Some Design Principles for Immune System Recognition

The immune system is a complex system that learns, remembers what it has learned, and acts to protect us from a variety of pathogens. Here we address the question of how the immune system is able to recognize and learn about pathogens that can rapidly evolve and, hence, potentially change so as to avoid immune recognition.

ALAN S. PERELSON AND FREDERIK W. WIEGEL

Alan S. Perelson is head of the Theoretical Biology and Biophysics Group and a Laboratory Fellow at Los Alamos National Laboratory. He is also an external professor, member of the Science Board, and director of the Theoretical Immunology program at the Santa Fe Institute. He obtained B.S. degrees in Life Science and Electrical Engineering from MIT and a Ph.D. in Biophysics from the University of California, Berkeley. He has been on the faculties of UC Berkeley and Brown University Medical School as well as a visiting professor of physics at the Ecole Normale Superieure and the University of Paris VII. His research interests are in the application of mathematical modeling to problems in virology and immunology.

Frederik W. Wiegel is a professor of theoretical physics at Twente University, The Netherlands. His publications range from quantum mechanics to theoretical biophysics. During the last decade his research interests have also expanded into the psychology of unconscious processes.

GENERAL CONSIDERATIONS

The immune system functions to protect us from disease-causing organisms. In order to perform this function the immune system must be able to recognize a potentially large variety of pathogens. Recognition in the immune system is performed on the basis of chemistry, and thus the immune system, rather than recognizing whole organisms, recognizes foreign molecules, or antigens as they are called by immunologists. The number of possible antigens is extremely large, and the immune system has only finite resources to devote to antigen recognition. Thus, devising efficient strategies for antigen recognition has been essential in the evolution of a functional immune system.

Our goal in this article is to elucidate some design principles for an immune system that has to cope with recognizing a large number of antigens. Because pathogens are living organisms, they too evolve. Thus, the task of the immune system is to recognize pathogens even if they evolve into somewhat changed forms. The signature of evolution in a pathogen is the change in some molecules; that is, antigens can change in time due to mutation.

Here we consider two problems, one in which the antigens remain fixed during the time of the immune response, and one in which the antigens change rapidly enough for that change to occur during an immune response. Space does not permit a thorough discussion of the multitude of facts and existing theoretical considerations about the immune system. For an introduction to some of the material, we refer the reader to a recent review [1].

Immune recognition involves a type of pattern recognition, and as such the principles that we discuss have application to problems outside the realm of immunology. For example, some of the ideas that we present have found application in computer security [2] and shape spaces, which are a key concept in the article, appear also in morphometry [3] and computer vision [4].

 Vol. 4, No. 5
CCC 1076-2787/99/05029-09

DESIGN PRINCIPLES

If one assumes that the number of possible antigens is sufficiently large that to a first approximation it is infinite, then one can immediately see that the immune system needs to build antigen detectors, called *receptors* by immunologists, that are sloppy in the sense that each detector must be able to recognize multiple antigens. If each receptor were able to detect only one antigen, as in the classical lock-and-key paradigm that has been used for decades to teach about the specificity of the immune response, then the number of detectors would need to equal the number of possible antigens. No one knows the number of possible antigens. Antigens are frequently proteins or parts of proteins. Proteins are strings of amino acids, of which there are 20 possible types. A protein composed on 100 amino acids can then be made in $20^{100} \cong 10^{130}$ possible ways. The genome of a pathogen such as a bacterium contains a few thousand genes, each of which encodes a protein. There are thousands of different bacteria. Thus, it appears obvious that it will be impossible to build a specific detector for each possible foreign protein. This argument has been greatly exaggerated, since the immune system recognizes the shape of molecules and not their sequence. But still one should get the sense of the enormity of the task of recognizing all possible foreign molecules.

Thus, the first design principle is that the immune system needs to use "sloppy" detectors, each of which can, in principle, recognize many objects. If the immune system uses sloppy detectors, how can it perform highly reliable recognition?

The first attempts to develop a quantitative description of the immune system date from the early 1970s. In the following quarter of a century these quantitative studies were extended both through the use of analytical methods and numerical simulations. As a result, there now exists an extensive literature that enables us to verify the consequences of certain assumptions concerning the functioning of the immune system [1]. However, to many researchers it is now becoming clear that detailed modeling and numerical simulations have to be counterbalanced by a more global perspective that looks at the design of the immune system and tries to recognize some quantitative arguments for an efficient design. In this article we shall discuss some design arguments.

The major components of the immune system are a subclass of white blood cells, known as lymphocytes: B lymphocytes, which secrete antibody molecules, and T lymphocytes, one type that secretes molecules that regulate the B-cell response and another type that secretes molecules that can kill virally infected cells.

The aspect of the immune system that we examine here is a consequence of the fundamental immunological phenomenon that B cells responding to an antigen mutate their immunoglobulin (or antibody) genes. The proteins formed under the direction of these genes are the B cell's antigen-binding receptors when expressed on the cell surface and are also the antibodies that activated B cells secrete. The mutation of immunoglobulin genes occurs in somatic cells, lymphocytes, and not in germ cells, the sperm and eggs that transmit genetic information from one generation to another. Moreover, the rate of mutation is greatly enhanced over the usual mutation rate that is seen in other proteins. For these reasons the mutation process is called *somatic hypermutation.* By the process of somatic hypermutation and the selection of "fitter" variants, the immune system can evolve. This gives rise to a type of immune learning. Further, the population of fitter variants is retained within an animal for long periods, possibly for the entire life of the organism. This imparts to the immune system the property of memory, which forms the physiological basis underlying the observation that we usually do not get many diseases, such as measles, twice.

