cho S, Yamaguchi H, et al. 2004. Binding specificity of multiprotein signaling complexes is determined by both cooperative interactions and affinity preferences. *Biochemistry* 43:4170–78
18. Lin J, Weiss A. 2001. Identification of the minimal tyrosine residues required for linker for activation of T cell function. *J. Biol. Chem.* 276:29588–95
19. Zhu M, Janssen E, Zhang W. 2003. Minimal requirement of tyrosine residues of linker for activation of T cells in TCR signaling and thymocyte development. *J. Immunol.* 170:325–33
20. Genot E, Cantrell DA. 2000. Ras regulation and function in lymphocytes. *Curr. Opin. Immunol.* 12:289–94
21. Ebinu JO, Bottorff DA, Chan EY, Stang SL, Dunn RJ, Stone JC. 1998. RasGRP, a Ras guanyl nucleotide-releasing protein with calcium- and diacylglycerol-binding motifs. *Science* 280:1082–86
22. Chardin P, Camonis JH, Gale NW, van Aelst L, Schlessinger J, et al. 1993. Human Sos1: a guanine nucleotide exchange factor for Ras that binds to GRB2. *Science* 260:1338–43
23. Margarit SM, Sondermann H, Hall BE, Nagar B, Hoelz A, et al. 2003. Structural evidence for feedback activation by Ras.GTP of the Ras-specific nucleotide exchange factor SOS. *Cell* 112:685–95
24. Sondermann H, Soisson SM, Boykevisch S, Yang SS, Bar-Sagi D, Kuriyan J. 2004. Structural analysis of autoinhibition in the Ras activator Son of sevenless. *Cell* 119:393–405
25. Freedman TS, Sondermann H, Friedland GD, Kortemme T, Bar-Sagi D, et al. 2006. A Ras-induced conformational switch in the Ras activator Son of sevenless. *Proc. Natl. Acad. Sci. USA* 103:16692–97
26. Irvine DJ, Purbhoo MA, Krogsgaard M, Davis MM. 2002. Direct observation of ligand recognition by T cells. *Nature* 419:845–49
27. Sykulev Y, Joo M, Vturina I, Tsomides TJ, Eisen HN. 1996. Evidence that a single peptide-MHC complex on a target cell can elicit a cytolytic T cell response. *Immunity* 4:565–71
28. Purbhoo MA, Irvine DJ, Huppa JB, Davis MM. 2004. T cell killing does not require the formation of a stable mature immunological synapse. *Nat. Immunol.* 5:524–30
29. van Kampen NG. 2007. *Stochastic Processes in Physics and Chemistry*. Amsterdam: Elsevier. 463 pp.
30. Gillespie DT. 1977. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81:2340–61
31. Bortz AB, Kalos MH, Lebowitz JL. 1975. New algorithm for Monte-Carlo simulation of Ising spin systems. *J. Comput. Phys.* 17:10–18
32. van Zon JS, ten Wolde PR. 2005. Green's-function reaction dynamics: a particle-based approach for simulating biochemical networks in time and space. *J. Chem. Phys.* 123:234910
33. Danos V, Laneve C. 2004. Formal molecular biology. *Theor. Comp. Sci.* 325:69–110
34. Faeder JR, Blinov ML, Hlavacek WS. 2009. Rule-based modeling of biochemical systems with BioNetGen. *Methods Mol. Biol.* 500:113–67
35. Gillespie CS, Wilkinson DJ, Proctor CJ, Shanley DP, Boys RJ, Kirkwood TB. 2006. Tools for the SBML community. *Bioinformatics* 22:628–29
36. Hattne J, Fange D, Elf J. 2005. Stochastic reaction-diffusion simulation with MesoRD. *Bioinformatics* 21:2923–24
37. Li H, Cao Y, Petzold LR, Gillespie DT. 2008. Algorithms and software for stochastic simulation of biochemical reacting systems. *Biotechnol. Prog.* 24:56–61

