**[Frontiers in Bioscience 1, d318-323, 1 October 1996]**

# THE IMMUNE SYSTEM: A LOOK FROM A DISTANCE

**Melvin Cohn[1] and Rodney E. Langman**

*The Salk Institute, P.O. Box 85800, San Diego, California, 92186-5800 USA*

**TABLE OF CONTENTS**



## 1. ABSTRACT

The self-nonself discrimination is germline encoded for defense mechanisms, but it is somatically learned for the immune system and this is the fundamental difference between the two. When referring to the defense mechanisms of vertebrates, immunologists like to use the term “innate immune systems” to describe the germline encoded class of defense mechanism. It was the acquisition of a somatically learned S-NS discrimination during vertebrate evolution that permitted the immune system to develop large recognitive repertoires compared to those of defense mechanisms. This seemingly boundless immune repertoire has fascinated immunologists for almost a century.

Today we have a better understanding of the size and function of the antibody repertoire. Humoral antibody effector functions depend upon secreted immunoglobulin and the concentration of antibody must reach a minimum effective threshold in a short enough time to stop a growing pathogen before it becomes lethal. This requires that initially an equivalent number of B-cells per ml respond to the pathogen. This number of B-cells must respond for each and every milliliter of animal. Consequently, the humoral immune system must be iterated. This straightforward conclusion has far reaching implications, some of which are explored in this review.

Received: 8/15/96; Accepted: 9/9/96

[1] To whom correspondence should be addressed at The Salk Institute, P.O. Box 85800, San Diego, California 92186-5800, Tel #: (619) 453-4100 Ext. 1351, FAX #: (619) 453-4133, E mail: cohn@sc2.salk.edu

## 2. INTRODUCTION

All organisms need mechanisms that provide protection against infectious pathogens. Prokaryotes and invertebrate eukaryotes have a variety of such mechanisms, including restriction enzymes, lectins, lytic peptides, phagocytes, etc.(1). We refer to these as defense mechanisms. Vertebrates also have these defense mechanisms, but, in addition, they have an immune system. Both defense mechanisms and the immune system must make a self-nonself (S-NS) discrimination because they link a recognitive element to a destructive and ridding set of effector functions. Any host that allowed the destruction and ridding of the pathogen to entrain a significant measure of destruction and ridding of the host, would obviously self-destruct (*i.e.*, be deleted by evolutionary selection). The self-nonself discrimination is germline encoded for defense mechanisms, but it is somatically learned for the immune system and this is the fundamental difference between the two. When referring to the defense mechanisms of vertebrates, immunologists like to use the term “innate immune systems” to describe the germline encoded class of defense mechanism.

The acquisition of a somatically learned S-NS discrimination during vertebrate evolution permitted the immune system to develop large recognitive repertoires compared to those of defense mechanisms. It is the seemingly boundless size of the immune repertoire that fascinates immunologists.

Recognition without any consequence would be evolutionarily unselectable. This indissoluble linkage is what drives the pathway of decisions that the immune system must make on encountering an antigen.

## 3. THE DECISION PATHWAYS OF AN IMMUNE RESPONSE

Decision 1, is the antigen self or nonself? If it is self, an immune response must be inactivated; if it is nonself, an immune response must be activated and further control passed on to Decision 2, in order to determine which effector class would be optimal in ridding the pathogen? This latter decision is needed to cope with multiple, often contradictory effector reactions. For any given pathogen, there are ineffective and effective effector functions. In many cases the ineffective effector functions can block the efficacy of the effective effector functions because both compete for the recognition of antigen. The S-NS discrimination, determines the specificity with which the effector response rids the inducing pathogen without self-destructing. The specificity of the effector response is composed of several elements, one of which is the specificity of the antigen-receptor on responsive cells itself.