From a more general perspective, somatic hypermutation is an attempt of the immune system to counterbalance the mutational activity of many pathogens. Recall that humans have a generation time of 20 years or so, while pathogens such as bacteria can have generation times as fast as 20 minutes. The rapid generation times of pathogens allow them to mutate and evolve at rates that are orders of magnitude faster than human evolution. The immune system can be viewed as an autonomous defense system in which generation times are fast (e.g., 6 to 7 hours) and in which mutation rates are boosted, thus allowing the immune system to evolve on a time scale that may be comparable to that of many pathogens.

Thus, a second design principle is that rather than having a static set of detectors for antigens, the immune system mutates, or changes some of its receptors in response to changes in the environment. Consequently, some receptors might be thought of as generalists that function as made, while others, specialists, are created by mutation to be highly specific for an encountered antigen. A third design principle is to retain for some long period of time the B-cell population that can produce these mutated, highly specific antibodies. This population forms the basis of immune memory and allows the immune system to respond to pathogens that it has previously seen with greater specificity and greater efficiency.

Below, we will quantitatively examine the ability of the immune system to recognize pathogens that mutate. We present two arguments about the global design of the immune system, which take the effects of mutations into account in a quantitative way. In order to do this we shall use the concept of shape space as introduced by Perelson and Oster [5]. A concise description of shape space would essentially run as follows: Take a particular type of antibody in the immune system (i.e., a particular clone of B cells) and specify the geometrical, physical, and chemical properties of the binding site by a sequence of real numbers $s_1, s_2, s_3, \ldots, s_n$. For example, s_1, s_2, s_3 might measure length, width, and height of the binding site; s_4 might give the electrostatic charge; s_5, s_6, s_7 might give the three components of the electric dipole mo-

ment; and so on. Shape space (to be denoted by S) is the linear space of all possible sequences $s_1, s_2, \ldots s_n$. Hence, each type of antibody (or each clone of B cells) is characterized by a specific point in S, which for convenience we will call its shape. An immune system with antibodies of N different shapes (i.e., consisting of N different clones of B cells) can be represented by a cloud of N representative points in shape space.

To be more specific we ask: What is a typical number for n, the dimension of shape space? What is the order of magnitude of N, the diversity of the immune system? Most estimates for N are of the order 10^7 [6–8]. The number n was estimated in Ref. 5 on theoretical grounds to be in the range 5 to 10. Recent analyses of experimentally derived antigen-antibody binding data suggest that the dimension of shape space may, in fact, be approximately 5 [9], a remarkably small number that suggests that the lock-key relation between antibody and antigen is not as specific as one would tend to think.

Since each receptor is sloppy, we can view it as covering a portion of shape space. Another design principle, on which we shall not elaborate, is to make a sufficiently large number of receptors, each with a pseudo-random shape, so that the ensemble of receptor shapes comes close to fully covering shape space [5].

ANTIBODIES MUTATE

During an immune response many antibody-producing cells (B cells) will mutate their antigen-binding receptor; roughly speaking, there will be one mutation per cell division, as the mutation rate is approximately 10^{-3} per base pair per generation [10], and about 700 base pairs make up the variable part of an antibody molecule. As a result, a clone of identical cells (represented by a single point in shape space) will proliferate into a large number of different points in shape space. In this section we model this phenomenon by Brownian motion or "diffusion" in shape space.

In order to become more quantitative let us assume that shape space is Euclidean, that the biologically relevant part of shape space,[1] S, is an n-dimensional sphere of radius R, and that the representative points of the N different antibody shapes diffuse through the interior of this sphere, all with the same diffusion coefficient, D. The assumptions of a Euclidean shape space and that mutation leads to isotropic diffusion are both rather drastic. In some related work, the shape of a molecule has been represented in Euclidean space [11]. However, in other models that have proven useful, particularly in simulations of large immune systems, shape has been represented by a string of digits. In this case, shape space becomes a hypercube if binary digits are used, or it becomes a generalization of the hypercube if a larger digit alphabet is used [12]. Mutations that lead to changes in the amino acids comprising the antibody molecule can either lead to small changes in structure, which could adequately be represented by diffusion in shape space, or, albeit with lower frequency, lead to large changes in structure, which would appear as a type of long jump in a Euclidean shape space that could not be represented easily by diffusion. The best representation for shape space is unknown. Thus, the calculations that follow are only a first attempt at studying the effects of mutation in shape space.

Now suppose A is a "new" antigen, that is, an antigen to which no antibodies bind with an accurate fit. The meaning of the loose phrase "accurate fit" can be made more precise as follows. If (the representative point of) an antigen is situated less than distance ε_1 from (the representative point of) an antibody, we shall assume that the antigen can be efficiently bound by the antibody. In other words, around each antibody we imagine a small sphere (an ε_1 sphere), and each antigen inside any of the different ε_1 spheres will be easily recognized and, under most circumstances, rapidly annihilated. Note that here by the term "antibody" we mean the shape of its binding site, and the same for "antigen." So the fact that A is a new antigen encountered by the immune system means that A "lands" outside of all the ε_1 spheres that cover much, but not all, of shape space and that make S look a bit like Swiss cheese. The B cells that mutate and hence move in shape space are those that are "stimulated" by the antigen. We assume stimulation involves weaker recognition than efficient antigen elimination, that is, that B cells within an ε_2 sphere of the antigen are stimulated by it. In immunology the relative distances in shape space are measured in terms of the affinity of the interaction between antigen and antibody. Hence, we are assuming that affinities above some low threshold are sufficient to *stimulate* a B cell, but that higher affinity B cells are needed to efficiently *eliminate* the antigen. The higher the affinity, the more efficient the elimination up to some point. For those readers who know chemistry, the affinity of an antibody is the name given by immunologists to the equilibrium binding constant between the antibody and antigen. Antibodies can either be free or be bound to the antigen. At equilibrium the higher the affinity, the larger the fraction of antibodies bound to the antigen. Put another way, the higher the affinity, the greater the fraction of antibodies that are engaged in fighting the antigen. At some point, enough antibodies are engaged in the fight that higher affinity contributes little more to the success of the response.

