22

The Immunological Homunculus and Autoimmune Disease

IRUN R. COHEN

Department of Cell Biology, The Weizmann Institute of Science, Rehovot, Israel

THE IMMUNE REFLEX

How we interpret the purpose of the immune system biases how we view self tolerance and autoimmunity. Although nature seems to behave as if she were oblivious to our expectations, our fundamental views are worth examining because they usually determine what we notice.

A prevalent view of the immune system is that its purpose is recognition, specifically recognition of the foreign. Recognition is central because of the premise that recognition of an antigen automatically triggers an attack, in the manner in which a stimulus leads to a response. Immunity, according to this view, is a reflex. The term immune *response* implies as much.

The notion of immunity as reflex is inherent in Burnet's interpretation of clonal selection: recognition, which is the purpose of the immune system, is also the regulator of the immune response (Burnet, 1959). An antigen becomes a stimulus once it is recognized by a lymphocyte clone whose response generates both effectors (antibodies and T cells) and memory (an amplified, persisting population of clonal descendants).

Systems with a direct reflex connection between stimulus and response often can be regulated by varying the strength of the stimulus. In the case of an immune response to a foreign invader, the invader introduces the antigens; the immune response produces its own negative feedback by eliminating the invader and so reducing the antigenic stimuli. Regulation of autoimmunity, in contrast, cannot be had by eliminating self antigens, tantamount to self-destruction. According to Burnet, it is the recognition of self antigens that is eliminated. If there is no recognition there will be no stimulus and hence there will be no response and no disease. The self-

MOLECULAR AUTOIMMUNITY
ISBN 0-12-682348-0

recognizers, according to orthodox Burnetians, must be eliminated; clonal deletion does the job (Burnet, 1959).

The logic of the Burnetian view of the immune reflex fostered the expectation that in the end clonal deletion would be shown to be the foundation of self-tolerance. This expectation was recently bolstered by the experimental demonstration of negative selection of some T cells in the thymus (Blackman *et al.*, 1990; von Boehmer, 1990). However, the thymus need not be the only site of lymphocyte inactivation; mice expressing transgenes have taught us that contact with antigens outside of the thymus can induce stable non-responsiveness, anergy, to the transgenic antigen (Burkly *et al.*, 1990; Goodnow *et al.*, 1990). Although their molecular mechanisms are obscure, anergy induced in the periphery is not different in purpose from negative selection occurring in the thymus: both processes rid the immune system of certain recognizers.

Autoimmune disease, according to Burnet and the neoburnetians, erupts from acts of self-recognition committed by forbidden clones of lymphocytes that have arisen, despite and distal to negative selection and anergy, by some chance mutation (Burnet, 1959). In addition, self-antigens that are not represented in sufficient concentration in the thymus or elsewhere may fail to affect their recognizers. Irrespective of the history of the self-recognizers, recognition of self is felt to foster autoimmune disease by a reflex autoimmune response, just as recognition of the foreign invader produces health through reflex rejection of the invader.

The notion of the immune reflex leads to two predictions about autoimmunity: (1) Self-recognition is incompatible with health; the immune system must attack what it notices. (2) Autoimmune diseases are unpredictable; forbidden clones arise at random. Negative selection notwithstanding, both of these predictions are refuted by experience.

SELF-RECOGNITION IS COMPATIBLE WITH HEALTH

Ample evidence attests to the existence of autoimmune B and T cells in health animals and humans. Natural and benign autoantibodies are detectable in every body (Avrameas *et al.*, 1983; Shoenfeld and Isenberg, 1989); their specificities range from ubiquitous target molecules (anti-DNA, anti-cytochromes) to tissue-specific molecules (anti-thyroglobulin, anti-myelin) and include elements of the immune system itself (rheumatoid factors, anti-idiotypes). Remarkably a distinct lineage of B cells, the Ly-1 or $CD5^+$ B cells, existing from birth or even from fetal life, appears to specialize in producing autoantibodies (Hayakawa and Hardy, 1988). One may quibble about whether the B cells producing natural autoantibodies are only 1% or as much as 30% of the B cell pool, or about what fraction

of the natural autoantibodies are IgG (small) and what fraction IgM (large); one may try to demean them as being non-pathogenic (another way of saying they are compatible with health); or disparage their affinity calling them polyreactive; the conclusion cannot be denied: the healthy immune system is busy making autoantibodies.

Much less is known about natural autoimmune T cells; there are fewer comprehensive studies. Nevertheless, from the earliest experiments done by me and Wekerle it was clear that naturally autoreactive cytotoxic lymphocytes (later known as T cells) were present in healthy rats (Cohen and Wekerle, 1977). Autoimmune T cells with reactivity to myelin basic protein (Wicherpfennig *et al.*, 1990), to human epitopes of the 65 kD heat shock protein (hsp65) (Munk *et al.*, 1989), or to insulin (Naquet *et al.*, 1988) have since been detected in healthy humans. Thus at least some autoimmune T cells, like many autoimmune B cells, are not deleted or anergic but are alive and active in the healthy immune system. Autoimmunity by itself does not necessitate disease.