Decision 2, the choice of class of effector function is related to the location and the nature of the pathogen, because these factors determine the ability of a particular effector function to destroy and rid the pathogen. Cell-bound pathogens such as viruses, intracellular bacteria, rickettsia, and certain protozoan parasites require a response in the cell-mediated effector class. In general, the cell-mediated mode is a delaying tactic. The infection is slowed down but not ridded. A virally infected cell that is lysed by a cytotoxic lymphocyte can liberate free virus capable of infecting other cells albeit at a much lower yield. To rid this virus, a humoral response is eventually required. In many cases of viral infection, the effector response is initially cell-mediated with a subsequent switching over to the humoral response. In a few cases, generally involving non-viral intracellular pathogens, the cell-mediated mode is sufficient to keep the infection in check.

Free pathogens, such as bacteria, initially require a humoral antibody response. In this case there are a handful of effector functions available to the immune system; including, complement lysis, antibody dependent cellular cytotoxicity, opsonization, chemical warfare (*e.g.*, histamine and serotonin release), neutralization of toxicity, and blockage of invasiveness. These effector functions are associated with different Ig isotypes, albeit with some overlap. A choice must be made between these isotypes that relates them to the effectiveness of ridding the pathogen.

There are three key questions to consider: What are the factors governing a learned S-NS discrimination? What does evolution look at when selecting upon the humoral response? What are the requirements for a regulation of class?

## 4. THE S-NS DISCRIMINATION IN A NUTSHELL

The S-NS discrimination, a somatic learning process, is dependent on:

a) Antigen-responsive cells (i-cells) being born in an initial i-state with no effector function and with two pathways, inactivation or activation, open to them upon encounter with antigen.

b) Self being defined as those antigens that are present when the immune system arises in ontogeny and which persist.

c) Nonself being defined as those antigens that appear after the immune system is mature and which are transient.

A learning or historical process means that the pathway taken by an i-cell upon encountering antigen depends on the prior experience of the immune system with respect to that antigen.

A decision between two pathways requires two signals. The interaction of the i-cell with antigen signals inactivation (Signal 1). If, in addition to Signal 1, the i-cell receives a second signal (Signal 2), then the cell will be activated and the further steps of division and differentiation to effectors (Decision 2) is put under the control of interleukins.

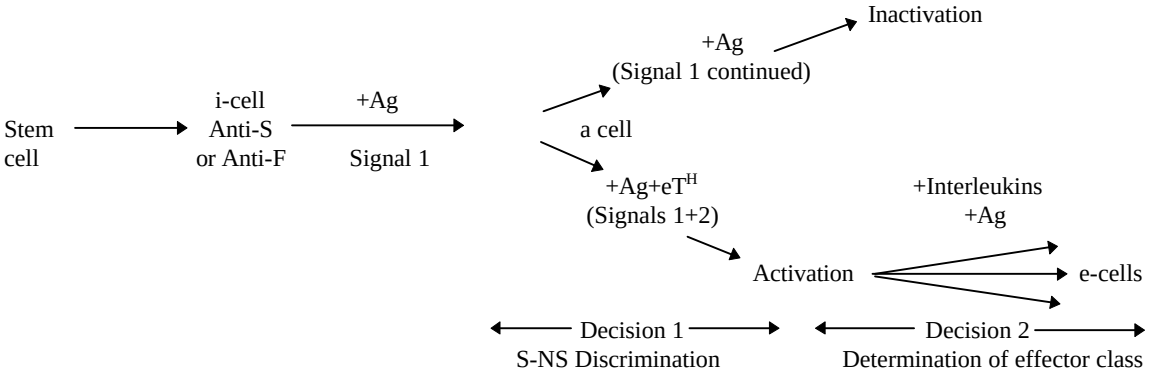


**Figure 1:** The pathway of induction of antigen-responsive (i-) cells to effectors (e-cells).

Once activated a S-NS discrimination is no longer possible. This pathway of induction is illustrated in Figure 1:

The presence or absence of Signal 2 determines whether or not a cell is activated or inactivated. Therefore, the central question is, "What regulates the delivery of Signal 2."