When a random antigen is injected into a mouse, it tends to stimulate about 10^{-4} to 10^{-5} of all B cells [6–8]. This suggests that the volume of an ε_2 sphere is of the order of 10^{-4} to 10^{-5} of the volume of shape space. Further, the total number of B cells in a mouse is about 10^8. Thus, on average a random antigen would be expected to stimulate 10^3 to 10^4 B cells. The total number of different types of B cells is about 10^7, suggesting that, on average, there are about 10 B cells of any given type in the B cell repertoire before immune learning takes place. If there are 10 copies of a B cell with a particular receptor specificity, then, of the 10^3 to 10^4 B cells stimulated, one expects 10^2 to 10^3 different types to be stimulated. Thus, on average, any antigen should simultaneously fall into 100 to 1000 ε_2 spheres,

and it would be extremely unlikely to avoid falling into any ε_2 sphere.

We view the process of somatic hypermutation and the subsequent selection of higher affinity variants as a process that allows the immune system to convert low-affinity antibodies, which can recognize an antigen, into high-affinity antibodies, which can more efficiently remove an antigen. In a sense the immune system converts "generalist" antibodies, which are good at recognizing many potential pathogens, into "specialists," which deal with particular pathogens efficiently. These specialist B cells that secrete high-affinity antibodies are retained and form part of the memory of the immune system. Because of limited resources the immune system cannot start with enough specialists to recognize every possible pathogen with high affinity—remember the immune system needs to deal with novel pathogens, some of which have never before appeared in all of evolutionary history.

Having introduced the idea of ε spheres, we can estimate the size of D relative to ε. During an immune response, B cells that recognize the antigen, that is, that are within the ε_2 sphere, may diffuse and improve their ability to recognize the antigen. If we assume that during the time of an immune response, which we call T_R, B cells that barely recognize the antigen improve so that they are very good at recognizing the antigen, then they would have diffused a distance of nearly $\varepsilon = \varepsilon_2 - \varepsilon_1$ in time T_R. In general, the average distance, x, that a B cell diffuses in an n-dimensional Euclidean space in time t is given by

$$x^2 = 2nDt. \tag{1}$$

Thus, if a B cell diffuses a distance ε in time T_R, we have

$$D = \varepsilon^2/2nT_R. \tag{2}$$

As we know, most common infections last a week or two, and this is a reflection of the time for a typical immune response. Assuming that the immune system "learns" during this period of time, it should take on the order of 2 weeks for a B cell to diffuse the distance ε. Letting T_R = 2 weeks and assuming the dimension of shape space to be n = 5, then $D \approx 0.05/\varepsilon^2/$ week^{-1}.

Let B be the B cell receptor that is nearest to antigen A in shape space. As a result of the mutational diffusion of B there is some probability that the ε_1 sphere around B will hit A at some later time, leading to the generation of high-affinity specialist antibodies that recognize A and that can be retained in the memory of the animal to prevent recurrence of the disease. We ask for the expected time (T_0) for the generation of these high-affinity antibodies under this mechanism. A rough calculation, which is reproduced in Appendix I, gives the approximate result

$$\frac{DT_0}{\varepsilon_1^2} \cong \phi\left(\frac{d}{2\varepsilon_1}\right). \tag{3}$$

The function $\varphi(x)$ is given by

$$\phi(x) = \frac{x^n}{n(n-2)} - \frac{x^2}{2(n-2)} + \frac{1}{2n}, \tag{4}$$

and the distance d, defined by

$$d = R\left\{\frac{2\pi^{n/2}}{nN\Gamma(n/2)}\right\}^{1/n}, \tag{5}$$

with Γ denoting the gamma function, is the average center-to-center distance between ε_1 spheres. Observe that both $DT_0/{\varepsilon_1}^2$ and $d/2\varepsilon_1$ are dimensionless quantities. As $d \geq 2\varepsilon_1$ (see Appendix I), we are interested in the regime $x \geq 1$. Note that $\varphi(1) = 0$ and that φ increases very rapidly for $x > 1$ (Fig. 1).

Obviously, the design of the immune system should be such that the mean time for learning about antigens and generating high-affinity memory B cells, T_0, which is roughly the time of the immune response, T_R, should be smaller than (and preferably even small as compared to) the time the antigen needs to proliferate sufficiently to cause disease.

FIGURE 1

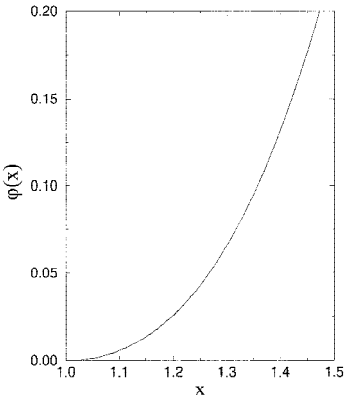

Figure 1. The function $\varphi(x)$ defined by equation (4).

Calling the latter time T_A, this criterion can be written in the form

$$T_0 \leq T_A. \tag{6}$$

For antigens, such as bacteria, that can proliferate rapidly, T_A is probably on the order of a few days. However, for other antigens, such as a tumor cell, that can potentially cause cancer, T_A may be on the order of months or years.

