



## How Diverse should the Immune System be?

Rob J. De Boer, Alan S. Perelson

*Proceedings: Biological Sciences*, Volume 252, Issue 1335 (Jun. 22, 1993), 171-175.

---

Your use of the JSTOR database indicates your acceptance of JSTOR's Terms and Conditions of Use. A copy of JSTOR's Terms and Conditions of Use is available at <http://www.jstor.org/about/terms.html>, by contacting JSTOR at [jstor-info@umich.edu](mailto:jstor-info@umich.edu), or by calling JSTOR at (888)388-3574, (734)998-9101 or (FAX) (734)998-9113. No part of a JSTOR transmission may be copied, downloaded, stored, further transmitted, transferred, distributed, altered, or otherwise used, in any form or by any means, except: (1) one stored electronic and one paper copy of any article solely for your personal, non-commercial use, or (2) with prior written permission of JSTOR and the publisher of the article or other text.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

*Proceedings: Biological Sciences* is published by The Royal Society. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org>.

---

*Proceedings: Biological Sciences*

©1993 The Royal Society

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact [jstor-info@umich.edu](mailto:jstor-info@umich.edu).

©2001 JSTOR

# How diverse should the immune system be?

ROB J. DE BOER<sup>1</sup> AND ALAN S. PERELSON<sup>2</sup>

<sup>1</sup> *Theoretical Biology, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*

<sup>2</sup> *Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, U.S.A.*

## SUMMARY

We develop a probability-based model that suggests a novel explanation for the enormous diversity of the lymphocyte repertoire in the immune system. Taking into account the fraction of the repertoire that is expected to be rendered tolerant because of potential self reactivity, we calculate  $P_E$ , the probability that a pathogen escapes recognition by the immune system. Obviously, for  $P_E$  to be low the diversity should be large. Our novel result is that the major factor in this relation between diversity and  $P_E$  is formed by the number of self antigens. Thus our model suggests that, rather than the number of foreign antigens the system is protecting against, the diversity reflects the number of self antigens the immune system needs to avoid reactivity with.

## 1. INTRODUCTION

The vertebrate immune system is composed of B and T lymphocytes. Each lymphocyte clone is identified by a unique antigen receptor. The diversity of lymphocyte receptors is so large that virtually any antigen is recognized by at least one of the lymphocyte clones. Thus, each pathogen evokes a specific immune response. It is in an open question why the immune repertoire is so large. A large repertoire is not required for the recognition of a large variety of antigens. This is illustrated by the major histocompatibility complex (MHC) molecules that play an essential role in the presentation of peptides to T lymphocytes. Although any individual has only a limited number of different MHC molecules, most individuals are perfectly capable of presenting peptides from virtually any antigen.

The question that we address here is how large the repertoire should be so that virtually all pathogens are recognized by at least one lymphocyte clone. Perelson & Oster (1979) have previously addressed this question for B cell repertoires. Here we expand their discussion to take into account self tolerance and the T cell repertoire. When the repertoire of lymphocyte receptors recognizes virtually any antigen it also contains specificities recognizing self antigens, i.e. molecules from within the organism. We here suppose that such 'auto-reactive' clones are functionally deleted from the repertoire. Several self-tolerance processes have been described for both B cells (Nossal 1983; Schwarz 1989; Hartley *et al.* 1991; Goodnow 1992) and T cells (Kappler *et al.* 1987, 1988; MacDonald *et al.* 1988; Mueller *et al.* 1989); see Kroemer & Martinez-A. (1992) for a review. The efficiency of self tolerance is still debated because the normal repertoire contains auto-reactive specificities (Avrameas 1991). This is especially true for B cells, i.e. for IgM antibodies. We will here assume that those lymphocytes that recognize any self antigen with high affinity are functionally removed from the repertoire, and that the functional

repertoire after tolerance induction is still sufficiently large that virtually all pathogens are recognized, i.e. that the repertoire is complete. We ignore idiotypic interactions: the concentrations of the idiotypes are considered to be too low for tolerance induction.