AUTOIMMUNE DISEASES ARE ORDERLY

The visitation of one or another autoimmune disease on a single person may seem to be random bad luck, the unpredictable misconduct of a forbidden clone. However, when we examine the population rather than the individual, we see that the clinical picture of the disease and its immunology are predictable to a high degree. Forbidden clones randomly appearing should cause autoimmune disease of random nature and random immunological specificity. Every sick person should have a disease tailored by the specificity of the forbidden clones that happen to emerge. No two diseases should be alike (except by chance) and autoimmunity to the same organ in different people should be associated with forbidden clones reactive to different target antigens within the particular tissue. But this is not what we observe.

What we observe in the population is no more than a few dozen autoimmune diseases, each marked by reactivity to a few characteristic self antigens (Shoenfeld and Isenberg, 1989). Each tissue of the body seems to have a particular cell-type and set of autoantigens destined to be the target of autoimmune disease arising in that tissue. For example, autoimmune disease affecting the pancreatic islets habitually attacks the β cells making insulin (Castano and Eisenbarth, 1990); the α cells making glucagon seem never to suffer autoimmune attack. Insulin-dependent diabetes mellitus (IDDM), the disease resulting from destruction of the β cells, seems to be associated with a standard set of autoimmune reactivities to a limited number of target antigens: a 64-kD enzyme (Baekkeskov *et al.*, 1990), a

65-kD heat shock protein (Elias *et al.*, 1990), and insulin (Palmer *et al.*, 1983). In general, it seems that the same autoantigens are habitually the targets of immune responses in different people, even in the face of other competing antigens. Antigens that are usually selected for special attention by the immune system may be termed dominant antigens. The target antigens of autoimmune diseases are immunologically dominant. Hence, the diseases and their immunology are not random events.

The regularity of autoimmunity is also evident in the shared idiotypes characteristic of autoantibodies (Shoenfeld and Isenberg, 1989) and in the restricted usage of T cell receptor genes expressed in some diseases (Acha-Orbea *et al.*, 1989).

These attributes of autoimmune disease are not compatible with the concept of forbidden clones; the clones are allowed, the diseases are predictable, the reactivities are stereotyped. Within the population autoimmune diseases express law and order despite the chaos they perpetrate in the individual life. Note that these observations about autoimmunity do not negate the fact of clonal selection; recognition is fundamental to all immune behavior. The regularity of autoimmune disease also does not contradict negative selection of T cells in the thymus, which is a fact. The regularity of autoimmunity does tell us that autoimmune disease is the expression of ordered components; it has structure. Regularity implies regulation. The immunology of the hsp65 molecule is illustrative (Young, 1990).

SELF–NON-SELF DISCRIMINATION AND hsp65

The hsp65 molecule is interesting as an antigen because of its great conservation throughout evolution: the hsp65 of prokaryotes and humans has about 50% homology overall, with many stretches of sequence identity (Jindal *et al.*, 1989). Consequently, the hsp65 molecule of any bacterial invader comprises epitopes that are foreign along with epitopes that are self to any mammalian host. A paradox of immunity to hsp65 is its association with both health and autoimmune disease.

Adjuvant arthritis can be induced in susceptible strains of rats by immunizing them with killed *M. tuberculosis* bacteria (Pearson, 1964). My associates and I discovered that a T cell clone capable of transferring arthritis recognized a nine amino acid peptide of mycobacterial hsp65 (van Eden *et al.*, 1988). Although this epitope is not present in mammalian hsp65 (see Jindal *et al.*, 1989), it seems to mimic a self epitope present in cartilage (van Eden *et al.*, 1985; Cohen, 1988).

Irrespective of the molecular details, adjuvant arthritis exemplifies an autoimmune disease triggered by antigenic mimicry between molecules of

host and microbe. The connection between hsp65 and rat arthritis led to the finding that rheumatoid arthritis of humans might also be associated with T cell immunity to hsp65 (Res *et al.*, 1988).

Unexpectedly, we discovered a second autoimmune disease involving T cell immunity to hsp65: the spontaneous autoimmune diabetes of NOD mice. IDDM in NOD mice (and in humans) is caused by destruction of the insulin-producing β cells of the pancreas (Castano and Eisenbarth, 1990). Serendipitously, we found that a key target epitope in NOD mice is expressed on the hsp65 molecule (Elias *et al.*, 1990). In this case the epitope is on the mouse's own hsp65 molecular; the mycobacterial sequence is only weakly cross-reactive with the true mammalian sequence (Elias *et al.*, 1991). The evidence for the causal connection between T cell immunity to the mammalian hsp65 epitope is as follows (Cohen, 1991):

(1) The onset of spontaneous insulitis at +–6 weeks of age is associated with spontaneous T cell reactivity to the antigen. T cell reactivity to hsp65 wanes after the onset of clinical IDDM (at +–6 months) and the loss of β cells.
(2) T cell clones responsive to the hsp65 epitope transfer diabetes.
(3) Active immunization to hsp65 induces diabetes.
(4) T cell vaccination using attenuated anti-hsp65 T cells downregulates the anti-hsp65 immunity and aborts the destruction of β cells.
(5) IDDM can also be cured by inducing tolerance to a 24 amino acid peptide of hsp65.

In short, T cell immunity to the hsp65 epitope is both necessary and sufficient for IDDM in the NOD mouse. Although the expression of the hsp65 epitope in the islets needs to be clarified, the association of hsp65 immunity with diabetes satisfies the logical requirement for identifying an etiological agent – Koch's postulates adapted to autoimmunity.