In the mouse, affinity maturation takes places over the course of a few weeks. Let us calculate, assuming $T_0 = T_R$ = 2 weeks, the relative volume of an ε_1 sphere that needs to be attained in order to generate high-affinity antibodies. Recall that the diffusion process underlying affinity maturation is assumed to start in an ε_2 sphere of relative volume of approximately 10^{-5}. From equation (2), with n = 5, we found that $D = 0.05\varepsilon^2 = 0.05(\varepsilon_2 - \varepsilon_1)^2$. Using this in equation (3),

$$\left(\frac{\varepsilon_2/R - \varepsilon_1/R}{\varepsilon_1/R}\right)^2 T_0 = 20\phi\left(\frac{d/R}{2\varepsilon_1/R}\right).$$

For a B-cell repertoire of size $N = 10^7$, the value estimated in mammals, and a Euclidean shape space of dimension n = 5, that is, $\mathbb{R}^5$, one finds from equation (5), d/R = 0.055. Substituting into the above equation the value 0.055 for d/R, T_0 = 2 weeks, and $\varepsilon_2/R = 10^{-1}$, the value needed so that the relative volume (ε_2/

$R)^n = 10^{-5}$, and solving numerically, we deduce that $\varepsilon_1/R = 0.01$. Finally, the relative volume of the ε_1 sphere in $\mathbb{R}^5$, $(\varepsilon_1/R)^5 = 10^{-10}$, is very small compared to the relative volume (10^{-5}) of an ε_2 sphere, and the total relative volume of all N ε_1 spheres is only 0.001. This implies that as the immune system encounters random antigens, only in 1 in 1000 cases will the response start with high-affinity antibodies, that is, those already within ε_1 of the antigen. However, after learning by somatic hypermutation has occurred, the immune system has the opportunity to retain these high-affinity cells in memory, and thus it routinely starts subsequent encounters with the same antigen with high-affinity antibodies.

ANTIGENS ALSO MUTATE!

Many antigens will mutate in order to escape immune detection (an example is HIV, the virus that causes AIDS). In view of the fact that B cells mutate in order to better recognize these antigens, one can wonder who will win this biological arms race. In this section we shall look at the design of the immune system from this point of view. A rough quantitative model should have the following features.

First, one should model what is meant by mutation of the antigen. We will assume that the shape $\vec{s}_A(t)$ of the binding site of a "new" antigen A at time t is a linear function of time

$$\vec{s}_A(t) = \vec{s}_A(0) + \vec{v}_A t, \tag{7}$$

that is, we assume that the shape of the antigen runs through n-dimensional shape space with a constant velocity $\vec{v}_A$. This assumption of a straight-line movement of the antigen's mutations might not be too realistic; one might actually ask how the antigen "knows" what a straight line is in shape space? A more realistic model would be one in which the antigen diffused, thus forming a cloud or "quasi-species" in shape space. However, the antigens that moved too close to antibodies would be quickly recognized and eliminated. Thus, the population would, in effect, move away from regions of high antibody density, and in the extreme case one might envision this as straight-line movement.

From the point of view of antigen survival, straight-line mutations are the fastest way to get away from the vicinity of some antibody, hence the Darwinian model of evolution would suggest that the antigen indeed knows what straight-line mutation is. There are interesting examples, such as the malaria parasite and trypanosomes, the parasite that causes African sleeping sickness, in which the organisms have the ability to change their surface proteins, the antigens that the immune system uses to recognize these organisms. In the case of trypanosomes, there are a set of genes that code for more than 100 different coat proteins, only one of which is expressed in large amounts on the surface of a trypanosome at a given time. The coat proteins are recognized by the immune system and hence act as antigens. Once a population of trypanosomes expressing mainly antigen A is detected by the immune system, the A variants are killed, and other variants grow out. The population of trypanosomes expressing antigen A is thus replaced by parasites expressing antigen B. When antigen B is detected, a population expressing antigen C arises and so forth. The order of appearance of variants is to a large extent genetically programmed (for a review of this interesting example see [13]), and the population appears to be escaping at constant velocity along some path. Whether the path in shape space is actually straight is not known since the shapes of proteins A, B, C, and so on, as perceived by the immune system have not been elucidated.

Second, we need an equation that describes the proliferation of antibody-producing cells and their diffusion in shape space due to antibody gene mutation. Here we restrict our attention to clones that have sensed the presence of the antigen and are stimulated into proliferation. Let $b(\vec{s},t)$ denote the density in shape space of B cells with shape $\vec{s}$ at time t that have been stimulated by the antigen. Previous modeling by Segel and Perelson [11] has considered how this density changes in a model that neglects somatic mutation. Here, this density changes as a result of three processes:

(1) Stimulated B cells proliferate at per capita rate $c(\vec{s},t)$. The function $c(\vec{s},t)$ is not known very well. While the affinity of interaction between the antigen and the B-cell receptor may play a role in the degree of activation of the B cell, to a first approximation it seems reasonable to assume all stimulated cells divide at the same constant rate c. Thus,

$$c(\vec{s}, t) = c. \tag{8}$$

(2) Stimulated B cells die, which is represented by a term $-d(\vec{s},t)\,b(\vec{s},t)$. Within germinal centers, the location in the body in which most somatic mutation occurs, dividing B cells appear to be programmed to die [14]. However, they are rescued from death in an affinity-dependent fashion [14]. Thus, cells that mutate to be closer in shape space to the antigen are preferentially spared from death. We model this by assuming

$$d(\vec{s}, t) \cong \hat{d} - f(|\vec{s} - \vec{s}_A(t)|), \tag{9}$$

where $\hat{d}$ is a constant and the function f increases when the distance $|\vec{s} - \vec{s}_A(t)|$ in shape space between the particular B cell and the antigen decreases.

(3) Stimulated B cells mutate; this leads to a term $+D\triangle b$, where D is the diffusion coefficient and $\triangle$, the Laplace operator.

