VIEWPOINT

Regulation of the Immune Response by Antigen

Rolf M. Zinkernagel and Hans Hengartner

How, why, and when specific T and B lymphocytes respond against infection follow explicit rules, but how this can be assessed experimentally depends crucially on the methodology used. In this Viewpoint, we discuss the parameters of receptor specificity and antigen that determine whether an immune response can be accurately measured against model antigens and how this relates to protection against a given pathogen. We suggest that antigen structure, localization, dose, and time during which antigen is available are all decisive factors in regulating an immune response.

The balance between host and infectious agents tends to evolve toward a state of mutual survival. For the pathogen, this balance depends on the level of direct cell and tissue damage caused by the infectious agent and on the kinetics of infection. For the host, survival depends on mechanisms of innate resistance and on the repertoire of T and B lymphocytes responsible for providing adaptive immunity (*1, 2*).

Raising an immune response can cost the host significantly because, to some extent, a degree of collateral damage to the host's own cells and tissues is an inevitable side effect and outcome of immunity. Therefore, evolution of beneficial immune protection has had to develop in equilibrium with the potentially lethal damage that immune responses can cause, namely, immunopathology (*3, 4*). In addition, the immune system needs to control foreign infections without reacting specifically against the host's own antigens and cells (*5, 6*). The balance of host and infectious agent might therefore be seen to reflect the strengths and weaknesses of a system that is "respected and exploited" by infectious agents to promote the coexistence of both. How such relations have evolved is illustrated by certain persistent viral infections. For example, the wart-forming papilloma viruses do not significantly harm the host and avoid generating immunity by evading the immune system's scope of detection for long periods of time. Noncytopathic hepatitis viruses have co-evolved with humans to infect offspring at birth via infected blood from carrier mothers. By infecting at a time when immune responsiveness of the newborn is virtually absent, disease-causing immunopathology is avoided or reduced to a level compatible with survival of the host (*1–3*).

How, then, does the immune system decide which antigens it should mount a full-scale immune response against and which it should not? Importantly, how can such decisions be accurately measured using existing experimental systems?

Institute for Experimental Immunology, University Hospital, 8091 Zurich, Switzerland.

Starting and Stopping an Immune Response

Whether the immune system reacts to antigen depends on the relative frequencies of responding T and B cells and on the thresholds of binding avidity their receptors display. Equally important are the levels of antigen present and the period during which the antigen remains in secondary organized lymphatic tissues, where primary immune responses are initiated (*7*).

T and B cells respond to antigens that become transiently localized within organized lymphatic tissues for at least 3 to 5 days. In contrast, T cells do not generally react against antigens, such as self proteins, that are continuously present at some level in blood and lymphoid organs (including thymus, spleen, lymph nodes, and bone marrow) (*2, 5, 6*). Principally, this is because such T cells have been functionally and physically eliminated by exhaustive induction or clonal deletion. Some noncytopathic persistent systemic infections that are transmitted from mother to offspring, as well as overwhelming chronic noncytopathic infections or lymphatic tumors in immunocompetent hosts, may also delete T cells (*8*). Thus, a common theme that has emerged from numerous studies is that persistence of antigen in the lymphohemopoietic system will eventually activate and delete all T cells specific for that antigen. A low precursor T cell frequency, together with persistent self antigen is also characteristic of the thymus. In the thymus, therefore, optimal conditions exist for early induction of self-reactive T cells, leading to deletion in the absence of immunopathology (*2, 5*). In contrast, adult hosts with sufficient numbers of reactive precursor T cells may induce immune responses that cause severe immunopathology when a noncytopathic virus has become widely disseminated within a particular organ (*3, 4, 9*).

Antigens that do not reach organized lymphatic tissue, irrespective of whether they are derived from self proteins or from infectious agents, not only fail to induce an immune response (*2, 7, 10*) but are also unable to delete T cells. For example, self antigens exclusively expressed in the brain (*11*) in pancreatic islet cells (including model antigens expressed after ectopic gene expression) (*10*) or infectious antigens expressed exclusively extra-lymphatically such as those of papilloma virus, which mature in keratinocytes of the skin, are all effectively ignored by the immune system. Nevertheless, potentially reactive T cells against these antigens do exist and can be induced if peripheral cells become damaged for prolonged periods of time. Under such conditions, self antigens that have evaded immunosurveillance may now reach secondary lymphoid organs and induce autoimmune T and B cells and cause disease if conditions prevail for long enough for an immune response to be induced (*4, 12*).

