Adorini L (ed): Selective Immunosuppression: Basic Concepts and Clinica Applications.
Chem Immunol. Basel, Karger, 1995, vol 60, pp 150–160

Treatment of Autoimmune Disease: To Activate or to Deactivate?

Irun R. Cohen[1]

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

An autoimmune disease is the mark of an immune system gone wrong. The autoimmune disease challenges clinical immunology to set the system right. In recent years, immunology has deciphered much of the basic chemistry of the cells, antibodies, antigen receptors, cytokines and cytokine receptors that comprise the immune system. This new molecular information is important because it should speed the design of therapies rationally targeted to critical elements in the pathologic autoimmune process. But to choose the most effective therapy, to foresee just what the novel therapies are likely to accomplish, we need to have some idea of the pathophysiology of autoimmunity: What is it that has gone wrong, what are we trying to fix? These questions are not mere philosophical niceties; they can guide concrete clinical decisions. And the strategies beyond the decisions may be poles apart. For example, some of the new therapies have been devised to inactivate immune responses by blocking MHC-T-cell receptor interactions or by killing specific T cells with toxin conjugates [1]. In contrast, other proposed therapies such as oral administration of antigen [2] or vaccinations with T cells or with peptides [3], have been devised to *activate* immune responses. The seemingly opposite strategies of *inactivation* and *activation* are each rational to their adherents because they rest on two different interpretations of autoimmunity: the clonal selection paradigm and the cognitive paradigm. The aim of this review is to articulate these paradigms and their implications for the therapy of autoimmune diseases.

[1] Incumbent of the Mauerberger Chair of Immunology and the Director of the Robert Koch-Minerva Center for Research in Autoimmune Diseases.

Clonal Selection

The clonal selection paradigm asserts that the echelons of lymphocytes that form the immune system are selected entirely by encounter with antigen [4]. This concept has wide currency and need not be described here. I will cite only two tenets of clonal selection relating to autoimmunity and autoimmune disease:

(1) The healthy immune system must not include lymphocytes that recognize self antigens; autoimmune lymphocytes, at least the T cells, are removed from the repertoire by processes of negative selection [5] and anergy [6] that take place in the thymus and in the periphery.

(2) Autoimmune diseases, therefore, can have no intrinsic regularity in their clinical or immunological expression; autoimmune diseases are caused by autoimmune lymphocytes that arise from random mutations or unstructured escape from deletion. At the risk of over-simplification, we can say that the clonal selection paradigm asserts that autoimmune disease develops because an entity that should be absent, an autoimmune lymphocyte clone, is present by accident. The failure of the immune system to have deleted the forbidden clone should then be rectified by blocking with various antibodies, for example, the clone's ability to recognize its antigen, or by medicinally killing the clone, for example by toxin conjugates targeted to activated T cells [1].

In other words, an autoimmune disease is caused, according to the clonal selection paradigm, by aberrant *activation* of the immune response [7]. The rational answer to harmful activation is to find a way to *deactivate* the pathogenic lymphocytes. Thus, the clonal selection paradigm argues for suppression of the immune system. Unlike the cytotoxic drugs and steroids now in use with their broad and uncontrollable effects, the new molecular immunology makes it feasible to design highly specific silver bullets that can target the forbidden clones of autoimmune lymphocytes and so de-activate the autoimmune agents of disease without inactivating the entire immune system.

The Cognitive Paradigm

Elsewhere, I have presented an alternative to the clonal selection paradigm called the cognitive paradigm [8, 9]. The cognitive paradigm takes account of two empirical observations that contradict the tenets of orthodox clonal selection relating to autoimmunity: (1) Healthy immune systems do in fact contain T cells and B cells that recognize certain dominant self antigens; some autoimmunity is natural. (2) Autoimmunity is not chaotic, but manifests quite predictable immunologic characteristics.