Each pathogen is recognized by any of several antigens. For example, enteric bacteria are recognized by O antigens, and group A streptococci by M protein. We define  $a$  as the average number of antigens per pathogen. We also define a 'probability of escape',  $P_E$ , which is the probability that no clone in the functional repertoire is able to recognize a given antigen. Thus, the probability of not recognizing a pathogen expressing  $a$  antigens is  $P_E^a$ . Evolution is expected to minimize  $P_E^a$ . This evolutionary minimization is expected to halt around a value of  $P_E^a$  at which the organism is very likely to survive into adulthood. Implicit in our model is the assumption that the antigens an individual encounters are random samples from a potentially very large antigenic universe. We do not consider the special case where repertoires have evolved to preferentially recognize specific foreign antigens.

This upper bound of  $P_E^a$  can be estimated from the total number of deleterious pathogens an organism is expected to encounter during its life. The first failure in recognition can be computed from the geometric distribution and is expected after  $P_E^{-a}$  pathogens. Thus, for an immune system in which 99% of all antigens are recognized, i.e.  $P_E = 0.01$ , and an environment in which  $a = 2$ , the first failure is expected after  $10^4$  pathogens. Thus  $P_E^a = 10^{-4}$  seems a reasonable order of magnitude. Below we show that our results are not sensitive to these estimates because our expressions depend only on the logarithm of  $P_E$ .

## 2. MODEL

In the immune system an antigen is recognized as 'epitopes', i.e. as regions of the antigen that bind B or

T cell receptors. Let  $P_R$  be the probability that a randomly chosen receptor recognizes a randomly chosen epitope with sufficient affinity to trigger the cell. Empirical estimates of  $P_R$  vary around  $P_R \approx 10^{-5}$  (Cancro *et al.* 1978; Milon *et al.* 1986). Because  $P_R \ll 1$  we will approximate  $\ln(1 - P_R)$  as  $-P_R$ , and hence  $(1 - P_R)^n$  as  $\exp[-nP_R]$ . Let  $R_0$  be the size of the 'pre-immune' repertoire, i.e. before tolerance induction, and let  $R$  be the 'functional' repertoire, i.e. after tolerization. Because we concentrate on initial recognition abilities we do not consider somatic mutation. For mice it is estimated that the functional repertoire of both T and B lymphocytes is comprised of the order of  $10^7$  specificities (Du Pasquier 1982; Berek & Milstein 1988; Davis & Bjorkman 1988). It is estimated that T cell tolerance involves 50–70% of the lymphocytes in the thymus (Scollay & Shortman 1985). For murine B cells this is unknown (Hartley *et al.* 1991; Goodnow 1992), but in the rat 75% of surface IgM<sup>+</sup> bone marrow B cells appear to be lost in the bone marrow (Deenen *et al.* 1990). We define  $\epsilon$  as the average number of epitopes per antigen. An order of magnitude is  $\epsilon = 10$ . Let  $n$  be the number of self antigens, hence,  $n\epsilon$  is approximately the number of self epitopes: if the same epitope appears in different self antigens then the number of self epitopes will be less than  $n\epsilon$ . However, as  $n$  is not known with great accuracy, a typical estimate for the number of self antigens is  $10^4 \leq n \leq 10^5$  (Cohn 1987); the potential discrepancy between  $n\epsilon$  and the true number of self epitopes is probably of no consequence.

### 3. B CELL REPERTOIRE

For reasons of simplification we first use our model for analysing the B cell repertoire. The analysis of the T cell repertoire is more complicated because it involves antigen presentation. Self tolerance has been shown for B cells (Nossal 1983; Schwarz 1989; Hartley *et al.* 1991; Goodnow 1992) but remains an issue of debate. Rather than being deleted, clones may be anergized or incorporated into a network that controls their self reactivity.

We calculate  $f$ , the fraction of receptors surviving the self-tolerance process. The probability that a randomly chosen receptor fails to recognize any of the  $n\epsilon$  self epitopes is

$$f = (1 - P_R)^{n\epsilon} \simeq \exp(-P_R n\epsilon). \quad (1a)$$

Assuming  $f$  is the fraction of receptors surviving self tolerance, the functional repertoire is

$$R = fR_0. \quad (1b)$$

We expect that natural selection acts upon the probability that an immune response to a pathogen fails. For a pathogen corresponding to  $a$  antigens, each with  $\epsilon$  epitopes, and for a functional repertoire of size  $R$ , the probability of escaping recognition by all  $R$  receptors is

$$P_E^a = (1 - P_R)^{R\epsilon a} \simeq \exp(-P_R R \epsilon a). \quad (2)$$

Substituting equation (1) into equation (2) we obtain,

$$R_0 \simeq -\ln(P_E) [\exp(P_R n\epsilon)] / P_R \epsilon. \quad (3)$$

For the empirical estimates given above, i.e.  $P_R = 10^{-5}$ ,  $10^4 \leq n \leq 10^5$ ,  $\epsilon = 10$ ,  $P_E = 10^{-2}$ , equation (3) predicts  $10^5 \leq R_0 \leq 10^9$ . Thus, this prediction is very sensitive to  $n$ .

