Immunol Res (2017) 65:573–591
DOI 10.1007/s12026-017-8908-0

INTERPRETIVE SYNTHESIS REVIEW ARTICLE

Learning from a contemporary history of immunology

Melvin Cohn[1]

Published online: 13 February 2017

Abstract This essay is a selected aspect of the history of contemporary immunology seen from a "what can we learn" point of view. It is limited to the ideas and experiments from which we might draw a take-home message. The emphasis is on the convoluted pathway that was actually used by immunologists to arrive at understanding compared to the direct pathway that could have been used given the knowledge at that time. It takes the reader through the instructionist era of the 1940s to the present by stressing the elements of thinking most conducive to the arrival at a default understanding of the intact immune system. It is a personalized account because the author participated directly in the debates that led eventually to agreed-upon or default conceptualizations. Given this, a peek at the future is attempted as a test of the validity of a Cartesian or reductionist approach to arriving at simplification of complexity and at the maximizing of generalization. A reasoned guess (i.e., a theory) is the only way we have to understand the world around us.

Keywords TCR · BCR · Self-nonself discrimination · Dendritic cells · Concepts

Introduction

We as scientists avoid the question, how to think, for a variety of reasons, the most intimidating being that it might be viewed as arrogance. However, this should not be an insurmountable impediment as I have never met a scientist who thought that he/she was average. In any case, the subject is usually left to philosophers who do not deal with it as a pragmatic question. Consequently, I will pose a challenging invitation to my colleagues by considering the question. I will not attempt a rigorous philosophical analysis but rather use examples to illustrate why there is such a thing as how not to think about the immune system. Subtracting how not to think, somewhere in the residue is how to think about the immune system. Many of my colleagues would prefer to be wrong but viewed by the community as being right, rather than be right but viewed by popular vote as being wrong. To this end, they defend unto irrationality their ideas, as I will demonstrate by example. Erroneous ideas die in our community more often by neglect than by reason, certainly rarely by "I change my mind." Yet, a disproven valid theory is an important step in "understanding"; it should not be treated pejoratively.

An analysis such as this one is based on history. What we can do in the future is dependent on what we did in the past. This past is laden with right conclusions for the wrong reason, a curious fact. We will see that individuals can make probing or profound statements that they, themselves, do not understand. There are experiments totally lacking in rationale and even erroneously interpreted that become major steps forward in the field. And then, there are ideas discarded because they were cast in the wrong framework but in the right framework are precious. Lastly, aspects of immunology can become the subject for semantics, philosophy, and blind intuition in ways that tend to derogate productive scientific progress. Yet surprisingly, in the long run, all this matters little. In the end, there is always someone who looks at the overwhelming quantity of jumbled random observation and intuitive speculation, sees through the errors of the past, and tells us how to think. There are those who believe that computer simulation will deal with most of this, forgetting that, in order for computers to be helpful, they must be told how

✉ Melvin Cohn
cohn@salk.edu

[1] Conceptual Immunology Group, The Salk Institute, 10010 N. Torrey Pines Rd., La Jolla, CA 92037-1002, USA

to think before they can tell us what to think. Eventually, new rounds of observation are triggered that are translated into a default conceptual framework. This is what we call understanding.

Lastly, my colleagues who are serious historians constantly insist that their favorite scientist could not have known at the time that his/her thinking was contradictory or that a given experiment did not warrant the conclusion. I have therefore been very careful to consider the knowledge of the time and to attempt to extract a take-home lesson. Many historians treat science as a storybook from which there is no take-home lesson that can guide our thinking today. We disagree!

Let us begin with a statement of principle, i.e., the ground rules. There exists a set of observations that evokes our curiosity. They have been described in detail but are not understood, meaning that they await explanation at different levels of organization from molecular to ecological in a way that tells us what to expect in the future. This requires a theory that ties all this together. A valid theory must have the potential to be experimentally disprovable. This, in turn, requires that it suggest predictions that can guide experiment. There is nothing more misleading than, "I believe what I see!" in the absence of a theory. Magicians depend on I believe what I see when they saw the lady in half. Description is not equivalent to understanding. One can describe in extenso and understand little. The optimal situation is when we are faced with two or more equally valid theories that predict different outcomes. Experiment disproves one or the other of them, a process that is repeated until we are left with a default theory, which awaits disproof and which, in the mean time, we call the "truth" or understanding. One can always find when considering a default theory elements of its origins in the past. Knowledge, in particular scientific knowledge, is a historical or learning process.

The immune system is the product of an evolutionary process that introduces variations in the genome that result in a distribution of abilities to adapt. Most are deleterious or neutral. What we see at the functional level are a subset of the total variants, namely, those that improved adaptability, assayed by a process of selection. This selection process can only operate to adequacy, not to perfection. We are easily able to detect and count most of the heritable variation as nucleotide replacements in the DNA sequence, but only in a limited number of situations can we count and detect the number of replacements that result in changes in selectable function (ability to adapt). The result is a misinterpretation of the observation because it is not the total variation that is essential to deciphering the biology; it is the functional variation.

The "instructionist" era

In 1940, a California Institute of Technology physical chemist, Nobelist Linus Pauling, interested in antibodies, can be imagined to have said to himself, Immunologists, for example Karl Landsteiner, tell me that the immune system can be coaxed to respond to any stable structure that exists and what is more surprising to one that it has never seen. My next-door neighbor, the geneticist George Beadle, tells me that one gene encodes for one polypeptide chain. As there must be more stable structures in the universe than there are genes in an individual, antibodies must originate by a mechanism that is not encoded in the germ line. Being resourceful and imaginative, Pauling proposed that a single gene produced a single polypeptide chain that underwent via a somatic antigen-driven template mechanism, a reconfiguration that produced the specific combining site [1]. He knew that each antibody had two combining sites that he referred to as "bivalence," a rather tongue-in-cheek nomenclature given that he was aware that the antigen-antibody interaction did not involve a chemical reaction. What he did not know was that both sites had to be identical, making his proposed mechanism implausible. As immunology at the time was a captive of clinical medicine, such considerations evoked little interest and Pauling's conceptualization was passively accepted or ignored for a period of close to 10 years. When it began to be challenged, it was the template mechanism that came under attack not what should have been the central theme, namely, that Pauling's argument was based on the overlooked assumption that the generation of the antibody repertoire, as a minimum, had a somatic component. The mechanism should have been the next question once it was agreed whether or not a somatic process was involved. The importance of the assumption that a somatic process was involved was missed by all, not only by Pauling. His suggestion of a mechanism was premature and a blind alley because, whether right or wrong, it acted as a dark hole into which all attention was drained. We will return to the question of somatic evolution as it became the central theme for debate years later. What is surprising is that the somatic aspect of Pauling's theory was not challenged even in his home base, the California Institute of Technology, the genetic capital of the world at the time.

However, before we discuss the "germ line-versus-somatic" preoccupation of the late 1960s, a discussion of the interval is quite illuminating. As I have developed my view of the history of this period previously [2, 3], here I only wish to illustrate why how-to-think is of interest.

In 1948, a major contribution was made by Burnet and Fenner [4], who introduced biology into the discussions of the template mechanism. It is amazing to me that at that time, they talked about "genes" and revived the concept, *horror autotoxicus*, under the camouflage, "self-nonself discrimination." Catchy as their new nomenclature was, it invited the most unfortunate muddling. Everyone, from philosopher to journalist, seized on the word "self," making its use in describing immune behavior a nagging unproductive battleground. Today, self is so engrained in the literature that one cannot communicate without using it. As my view of the

subsequent history depends on understanding the meaning of self and "nonself" as elements of immune responsiveness, let me digress to define them. Burnet and Fenner never defined them, although I am quite certain from their discussion that they, as do most immunologists today, viewed self as any component autogenously generated. This being untenable, I am obliged to define it here.

Self is an endogenous component (not just any), a specific immune response to which is sufficiently debilitating to be an object of evolutionary selection (i.e., *horror autotoxicus*). This selection pressure operates on the immune system, not the self-component, which is selected upon to function in the physiology of the organism. All the rest in the universe is nonself, whether or not it is endogenous. For example, autogenously generated waste due to cell necrosis is ridded by a salutary immune response; this autogenous waste is nonself to the immune system no matter how the immunologist may view it. It is the immune system not the immunologist that defines self, consequent to a somatic historical or learning process. The self-nonself discrimination is nothing more than the mechanism used by the adaptive immune system to sort a somatically generated, random combining site (paratopic) repertoire into anti-self and anti-nonself, the goal being to purge the anti-self. In sum, "tolerance" is not broken to autogenous waste (autoreactivity); it is broken to self when a specific attack on this constituent is directly debilitating (autoimmunity). In sum, any antigen not encountered, while the developmental time window is open, is nonself to the immune system. If the antigens in privileged or sequestered sites (eye, brain, gut, skin) are not encountered by the immune system during its ontogeny, they would be treated as nonself, where they are to be encountered when the system is mature. Lastly, the interactions involved in defining self must engage specific recognition by the antigen-specific receptors, T cell receptor (TCR)/B cell receptor (BCR). The selection pressure on the immune system to reduce autoimmunity to an acceptable level is distinct from that which operates to reduce the frequency of innocent bystander pathology. This latter is not an element in the self-nonself discrimination.

Burnet and Fenner's subsequent detailed discussion [5] illustrates a remarkable example of intuition. They latched onto the finding of Owen [6] that fraternal bovine twins are chimeric for each others hematopoetic system. In 1949, this observation was hardly noticed by the immunological community. All the more impressive is that they correctly emphasized that this finding implied a developmental time period during which tolerance was established. Further, they insisted that the problem of the self-nonself discrimination had to have a solution, if ever one was to understand the immune response. Given these startling insights, one might imagine that they were on the road to a reasonable solution to the problem of the self-nonself discrimination. This is clearly not the case as their subsequent discussion exemplifies.

A role for developmental time implies a somatic selection process, which was entirely missed by Burnet and Fenner, as it was by immunologists. Further, although the self-nonself discrimination occupied most of their discussion, they never used it as an argument to challenge instructionist theory. Simply put, there is no way to make a self-nonself discrimination if both self and nonself are templates. Something had to be added to sort that repertoire. Lastly, and most revealing, they produced a model for the self-nonself discrimination that was in flat contradiction with their own conclusion that the Owen finding implies a developmental time window where tolerance is established, as well as with the conclusion that they missed, namely, that a somatic process was required. Their model was that (1) a random repertoire is somatically generated (Pauling's unwitting assumption) and (2) self is tagged with a self-marker that is read by the immune system as not to be attacked. It is of no interest to detail their self-marker model because whatever credibility factor one places on this model, it illustrates that they did not understand their own conclusion about the self-nonself discrimination. A self-marker had to be germ line encoded, making a role for developmental time irrelevant. No matter when a self-marked component appeared in the individual, it would be treated as self. Further, transplantation studies had made it clear at the time that what is self for one individual of a species is nonself for another, raising questions about the possible distribution of these markers in the population. They tried to use the self-marker concept to explain the ABO blood group system, a perfect choice to test any theory of the self-nonself discrimination. They concluded, after rejecting other assumptions, that two germ line-encoded systems are involved, one is ABO and the other anti-ABO. They add however that it is an unsatisfactory solution without telling us why? So let me do that, as it is illustrative. Given that the two germ line-encoded systems must signal each other, the simplest assumption would be that if A is inherited, it tells the other system to turn on anti-B; if B is inherited, it signals turn on anti-A. If O is inherited, it must tell the system to turn on both anti-A and anti-B. If AB is inherited, we face a contradiction. How is the observed failure to produce either anti-A or anti-B explained? A should turn on anti-B; B should turn on anti-A, and the animal would suffer autoimmunity. This of course would be an absurdity or in more polite language, an unsatisfactory conclusion.