The rules for B cell (antibody) responses differ from those of T cells. B cells are induced very efficiently when an antigen is presented in a repetitive rigid form, such as found on the surface of infectious agents (Fig. 1) (*13*) or when linked to polyclonal B cell activators such as lipopolysaccharides (LPS) (*14*). Such antigens can induce immunoglobulin M (IgM) B cell responses independently of T cell help. Other antigen configurations, including multimeric antigens present on flexible backbones (e.g., flagellae of bacteria) or inserted into infected cell membranes, as well as monomeric or oligomeric protein antigens, induce B cells only if helped by specific $CD4^+$ T cells. Antigen concentration is also important in determining the level of help appropriated from T cells. Thus, the smaller the amount of antigen in absolute and local concentrations, the more directly dependent the B cell response is on T cell help. In addition, efficient switching from short-lived IgM production (IgM half life is about 24 hours) to other classes in particular long-lived immunoglobulin G (IgG) responses (IgG half life is about 20 days) requires conventional T helper cell activity.

B cell tolerance also differs from that of T cells. Experiments with transgenic B cells have suggested that deletion of autoreactive transgenic B cells can occur, as it does for

many self-specific T lymphocytes (*15*). However, several clinical and experimental observations indicate that autoantibody responses are readily induced but that they depend on the nature of the antigen (*2*, *6*). Thus, B cell autoreactivity may develop readily if antigens are highly repetitive in structure (common structures to which autoantibodies can develop include DNA, collagen, and the acetylcholin receptor) and if there are available helper T cells specific for a linked immunogenic helper T cell epitope (*6*, *13*).

How, then, is the immune system regulated or terminated so that immunopathology and autoimmunity may be avoided? As previously discussed, antigen dose and kinetics influence the duration and extent of an immune response. Accordingly, reduction of antigens below a minimal threshold by elimination of the infectious agent and of the antigen through opsinization and phagocytosis, will bring the response to a halt (*2*, *8*, *16*). In contrast, persistence of systematic antigen throughout the body may stop the immune response through the deletion of T cells (*8*). Persistence of very low levels of antigen in the periphery with periodical spreading offers a means of maintaining protective immunity, as observed in tuberculosis (*17*). But persistence of substantial localized amounts of antigen, together with an ongoing immune response, causes immunopathology, as seen with AIDS or hepatitis (*2*, *3*, *9*). Many do not believe the regulation of immune responses is invoked by antigen directly, but may instead rely on inherent mechanisms that enhance (positive regulation) or suppress (negative regulation) immune responses. Negative regulation has, however, been difficult to elucidate independently of antigen despite the many attractive hypotheses and experiments offered in support of this concept. For example, the existence of idiotypic networks, suppressor T cells, regulatory cytokine networks, or inhibition or deletion of lymphocytes by antigen encountered in the absence of costimulation (including so-called cross tolerance) have all been proposed as mechanisms of down-modulating immune responses (*18*–*25*). We would argue that although such mechanisms cannot be excluded completely, immune responses are generally brought to a halt primarily when antigen has been eliminated or has become segregated to specific peripheral sites.

