Biology and Philosophy (2006) 21:235–252
DOI 10.1007/s10539-005-8621-6

The self model and the conception of biological identity in immunology

THOMAS PRADEU[1,*] and EDGARDO D. CAROSELLA[2]
[1]*Institut d'Histoire et de Philosophie des Sciences et des Techniques (IHPST) Department of Philosophy, Paris1-Sorbonne University, Paris, France;* [2]*Service de Recherches en Hémato-Immunologie, Commissariat à l'Énergie Atomique CEA-DRM-DSV, Hôpital Saint-Louis, Paris, France;* **Author for correspondence (e-mail: thomas.pradeu@ens.fr)*

Received 29 November 2004; accepted in revised form 10 June 2005

Key words: Continuity, Identity, Immunology, Organism and environment, Self

Abstract. The self/non-self model, first proposed by F.M. Burnet, has dominated immunology for 60 years now. According to this model, any *foreign* element will trigger an immune reaction in an organism, whereas *endogenous* elements will not, in normal circumstances, induce an immune reaction. In this paper we show that the self/non-self model is no longer an appropriate explanation of experimental data in immunology, and that this inadequacy may be rooted in an excessively strong metaphysical conception of biological identity. We suggest that another hypothesis, one based on the notion of continuity, gives a better account of immune phenomena. Finally, we underscore the mapping between this metaphysical deflation from *self* to *continuity* in immunology and the philosophical debate between substantialism and empiricism about identity.

Introduction

The immunological self/non-self model, proposed by F.M. Burnet in the 1940s, is almost universally accepted nowadays among immunologists (Langman and Cohn 2000). This model was built to offer a *criterion of immunogenicity*, that is, to answer the fundamental question 'when (in what circumstances) does an immune reaction occur in a given organism?'. The self/non-self model provides the following answer: any element which is *foreign* (non-self) to an organism will trigger an immune reaction if introduced to it, whereas endogenous elements (self) do not, in normal circumstances, induce an immune reaction. Because immunology has been one of the key domains of the molecularization of biology since the 1960s, the discipline is currently torn between a very strong molecular programme, and a theoretical framework borrowed mainly from psychology[1] and philosophy (Burnet 1965). It seems worthwhile to ask on which principles the self/non-self model was built, what its origins are, and whether it is a satisfactory model.

[1] Burnet said explicitly that he had been inspired by Wells et al.'s (1929) *The Science of Life*, which uses the term 'self' to refer to psychological identity: see Tauber 1994, pages 94–98.

Here we show, using experimental evidence, that the self/non-self model is no longer appropriate to account for the available immunological data as a whole, and that this inadequacy may be rooted in an excessively strong metaphysical conception of biological identity. We suggest that another theoretical model, one based on the notion of *continuity*, gives a better account of immune phenomena. As we will see, this model may be conceived of as the immunological part of biological genidentity[2], which has been recently renewed by Boniolo and Carrara (2004).

We analyze first the standard view, that is, the self/non-self model (Section 2); then we present experimental arguments showing the inadequacy of this model (Section 3); after which we argue that the *continuity* hypothesis gives a better account of immunological data (Section 4); and finally we discuss the metaphysical roots of the two conceptions of biological identity underlying these two models, that is, *substance* versus *continuity* (Sections 5 and 6).

The standard view: the immune 'self'

The theory of the immune self has dominated immunology for 60 years. Leaning on Ehrlich's principle of *horror autotoxicus* (Ehrlich 1900), F.M. Burnet suggested the immune self/non-self hypothesis for the first time in 1949 (Burnet and Fenner 1949). The origins and developments of this model were rich and complex (Löwy 1991; Tauber 1994; Silverstein and Rose 1997), but the general theoretical framework remained the same. According to the self model, every *foreign* element ('non-self') triggers an immune reaction of defence from the organism, whereas no component of the organism ('self') triggers an immune reaction (Burnet and Fenner 1949, 100–102; Burnet 1960). In other words, everything that comes from the *inside* (endogenous) is tolerated and preserved, whereas everything that comes from the *outside* (exogenous) is attacked. Hence, the *integrity* of the organism is maintained (Burnet 1962). A critical example of this self/non-self discrimination is the fact that an organism accepts a graft of its own tissues (autograft), whereas it rejects a graft from a foreign organism (allograft). The 'self' can thus be defined as a *closed fortress* (Wilson 1972, 8–10). The principles of the self model have been based on observations dealing with pathogens (Burnet and Fenner 1949), grafts rejection/acceptance (Murphy 1913, Billingham et al. 1953), thymectomy (Miller 1961), and MHC (Major Histocompatibility Complex) presentation (Zinkernagel et al. 1978). Particularly, the demonstration that, in a given organism, lymphocytes bearing receptors specific for an antigen are deleted if the genes that code for this antigen are present in the genome of this organism (*negative selection*) (Kappler et al. 1987) was seen as a strong confirmation of the self model.

[2] According to the genidentity criterion, 'two' biological entities are the same entity if they occupy the same space-time region.

A critical question here is: how is the immune 'self' defined in the self/non-self theory? The boundaries of the immune 'self' are not the same as those of the organism: since a graft between two identical twins is tolerated, these two individuals have the same immune 'self'. Hence, the immune self is not equivalent to what is 'under the skin' of an organism. Nonetheless, the immune self is not equivalent to the genomic self either, because the immune self is *acquired* and not innate: it is defined in the process of *selection* of immunocompetent cells, that is, as suggested above, the elimination of all the immune cells that react to the components of the organism that are presented to them (Burnet 1957, 1959; for an example in contemporary research, see Walker and Abbas 2001). This selection occurs in the foetal or immediately post-natal period, depending on species. If, during this period, some entities genetically different from the organism are introduced in it, it will not launch any immune reaction against them later: "cells 'foreign' to the host may be tolerated indefinitely provided that they are implanted early in embryonic life" (Burnet and Fenner 1949, 76). Hence, the immune self is *based* on the genomic self (because most of the time, without any external intervention, the molecular patterns on the basis of which immune cells are selected come from the host) but is not *equivalent* to it. Thus, the basic rule of the self model is the following: molecular patterns originating in the genome of the organism, those presented to it during the fœtal/immediately post-natal period, and those identical to them, define the immune 'self'. Accordingly, they do not trigger an immune reaction. All other patterns constitute the immune 'non-self', and as a consequence activate a reaction.