A bacterial hsp65 molecule that looks like self and a mammalian hsp65 molecule that is self are plausible agents for inciting autoimmune disease, the difficulty of explaining tissue specificity notwithstanding. According to Burnetian ideas, a molecule of proven autoimmune potential such as hsp65 should be forbidden as a target in health. Nevertheless, it is a fact that hsp65 is an immunologically dominant antigen in health; one of every five T cells responsive to mycobacteria in immunized mice is directed to hsp65 (Kaufmann *et al.*, 1987). The immunological dominance of hsp65 is evident in immune responses to many different microbes: hsp65 has been called a common bacterial antigen (Young *et al.*, 1987). But hsp65 is also a common autoantigen; in the absence of overt immunization, T cells of healthy humans respond to self epitopes of hsp65 (Munk *et al.*, 1987).

Thus, a self or self-like molecule can cause autoimmune disease and yet

can serve as a dominant antigen in infection and a natural antigen in health. How can autoimmune disease be avoided if hsp65 is self-like, ubiquitous and immunologically dominant? What is dominance and how is it regulated?

IMMUNOLOGICAL DOMINANCE IS NECESSARY

Immunological dominance is the expression of focus: the ability of the immune system to direct its attention to a limited segment of the recognizable antigenic universe. Systems which process information, notably the nervous and immune systems, must be able to extract meaningful signals from the cacophony of stimuli impinging upon them. Just as a person cannot read a book and watch television and understand both simultaneously, the immune system cannot effectively respond to more than a limited number of the potential epitopes present on even a simple macromolecule. Without focus there is no discrimination between the ground and the subject, between the signal and the noise. Focus is essential to the creation of meaning. The immune system, like the nervous system, jams when it fails to restrain itself, to limit its field of attention (Cohen *et al.*, 1985). Polyclonal activation, a kind of super-response, is the equivalent of no response at all; a fact well-known to a number of successful parasites who avoid rejection by triggering polyclonal activation (Grossman *et al.*, 1986).

The astronomical number of antigen receptors comprising the immune repertoire ensures that it will be able to recognize large numbers of epitopes on any biological macromolecule, present or future. Polyclonal actiation is therefore inherent in the repertoire and could take place during any immune response. Focus is required to make sense out of the enormity of the antigenic world perceivable by the immune repertoire (Cohen and Young, 1991). Immunological dominance is the functional antithesis of chaotic polyclonal activation.

Immunological dominance, however, denotes more than clonal restraint, a goal that could be achieved by focus on any antigen among the many. Dominance indicates that a certain antigen tends to be the subject of focus; the focus is predictable. Predictability implies the existence of an ordering principle. What might it be?

The first ordering principle likely to come to mind is the MHC. By serving to present epitopes to T cells, the structure of the class I and II MHC molecules imposes structure on the immune response: some processed peptides may be favored and others may be inadmissible by particular allelic products (Bjorkman *et al.*, 1987). In this way allelic variations in the MHC can effectively filter much of what is available for T cell recognition.

Note, however, that the filter-function of the MHC operates on processed peptide fragments of antigens, it does not operate on whole antigens. Myelin basic protein (MBP), a dominant antigen in the central nervous system (CNS), illustrates this point well. MHC alleles are critical to autoimmunity (de Vries and van Rood, 1988) in that they determine which portion of a molecule will be recognized by T cells. For example Lewis rats see a peptide in the MBP sequence 68–88; BN rats see a peptide in the sequence 45–68. Since the 68–88 peptide contains the major encephalitogenic epitope of MBP, the Lewis rats develop experimental autoimmune encephalomyelitis (EAE) while the BN rats do not (Beraud *et al.*, 1986). However, the dominance of the MBP molecule as a whole is indifferent to allelic and species differences in MHC molecules; from mouse to man MBP is the dominant neuroantigen.

If the MHC encodes the dominance only of epitopes, what encodes the dominance of whole antigens? The results of study of T cell vaccination were helpful in probing this question.

THE PHYSIOLOGY OF T CELL VACCINATION

T cell vaccination is a procedure for prevention or therapy of autoimmune disease in which the subjects are vaccinated with their own autoimmune T cells (Cohen, 1986). The virulence of a sample of the autoimmune T cells is attenuated *in vitro* by treatment with irradiation (Ben-Nun *et al.*, 1981), hydrostatic pressure (Lider *et al.*, 1986) or chemical cross-linkers (Lider *et al.*, 1987). Effective T cell vaccination without transfer of the autoimmune disease may also be induced by inoculating animals with numbers of viable T cells below the threshold number needed to produce the disease; low-dose vaccination (Lider *et al.*, 1988; Beraud *et al.*, 1989). In all cases the T cells used for vaccination need to have been activated using the specific antigen or a mitogen within several days before administration (Naparstek *et al.*, 1983). Non-activated T cells do not effectively vaccinate.

T cell vaccination in its various forms has been successful in preventing EAE (Ben Nun *et al.*, 1981) or experimental autoimmune thyroiditis (EAT) (Maron *et al.*, 1983), and in preventing and treating adjuvant arthritis (Holoshitz *et al.*, 1983) or the IDDM of NOD mice (Elias *et al.*, 1991). It is effective against autoimmune disease induced by active immunization to the specific autoantigen as well as against disease produced by adoptive transfer of the autoimmune T cells in a virulent state.