We have always insisted that Signal 2 must be delivered by a regulatory effector T-helper ($eT^H$), a cell that has the same degree of antigen specificity as all other immune effectors and has itself undergone a S-NS discrimination. The delivery of Signal 2 must be short range (a cell-cell interaction between $eT^H$ and an i-cell), and require associative recognition of antigen (*i.e.*, two or more linked determinants on the antigen must be recognized, one by the $eT^H$ cell and the other by the i-cell). Associative recognition of antigen is the only way to assure a coherent response of i-cells to any of the epitopes linked on the antigen.

This raises the following question:

If $eT^H$ are required to activate all i-cells including $iT^H$ itself, where does the first $eT^H$ come from?

This is referred to as the "primer question."

Over the years there have been two answers to the primer question.

First, $eT^H$ are required for the activation of all i-cells, except $iT^H$. These latter are activated to $eT^H$ upon receiving Signal 1 in the presence of non-specific "inflammatory" agents, adjuvanticity, danger, harm, costimulation, cell necrosis, etc. We will comment on this position later.

Second, $eT^H$ are required for the activation of all i-cells, including $iT^H$ and, therefore, there must be a nonself antigen-independent pathway from $iT^H$ to $eT^H$ that has undergone a S-NS discrimination. In this case, a steady state production of $eT^H$ anti-nonself primes the immune response.

Under the first view, the non-specific activating event has nothing to do with the S-NS discrimination. The S-NS discrimination must be made by deleting all anti-S from the $iT^H$-cell population prior to being activated by the nonspecific signal. This nonspecific activation rule applies only to the generation of all $eT^H$.

Under the second model, the rules of associative recognition of antigen are universal and include $iT^H$. The antigen-independent pathway from $iT^H$ to $eT^H$ provides the priming level of $eT^H$ anti-nonself.

The first model for the origin of $eT^H$ has several variants. The best formulated is referred to as the "danger model."

Like all primer models based on an inductive Signal 2 that is delivered by a source that itself has not undergone a S-NS discrimination (referred to as "nonspecific"), a set of filtering and deletion steps must be proposed to rid those $iT^H$ that are anti-self prior to their nonspecific activation. Under the "danger" model, this is accomplished in two stages. The $iT^H$ are born in the thymus where most of the self is presented and where most of the $iT^H$ anti-self are deleted. For those self-antigens present uniquely in the periphery (*i.e.* not presented in the thymus) another mechanism operates that is based on the partitioning of self on uniquely tolerigenic antigen-presenting cells and of nonself on uniquely inductive antigen-presenting cells (activated by danger, a unique property of nonself).

While we argue that the first model ("danger") contributes nothing to the S-NS discrimination and cannot account for the origin of effector T-helpers, there is merit in putting an emphasis on the role of inflammatory factors in modulating immune responsiveness. These factors play their role in Decision 2 by modulating the quantity and quality of the effector response. Many are known and referred to as interleukins and cytokines. Important here is that they do not contribute to Decision 1.

## 5. THE ESSENCE OF THE HUMORAL RESPONSE, THE CONCEPT OF A PROTECTON

What activities constitute the evolutionary selection pressure that shaped the humoral response?

The analysis of this question leads to a new concept that will appear at first somewhat strange. Humoral antibody effector functions depend upon secreted immunoglobulin. The concentration of antibody must reach a minimum effective threshold in a short enough time to stop a growing pathogen before it becomes lethal. This requires that initially an equivalent number of B-cells per ml respond to the pathogen. This number of B-cells must respond for each and every milliliter of animal. Consequently, the humoral immune system must be iterated. This straightforward conclusion has far reaching implications.

Before discussing the implications, let us give some rough numbers that would illustrate this concept. First, is that a threshold antibody concentration of 100ng/ml must be reached within 5 days to protect against the 'worst case' pathogen. This would require that roughly 200 B-cells per ml specific for the pathogen be present initially. This

applies to each and every milliliter of animal. Second, the iterated unit must be sufficiently diverse to be protective against a variety of pathogens - missing, say, 3 in every $10^3$ pathogens. Third, there is a limit to the total number of B cells per ml that is around $10^7$/ml for most species. The iterated unit, then, must have a minimum total size and a concentration parameter. Our best estimate is that the iterated unit is a total of $10^7$ B-cells, at a concentration of $10^7$ B-cells per ml with roughly 200 B-cells per ml responsive per pathogen. We refer to this iterated unit of protection as a Protecton. The Protecton is the target of evolutionary selection on the humoral immune system.