However, in the framework of a self-nonself discrimination, a reasonable solution emerges. If the production of anti-A and of anti-B were not germ line encoded but somatically selected, for example, induced by the antigenic load, then A would have to turn off anti-A, B would turn off anti-B, O would turn off neither, and AB would turn off both. Given only a modicum of probing, Burnet and Fenner would have discovered the starting point for the analysis of their favorite subject, the self-nonself discrimination, namely, that the self-component inactivates any response to itself. Burnet

simply did not understand his own interpretation of the Owen experiment, *to wit*, a developmental time window plays a role in the sorting of the random repertoire and that this had to be a somatic process.

What is striking is that even today, the lessons from self-marker theories have not been learned. Today, a major proportion of the immunological community, presumably more sophisticated, denies a role for developmental time in establishing the self-nonself discrimination by substituting germ line-encoded nonself-markers (danger, pathogenicity, discontinuity, context, tuning, etc.) for Burnet's self-marker. The arguments ruling out self-marker theories also rule out nonself-marker theories. This point will be detailed later as, unlike self-marker theory, in the proper context, one can extract a positive value from the nonself-marker theories.

In 1955, a paper by Jerne [7] provided a proposal as to how the combining site (paratopic) repertoire arises and is sorted into anti-self and anti-nonself (i.e., the self-nonself discrimination). He invoked Burnet's assumption of a developmental time window adding that, when open, the total repertoire is expressed as secreted antibody and the anti-self is subtracted from it. When the developmental time window closes the interaction with antigen, instead of being deletional, results in replication of the antibody-antigen complex producing a protective response. Everybody recognized his overlooking of what was on every molecular biologist's blackboard, namely, DNA → RNA → protein. What is more subtle and more interesting were the unrecognized assumptions. The corollaries of Jerne's proposal were first, that the immune system must be born, inactivatable only. Whether this state is a property of the cell or of the organism became a subject of debate later on when Burnet put Jerne's immunoglobulin as a receptor on cells. Second, during the period when the developmental time window was open, all self had to be present and no nonself; when the window closed, no new self is allowed to appear because it would be treated as nonself. Lastly and most important, it should have been clear that the self-nonself discrimination cannot be regulated at the level of the effector output. Any interaction avid enough to initiate a purge signal would be avid enough to activate the biodestructive and ridding effector functions. Yet, throughout his career, Jerne maintained the untenable position that the self-nonself discrimination was regulated at the level of effector output (e.g., idiotype network theory). I have always viewed this as irrational.

In 1957, Burnet published his "clonal" selection theory [8]. What Burnet meant by clonal selection is obscure. In 1961, he [9] wrote, "*the theory is called clonal selection theory because the action of antigen is simply to select for proliferation that particular clone of cells that can react with it.*" Antibodies see antigenic determinants (epitopes) not antigens, which are collections of linked epitopes. Antigens do not select clones but families of cells. Talmage and Lederberg correctly referred to this as "cellular selection." I, not knowing how to discuss the relationship of clonal to one cell–one antigen receptor introduced the distinction "unispecific" clonality and "multispecific" clonality, an all too-poor nomenclature. In any case, both Burnet and Talmage proposed that one cell expressed only one antigen receptor but neither added a reason why. Burnet [8] did however make a guess as to the mechanism establishing "unispecific clonality." He suggested that "*[t]he clonal selection theory requires at some stage of early embryonic development a randomization of the coding responsible for part of the specification of gamma globulin molecules so that…there are specifications in the genomes for virtually every variant that can exist….*" Thus, Burnet tried to answer the right question but inverted the logic in his answer. If all men are mortal, it does not mean that all mortals are men.

If the generation of diversity were due to "randomization" of diploid genes, then unispecific clonality (haplotype exclusion) might be predicted dependent on a great many additional (and unlikely) assumptions that would have had to be added. However, if unispecific clonality is the case, it does not mean that the generation of diversity is due to the somatic randomization of genes, as the heated discussions in the late 1960s to mid-1970s on the germ line versus somatic origins of the repertoire illustrate. In the end, Burnet's guess was wrong; one cell–one antigen receptor (haplotype exclusion) is driven during evolutionary selection by two factors, the necessity to make a self-nonself discrimination and to minimize the flooding of the antibody response with nonfunctional molecules. All this having been said, the hypothesis, one cell–one antigen receptor, was a major advance in 1957, as well I know as I was in the middle of an experiment to test it [2, 10].

In 1959, Burnet [11] corrected the Jerne assumption of a self-replicating protein by putting it as a signaling receptor on cells but kept all of the other components of the Jerne formulation. Neither Jerne nor Burnet attempted to look more deeply at their proposals. Lederberg [12] in his 1959 paper rationalized their model by abandoning the window during which the repertoire was generated and by substituting that antigen-responsive cells are born in a steady state throughout life as inherently inactivatable only and then differentiate antigen independently to activatable only. I will refer to this class of models as "one-signal models" because the same signal is read as inactivation at the initial stage of differentiation and as activation at a later stage, but more of this later.

In 1953, Medawar and associates [13] entered the arena. He was a great admirer of Burnet and was justifiably taken by the assumption of a developmental time component in the establishing of a self-nonself discrimination. So they decided to test it. They injected into newborn mice spleen cells of the same foreign genotype as the skin they would later transplant. As skin transplants between individuals differing at the major histocompatibility complex (MHC) were known to be rejected, it was a dramatic finding that in several cases, the transplant was accepted. This was so startling that it catapulted

Medawar and Burnet into the arms of a Nobel Prize, Burnet for his theoretical insights as described above and Medawar for his experimental confirmation of the theory. This prize served well the immunological community because it aroused interest in a fundamental requirement of the self-nonself discrimination, namely, a developmental time window. However, the speculations of Burnet show that he did not understand what the developmental time window entailed and Medawar did not appreciate that his interpretation of that experiment should have raised eyebrows; in fact, oddly enough, no one has ever questioned the interpretation of this experiment.

One property of the self-nonself discrimination was obvious, namely, that it had to be antigen specific. The ability to distinguish self from nonself has specificity as a minimum requirement, and of course, this is equivalent to distinguishing one nonself determinant from another. The necessity to distinguish self from nonself is the evolutionary selection pressure determining the degree of specificity of its antigen receptors. Consequently, there was no reason to expect that lymphocytes injected into newborns would "tolerize" against the antigens specific to skin. Yet, the experiment worked in a few cases. The lymphocytes of the donor are obviously blind to the recognition of the donor skin but perfectly reactive to the antigens of the host and could easily have set up a graft-versus-host reaction. The lymphocytes of the host could recognize not only the injected foreign lymphocytes but also the skin antigens of the donor and set up a rejection of skin, yet neither occurred. Medawar might have asked himself, why? Simply posing this question, in and of itself, would have cast doubt on the interpretation that this experiment dealt with developmental time as a factor in the self-nonself discrimination.

Today, we might reinterpret this experiment as follows. Within a certain ratio of donor to host lymphocytes, an *entente cordiale* might well have been established based on the mutual nonspecific suppression of each population. This would permit both the donor skin to be accepted and the host to be protected from graft lymphocyte attack. Such a scenario would have predicted that the animal accepting the graft would also be unresponsive to third-party grafts. Unfortunately, such specificity studies were not carried out. Lastly, had it been technically possible, a direct graft of skin on the neonate would have been predicted from Medawar's interpretation of the experiment to be accepted, whereas the above suppressor model would have predicted that it would be rejected. This experiment did not demonstrate a role for developmental time in establishing the self-nonself discrimination but rather suggested a role for "suppression" in regulating the magnitude of the effector response. Although the immunologist could manipulate an animal to become unresponsive by a given assay and this could even be of enormous practical consequence, there is no way that the Medawar et al. observation could be extrapolated to a mechanism that establishes the self-nonself discrimination.

Even today, the Owen observation that bovine fraternal twins can share their hematopoietic systems has loose ends. Medawar and colleagues [14, 15] extended the observation by showing that fraternal twins accept reciprocal skin grafts. Bovine fraternal twins anastomose their placentas allowing exchange of blood-related elements and possibly skin antigens. However, one would not expect sharing of all tissues, predicting that grafts from other organs would be rejected. If this was not the case, and grafts from all other organs are accepted, then either all self-antigens are exchanged as tolerigens via the placenta or the animal is totally suppressed (i.e., unresponsive to all antigens). Both appear as very unlikely, predicting that some tissues will be rejected and some not.

It is hard to say what we learned from this experiment about how to think. All of the information was available at the time to question its interpretation, yet no doubt was voiced, illustrating how important a pinch of skepticism is. Asking what would an alternative interpretation entails is always a valuable question.

This failure to distinguish tolerance from "unresponsiveness" plagues immunology even today. Unresponsiveness describes an observation in which a manipulated animal specifically does not respond to an antigenic challenge to which an unmanipulated animal would respond. Tolerance is the interpretation of this observation to explain how the repertoire is sorted into anti-self and anti-nonself. Unresponsiveness is observation; tolerance is theory. For example, a graft can be accepted by switching the effector response from effective to ineffective or by suppressing the magnitude of the response to below effective, but neither of these cases of unresponsiveness can be extrapolated to how the repertoire is sorted (tolerance).

In 1959, when the Lederberg paper [12] appeared, the emphasis was still on Pauling's template mechanism. Given that framework, Lederberg categorized the two opposing theories as "instructionism" and "selectionism." However, his emphasis was misplaced. That the variation precedes the selection had been firmly established experimentally, as well he knew since one of the most elegant demonstrations of it was Lederberg's replica plating experiment [16]. The variant is selected by the stimulus, not induced by it. Pauling's hidden assumption of a somatic process should have been the emphasis as illustrated by the next era, the late 1960s. The two opposing theories that had to be faced were "somatic" versus "germ line" generation and selection of the combining site repertoire. In 1959, an attack on instructionism was only the breaking down of an open door. In any case, it should be pointed out that throughout the 1940s up to the late 1950s, the vast majority of immunologists were steadfastly, whether *laissez faire* or not; "instructionists" and the inability to explain the self-nonself discrimination in that framework evoked little interest.