Investigation of the resistance to autoimmune disease induced by T cell vaccination indicated the involvement of anti-idiotypic T cells directed to the specific autoimmune T cells (Lider *et al.*, 1988). Anti-ergotypic T cells

responsive to activation markers on effector T cells also play a role (Lohse *et al.*, 1989).

Effective T cell vaccines can be made using antigen-specific T cells obtained from immunized or spontaneously sick animals, and they can be made from T cell lines or even from T cell clones which recognize single defined epitopes on the target self antigen. Vaccination against EAE may also be achieved by immunization with peptides representing parts of the sequence of the β chain of the dominant anti-MBP T cell receptor (Howell *et al.*, 1989; Vandenbark *et al.*, 1989).

The order inherent in autoimmunity is expressed by these observations: if a T clone vaccinates individual animals against disease by inducing anti-idiotypic T cells, then the autoimmune T cells in the diseased animals and the T cell vaccine must all share T cell receptor idiotypes. In other words, antigenic dominance and idiotypic dominance are related: the T cells defining antigenic dominance share dominant idiotypes.

This is fortunate for T cell vaccination because the procedure for raising autoimmune T cell lines and clones is biased towards selecting T cells responsive to the most dominant antigens. Quite simply, the most activated T cells, those that define dominance, are the most easily grown *in vitro* (Mor *et al.*, 1990).

The roots of the immunologic dominance of an antigen can be seen to be a derivative of the idiotypic dominance of the set of T cells recognizing the particular antigen. Idiotypic dominance in turn is a reflection of the genes used to construct the antigen receptor which encodes the T cell idiotype. This dominance can be attributed, at least in some cases, to a restricted use of T cell receptor genes (Acha-Orbea *et al.*, 1989).

What mechanism determines the dominance of T cell idiotypes? The answer is not yet clear but it seems that natural anti-idiotypic T cell networks may do the job. The evidence for this conclusion is both circumstantial and direct.

(1) T cell vaccination with as few as 10^4, 10^3 or even 10^2 cells of an anti-MBP T clone can activate anti-idiotypic T cells within 5 days and protect rats against EAE (Beraud *et al.*, 1989; Lider *et al.*, 1988). The anti-idiotypic T cells must exist naturally to explain how so few vaccinating T cells could activate so powerful a response so quickly.

(2) Rats immunized with killed *M. tuberculosis* within 4 days spontaneously activate T cells specific for an anti-hsp65 T cell idiotype (N. Karin and I.R. Cohen, in preparation). In other words, the hsp65 target antigen triggers an anti-idiotype to the dominant anti-hsp65 idiotype. Moreover, a low but significant level of anti-anti-hsp65 T cell activity exists in Lewis rats before immunization.

(3) NOD mice developing IDDM spontaneously develop anti-idiotypic T cells reactive to the disease-causing T cell idiotypes (D. Elias and I.R. Cohen, in preparation). The outbreak of overt IDDM is preceded by a fall in the activity of the anti-idiotypic T cells. Mice maintaining a high level of specific anti-idiotypic T cells do not develop IDDM.

Thus natural anti-idiotypic T cell networks can regulate the expression of autoimmune disease. Indeed, it seems that T cell vaccination is most effective when it succeeds in boosting these natural networks. T cell vaccination may have a physiological foundation (Cohen, 1989a,b).

(Parenthetically, it remains to be seen whether or not vaccination against autoimmune disease using particular peptides of the dominant T cell receptor idiotype (Vandenbark *et al.*, 1989; Howell *et al.*, 1989) also invokes natural regulatory networks.)

The natural networks of idiotypic and anti-idiotypic T cells surrounding the dominant autoantigens appear to function in at least two ways. As illustrated both by premeditated T cell vaccination and by the spontaneous anti-idiotypes in NOD mice, such networks can control autoreactivity and prevent or abort autoimmune disease.

The second function of the networks – less intuitively obvious but no less important than the first – is to encode immunological dominance. My appreciation of this output of the networks came about through two unexpected observations. The first observation arose when Ofer Lider and his associates cloned the anti-idiotypic T cells induced by T cell vaccination with a subencephalitogenic dose of a live virulent anti-MBP clone (Lider *et al.*, 1988). Anti-idiotypic T cells of two phenotypes were isolated: $CD4^+8^-$ clones and $CD4^-8^+$ clones. The $CD4^-8^+$ suppressor/cytotoxic clones acted *in vitro* to suppress the response of the virulent T cells to MBP. At about the same time Sun and associates showed that an anti-idiotypic clone of $CD4^-8^+$ phenotype could suppress EAE *in vivo* (Sun *et al.*, 1988). The results of the Lider and Sun experiments satisfied the expectation that the anti-idiotypic network suppressed disease through the agency of anti-idiotypic suppressor/cytotoxic T cells.

However, the $CD4^+8^-$ clones of Lider were a problem: *in vitro* they actually stimulated the virulent anti-MBP T cells as well as did the MBP antigen (Lider *et al.*, 1988). The $CD4^+8^-$ anti-idiotypic clones, in short, behaved *in vitro* as enhancers of the autoimmune response; yet *in vivo* the rats that bore these cells were refractory to the disease EAE. How could we reconcile enhanced autoimmunity with suppressed autoimmune disease?