38. Lis M, Artyomov MN, Devadas S, Chakraborty AK. 2009. Efficient stochastic simulation of reaction-diffusion processes via direct compilation. *Bioinformatics* 25:2289–91
39. Coombs D, Kalergis AM, Nathenson SG, Wofsy C, Goldstein B. 2002. Activated TCRs remain marked for internalization after dissociation from pMHC. *Nat. Immunol.* 3:926–31
40. Cemerski S, Das J, Locasale J, Arnold P, Giurisato E, et al. 2007. The stimulatory potency of T cell antigens is influenced by the formation of the immunological synapse. *Immunity* 26:345–55
41. Cemerski S, Das J, Giurisato E, Markiewicz MA, Allen PM, et al. 2008. The balance between T cell receptor signaling and degradation at the center of the immunological synapse is determined by antigen quality. *Immunity* 29:414–22
42. Lee KH, Dinner AR, Tu C, Campi G, Raychaudhuri S, et al. 2003. The immunological synapse balances T cell receptor signaling and degradation. *Science* 302:1218–22
43. Grakoui A, Bromley SK, Sumen C, Davis MM, Shaw AS, et al. 1999. The immunological synapse: a molecular machine controlling T cell activation. *Science* 285:221–27
44. Monks CR, Freiberg BA, Kupfer H, Sciaky N, Kupfer A. 1998. Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature* 395:82–86
45. Qi SY, Groves JT, Chakraborty AK. 2001. Synaptic pattern formation during cellular recognition. *Proc. Natl. Acad. Sci. USA* 98:6548–53
46. Mossman KD, Campi G, Groves JT, Dustin ML. 2005. Altered TCR signaling from geometrically repatterned immunological synapses. *Science* 310:1191–93
47. Irvine DJ, Doh J, Huang B. 2007. Patterned surfaces as tools to study ligand recognition and synapse formation by T cells. *Curr. Opin. Immunol.* 19:463–69
48. Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP. 2007. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput. Biol.* 3:1871–78
49. Huseby ES, Crawford F, White J, Marrack P, Kappler JW. 2006. Interface-disrupting amino acids establish specificity between T cell receptors and complexes of major histocompatibility complex and peptide. *Nat. Immunol.* 7:1191–99
50. Huseby ES, White J, Crawford F, Vass T, Becker D, et al. 2005. How the T cell repertoire becomes peptide and MHC specific. *Cell* 122:247–60
51. Hemmer B, Vergelli M, Gran B, Ling N, Conlon P, et al. 1998. Predictable TCR antigen recognition based on peptide scans leads to the identification of agonist ligands with no sequence homology. *J. Immunol.* 160:3631–36
52. Kersh GJ, Allen PM. 1996. Essential flexibility in the T-cell recognition of antigen. *Nature* 380:495–98
53. Misko IS, Cross SM, Khanna R, Elliott SL, Schmidt C, et al. 1999. Crossreactive recognition of viral, self, and bacterial peptide ligands by human class I-restricted cytotoxic T lymphocyte clonotypes: implications for molecular mimicry in autoimmune disease. *Proc. Natl. Acad. Sci. USA* 96:2279–84
54. Sloan-Lancaster J, Allen PM. 1996. Altered peptide ligand-induced partial T cell activation: molecular mechanisms and role in T cell biology. *Annu. Rev. Immunol.* 14:1–27
55. Fischer E. 1894. Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.* 27:2985–93
56. Starr TK, Jameson SC, Hogquist KA. 2003. Positive and negative selection of T cells. *Annu. Rev. Immunol.* 21:139–76
57. von Boehmer H, Aifantis I, Gounari F, Azogui O, Haughn L, et al. 2003. Thymic selection revisited: How essential is it? *Immunol. Rev.* 191:62–78
58. Werlen G, Hausmann B, Naeher D, Palmer E. 2003. Signaling life and death in the thymus: Timing is everything. *Science* 299:1859–63
59. Siggs OM, Makaroff LE, Liston A. 2006. The why and how of thymocyte negative selection. *Curr. Opin. Immunol.* 18:175–83
60. Daniels MA, Teixeiro E, Gill J, Hausmann B, Roubaty D, et al. 2006. Thymic selection threshold defined by compartmentalization of Ras/MAPK signalling. *Nature* 444:724–29
61. Prasad A, Zikherman J, Das J, Roose JP, Weiss A, Chakraborty AK. 2009. Origin of the sharp boundary that discriminates positive and negative selection of thymocytes. *Proc. Natl. Acad. Sci. USA* 106:528–33
62. Košmrlj A, Chakraborty AK, Kardar M, Shakhnovich EI. 2009. Thymic selection of T-cell receptors as an extreme value problem. *Phys. Rev. Lett.* 103:068103