The 'self' framework still constitutes today the standard view in all the aspects of immunology: innate immunity (Medzhitov and Janeway 2002), selection of immunocompetent lymphocytes (Walker and Abbas 2001), MHC presentation (Garcia et al. 1998), autoimmunity (Pedotti et al. 2001), tumour immunity (Nanda and Sercarz 1995), HIV infection (Douek et al. 2003), etc.

Problems with the standard view: the inadequacy of the self model

The self/non-self has been called into question both conceptually (Jerne 1974; Tauber 1994; Tauber 1999; Cohen 2000), and experimentally (Matzinger et al. 1999; Cohen 1992). Alfred Tauber made a critical contribution on this issue: he showed the relation between the definition of *immunity* and the various conceptions of *identity*, and hence underlined specifically the metaphysical roots of the self/non-self model, and also its vagueness. The immunological 'self', remarked Tauber in his incisive book (Tauber 1994), cannot be considered as a scientific concept, it is no more than a *metaphor*.

Though our method is quite different, because we start from experimental data and only then examine the metaphysical aspects, we believe that it is possible to go further in this critique, and that the data accumulated in very recent immunology clearly prove the inadequacy of the self/non-self model

(Pradeu and Carosella 2004). Here we demonstrate that, contrary to the assertion of the self/non-self model, 'self' components do induce immune reactions, and many 'non-self' components do not, and consequently that the distinction between *internal* and *external* origins of entities is not adequate for a proper understanding of immune phenomena.

Falsity of the principle 'no entity originating from the organism will trigger an immune reaction'

From its birth, the self model, built on the foundations of Ehrlich's theory of *horror autotoxicus* (Ehrlich 1900), defined autoimmunity as a pathological exception to the principle of absence of immune attack against the 'self': autoimmunity was seen as a disorder with regard to normal immune functioning. By contrast, Jerne suggested that immune cells could react constantly to components of the organism, defining autoimmunity as an ongoing *surveillance* process and not as an abnormal destruction (Jerne 1974). Nowadays, a revised version of this thesis is supported by strong evidence. During the selection of lymphocytes in primary lymphoid organs (that is, thymus for T lymphocytes, and bone marrow for B lymphocytes), cells which react strongly to the patterns presented to them *and* those which do not react to these patterns are eliminated. Hence, a lymphocyte survives in primary lymphoid organs if, and only if, it reacts weakly to the 'self' constituents (Ashton-Rickardt et al. 1994), and not if it does not react at all. Furthermore, this selection continues throughout the lifetime of the organism: in peripheral organs (spleen, lymph nodes, etc.) circulating lymphocytes which do not react to 'self' antigens die (Freitas and Rocha 1999). Immune reaction to the 'self' is not only possible, but also necessary for a sound immune system. What do we mean in this case when we say that immune cells 'react' to the self? We mean that, constantly, some immune cells *interact* with normal, endogenous components of the organism (most of the time by binding to them). This *interaction* (the affinity and specificity of which can be evaluated) between the immune receptors and the 'self' ligands, nonetheless, is a necessary, but not a sufficient, condition for the immune *response*, that is, the *activation* of immune cells, or, in other words, the triggering of a cascade of modifications and *effector* mechanisms (such as cytokines production, phagocytosis, lysis, etc.).

However, immune reaction to 'self' components is not limited to mere *interaction*: we can also observe immune *responses* to 'self' components, that is, the *activation*, by 'self' components, of immune cells and immune molecules. One critical case of such activation is the phagocytosis (internalization and destruction) of cells of the organism which undergo changes in their patterns, particularly dead cells (by apoptosis, that is, programmed cell death). These cells are 'self' cells, but firstly they are recognized by immune cells as entities to be destroyed, and secondly some immune effector functions are triggered (Savill et al. 2002). Moreover, research on regulatory T cells (T_{Reg}) undertaken

during the past two decades points in the same direction. T_{Reg} are lymphocytes that *respond* to other, normal, lymphocytes in order to control their activation. In other words, the role of T_{Reg} is to stop or slow down immune reactions. T_{Reg} are 'self' cells that respond to other 'self' cells, and they are thought to be involved in the balance of autoimmunity, in tolerance of tumours, etc. (Sakaguchi 2004). Phagocytosis of dead cells and auto-regulation of T lymphocytes illustrate clearly the fact that the immune system is a set of homeostatic processes, in which reactions to 'self' components are indispensable and involve most of the time effector mechanisms similar to those responsible for reactions to pathogens (Savill et al. 2002). It is very likely that other components of the immune system have the same capacity to be activated by immune 'self' constituents, with the role of controlling (downregulating) their functioning.

Thus, immune reactions (interaction and response) to 'self' constitunts are not only possible, but also necessary. We can conclude that the principle "the immune system does not react to the 'self'" is false.

Falsity of the principle 'every foreign (originating from the outside) entity will trigger an immune reaction'

Here we lean on the concept of *immune tolerance*, which is defined as the *absence of immune response towards an antigen*. Naturally, 'self' has long been described as inducing tolerance, and, by contrast, 'non-self' is supposed not to trigger tolerance, except in a few cases (Burnet 1970, 46–49). The problem is that the more one looks at this question, the more one finds exceptions to the principle:

(i) *Tolerance of bacteria*. Many bacteria live in organisms without inducing immune effector reactions, and, in some cases, they are even beneficial to the host, especially on the mucosal surfaces (lungs, gut, sensory organs, organs of reproduction). The gut is thought to be colonized by 10^{14} commensal microorganisms, which contribute to the defence of the host, and to its digestive capacity (Berg 1996). The surface of the skin is also rich in bacteria.