Consider a pygmy shrew with $10^7$ total B-cells, a mouse with $10^8$ B-cells, a human with $10^{12}$ B-cells and an elephant with $10^{14}$ B-cells. This translates into a pygmy shrew with 1 Protecton, a mouse with 10 Protectons, a human with $10^5$ Protectons and an elephant with $10^7$ Protectons. These animals are equally protected against their pathogenic universes by their humoral immune systems. They are protected per milliliter not per animal. All Protectons are equivalent in function.

There are four points to make before confronting several implications of the concept of a Protecton.

1) The Protecton is defined as the smallest sample of the humoral immune system that retains all of the evolutionary selectable protective properties of the whole.

2) While, for simplicity, we have treated Protectons as independent units, there is a factor of cooperativity between them, but this is second order for our discussion.

3) The minimum Protecton unit we have introduced is based on protection against a 'worst-case' bacterial infection. On the one hand, there are some exceptionally lethal pathogens, especially those producing potent toxins, that are largely beyond immune control. These exceptionally fast growing virulent pathogens are rarely encountered and must be self-limiting for non immune reasons—otherwise there would be no surviving vertebrates with immune systems. On the other hand, there are pathogens that grow relatively slowly but require rare and highly specific antibodies to limit their growth. A Protecton can be defined on the basis of the immune response to a particular infection and in the case of a slow growing pathogen it could take as long as 10 or even 20 days before the threshold concentration is reached. A Protecton for a slower growing pathogen might be $10^6$ cells or for a faster growing pathogen, closer to $10^8$ cells. Thus, a Protecton is a vectorial quantity that is directed against each individual pathogen as it is encountered.

4) The Protecton characterizes the primary encounter with a pathogen. This is the key step in evolutionary selection. If the response of the virgin immune system cannot protect against a pathogen then the response on secondary encounter would be of little interest. The secondary response is essentially a byproduct of an effective primary response, the direct target of evolutionary selection.

### 5.1. Some consequences of Protecton theory

Now let us look at the consequences of this concept that the humoral immune system is iterated.

Immunologists have always viewed the immune system as being able to call upon a transcendental repertoire for an effective response. The early models of diversification might best be described as "big bang." The repertoire was viewed as being expressed in its totality in one step whether this step was the combinatorial expression of many germline V-genes segments or this step was the hyper recombined V-gene segments or of hyper varied V-gene segments by random replacement by minigenes of their complementarity-determining regions. While big-bang seemed to describe the observations it lacked any credible arguments of evolutionary necessity. We have argued that the repertoire must be expressed in two stages. STAGE I is a small germline encoded repertoire that is represented in high copy number, and also acts as a substrate for STAGE II, which is generated by somatic diversification (hyper mutation) and is in single copy. This view has met with strong resistance largely because it was derived as an evolutionary necessity not as a direct observation.

When it was learned that a relatively small number of V-gene segments were present in the genome, big-bang fell briefly from favor. A short time later the "big bang" model was reborn like the Phoenix when junctional diversity and an extra D-gene segment was discovered. This allowed enormous repertoires to be derived by multiplication of numbers of rearranging gene segments by functional joining variation by subunit complementation to arrive at repertoire sizes in excess of $10^{10}$. Every review and textbook covering repertoires carries this calculation and the term 'complete' has become popular to describe the range of these repertoires.

The concept of a Protecton has made this calculation misleading. Clearly the size of the available repertoire cannot be larger than the number of B-cells per Protecton, that is, $10^7$. To illustrate, consider a mouse with $10^8$ total B-cells. If the repertoire were $10^{10}$ (the usual minimum estimate) and any one of those specificities were important for the protection of an individual, then only one in 100 mice would express that specificity at any given moment in time and even that mouse would be unprotected unless it was allowed to take almost 30

days for the one B-cell to multiply to a protective level. The latter would even be true for a human with $10^{12}$ total B-cells. The individual would express $10^2$ total B-cells specific for the pathogen but being too few they would respond too slowly to protect. Vast, transcendental repertoires are evolutionarily unselectable as such because they are of a nonfunctional size.