Collecting these results, one obtains the equation

$$\frac{\partial b}{\partial t} = D\Delta b + \{f(|\vec{s} - \vec{s}_A|) - d_0\}b, \tag{10}$$

where $d_0 \equiv \hat{d} - c$.

Note that we have neglected a source of new B cells created from the bone marrow. Over the short times of a hunt for antigen we assume that the *de novo* generation of new B cells provides a negligible contribution to the population density in shape space.

Our procedure is now as follows (the details of the calculation can be found in Appendix II). First, one writes the solution of equation (10), subject to the proper initial and boundary conditions, as a path integral. Second, one evaluates the path integral approximately with the method of Laplace. This procedure gives the leading term in the asymptotic expansion of b in the limit $D \to 0$. It also leads to the exact result in the case in which the function $\{f(|\vec{s} - \vec{s}_A| - d_0\}$ is quadratic in the deviation from the classical trajectory. In the present case, the method can be expected to yield a fair approximation to the exact solution of equation (10). In this way, we find an approximate expression for the time development of the function $b(\vec{s},t)$. From this we calculate the number, $N_B(t)$, of B cells with receptors, the shapes of which are located inside a sphere of radius ε_1 around the point that represents the shape of the antigen, at time t:

$$N_B(t) = \frac{2\pi^{n/2}}{n\Gamma(n/2)}\,\varepsilon_1{}^n b(\vec{s}_A(t),\, t), \tag{11}$$

where the asymptotic behavior of $b(\vec{s}_A(t),t)$ is given by

$$b(\vec{s}_A(t),\, t) \equiv (constant)\, \exp\left[\left\{f(0) - d_0 - \frac{v_A^2}{4D}\right\} t\right]; \tag{12}$$

as shown in Appendix II.

When immune responses to molecular antigens, such as proteins, are studied, it is usually found that by the time high-affinity antibodies are generated the antigen has been eliminated. However, if the antigen is a pathogen that grows, and especially one that mutates, it may be very important to generate high-affinity antibodies. In the case of HIV infection, high-affinity antibodies are not generated, and the antigen is not eliminated. Thus, it is of interest to estimate the life expectancy, T_0, of a growing antigen when we assume that the antigen will be eliminated only if N_B exceeds some critical number N_0, so

$$N_B(T_0) = N_0. \tag{13}$$

If one pursues this calculation in detail (cf. equation (II.20)) one finds two regimes:

I. A regime in which the parameter values are such that

$$p_0 \equiv f(0) - d_0 > \frac{v_A^2}{4D}, \tag{14}$$

where p_0 is the net proliferation rate of the best antigen-recognizing B cell. In this case the function $N_B(t)$ increases exponentially with time for large values of t. Hence equation (12) will have a solution, and the antigen will be eliminated after some time T_0.

II. A regime of parameter values such that

$$p_0 < \frac{v_A^2}{4D}. \tag{15}$$

In this case, the function $N_B(t)$ will decrease exponentially for large values of t. Hence, unless the antigen is detected after a short time, it will escape elimination in the long run.

The rates of proliferation and death of somatically mutating B cells have been measured. The cells can divide every 6 to 7 hours [15], that is, the population can double, say, every 7 hours, implying that $c = \ln 2/7 = 0.1$ hour$^{-1} \simeq 17$ week^{-1}. If the best-fitting B cells do not die, then $p_0 = 17$ week^{-1}. How rapidly an antigen mutates will depend on the type of antigen. RNA viruses, such as HIV and influenza, mutate rapidly, whereas DNA viruses, such as polio, mutate slowly. Bacteria mutate more slowly than viruses, and multicellular parasites probably mutate their surface antigens even more slowly. With $D = 0.05\varepsilon^2$, we have $v_A^2/(4D) = 5(v_A/\varepsilon)^2$ week^{-1}. If an antigen took 1 week to move sufficiently far in shape space that an antibody that recognized it would no longer be able to do so, then the distance it moved should be more than ε, the distance separating low- and high-affinity antibodies in shape space. Let us assume that movement of a distance 2ε in a week would suffice for nonrecognition. Then, $v_A \simeq 2\varepsilon$ and $v_A^2/(4D) = 20$ week^{-1}. This is slightly greater than p_0, and we would predict that the organism would escape immune elimination. However, if it took more than a week to mutate sufficiently to be unrecognizable to the initial antibodies, then the immune system should be able to generate antibodies that keep up with the movement of the organism in shape space. Thus, rapid movement in shape space is required to escape immune elimination.

The immune system is not perfect, but even in this day and age of miracle drugs the immune system extends life by allowing us to survive many infectious diseases. Those who doubt this can look to the example of people whose immune systems fail, such as those with AIDS. In this article we have tried to elucidate some of the design principles underlying the recognition abilities of the immune system. We have argued that the immune system is designed to be able to recognize an enormously large variety of antigens. This requires the use of generalist antibodies. We then argued that by using the process of somatic evolution the immune system learns to improve its responses against pathogens and, in essence, creates specialists that can deal more effectively with particular antigenic challenges. Last, we have presented calculations that suggest that the somatic evolution of B cells also endows the immune system with the capability of defending against organisms that evolve during the course of infection, but such protection has its limits and fails when organisms evolve extremely rapidly.

ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy and the Santa Fe Institute (SFI). We thank Lee A. Segel for helpful comments. The work was supported by National Institutes of Health grant AI28433 and by a grant from the Joseph P. and Jeanne M. Sullivan Foundation. F. Wiegel thanks the Theoretical Immunology Program at the Santa Fe Institute for supporting his various visits to SFI in order to foster the collaboration that led to this work.