(ii) *Tolerance of parasites*, i.e. protozoan parasites and parasitic worms (helminths). Parasites display, most of the time, large quantities of antigens at their surface, and yet in many cases they induce either no immune response, or no effective immune response (e.g. Malaria, which is caused by various species of the genus *Plasmodium*, and infects nearly 10% of the world's population). Many parasites, such as *Trypanosoma Cruzi*, remain several years in the body without being eliminated by the immune system (Buscaglia and Di Noia 2003)[3].

[3] At some point, an immune response against *Trypanosoma* cruzi can be triggered, but one cannot explain it by the self/non-self hypothesis, since the parasite is no more 'non-self' when the response occurs than before.

(iii) *Tolerance of grafts*. Some organs, called *immunoprivileged organs* (brain, eye, testis), do not reject alloantigens in case of transplantation (Ferguson and Griffith 1997).

(iv) *Foetomaternal tolerance*. Pregnancy is the most common, and the only non-artificial, graft, and we can observe that, in the great majority of situations, the foetus is not rejected by the mother. Although its genome is semi-different from that of the mother, the foetus does not trigger an immune reaction, or is protected against such a reaction. Induction of tolerance mechanisms, such as those due to the molecule of the Human Leukocyte Antigen called HLA-G (Carosella et al. 2003) and regulatory T cells (Aluvihare et al. 2004) have been proved to play a critical role in the acceptance of the foetus.

(v) *Chimerism*. The term *chimerism* refers to the process by which cells are exchanged between two organisms and maintained in at least one of them, in spite of their 'foreign' character. Different forms of chimerisms exist – twin cattle (Owen 1945), marmosets (Haig 1999), etc. – but the most striking example is foetomaternal chimerism: components originating from the child have been found in the mother's organism up to 27 years after birth (Bianchi et al. 1996).

Thus, 'self' does sometimes trigger an immune reaction; 'non-self' very often does not trigger any immune reaction. The self/non-self model is compelled therefore to define a series of exceptions to its own principles: tumour cells, apoptotic cells, for instance, are said to belong to the 'non-self', and commensal bacteria or the foetus are conceived of as parts of the 'self', which makes no sense with regard to the initial meaning of these two concepts, and creates a constant obscurity. Immunologists can ultimately have no other choice than to give, as an *explanation* of immune triggering, the law 'only non-self components induce an immune reaction', and, as a *definition* of the self, 'that which does not trigger any immune reaction' (while the immune non-self is that which does trigger an immune reaction)[4]. Obviously this is no explanation at all, but rather a purely logical circle, or a mere name for a phenomenon (which is far removed from Burnet's aim to find the best and easiest explanation for immune phenomena). Every scientific hypothesis can admit exceptions, but once the exceptions become too numerous, we have to accept the possibility that the general principles are flawed. We believe that another conception permits us to evade these exceptions, and constitutes therefore a better hypothesis for immunology.

The advantages of an alternative view: the continuity hypothesis

The aim of the continuity hypothesis is to offer an alternative *criterion of immunogenicity* with regard to the self hypothesis, that is, an actual explanation

[4] Criticizing the overestimation of the 'self', Silverstein and Rose (2000) write: "Self is only that collection of potential immunogens that cannot stimulate a response at that time and place!"

(and not a mere description) of what triggers an immune response and what does not. The continuity hypothesis relies on two main observations, which we have already mentioned. First, the immune system is acquired and not innate: everything that is *present* when the lymphocytic selection occurs will not trigger an immune reaction later (see Section 2 above). Second, *autoreactivity* is a constant, normal and necessary process in every organism. Here is the principle of the continuity hypothesis: *every strong discontinuity in the interactions between immune receptors and their targets triggers an immune response*. What does this principle mean exactly? This section endeavours to answer this question.

First, the *immune receptors* involved are those of lymphocytes, dendritic cells, macrophages, etc. Their *targets* are epitopes (i.e. molecular patterns), to which the immune receptors bind; these epitopes can be either *exogenous* ('non-self': pathogenic patterns on bacteria or viruses, alloantigens in a transplantation, etc.), or *endogenous* (tumour markers, molecular patterns expressed on apoptotic cells, patterns recognized by regulatory T cells, etc.) According to the continuity hypothesis, an immune response is induced by the appearance of epitopes (molecular patterns) that are different from those to which the receptors of the immune system *constantly* react. The immune receptors interact strongly with these abnormal epitopes, that is, with a strong affinity and/or specificity. Hence, the key is the advent of something in the immune system that bears *unusual* epitopes: the immune cells constantly react in a weak manner to the epitopes that remain the same or change very slowly[5] An immune response occurs when an epitope that is strongly different from those to which the immune system constantly reacts appears suddenly[6]. Thus, according to the continuity hypothesis, the immune system does not discriminate between 'self' and 'non-self', but between epitopes (whether endogenous or exogenous) that are constantly present in the organism – and hence induce constant interactions with the immune receptors of this organism – and epitopes (whether endogenous or exogenous) that strongly break the continuity of these interactions. This hypothesis requires us to understand more precisely the molecular nature of changes in cells.

So, in what circumstances does an immune response occur, according to our hypothesis? It occurs when, and only when, there is a strong discontinuity in

---

[5] Note that our criterium of immunogenicity can be formulated in terms of *usual* versus *unusual* epitopes, but not in terms of *never before seen* versus *new* epitopes, because it does not conflict with immune memory: a second encounter with a given antigen provokes a stronger and quicker response. An antigen already seen by the immune system triggers a stronger and quicker response if, when the first encounter took place, it induced a strong reaction, followed by an immune activation and a differentiation.

[6] Because it says that an immune reaction is due to a *strong* discontinuity, the continuity hypothesis takes into account three characteristics of the antigen: its degree of molecular difference (with regard to what the immune system constantly interacts with), the progressiveness (speed) of its appearance, and the duration of the interaction. The continuity and discontinuity are thus evaluated in spatio-temporal terms.

the interactions between immune receptors and epitopes, that is, each time the four following conditions are satisfied:

(i) An entity (whether endogenous or exogenous) that breaks the continuity of interactions appears and interacts with the immune system
(ii) This entity is in sufficient quantities (very small quantities of antigens do not induce an immune response, see 'insufficient discontinuity' below)
(iii) The discontinuity due to this entity is perceived by several components of the immune system, so there are several signals of activation
(iv) This entity triggers stress signals and/or proinflammatory signals.

