Cell. Mol. Life Sci. (2010) 67:2851–2862
DOI 10.1007/s00018-010-0438-z

Cellular and Molecular Life Sciences

VISIONS & REFLECTIONS

The evolutionary context for a self–nonself discrimination

Melvin Cohn

Received: 1 June 2010 / Accepted: 10 June 2010 / Published online: 24 June 2010

Abstract This essay was written to illustrate how one might think about the immune system. The formulation of valid theories is the basic component of how-to-think because the reduction of large and complex data sets by the use of logic into a succinct model with predictability and explanatory power, is the only way that we have to arrive at "understanding". Whether it is to achieve effective manipulation of the system or for pure pleasure, "understanding" is a universally agreed upon goal. It is in the nature of science that theories are there to be disproven. An experimentally disproven theory is a successful one. As they fail experimental test one by one, we end up with a default theory, that is, one that has yet to fail. Here, using the self–nonself discrimination as an example, how-to-think as I see it, will be illustrated.

Keywords Self–nonself discrimination · Autoimmune disease · Repertoire · T cell · Immunopathology

Introduction

When Prof. Eichmann invited me to write a paper on my experiences that might be useful to the next generation in pursuing a scientific career, I hesitated and for good reason. The society in which we live today has evolved and its values and goals have changed. The aspirations of my generation were quite different from those of the generation likely to be reading this essay. This is not a question of good and bad, or moral and immoral. It is purely an observation. Consequently, there is little of practical value that my experiences can teach the contemporary reader.

My generation went through the Great Depression of the 1930s. Getting an education was hard fought. We cared about justice and were at the forefront of the struggle against religious and racial bigotry on the one hand, and fanaticism on the other. My generation suffered huge casualties in the war against dictatorships, and in the end paid the price of eternal guilt for developing and using the atomic bomb. When we returned to civilian life to face a career, we were treated as having fallen far behind but that made us much more determined to catch up. Here, providence smiled as the war in Europe had produced refugees among which were the intellects of a previous generation, some of whom I was fortunate to have had as my teachers. Hopefully, names such as Fritz Lipmann, René Dubos, Ephriam Racker, Severo Ochoa, Otto Loewi, Max Delbruck, Salvatore Luria will still be remembered. And, of course, among my mentors were those who were not refugees but had suffered immense gaps in their scientific careers during the World War II years either by serving in the armed forces or in activities allied to it. To cite a few, A.M. Pappenheimer Jr., Elvin Kabat, Michael Heidelberger, Oswald Avery, Colin Macleod, and Mark Adams.

We were, as a result, an ascetic generation; we suffered the onslaught of political intolerance and ignorance characterizing the cold war of the 1950s and 1960s. In spite of it, we cared about and were excited by "understanding" and searched for the most general explicative concepts that we unashamedly referred to as "truths". We respected

M. Cohn (✉)
Conceptual Immunology Group,
The Salk Institute for Biological Studies,
10010 North Torrey Pines Road, La Jolla,
CA 92037, USA
e-mail: cohn@salk.edu

"thinking" as a way of life and resisted all manners of fanaticism by our rational personal behavior. In general we were open, communicative, transparent, and engaged. We built new institutions of learning and nervously watched them evolve, from reflective to executive. The Salk Institute, of which I am a Founder, being but one example. We created new fields of investigation, molecular biology and regulatory biology being the two most obvious.

The generation who might read this essay grew up in an affluent society and was handed most of the benefits we struggled to have. They are producing scientists just as creative as we were, probably even more so. However, their values are different and my past experience has little to teach them. Their emphasis in science is to investigate that which is easily translatable into application and to thereby amass personal wealth. Seems healthy enough, as long as ethical standards, the lack of which did not plague my generation as it does this one, are agreed upon and met. Translational science requires a knowledge-base from which to translate; that was our emphasis. This generation favors "big" science that collects massive amounts of data free of hypotheses. We favored "small" science that was hypothesis-driven, a philosophy that has proved its value. Thus far big science justified by the philosophy of systems biology has yet to prove its mettle and much of the "omics" data gathered in the absence of a concept is likely to be lost in archives.

Consequently, I feel that it would be more productive for me to use this essay to illustrate by example how I believe that we should think about a biological system, in particular, the immune system. "How to think" is a nagging question rarely considered by scientists and when addressed, is generally viewed as arrogance. As living systems are the product of evolutionary selection, our analysis of them must be based on evolutionary principles. The artificial immune systems so popular with bioengineers and computer modelers tell us how the immune system might work and are sometimes a source of ideas, but if you wish to know how the immune system does work, then valid theories based on evolutionary principles should guide the computer modeling. A computer is a tool that must be told how to think before it can tell you what to think. Given the pace of the growth of today's technological wonderland, in which the brain and the computer are in an interactive competitive evolution, it might be valuable to illustrate "how-to-think" while the brain is still useful or rather before the brain becomes obsolete.

I would like to pick the self–nonself discrimination for my illustration because it has been a quagmire of semantics, analogies, and all-too-profound conceptualizations, not particularly illuminated by the contributions of historians, philosophers, and cabalistic immunologists.

The self (S)-nonself (NS) discrimination has a logic

Defining self

It does not clarify to complain about the terms self and nonself. They have found their way into the scientific literature such that their usage is directly or indirectly required in order to communicate. It is by far better to define them than to try to invent competing terms, although I have tried unsuccessfully [1, 2]. This is not simple and must be approached stepwise and cautiously because self and nonself are, in fact, defined by the very same immune system that we are trying to interrogate.

The output of the immune system is biodestructive and ridding. When the target is a pathogen, this is obvious. When the target is autogenously generated (e.g., products of cell necrosis, denatured or enzymatically inactivated protein), then this ridding function has salutary consequences like improved wound healing or regulation of inflammation. This housekeeping role should not be a source of debate; the output, biodestructive and ridding (complement lysis, cytotoxicity, phagocytosis, chemical warfare) is evolutionarily selected because it is salutary for the host, whether the target is a pathogen, autogenous waste, or a traumatized tissue.

A biodestructive and ridding effector output requires a way to distinguish what is to-be-ridded (nonself) because it tends to be lethal from that which is not-to-be-ridded (self) because ridding of it would be debilitating; this is what we mean by a self–nonself discrimination. The effector mechanisms have the potential to be just as harmful for the host as they are for pathogens. As a minimum, they must be directed with specificity and regulated in magnitude. Whatever else immune-related cells might do is not relevant; it is the biodestructive and ridding effector activity that demands a self–nonself discrimination.