Equation (3) gives the size of the pre-immune repertoire required to insure a complete repertoire, i.e. one in which the probability of antigen escape is  $P_E$ . It predicts that the larger the complexity of the individual, i.e. the larger the number of self antigens,  $n$ , the larger the required repertoire  $R_0$ . Additionally, if evolution selects for a lower  $P_E$ , the repertoire  $R_0$  should also increase. However, because the latter depends on  $\ln P_E$ , whereas the former depends exponentially on  $n$ , we conclude that immune systems are diverse principally because of the large number of self antigens. This is the main conclusion of the paper.

In figure 1a we plot the pre-immune repertoire  $R_0$  as a function of  $P_R$  and  $n$ . It is a function with a single minimum, and the location of the minimum is independent of  $P_E$  and  $a$ . It is possible to proceed from here by making a disputable evolutionary assumption. Because  $R_0$  as a function of  $P_R$  has a minimum, evolution may select for a receptor specificity  $P_R$  that requires the smallest repertoire. This minimum, found by setting  $\partial R_0 / \partial P_R = 0$ , is at

$$\bar{P}_R = 1/n\epsilon, \quad (4)$$

where the overbar denotes an optimal value. (Interestingly, the same optimal receptor specificity is predicted if one assumes that evolution minimizes  $P_E$  at

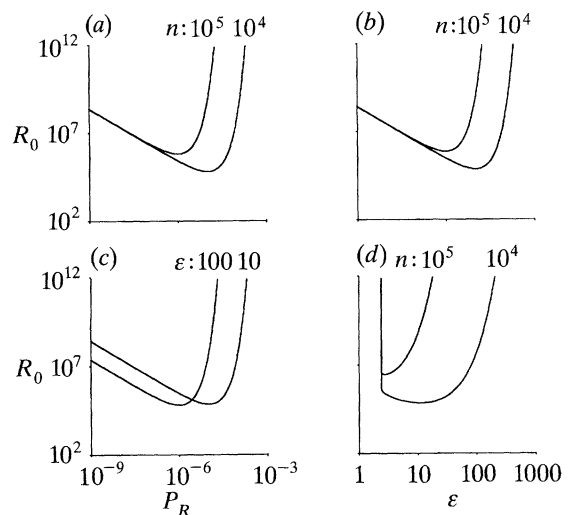


Figure 1. Pre-immune repertoire diversity,  $R_0$ , as a function of lymphocyte specificity,  $P_R$ , or of the number of MHC-peptide complexes,  $\epsilon$ . (a) Equation (3), i.e. the B cell repertoire  $R_0$ , for  $P_E = 10^{-2}$ ,  $\epsilon = 10$ ,  $n = 10^4$  and  $n = 10^5$ . This is a function with a single minimum at  $P_R = 1/(n\epsilon)$ . Panels (b–c) plot equation (11), i.e. the T cell repertoire  $R_0$ . (b) Equation (11), i.e. the T cell repertoire  $R_0$ , for  $P_E = 10^{-2}$ ,  $\epsilon = 10$ ,  $n = 10^4$  and  $n = 10^5$ . On this scale it is hardly visible that the T cell repertoire is larger than the B cell repertoire. The minimum is at the same specificity  $P_R = 1/(n\epsilon)$ . (c) Equation (11) for  $P_E = 10^{-2}$ ,  $n = 10^4$ ,  $\epsilon = 10$  and  $\epsilon = 100$ . The location of the minimum depends on  $\epsilon$ . (d) Equation (11) for  $P_E = 10^{-2}$ ,  $P_R = 10^{-5}$ ,  $n = 10^4$  and  $n = 10^5$ . Because of the minimum there is an optimal number of MHC molecules at which the required repertoire is minimal.