When, by contrast, is there no effector immune response? In six kinds of circumstances:

(i) *Continuity*: There is no change in the molecular patterns to which immune components react. Hence, we can speak about 'actual continuity' or 'continuity by full absence of change'. An example is the normal functioning of an immune system. In this situation, there are no proinflammatory signals, and no modification of tissues.

(ii) *Imperceptible discontinuity*: Changes occur, but the immune system cannot perceive them, because there is no molecular difference between the two states. We can speak of 'continuity by similarity'. Thus, in this case, there is *immune continuity* in spite of the appearance of discontinuity from our point of view. Examples of this phenomenon include the acceptance of grafts between two identical twins, the acceptance of autografts, but also the preservation of continuity by molecular mimicry in some pathogens (e.g. the parasite *Trypanosoma Cruzi*: see Girones and Fresno 2003). There are no proinflammatory signals, and no modification of tissues.

(iii) *Ignorance*: There is no *interaction* between the immune receptors and the molecular patterns of the entity. This can be illustrated by the processes of isolation in 'immuno-privileged' sites: some immune components cannot *enter* the site, and therefore no recognition is possible (Ferguson and Griffith 1997).

(iv) *Insufficient discontinuity*: First, if antigens are in insufficient quantities, no immune reaction occurs (even if there is a discontinuity); second, immune tolerance is usually the default situation, which means that several components of the immune system must react to a given antigen for an immune response to occur. For instance, if a T lymphocyte recognizes its specific antigen on an antigen-presenting cell which is not itself activated, it dies; it is only if *both* cells (APC and T lymphocyte) react strongly to the antigen that an immune response is triggered.

(v) *Inhibition of the response*: Some pathogens and some tumours induce a decrease in the production of proinflammatory signals or actively deceive immune cells (Marincola et al. 2000).

(vi) *Induction of tolerance*: In this case, the entity is perceived by the immune system and reacts with several of its components, but nonetheless the antigenic patterns involved in the reaction are here considered as normal by the immune cells and are tolerated. (e.g., tolerance of the foetus which is semi-allogenic by the mother). In some cases, it even appears that, contrary to what usually

happens in the immune system, the more the organism encounters these patterns, the less it responds to them. Example: better tolerance of a graft if antigens from the donor are injected before the transplantation (Seung et al. 2003), as if the organism could get 'used' to some antigens. Why does this happen? First, it is important to understand that induction of tolerance involves *regulatory* immune mechanisms: an effector immune response can be initiated, but it is subsequently downregulated by the regulatory mechanisms, mainly the regulatory T cells (Belkaid and Rouse 2005), but also, for instance, the HLA-G molecule of histocompatibility (Lila et al. 2001) and the regulatory properties of dendritic cells (Smits et al. 2005). Hence, it is actually a form of induction of tolerance by *induction of continuity*, because in these circumstances effector cells which are likely to respond to the antigen are downregulated, killed, or rendered anergic by the regulatory components of the immune system. Thus, a new continuity is established between the immune receptors and this antigen (the immune cells with the strongest affinity or specificity being ruled out). Now, the question is: in what circumstances do such regulatory mechanisms outweigh the effector ones? Such a process, in which a discontinuity leads not to the destruction of an antigen but rather to tolerance, has both evolutionary (e.g. tolerance of foetus) and co-evolutionary (e.g. tolerance of some bacteria) reasons. More precisely, though, we can say that the conditions of this process are probably the following: small quantities of antigen (or progressive increasing of antigenic quantities), progressive contacts between the antigen and the immune receptors, no proinflammatory signals, no stress signals, no tissue damage. The absence of damage is critical here, since immunologists have long known that the same antigen may or may not trigger an immune reaction, depending on whether damage occurs or not. We believe that foetomaternal tolerance, foetomaternal chimerism, induction of tolerance in tumour cells, and some kinds of tolerance to pathogens (especially to some parasites), could all be examples of induction of continuity.

So, what is exactly the difference between the continuity hypothesis and the self hypothesis? Contrary to the self hypothesis, the continuity hypothesis says that 'self' components are perfectly capable of triggering an immune activation, provided that there is a break of continuity (e.g. phagocytosis of dead cells), and asserts that 'non-self' components can perfectly be tolerated by the organism, provided that either the break of continuity is not perceived by the immune components (*ignorance*), or the immune response is prevented (*inhibition*), or an active acceptation is induced (*induction of tolerance by continuity*).

Nonetheless, it is necessary not only to emphasize the differences between the two hypotheses, but also to prove that the continuity hypothesis possesses advantages in comparison with the self hypothesis. Several domains tend to prove that the continuity hypothesis is a better explanation of immune phenomena than the 'self' hypothesis:

Regulation of immunity: the functioning of the immune system as an homeostasis. The continuity hypothesis subsumes under a unique explanation the phagocytosis of dead or abnormal cells and immune reaction to pathogens: in

each case, it is the discontinuity in the molecular patterns expressed on the cell surface that triggers the immune response[7] It offers also an explanation for normal autoreactivity, that is, the necessity for immune cells to be stimulated constantly by endogenous components, and the role of regulatory cells, especially regulatory T cells (T_{Reg}). The case of T_{Reg} is particularly striking. While the self/non-self model has great difficulty in explaining the triggering of regulatory T cells (which can respond to 'self' as well as to 'non-self'), there is no such difficulty with the continuity hypothesis. T_{Reg} respond to a strong discontinuity in the interactions between their receptors and the epitopes to which they react (whether 'self' or 'non-self'), exactly like the other immune cells do. The only difference is that T_{Reg} downregulate the response mounted by other components of the immune system. When, in the balance between effector and regulatory mechanisms, the latter win, we have an induction of tolerance.