63. Košmrlj A, Jha AK, Huseby ES, Kardar M, Chakraborty AK. 2008. How the thymus designs antigen-specific and self-tolerant T cell receptor sequences. *Proc. Natl. Acad. Sci. USA* 105:16671–76
64. Chao DL, Davenport MP, Forrest S, Perelson AS. 2005. The effects of thymic selection on the range of T cell cross-reactivity. *Eur. J. Immunol.* 35:3452–59
65. Detours V, Mehr R, Perelson AS. 1999. A quantitative theory of affinity-driven T cell repertoire selection. *J. Theor. Biol.* 200:389–403
66. Detours V, Perelson AS. 1999. Explaining high alloreactivity as a quantitative consequence of affinity-driven thymocyte selection. *Proc. Natl. Acad. Sci. USA* 96:5153–58
67. Leadbetter MR, Lindgren G, Rootzen H. 1983. *Extremes and Related Properties of Random Sequences and Processes*. New York: Springer-Verlag. 336 pp.
68. Burroughs NJ, de Boer RJ, Kesmir C. 2004. Discriminating self from nonself with short peptides from large proteomes. *Immunogenetics* 56:311–20
69. Rao X, Costa AI, van Baarle D, Kesmir C. 2009. A comparative study of HLA binding affinity and ligand diversity: implications for generating immunodominant $CD8^+$ T cell responses. *J. Immunol.* 182:1526–32
70. Garboczi DN, Ghosh P, Utz U, Fan QR, Biddison WE, Wiley DC. 1996. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* 384:134–41
71. Garcia KC, Degano M, Stanfield RL, Brunmark A, Jackson MR, et al. 1996. An $\alpha\beta$ T cell receptor structure at 2.5 A and its orientation in the TCR-MHC complex. *Science* 274:209–19
72. Kardar M. 1983. Phase transitions in new solvable Hamiltonians by a Hamiltonian minimization. *Phys. Rev. Lett.* 51:523–26
73. Burnet FM. 1957. A modification of Jerne's theory of antibody production using the concept of clonal selection. *Aust. J. Sci.* 20:67–69
74. Talmage DW. 1957. Allergy and immunology. *Annu. Rev. Med.* 8:239–56
75. Tonegawa S. 1983. Somatic generation of antibody diversity. *Nature* 302:575–81
76. Oprea M, Perelson AS. 1997. Somatic mutation leads to efficient affinity maturation when centrocytes recycle back to centroblasts. *J. Immunol.* 158:5155–62
77. Kepler TB, Perelson AS. 1993. Somatic hypermutation in B cells: an optimal control treatment. *J. Theor. Biol.* 164:37–64
78. Oprea M, van Nimwegen E, Perelson AS. 2000. Dynamics of one-pass germinal center models: implications for affinity maturation. *Bull. Math. Biol.* 62:121–53
79. Moreira JS, Faro J. 2006. Re-evaluating the recycling hypothesis in the germinal centre. *Immunol. Cell Biol.* 84:404–10
80. Kesmir C, De Boer RJ. 2003. A spatial model of germinal center reactions: cellular adhesion based sorting of B cells results in efficient affinity maturation. *J. Theor. Biol.* 222:9–22
81. Celada F, Seiden PE. 1996. Affinity maturation and hypermutation in a simulation of the humoral immune response. *Eur. J. Immunol.* 26:1350–58
82. Kauffman S, Levin S. 1987. Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol.* 128:11–45
83. Kauffman SA, Weinberger ED. 1989. The NK model of rugged fitness landscapes and its application to maturation of the immune response. *J. Theor. Biol.* 141:211–45
84. Meyer-Hermann M. 2002. A mathematical model for the germinal center morphology and affinity maturation. *J. Theor. Biol.* 216:273–300
85. Meyer-Hermann ME, Maini PK. 2005. Cutting edge: back to "one-way" germinal centers. *J. Immunol.* 174:2489–93
86. Perelson AS, Oster GF. 1979. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. *J. Theor. Biol.* 81:645–70
87. De Boer RJ, Segel LA, Perelson AS. 1992. Pattern formation in one- and two-dimensional shape-space models of the immune system. *J. Theor. Biol.* 155:295–333
88. Kepler TB, Perelson AS. 1995. Modeling and optimization of populations subject to time-dependent mutation. *Proc. Natl. Acad. Sci. USA* 92:8219–23
89. Berek C, Milstein C. 1987. Mutation drift and repertoire shift in the maturation of the immune response. *Immunol. Rev.* 96:23–41