In the 1960s a new generation of molecular immunologists entered the arena and their focus was on whether the repertoire

is entirely germ line selected or in large measure somatically selected. It should be clear that there is an asymmetry between the two views. Germ line selection does not require a somatically selected component, whereas somatic selection is built on a germ line-selected base. The problem of the self-nonself discrimination during this period was viewed as a parallel phenomenon and, incorrectly, as a somewhat disconnected subject.

What we should learn from this saga is that one must always ask what would the next step look like if a given interpretation was correct? Abandoning thinking with, I believe what I see is all too frequently misleading.

The next era (1960s) was dominated by attempts to characterize the combining site (paratopic) repertoire, the substrate of the self-nonself discrimination. This emphasis automatically divided the immune system into three modules:

1- The generation and characterization of the combining site (paratopic) repertoire;
2- The sorting of the repertoire (the self-nonself discrimination);
3- The choice and regulation of the coupling of the paratope to a biodestructive and ridding effector class.

Each module has its own database and logic. The strength of this simplification is that the three modules could, in principle, be linked to understand the behavior and selection pressures that characterize the integrated immune system that we see today. One of the best ways to deal with complexity is to modularize.

Introducing the combining site (paratopic) repertoire of the immune system

In the 1960s through the 1970s, almost all immunologists were supporters of a repertoire that was derived uniquely by germ line selection. Jerne in 1971 [17] joined the minority of defenders of somatic models [18] by proposing a mechanism for the generation of the repertoire. Once again, he relied solely on his intuition. He proposed that cells expressing a germ line-encoded anti-self receptor interacted with a self-component and were deleted. The mutants of this receptor that failed to recognize the self-component formed the anti-nonself repertoire. He ignored that most mutations would be inactivating and many would be anti-self raising doubts that mutation by escape from purging of a germ line anti-self specificity could yield a functional anti-nonself repertoire in a short enough time. However, this is a minor consideration compared to the entrained assumption that evolution could select in the germ line for an interaction, which would be debilitating in order to require a somatic process to counteract it. In other words, there is no way to select for recognition of self in the germ line, if a somatic self-nonself discrimination deletes its expression. And then of course, how do you deal with the somatic mutants to anti-self?

Jerne guessed that the self-component in question was encoded in the MHC, a very popular locus at the time, but he gives us no rationale, as any self-component would have satisfied his theory. There is a lesson to be learned in all of this. Today, many authors credit Jerne with having predicted that recognition of the "MHC," as it applies to restrictive recognition by the T cell antigen (TCR), is germ line encoded. This becomes an intellectual roller coaster based on the failure to parse the paper being cited. The MHC when functioning in restrictive recognition of peptide is not functioning as a self-component; it is functioning as an element in the physiology of the animal. The result is the irony that Jerne's correct guess is rejected today by these same authors who now favor the unjustified guess that establishing the allele-specific recognition of the MHC required for restrictive recognition is a somatic process [19]. The take-home lesson is that a consideration of the correctness of the reasoning leading to the conclusion is as important as the correctness of the conclusion itself.

My last point is that during this period, the transplantation biologists [20] had given us a precious datum, which I have referred to as the F1 test [3]. Considering a graft of F1AB onto parental AA or BB, it is rejected. The interpretation is that AA is induced to express anti-B and BB is induced to express anti-A, explaining why the F1AB graft is rejected. It is well to recall that AA rejects BB grafts as well as CC, DD, and EE grafts. The F1AB does not self-destruct; further, it accepts grafts from AA and BB, while still rejecting CC, DD, or EE. This F1 test shows that somatic selection depends on the specificity of recognition of self and by difference, nonself. Self-recognition is deleted, leaving nonself-recognition. The adaptive immune system responds to the difference between the self of the individual and nonself. This is to be contrasted to the innate system, which responds to the difference between the self of the species and nonself.

Returning now to an earlier discussion, if an antigen-related factor such as a self- or nonself-marker (danger, pathogenicity, discontinuity, context, etc.) was the determining element, we would not expect the graft of F1AB onto AA or BB to be rejected when grafts of AA onto AA or BB onto BB are accepted. Clearly, recognition of nonself-B by AA is the determining factor in the response. The ability to distinguish a self-determinant from a nonself-determinant is equivalent to distinguishing two nonself-determinants. This is what gives us a uniquely meaningful definition of specificity [21–23] and explains why transplants of tissues between individuals of a randomly mating population are rejected.

It is important to appreciate that the F1 test defines the adaptive immune system, distinguishing it from the innate system. If the antigen-recognition system was entirely germ

line encoded (i.e., innate), AA and BB would accept an F1AB graft (see below). Individuals with innate systems only (e.g., RAG-minus mutants, not to mention invertebrates and plants) accept grafts from other individuals of the species.

I do not think that either Medawar or Burnet was aware of these studies or even if they were, would have analyzed them from this perspective.

The era of the paratopic repertoire

The germ line versus the somatic origin of the paratopic repertoire of BCRs and antibodies

How does one discuss a short-lived but passionate debate about a phenomenon where everyone had to be "a little bit right"? All conceptualizations of the origin of the combining site (paratopic) repertoire had to begin by telling us what is encoded in the germ line ("the little bit right"). The germ line theorists put the encoding of the entire combining site repertoire in the germ line, whereas the somatic theorists put a handful of specificities encoded in the germ line and varied them somatically. The vast majority of the immunological community supported the germ line theory as it appeared to be the most simple, and the minority who favored a somatic theory never properly rationalized why they felt that a somatic component was required.

A germ line theory was totally deceptive as it required a solution to the pathway of selection of the encoding of a considerable number of germ line genes. So let us look at that more closely.

Immunologists divide the immune system into two classes, innate and adaptive. The innate system is entirely germ line encoded. The selection on the specificity of the innate combining site is twofold: (1) it must distinguish nonself from the self of the mating pool (i.e., the self of the species) and (2) it must be optimized for recognition of determinants common to as many pathogens as possible. The adaptive system also has a germ line-encoded repertoire that in principle is selected as is the innate repertoire and where it is not for the fact that it is the substrate for a somatic diversification mechanism, could well be classified as part of the innate system. The debate then in the early 1960s before somatic diversification was demonstrated might be restated. Is the entire repertoire part of the innate system or does a superimposed somatic diversification mechanism exist (i.e., the adaptive system)?

Although the information to answer this question was there, no one had the acumen to put it together. Selection in the germ line would result in a repertoire that was blind to the self of the species, whereas selection in the soma would be blind to the self of the individual. This meant that under the germ line hypothesis, transplants between individuals would be accepted, when in fact they are rejected. Consider the following scenario: a duplicated gene mutates to encode recognition of a determinant on a self-component in the species that is absent in the individual. A mating of that mutant with an individual expressing that self-component would result in the debilitation of the offspring, with the result that the mutant gene would be eliminated from the gene pool. A species with no adaptive system, only with an innate system, accepts grafts between members of the species. The appearance of an adaptive system results in their rejection.

All this having been said, in the minds of immunologists, the germ line hypothesis was not disproven by reasoning, as it might have been, but by direct experiment [24, 25]. The somatic mutants of a single gene were superimposable on the regions defined as complementarity determining by the Kabat-Wu hypervariability plots [26]. This demonstrated somatic mutation followed by antigenic selection. Unfortunately, by resting on one's laurels, the incompleteness of the somatic hypothesis is ignored or unappreciated even today.

Still missing today is an understanding of how the germ line-encoded portion of the adaptive repertoire is maintained by evolutionary selection. To immunologists, this is viewed as a peripheral question. And well it may be, but there is always someone driven by curiosity that would like to see where an answer to that question leads.

The best estimate of the number of VLVH pairs maintained in the germ line by the specificities they encode is ~40. This is based on direct sequencing of the VH locus in humans [27] and the repeat specificities of immunoglobulins arising in the murine myeloma libraries [28]. These specificities include recognition of a variety of carbohydrates as well as housekeeping products of cell necrosis (e.g., DNA, RNA, nucleoprotein, and myosin). These ligands are those that vary sufficiently slowly to permit selection for their recognition over evolutionary time. The selection pressure would be that these 40 specificities confer a protective advantage to the individual expressing them. This poses an intriguing paradox.

The expressed combining site (paratope) repertoire is made up of complementation between VL and VH with a contribution by the joining region (NDN or CDR3). In order to select germ line for the 40 primordial pairs, they must be displayed in a unique way. Yet, it appears that the VL and VH of these 40 pairs are expressed randomly complemented to yield 1560 (40^2–40) antibodies with new specificities. If, to this, one adds a dead reckoning estimate of the contribution of NDN (CDR3), the 40 pairs presumably under selection would be diluted in a sea of somatically derived specificities that could well be of the order of 10^4. What is the distinguishing feature that tells evolution which 40 of the expressed 10^4 to select? As I pointed out above, this question could have been asked by the supporters of somatic diversification in the early 1970s.

The most reasonable solution would have been that, independently of the scrambling of specificities by

complementation and NDN (CDR3), the 40 are expressed separately, undiluted in a special class of B cell that would make them selectable. As they are selected in the germ line, these antibodies would have to be directed against epitopes that are absent in the self of the species and therefore would not require a somatically derived self-nonself discrimination before expression as effector molecules. Therefore, this antibody could be produced constitutively or secreted simply after interaction of the cell with ligand. This speculation has only minimal support today as no direct effort has been expended to test it. Possibly, the B1a cell subset is this postulated special class, and possibly, "the 40" are expressed as so-called "natural" antibody. Even if we knew the answers at this level, we would not be out of the woods as this solution poses the problem of the mechanism of antibody expression by the postulated B1a cell.

When a given VH is rearranged into the H-chain subunit, it would have to select the appropriate VL that yields the corresponding specificity. As the given VH has at least 40 VL to select from, how is the correct choice orchestrated? And lastly, the NDN (CDR3) region must not be allowed to affect the specificities of the 40 pairs. The only way to accomplish that is to not or ineffectually express this region in these antibodies.

I have indulged in this digression to illustrate why earlier I said that one should always ask, what would the next step look like? The failure to answer these questions could 1 day topple the whole framework for visualizing what is carried in the germ line under a somatic model. Lastly, this problem will arise again, if one considers the TCR to be a BCR analog, as do most, if not all immunologists today.

Is the paratopic repertoire of the TCR or germ line somatically selected?