Self is selected to function in the physiology of the host, *not* to escape recognition by the immune system. The immune system is selected not to attack self. Revealing the mechanism of this discrimination is our goal.

Self is defined by the immune system, not by the immunologist, the philosopher, or the historian. How the immune system learns this definition is a question for scientific investigation, not poetics. Scientific investigation begins with a consideration of validly competing theories and their experimental tests. However, any theory that is envisaged as reflecting the real-world immune system must be based on evolutionary thinking. So let us stop to give this thought a background.

The two recognitive repertoires

All free-living organisms have germline-selected recognitive repertoires coupled to ridding effector mechanisms

that protect them against parasitism. As the evolutionary selection pressure operating between pathogen and host is interactive, in most cases neither wins, and their relationship approaches asymptotically an *entende cordiale*. Each paratope of the host recognitive repertoire is germline-selected to recognize an epitope that is common to as many different pathogens as possible, yet absent from the self-of-the-species (i.e., the mating pool). The term "innate" is used to describe this germline-selected immune system.

As the interactive evolution between hosts and pathogens progressed there came a point when the germline evolution of the eukaryotic host to defend itself became too slow to match that of the prokaryotic pathogenic universe to escape the defense. At this point the selection pressure on the host shifted from germline to somatic, which was only possible because the host was multicellular. The effector outputs of the innate system were adequate by evolutionary criteria because they do not make a self–nonself discrimination; pathogens escape most frequently from recognition by the innate repertoire. A key component in this escape was the appearance of monomeric toxins produced by most bacterial pathogens and without which they are essentially controlled by the innate system or are harmless. The solution of evolution was to generate somatically a random repertoire using, in most cases, the innate recognitive system as a base or substrate and to couple this repertoire to the same biodestructive and ridding effector mechanisms as those used by the innate system. This was the only way to anticipate for recognition any new shape or epitope appearing in the pathogenic universe.

A word on the selection pressure is helpful here. In order for an epitope to be interactively selective on the germline it must remain unchanged over many generations; the epitope can act as a selection pressure only as long as it remains invariant. A protein epitope expressed by a viable pathogen escapes recognition too easily; a carbohydrate epitope is essentially invariant because in order to change its specificity, the specificity of a synthetic enzyme must be changed and that is sufficiently rare so that germline-selection can track it. Protein epitopes no longer associated with cells (secreted or products of necrosis or denaturation) are sufficiently invariant to act as selective determinants and, in fact play a role in maintaining a germline-selected housekeeping function.

The somatic diversifying mechanism operates on the germline-selected innate repertoire in a variety of ways dependent in part on the species [3]. Its function is to generate a random recognitive repertoire, one that is random with respect to the recognition of self and nonself epitopes. Species that express this somatically generated random recognitive repertoire are said to have an "adaptive" immune system.

All free-living organisms have "innate" systems; only vertebrates have "adaptive" systems. The pathogenic universe of vertebrates is such that mutations which inactivate the "adaptive" system results in their death by infection. The "innate" system is inadequate to protect against the infectious world selecting on vertebrates because its recognitive repertoire is too limited.

The germline-selected or innate recognitive repertoire is blind to the self-of-the-species. Individuals expressing only innate systems will accept grafts from other individuals of the same species. By contrast, individuals with adaptive systems function blind to the self-of-the-individual and reject as nonself, grafts from other individuals of the same species.

A meaningful definition of the self–nonself (S–NS) discrimination

The S–NS discrimination is the mechanism by which the adaptive immune system sorts its paratopic repertoire. It does this by deleting those specificities (anti-self) which, if expressed, would debilitate the host leaving as a residue those specificities (anti-nonself) which, if not appropriately expressed, would result in the death of the host by infection. The necessity to sort the repertoire (the S–NS discrimination) is the sole selection pressure for the specificity of paratopes. A theory of the self–nonself discrimination is then a theory of how the somatically generated repertoire is sorted.

What are the requirements of a theory defining such a sorting mechanism?

It might be well to begin by pointing out that evolution only selects to "adequacy". Perfection, which is a human value, is not selectable. Consequently, there is an acceptable frequency of hosts that succumb to infection on the one hand and to autoimmunity on the other hand. These frequencies become unselectable when they are no longer limiting to the procreation of the species. We, as humans, wish to understand the immune system so as to be able to manipulate it in a way that approaches perfection. This effort defines clinical immunology.

As what is self for one individual is nonself for another, and as the unsorted somatically generated paratopic repertoire is random with respect to self and nonself, the mechanism must involve a somatic learning process. Learning means that the response of the system depends upon its previous experience with respect to that antigen; it is a historical process. Any theory that ignores or fails to deal with the requirement for a learning or historical process is justifiably put aside.

At this point we must distinguish two terms, "*unresponsiveness*" and "*tolerance*". In the present context, both terms refer to epitope-specific events only. The experiment is to manipulate an animal with respect to an antigen to which it normally responds, in such a way that it becomes unresponsive specifically to that antigen. This observation is now conceptualized as a theory of the mechanism by which the animal normally learns to become unresponsive to self. We refer to the unresponsiveness defined by the theory as "tolerance". In sum, "unresponsiveness" is the observation; "tolerance" is the extrapolation of the observation to a theory of how the S–NS discrimination is accomplished. There are many ways to render an animal unresponsive; there is only one way to render it tolerant, a conclusion that evolution learned the hard way.

A theory of the sorting of the adaptive repertoire (i.e., the S–NS discrimination) then is the proposed mechanism for (1) a somatic learning process and (2) the signaling consequence for the cell, of the paratope–epitope interaction. Logic goes a long way here.

The associative recognition of antigen (ARA) theory

As the paratope has no way of knowing whether it is anti-S or anti-NS, the sorting of its repertoire requires the prior sorting of the epitopic universe into self and nonself [4, 5]. The adaptive immune system of each individual learns whether an epitope is self or nonself. More precisely, the immune system learns which epitopes are self; the residue is perforce defined as nonself-epitopes. However, a key point must be kept in mind. Paratopes recognize epitopes, not antigens. Antigens are collections of linked epitopes. The linked set of epitopes (an antigen) can be all self, all nonself or a combination of self and nonself. Roughly 10% of antigens are in this latter category. This has consequences!