90. Jacob J, Kelsoe G, Rajewsky K, Weiss U. 1991. Intraclonal generation of antibody mutants in germinal centres. *Nature* 354:389–92
91. Kelsoe G. 1996. Life and death in germinal centers (redux). *Immunity* 4:107–11
92. Przylepa J, Himes C, Kelsoe G. 1998. Lymphocyte development and selection in germinal centers. *Curr. Top. Microbiol. Immunol.* 229:85–104
93. Choe J, Li L, Zhang X, Gregory CD, Choi YS. 2000. Distinct role of follicular dendritic cells and T cells in the proliferation, differentiation, and apoptosis of a centroblast cell line, L3055. *J. Immunol.* 164:56–63
94. Lindhout E, Koopman G, Pals ST, de Groot C. 1997. Triple check for antigen specificity of B cells during germinal centre reactions. *Immunol. Today* 18:573–77
95. Manser T, Tumas-Brundage KM, Casson LP, Giusti AM, Hande S, et al. 1998. The roles of antibody variable region hypermutation and selection in the development of the memory B-cell compartment. *Immunol. Rev.* 162:183–96
96. Radmacher MD, Kelsoe G, Kepler TB. 1998. Predicted and inferred waiting times for key mutations in the germinal centre reaction: evidence for stochasticity in selection. *Immunol. Cell Biol.* 76:373–81
97. Kleinstein SH, Singh JP. 2001. Toward quantitative simulation of germinal center dynamics: biological and modeling insights from experimental validation. *J. Theor. Biol.* 211:253–75
98. Berek C, Berger A, Apel M. 1991. Maturation of the immune response in germinal centers. *Cell* 67:1121–29
99. Pontryagin LS, Boltyanskii VG, Gamkrelidze RV, Mischenko EF. 1962. *The Mathematical Theory of Optimal Processes*. New York: Wiley. 360 pp.
100. Heo M, Zeldovich KB, Shakhnovich EI. Unpublished manuscript
101. Shakhnovich E. 2006. Protein folding thermodynamics and dynamics: where physics, chemistry, and biology meet. *Chem. Rev.* 106:1559–88
102. Liu YJ, Barthelemy C, de Bouteiller O, Banchereau J. 1994. The differences in survival and phenotype between centroblasts and centrocytes. *Adv. Exp. Med. Biol.* 355:213–18
103. Bogarad LD, Deem MW. 1999. A hierarchical approach to protein molecular evolution. *Proc. Natl. Acad. Sci. USA* 96:2591–95
104. Deem MW, Lee HY. 2003. Sequence space localization in the immune system response to vaccination and disease. *Phys. Rev. Lett.* 91:068101
105. Gupta V, Earl DJ, Deem MW. 2006. Quantifying influenza vaccine efficacy and antigenic distance. *Vaccine* 24:3881–88
106. Munoz ET, Deem MW. 2005. Epitope analysis for influenza vaccine design. *Vaccine* 23:1144–48
107. Yang M, Park JM, Deem MW. 2006. A theory of multi-site vaccination for cancer. *Phys. A Stat. Mech. Appl.* 366:347–64
108. Fazekas de St Groth S, Webster RG. 1966. Disquisitions on original Antigenic Sin. I. Evidence in man. *J. Exp. Med.* 124:331–45
109. Fazekas de St Groth S, Webster RG. 1966. Disquisitions on original Antigenic Sin. II. Proof in lower creatures. *J. Exp. Med.* 124:347–61
110. Schreiber H, Wu TH, Nachman J, Kast WM. 2002. Immunodominance and tumor escape. *Semin. Cancer Biol.* 12:25–31