Tumour cells. Tumour cells, except perhaps those due to oncogenic viruses, are 'self' cells, in so far as they come from the genome of the individual and are components of the organism. The continuity hypothesis asserts that tumour cells, in most circumstances, trigger an immune response, because the molecular patterns expressed at their surface change considerably and this change thus constitutes a break of continuity[8].

Tolerance of pathogens, such as commensal bacteria or some parasites: these pathogens, especially when they do not harm the organism and even play a useful role (example: bacteria in the gut facilitate digestion), induce a tolerance by progressive induction of continuity (Hooper and Gordon 2001). *Leishmania major* actively induces interleukin 10 (IL-10) producing T_{Reg}, and these T_{Reg} prevent the clearance of the parasite by other immune cells (Belkaid et al. 2002).

Foetomaternal tolerance and chimerism. As suggested previously, in this case, induction of tolerance (Claas 2004) may be realized by induction of continuity.

Thus, the continuity hypothesis attempts to give an account of phenomena that the immune self/non-self model does not explain, or explains only by using *ad hoc* hypotheses. The continuity hypothesis offers a more comprehensive and less metaphysically loaded explanation of immunity than the self/non-self model; hence we think it can prepare the ground for other ways in which to understand the functioning of the immune system.

This alternative hypothesis based on the notion of continuity is rooted in a metaphysical definition of identity different from the one of the self hypothesis. Indeed, the conceptual shift from *self* to *continuity* in immunology has a

---

[7] There is clear evidence for this discontinuity in dead cells: see Albert (2004).

[8] Changes in tumour cells are indeed very different from changes in normal cells: the genome of normal cells is stable whereas cancer cells undergo multiple genetic alterations; the transcriptome in normal cells is stable, whereas cancer cells are characterized by a major epigenetic instability; no tissue invasion occurs with normal cells, whereas there is invasion and metastasis with cancer cells; normal cells have a stable pattern of cytokine and growth factor expression, by contrast with cancer cells which have an abnormal expression of cytokines and growth factors (Pardoll 2003).

philosophical analogue in the metaphysical deflation from a definition of identity based on *substance* (identity-substance) to a definition of identity based on *continuity* (identity-continuity). The process of metaphysical deflation was operated in philosophy by authors usually, although questionably, subsumed under the term of 'empiricism'. In the next section we argue that there is a rich mapping between the immunological opposition *continuity versus self* and the philosophical opposition *empiricism versus substantialism* about identity, and that these two rival metaphysical conceptions underlie implicitly the two scientific hypotheses. Our claim will be that immunology must operate a shift from identity-substance (self hypothesis) to identity-continuity (continuity hypothesis).

The opposition between two metaphysical conceptions of identity: substance *versus* continuity

First, what do we call 'identity'? A being is defined by two aspects: first, the individual characteristics that make it distinct and different from everything else (its location in space and time, its physical and psychological characteristics, etc.); second, the fact that, in spite of the changes that occur to it, it can be said to remain the 'same' being. These two aspects are *individuality* on the one hand, and *sameness* on the other hand (Wiggins 2001).

There are two rival definitions, an *internalist* one and an *externalist* one, of the identity of an organism. Since both are extreme conceptions, a wide range of conceptions naturally exists in-between. According to the internalist conception, the fundamental characteristics that constitute the identity of an organism (that is, its individuality and its sameness) are endogenous (that is, they come from the inside of the organism), and everything exogenous (that is, coming from the outside of the organism) represents a threat to its integrity (that is, to the maintaining of this self-defining process). According to the externalist conception, on the other hand, the fundamental characteristics that constitute the identity of an organism are exogenous, and nothing happens inside an organism without interacting in many ways with the environment. This opposition is sketched out clearly by Richard Lewontin (2000). In this section, we suggest that the immunological self hypothesis tends to belong to the internalist conception, whereas the immunological continuity hypothesis constitutes a strong suggestion to move towards an externalist conception, and we show that these two views are rooted in the metaphysical opposition between *substance* on the one hand, and *continuity* on the other hand.

At least since the seventeenth century, two philosophical definitions of identity as sameness have fought: according to the first (identity-substance), a being is the same being by remaining the same *substance*, that is by virtue of maintaining a metaphysical core throughout changes; according to the second (identity-continuity), a being remains the *same* simply because of the spatio-temporal connexion between its constituents (if the connexion is broken,

identity vanishes). More precisely, identity as *substance* is based on the idea that, in spite of the changes that affect any individual entity, there is a metaphysical *core* that is preserved throughout time. This core is supposed to explain why an adult is the *same* being as the child he or she was, even if there are very few similarities between his or her two conditions, or why an acorn is the same being as the oak tree it becomes. Thus, the concept of substance is a way to understand *sameness*: the immutability of the metaphysical core of each individual assures its identity and preserves its integrity. Such a conception of identity can be found in Aristotle's *Categories* (Chapter 5, 3b22–33 and 4b18–19) and *Metaphysics* (Book D). It can be found in contemporary philosophy, also. For instance, Wiggins leans on Aristotle's thought to suggest "an unmysterious but pre-empiricist notion of substance" (2001: 80): continuity of states by itself does not define identity, there is a *something* to which these states belong[9]. The immune self/non-self model can be said to uphold a *substantialist* view of identity, because the immunological conception of the self is grounded on the idea of preservation of the integrity of the organism, integrity which must be maintained against any external threat, exactly as individual substance is defined on the basis of its always preserved metaphysical core. More than this: the immune self/non-self model holds a *monadologic* view of identity Leibniz 1714), that is, it conceives of identity as *self-defining* and *enclosed*: the immune system "knows" itself, and every change that comes from the inside is tolerable (acceptable), whereas every change that comes from the outside ('foreign') is to be rejected. Thus, in several of his books (see particularly Burnet 1962 and Burnet 1969), Burnet emphasizes the two main ideas of *self-knowledge* and *preservation* (exactly like a monad):

> It is one of the concise statements of modern immunology that the body will accept as itself only what is genetically indistinguishable from the part replaced… It is as if the body can recognize its own individuality and will accept nothing that is inconsistent with that individuality (1962: 13–14).