Healthy humans and healthy rodents bear natural autoantibodies and autoimmune T cells to similar sets of self antigens. Indeed, many of the same self antigens attacked in diseases are also recognized in health. Moreover, a small number of diseases encompasses the majority of autoimmunity patients: rheumatoid arthritis, type I diabetes mellitus, multiple sclerosis, systemic lupus erythematosus. Suffering a disease may be accidental; but the disease visited upon the patient is highly structured. This is true even across species: systemic lupus erythematosus and type I diabetes are quite similar in their immunology in both mice and humans. Order and predictability imply that natural autoimmunity and autoimmune diseases are governed by rules. But what function is served by having autoimmunity built into the system?

The Internal Physician and the Problem of Ambiguity

One way to think about the functions of the immune system is to consider it, metaphorically, as the body's internal physician, a physician specializing in infectious diseases (and perhaps in some types of cancer). This internal physician, like any external physician, performs two functions: diagnosis and treatment. The immune system has to diagnose the nature of a threat to the integrity of the body (virus, bacterium, higher parasite) and it has to prescribe the mix of therapeutic responses (types of T cells, B cells, antibodies, cytokines, inflammation) best suited to destroy the invader.

The real problem of the immune system, the one that makes simple clonal selection unworkable, is antigenic ambiguity: many of the key protein antigens of invading microbes and other parasites contain amino acid sequences that are identical to those of the host body. Molecular biology has demonstrated that evolution conserves useful genes. Nature does not reinvent the wheel each time she designs a new creature. Once evolution comes upon a molecule that performs an essential function, it conserves the molecule's genes for further use in other instances. Indeed the more important the molecule, the more its sequence is conserved across the epochs. Heat shock protein (hsp) molecules, essential for all cells, are a telling example: humans and bacteria manifest 50–80% identity of amino acid sequences in the various members of the hsp family, nevertheless, hsp molecules are recognized as dominant antigens of many infectious agents [10]. These conserved molecules are also among the self antigens recognized in healthy, natural autoimmunity [8, 9]. How does the immune system know how to respond appropriately to hsp and other molecules that antigenically are both self and not-self and that may accompany either health or infection? Recognition alone is not sufficient. The immune system must interpret the meaning of an antigen; to interpret, it has to resort to cognition.

Diagnostic Interpretation

The cognitive paradigm proposes that the immune system, the internal physician, interprets the meaning of antigens using a strategy similar to that used by an external physician to interpret the meaning of clinical signs. Interpretation is brought about using the cognitive process called *differential diagnosis*. Differential diagnosis is a way of making clinical decisions based on searching for a string of relevant information (history, physical examination, laboratory tests) and on organizing this information in the light of the physician's knowledge. The external physician will interpret the meaning of a fever, for example, according to what he knows about fever and what he learns about the patients. Diagnosis, in essence, is paradigmatic; it is the process of fitting the actual case to one of the archetypal models that the physician has in mind [11]. Implicit in the process are *memory*, the storage of information, and *learning*, the adjustment, through experience, of the archetypal models and the program of assembling information. The cognitive paradigm teaches that the immune system too interprets antigens by building internal representations (archetypal categories) and assembling strings of relevant signals (fig. 1).

Images of Self

Physicians know that internal images exist; they use mental images when they practice differential diagnosis. Cognitive scientists talk about internal representations of the external world that function in mental cognition [12], but nobody knows how these representations are actually encoded in the mind. Immunologists appear to be better informed; they know the molecular basis of at least some of the images in the immune system. The idea of immune internal images was made explicit by Jerne [13] in his anti-idiotypic theory of immune regulation. The antigen receptors of T and B cells are structurally complementary to the epitopes (sites) they recognize on their antigens. Thus, antigen receptors structurally are negative images of parts of antigens, just as locks and keys, in part, are negative images of each other. Anti-idiotypic antibodies that bind to the antigen-binding sites of other antibodies are structurally complementary to the negative images of antigens; thus, anti-idiotypic antibodies can be positive images of antigens (a negative of a negative can create a positive). The images created by T-cell idiotypes and anti-idiotypes are complicated by the fact that T-cell receptors see processed antigen fragments in the clefts of major histocompatibility complex (MHC) molecules; nevertheless, T-cell receptors are images all the same.