Annual Review of
Physical Chemistry
Volume 61, 2010

Contents

On Walking in the Footprints of Giants
Marilyn E. Jacox 1

Novel Computational Methods for Nanostructure Electronic Structure Calculations
Lin-Wang Wang 19

Hyper-Raman Scattering by Molecular Vibrations
Anne Myers Kelley 41

Chemistry of Hofmeister Anions and Osmolytes
Yanjie Zhang and Paul S. Cremer 63

Tuned Range-Separated Hybrids in Density Functional Theory
Roi Baer, Ester Livshits, and Ulrike Salzner 85

Subcellular Dynamics and Protein Conformation Fluctuations Measured by Fourier Imaging Correlation Spectroscopy
Eric N. Senning and Andrew H. Marcus 111

Oxide Surface Science
Ulrike Diebold, Shao-Chun Li, and Michael Schmid 129

The Diabatic Picture of Electron Transfer, Reaction Barriers, and Molecular Dynamics
Troy Van Voorhis, Tim Kowalczyk, Benjamin Kaduk, Lee-Ping Wang, Chiao-Lun Cheng, and Qin Wu 149

Electrostatics of Strongly Charged Biological Polymers: Ion-Mediated Interactions and Self-Organization in Nucleic Acids and Proteins
Gerard C.L. Wong and Lois Pollack 171

Dynamics on the Way to Forming Glass: Bubbles in Space-Time
David Chandler and Juan P. Garrahan 191

Functional Motifs in Biochemical Reaction Networks
John J. Tyson and Béla Novák 219

Electronic Properties of Nonideal Nanotube Materials: Helical Symmetry Breaking in DNA Hybrids
Slava V. Rotkin 241

Molecular Structural Dynamics Probed by Ultrafast X-Ray Absorption Spectroscopy
Christian Bressler and Majed Chergui 263

Statistical Mechanical Concepts in Immunology
Arup K. Chakraborty and Andrej Košmrlj 283

Biological Cluster Mass Spectrometry
Nicholas Winograd and Barbara J. Garrison 305

Bio-Enabled Synthesis of Metamaterials
Christopher C. DuFort and Bogdan Dragnea 323

Superresolution Imaging using Single-Molecule Localization
George Patterson, Michael Davidson, Suliana Manley, and Jennifer Lippincott-Schwartz 345

From Artificial Atoms to Nanocrystal Molecules: Preparation and Properties of More Complex Nanostructures
Charina L. Choi and A. Paul Alivisatos 369

Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events
Weinan E and Eric Vanden-Eijnden 391

Complex Fluids: Probing Mechanical Properties of Biological Systems with Optical Tweezers
H. Daniel Ou-Yang and Ming-Tzo Wei 421

Enhanced Sampling of Nonequilibrium Steady States
Alex Dickson and Aaron R. Dinner 441

Fluctuations in Biological and Bioinspired Electron-Transfer Reactions
Spiros S. Skourtis, David H. Waldeck, and David N. Beratan 461

Indexes

Cumulative Index of Contributing Authors, Volumes 57–61 487

Cumulative Index of Chapter Titles, Volumes 57–61 490

Errata

An online log of corrections to *Annual Review of Physical Chemistry* articles may be found at http://physchem.annualreviews.org/errata.shtml