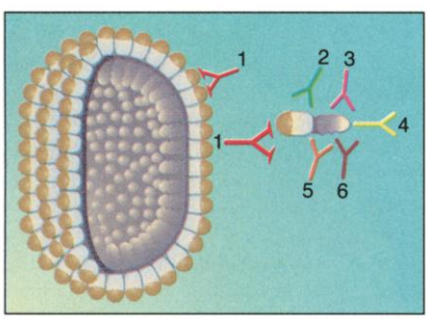

Fig. 1. Antigenic sites on a virus compared to an isolated protein. On many viruses (e.g., rhabdovirus VSV, vesicular stomatitis virus), the only determinant exposed on the surface is the tip of the glycoprotein to which the neutralizing antibody binds and because of size restrictions, only one antibody can bind at a time. Other antibodies cannot bind determinants that lie between the glycoproteins or inside the virus because these determinants are not accessible on the intact virus particle. Isolated purified glycoprotein of VSV has several antigenic sites, including the tip exposed on the intact virus particle (1). The other determinants (2, 3, 4, etc.) can bind antibodies that have been generated during a VSV infection against fragments of the envelope and against released glycoprotein. These additional antibody specificities (2, 3, 4, etc.) can be measured experimentally in a binding assay. However, only antibodies specific for determinant 1 are protective. Thus, antibody responses measured to isolated proteins (including model protein antigens such as hen egg lysozyme or ovalbumine) measure polyspecific and polyclonal antibody responses. In contrast, virus-neutralizing antibody responses are mono- or oligospecific and often oligoclonal (*2*, *13*, *26*, *33*).

Protective Immunity and Model Antigens

A major aim of immunological research is to unearth general rules for inducing protective immune responses and for preventing unwanted immune responses, such as in autoimmunity, immunopathology, or graft rejection. To do this, we often rely on model antigens or infections. Unfortunately, experimental observations and their interpretation often differ between these approaches and may even confuse our understanding of how the immune system operates.

To understand why this is so, we must critically re-evaluate not only the basic (and perhaps even dogmatic) assumptions that exist in immunology, but also methodologies applied to dissect immune responses. As in any science, the experimental methods used must hold some influence over results obtained (*2*). In immunology, protection against infection or against established tumors might be considered the most meaningful readout for immunity. Nevertheless, in vitro measurements including cellular proliferation, interleukin production, cytotoxic activity, and tetramer or antigen binding all provide relatively reliable surrogates for measuring immunity, providing we keep in mind that each measures distinct parameters and each has a different threshold of detection.

One potential cause of discrepant experimental results in measuring immune responses, is that often neither the number of antigenic determinants measured and recognized by B or T cells nor the involved receptor specificities are well defined. Although infectious agents express hundreds to thousands of determinants, they expose on the surface of infected cells only one or a few antigenic sites that are relevant for recognition by the host's T cell pool. Also, on the intact infectious agents, only one or few antigenic determinants are exposed that are both important for virus infection and are accessible to B cell receptors and protective antibodies (*2*, *26*) (Fig. 1). For example, immunity against polio virus strain 1 cannot protect against strains 2 or 3. Therefore, although most specificities of B and T cells induced by polio virus are the same for all three stains, only the neutralizing (protective) antibodies define specificity of immunity. In contrast to protective determinants on infectious agents or toxins, model antigens often used in immunological research expose multiple (between 10 and 1000) antigenic sites on particular proteins, or in mixed antigen preparations (*27*). Similarly, although only one or few specific T cell epitopes are expressed by infectious agents in the context of one individual's major histocompatibility complex (MHC) haplotype (*2*), basic immunological studies often examine T cell responses against multiple minor or major histocompatibility differences that represent hundreds to thousands of different T cell epitopes (*19*, *28*, *29*). Thus, polyspecificity and polyclonality of T or B cell responses may suggest specific responses that in fact reflect cross-reactivities (*30*).

A second experimental factor that may confuse the analysis of an immunological response, is the issue of the available T cell and B cell receptor repertoire and respective receptor frequencies. The immune repertoire can probably offer efficient protection against about 10^3 to 10^4 distinct infections relevant for the survival of a given species (*1*, *2*, *16*). Thus, in mice about 1 to 5×10^7 each of mature $CD8^+$ T, $CD4^+$ T, or B cells exist and immunity is generated from a starting number of about 100 to 1000 antigen-specific precursor T and B cells. If the frequency of specific lymphocytes were much lower, induction of an immune response would be too slow to be protective. On the other hand, if it increases above 10^{-3} to 10^{-2}, then activation becomes less controllable because of nonspecific bystander effects, which can influence the immune response. This potential problem is borne out by studies using experimental mice that express very high precursor frequencies of specific transgenic T cells to study crosspriming, tolerance, or memory (*19*–*23*, *31*). For example, bacterial infection of T cell receptor transgenic mice bearing an unrelated receptor specificity can result in T cell activation via nonspecific bystander effects (*22*, *31*, *32*).