Given the above, what are the steps in formulating a theory for the somatic learning process that sorts the paratopic repertoire?

The sorting of the repertoire (S–NS discrimination) is logically translated into a decision process by single cells which are born with two pathways open to them upon interaction with antigen, inactivation or activation. Activation is either an effector state or an intermediate on the pathway to becoming an effector. Let us refer to these naive/virgin antigen-responsive cells as *initial state* or i-cells. The two pathways (see Fig. 1) can be configured either as a fork-in-the-road [4, 6], or as sequential [7] (i.e., born inactivatable-only on interaction with antigen and over time differentiating antigen-independently to activatable-only on interaction with antigen).

The two models are not mutually exclusive as we will see but operate at different levels in the S–NS discrimination. The sequential model is a necessary adjunct of the fork-in-the-road model, in essence, predicted by it.

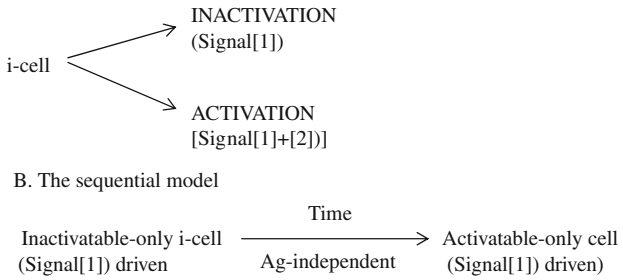

Fig. 1 The two configurations of the inactivation-activation decision

Two pathways: two signals

An inactivation–activation decision by an i-cell requires two signals as a matter of logic. The interaction of an epitope with the cell-receptor is referred to as Signal[1] and is inactivating. This is independent of whether the epitope is self or nonself. A second signal (Signal[2]) delivered to the i-cell receiving Signal[1] is required to tell the cell if it is interacting with a nonself-antigen (antigen, not epitope). It is essential to appreciate that inactivation is mediated epitope-by-epitope; activation is mediated antigen-by-antigen. Signal[2] defines the nonself-antigen as one that possesses a nonself-epitope. The requirement for an activating Signal([1 + 2]) assures that no cell can be activated that, in principle, could not have been inactivated.

Signal[2] is unique

The above highlights the two properties of the effector cell delivering Signal[2].

1. The effector, known experimentally to be a T helper (eTh), a very poor but entrenched term, must have had its repertoire presorted (i.e., made an S–NS discrimination) before it can deliver Signal[2]. The sorted repertoire of the eTh must be and is known to be anti-nonself.
2. The eTh interacting with one epitope on an antigen can only deliver Signal[2] to an i-cell interacting with another epitope derived from that same antigen. This is referred to as associative (linked) recognition of antigen (ARA). Activation is the first step on the pathway to effectors and is part of the process that determines and regulates the effector class. The defensive mechanisms are directed against antigens, not epitopes. A coherent and independent ridding response to antigens requires that the cells receive Signal[2] via interactions with epitopes that were structurally part of the given antigen (ARA). This

establishes the relationship between what is to be ridded and what is to be activated. The inactivation step which purges anti-self is the key element in the self–nonself discrimination resulting in a host tolerant of self.

The somatic learning of what is self

Now we are in a position to consider the learning mechanism which is a subject of debate [8, 9]. The i-cells must arise during ontogeny in the presence of all "self" and no "nonself" under conditions where they are inactivatable-only upon interacting with the available epitopes. The i-cells are potentially inactivatable and activatable but when this developmental time window is open, the effective absence of Signal[2], that is, an insufficiency of effector T helpers, results in all interactions with epitopes being inactivating (Signal[1]). As long as the self-epitope persists, even when the system becomes responsive, tolerance of that self-epitope is maintained because the iTh anti-self are inactivated as they arise. During this ontogenetic period, establishing and maintaining tolerance is the same. The cells that do not interact with an epitope during this inactivatable-only period of ontogeny are defined as anti-nonself. There are several problems that need to be addressed.

1. All Self is in "tres partes divisa est".

The TCR recognizes as ligand [PR], peptide (P) complexed to an MHC-encoded restricting element (R). There are two categories of [PR] dependent on the class of R. Class I (RI) presents peptide to cytotoxic T cells or, more general, $CD8^+$ T cells. Class II (RII) presents peptide to helper T cells or, more general, $CD4^+$ T cells.

The BCR recognizes as ligand a shape patch (determinant, D) on the surface of intact molecules.

The consequence is that there are three families of self-ligand, [Pself-RI], [Pself-RII], and Dself. Any autogenously generated component not presented as a ligand when the developmental time window is open (insufficiency of eTh) would be defined as nonself, if it were presented after the window closes and the system is responsive (sufficiency of eTh). In other words, a self-component that appears de novo as a ligand when the system is responsive (e.g., postnatally) cannot be distinguished from nonself.

There is an asymmetry to note. Any component that is not presented as [P-RII], the helper T cell ligand, but is presented as [P-RI] or D, is obligatorily tolerigenic for $CD8^+$ T cells and B cells, respectively. As RII is expressed on specialized antigen-presenting cells (e.g., dendritic and B cells), which may not encounter and process many self-antigens, whereas RI is expressed on all cells, there may be a subset of self, in essence tolerigenic-only for $CD8^+$ T cells. B cells interacting with surface self-components may receive Signal[1] but be unable to extract them for processing; hence, these components would be tolerigenic-only for B cells.

2. As self-components are selected to function in the physiology of the organism, it is to be expected that some of them would be expressed late, that is, when the immune system is responsive. As delayed expression self would be treated as nonself and become an autoimmune target, only one selectable solution presented itself. The delayed expression self-components had to be presented as [Pself-RII] to initial state T helpers (iTh) while the developmental time window is open (i.e., the system is unresponsive due to a lack of eTh). This results in the deletion of iTh anti-S as they arise establishing tolerance to the delayed expression self. The presentation as [Ps-RII] has to occur well before the self-component is expressed as a functional physiological element.

Such a situation appears to obtain (discussed in [10]). There is a family of peripheral self-components that are ectopically expressed in thymus as mRNA (presumably also as [Ps-RII]) under the control of the transcription factor, *Aire*. *Aire* knockouts display a variety of autoimmune disorders at variable periods after birth (in mice 2–3 weeks). This implies that the ectopic expression in thymus is tolerigenic for iTh. If *Aire* function is experimentally inactivated after the peripheral self-component is expressed, autoimmunity is not manifested [11] as is predictable because the appearance of the delayed peripheral self-component expressed as [Ps-RII] maintains the tolerance that was established while the developmental window was open.