The division of immune effector functions into cell mediated and humoral was established over many years and was certainly a comfortable concept in the 1960s. The general picture was that immunoglobulin functioned as a receptor on the cell-mediated class and as secreted antibody by the humoral class. In the 1970s, the roles of T cells and B cells were defined. The T cells were pictured to recognize a protein encoded in the MHC (symbolized R) and cell-bound antigen. The recognition of R told the T cell that it is dealing with a cell-related antigen. The TCR was viewed as recognizing cell-bound antigen in the context of the recognition of R. The antigens involved in these studies were minor histocompatibility antigens or viruses that expressed intact molecules on the cell surface. This spawned two classes of model. Matzinger and Bevan [29] proposed that the R element formed a meld epitope with the cell-bound antigen that is recognized by a BCR-like or single combining site. This meld epitope was referred to as "altered self," "interaction antigen," or "new antigenic determinant" (NAD). Langman [30] and I [31] proposed a competing two receptor model, one receptor recognizing R and the other recognizing the cell-bound antigen. The two linked receptors, of course, had to signal the T cell by acting in concert. When Zinkernagel and Doherty [32] demonstrated that the recognition of R was allele specific, the two models were easily adjusted to accommodate the finding. Under the NAD model, the meld ligand would be formed by an interaction between the allele-specific determinant on R and the cell-bound antigen. Under the two receptor model, the anti-R receptor would be germ line selected to be specific for an allelic determinant on R and the receptor for cell-bound antigen would be somatically generated like the BCR. However, when it was demonstrated that R is a presenter of peptide [33, 34], the two receptor model was ruled out but the NAD model remained unaffected. The demonstration that the ligand for the TCR is peptide (P) presented by R (PR) told us that the TCR is not interested in *cell-bound* antigen. It was sculpted to deal with *intracellular* antigen. I am certain that if the difference between cell-bound and intracellular antigens had been appreciated, the role of peptide would have been predicted. In any case, at the time, all two receptor models of the TCR function-structure relationship were ruled out.

Today, the BCR-like TCR as envisioned by the NAD model reigns supreme. I know of no immunologist who views as tenable the competing model, which I will rationalize here. It is a reformulation of the two-receptor model. The reason that the competing model (Tritope) is germane is that it is a chance to balance the discussion and see which model the future will favor. Is the logic of evolution or description by analogy, the better way to think about the immune system. Competing models should be appreciated and respected as they sharpen the question.

We will be analyzing two models of the TCR function-structure relationship. Model I is based on an analogy between the BCR and the TCR. The TCR has a single combining site that is composed by complementation between Vα and Vβ and which recognizes a composite epitope formed between the peptide (P) and some determinant on R (NAD). Model II envisions a TCR with three sites, anti-P, anti-R_{host} allele, and anti-R_{allo} allele. Under model II, it is a given that the TCR recognizes an allele-specific determinant on R (i.e., restrictive recognition of peptide and alloreactivity). Of course, the TCR must also be able to recognize the peptide (anti-P). Further, extrapolating from limited data, every restricted TCR is also alloreactive.

It follows then that

1- The recognition of the host R allele must be germ line selected and encoded;
2- There must be a combining site on the TCR (anti-R_{host}) that is the target of this germ line selection;

3- There must be a combining site (anti-P) on the TCR that recognizes peptide;
4- Restrictive signaling via the TCR must engage a coordinated interaction between anti-R_{host} and anti-P; and
5- As allorecognition must also be germ line selected and encoded, there must be a combining site on the TCR (anti-R_{allo}) for it.

The fact that the complementation of subunits of class II R does not produce hybrid alleles tells us that single V-gene segments specify recognition of each allele. The TCR, which is composed of Vα and Vβ, must have two sites, one used for restrictive and the other for allorecognition.

What is allorecognition for one individual of a species is restrictive recognition for another. Therefore, a given anti-R_{host} and anti-R_{allo} in one individual can reverse roles and be anti-R_{allo} and anti-R_{host} in another individual. As the only germ line encoded part of the TCR that can be engaged in this recognition is the V region, one can conclude that single V domains (Vα or Vβ) specify the recognition of R alleles. The Vα- and Vβ-gene segments are each germ line selected to recognize an allele of R. The TCR must restrictively signal the cell only when the two sites, anti-R_{host} and anti-P, are engaged. In contrast, the TCR must be able to alloreactively signal the cell upon engagement of anti-R_{allo} alone. The signal from either source is the same (signal 1). The mechanism or structural basis for this is open and merits some attention.

It might be noted in passing that the anti-P repertoire of the TCR is entirely somatically derived in contrast to the BCR, which depends on a germ line-selected set of VLVH pairs that is the substrate for somatic variation by random complementation, gene conversion, and mutation. This difference between the TCR and the BCR might reflect the inability to select in the germ line for recognition of peptides from pathogens because they are not stable enough for germ line selection (i.e., too easily escape recognition by mutation).

Model I has a competing logic. As a BCR-like combining site made up by complementation between Vα and Vβ is envisaged, a meld between the peptide and a determinant on R must be the ligand for the TCR. As a model I TCR has no way to know which part of the meld ligand is P and which is R, it is postulated that TCRs comprise a family ranging from R centric to P centric. TCRs that at the extreme see P centric without engaging R (unrestricted signaling) do not exist. TCRs that at the other extreme see R centric without engaging P would be lethal. Therefore, the TCR must be somehow limited to signaling only when elements of both P and R (the meld or NAD) are recognized. This makes it difficult to believe that an allele-specific determinant on R is uniquely involved in the meld. If an allele-specific determinant was an element in the PR-composite epitope and there are thousands of peptides with which it melds to create a unique NAD, allele-specific recognition would be unselectable in the germ line. Presumably, it is for this reason that the supporters of model I either challenge a role for allele-specific recognition in TCR function or bury the problem under such rhetoric, as the TCR has a "bias," "obsession," "predilection," or "preference" for R. As R itself must be recognized during thymic positive selection in order to establish the class of R (RI or RII)–effector function relationship, the only alternative candidate determinant would be an invariant site on R that is common to all RIs or to all RIIs of the species. Consequently, it would be predicted that no given single TCR can be signaled by both RI or RII. This is contrary to fact, as many examples exist of TCRs restricted to RI and alloreactive to RII or vice versa.

However, more interesting for me is how model I would deal with the germ line-encoded recognition of R, whether it be allele specific or class of R specific. The combining site is envisaged to be formed by complementation of the α and β subunits of the TCR. The recognition of R itself must be germ line encoded by the complementation of given Vα-Vβ pairs. As there are in mouse ~20 Vβ and ~80 Vα, random complementation would yield ~1600 Vα-Vβ pairs, of which only 20 could be functional in the recognition of R, whether it be allele or class specific. If class of R is recognized, then only 2 Vα-Vβ pairs would be needed and the 20 might be viewed as a physiologically valuable redundancy, raising the question as to what property of this large redundancy makes it evolutionarily selectable. If allele-specific determinants on R are recognized, then there would be only 20 in the species, clearly too few. This is why the assumption of model II that single V domains recognize is by far more likely.

To appreciate these above arguments, it is necessary to distinguish peptide functioning as a specificity ligand and peptide playing a role as a structural element in the PR-complex essential for the stability and conformation of R (i.e., in particular, for the expression of the allele-specific determinant). Under model I, every role of the TCR, positive selection, alloreactivity, and of course negative selection, depends on peptide functioning as a specificity ligand. Under model II, peptide functions as a specificity element only during negative selection (i.e., restrictive recognition of peptide). For positive selection and alloreactivity, peptide functions as a structural element maintaining the stable expression of the allele-specific determinant on R; it does not function as a specificity ligand [35].

In my view, the need to challenge allele-specific recognition by the TCR in order to be viable is sufficient reason to rule out model I. Of course, it is incumbent on the supporters of model II to provide an evolutionary selection pressure for allele-specific recognition. This would be the next step in characterizing model II.

As a preliminary attempt to approach this next step, consider a primordial TCR that recognizes a determinant on R itself. The species expressing this TCR would see R as a monomorphic

element (R_{mono}). There are species today for which R appears to have remained monomorphic (e.g., [36]). One has to be cautious as the assay for monomorphism must be able to examine each R element, not just the haplotype of the locus. Parenthetically, an analysis of the TCR-gene family encoding monomorphic recognition could topple model II.

The recognition of a monomorphic R under model II is indistinguishable from the recognition of an invariant site on R under model I. This postulated primordial TCR-R_{mono} complex is in a seesawing relationship with the pathogenic family against which it protects. There are two pathways of evolution that this relationship can be envisaged to follow:

1- Any mutation in the pathogen that directly mimicked the determinant on R recognized by the TCR would act as a decoy to inactivate the functioning of the immune system. This would necessitate a countermutation in R that created a new determinant (i.e., an allele). The subsequent selection for this protective mutant of R would make it polymorphic.
2- It would be reasonable to assume that the selection pressure is on the repertoire of peptides anchored by the binding groove on R. In this case, the mutations in R affecting the specificity of anchoring of a P family would be assumed to allosterically alter the determinant on R recognized by the TCR. The pathogen would mutate the sequence of the bound target peptide so that it is no longer anchored. The countermutation in the binding groove on R that permitted the new peptide from the pathogen to be presented would create a new allele-specific determinant. The subsequent selection would make it polymorphic.

An important conceptualization, the protecton; what is the size of the repertoire?

Lederberg in a footnote to his 1959 paper makes a provocative comment. He argues that it would embarrass any theory of the size of the repertoire, if it were larger than the number of antigen-responsive cells in the animal. As it was widely felt that the repertoire is open-ended, transcendental, or "complete," Lederberg's putting this cap on the size of the repertoire was a bold advance, but on reflection, we [21, 23] found it quite paradoxical. Let us use order of magnitude numbers to illustrate the paradox. If we assume that a pygmy shrew (10^7 cells), a mouse (10^8 cells), a human (10^{12} cells), and an elephant (10^{14} cells) are equally protected against their pathogenic universes, then one can envisage that 10^7 elephant cells transplanted into an empty pygmy shrew would protect it. However, it would take the cells from 10^7 pygmy shrews to protect the elephant. Of course if the repertoire were open-ended ($>10^{14}$), nobody would be protected, the pygmy shrew because it would miss over 99.999% of the repertoire and the elephant because no single specificity would be at a sufficient concentration to respond in a short enough time. The size of the repertoire has as its boundary conditions being sufficiently small to be adequately responsive in a short enough time, yet sufficiently large to be adequately anticipatory. This implies that the immune system is made up of iterated units, each unit totally functional. The unit, a protecton, is the smallest segment of an immune system that retains all of the recognitive properties of the whole. If this unit is 10^7 cells at a concentration of 10^7 cells/cm^3, then a pygmy shrew would possess one protecton, a mouse 10 protectons, a human 10^5 protectons, and an elephant 10^7 protectons. They would all be equally well protected against their pathogenic universes. The protection is per cubic centimeter, not per animal. Most importantly, the paratopic repertoire cannot be larger than the number of cells in a protecton, not in an animal. In this illustration, the upper limit would be 10^7. As we will see, this cap on the repertoire is probably closer to 10^6 [22].

A model of paratopic recognition; introducing the specificity index

There is compelling evidence supporting the view that the TCR/BCR is polyreactive or polyspecific [37]. This means that a given TCR/BCR appears to recognize as ligand a variety of epitopes that are chemically distinguishable. For over 50 years, the interaction between a combining site (paratope) and its ligand (epitope) has been viewed with "lock and key" imagery. Polyreactivity in that framework is referred to as "mimicry" or "cross-reactivity," but that description does not give us any insight into the basis of the signaling interaction and in a sense is misleading. For this reason, I developed a model [22] that, whether valid or not, highlights the problems to be faced. This model is applicable to both the TCR and BCR, but because of ease of presentation will be summarized here using as example the TCR.