APPENDIX I

In order to calculate T, the mean time that passes until the new antigen A comes within the ε_1 sphere around the mutating B cell B, we note that one might as well keep B fixed and endow A with a diffusional movement, with diffusion coefficient D. In Ref. 16 one finds a discussion of the function $T(\vec{r})$, which is defined as the mean time till A—when located at position $\vec{r}$ at time 0—hits the ε_1 sphere around B for the first time. It is shown there that this function is a solution of the partial differential equation

$$\Delta T = -\frac{1}{D}, \tag{I.1}$$

where $\triangle$ denotes the Laplace operator. Because the system has spherical symmetry the solution must be a function $T(r)$ of the radial coordinate r; hence the equation becomes

$$\left(\frac{d^2}{dr^2} + \frac{n-1}{r}\frac{d}{dr}\right) T(r) = -\frac{1}{D}. \tag{I.2}$$

This is now a second-order ordinary differential equation, which needs two boundary conditions in order to have a unique solution. The first one is

$$T(\varepsilon_1) = 0, \tag{I.3}$$

and it expresses the assumption that when the antigen A is originally placed at the surface of the ε_1 sphere around B it will immediately be "killed" by B. The other boundary condition

$$\frac{dT}{dr} = 0 \text{ for } r = R \tag{I.4}$$

expresses the fact that the outer edge ($r = R$) of the whole shape space, assumed to be Euclidean, can be modeled as a hard wall that will reflect the mutational movement of the antigen. The solution of equations (I.2) to (I.4) is

$$T(r) = \frac{R^n}{n(n-2)D}\left({\varepsilon_1}^{2-n} - r^{2-n}\right) - \frac{1}{2nD}\left(r^2 - {\varepsilon_1}^2\right), \tag{I.5}$$

as can be verified by substitution. Note that if shape space is assumed to be Euclidean, then the volume of shape space is $V_n R^n$ and of an ε_1 sphere $V_n {\varepsilon_1}^n$, where

$$V_n = \frac{2\pi^{n/2}}{n\Gamma(n/2)}. \tag{I.6}$$

If there are N ε_1 spheres in shape space, their average center-to-center distance, d, follows from $Nd^n = V_n R^n$, so

$$\frac{d}{R} = \left(\frac{V_n}{N}\right)^{\frac{1}{n}}. \tag{I.7}$$

Hence, an antigen A that lands outside of all ε_1 spheres will have a distance of the order

$$\frac{d}{2} = \frac{R}{2}\left(\frac{V_n}{N}\right)^{\frac{1}{n}} \tag{I.8}$$

to the center of the nearest ε_1 sphere. Although, on average, this point will simultaneously be in 100 to 1000 overlapping ε_2 spheres, we calculate the time to reach the nearest ε_1 sphere. Its expected lifetime, T_0, can be calculated from equation (I.5) for $r = d/2$ and $R = d/2$. We take $R = d/2$ because when this antigen diffuses to a distance larger than $d/2$ from the center of the original ε_1 sphere, another ε_1 sphere becomes its nearest ε_1 sphere. One finds in this way

$$DT_0 \cong \frac{{\varepsilon_1}^2}{n(n-2)}\left(\frac{d}{2\varepsilon_1}\right)^n + \frac{{\varepsilon_1}^2}{2n} - \frac{1}{2(n-2)}\left(\frac{d}{2}\right)^2, \tag{I.9}$$

which gives the approximate results of equations (3) to (5).

APPENDIX II

In this appendix we outline the approximate solution of equation (10) under the initial condition

$$b(\vec{s}, 0) = \delta(\vec{s} - \vec{s_B}), \tag{II.1}$$

where $\vec{s_B}$ denotes the position in shape space of that B cell that is nearest to A at time 0. The exact solution can be represented by a path integral of the form [17]

$$b(\vec{s}, t) = \int_{\vec{s_B},0}^{\vec{s},t} \exp\left[-\frac{1}{4D}\int_0^t \left(\frac{d\vec{s}}{d\tau}\right)^2 d\tau + \int_0^t \{f(|\vec{s} - \vec{s_A}(\tau)|) - d_0\}\, d\tau\right] d[\vec{s}(\tau)]. \tag{II.2}$$

An approximate evaluation of the path integral proceeds as follows. We write the integrand in the form

$$\exp\left[\int_0^t L d\tau\right], \text{ where } L\{\vec{s}(\tau)\} = -\frac{1}{4D}\left(\frac{d\vec{s}}{d\tau}\right)^2 + f(|\vec{s} - \vec{s_A}(\tau)|) - d_0. \tag{II.3}$$

A rough approximation is to put

$$b(\vec{s}, t) \cong C \exp\left[\int_0^t L\{\vec{s^*}(\tau)\} d\tau\right], \tag{II.4}$$

where $\vec{s^*}(\tau)$ is the path through shape space that makes the integral $\int_0^t L\{\vec{s}(\tau)\}\, d\tau$ as large as possible. This path must pass through the initial and final positions, so $\vec{s^*}(0) = \vec{s_B}$, $s^*(t) = \vec{s}$. Moreover, C is a normalization constant that has to be determined in a separate way; in this article we shall not need its precise value.

The optimal path is a solution of the Euler-Lagrange equations. Using the fact that

$$\frac{\partial L}{\partial s_i}=\frac{\partial f}{\partial s_i};\quad \frac{\partial L}{\partial \dot{s}_i}=-\frac{\dot{s}_i}{2D};\quad \frac{d}{d\tau}\frac{\partial L}{\partial \dot{s}_i}=-\frac{\ddot{s}_i}{2D} \tag{II.5}$$

one finds

$$\frac{1}{2D}\frac{d^2 s_i}{d\tau^2}=-\frac{\partial f}{\partial s_i}. \tag{II.6}$$

In these equations, $i = 1, 2, \ldots .n$, corresponding to the different directions in shape space; the dot denotes the derivative with respect to τ, and the star on s_i^* has been omitted because confusion is unlikely. The Euler-Lagrange equations (II.6) are formally identical with the equations of motion in shape space of a classical particle with a mass equal to $1/2D$, moving in an external force field with a potential equal to $f(|\vec{s} - \vec{s}_A(\tau)|)$. The mechanical analogue will turn out to be very useful, so we shall use the corresponding terminology for a while.