Taken together, caution needs to be exercised when interpreting results from some experimental model systems. Thus, studies of

immune responses using model antigens that may not reflect structural features of antigens resulting from co-evolution, or those that use systems where an abnormally high frequency of specific T or B cells exist, may not be representative for understanding immune tolerance or immunity against infections and tumors. In the latter case, antigens generally expose to the immune system only one or few essential antigenic sites or peptides to relatively low frequencies of T and B cells (Fig. 1) (*1*, *13*, *26*). Thus, discrepancies between the immunology of model antigens and immunity against infections may eventually be resolved by more stringent definition of relevant characteristics of the chosen experimental system (*10*, *12*, *26*, *33*). Understanding these critical parameters will enhance our understanding of basic immunology and will not only help predict the rules of why, how, and when the immune system reacts but will also enable us to better explain the pathophysiology of infectious disease. Ultimately, this will vastly improve rationales on how to offer protection through vaccination.

References and Notes

1. C. A. Mims, *Pathogenesis of Infectious Disease* (Academic Press, London, 1987).
2. R. M. Zinkernagel, *Science* **271**, 173 (1996).
3. J. Hotchin, *Cold Spring Harbor Symp. Quant. Biol.* **27**, 479 (1962).
4. R. M. Zinkernagel *et al.*, *Immunol. Rev.* **122**, 133 (1991).
5. P. Kisielow, H. von Boehmer, *Adv. Immunol.* **58**, 87 (1995).
6. W. O. Weigle, *Adv. Immunol.* **16**, 61 (1973).
7. U. Karrer *et al.*, *J. Exp. Med.* **185**, 2157 (1997).
8. D. Moskophidis, F. Lechner, H. P. Pircher, R. M. Zinkernagel, *Nature* **362**, 758 (1993).
9. F. V. Chisari, C. Ferrari, *Annu. Rev. Immunol.* **13**, 29 (1995).
10. P. S. Ohashi *et al.*, *Cell* **65**, 305 (1991).
11. W. E. Klinkert *et al.*, *J. Neuroimmunol.* **72**, 163 (1997).
12. P. S. Ohashi *et al.*, *J. Immunol.* **150**, 5185 (1993).
13. M. F. Bachmann *et al.*, *Science* **262**, 1448 (1993).
14. A. Coutinho, *Transplant. Rev.* **23**, 49 (1975).
15. J. I. Healy, C. C. Goodnow, *Annu. Rev. Immunol.* **16**, 645 (1998).
16. M. Cohn, R. E. Langman, *Immunol. Rev.* **115**, 11 (1990).
17. T. M. Kundig *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9716 (1996).
18. G. Moller, *Immunol. Rev.* **110**, 1 (1989).
19. P. Matzinger, *Annu. Rev. Immunol.* **12**, 991 (1994).
20. E. M. Shevach, *J. Exp. Med.* **193**, F41 (2001).
21. The popular postulate of cross priming assigns to dendritic cells (DC), those cells that carry antigen to lymph nodes, the exclusive and essential capacity to induce helper and cytotoxic T cells. Intracellular antigens are generally presented via MHC class I proteins and phagocytized cell-external antigens via MHC class II, for recognition by lytic $CD8^+$ T cells and nonlytic $CD4^+$ helper T cells, respectively. To induce $CD8^+$ killer T cells against infections that avoid DCs and against extralymphatic peripheral tumours (carcinomas or sarcomas), it has been proposed that DCs are also uniquely capable of activating naïve $CD8^+$ T cells after processing phagocytozed antigens via MHC class I. These processes have been termed cross-presentation and cross-priming (*22*, *23*). However, we suggest that the assumption that only DCs can induce $CD8^+$ T cells have been overstated because antigenic fibroblasts, or epithelial cells that lack conventional second signals, efficiently induce $CD8^+$ T cells if they reach draining lymph nodes or the spleen and if helper T cells are available (*24*, *25*). Therefore, we suggest that cross-priming is not essential and, at best, is inefficient under physiological conditions in vivo. This makes sense because otherwise B cells or DCs picking up soluble antigens may generally become over-susceptible to elimination by cytotoxic T cells, a process that would prematurely stop immune responses. Nevertheless, with some manipulation certain limitations of cross-priming may be overcome and might, perhaps, become exploitable therapeutic setting. Last but not least, our tendency to use experimental animals raised under nonphysiologically clean conditions (so called specific pathogen-free) may skew the specificity of reactivity due to reduced nonspecific, concomitant immunity and unwanted bystander effects.
22. F. R. Carbone, C. Kurts, S. R. Bennett, J. F. Miller, W. R. Heath, *Immunol. Today* **19**, 368 (1998).
23. R. M. Steinman, S. Turley, I. Mellman, K. Inaba, *J. Exp. Med.* **191**, 411 (2000).
24. T. M. Kundig *et al.*, *Science* **268**, 1343 (1995).
25. A. F. Ochsenbein *et al.*, *Nature* **411**, 1058 (2001).
26. H. P. Roost *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1257 (1995).
27. M. A. Newman, C. R. Mainhart, C. P. Mallett, T. B. Lavoie, S. J. Smith-Gill, *J. Immunol.* **149**, 3260 (1992).
28. P. Matzinger, M. J. Bevan, *Cell Immunol.* **29**, 1 (1977).
29. M. Simonsen, *Transplant. Rev.* **3**, 22 (1969).
30. T. Hunig, M. J. Bevan, *Nature* **294**, 460 (1981).
31. R. Ganss, A. Limme, T. Sacher, B. Arnold, G. J. Hammerling, *Immunol. Rev.* **119**, 203 (1999).
32. S. Ehl *et al.*, *J. Exp. Med.* **187**, 763 (1998).
33. M. F. Bachmann *et al.*, *Science* **276**, 2024 (1997).