The ligand for the TCR is a peptide bound in a groove on R. The average peptide presented is ~10 amino acids in length, 5 of which are anchored in the binding groove and unavailable for recognition by the anti-P site of the TCR. Five are exposed and potentially recognizable by the TCR. This means that the maximum epitopic repertoire available to the TCR is 20^5 or 3.2×10^6. If only 17 amino acids contribute to this epitope in a distinguishable way, then the maximum epitopic repertoire would be 17^5 or 1.4×10^6.

Now, let us consider five classes of TCRs that can be signaled (signal 1) by recognition of either one or two or three or four or five amino acids. This defines five categories of TCR looking at various combinations of the target amino acids, which will be identified as being in positions 1 to 5. As an example, let us look at a TCR that can be signaled by interaction with a single amino acid (e.g., a TCR that is triggered by tyrosine at position 2 in P). From the point of view of that

TCR, it is absolutely specific (i.e., unireactive). However, from the point of view of the immune system, that TCR is polyreactive. It can recognize 20^4 peptides presented by R. If there is 1 peptide out of the 20^4 peptides with tyrosine at position 2 that is a self-peptide, then that TCR will be negatively selected. It comes as no surprise that a TCR recognizing one amino acid could never survive a self-nonself discrimination. As an aside, from the viewpoint of the experimental immunologist, this TCR can recognize 20^9 peptides. Fortunately, it is only the opinion of the immune system that is germane.

The extrapolation of this example to all possible combinations of TCRs and their ligands is detailed in ref. [22]. Here, I will only discuss the general picture. There are five categories of TCR; those signaled by one amino acid recognize 20^4 peptides, by 2 amino acids recognize 20^3 peptides, 3 amino acids recognize 20^2 peptides, 4 amino acids recognize 20^1 peptides, and 5 amino acids recognize 20^0 peptides. The number of peptides recognized by a polyreactive TCR will be symbolized, *n*. Given this picture of recognition of the epitopic repertoire, what are the consequences for the paratopic repertoire?

In order to analyze this, we must introduce a parameter, the epitopic specificity index, *si*, which is the probability that a given epitope is a self-epitope. This permits us to calculate the probability that a given paratope is anti-self, *SI*, the paratopic specificity index, $SI = 1 - (1 - si)^n$. For illustration, after examining a wide range of values for *si*, 0.01 appeared to be the most likely value. Given this, *SI* for each category of TCR can be calculated and the total summed. This sum, the anti-self repertoire, is subtracted and the residue is the anti-nonself repertoire. As shown in ref. [22], 63% will be purged as anti-self, leaving 37% anti-nonself as the functioning repertoire. A value si = 0.01 yields a TCR repertoire that on average binds an optimal number of peptides as calculated independently [38–41]. In a protecton of 10^6 cells, each nonself epitope would be recognized with a multiplicity of 3.8, missing therefore 0.02 ($e^{-3.8}$) epitopes at any moment in time. This would be the minimum size protecton that possesses a functional anti-nonself repertoire. The size of the protecton is estimated, therefore, to be between 10^6 and 10^7 cells at a density of roughly10^7 cells/cm^3.

Lastly, for the T cell, which needs to recognize only one epitope to mediate effector function, in order to miss less than 1 per 1000 antigens, each antigen must on average be presented as 7 epitopes (e^{-7}). For the B cell, which must see each antigen in 3 or more ways to carry out effector function, in order to miss less than 1 in 1000 antigens, it must recognize on average 10 epitopes per antigen [42]. However, in order to do this, the minimum size of the B cell repertoire is expected to be slightly larger than that of the T cell repertoire.

Evolution had no choice. Polyreactivity is expressed at the level of the immune system, and unireactivity is expressed at the level of the TCR. The paratope (anti-P) operates at the molecular level (recognition of combinatorial amino acid side chains), whereas the S-NS discrimination depends on the presentation of self-epitopes. These include amino acid residues not seen by the anti-P of the TCR.

One's conceptualization of polyreactivity is important because it bears on every aspect of immune behavior. Any solution to the self-nonself discrimination must begin by characterizing the paratopic repertoire (module I) as that is the substrate to be sorted.

A change of approach

At this point, I would like to introduce a change of perspective. Thus far, unavoidably, I have viewed the immune system with a knowledge of today's default concepts, although I have been as careful in analyzing early data to consider only the knowledge of the time. Historians dispute whether objectivity can be achieved by such an observer, although there is no other choice. Respecting their arguments, I decided to look at the immune system with no such knowledge by analyzing contemporary competing models between which only future experiments will decide. In this way, we will eventually be able to see whether concept or tinkering will prove more fruitful?

The era of the self-nonself discrimination

The two-signal model

I would like to begin this discussion with our 1970 proposal for the mechanism of the self-nonself discrimination [43]. The reason is that it tried for the first time to deal rationally with the real-world immune system. The model included (1) a developmental time component; (2) a repertoire that was generated continuously as a steady state throughout life; (3) negative selection to inactivate anti-self (signal 1) and an activating anti-nonself signal (signals 1 + 2) putting the activated cell on first step of the pathway to effectors; (4) a signal 2 that was delivered by a thymus derived cell (today referred to as a T helper, Th); and (5) most importantly, it highlighted the role of associative (linked) recognition in the mechanism. I am amazed today at how prescient that 1970 proposal was. Yet, it was received hostilely by the immunological community [44], forcing us over the years to engage in lively debates that permitted important extensions and clarifications of what had become known as the two-signal theory. In order to emphasize the clarifications and extensions of the original theory resulting from these debates, I renamed it, the Associative Recognition of Antigen (ARA) model [45–48]. I view this ARA model today as a default model [49–52]. Before facing

the debates and the present-day unresolved aspects of the model, let us look at the pathway under discussion (Fig. 1). We need a nomenclature to discuss its ramifications.

Initial state cells (i-cells) that are activatible or inactivatible are generated in a steady state throughout life. These cells express antigen receptors that are either anti-S or anti-NS. Upon interaction with an antigenic determinant (epitope), signal 1 is delivered to the i-cell, which drives its differentiation to an anticipatory cell (a-cell). It is so named because it is at this stage that the decision between inactivation and activation is made. If no additional information (e.g., signal 2) is given to this cell, then at a defined pace with increasing refractoriness or anergy to reversibility, the a-cell proceeds on to irreversibility (inactivation). If, on the other hand, signal 2 is delivered early enough to the a-cell receiving signal 1, then it is diverted to the activated or g-state, the first step on the pathway to effectors (e-cell). In order to maintain specificity of activation, signal 2 must be delivered to the a-cell via an epitope from the same antigen delivering signal 1 (ARA). This pathway has been discussed with increasing sophistication over the years, and I will leave this to the references [49, 53–57].

Here, I would like to deal with the challenges to the ARA model and to its unresolved aspects. It will allow us to view with foresight why independent as opposed to consortium thinking still has value.

Challenges to the ARA model

Idiotype network theory

Jerne's "*Eureka*" while in his bathtub [58] that the immune system was a network of immunoglobulin interactions spread through the immunological community like a hurricane spawning in its wake a vast amount of trivial and, frankly, questionable experimental work. As it was a blunt challenge to the ARA model, it became important for us to analyze it, and as far as I can tell, only Langman and I undertook its defense [59–61]. If I may indulge in a personal comment, in 1981, I wrote an invited article for Jerne's 70th birthday *festschrift* volume, in which I analyzed the distinction between and consequences of associative (linked) and nonassociative (unlinked) recognition. This latter is at the heart of idiotype network theory and what makes it untenable because in that framework neither the self-nonself discrimination nor the regulation of effector class is explicable. This article was considered by the editors of the volume too critical for a *festschrift* volume and they rejected it, so I sent it elsewhere [46]. This is an example where scientific discourse looks more like political discourse.

It is no wonder then that in 1981 at a workshop organized at the Basel Institute and entitled "Idiotypes: antigens on the inside," Jerne declared in response to my criticism of such networks based on their inability to deal with the self-nonself discrimination that "*There is no theory about self/nonself discrimination, as far as I know. Nobody understands it* ([62], p. 140)." The question was not do we understand the self-nonself discrimination but rather is a theory that is incompatible with the existence of such a discrimination a priori viable? Clearly, nothing had been learned from the instructionist era where its demise, in part, resulted from this incompatibility being shrugged off. However, more informative for me is that Jerne and his followers refused to recognize the ARA model as a valid scientific theory.

This led to a series of debates on the validity of network theory [59–61], in which our criticism of untenability was left unanswered (i.e., ignored). This provoked our final paper entitled, "*The 'complete' idiotype network is an absurd immune system*," in which we spelled out the argument [63]. The most interesting claim of the network theorists was that the recognitive repertoire of the immune system is open-ended, transcendental, or as they put it, complete. Therefore, they argued that such a repertoire could not help but recognize

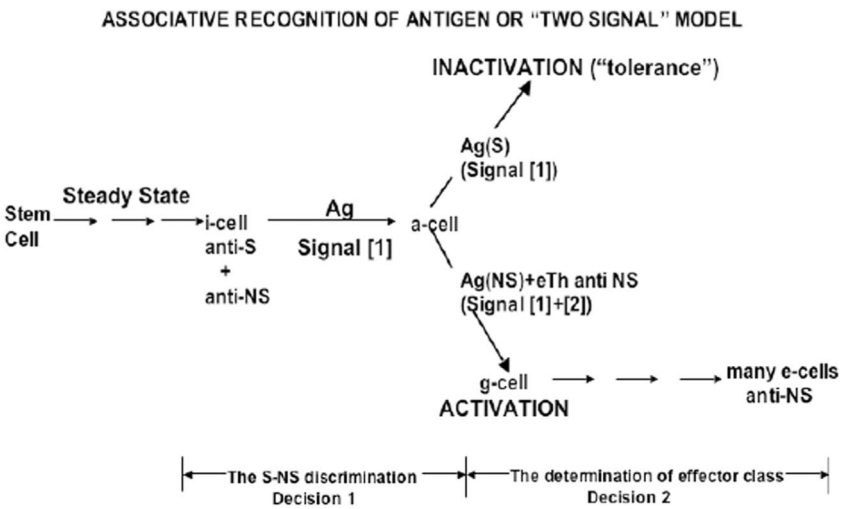

Fig. 1 The pathway of the self (S)-nonself (NS) discrimination (taken from ref. [50])

itself, making an idiotype network inevitable. There were even those who argued, using the premise of "completeness," that the immune system recognizes every bodily constituent, making it the major regulator of the animal's physiology. This is why earlier I presented the more rational protecton theory and suggested a model of the paratope that put a cap on the size of its repertoire, making the concept of completeness irrational.