3. The immune system is not just a bag of cells but has important organ structures. These permit cell–cell interactions to take place at functional rates and allow for an orderly controlled process of differentiation. For our discussion, the thymus is central and need only be distinguished from the periphery. The thymus is the organ in which iT-cells arise, are differentiated into $CD8^+4^-$ cytotoxic T cells (iTc) and $CD4^+8^-$ helper T cells (iTh) and undergo some sorting of their repertoires, a process referred to as negative selection (sorting) against anti-self. The thymus is a space that lacks eTh, a key point; it is a tolerigenic enclave. The repertoire of the iT-cells that leave the thymus to the periphery is divisible into three categories:

- a major population anti-nonself, the protective repertoire
- those iT-cells anti-peripheral self not ectopically expressed in thymus but expressed only in the periphery while the window is open (peripheral tolerance is required)

- iT-cells anti-self of all types that "escape" inactivation.

This latter category needs discussion as the term "escape" that fills the literature is misleading. The delivery of Signal[1] cannot be instantaneously inactivating. A paratope–epitope interaction is rapid compared to cell–cell interaction required to deliver Signal[2]. If Signal[1] were instantaneously irreversible, no cell would be activatable. This results in a steady-state population of anti-self cells leaving the thymus into the periphery as well as a population encountering peripheral self for the first time. This steady-state population of rescuable T cells anti-self on the pathway to death is referred to as the autoimmune boundary [12–14]. It is an unavoidable concomitant of the inactivation–activation decision. The level of cells in this boundary is a function of the length of time that it takes Signal[1] to become irreversible. Evolution walked a tightrope here. If the time it takes to become irreversible were too short, too few cells anti-nonself could be activated and the organism would die of infection. If it were too long, the accumulated high level of anti-self cells would make the frequency of autoimmunity too high for reasons we will analyze next. The time period for Signal[1] to become irreversible was set by evolutionary selection such that the protective level was adequate and the frequency of autoimmunity acceptably low.

4. Signal[2] delivered by an eTh is required on logical grounds for the activation of all i-cells, including iTh itself. This posed a chicken-and-egg problem that for years was and probably still is viewed as the Achilles heel of the ARA model. If eTh anti-nonself are required for the activation of iTh anti-nonself, where do the primer eTh anti-nonself come from? This is a question with many ramifications yet it has not been addressed experimentally. Although a solution was proposed over 25 years ago [15] and repeatedly updated [13, 14, 16–18], it still remains ahead of its time and is ignored.

The primer eTh must be derived by an antigen-independent pathway. If the rate of inactivation of iTh anti-self is rapid compared to the rate of antigen-independent differentiation to primer eTh, then the primer population will be effectively anti-nonself meaning that the presence of primer eTh anti-self will be too low to initiate an autocatalytic or self-sustaining anti-self response (Fig. 2). Thus we invent the essence of the sequential model (Fig. 1) introduced earlier. The iTh anti-self are purged in thymus and in the periphery while the developmental time window is open (insufficiency of eTh). The residue iTh anti-nonself undergoes the slow differentiation antigen-independently to provide a priming level of eTh anti-NS.

What about the iTh anti-self in the autoimmune boundary?

If the level of iTh anti-self becomes too high then the primer eTh derived by the antigen-independent pathway can reach a level resulting in a self-sustaining autoimmunity. This pathway can also be mobilized in recovery from lymphopenia (e.g., neonatal thymectomy or restoration of an immune-blank animal) because the iTh rapidly dividing to restore homeostasis resets the balance between anti-self going to death and to primer eTh. The observed result is autoimmunity.

5. I introduced the term "autoimmunity" into our discussion and this needs precision to distinguish it from "immunopathology".

"Autoimmunity" results from the breaking of tolerance to a self-component with debilitating consequences. It is an antigen-specific attack that can involve one or many self-components, depending on the way in which tolerance was broken. In order for the immune system to be under evolutionary selection not to attack self, the attack must have

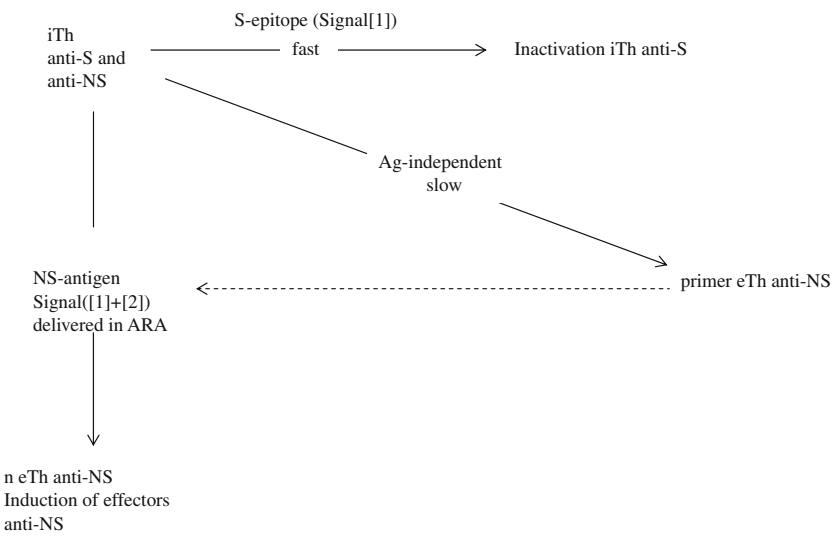

Fig. 2 The nonself antigen independent pathway to primer effector T helpers (eTh) anti-nonself

debilitating consequences. Housekeeping or the ridding of autogenous waste is not a selective pressure on the mechanism for the sorting of the repertoire. Autogenously generated garbage is not self to the immune system; it is nonself.

"Immunopathology" arises as a consequence of a normal effector response to nonself (or to self) during which the biodestructive and ridding mechanisms spill over to attack innocent bystanders. This problem, which is not our concern here, is regulated in a variety of ways that have nothing to do with the S–NS discrimination.

Any manipulation that results in a pathology or that arises spontaneously must distinguish autoimmunity from immunopathology, if it is to be interpretable. Much of what immunologists refer to as autoimmunity could well be immunopathology and not relevant to the S–NS discrimination.