Returning now to a more realistic estimate of the size of the functional (or available) repertoire, an upper limit in principle is $10^7$ based on $10^7$ B-cells/ml; but, this too is a substantial overestimate. An analysis of the pathway of expression of the Protecton places the repertoire at about $5x10^4$. This repertoire is composed of a germline (STAGE I) repertoire of ~$10^4$, but each specificity is present in high copy number (~$10^2$ B-cells per specificity per Protecton) and a STAGE II somatic mutationally derived repertoire of ~$4x10^4$, which is in low copy number (1 B-cell per specificity per Protecton). These two repertoires interact synergistically to provide a sufficiently rapid response to a large enough family of pathogens.

**5.2. The primary repertoire and the pathogenic universe**

As a rough estimate, this virgin repertoire protects the individual at the 99% level. This is the limit to evolutionary selection because other factors such as the probability of being eaten by a predator or of starving becomes the limiting factors for survival. What the immune system really does is seen in immune deprived individuals where a surprisingly large family of pathogens are revealed as 'opportunistic'.

Protecton theory highlights a detail of effector function that is very important for the design of vaccines and passive antibody treatment. A monoclonal antibody may neutralize a pathogen or toxin by blocking attachment to its target or by inactivating an enzymatic activity, but it is ineffective in ridding the antigen. Ridding is largely a function of opsonization by macrophages and this requires the formation of a three dimensional aggregate of antibody.

By way of illustration, consider a monomeric antigen like diphtheria, tetanus or cholera toxins. A monoclonal antibody might neutralize its toxicity, but because it cannot form a aggregate with the antigen, the toxin would not be effectively ridded. Two monoclonal antibodies reacting with different determinants on the monomers would form a linear chains of immunoglobulin, and that too is inefficiently opsonized. It takes 3 or more antibodies reacting with different determinants to form the three dimensional aggregate that is ridded efficiently. Neutralization does play an important role by giving the immune system more time to respond and produce the ridding antibodies.

Because evolution selects on the limiting case, on average three or more antibodies would be induced by polymers even though a monoclonal antibody reacting with a polymer might be sufficient to allow aggregation. However, antibody aggregated on a virion is less effective in ridding the virus than virions aggregated by antibody. Whether a monoclonal antibody interacting with virions will cross link or bivalently bind depends on the spacing of the ligand recognized. If 3 or more antibodies bind, cross linking is assured and ridding is effective.

The repertoire of ~$5x10^4$ specificities in the Protecton divides the antigenic universe into epitopes distributed randomly and combinatorially on antigens, ten at a time. The total number of antigens distinguishable by this repertoire is ~$10^{43}$ ($_{5x10^4}C_{10}$), a big enough number for any theory. This repertoire will "miss" 3 in 1000 antigens because they will be seen in less than 3 ways by the Protecton. The number 10 epitopes per antigen is an estimate based on a computer modeling study of the Protecton.

It might seem surprising that a small repertoire can deal with a large antigenic universe. An understanding of how this works begins with four points.

1) It is the paratope (combining site) that is primary as the target of evolutionary selection. Only paratopes, define epitopes.

2) A given paratope (antibody) that reacts with several epitopes distinguishable by the immunologist (referred to as crossreactivity) treats these epitopes as a single epitope functionally. Any antibody that recognizes an epitope present on a self and a nonself antigen is defined by the immune system as anti-self and the epitope as a self-epitope.

3) Symmetrically, if a given epitope is recognized by several paratopes (antibodies) distinguishable by the immunologist (referred to as degeneracy), the immune system treats this family of antibodies as one antibody functionally.