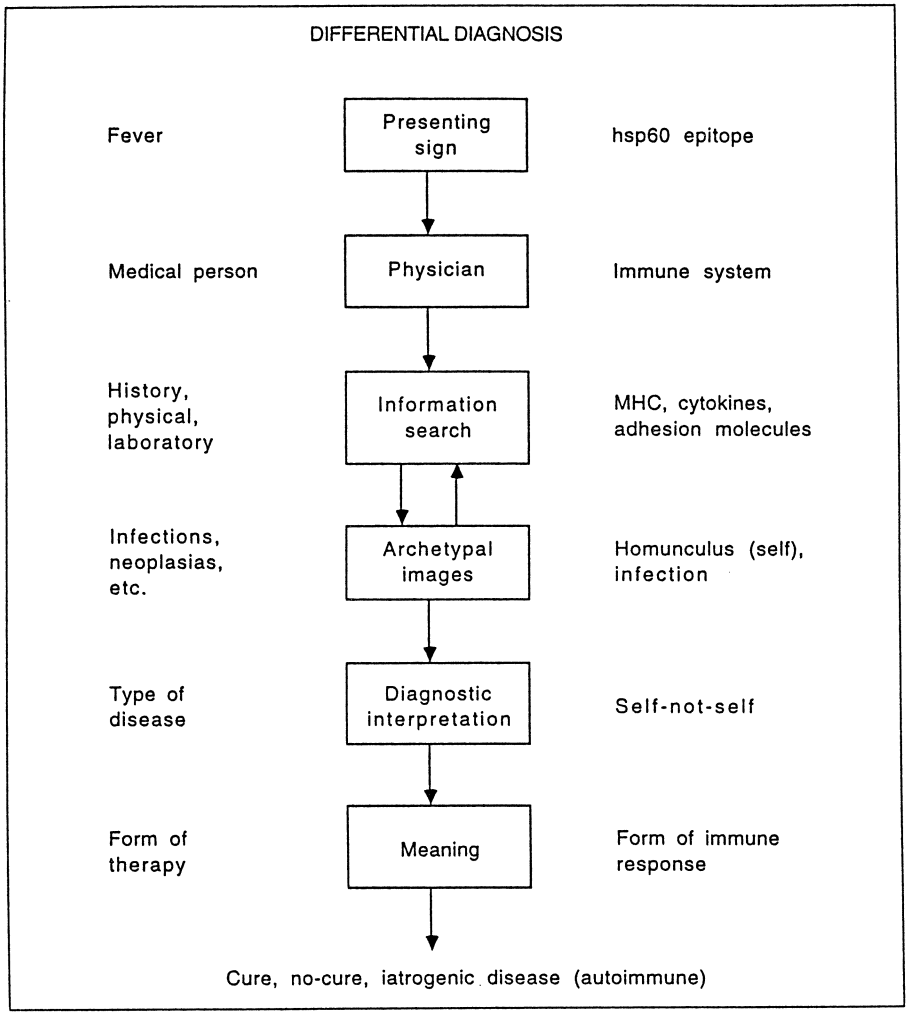

Fig. 1. The process of differential diagnosis – medical and immunological. The meaning of a presenting sign, such as a 'fever' for a medical person or an 'hsp60 epitope' for the immune system, is decided in both cases by a search for the appropriate additional information needed to define the context of the sign. The person and the immune system are each guided by internal archetypal images. The medical person's internal images of diseases are the accumulated useful experience of the profession taught in medical school and modified by the person's actual experience. The immune system's internal images of the self (the homunculus) and of infection in general are formed by the accumulated, evolutionary experience of the species carried in the germ line combined with the individual's somatic experience recorded in the antigen-receptor repertoire. The diagnostic interpretation of the person and of the system is the outcome of the interaction between the information search and the preformed archetypal images. The medical person decides whether

Because we humans are so attuned to visual information, we tend to think of images as organized pictures. Nevertheless, receptors on wandering cells or cell networks can also create images, albeit images that are distributed about the body.