6. Is inactivation resulting in tolerance due to negative sorting (Signal[1]) or to positive sorting (suppression)?

Regulatory T cells include both T helpers (Th) and T suppressors (Tsu). The use of the symbol Tregs to mean Tsu is just one more example of sloppy terminology. However, Tregs $\cong$ Tsu are the rage today as postulated mediators of the S–NS discrimination. As this runs in direct contradiction with the model I have developed thus far, I must stop to deal with them.

The models using either negative or positive sorting to purge anti-self from the repertoire are symmetrically converse. If negative sorting obtains, then Signal[1] is inactivating (epitope-by-epitope) and Signal[2] delivered by an eTh in ARA is activating (antigen-by-antigen). If positive sorting obtains, then Signal[1] is activating (epitope-by-epitope) and Signal[2] delivered by an eTsu in ARA is inactivating (antigen-by-antigen). There is no a priori reason for eTh to play an activating role under a positive sorting model. However, given the strong evidence that eTh plays an activating role, most supporters of positive sorting would logically propose no signal via the TCR/BCR on binding ligand, and have an activating signal delivered by an eTh and a competing inactivating signal delivered by an eTsu, both delivered in ARA to the i-cell. As this latter configuration of positive sorting is oft presented in lectures but is absent in print, I will not discuss it further.

The positive sorting model faces four failures [19–21]:

- The eTsu (eTregs) must have a repertoire that is sorted to be anti-self (contrary to fact).
- Nonself antigens (pathogens) that share epitopes with self will be tolerigenic-only (a lethal situation).
- Activation, epitope-by-epitope, and inactivation, antigen-by-antigen, would make coherent and independent regulation of the effector class impossible. Effector mechanisms rid antigens not epitopes.
- The antigen-receptors, TCR/BCR, are themselves self-components, which, if targets of eTsu, would turn off the entire immune response.

I, therefore, conclude that suppression (positive sorting) is ruled out as the mechanism of tolerance. It of course can be manipulated experimentally to establish unresponsiveness.

If an eTsu anti-Pns population is isolated experimentally and added to a responding system anti-Pns, the antigen-specific response will be inhibited. This can be used clinically to downregulate the response to a self or nonself antigen. The normal role of eTsu $\cong$ eTreg is to regulate the magnitude of the effector response anti-NS. In the absence of feedback regulation, immunopathology would be manifest. As the repertoire of iTsu is sorted to be anti-NS, regulation of a self-specific autoimmune response would require that tolerance in the iTsu population be broken. If the iTsu repertoire were unsorted, it could not contribute to a S–NS discrimination. If it were sorted to be anti-S, then it cannot regulate the magnitude of the normal anti-NS effector response. Of course, pathogens that share epitopes with self would escape immune attack. If eTsu are sorted to be anti-NS, then they play no role in the sorting of the repertoire (S–NS discrimination). Thus this latter appears to obtain.

7. The role of the antigen-presenting cell (APC) is an unsettled question of mechanism.

In order for an eTh to deliver Signal[2] to an iT-cell (T–T interaction) both cells must interact with an APC that presents processed peptides (P) from the antigen on the appropriate MHC-encoded R-elements, RII for eTh and RI/RII for the defensive iT. As the APC cannot make a S–NS discrimination, so-called "costimulation" cannot be the source of Signal[2]. Further, there is no experimentally revealed pathway of uptake and processing of antigen that would permit an eTh–APC–iTh/c interaction of ARA. In the absence of an eTh driven Signal[2], the APC-iTh/c interaction is tolerigenic.

Thus, two questions remain to be answered. First, how is signaling ARA mediated during the eTh–APC–iTh/c interaction? Second, how is the uptake of monomeric nonself protein to which the innate system is blind, accomplished? This latter is a problem best appreciated when one considers that a bacterial toxin is lethal at ng/ml concentrations and it is present in bodily fluids which have close to 100 mg/ml monomeric self-protein (e.g., serum albumin). Unspecific pinocytosis is not a solution. A few immunologists have appreciated and dealt with these questions, so I will cite their analyses and leave the problem here as food for thought [4, 5, 22–25].

8. How does the immune system learn that autogenously generated waste is a nonself target?

During ontogeny, cells die by apoptosis and the resultant granules are ridded by the innate system without their contents encountering the newly arising antigen-responsive i-cells. When the immune system becomes responsive, in most species postnatally and the cells are subjected to a variety of traumas from pathogenesis to ageing, they die by necrosis, spilling their contents as nonself targets. This autogenous garbage has acted as a selective pressure on the germline as it is recognized both by the innate system and by germline-encoded antibodies of the adaptive system. The immunologist may refer to this autogenously generated garbage as self, but to the immune system it is nonself and in no way does it present a challenge to the ARA model's definition of self.

The competing views

I have tried thus far to avoid facing who said what and when because I wanted to get what in my mind is a clear conceptualization of the S–NS discrimination free of priority and hindsight reinterpretation of data considerations. The history of thinking about the problem has been comprehensively reviewed [26] and the reader is referred to it.

Here I would like to face the more recent claims of a "theory" when none exists, as well as the theories that I would argue do not deal with the sorting of the adaptive repertoire (i.e., the self–nonself discrimination).

The black box theories of the S–NS discrimination

I view the ARA model of the S–NS discrimination as a default position. After evaluating all theories and eliminating all but one, the latter becomes a default theory until, of course, a new theory is proposed that supersedes it. This is what I would like to examine now.

The response of the adaptive immune system can be operationally divided into two sets of decision processes:

Decision 1 The sorting of the repertoire (i.e., the S–NS discrimination)

Decision 2 The choice of effector class and the regulation of its magnitude

Decision 1, as discussed here, is a somatic learning process which results in a sorted repertoire anti-nonself.

Decision 2 is initiated by activation of the i-cell that then undergoes a series of steps of differentiation to become an appropriate effector. These steps are regulated by germline-selected mechanisms [27].

The experimentalist looks at the effector output of Decision 2 after manipulating and challenging an animal with antigen. The observation of unresponsiveness at the level of Decision 2 is then conceptualized as a mechanism of tolerance (i.e., the S–NS discrimination). A valid theory of tolerance at the level of Decision 2 would make the sorting of the repertoire neither necessary nor sufficient to account for a S–NS discrimination. Thus far none exists.