Furthermore, contemporary immunology provides constant illustrations of this monadologic conception. Here follows one example:

> Thus, we envision self-nonself discrimination as being mediated by both arms of the immune response [that is, innate immunity and adaptative immunity]. The innate immune response to pathogen-associated molecular patterns (PAMPs) distinguishes non-infectious self from infectious

[9] "Wherever the sortal concept under which *a* and *b* coincide is the sortal concept for a kind of continuant and one can ask 'what is it for an f to persist?', it is the idea of a sequential history of a thing's doings and undergoings that comes into consideration. … As Leibniz puts the point that I too have wanted to insist upon, 'By itself continuity no more constitutes substance than does multitude or number … Something is necessary to be numbered, repeated and continued" (Wiggins 2001: 57).

> nonself, and does so with great accuracy. The adaptive immune system of lymphocytes is *basically self-referential*, being selected positively for the recognition of self-peptide: self-MHC molecules, and negatively selected by those self-ligands that are able to activate developing T cells to induce programmed cell death… However, for full activation in the peripheral tissues, they need to see antigen on the same cell as the expression of the co-stimulatory molecules induced by innate immune recognition of *foreign* PAMPs. (Medzhitov and Janeway 2000, our emphasis).

Here, the ideas that the immune system is *self-centred*, and that only *foreign* antigens can induce an actual immune response, are clearly expressed. As we see, the immune 'self' consists of the unfolding of internal processes (*self-definition*), the result of which must always be defended against any external presence (*enclosure*). In other words, the immune self/non-self model can be said to support a monadologic conception of identity in the sense that it is based on *integrity*, *internalism* and *enclosure*. In the self/nonself framework, the self production of any organism is understood as the harmonious unfolding of individuality, whereas foreignness appears as a threat on the integrity of that individuality (Klein 1982)[10].

The continuity hypothesis, on the other hand, conceives of identity as an identity-continuity, since it claims that nothing more than the spatiotemporal continuity of adhesions between immune receptors and ligands defines immune identity. This hypothesis can therefore be seen as the immunological point of view on the identity of organisms, as Boniolo and Carrara (2004) very recently defined it[11]. According to the continuity hypothesis, nothing like a permanent 'core' to be preserved against all foreign threats is presupposed and thought to define immunity. Changes from the inside and changes from the outside equally can trigger an immune response, depending on the conditions of encounter we described above.

In the next section, we show that the continuity hypothesis tends to belong to the externalist views of biological identity, and we suggest that adopting these views offers many advantages to current biology.

---

[10] "An organism is a miracle of harmony. Its individual parts have divided among themselves the functions necessary for the sustenance of life and interact with one another in a complex but precisely defined manner. However, this delicate system of interactions can easily be disturbed by any agent that endangers the body's integrity." (Klein 1982, 3). Wiggins himself attempts to suggest an internalist description of biological substances: "But now we are led by simple conceptual considerations to precisely the account of living substances that biologists can fill out *a posteriori* by treating them as systems open to their surroundings, not in equilibrium with those surroundings, but so constituted that a delicate self-regulating balance of serially linked enzymatic degradative and synthesizing chemical reactions enables them to renew themselves" (Wiggins 2001: 86).

[11] In their paper, Boniolo and Carrara (2004) even deal with immunological identity, but, since it is not specifically their problem, they do not elaborate the point. We believe that our definition of immune identity as based on the continuity hypothesis can be seen as an instance of their suggestion: in their terms, the continuity hypothesis reflects the particular immunological point of view V on organisms and, therefore, the particular isolation of N immune properties P_i ($i = 1,\ldots,N$).

Theoretical status of the continuity hypothesis

The continuity hypothesis shares a feature with the theories of the immune 'system' or 'network' (Jerne 1974; Cohen 1992; Atlan and Cohen 1998; Tauber 2000, and, for a use of the very questionable concept of 'autopoiesis', Maturana and Varela 1980): immune reactions are defined as *perturbations of the system*. Within the continuity hypothesis, an immune reaction follows the arrival of an *unusual* entity (break of continuity). Nonetheless, there is a critical difference between the two theoretical frameworks: in the theories of the immune network, the critique of the distinction between 'self' and 'non-self' relies on the assertion that there is only the 'self'. Indeed, autoreactivity is seen as essential and as the basic definition of immunity: the immune system constantly reacts to self constituents, hence autoimmunity is not a dysfunction, but the basis of normal immunity (Cohen 2000). In other words, the immune system always reacts to itself because it 'sees' only itself (Jerne 1974). Thus, network theories hold an even stronger *internalist* view than the 'self' model. By contrast, the continuity hypothesis moves towards *externalism*: relying on the *induction of continuity*, this hypothesis attempts to explain the relative *openness* of organisms. Biological identity is seen as *continuous* and *open*, which means that it is defined as a succession of states without a permanent core, and that the integration of 'foreign' elements is in many circumstances a normal and necessary process in organisms. This relative openness, explained here by the induction of continuity, reflects a great variety of known phenomena, such as the bacterial origin of mitochondria in our cells (Margulis and Chapman 1998, Martin and Müller 1998), all kinds of chimerisms and particularly foetomaternal chimerism (Bianchi et al. 1996), and environmentally dependent development (Gilbert 2002). This view provides a basis for the externalist or rather 'heterogeneous' view of biological identity (supported by Lewontin 2000[12]) in the field where it is the least expected, because of the vision of organisms as fortresses, the integrity of which has constantly to be preserved, that is, immunology.

Conclusion

The self/non-self model and its vocabulary are inadequate and misleading. The continuity hypothesis offers a more comprehensive and less metaphysically loaded theory to give a proper account of immunological phenomena. It is time to apply to immunology the principle of metaphysical deflation, that is, to operate a transition from identity-substance (self hypothesis) to identity-continuity

[12] Lewontin's dialectical conception of organisms rejects externalism as far as *evolution* is concerned and internalism as far as *development* and *maintenance* of organisms are concerned (Godfrey-Smith 2001). Our use of Lewontin's thought deals with the second aspect. In this sense, Lewontin writes: "Organisms are internally heterogeneous open systems" (2000: 114).