The second paradoxical observation was made by Nathan Karin in the course of studying the effects of T cell vaccination on adjuvant arthritis and the immune response to hsp65 (N. Karin and I.R. Cohen, in

preparation; Cohen, 1989b). Karin vaccinated a group of rats with a clone of anti-hsp65 T cells in a way that induced resistance to arthritis induced later by active immunization to *M. tuberculosis* (Lider *et al.*, 1987). At regular intervals after immunization with the killed bacteria we assayed the T cell responses of the rats. As expected, the vaccinated rats showed a strong and early (within 4 days) anti-idiotypic reaction to the anti-hsp65 clones. In other words the antigen elicited the anti-idiotypic network that we had amplified earlier by T cell vaccination. Since the vaccinated rats were now refractory to arthritis and since arthritis seemed to result from anti-hsp65 immunity, we expected the protected rats to show depressed T cell responses to hsp65: no diseases ought to mean no autoimmune response. Contary to this simple logic, the vaccinated rats showed an accelerated T cell response to hsp65. The response was evident by day 4 after immunization, a time at which the control rats on their way to developing arthritis, showed only slight activity to the inciting antigen. Thus resistance to an autoimmune disease was heralded by accelerated autoimmunity, not by depressed autoimmunity.

At later time points, however, the T cell autoimmune responses of the vaccinated rats decayed while the responses of the control arthritic rats waxed and predominated. Nevertheless disease or health were not monotonic functions of the autoimmune response; the nature of the autoimmune response, including its kinetics, were critical.

Although the details were not yet clear, these observations could be amalgamated into the idea that one anti-idiotypic network could both prevent clinical disease while enhancing, at least transiently, the autoimmune response to the antigen. The interplay between the $CD4^-8^+$ anti-idiotypic suppressors and the $CD4^+8^-$ anti-idiotypic stimulators could encode the immunological dominance of the autoantigen while controlling the disease inherent in the response to that antigen.

THE IMMUNOLOGICAL HOMUNCULUS

The fare elaborated above boils down to an essential relationship between immunologically dominant self antigens, shared idiotypes and the regulation (or dysregulation) of autoimmunity. This relationship can be described as an immunological homunculus (Cohen, 1989c).

It would seem that the $CD5^+$ B cells and other B cells involved in the production of the natural autoantibodies are organized, like the autoimmune T cells we have studied, into connected anti-idiotypic networks (Coutinho and Bandeira, 1989; Kocks and Rajewsky, 1989). Extrapolating from these empirical observations, one could consider the possibility that all the dominant self antigens are outfitted by such networks. These natural

networks, as we saw above, may encode the dominance of their antigens: the anti-idiotypes, by interacting with the idiotypes, prime the system for its response to the antigen. The network in this way anticipates the antigen; consequentially, the network encodes a representation of the antigen. By its networks the immune system contains within it a picture of some of the molecules of the self. I have termed the collective set of these natural lymphocyte networks the immunological homunculus (Cohen, 1989c; 1991). I chose this term by analogy to the neurological homunculus; the picture the brain has of the individual's body encoded in a set of neural networks. The neurologists appropriated the term homunculus from the early embryologists. Niklass Hartsoeker (1656–1725) claimed that sperm contain a miniscule primordial man (homunculus = little man), from which a person later developed by the accretion of matter (Lyons and Petrucelli, 1978).

The function of the neurological homunculus, unlike the fanciful embryological homunculus, is not to build the body but to regulate the body, to organize information moving to and from the various parts of the body. The features of the body encoded in the neural networks reflect neurological function: for example, the areas devoted to vital interests such as speech in man or smell in dog dominate over functionally less exacting organs. The neurological homunculus is operationally accurate, though topologically distorted – consult any basic neurological text to see the bizarre picture imprinted in the sensory and motor cortices.

The immunological homunculus is like the neurological homunculus in that it contains a distorted picture of the self; only certain self antigens seem to be included. Why this is so is not clear, but drawing on the logic of the neurological homunculus it is likely that the functional roles of certain self molecules determine or influence their immunological interest (Cohen, 1991).

Another feature of both the neurological and immunological homunculi is their dependence on experience. Neural networks do not develop properly unless the connections are consolidated through use (Changeux and Danchin, 1976). Similarly, immunological networks appear to be molded by immunological experience. Self-mimicking antigens bombard the system; we are infected with them, we eat them, we breathe them. Metabolism, turnover and damage constantly release self antigens in various forms. These constant inputs not only create and consolidate the natural immunological networks, they probably necessitate the networks.

Elsewhere Douglas Young and I describe the advantages of the immunological homunculus: by encoding dominance the homunculus serves to attract autoimmune responses to the very same antigens for which network regulation already exists (Cohen and Young, 1991). Dominance

and regulation are wedded partners. For example, MBP is an attractor that channels immune responses away from other neural self antigens for which regulation does not exist. Cardiac myosin dominates the many other self antigens in heart muscle and so protects the mycocardial infarct patient from uncontrolled autoimmunity to the other self antigens. The autoimmune response to cardiac myosin in the great majority of instances is thus benignly regulated and transient (Kuch, 1973). Attracting autoimmunity to a few antigens spares the immune system the need to delete or anergize all the T and B cells that could possibly recognize any of the enumerable epitopes of the tens of thousands of macromolecules composing the self. Such an overkill of self-recognizers would not only be awkward to carry out, it would probably erase most of the immune repertoire.