Consider a classic experiment [26] in which fetal or neonatal mice (H-2^a) are injected with splenic lymphocytes (H-2^b). At a later time they are grafted with H-2^b skin which is specifically accepted. Control mice (H-2^a) reject H-2^b skin. What can be concluded from this information concerning the mechanism for establishing "tolerance" (i.e., the S–NS discrimination or Decision 1)?

If the acceptance of the graft is due to suppression by the H-2^b lymphocytes of the H-2^a effector attack on the H-2^b skin, then clearly the acceptance of the graft is a property of Decision 2 and nothing can be said about the establishment of tolerance. If acceptance were due to negative sorting initiated by the antigens expressed on the H-2^b lymphocytes, an extrapolation to tolerance would be possible but it would remain unsettling to have to assume that lymphocytes would express all the target self-antigens present in skin. Besides, one wonders why a neonatal skin graft itself isn't accepted. Lastly, graft-acceptance cannot be equated to unresponsiveness. A response in an ineffective humoral class would go undetected by an assay of graft acceptance. I, therefore, question whether this class of experiment deals with tolerance at all. It possibly tells us about the role of suppression as a feedback mechanism at the level of Decision 2, the rejection of a graft being an assay of an effective effector response.

This black box view of the S–NS discrimination has spawned a family of theories that can be summarized as being based on germline-selected mechanisms at the level of Decision 2. This, in itself, rules them out as models for the S–NS discrimination.

Those theories that are based on the requirement for a signal from the pathogen–host interaction in order to initiate a response are clearly theories of Decision 2. The fact that the effector response to some nonself-antigens is dependent on such properties as danger [28, 29], pathogenicity [30, 31], cytopathicity [32], localization [33], integrity [34, 35], ecoimmunity [36], morphostasis [37], tuning [38], etc. does not imply that the failure to respond to self-antigens is due to their lack of these properties. Further, given that what is self for one individual is nonself for another, it should be clear that there is no physical or chemical property of an antigen that can be used by the immune system to distinguish self from nonself as classes. Lastly, the recognition of these properties is germline-selected, whereas the sorting of the repertoire (Decision 1) is a somatic learning process. The casting of these theories as mechanisms of the self–nonself discrimination is inappropriate. At best, they are elements that may be used to develop a theory of the regulation of class (Decision 2).

They are, therefore, theories that are non-competing with the ARA model.

The theory that the unsorted repertoire is anti-self

The postulate that the unsorted repertoire was selected in the germline to be anti-self [39, 40] achieved great popularity although I have always viewed it as not just erroneous, but rather as irrational. There is no way to select for either a silent or a debilitating recognition of self in the germline. In any case, the specificity of the innate system's repertoire is anti-nonself. A mutation in the germline to recognition of self would be lethal to the offspring of a mating that included the target self-component. The example of allele-specific recognition of MHC-encoded restricting elements by the TCR is misplaced, and the citing (see table in Ref. [41]) of housekeeping targets as self needs rationalization (both discussed in Ref. [17]).

The cognitive paradigm as a denial of a need to sort the repertoire

The cognitive paradigm is viewed as a challenge to "clonal selection theory" [41, 42]. As I don't know what is meant or encompassed by clonal selection theory, I will translate it into the ARA model by postulating that:

1. only i-cells that recognize and are signaled by interactions with appropriate ligands undergo a Decision 1 step of inactivation or activation (cellular selection); and
2. autoimmunity is acceptably minimized by the purging of anti-self, epitope-by-epitope, by negative sorting and the activation of anti-nonself, antigen-by-antigen, by an eTh anti-nonself functioning in ARA. In other words, the random repertoire is sorted by deletion of anti-self leaving as a residue anti-nonself to protect the individual (cellular selection theory).

Using this translation I will equate the ARA model with clonal (cellular) selection theory. The comment that there are immunologists who believe that "the primary function of the immune system is to distinguish between the self and the foreign" is as attention grabbing as it is questionable. If anything, it is believed that "the primary function" is to protect the organism from infection without autoimmunity.

The ARA model or "clonal selection paradigm" does not "identify autoimmune disease as an accident of self-recognition originating from a random mutation….or from failure to delete" an anti-self clone. As discussed here, it is somewhat more subtle.

Autoreactivity is *not* autoimmunity and self is *not* anything autogenously generated. Autoreactivity does not negate the clonal selection paradigm; it negates the definition of "self" used by some immunologists, not the definition used by the immune system. It is because what is "self" is somatically learned by the immune system that autoreactivity to autogenous nonself is to be distinguished from autoimmunity to self. To argue that "autoimmunity is not an aberration but is a property of all healthy immune systems [41]" is no more than debatable semantics.

In any case, how does the cognitive paradigm solve these "flaws in the clonal selection paradigm". As the "flaws" are slowly transformed during the discussion into "incompletenesses", the challenge to the ARA model can be viewed as a difference of opinion as to what it was designed to explain.

The clonal selection theory is argued [41] to have the following flaws because it fails to encompass, require or explain [41]:

- antigen processing and presentation
- structure and function of MHC
- restrictions in TCR/BCR gene usage
- superantigens
- cytokine networks
- anti-idiotypes.

Clearly clonal selection theory was not designed to explain everything, only the sorting of the repertoire. Accepting this incompleteness, we would expect the proposed cognitive paradigm to solve these flaws. Three problems are defined, which if solved, would account for the above flaws:

- the signal/noise problem (focus)
- the context problem (context)
- the response problem (response).

The "focus" problem is Decision 1, while "context" and "response" are problem of Decision 2 not germane for our discussion of the sorting of the repertoire that deals with antigen-specific interactions only. Yet the cognitive paradigm which is not a theory but a milky description of Decision 2 makes comment irresistible.

The cognitive paradigm provides the framework for a conceptualization referred to as the "immunological homunculus". This framework derives from the idiotype network era, the history of which has been thoughtfully discussed [43].

> The immunological homunculus is an internal image of the self acquired by early recognition of self antigens…. The self image is, in fact, composed of committees of T and B cells that deal with dominant self antigens [42].