Today, idiotype network theory, except for a few "defenders of the Alamo" [64], has disappeared silently from the literature [65] with no one having changed their mind either by accepting or denying the criticisms. This is why nothing has been learned by its defenders despite a vast amount of experimental work and debate. Over the years, I have constantly asked what experiment could be done that would disprove this theory. In this regard, it is not a valid theory. In sum, this era illustrates a situation in which an idea took over the mind, rather than the mind taking over the idea.

Nonself marker theories

In the context of the self-nonself discrimination, self and nonself marker theories are ruled out. There is no physical or chemical property of antigens that can be used by the immune system to separate self from nonself as classes. What is self for one individual of a species is nonself for another. Therefore, germ line-encoded markers (danger, pathogenicity, discontinuity, tuning, context, etc.) cannot sort a somatically generated random (with respect to self and nonself) paratopic repertoire. Further, self or nonself markers do not predict any need to sort the paratopic repertoire.

Before enlarging on this question, I would like to make a comment of semantics. Terms like "danger" and "pathogenicity" are poorly chosen because the immune system has no way to recognize and assay those properties. The immune system only becomes aware of danger and pathogenicity when they "stress" or do "harm." One cannot disprove the danger/pathogenicity theory by experiment because the immune system is blind to these attributes, and therefore, they can be invoked by the protagonist at any time to validate the theory. There is no way to demonstrate whether or not danger/pathogenicity is present, as long as one cannot assay it.

The sorting of the paratopic repertoire requires the prior somatic sorting of the epitopic repertoire into self and nonself. The only solution is that this is the role of developmental time. There must be a window during ontogeny when all self is expressed and no nonself. Further, self must persist throughout life to maintain the sorting process and no new self can appear after the window closes (i.e., the immune system becomes responsive). The *Aire*-controlled ectopic expression of peripheral antigens in thymus is part of this mechanism, presumably to deal with antigens that are expressed delayed in the periphery [66, 67]. The presently wide spread denial of a role for developmental time (e.g., [51]) is, in part, based on a handful of failures to experimentally reveal it. These are negative results explicable in many ways, the most likely being the contamination of the antigen with nonspecific activating agents like lipopolysaccharide or an eclipse period in the expression of the antigen. Lastly, tolerance must be mediated epitope by epitope and nonself markers mediate tolerance, antigen by antigen, making impossible nonself marker regulation of antigens that share epitopes with self. These latter comprise about 10% of all antigens and cannot be neglected. A role for self and nonself markers in the self-nonself discrimination has not been disproven by experiment; it has been disproven by logic in the context of the ARA model. A successful theory should challenge observation with the same validity that observation challenges theory. Observation and theory must be reciprocally interactive.

This does not mean that the concept of a nonself marker has no role in immune responsiveness. Germ line-encoded nonself markers may potentially play a role in module 3, the regulation of effector class (see below). However, the description of these ideas would be inappropriately referred to as "nonself markers." They are elements in germ line-encoded signaling pathways that tie the site and nature of the harm to the appropriate immune response. One should not throw out the baby with the bathwater (see regulation of effector function), although the nonself-marker theorists are trying hard to do just that by insisting that they play their roles in the self-nonself discrimination (module 2).

Suppression versus negative selection as elements in the self-nonself discrimination

This is a popular subject at the moment. There is today not a single publication in which the authors do not view suppression as a mechanism of tolerance (i.e., an element in the sorting of the paratopic repertoire, the self-nonself discrimination). By way of history, Gershon [68, 69] revealed the phenomenon in 1970 and extended it by analyzing a phenomenon known as "infectious tolerance." This was followed by an era lasting to today, in which the role of suppression has been debated. Once again, we ended up as a unique dissenting voice. Suppression plays its role in regulating the magnitude of the effector response, not as a determinant in the self-nonself discrimination. As we [56, 70–74] have analyzed many times, the various arguments ruling out suppression as a determinant in the self-nonself discrimination, here I would like to try a new tack by first developing a model in which suppression appears to play a central role in the self-nonself discrimination. I do this to illustrate the importance of placing two competing models on the table.

Under the ARA model, purging of anti-self (negative selection) is the outcome of signal 1; the residue, anti-nonself is activated to enter the pathway to effectors on receiving signals 1 + 2. Signal 2 is delivered by the Th anti-nonself. This is a

property of all i-cells, including iTsu and iTh themselves. No role for Tsu is envisaged in this pathway, the sorting of the repertoire.

Under a suppressor model, signal 1 is activating and signals 1 + 3 are inactivating; signal 3 is delivered by a T suppressor (Tsu), popularly referred to as Treg. No role for Th is envisaged in this pathway, the sorting of the repertoire.

A model for the self-nonself discrimination mediated by suppression is best initiated by a comparison of the origins of the paratopic specificities of the Th versus Tsu. Th are negatively selected in the thymus and periphery to be anti-nonself. By way of contrast, in order for Tsu to regulate the self-nonself discrimination, they must be positively selected to be anti-self in the thymus and periphery. The newborn Tsu anti-nonself dies by neglect in the thymus and periphery. The Tsu anti-nonself must, therefore, be derived independently and peripherally as a separate lineage, given that they carry out a quite different regulatory role.

The Tsu anti-self completely shut off any Th-driven anti-self response by acting via ARA to maintain specificity. This immediately poses two questions:

1- Where do the Tsu anti-nonself come from that regulates the response to nonself?
2- How is the response regulated for nonself-antigens that share epitopes with self?

The answer to the first question could be that the Tsu anti-nonself can be derived peripherally as a separate lineage from $CD4^+$ Th that had been negatively selected in thymus to be anti-nonself, and there is evidence for this [75–78]. The two lineages of Tsu would have to express distinctly different regulatory properties. The Tsu anti-self would have to shut off the response to self essentially completely, whereas the Tsu anti-nonself must fine-tune the response to nonself so that it is large enough to be adequately protective but not so large that it triggers a debilitating level of innocent bystander pathology.

The second question poses a regulatory problem faced also by the ARA model. It would be lethal if the Tsu anti-self shut off the response to nonself that is cross-reactive with self. This implies a Yin-Yang relationship between eTh and eTsu that permits a response to these antigens [56]. Parenthetically, under the ARA model, all classes of i-cell including iTh and iTsu themselves require an eTh-delivered signal 2 to be activated to become e-cells (effectors).

Given this coherent model of Tsu ontogeny, why argue that the postulated thymically derived Tsu anti-self cannot regulate the self-nonself discrimination; that is, they cannot be selected to be anti-self in the thymus? In other words, as normal functional entities, Tsu anti-self do not exist.

1- Returning now to the polyreactive TCR, if Tsu were positively selected in thymus to be anti-self, then the paratopic repertoire would be skewed towards high polyreactivity. If they are negatively selected to be anti-nonself like all of the other classes of T cell, the repertoire would be skewed towards low polyreactivity. As most of the ligands recognized by a polyreactive TCR are nonself, the selection to be anti-self would result in a repertoire of TCRs, which expresses a large anti-nonself library. As these anti-self TCRs are postulated to shut down any response to self, they would also shut down the response to nonself, a lethal situation.
2- The above argument can be made more specific. If Tsu (Tregs) determined tolerance, grafts between individuals of a species would be accepted, not rejected. It should be recalled that tolerance must be antigen specific, and to accomplish this, the Tsu must function via ARA. Individuals in a species express self-antigens that share epitopes.
3- The T helper, deregulation of which is responsible for the initiation of autoimmunity, would have to be the major target of Tsu suppression. This requires that the Tsu and Th display the same repertoire. This would not be possible if Tsu are positively selected to be anti-self and Th are negatively selected to be anti-nonself.
4- If tolerance to the self-antigens on T and B cells themselves were due to Tsu, no immune response would be possible. Obviously, suppressive tolerance to the self-components of Tsu themselves cannot be due to Tsu. Tsu (Tregs) cannot suppress Tsu (Tregs) and be functional. Tolerance to the self-components of Tsu itself must be due to negative selection.

If Tsu anti-self cannot be the mediators of tolerance, then Tsu are expected to be negatively selected in thymus as are all other T cells and leave to the periphery as Tsu anti-nonself. This is crucial as Tsu perform a central role in regulating the magnitude of the effector response. They play no role in the self-nonself discrimination. Tsu can be manipulated by the experimenter to dampen an autoimmune response or permit graft acceptance, a subject in and of itself, but such findings are not extrapolatable to the mechanism used to sort the repertoire (i.e., the self-nonself discrimination).

One-signal versus two-signal models

Lederberg was the first to propose a one-signal model. As mentioned above, he postulated that cells are born inactivatable only and, after a suitable time, differentiate to activatable only. This model requires, as does the two-signal model, a period in developmental time when all self is present and no nonself, in order to get started. The purging of anti-self occurs as long as the self-component persists. An anti-nonself repertoire accumulates during this period, reaching a steady state level that is maintained throughout life. When nonself

appears in the system, all de novo generated anti-nonself cells being inactivatable only are deleted by the nonself-antigen. The animal responds to nonself using the anti-nonself repertoire generated in its absence. The time spent in differentiating to activatable only is a key parameter. If it is too long, then the response to nonself will be compromised; if it is too short, anti-self will sneak through as an inducible-only cell.

This elegant model was ruled out when it was discovered that antigen-responsive B and T cells undergo mutation or receptor editing (revision) that would generate anti-self cells when they would be in the inducible-only state. It could not deal with de novo anti-self generated in cells in the inducible-only state.

This one-signal model in a somewhat circumscribed form turned out to be a crucial element in the ARA model in order to deal with the primer problem and ended up being swallowed up by it (see below).

Now, let us consider two unresolved elements of the ARA model itself [57].

The mechanism of ARA

It is embarrassing that a central element of the default ARA model has yet to be convincingly reduced to mechanism. In order to mediate ARA, the eTh must communicate signal 2 to the a-cell, T or B, via a platform that contains only one antigen, referred to as a signaling patch. Two signaling patch models have been proposed:

1- The APC picks up antigen as an antigen-antibody complex that is processed into a signaling patch across which the eTh delivers signal 2 to the aT cell [79].
2- The B cell is the signaling patch and the only "APC" that can present antigen for an iTh-eTh signaling interaction [80].

In both cases, the antigen must be processed to peptides that are kept together in the patch and the antigens endogenous to the B cell or APC must somehow be blocked from expression when a patch is being presented.

Whatever be the merits or demerits of each solution, they both appear to be ruled out by the finding that immunoglobulin-negative or B-less animals can mount a T cell response [81–83]. There are only two escapes from this impasse. Either the cited experiments can be argued to be not good enough to rule out signaling patches as they were performed in a different framework. In this case, we must await the definitive experiment. Or we must try to see if there is an equivalent way to achieve ARA that does not require a signaling patch [57].

The origin of primer eTh

All primers must arise independently of the process that they initiate. For example, the synthesis of glycogen in the liver requires a primer molecule of glycogen to get started. The synthesis of DNA requires a molecule of DNA to get started. And so, the initiation of an immune response requires an antigen-independent source of primer eTh anti-nonself.