Now, if the healthy immune system naturally contains antibodies, B cells and T cells capable of recognizing a standard set of self antigens, and if the system also contains anti-idiotypes to these elements, then we can state that the immune system contains both positive and negative internal images of particular self molecules. These images constitute what I call the *immunological homunculus* [8–10]. I borrowed the term homunculus from neurology where it refers to representations of the body encoded in the brain. Since self molecules are the first antigens the immune system meets, then it follows that the formation of the immunological homunculus precedes actual contact with foreign antigens. The T-cell imunological homunculus probably develops through the positive selection of some autoimmune T cells in the thymus, but much remains to be discovered about its origins; are particular self antigens programmed to be expressed in the thymus [14] to positively select autoimmune T cells? The homunculus formed by the natural autoantibodies is now being studied using a new technology [15].

It appears that many of the self molecules in the homunculus set, such as the hsp molecules, have been highly conserved in evolution and are shared by hosts and parasites [10]. It can be argued that autoimmunity to highly conserved molecules expressed both by host and by parasites might serve to prime the host immune response against the parasites. Note that tight regulation by anti-idiotypic and other regulatory cells is required to avoid the development of autoimmune diseases when the system transiently exploits natural autoimmunity to hsp and other conserved molecules to fight infection [10]. Thus, the immunological homunculus, regulated autoimmunity, is one of the ways evolved by the immune system to deal with the problem posed by conserved

the case fits any of the recognized disease categories in his stock of internal mental images. The immune system automatically interprets whether the sign relates to self or not-self by integrating the string of acquired information with the archetypal images. (Note that the distinction between the self and the not-self is a matter of interpretation. The self is defined by a set of interactions; it is not an immutable given.) The functional meaning of the presenting sign is the form of treatment prescribed by the medical person or the type of immune response deployed by the immune system. The diagnosis is correct if the therapeutic response cures the patient. A wrong diagnosis may result in no cure or, unfortunately, in an iatrogenic disease. In the case of the immune response, a wrong diagnosis may mean an autoimmune disease. Wrong diagnoses of both sorts can be corrected, according to the cognitive paradigm, by supplying the right information.

molecules shared by host and infectious agents, the problem of ambiguous self-not-self [16]. Of course, we must still explain the need for natural autoimmunity to self antigens that are not shared with parasites.

Images of Infections

In addition to the immunological homunculus, the immune system contains primordial internal images of infectious agents encoded by the many cell receptors specific for microbial products: components of bacterial cell walls, lipopolysaccharides, viral nucleic acids [8, 9]. Receptors for the cytokines triggered by infection can also be viewed as internal images of infection, images which reflect not the structure of infectious agents but their presence. These types of receptors for infectious agents are invariant; they are determined by germ-line genes and exemplify a primitive adaptation of the species to pervading features of infectious agents. Antibodies, B cells and T cells specific for microbial antigens are also internal images of the microbes, but these specific images are learned through the actual infectious experience of the individual. Thus, immune adaptation is expressed at two levels: a first, crude image of infection is encoded in the germ line; a second, refined image is encoded by the antigen-specific T and B cells activated in each individual through experience.