Moreover, the regulated self-recognizers can also be used to protect the host against microbes. The immune system can safely respond to microbial self-like antigens, for example hsp65, without paying the price of autoimmune disease while enjoying the benefits of protection (Cohen and Young, 1991).

AUTOIMMUNE DISEASE

Benign autoimmunity, which is unexplained by Burnet and the neo-burnetians, is a natural product of the networks constituting the homunculus. Disease does not develop as a reflex response to self-recognition; disease develops through some weakness in the regulation of autoimmunity to the dominant self-antigens. In other words, the common autoimmune diseases may occur through a schism between the twin functions of the homunculus: dominance and regulation (Cohen and Young, 1991). The networks succeed in directing autoimmunity to the few dominant self antigens, but fail to append effective control. Elsewhere, H. Atlan and I present a formal analysis of how a single autoimmune network could produce health or autoimmune disease depending on the strength of its connections (Cohen and Atlan, 1990).

The history of the term homunculus also provides a metaphor for appreciating disease as a fault in the regulation of an otherwise useful activity. Before the sperm homunculus was proposed by embryologists as a means to explain the mysteries of growth and differentiation, the word was used by others.

Paracelsus (1493–1543), the 'father of pharmacology', used the word homunculus to describe a humanoid automaton designed to perform sterotyped tasks for a human master (Scholem, 1965). (The Latin *homunculus* of Paracelsus was begot from the Hebrew *golem* of the Kabbalists, preceded by the *golem* legend of the Talmud.) A popularization of the homunculus as automaton is the story of the sorcerer's apprentice, an homunculus

created by a sorcerer to fetch water. Happily the apprentice draws water in proper measure until a defect in regulation allows the homunculus to run wild and inundate its human master. The immunological homunculus, which can also be described as an automaton (Cohen and Atlan, 1990), may likewise through a loss of regulation exaggerate the performance of its beneficial function to the extent that it damages the individual it was designed to serve. Note that, like the sorcerer's water-bearer, the dysregulated immunological homunculus does its damage in disease with the instrumentality which, properly regulated, it uses for good health: the dominant self antigens selected by the network as targets of benign autoimmunity become the targets of pernicious autoimmunity. According to homunculus theory, the target self antigens in health and disease should be similar. The difference between health and disease will be due not to a change in recognition but to a change in cell proliferation and differentiation: the numbers of autoreactive T and B cells and the functional programs they express; affinity maturation, autoantibody titers and isotypes; the amounts and kinds of lymphokines, cytokines and cytotoxins.

If the root problem of autoimmune disease is self-regulation rather than self-recognition, than the root solution is restoration of regulation by specific means such as T cell vaccination (Cohen, 1986, 1989b) or oral tolerance (Thompson and Staines, 1990). Blocking of recognition by peptides or antibodies, or destruction of T cell sets (Samvil and Steinman, 1990) will be less effective, unless these procedures too turn out to reinstate network regulation.

IMMUNE TRANSFORMATIONS

Recognition is elemental to the immune system in the way that sensory input is elemental to the nervous system: it defines the perceptual universe upon which the system acts. Nevertheless, to attribute the versatility and competence of the immune system to a recognition reflex is akin to attributing the dexterity and virtuosity of a Pele to the realm of the knee-jerk, of a Heifetz to the domain of the grasp-reflex.

The objective of the immune system (that for which it has been selected by evolution) is not recognition, but transformation: the transformation of environmental input into useful output.

The nervous system too deals with transformation; it transforms sensory input into behavior that adapts the organism to changes in the internal and external environments (Ilinas, 1987). The nervous system fashions sensory input into information that tells the organism not only what is presently going on but what according to past experience is likely to happen

in the future. The successful adaptation of behavior to predicted changes in the environment aids survival.

The immune system senses a world of molecular structure that is far different from the world seen by the nervous system; nevertheless both systems use conceptually similar strategies.

The sensory input to the immune system is composed of the epitopes it recognizes. Recognition, the structural complementarity between lymphocyte receptors and antigenic ligands, is the raw material of the system. The output of the immune system is an immune response appropriate to the situation: tolerance for the harmless and attack against the threatening. However, not all that looks like self is harmless and not all that is not self is threatening: compare transformed self cells and self-mimicking microbes (attack) with inhaled or ingested potential allergens (ignore). Hence, the adaptive benefit of attack or non-attack is not congruent with a simple discrimination between non-self and self.

Moreover, even when it pays to attack, each type of target is best rejected by different sets and combinations of effectors. Lepromatous leprosy is not the result of a non-response, the disease is the result of an inappropriate response (Godel, 1980). The appropriate response cures the infection.

In short, the objective of the immune system is to transform recognition into adaptive immune behavior, behavior that aids survival. Immunity functions to turn molecular recognition into information and information into action. How is this transformation done?

The nervous system performs its tranformations by channeling signals through interneurons placed between the sensory neurons and the effector neurons (Llinas, 1987). These networks of interneurons elevate the nervous system from the level of simple reflex to the level of complex behavior. How neural networks accomplish this is the subject of much research and thought in the neurological sciences.