From its very inception, the ARA model ran into what was referred to as the "chicken and egg problem." At the Brooks Lodge meeting in 1968, the failure to present a solution to the origin of the primer necessary to initiate responsiveness meant that no one took the model seriously [44]. The problem arises under the ARA model because all i-cells, T or B including iTh itself, require an eTh-delivered signal 2 to be activated. What is the origin of this primer eTh anti-nonself?

It is a matter of simple logic that any and all solutions require an antigen-independent step in the differentiation pathway from iTh to eTh. This recalls the Lederberg model, albeit greatly limited and circumscribed. The antigen-independent pathway must be a property of the Th lineage only. As an antigen-independent step cannot distinguish iTh anti-self from iTh anti-nonself, the pathway must include a discriminatory mechanism. We have proposed that the generation of a primer repertoire of eTh anti-nonself is the resultant of a slow antigen-independent pathway to eTh compared to a rapid blocking of this pathway by interaction with self (signal 1) [53, 57, 66, 84–86]. This would result in a priming eTh anti-nonself population sufficiently devoid of anti-self.

The regulation of effector function (module 3)

This aspect of the immune response is rich in data but limited in concept. Module 3 is crucial in that evolutionary selection operates at the level of effector output. Modules 1 and 2 are sculpted during evolution by their contribution to the efficiency and safety of the effector output. Given this, it is important to keep in mind that modules 1 and 2 result from somatic evolution, whereas module 3 is determined by germ line evolution.

Given an anti-nonself paratopic repertoire, function requires that it be appropriately coupled to an array of effector activities. This necessitates mechanisms to (1) home this activity to the site of infection, (2) to select the expression of the optimal class of effector activity, and (3) to feedback regulate the magnitude of the response.

Homing to the site of infection must involve recognition by circulating i-cells of antigens specific to the harmful agent or of a signal from the harmed target. Selection of effector class involves answering two questions: (1) is the pathogen intracellular or extracellular and (2) what signals does the activated g-cell receive from the injured tissue?

In order for an intracellular pathogen to become immunogenic, it must be processed to peptide presented by RII (PRII) the ligand for Th. Most cells parasitized by a virus do not express RII making the pathway to becoming a PRII ligand, a necessity. As dominantly B cells and professional antigen-

presenting cells (APCs) express RII, transport of the antigens of intracellular pathogens to B cells/APC is mandated. It seems likely that the virus (or other) is released by the infected cell and is picked up specifically (via the BCR) by B cells acting as APC that present PRII. The primer eTh interacting with PRII expressed by these B cells deliver signal 2 that initiates the pathway to eB-secreting antibody. This antibody forms complexes with the virus that permit optimal uptake by the APC, where a signaling patch expressing a unique antigen as PRII and PRI permits induction to eTh and eTc (cytotoxic T cells) in ARA. The amplification of this whole process can be viewed as autocatalytic and is dependent on the rate of increase of eTh.

There are two comments relative to this scenario that merit citing. First, the B cell involved in the initiation of this pathway can be justifiably described as primer eB in the same sense that we refer to primer eTh [79]. Two primers are needed to initiate the response that the APCs take over by amplifying it to an effective level. Second, in order to explain induction of T cells in B-less mice (other than its being an experimental glitch), one can envisage that the APC has an inefficient antibody-independent uptake system derived from its evolutionary history when it was a simple macrophage [87]. This particle-sensing mechanism would be unspecific, unable to distinguish self from nonself. However, thanks to module 2, the sorting of the repertoire, there would be an insufficiency of i-cells anti-self, and in a short-term experiment, the response to a nonself-antigen, as observed, would dominate. In the long run, B-less animals would be expected to suffer T cell-driven autoimmunity.

Lastly, it might be pointed out that effector T cells can be biodestructive but they are unable to rid the resultant debris. The debris is ridded by a combination of innate driven effector activity and antibody-driven phagocytosis. The system must switch from cell-mediated to humoral if the debris from its biodestructive activity is to be ridded by the adaptive system.

Up to this point, we have explained the induction of virgin iT/iB cells to effectors, but we have not considered what the signals are that direct these cells to undergo further differentiation to the various subclasses or isotypes. Most reasonable would be that the activated cells receive messages (chemokines, interleukins, hormones) from the harmed tissues [88]. There is a hint that this be the case as, for B cells, in-frame and out-of-frame rearrangements appear to switch to the same isotype ([89], hypothesis VII). If the switch were random and the isotype chosen by subsequent selection, both chromosomes would only rarely (stochastically) switch to the same isotype. Solidly establishing this point experimentally is crucial. This is where danger theory (more precisely the trauma model) becomes helpful because it predicts messages from the injured tissue that direct class determination [88].

Lastly, as pointed out above, the magnitude of the effector response must be regulated at a level high enough to efficiently rid the "pathogen," yet maintained low enough to minimize innocent bystander pathology. This is the postulated role of the T suppressor (Treg), the mechanism of which needs attention.

All that I tried to do in this section is illustrate how knowledge of modules 1 and 2 affect the way one looks at module 3.

Intracellular signaling pathways

This is the most popular subject of today's research activity. The reason is that deletion or regulation of one or the other of these signaling factors has the potential to permit control of a deregulation resulting in immune pathology. So, a comment on the basic biology of intracellular signaling pathways might be relevant.

A cell can be envisaged to express on its surface between 10 and 100 signaling receptors. Each input is coupled to a unique output. This latter is determined by which genes are transcriptionally activated. In the extreme, possibly the most frequent relationship between signaling input and output would be one signaling input–one unit of activated genes (i.e., a "regulon"). This would require a step in the pathway that ties input to output in a unique way (one input–one output). Let us look at a simple example taken from the regulation of effector function. The B cell is activated expressing an IgM BCR and awaits a signal telling it to which isotype to switch. It has the potential to switch to any one of ~10 isotypes. A given infection in the gut or lung signals switch to IgA; in the liver signals IgG2; and at the maternal–fetal interface signals IgG3, i.e., three inputs–three outputs. What steps in each signaling pathway permits this one-to-one relationship? Until we understand or at least can model the steps that govern the intracellular signaling factors that link in a specific way input to output, no unifying concept will emerge.

A general comment

This essay is a personalized view that takes us from the past to the present. The errors of the past are used to criticize the conceptualizations of the present. As was also true in the past, there is today a majority position that ignores the minority position instead of engaging it. We have seen that a majority position formed by pure observation has all too often proven wrong. Being wrong, in and of itself, would be a major step forward, if it were also rationalized. In general, as we have seen, the majority position can lack that essential element. The only way to appreciate this is to put on the table for discussion, two competing conceptualizations.

A personal note

All of the players in this story were known to me either as friends or acquaintances. I have tried to separate the science from the individual but to the naïve reader this attempt is all too often misconstrued. My only way to keep this clear is to state it frontally.

During my scientific career, I have had the privilege to work with and appreciate the major contributions of three different types of scientist. Burnet, Jerne, and Lwoff were essentially intuitive arriving at important conceptualizations that they, themselves found difficult to rationalize. Monod, Jacob, Lederberg, and Pauling were essentially Cartesian relying on: I do not believe what I see unless it is supported by a rationalized theory. Landsteiner, J. R. Marrack, and Benacerraf were essentially empiric advancing in small steps from X to Y. Kabat changed during his career from empiric to Cartesian. Of course, there is a degree of overlap between these characterizations as the extremes are not viable. Lastly, with the appearance of computer simulation modeling, there is a new category of scientist emerging, namely, the empiricist in theoretician's clothing.

Acknowledgement The author is deeply appreciative of the assistance of Bernice Walker in the preparation of this manuscript.

Compliance with ethical standards

Conflict of interest The author declares that they have no conflict of interest.

References

1. Pauling L. A theory of the structure and process of formation of antibodies. J Am Chem Soc. 1940;62:2643–57.
2. Cohn M. The wisdom of hindsight. Ann Rev Immunol. 1994;12:1–62.
3. Cohn M. Logic of the self-nonself discrimination: principles and history. In: Cambrosio A, Moulin A, editors. Dialogues with selves. Historical issues and contemporary debates in immunology. Paris: Editions Elsevier; 2001. p. 53–85.
4. Burnet FM, Fenner F. Genetics and immunology. Heredity. 1948;2:289–325.
5. Burnet FM, Fenner F. The production of antibodies. Melbourne: Macmillan and Company Limited; 1949. p. 142.
6. Owen RD. Immunogenetic consequences of vascular anastomoses between bovine twins. Science. 1945;102:400–1.
7. Jerne NK. The natural selection theory of antibody formation. Proc Natl Acad Sci U S A. 1955;41:849–56.
8. Burnet FM. A modification of Jerne's theory of antibody production using the concept of clonal selection. Aust J Sci. 1957;20:67–9.
9. Burnet FM. The mechanism of immunity. Sci Am. 1961;204:158–64.
10. Attardi G, Cohn M, Horibata K, Lennox ES. Symposium on the biology of cells modified by viruses or antigens. II. On the analysis of antibody synthesis at the cellular level. Bacteriol Rev. 1959;23:213–23.
11. Burnet FM. The clonal selection theory of acquired immunity. Nashville: Vanderbilt University Press; 1959. p. 209.
12. Lederberg J. Genes and antibodies. Science. 1959;129:1649–53.
13. Billingham RE, Brent L, Medawar PB. Actively acquired tolerance of foreign cells. Nature. 1953;172:603–6.
14. Anderson D, Billingham RE, Lampkin GH, Medawar PB. The use of skin grafting to distinguish between monozygotic and dizygotic twins in cattle. Heredity. 1951;5:379–97.
15. Billingham RE, Lampkin GH, Medawar PB, Williams HR. Tolerance to homografts, twin diagnosis and the freemartin condition in cattle. Heredity. 1952;6:201–12.
16. Lederberg J, Lederberg EM. Replica plating and indirect selection of bacterial mutants. J Bacteriol. 1952;63:399–406.
17. Jerne NK. The somatic generation of immune recognition. Eur J Immunol. 1971;1:1–9.
18. Lennox ES, Cohn M. Immunoglobulins. Ann Rev Biochem. 1967;36:365–406.
19. Bretscher PA. Is the framework of Cohn's "Tritope model" for how T cell receptors recognize peptide/self-MHC complexes and allo-MHC plausible? Scand J Immunol. 2016;83:311–5.
20. Brent L. A history of transplantation immunology. San Diego: Academic Press; 1997. p. 482.
21. Langman RE, Cohn M. The E-T (elephant-tadpole) paradox necessitates the concept of a unit of B-cell function: the protecton. Mol Immunol. 1987;24:675–97.
22. Cohn M. A stepwise model of polyreactivity of the T cell antigen-receptor (TCR): its impact on the self-nonself discrimination and on related observations (receptor editing, anergy, dual receptor cells). Cell Mol Life Sci. 2014;71:2033–45.
23. Cohn M, Langman RE. The protecton: the evolutionarily selected unit of humoral immunity. Immunol Rev. 1990;115:1–131.
24. Weigert MG, Cesari IM, Yonkovich SJ, Cohn M. Variability in the lambda light chain sequences of mouse antibody. Nature. 1970;228:1045–7.
25. Cohn M. Antibody diversification: the somatic mutation model revisited. In: Pollak JK, Lee JW, editors. The biochemistry of Gene expression in higher organisms. Dordrecht: Reidel Publishing Co.; 1973. p. 574–92.
26. Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C (eds). Sequences of proteins of immunological interest (NIH Publication No. 91–3242), (1991), 5th Ed, Pages.
27. Matsuda F, Ishii K, Bourvagnet P, Kuma K, Hayashida H, Miyata T, Honjo T. The complete nucleotide sequence of the human heavy chain variable region locus. J Exp Med. 1998;188:2151–62.
28. Hirst JW, Jones GG, Cohn M. Characterization of a BALB/c myeloma library. J Immunol. 1971;107:926–7.
29. Matzinger P, Bevan MJ. Induction of H-2 restricted cytotoxic T-cells: in vivo induction has the appearance of being unrestricted. Cell Immunol. 1977;33:92–100.
30. Langman RE. (1978) Cell-mediated immunity and the major histocompatability complex. In: Rev. Physiol. Biochem. Pharmacol., Vol. 81. Springer-Verlag, New York, 0 Ed, 1–34.
31. Cohn M, Epstein R. T-cell inhibition of humoral responsiveness. II. Theory on the role of restrictive recognition in immune regulation. Cell Immunol. 1978;39:125–53.
32. Zinkernagel RM, Doherty PC. Immunological surveillance against altered self components by sensitised T lymphocytes in lymphocytic choriomeningitis. Nature. 1974;251:547–8.
33. Babbitt BP, Allen PM, Matsueda G, Haber E, Unanue ER. Binding of immunogenic peptides to Ia histocompatibility molecules. Nature. 1985;317:359–61.
34. Townsend AR, McMichael AJ, Carter NP, Huddleston JA, Brownlee GG. Cytotoxic T cell recognition of the influenza nucleoprotein and hemagglutinin expressed in transfected mouse L cells. Cell. 1984;39:13–25.