Immune Interpretation: Strings of Information

Table 1 lists four categories of information which the immune system uses to make diagnostic interpretations and therapeutic decisions. The first category (table 1) includes the receptor molecules that recognize antigens, what is usually called the repertoire of the immune system. These receptors are generated in each individual by random somatic processes. The original clonal selection paradigm focused its attention on this category of information, almost exclusively. The clonal selection paradigm was formulated early, at the time when immunologists still were unaware of the considerable resources of immune information encoded within the germ line: networks of cytokines, cell interactions, and antigen processing and presentation (table 1). Sections III and IV in table 1 include entities also unknown or disregarded by the formulators of clonal selection: the internal image of self (the immunological homunculus) and the primordial image of infection. The cognitive paradigm is a global view of the immune system up-dated to encompass the additional information that immunologists have discovered since the clonal selection paradigm.

Table 1. Immune information

I	Clonal recognition of antigen by way of somatic genetic recombination The repertoire of antibodies and antigen receptors on T and B cells
II	Germ-line recognition of context: accessory signals
	1 Cytokines [26]: Interleukins 1a, 1b, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12; TNFα, TNFβ, TGF-β1, TGF-β2, TGF-β3; interferons α, β, γ; etc.
	2 Ligand pairs in inter-leukocyte interactions [27]: MHC class I, class II; CD3, CD4, CD8, LAG-3, LFA-1, ICAM-1, CD2, LFA-3, CD28, B7, CTLA-4, CD40, CD40 ligand, CD45Ro, CD22, CD5, CD72; etc.
	3 Ligand pairs in leukocyte-tissue interactions [28]: Selectins (E, P, L) – carbohydrate ligands; IgCAM integrins; addressins; CD31; LFA-3-CD2; etc.
	4 Molecules involved in the processing and transport of antigens and antigen fragments
III	Primordial internal image of infection: receptors for microbial products [9]
IV	The immunological homunculus: natural autoimmunity [9]

Many of the elements listed in table 1 are membrane receptors connected directly or indirectly to protein kinases and other mechanisms of signal transduction. It is obvious from the large number of these receptors that lymphocytes are capable or receiving ancillary information from many sources in additon to recognizing antigens. It is the integration of this information which tells the lymphocytes what to do about the antigens it sees. Integration is interpretation; the nature of the response follows. For example, T cells of the Th1 type are involved in causing cell-mediated tissue damage; T cells of the Th2 type are involved in inducing antibody production by B cells. Whether a naive T cell becomes a Th1 or Th2 cell in response to a given antigen depends on many factors including the cytokine environment (IL-2 and interferon-γ vs. IL-4 and IL-10), the presence or absence of CD8 cytotoxic T cells, the nature of the antigen-presenting cell (macrophage vs. B cell), the concentration of the antigen and its mode of entry into the antigen-presenting cell (phagocytosis vs. pinocytosis) [17]. Important biological consequences depend on whether the immune system decides to generate Th1 or Th2 T cells to a particular antigen, but these are only two of many different possible effector agents that can result from antigen recognition.

Here lies the basic difference between the clonal selection and the cognitive paradigms: the former entrusts the antigens to organize the immune

system; the latter proposes that the immune system has been programmed by evolution to interpret antigens, albeit automatically and unconsciously, like a very complex computer.

Autoimmune Disease: A Misinterpretation

The cognitive view of autoimmune diseases differs from that of the clonal selection paradigm. The autoimmune clones specific for the dominant self antigens are naturally present in the homunculus; nonetheless, these autoimmune clones do not normally cause disease for at least two reasons: the clones are tightly regulated by specific anti-idiotypic and other mechanisms, and the autoimmune clones see their target self antigens in contexts that do not incite an effector response (for example, in the absence of inflammatory cytokines, or in the presence of suppressive signals). The transition from natural, benign autoimmunity to autoimmune disease follows two circumstances: the self antigen is presented inadvertently in a context that drives a damaging effector response and the anti-idiotypic and other regulatory connections are too weak to counteract the stimulus and restore a healthy equilibrium [8–10].