Unclear is whether the "internal image of self" that is somatically "acquired by early recognition of self" operates in the presence of a random somatically generated or

an anti-self germline-selected repertoire or both. If the repertoire is germline-selected to be anti-self, then the assumption of somatic selection for "early recognition of self" is gratuitous. If the repertoire is random and somatically generated, then the assumption of somatic selection for self recognition implies that the repertoire is sorted to amplify anti-self recognition rather than delete it raising the question of how autoimmunity is controlled. The proposed selection pressure for the homunculus is that the immune system is "bombarded by self-like foreign molecules" which are presented "in the context of infection" with the result that "autoimmunization cannot be avoided". If "autoimmunization" means "housekeeping", it is an aside for this discussion. If it means autoimmunity as a pathology then we have a problem. The importance of confronting competing theories, in this case, the ARA model becomes obvious.

How does the homunculus solve this problem of avoiding autoimmunity?

> Natural autoimmunity is benign because the immunological dominance of the major self-antigens comprising the homunculus is encoded by two committees of cells: naturally autoimmune T and B cells and their anti-idiotypic regulatory cells [42].

Implied is that there is somatic selection for the expression of autoimmune disease directed at a family of dominant self-components, accompanied by somatic selection for a suppressive mechanism to limit autoimmune disease based *not* on recognition of the dominant self-antigens but on the idiotypes of the receptors that recognize them. As the idiotypes are themselves self-epitopes, this proposal faces infinite regress.

Such a double bluff proposal requires a detailed, stepwise rationalization of a pathway or mechanism to accompany the bare assertion. Lastly, given a random repertoire, the non-dominant self-antigens are ignored or do not participate in a somatic S–NS discrimination. As I consider this view to be irrational, there is no way to refute it.

How does the ARA model solve this problem of avoiding autoimmunity?

The individual is tolerant of all self-epitopes at the level of the autoimmune boundary. A foreign antigen that shares epitopes with self (i.e., "self-like foreign molecules") can break tolerance and establish autoimmunity. What does it take to do this?

The self-antigen competes with the cross-reactive nonself-antigen to prevent the breaking of tolerance. In order to override this inhibition, the eTh anti-self must be induced to a self-sustaining or autogenerative level. If the cross-reactive nonself-antigen is close to self, it will be unable to induce a high enough level of eTh anti-self and may even establish unresponsiveness to the nonself-epitopes. As the cross-reactive nonself-antigen becomes more and more foreign to self, it has an increasing probability of breaking tolerance. If the cross-reactive nonself-antigen is ridded rapidly enough, the eTh anti-self never reach a high enough level to break tolerance and the system remains tolerant of the self in question.

In sum then, the cognitive paradigm is just another way of stating that an immune response is complex. Complexity of the system does not negate any theory that is used to explain a segment of the response. How we deal with complexity is a fundamental question that the cognitive paradigm courageously, albeit unsuccessfully, addresses. Modularizing the system into manageable units like the generation of the repertoire, the sorting of it and the appropriate coupling of the sorted repertoire to effector mechanisms, is another approach. Modularization, unlike the "cognitive paradigm", makes it not only possible for the computer but also for the immunologist to understand the system.

The scholars examine "Self": a clarification or a turbidification?

A catchy way to describe the immune system's definition of "self" is that it is "prior and persistent" whereas "nonself" is "posterior and transient". The ARA model based on a somatic learning mechanism makes this description clear.

Self to the immune system is a family of ligand/epitopes present when the i-cells of the system arise in developmental time and, for which interaction with ligand is inactivating. Self remains self as long as it persists as a ligand. No historical account or philosophical perspective or analysis by analogy with the nervous system can change the immune system's definition of self. The selection pressure defining self is that an attack on self is debilitating and what must not be attacked is somatically learned.

The historians and philosophers have latched onto "self" as a concept that can be rejected or ignored based on three arguments.

1. The existence of allele-specific recognition of the MHC-encoded restricting element presenting peptide. This is *not* an example of a self-marker involved in a self–nonself discrimination. It is the recognition of the peptide not the MHC that is the ligand for the sorting of the repertoire. When the MHC is acting as a presenter of peptide, it is not functioning as a self-ligand [17].
2. Theories based on danger, pathogenicity, integrity, etc. require a germline-selected recognitive system which, to the extent that they are valid, automatically places

them at the level of Decision 2, the regulation of effector class, not Decision 1, the sorting of the repertoire.

3. To the recognitive repertoire (BCR/TCR), self is an epitope, not an antigen. A self-antigen is a linked collection of self-epitopes. A nonself-antigen can share epitopes with self or be made up of nonself-epitopes only. The immune system must and does deal with these cross-reactive nonself-antigens to keep autoimmunity at an evolutionarily acceptable level, as discussed earlier.

The philosopher debates the historian

Howes [44] as a philosopher posits that the "self" as it is treated in philosophy and immunology, "must have some stable core, essence or foundation that enables it to be reidentified through time".

For the adaptive immune system, what is self is learned during ontogeny and is individual specific. As self is prior and persistent, it may be described as "reidentified through time". The concept of "a stable core, essence or foundation" is valid but trivial when treated as nothing more than "persistent". What is missing is the requirement that the immune system learn somatically what is the self-of-the-individual. This is certainly also true for the nervous system. It is not the persistence of self that is paradoxical. It is the somatic learning of what is self that needs attention because that is how the immune system defines self.

For the innate immune system, what is self is "learned" during germline evolution and is as a minimum, species-specific (species being the mating pool). The repertoire of the innate system is germline-selected (sorted) to recognize the pathogenic universe with which it is interactive to the exclusion of the self-of-the-species.

Tauber [45] as a historian latches onto all views of self from its not being a substance, but rather a process, to its *not* being a problem that the immune system faces. The one view that he rejects by its absence from discussion is the definition of self derived from the ARA model. This permits the "self" to be dethroned as no more than a useful metaphor, certainly of no import for immunologists. Howes [44] on the other hand, sits on a fence, concluding that although "the self" is still a viable concept, it may no longer be of use to immunologists".

All anticipatory mechanisms to deal with unexpected stimuli, require the somatic generation of a recognitive repertoire that focuses the response on the specific target. Such repertoires are random with respect to what must and what must not elicit a response. This, in turn, requires a somatic learning or historical process. The ARA model at the moment is the only one that confronts the mechanism for the immune system. Models that deny the need for a somatic learning process or that bury the problem in complexity, or in semantics, or in analogies are destined for the archives of history [18] or the debates of philosophy [44, 45]. For the immune system, when the self becomes an autoimmune target, evolutionary selection operates to correct the mechanism that attacks it. The immunologist is the observer and the interpreter of the process. Self (or whatever else you wish to call it) is still an unavoidable element shaping an immune response.