The hypothesis of the immunological homunculus suggests that the immune system generates complex behavior by interposing immunological networks between recognition and effect. The homunculus allows the system to sort, integrate, modify and regulate the outcome of the acts of self-recognition. The homunculus is an internal reference for creating useful information out of the disorderly barrage of self-epitopes present in the body.

Jerne's notion of anti-idiotypic networks provides a general idea of how the immune system could look into itself (Jerne, 1974). The idea of the immunological homunculus is a special case that owes much to Jerne's teachings. However, the homunculus differs in several aspects from Jerne's original formulation. The Jernean Networks make no distinction between self and foreign, are not specialized to deal with autoimmune disease, are

open-ended (anti-idiotypes generate anti-anti-idiotypes which in turn generate anti-anti-anti-idiotypes and so on), and relate primarily to antibodies. The experimental foundation of Jerne's type of open antibody network was the observation that it is possible to generate a fairly long chain of anti-antibodies by isolating the antibodies of one set of animals and immunizing the next set, and so forth. According to Jerne, every immune response will set up a wave motion of anti-idiotypes that continues until it ultimately peters out. Jerne teaches a much more complex vision of the immune system than the simple reflex proposed by Burnet. However, Jerne's open networks have no intrinsic structure of hierarchy and no fixed reference points that might explain immunological dominance, natural autoimmunity, or the regularity of autoimmune disease.

The immunological homunculus can be seen as the application of Jerne's thought to a restricted set of observations related to autoimmunity (Cohen, 1989c). Empirically we observe a hierarchy of dominance, restriction of natural networks to a limited number of self or self-like molecules, T cell networks that are closed and do not extend beyond the level of the anti-idiotypes, and a special class of idiotypically connected B cells making autoantibodies. The important details of the homunculus network are yet unknown: when and how the networks are formed, the identity of the cellular and molecular elements, the nature of the connections and how the connections translate recognition into immune behavior. Now that immunology is beginning to understand the molecular basis of immunity (antigen processing, presentation and recognition, cytokine signals and lymphocyte activation), the questions of network organization can be studied with greater precision. Bear in mind that however the immunological homunculus is put together it fulfills a principle requirement of all information systems: the homunculus supplies the system with an internalized picture of part of its sensory world. If you don't have a general idea of what you are looking for, you will never know it when you see it.

ACKNOWLEDGEMENTS

The work reviewed here was supported by grants from the National Institutes of Health (USA), the Juvenile Diabetes Foundation International, The Minerva Foundation, and Institut Merieux. Mrs Doris Ohayon helped prepare the manuscript.

REFERENCES

Acha-Orbea, H., Steinman, L. and McDevitt, H. (1989). *Ann. Rev. Immunol.* 7, 371–405.

Avrameas, S., Dighiero, G., Lymberi, P. *et al* (1983). *Ann. Immunol.* **134**, 103–13.
Baekkeskov, S., Aanstoot, H-J., Christgau, S. *et al* (1990). *Nature* **347**, 151–156.
Ben-Nun, A., Wekerle, H. and Cohen, I.R. (1981). *Nature* **292**, 60–1.
Beraud, E., Reshef, T., Vandenbark, A. *et al.* (1986). *J. Immunol.* **136**, 511–15.
Beraud, E., Lider, O., Baharav, E., Reshef, T. and Cohen, I.R. (1989). *J. Autoimmun.* **2**, 75–86.
Bjorkman, P.J., Saper, M.A., Samraoui, B. *et al.* (1987). *Nature* **329**, 506–11.
Blackman, M., Kappler, J. and Marrack, P. (1990). *Science* **248**, 1335–41.
Burkly, L.C., Lo, D. and Flavell, R.A. (1990). *Science* **248**, 1364–8.
Burnet, F.M. (1959). *The Clonal Selection Theory of Acquired Immunity*. Cambridge University Press, New York.
Castano, L. and Eisenbarth, G.S. (1990). *Ann. Rev. Immunol.* **8**, 647–79.
Changeux, J.P. and Danchin, A. (1976). *Nature* **264**, 705–12.
Cohen, I.R. (1986). *Immunol. Rev.* **94**, 5–21.
Cohen, I.R. (1988). *Scient. Am.* **256**, 52–60.
Cohen, I.R. (1989a). In *Progress in Immunology VII. Proceedings of the 7th International Congress of Immunology*, ed. F. Melchers, pp. 867–73. Springer-Verlag, Berlin.
Cohen, I.R. (1989b). *Cold Spring Harbor Symposia on Quantative Biology* **LIV**, 879–84.
Cohen, I.R. (1989c). In *Theories of Immune Networks* eds H. Atlan and I.R. Cohen, pp. 6–12. Springer-Verlag, Berlin.
Cohen, I.R. (1991). *Ann. Rev. Immunol.* ~~(in press)~~ 9, 567–89.
Cohen, I.R. and Wekerle, H. (1977). In *Autoimmunity: Genetic Immunologic, Virologic and Clinical Aspects*, ed. N. Talal, pp. 231–65. Academic Press, New York.
Cohen, I.R. and Atlan, H. (1990). *J. Autoimmun.* **2**, 613–25.
Cohen, I.R. and Young, D.B. (1991). *Immunol. Today* (~~in press~~) 12, 105–110.
Cohen, I.R., Altmann, D. and Friedman, A. (1985). *Immunol. Today* **6**, 147–8.
Coutinho, A. and Bandeira, A. (1989). *Immunol. Today* **7**, 363–4.
de Vries, R.R.P. and van Rood, J.J. (1988). In *Perspectives on Autoimmunity*, ed. I.R. Cohen, pp. 2–17. CRC Press, Boca Raton.
Elias, D., Markovits, D., Reshef, T., van der Zee, R. and Cohen, I.R. (1990). *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1576–80.
Elias, D., Reshef, T., Birk, O.S., van der Zee, R. and Cohen, I.R. (1991). (~~Submitted for publication~~) Proc. Natl. Acad. Sci. U.S.A. (in press).
Godel, T. (1980). *Prog Allergy* **25**, 211–42.
Goodnow, C.C., Adelstein, S. and Basten, A. (1990). *Science* **248**, 1373–9.
Grossman, Z., Greenblatt, C.L. and Cohen, I.R. (1986). *J. Theor. Biol.* **121**, 129–39.
Hayakawa, K. and Hardy, R.R. (1988). *Ann. Rev. Immunol.* **6**, 197–218.
Holoshitz, J., Naparstek, Y., Ben-Nun, A. and Cohen, I.R. (1983). *Science* **219**, 56–8.
Howell, M.D., Winters, S.T., Olee, T. *et al.* (1989). *Science* **246**, 668–70.
Jerne, N.K. (1974). *Ann. Immunol.* **142C**, 373–89.
Jindal, S., Dudani, A.K., Singh, B., Harley, C.B. and Gupta, R.S. (1989). *Molec. Cell. Biol.* **9**, 2279–83.
Kaufmann, S.H.E., Vath, U., Thole, J.E.R., van Embden, J.D.A., and Emmrich, F. (1987). *Eur. J. Immunol.* **17**, 351–7.
Kocks, C. and Rajewsky, K. (1989). *Ann. Rev. Immunol.* **7**, 537–59.
Kuch, J. (1973). *Cardiovas. Res.* **7**, 649–54.
Lider, O., Karin, N., Shinitzky, M. and Cohen, I.R. (1987). *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4577–80.