The transition from benign, natural autoimmunity to an autoimmune disease has been studied in rodents in the disease called experimental autoimmune encephalomyelitis (EAE), which models aspects of multiple sclerosis (MS). EAE is caused by activated T cells with receptors for the self-antigen myelin basic protein (MBP) [18]. Anti-MBP T cells are also detectable in MS patients [19]. However, it has been discovered that potentially virulent anti-MBP T cells are present in healthy rats that will never develop EAE spontaneously [20], and also in healthy people who will probably never develop MS [21]. But clinical EAE can be induced readily by immunizing rats with MBP emulsified in oil containing dead mycobacteria, a material called complete Freund's adjuvant [18]. Apparently, the adjuvant supplies MBP with the ancillary signals indicative of an infection, and the naturally quiescent anti-MBP T cells are driven to differentiate into activated effector T cells, causing clinical EAE. Thus, the transition from benign to noxious autoimmunity in EAE and in other experimental diseases is induced by signals that dress MBP or the other self antigens in the context of infection [9].

An aberrant expression of MHC class II molecules has been proposed as a mechanism to explain the induction of clinical autoimmunity [7]. The cognitive paradigm would consider expression of MHC molecules as only one of many elements capable of defining the context of an antigen (table 1). Be that as it may, the transition to disease is the result of inappropriate contextual information.

Note, however, that the perpetuation of autoimmune disease differs from the induction of autoimmune disease. Despite the inciting context, the healthy immune system of the rat usually learns to resist the inciting context of MBP in adjuvant, probably by activating and amplifying the anti-idiotypic regulatory T cells inherent in the immunological homunculus [8–10]. EAE cannot be induced a second time, unless the regulatory T cells have been inhibited. Relapsing or progressive MS would also seem to require some insufficiency of the regulatory cells.

According to the cognitive paradigm, medicine might mimic nature by activating and strengthening the regulatory cells connected to the specific autoimmune lymphocytes through treatments such as T-cell vaccination, found to be effective in experimental animals [22] and recently in MS [23]. Another way to stimulate regulatory cells is by oral tolerance [2]. These simple forms of treatment would seem to work because, to avoid chaos, complex biological systems must organize themselves to focus on specific bits of information. Such bits of information function as regulatory elements because they influence the state of the system. The scope of this article does not allow a discussion of the self-organization of systems [24], but the power of small bits of information to control complex behavior is an observable fact of life: witness the response of the central nervous system to a certain smile, to a kind word. Noxious behaviour by the immune system too is susceptible to modification by communicating with the system using signals the system understands. Even a spontaneous autoimmune disease such as the diabetes of the NOD mouse can be treated by vaccination with a suitable T-cell clone or with an appropriate peptide [3, 25]. Thus, safer and more effective therapies for autoimmune diseases will emerge as immunology uncovers and exploits the critical bits of information that the immune system has been adapted to interpret cognitively. Like the nervous system, the immune system can learn to behave itself. Properly informed, the internal physician will make it right.

References

1 Bach JF: Immunosuppressive therapy of autoimmune diseases. Immunol Today 1993;14: 322–325.

2 Weiner HL, Mackin GA, Matsui M, Orav EJ, Khoury SJ, Dawson DM, Hafler DA: A double-blind pilot trial of oral tolerization with myelin antigens in multiple sclerosis. Science 1993;259: 1321–1324.

3 Elias D, Reshef T, Birk OS, van der Zee R, Walker MD, Cohen IR: Vaccination against autoimmune mouse diabetes with a T-cell epitope of the human 65-kDa heat shock protein. Proc Natl Acad Sci USA 1991;88:3088–3091.