(continuity hypothesis). The *monadologic* conception of identity that underlies implicitly the self/non-self model is based on self-definition and enclosure. This monadologic (ultra-internalist) conception can be found, in addition to immunology, in other parts of biology, particularly genetics, with the idea that the individual can be defined as the unfolding of the genetic information contained in the ADN of the nucleus of his cells, and development (Oyama 1985; Lewontin 2000). By contrast, the continuity hypothesis enables us to understand identity as an open and integrative identity, constructed by the interrelations of an organism and its environment.

Acknowledgments

We thank Peter Godfrey-Smith, Anouk Barberousse, Alfred Tauber, Richard Lewontin, Giovanni Boniolo, Jean Gayon, Guy-Cédric Werlings, and Hannah-Louise Clark for their help and comments.

References

Albert M.L. 2004. Death-defying immunity: do apoptotic cells influence antigen processing and presentation? Nat. Rev. Immunol.4: 223–231.

Aluvihare V.R., Kallikourdis M. and Betz A.G. 2004. Regulatory T cells mediate maternal tolerance to the fetus. Nat. Immunol. 5(3): 266–271.

Aristotle 1984. Categories. In: Barnes J. (eds), The Complete Works of Aristotle, the revised Oxford translation. Princeton University Press, Princeton.

Aristotle 1984. Metaphysics. In: Barnes J. (eds), The Complete Works of Aristotle, the revised Oxford translation. Princeton University Press, Princeton.

Ashton-Rickardt P.G., Bandeira A., Delaney J.R., Van Kaer L., Pircher H.P., Zinkernagel R.M. and Tonegawa S. 1994. Evidence for a differential avidity model of T-cell selection in the thymus. Cell 76(4): 651–663.

Atlan H. and Cohen I.R. 1998. Immune information, self-organization and meaning. Int. Immunol. 10(6): 711–717.

Belkaid Y., Piccirillo C.A., Mendez S., Shevach E.M. and Sacks D.L. 2002. $CD4^+CD25^+$ regulatory T cells control Leishmania major persistence and immunity. Nature 420: 502–507.

Belkaid Y. and Rouse B.T. 2005. Natural regulatory T cells in infectious disease. Nat. Immunol. 6(4): 353–360.

Berg R.D. 1996. The indigenous gastrointestinal microflora. Trends Microbiol. 4: 430–435.

Bianchi D.W., Zickwolf G.K., Weil G.J., Sylvester S. and DeMaria M.A. 1996. Male fetal progenitor cells persist in maternal blood for as long as 27 years postpartum. Proc. Natl. Acad. Sci. 93: 705–707.

Billingham R.E., Brent L. and Medawar P.B. 1953. Actively acquired tolerance of foreign cells. Nature 172: 603–606.

Boniolo G. and Carrara M. 2004. On Biological Identity. Biol. Phil. 19: 443–457.

Burnet F.M. and Fenner F. 1949. The Production of Antibodies, 2nd edn. Macmillan, Melbourne.

Burnet F.M. 1957. A modification of Jerne's theory of antibody production using the concept of clonal selection. Aust. J. Sci. 20: 67–69.

Burnet F.M. 1959. The Clonal Selection Theory of Acquired Immunity. Cambridge University Press, Cambridge.

Burnet F.M. 1960. Immunological recognition of self. Nobel Lect. Physiol. Med. 3: 689–701.

Burnet F.M. 1962. The Integrity of the Body. Harvard University Press, Cambridge, Massachusetts.
Burnet F.M. 1965. The Darwinian approach to immunity. In: Sterzl J. (eds), Molecular and Cellular Basis of Antibody Formation. Academic Press, New York, pp. 17–20.
Burnet F.M. 1969. Self and Not-self. Melbourne University Press and Cambridge University Press, Victoria.
Burnet F.M. 1970. Immunological Surveillance. Pergamon, Sydney.
Buscaglia C.A. and Di Noia J.M. 2003. Trypanosoma cruzi clonal diversity and the epidemiology of Chagas' disease. Microbes and Infection 5: 419–427.
Carosella E.D., Moreau P., Le Maoult J., Le Discorde M., Dausset J. and Rouas-Freiss N. 2003. HLA-G molecules: from maternal-fetal tolerance to tissue acceptance. Adv. Immunol. 81: 199–252.
Claas F. 2004. Chimerism as a tool to induce clinical transplantation tolerance. Curr. Opin. Immunol. 16: 578–583.
Cohen I.R. 1992. The cognitive paradigm and the immunological homunculus. Immunol. Today 13: 490–494.
Cohen I.R. 2000. Tending Adam's Garden – Evolving the Cognitive Immune Self. Academic Press, San Diego.
Douek D.C., Picker L.J. and Koup R.A. 2003. T Cells Dynamics in HIV-1 Infection. Annu. Rev. Immunol. 21: 265–304.
Ehrlich P. 1900. On immunity with special reference to cell life. Proc. Royal Soc. 66: 424–448.
Ferguson T.A. and Griffith T.S. 1997. A vision of cell death: insights into immune privilege. Immunol. Rev. 156: 167–184.
Freitas A.A. and Rocha B. 1999. Peripheral T cell survival. Curr. Opinion Immunol. 11: 152–156.
Garcia K.C., Degano M., Pease L.R., Huang M., Peterson P.A., Teyton L. and Wilson I.A. 1998. Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. Science 279: 1166–1172.
Gilbert S.F. 2002. The genome in its ecological context. Ann. N.Y. Acad. Sci. 981: 202–218.
Girones N. and Fresno M. 2003. Etiology of Chagas disease myocarditis: autoimmunity, parasite persistence, or both? Trends Parasitol 19(1): 19–22.
Godfrey-Smith P. 2001. Organism, environment, and dialectics. In: Singh R.S., Krimbas K., Paul D.B. and Beatty J. (eds), Thinking About Evolution: Historical, Philosophical and Political Perspectives. Cambridge University Press, New York.
Haig D. 1999. What is a marmoset? Am. J. Primatol. 49: 285–296.
Hooper L.V. and Gordon J.I. 2001. Commensal host-bacterial relationships in the gut. Science 292: 1115–1118.
Jerne N.K. 1974. Towards a network theory of the immune system. Ann. d'immunologie 125(C): 373–389.
Kappler J.W., Roehm N. and Marrack P.T. 1987. T cell tolerance by clonal elimination in the thymus. Cell 49: 273–280.
Klein J. 1982. Immunology: The Science of Self-Nonself Discrimination. Wiley, New York.
Langman R.E. and Cohn M. 2000. A minimal model for the self-nonself discrimination: a return to the basics. Sem. Immunol. 12: 189–195.
Leibniz G.W. 1714. Monadologie, translation Monadology and Other Philosophical Essays. edited and trans. P. and A.M. Schrecker, 1965. Bobbs-Merrill, Indianapolis.
Lewontin R. 2000. The Triple Helix: Gene, Organism, and Environment. Harvard University Press, Cambridge, Massachusetts.
Lila N., Rouas-Freiss N., Dausset J., Carpentier A. and Carosella E.D. 2001. Soluble HLA-G protein secreted by allo-specific $CD4^+$ T cells suppresses the allo-proliferative response: a $CD4^+$ T cell regulatory mechanism. Proc. Natl. Acad. Sci. 98(21): 12150–12155.
Löwy I. 1991. The immunological construction of the self. In A.I. Tauber, ed., Organism and the Origins of Self, Kluwer, Boston, (1991), pp. 43–75.
Margulis L. and Chapman M. J. 1998. Endosymbioses: cyclical and permanent in evolution. Trends Microbiol. 6: 342–345.