4) The paratope looks at "shape" and a given paratope-defined shape can be created from many different chemical structures. For example, an anti-carbohydrate paratope can be found to react with a peptide. It is this recognition of shape, not chemistry, that allows the paratopic repertoire to divide the universe of chemically different antigens into a limited number of epitopes. Paratopes define antigens as collections of linked epitopes combinatorially distributed.

## 6. THE CLASS OF THE RESPONSE (DECISION 2), AN UNRESOLVED QUESTION

In general a functional immune system responds to a given pathogen in a class that is effective in destroying and ridding it. It seems obvious that this should be the case because there are effective and ineffective classes and a random response in all classes would result in the ineffective classes blocking the function of the effective classes. Implied is that regulation of the class of the response is required and this regulation must relate recognition of some property of the given pathogen to the induction of an effective class.

The immune system might look at the pathogen in a stereotyped or a learned fashion or both. As an example of a stereotyped response the immune system might respond in the cell-mediated category to all cell-bound pathogens (*e.g.* viruses) and in the humoral category to all free pathogens (*e.g.* bacteria). Cell bound pathogens would be recognized by their presentation with restricting elements encoded in the major histocompatibility complex. As an example of a learned response the immune system during infection might assay which class is effective in ridding the pathogen and suppress all other classes as ineffective.

A solution to the problem of regulation of class will have major practical consequences. The ability to manipulate the class will permit direct control of many dyscrasias. For example, switching a response from an effective to an ineffective class will permit transplantation of tissues, as well as control of autoimmunity and allergy. Switching from an ineffective to an effective class will permit treatment of infections where the pathogen subverts the response of the immune system by inducing an ineffective class as well as the rational design of effective vaccines that induce an effective class. Recognition is not enough; it must be coupled to an effective effector function.

## 7. THINKING ABOUT THE IMMUNE SYSTEM

The guiding principle must be based on evolutionary considerations. There is a tug-of-war relationship between a mutationally derived increase in the immune protection against the pathogenic load and the mutational escape of the pathogen from destruction. This process reaches an apparent steady state when the level of protection is no longer limiting to the procreation of the species. The consequence of this limit is that no property is absolute or perfect. There is a limit to the degree of specificity of the receptors, the completeness of haplotype exclusion, the accuracy of signaling between cells, the black-and-whiteness of the Self-Nonself discrimination, the efficacy of effector function, and so on. In the end the immune system fails to protect for two reasons:

1) The response is too slow because the number of cells per ml that respond to the pathogen are too few.
2) The class of effector function that is induced is ineffective.

Vaccination and passive antibody treatments deal with the first problem by calling on specificities in the virtual or potential repertoire that are in too low frequency to be protective in a primary response. This is possible because the animal is being vaccinated or immunized under non-threatening conditions, thus allowing weeks and even months to pass while the response of low frequency specificities is being amplified. Passive antibody treatments can call upon antibodies isolated by hi-tech hybridoma and cloning methodologies or by combinatorial libraries. In this case, a mixture of antibodies can be used for treatment that are not only too rare to be induced by vaccination but may even be non-inducible yet functional as effector molecules (*e.g.*, antibodies with mutations in the framework or with a DN region that creates a nonfunctional signaling antigen-receptor). Obviously these antibodies in the potential repertoire are available to us but unavailable to evolution.

Vaccination and passive antibody treatments require that the specificities involved be linked to effective effector functions. If there is no class or classes of response that would be effective in ridding the pathogen, then manipulation of the immune response would be useless. The effective class of response must be known if one is to design effective vaccines and antibody treatments. Clearly the treatments must be in the effective class.

The above interventions depend on understanding and knowledge. The more encompassing our understanding, the greater the probability that we will be able to creatively intervene.

## 8. ACKNOWLEDGMENTS

Supported by NIH grant RR07716. This paper was written while one of us, Melvin Cohn, was a Scholar-in-Residence at the Fogarty International Center for Advanced Study in the Health Sciences, National Institutes of Health, Bethesda, Maryland, USA.

## 9. REFERENCES

1. RE Langman: The Immune System, Academic Press, San Diego, California (1989)