VIEWPOINT

Sensing Pathogens and Tuning Immune Responses

Bali Pulendran,* Karolina Palucka, Jacques Banchereau

The immune system is capable of making qualitatively distinct responses against different microbial infections, and recent advances are starting to reveal how it manages this complex task. An integral component of the immune system is a network of cells known as dendritic cells (DCs), which sense different microbial stimuli and convey this information to lymphocytes. A better understanding of DC biology has allowed a model to be constructed in which the type of immune response to an infection is viewed as a function of several determinants, including the subpopulation of DCs, the nature of the microbe, microbe recognition receptors, and the cytokine microenvironment.

When a microbe enters the body, the immune system is faced with a series of challenges. First, a decision needs to be made as to whether to respond to that specific microbe or not. Second, if a response is made, it must be tailored to fight that particular microbe. For example, in response to intracellular microbes, such as viruses and certain bacteria, $CD4^+$ T helper (T_H) cells differentiate into T_H1 cells, which secrete interferon-γ (IFN-γ) and possess a specific range of functions. In contrast, extracellular pathogens such as helminths induce the development of T_H2 cells, whose cytokines [interleukin 4 (IL-4), IL-5, and IL-10] direct immunoglobulin E– and eosinophil-mediated destruction of the pathogens (*1*). Generating the right class of immune response can be a matter of life and death itself. Thus, in leprosy, the tuberculoid form of the disease is characterized by a protective type 1 response, but the lepromatous form induces an often lethal type 2 response.

Although B and T lymphocytes respond to antigens with high specificity, they alone are not capable of making these complex decisions. These choices are made jointly by the nature of the microbe and by dendritic cells (DCs). DCs are scattered throughout the body, including the various portals of microbe entry, where they reside in an immature form (*2*–*5*). Immature DCs can be considered "immunological sensors," alert for potentially dangerous microbes, and are capable of decoding and integrating such signals. They then ferry this information to naïve T cells in the T cell areas of secondary lymphoid organs, undergoing a maturation process en route. Here, the mature DCs present this information to T cells, thus launching an immune response and immune memory through which the antigenic encounter can be remembered even for a lifetime (*4*). DCs can also tune the immune response by modulating either the amplitude or the class of the response

Baylor Institute for Immunology Research, 3434 Live Oak, Dallas, TX 75204, USA.

*To whom correspondence should be addressed. E-mail: balip@baylordallas.edu