Lider, O., Reshef, T., Beraud, E., Ben-Nun, A. and Cohen, I.R. (1988). *Science* **239**, 181–3.

Lider, O., Shinitzky, M. and Cohen, I.R. (1986). *Ann. N.Y. Acad. Sci.* **457**, 267–73.

Llinas, R. (1987). In *Mindwaves: Thoughts on Intelligence, Identity and Consciousness*, eds C. Blakemore and S.A. Greenfield, pp. 339–57. Basil Blackwell, Oxford.

Lohse A., Mor, F., Karin, N. and Cohen, I.R. (1989). *Science* **244**, 820–2.

Lyons, A.S. and Petrucelli, R.J. (1978). *Medicine: An Illustrated History*, p. 437. Harry N. Abrams, New York.

Maron, R., Zerubavel, R., Friedman, A. and Cohen, I.R. (1983). *J. Immunol.* **131**, 2316–22.

Mor, F., Lohse, A., Karin, N. and Cohen, I.R. (1990). *J. Clin. Invest.* **85**, 1594–8.

Munk, M.E., Schoel, B., Modrow, S. *et al.* (1989). *J. Immunol.* **143**, 2844–9.

Naparstek, Y., Ben-Nun, A., Holoshitz, J. *et al.* (1983). *Eur. J. Immunol.* **13**, 418–23.

Naquet, P., Ellis, J., Tibensky, D. *et al.* (1988). *J. Immunol.* **140**, 2569–78.

Palmer, J.P., Asplin, C.M., Clemons, P. *et al.* (1983). *Science* **222**, 1337–9.

Pearson, C.M. (1964). *Arthr. Rheum.* **7**, 80–6.

Res, P.C.M., Schaar, C.G., Breedveld, F.C. *et al.* (1988). *The Lancet* **II**, 478–80.

Scholem, G. (1965). *On the Kabbalah and its Symbolism*, pp. 197–9. Schocken Books, New York.

Shoenfeld, Y. and Isenberg, D. (1989). *The Mosiac of Autoimmunity (The Factors Associated with Autoimmune Disease)*. Elsevier, Amsterdam.

Sun, D., Qin, Y., Chluba, J., Epplen, J.T. and Wekerle, H. (1988). *Nature* **332**, 843–5.

Thompson, H.S.G. and Staines, N.A. (1990). *Immunol. Today* **11**, 396–9.

van Eden, W. Thole, J., van der Zee, R *et al.* (1988). *Nature* **331**, 171–3.

van Eden, W., Holoshitz, J., Nevo, Z *et al.* (1985). *Proc. Natl. Acad. Sci. U.S.A.* **82**, 5117–20.

Vandenbark, A.A., Hashim, G. and Offner, H. (1989). *Nature* **341**, 541–94.

von Boehmer, H. (1990). *Ann. Rev. Immunol.* **8**, 531–56.

Wicherpfennig, K.W., Ota, K., Endo, N. *et al.* (1990). *Science* **248**, 1016–19.

Young, D.B. (1990). *Semin. Cell Biol.* **1**, 27–35.

Young, D.B., Ivanyi, J., Cox, J.H. and Lamb, J.R. (1987). *Immunol. Today* **8**, 215–19.

Zamvil, S.S. and Steinman, L. (1990). *Ann. Rev. Immunol.* **8**, 579–621.