Marincola F., Jaffee E., Hicklin D. and Ferrone S. 2000. Escape of human solid tumors from T-cell recognition: molecular mechanisms and functional significance. Adv. Immunol. 74: 181–273.

Martin W. and Müller M. 1998. The hydrogen hypothesis for the first eukaryote. Nature 392: 37–41.

Maturana H.R. and Varela F.J. 1980. Autopoiesis and Cognition – The Realization of the Living D. Reidel. Boston Studies in the Philosophy of Science 42: Boston.

Matzinger P., Gallucci S. and Lolkema M. 1999. Natural adjuvants: Endogenous activators of dendritic cells. Nat. Med. 5((11): 1249–1255.

Medzhitov R. and Janeway C. 2000. How does the immune system distinguish self from non-self? Sem. Immunol. 12: 185–188.

Medzhitov R. and Janeway C. 2002. Decoding the Patterns of Self and Nonself by the Innate Immune System. Science 296: 298–300.

Miller J.F.A.P. 1961. Immunological function of the thymus. The Lancet ii: 748–749.

Murphy J.B. 1913. Transplantability of tissues to the embryo of foreign species. J. Exp. Med. 17: 482–492.

Nanda N. and Sercarz E. 1995. Induction of anti-self-immunity to cure cancer. Cell 82: 13–17.

Owen R. 1945. Immunogenetic consequences of vascular anastomoses between bovine twins. Science 102: 400–401.

Oyama S. 1985. The Ontogeny of Information – Developmental Systems and Evolution. Cambridge University Press, Cambridge. 2nd ed. 2000.

Pardoll D. 2003. Does the immune system see Tumors as Foreign or Self? Ann. Rev. Immunol. 21: 807–839.

Pedotti R. and Mitchell D. 2001. An unexpected version of horror autotoxicus: anaphylactic shock to a self-peptide. Nat. Immunol. 3: 216–222.

Pradeu T. and Carosella E.D. 2004. Analyse critique du modèle immunologique du soi et du non-soi et de ses fondements métaphysiques implicites. Comptes Rendus de l'Académie des Sciences Biologies 327(5): 481–492.

Sakaguchi S. 2004. Naturally arising $CD4^+$ regulatory T cells for immunologic self-tolerance and negative control of immune responses. Ann. Rev. Immunol. 22: 17.1–17.32.

Savill J., Dransfield I., Gregory C. and Haslett C. 2002. A blast from the past: clearance of apoptotic cells regulates immune responses. Nat. Rev. Immunol. 2: 965–975.

Seung E., Mordes J.P, Rossini A.A. and Greiner D.L. 2003. Hematopoietic chimerism and central tolerance created by peripheral-tolerance induction without myeloablative conditioning. J. Clin. Invest. 112: 795–808.

Silverstein A.M. and Rose N.R. 1997. On the mystique of the immunological self. Immunol. Rev. 159: 197–206.

Silverstein A.M. and Rose N.R. 2000. There is only one immune system! The view from immunopathology. Sem. Immunol. 12: 173–178.

Smits H.H., de Jong E.C., Wierenga E.A. and Kapsenberg M.L. 2005. Different faces of regulatory DCs in homeostasis and immunity. Trends Immunol. 26(3): 123–129.

Tauber A.I. 1994. The Immune Self: Theory or Metaphor? Cambridge University Press, Cambridge.

Tauber A.I. 1999. The elusive immune self: a case of category errors. Perspect. Biol. Med. 42: 459–474.

Tauber A.I. 2000. Moving beyond the immune self? Sem. Immunol. 12: 241–248.

Walker L.S.K. and Abbas A.K. 2001. The enemy within: keeping self-reactive T cells at bay in the periphery. Nat. Rev. Immunol. 2: 11–19.

Wells H.G., Huxley J.S. and Wells G.P. 1929. The Science of Life. The Literary Guild, New York. (1934).

Wiggins D. 2001. Sameness and Substance Renewed. Cambridge University Press, Cambridge.

Wilson D. 1972. The Science of Self – A Report of the New Immunology. Longman, London.

Zinkernagel R.M., Callahan G.N., Althage A., Cooper S., Klein P.A. and Klein J. 1978. On the thymus in the differenciation of ''H-2 self-recognition'' by T cells: evidence for dual recognition? J. Exp. Med. 147: 882–896.