35. Cohn M. On the logic of restrictive recognition of peptide by the T-cell antigen receptor. Immunol Res. 2011;50:49–68.
36. Zy-Tmg E, Chu C, Carswell C, Cole BC, Jones PP. The minimal polymorphism of class II E alpha chains is not due to the functional neutrality of mutations. Genetics. 1994;40:9–20.
37. Wucherpfennig KW, Allen PM, Celada F, Cohen IR, De Boer RJ, Garcia KC, Goldstein B, Greenspan R, Hafler D, Hodgkin PD, Huseby ES, Krakauer DC, Nemazee D, Perelson AS, Pinilla C, Strong RK, Sercarz EE. Polyspecificity of T cell and B cell receptor recognition. Sem Immunol. 2007;19:216–24.
38. Mason D. A very high level of crossreactivity is an essential feature of the T-cell receptor. Immunol Today. 1998;19:395–404.
39. Nemazee D. Antigen receptor "capacity" and the sensitivity of self-tolerance. Immunol Today. 1996;17:25–9.
40. Percus JK, Percus OE, Perelson AS. Predicting the size of the T cell receptor and antibody combining region from consideration of efficient self-nonself discrimination. PNAS. 1993;90:1691–5.
41. Whitaker L, Renton AM. On the plausibility of the clonal expansion theory of the immune system in terms of the combinatorial possibilities of amino-acids in antigen and self-tolerance. J Theor Biol. 1993;164:531–6.
42. Langman RE. The specificity of immunological reactions. Mol Immunol. 2000;37:555–61.
43. Bretscher P, Cohn M. A theory of self-nonself discrimination. Science. 1970;169:1042–9.
44. Cohn M. Speculation on cellular and molecular mechanisms involved in immunological responsiveness and tolerance. In: Immunological Tolerance. New York: Academic Press; 1969. p. 281–338.
45. Cohn M, Blomberg B. The self-nonself discrimination: a one- or two-signal mechanism? Scand J Immunol. 1975;4:79–107.
46. Cohn M. Conversations with Niels Kaj Jerne on immune regulation: associative versus network recognition. Cell Immunol. 1981;61:425–36.
47. Cohn M. What are the "must" elements of immune responsiveness? In: Neural modulations of immunity. New York: Raven Press; 1985. p. 3–25.
48. Cohn M, Langman RE. To be or not to be ridded?—that is the question addressed by the associative antigen recognition model. Scand J Immunol. 2002;55:318–23.
49. Cohn M. The common sense of the self-nonself discrimination. Springer Semin Immunopathol. 2005;27:3–17.
50. Cohn M. Conceptualizing the self-nonself discrimination by the vertebrate immune system. In: Timmis J, Flower D, editors. In Silico Immunology. New York: Springer; 2007. p. 375–98.
51. Cohn M. On the opposing views of the self-nonself discrimination by the immune system. Immunol Cell Biol. 2009;87:113–9. Discussion 120-121
52. Cohn M. The discussion with Jacques Miller: illustrating the limitations of pure empiricism. Immunol Cell Biol. 2009;87:435–7.
53. Cohn M. The self-nonself discrimination: reconstructing a cabbage from sauerkraut. Res Immunol. 1992;143:323–34.
54. Cohn M. The evolutionary context for a self-nonself discrimination. Cell Mol Life Sci. 2010;67:2851–62.
55. Cohn M. What is so special about thinking; after all, we all do it! Exp Mol Pathol. 2012;93:354–64.
56. Cohn M. Autoimmunity: comments on the pathway from initiation to disease. J Theor Biol. 2015;375:40–51.
57. Cohn M. Two unresolved problems facing models of the self-nonself discrimination. J Theor Biol. 2015;387:31–8.
58. Soderquist T. The troubled life of Niels Jerne. New Haven: Yale University Press; 2003.
59. Cohn M. Comments for 10th forum on immunology, "is the immune system a functional idiotypic network?". Ann Inst Pasteur/Immunol. 1986;137C:173–88.
60. Cohn M. Idiotype network views of immune regulation: for whom the bell tolls. In: Capra JD, editor. Idiotypy and medicine. New York: Academic Press; 1986. p. 321–99.
61. Cohn M. The concept of functional idiotype network for immune regulation mocks all and comforts none. Ann Inst Pasteur/Immunol. 1986;137C:64–76.
62. Berek C, Etlinger H, Julius M. Idiotypes: antigens on the inside. (Editiones < Roche > Basel Switzerland). 1981.
63. Langman RE, Cohn M. The "complete" idiotype network is an absurd immune system. Immunol Today. 1986;7:100–1.
64. Cohn M. Analysis of Paris meeting redefining the "self" of the immune system. Immunol Res. 2015;62:106–24.
65. Eichmann K. The network collective, rise and fall of a scientific paradigm. Berlin: Birkhaauser; 2008. p. 274.
66. Langman RE, Mata JJ, Cohn M. 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 Immunol. 2003;15:593–609.
67. Cohn M. Why Aire? Compensating for late bloomers. Eur J Immunol. 2009;39:2969–72.
68. Gershon RK, Kondo K. Cell interactions in the induction of tolerance: the role of thymic lymphocytes. Immunology. 1970;18:723–5.
69. Gershon RK. T-cell control of antibody production. Contemp Top Immunobiol. 1974;3:1–49.
70. Langman RE. The self-nonself discrimination is not regulated by suppression. Cell Immunol. 1987;108:214–9.
71. Cohn M. Whither T-suppressors: if they didn't exist would we have to invent them? Cell Immunol. 2004;227:81–92.
72. Cohn M. What roles do regulatory T-cells play in the control of the adaptive immune response? Int Immunol. 2008;20:1107–18.
73. Cohn M. Musings about regulation by T-suppressors: a response to the commentary by Kristofor Ellestad on "Meanderings into the regulation of effector class by the immune system: derivation of the trauma model". Scand J Immunol. 2012;76:92–8.
74. Cohn M. What would Treg-cell biology look like when viewwed from a rationalized perspective? Eur J Immunol. 2015;45:3002–9.
75. Chen W, Jin W, Hardegen N, Lei KJ, Li L, Marinos N, McGrady G, Wahl SM. Conversion of peripheral $CD4^+CD25^-$ naive T cells to $CD4^+CD25^+$ regulatory T cells by TGF-beta induction of transcription factor Foxp3. J Exp Med. 2003;198:1875–86.
76. Zheng SG, Wang JH, Gray JD, Soucier H, Horwitz DA. Natural and induced $CD4^+CD25^+$ cells educate $CD4^+CD25^-$ cells to develop suppressive activity: the role of IL-2, TGF-beta, and IL-10. J Immunol. 2004;172:5213–21.
77. Kretschmer K, Apostolou I, Hawiger D, Khazaie K, Nussenzweig MC, von Boehmer H. Inducing and expanding regulatory T cell populations by foreign antigen. Nat Immunol. 2005;6:1219–27.
78. Hall JA, Grainger JR, Spencer SP, Belkaid Y. The role of retinoic acid in tolerance and immunity. Immunity. 2011;35:13–22.
79. Cohn M. A biological context for the self-nonself discrimination and the regulation of effector class by the immune system. Immunol Res. 2005;31:133–50.
80. Bretscher PA. A two-step, two-signal model for the primary activation of precursor helper T cells. Proc Natl Acad Sci U S A. 1999;96:185–90.
81. Dittel BN, Urbania TH, Janeway CAJ. Relapsing and remitting experimental autoimmune encephalomyelitis in B cell deficient mice. J Autoimmun. 2000;14:311–8.
82. Epstein MM, DiRosa F, Jankovic D, Sher A, Matzinger P. Successful T cell priming in B cell-deficient mice. J Exp Med. 1995;182:915–22.
83. Brundler M-A. Aichele P, Bachmann mea. Immunity to viruses in B cell-deficient mice: influence of antibodies on virus persistence and on T cell memory. Eur J Immunol. 1996;26:2257–62.

84. Cohn M. Antibody diversity 1983: some elementary considerations. In: Yamamura Y, Tada T, editors. Progress in immunology V. Orlando: Academic Press; 1983. p. 839–51.
85. Cohn M. How does the immune response get started? Cell Immunol. 2009;254:91–3.
86. Cohn M. Core principles characterizing immune function. Eur J Immunol. 2016; doi:10.1002/eji.201646706.
87. Langman RE. The immune system. San Diego: Academic Press; 1989. p. 207.
88. Cohn M. Meanderings into the regulation of effector class by the immune system: derivation of the trauma model. Scand J Immunol. 2012;76:77–88.
89. Cohn M. Ten experiments that would make a difference in understanding immune mechanisms. Cell Mol Life Sci. 2012;69:405–12.