4 Burnet FM: The Clonal Selection Theory of Acquired Immunity. Cambridge, Cambridge University Press, 1959.

5 Marrack P, Kappler JW: How the immune system recognizes the body. Sci Am 1993;269:49–55.

6 Schwartz RH: A cell culture model for T lymphocyte clonal anergy. Science 1990;248: 1349–1356.
7 Bottazzo JF, Pujol-Borrell R, Hanafusa T, Feldman M: Hypothesis: Role of aberrant HLA-DR expression and antigen presentation in the induction of endocrine autoimmunity. Lancet 1983;ii:1115–1119.
8 Cohen IR: The cognitive principle challenges clonal selection. Immunol Today 1992;13: 441–444.
9 Cohen IR: The cognitive paradigm and the immunological homunculus. Immunol Today 1992; 13:490–494.
10 Cohen IR, Young DB: Autoimmunity, microbial immunity, and the immunological homunculus. Immunol Today 1991;12:105–110.
11 Kassirer JP: Diagnostic reasoning. Ann Intern Med 1989;110:893–900.
12 Shanon B: Representations: senses and reasons. Philos Psychol 1991;4:355–374.
13 Jerne NK: Toward a network theory of the immune system. Ann Immunol (Paris) 1974:125C: 373–89.
14 Kourilsky P, Claverie JM, Prochnicka-Chalufour A, Spetz-Hagberg AL, Larsson-Sciard EL: How important is the direct recognition of the polymorphic MHC residues by TCR in the generation of the T-cell repertoire? Cold Spring Harbor Symp Quant Biol 1989;54:93–103.
15 Nobrega A, Haury M, Grandien A, Malanchère E, Sundblad A, Coutinho A: Global analysis of antibody repertoires. II. Evidence for specificity, self-selection and the immunological 'homunculus' of antibodies in normal serum. Eur J Immunol 1993;23:2851–2859.
16 Cohen IR: Kadishman's tree, Escher's angels and the immunological homunculus; in Coutinho A, Kazatchkine M (eds): Autoimmunity: Physiology and Disease. New York, Wiley-Liss, 1994, pp 7–18.
17 Fitch FWE, Mckisic MD, Lanck DW, Gajewski TF: Differential regulation of murine T lymphocyte subsets. Ann Rev Immunol 1993;11:29–48.
18 Ben-Nun A, Wekerle H, Cohen IR: The rapid isolation of clonable antigen-specific T lymphocyte lines capable of mediating autoimmune encephalomyelitis. Eur J Immunol 1981;11:195–199.
19 Ota K, Matsui M, Milford EL, Mackin GA, Weiner HL, Hafler DA: T-cell recognition of an immunodominant myelin basic protein epitope in multiple sclerosis. Nature 1990;346:183–187.
20 Mor F, Cohen IR: Shifts in the epitopes of myelin basic protein recognized by lewis rat T cells before, during and after the induction of experimental autoimmune encephalomyelitis. J Clin Invest 1993;92:2199–2206.
21 Burns J, Rosenzweig A, Zweiman B, Lisak RP: Isolation of myelin basic protein-reactive T-cell lines from normal human blood. Cell Immunol 1983;81:435–440.
22 Cohen IR: T-cell vaccination in immunological disease. J Intern Med 1991;230:471–477.
23 Zhang J, Medaer R, Stinissen P, Hafler D, Raus J: MHC-restricted depletion of human basic protein-reactive T cells by T cell vaccination. Science 1993;261:1451–1454.
24 Atlan H: Self-creation of meaning. Phys Scripta 1987;36:536–576.
25 Elias D, Cohen IR: Peptide therapy for diabetes in NOD mice. Lancet 1994;343:704–706.
26 Miyajima A, Kitamura T; Harada N, Yokata T; Arai K-I: Cytokine receptors and signal transduction. Annu Rev Immunol 1992;10:295–331.
27 Parker DC: T cell dependent B cell activation. Annu Rev Immunol 1993;11:331–360.
28 Bevilacqua MP: Endothelial-leukocyte adhesion molecules. Annu Rev Immunol 1993;11:767–804.

Prof. Irun R. Cohen, Department of Cell Biology, The Weizmann Institute of Science, Rehovot 76100 (Israel)