We want to find the classical trajectory along which the particle moves from position $\vec{s}_B$ at time $\tau = 0$ to position $\vec{s}_A = \vec{s}_A(0) + \vec{v}_A t$ at time $\tau = t$. While the particle approaches its final position it has to climb up the potential hill $f(|\vec{s} - \vec{s}_A(\tau)|)$, which is centered at the moving point $\vec{s}_A(\tau)$. It will slow down during this process, losing kinetic energy while gaining potential energy. Hence, the early part of the movement from $\vec{s}_B$ to $\vec{s}_A(t)$ will be relatively fast, and the final part will be relatively slow. The initial velocity has to be picked in such a way that the whole trajectory is completed in time t. As the potential $f(|\vec{s} - \vec{s}_A(\tau)|)$ is repulsive [c.f. the comments after eq. (8)] the movement of the particle will be such that its velocity at time τ will be directed toward some point on the line that connects the position $\vec{s}_A(\tau)$ of the antigen at time τ and the aimed-for position $\vec{s}_A(t)$. Obviously, the whole classical trajectory $\vec{s}(\tau)$ will be located in a two-dimensional subspace of shape space: the plane through the line $\vec{s}_A(0) + \vec{v}_A t$ and through $\vec{s}_B$.

The mechanical problem can be simplified even further by describing it in a system of coordinates that moves with a constant velocity $\vec{v}_A$ with respect to the original coordinates. This means that one writes

$$\vec{s}(\tau) = \vec{s}_A(0) + \vec{v}_A\tau + \vec{\sigma}(\tau). \tag{II.7}$$

The equation of motion (II.6) now becomes

$$\frac{1}{2D}\frac{d^2\sigma_i}{d\tau^2}=-\frac{\partial f}{\partial \sigma_i}\,(i=1,2) \tag{II.8}$$

where f is a function $f(\sigma)$ of the length σ of the vector $\vec{\sigma}$. The solution of the last equation should pass through $\vec{\sigma}(0) = \vec{s}(0) - \vec{s}_A(0) = \vec{s}_B - \vec{s}_A(0)$ at $\tau = 0$ and $\vec{\sigma}(t) = \vec{s}(t) - \vec{s}_A(0) - \vec{v}_A t = 0$ at $\tau = t$. Because of the spherical symmetry of $f(\sigma)$ the trajectory $\vec{\sigma}(\tau)$ moves radially inwards, along a straight line, from the original value $\sigma = \sigma_0 \equiv |\vec{s}_B - \vec{s}_A(0)|$ to the origin. This one-dimensional movement has a conserved energy E, which is given by

$$E=\frac{1}{4D}\left(\frac{d\sigma}{d\tau}\right)^2+f(\sigma). \tag{II.9}$$

The velocity is

$$\frac{d\sigma}{d\tau}=-\sqrt{4D(E-f)}, \tag{II.10}$$

and, hence, the time to move from σ_0 to 0 is

$$t=\int_0^{\sigma 0}\frac{d\sigma}{\sqrt{4D(E-f)}}. \tag{II.11}$$

The last equation fixes the value of E, given t and σ_0. For a fixed value of σ_0 and long times t, the value of E will be slightly larger than $f(0)$; the limit $t \to \infty$ corresponds to the limit $E \to f(0)$.

We now calculate the value of the time integral of L for the most likely path, which was denoted by $\vec{s}^*(\tau)$ in equation (II.4) and by $\vec{s}(\tau)$ in all subsequent equations. Substitution of equation (II.7) into equation (II.3) gives

$$L\{\vec{\sigma}(\tau)\}=-\frac{1}{4D}\left\{\left(\frac{d\vec{\sigma}}{d\tau}\right)^2+2\vec{v}_A\cdot\frac{d\vec{\sigma}}{d\tau}+\vec{v}_A^{\,2}\right\}+f(\sigma)-d_0. \tag{II.12}$$

Now integrate this over τ from 0 to t. There are two constant terms on the right hand side; upon integration they give a term

$$\int_0^t\left(-d_0-\frac{v_A^2}{4D}\right)d\tau=-\left(d_0+\frac{v_A^2}{4D}\right)t. \tag{II.13}$$

The second term gives upon integration

$$\begin{aligned}-\frac{1}{2D}\int_0^t\vec{v}_A\cdot\frac{d\vec{\sigma}}{d\tau}\,d\tau&=-\frac{1}{2D}\vec{v}_A\cdot\{\vec{\sigma}(t)-\vec{\sigma}(0)\}\\&=-\frac{1}{2D}\vec{v}_A\cdot\{\vec{s}(t)-\vec{v}_A t-\vec{s}_B\}.\end{aligned} \tag{II.14}$$

Note that we are especially interested in the case where $\vec{s}(t) = \vec{s}_A(0) + \vec{v}_A t$, so for this specific case the previous equation simplifies to

$$-\frac{1}{2D}\int\vec{v}_A\cdot\frac{d\vec{\sigma}}{d\tau}\,d\tau=\frac{1}{2D}\vec{v}_A\cdot\{\vec{s}_B-\vec{s}_A(0)\}. \tag{II.15}$$