Acknowledgments This work was supported by a grant (RR07716) from the National Center For Research Resources (NCRR), a component of the National Institutes of Health (NIH) and its contents are solely the responsibility of the authors and do not represent the official view of NCRR or NIH.

References

1. Cohn M, Langman RE (2002) To be or not to be ridded?—That is the question addressed by the associative antigen recognition model. Scand J Immunol 55:318–323
2. Cohn M (2002) The immune system: a weapon of mass destruction invented by evolution to even the odds during the war of the DNAs. Immunol Revs 185:24–38
3. Langman RE, Cohn M (1993) A theory of the ontogeny of the chicken humoral immune system: the consequences of diversification by gene hyperconversion and its extension to rabbit. Res Immunol 144:421–446
4. Cohn M (2007) Conceptualizing the self–nonself discrimination by the vertebrate immune system. In: Timmis J, Flower D (eds) In silico immunology. Springer, Berlin Heidelberg New York, pp 375–398
5. Cohn M (2005) The common sense of the self–nonself discrimination. Springer Semin Immunopathol 27:3–17
6. Bretscher P, Cohn M (1970) A theory of self–nonself discrimination. Science 169:1042–1049
7. Lederberg J (1959) Genes and antibodies. Science 129:1649–1653
8. Cohn M (2009) On the opposing views of the self–nonself discrimination by the immune system. Immunol Cell Biol 87:113–119
9. Cohn M (2009) The discussion with Jacques Miller: illustrating the limitations of pure empiricism. Immunol Cell Biol 87:435–437
10. Cohn M (2009) Why Aire? Compensating for late bloomers. Eur J Immunol 39:1–4
11. Guerau-de-Arellano M, Martinic M, Benoist C, Mathis D (2009) Neonatal tolerance revisited: a perinatal window for Aire control of autoimmunity. JEM 206:1245–1252
12. Cohn M, Langman RE (1990) The protection: the evolutionarily selected unit of humoral immunity. Immunol Rev 115:1–131
13. Cohn M, Langman RE, Mata J (2002) A computerized model for the self–nonself discrimination at the level of the T-helper (*Th-genesis*). I. The origin of "primer" effector T-helpers. Int'l. Int Immunol 14:1105–1112
14. Langman RE, Mata JJ, Cohn M (2003) A computerized model for the self–nonself discrimination at the level of the T-helper (*Th genesis*) II. The behavior of the system upon encounter with nonself antigens. Int Immun 15:593–609

15. Cohn M (1983) Antibody diversity 1983: some elementary considerations. In: Yamamura Y, Tada T (eds) Progress in immunology V. Academic Press, Orlando, FL, pp 839–851
16. Langman RE (1989) The immune system. Academic Press, San Diego
17. Cohn M (1992) The self–nonself discrimination: reconstructing a cabbage from sauerkraut. Res Immunol 143:323–334
18. Cohn M (2001) Logic of the self–nonself discrimination: principles and history. In: Cambrosio A, Moulin A (eds) Dialogues with selves. Historical issues and contemporary debates in immunology. Editions Elsevier, France, pp 53–85
19. Langman RE (1987) The self–nonself discrimination is not regulated by suppression. Cell Immunol 108:214–219
20. Cohn M (2004) Whither T-suppressors: if they didn't exist would we have to invent them? Cell Immunol 227:81–92
21. Cohn M (2008) What roles do regulatory T-cells play in the control of the adaptive immune response? Int Immunol 20:1107–1118
22. Mitchison NA, O'Malley C (1987) Three-cell-type clusters of T cells with antigen-presenting cells best explain the epitope linkage and noncognate requirements of the in vivo cytolytic response. Eur J Immunol 17:1579–1583
23. Bretscher PA (1999) A two-step, two-signal model for the primary activation of precursor helper T cells. Proc Natl Acad Sci USA 96:185–190
24. Cohn M (2009) How does the immune response get started? Cell Immunol 254:91–93
25. Cohn M (2006) Does the signal for the activation of T cells originate from the antigen-presenting cell or the effector T-helper? Cell Immunol 241:1–6
26. Brent L (1997) A history of transplantation immunology. Academic Press, San Diego
27. Cohn M (2008) A rationalized set of default postulates that permit a coherent description of the immune system amenable to computer modeling. Scan J Immunol 68:371–380
28. Matzinger P (2003) The real function of the immune system or tolerance and the four D's (danger, death, destruction, and distress)
29. Matzinger P (2002) The danger model: a renewed sense of self. Science 296:301–304
30. Janeway CA, Goodnow CC, Medzhitov R (1996) Immunological tolerance: danger—pathogen on the premises! Curr Biol 6:519–522
31. Janeway CA (1992) The immune system evolved to discriminate infectious nonself from noninfectious self. Immunol Today 13:11–16
32. Zinkernagel RM (2004) On 'reactivity' versus 'tolerance'. Immunol Cell Biol 82:343–352
33. Zinkernagel RM, Hengartner H (2004) On immunity against infections and vaccines: credo 2004. Scand J Immunol 60:9–13
34. Dembic Z (2000) Immune system protects integrity of tissues. Mol Immunol 37:563–569
35. Dembic Z (1996) Do we need integrity? Scand J Immunol 44:549–550
36. Nevo U, Hauben E (2007) Ecoimmunity: immune tolerance by symmetric co-evolution. Evol Dev 9:632–642
37. Cunliffe J (2006) Tissue homeostasis and immunity—more on models. Scand J Immunol 64:172–176
38. Grossman Z, Paul WE (2001) Autoreactivity, dynamic tuning and selectivity. Curr Opin Immunol 13:687–698
39. Jerne NK (1971) The somatic generation of immune recognition. Eur J Immunol 1:1–9
40. Jerne NK (1970) Generation of antibody diversity and self tolerance. In: Smith RT, Landy M (eds) Immune surveillance. Academic Press, Augusta, MI, pp 343–436
41. Cohen IR (1992) The cognitive principle challenges clonal selection. Immuol Today 13:441–444
42. Cohen IR (1992) The cognitive paradigm and the immunological homunculus. Immunol Today 13:490–494
43. Eichmann K (2008) The network collective, rise and fall of a scientific paradigm. Birkhauser, Berlin
44. Howes M (1998) The self of philosophy and the self of immunology. Perspect Biol Med 42:118–131
45. Tauber AI (1997) Historical and philosophical perspectives concerning immune cognition. J Hist Biol 30:419–440