The two remaining terms give, with equation (II.9)

$$\int_0^t\left\{f(\sigma)-\frac{1}{4D}\left(\frac{d\sigma}{d\tau}\right)^2\right\}d\tau=\int_0^t\{2f(\sigma)-E\}d\tau. \tag{II.16}$$

Using equation (II.10) this can also be written as

$$\begin{aligned}\int_0^{\sigma 0}(2f-E)\frac{d\sigma}{\sqrt{4D(E-f)}}&=\frac{E}{\sqrt{4D}}\int_0^{\sigma 0}\frac{d\sigma}{\sqrt{E-f}}\\&\quad-\frac{1}{\sqrt{D}}\int_0^{\sigma 0}\sqrt{E-f}\,d\sigma.\end{aligned} \tag{II.17}$$

Recalling equation (II.11), the first term on the right-hand side is seen to equal Et. In the second term one may put $E \cong f(0)$, provided t is large, so one finds

$$\int_0^t\left\{f(\sigma)-\frac{1}{4D}\left(\frac{d\sigma}{d\tau}\right)^2\right\}d\tau\cong f(0)t-\frac{1}{\sqrt{D}}\int_0^{\sigma 0}\sqrt{f(0)-f(\sigma)}\,d\sigma,\ (t\to\infty). \tag{II.18}$$

Collecting the results [equations (II.13), (II.15), and (II.18), one finds the asymptotic formula

$$\int_0^t L\{s^*(\tau)\}d\tau \cong \left\{f(0) - d_0 - \frac{v_A^2}{4D}\right\}t + \frac{\vec{v_A}}{2D}\cdot\{\vec{s_B} - \vec{s_A}(0)\} - \frac{1}{\sqrt{D}}\int_0^{\sigma 0} \sqrt{f(0) - f(\sigma)}d\sigma, \ (t \to \infty). \tag{II.19}$$

For finite time it is straightforward to give the exact expression by integration of equation (II.11) once $f(\sigma)$ is known explicitly.

Substituting equation (II.19) into equation (II.4) one concludes that the $t \to \infty$ behavior of $b(\vec{s_A}(t),t)$ is given by

$$b(\vec{s_A}(t),\ t) \cong (\text{constant}) \exp\left[\left\{f(0) - d_0 - \frac{v_A^2}{4D}\right\}t\right]; \tag{II.20}$$

the precise value of the constant

$$(\text{constant}) \cong C \exp\left[\frac{\vec{v_A}}{2D}\cdot\{\vec{s_B} - \vec{s_A}(0)\} - \frac{1}{4D}\int_0^{\sigma 0} \sqrt{f(0) - f(\sigma)}d\sigma\right] \tag{II.21}$$

is only of quantitative significance because the expression (II.20) is dominated by the exponential factor.

NOTES

1. Antibodies, which are molecules, can recognize only small objects. Thus, sizes on the order of 1 meter would be outside the biologically relevant part of shape space.

REFERENCES

1. Perelson, A.S.; Weisbuch G. Rev Mod Phys 1997, 69, 1219–1267.
2. Forrest, S.; Perelson, A.S.; Allen, L.; Cherukuri, R. Self-Nonself Discrimination in a Computer. Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy. IEEE Computer Society Press: Los Alamitos, CA, 1994; 202–212.
3. Bookstein, F. Bull Math Biol 1996, 58, 313–365.
4. Edelman, S. Behav Brain Sci 1998, 21, 449–498.
5. Perelson, A.S.; Oster, G.F. J Theoret Biol 1979, 81, 645–670.
6. Klinman, N.R.; Press, J.L.; Sigal, N.H.; Gerhart, P.J. The Acquisition of the B Cell Specificity Repertoire: The Germ-Line Theory of Predetermined Permutation of Genetic Information, in The Generation of Antibody Diversity. Cunningham, A.J., ed. Academic Press: New York, 1976; 127–150.
7. Klinman, N.R.; Sigal, N.H.; Metcalf, E.S.; Gerhart, P.J.; Pierce, S.K. Cold Spring Harbor Symp Quant Biol 1977, 41, 165–173.
8. Köhler, G. Eur J Immunol 1976, 6, 340–347.
9. Lapedes, A.; Farber, R. The Geometry of Immunological Shape Space. (submitted for publication).
10. McKean, D.; Huppi, K.; Bell, M.; Staudt, L.; Gerhard, W.; Weigert, M. Proc Natl Acad Sci USA 1984, 81, 3180–3184.
11. Segel, L.A.; Perelson, A.S. Computations in Shape Space: A New Approach to Immune Network Theory, in Theoretical Immunology, Part Two, SFI Studies in the Sciences of Complexity. Perelson, A.S., ed. Addison-Wesley: Reading, MA, 1988; 321–343.
12. Smith, D.J.; Forrest, S.; Hightower, R.R.; Perelson, A.S. J Theoret Biol 1997, 189, 141–150.
13. Borst, P.; Rudenko, G.; Blundell, P.A.; van Leeuwen, F.; Cross, M.A.; McCulloch, R.; Gerrits, H.; Chaves, I.M.E. Behring Inst Mitt 1997, 99, 1–15.
14. Liu, Y.-J.; Zhang, J.; Lane, P.; Chan, E.Y.-T.; MacLennan, I.C.M. Eur J Immunol 1991, 2, 2951–2962.
15. Zhang, J.; MacLennan, I.C.M.; Liu, Y.-J.; Lane, P.J.L. Immunol Lett 1988, 18, 297–299.
16. Wiegel, F.W. Physical Principles in Chemoreception; Springer: Berlin and New York, 1991.
17. Wiegel, F.W. Introduction to Path-Integral Methods in Physics and Polymer Science; World Scientific: Singapore, 1986; Chapters I and II.