fixed  $R_0$ .) Note that for our empirical estimates of  $P_R$ ,  $n$  and  $\epsilon$ ,  $P_R$  is close to its optimal value. Substituting equation (4) into equation (3) gives the 'minimal' pre-immune repertoire

$$\bar{R}_0 \simeq -e \ln(P_E) n, \quad (5)$$

where  $e$  is the base of the natural logarithm. Substituting equation (4) into equation (1a), we see that with the optimal value of  $P_R$  the fraction of B cells surviving self tolerance is  $\bar{f} \simeq \exp[-1] \simeq 0.37$ . Thus, roughly 63% of the specificities are expected to be tolerized. In the rat, about 75% of the bone marrow surface IgM<sup>+</sup> B lymphocytes are lost (Deenen *et al.* 1990).

For the minimal functional B cell repertoire after tolerance induction we obtain

$$\bar{R} = \bar{R}_0 \bar{f} \simeq -\ln(P_E) n. \quad (6)$$

Thus, the diversity of a 'minimal' immune system is a linear function of the number of self epitopes. Substituting our empirical estimates for mice, i.e.  $P_E = 0.01$ , and  $10^4 \leq n \leq 10^5$ , we obtain  $5 \times 10^4 \leq \bar{R} \leq 5 \times 10^5$ . This is 100–1000-fold lower than empirical estimates (Du Pasquier 1982; Berek & Milstein 1988; Davis & Bjorkman 1988).

#### 4. T CELL REPERTOIRE

Generically, T cells respond to peptides presented on MHC molecules. We assume that, on average, each antigen is degraded into  $\rho$  different peptides. Let  $m$  be the number of different MHC molecules, and  $P_p$  the probability that a given peptide is presented by a given MHC molecule. As each peptide has a probability  $P_p$  to get presented by one of the  $m$  MHC molecules, each antigen is expected to be presented as

$$\epsilon = \rho P_p m \quad (7)$$

different MHC-peptide complexes. We estimate  $m = 4$  because for each class of MHC molecules there are four major alleles expressed per individual. In combination with the empirical estimates  $\rho = 10$  and  $P_p = 0.2$  (Schaeffer *et al.* 1989) we arrive at an order of magnitude of  $\epsilon = 10$ . With  $\epsilon$  as defined above, equation (1) now can be applied to the T cell repertoire. Agreement with the empirical estimate of 50–70% deletion by tolerance (Scollay & Shortman 1985) requires that  $P_R n \epsilon$  should be of order one. Hence receptor specificity,  $P_R$ , should be the inverse of the number of self epitopes. This is in agreement with equation (4).

Consider the probability of escape from T cell recognition. For T cells, recognition involves: (i) the presentation of peptides by MHC molecules; and (ii) the recognition of peptide-MHC molecules by T cell receptors. For a pathogen consisting of  $a$  antigens, each consisting of  $\rho$  peptides, the probability of escape is

$$P_E^a = [1 - P_p + P_p(1 - P_R)^R]^{m\rho a}, \quad (8)$$

where the term  $1 - P_p$  corresponds to a failure of presentation, and the term  $P_p(1 - P_R)^R$  corresponds to

a failure of T cell recognition given that presentation has occurred. (This expression can be derived via a binomial expression for the number of MHC molecules that may present each peptide.) Simplifying, we obtain

$$\begin{aligned} \ln P_E &\simeq m\rho \ln\{1 - P_p[1 - \exp(-P_R R)]\} \\ &\simeq -\epsilon[1 - \exp(-P_R R)]. \end{aligned} \quad (9)$$

This expression allows us to determine  $\bar{m}$ , the minimal number of MHC molecules required for recognizing foreign antigens at probability of escape level  $P_E$ . Letting  $R \rightarrow \infty$  in equation (9), i.e. assuming that any presented peptide will be recognized, we obtain  $-\ln P_E \simeq \bar{\epsilon}$  or

$$\bar{m} \simeq -(\ln P_E)/\rho P_p. \quad (10)$$

For the typical estimates we used above, i.e.  $P_E = 10^{-2}$ ,  $\rho = 10$ ,  $P_p = 0.2$ , this gives  $\bar{m} \simeq 3$ , which is the right order of magnitude.

To determine the pre-immune T cell repertoire, we substitute equation (1) into equation (9), obtaining

$$R_0 \simeq -\ln[1 + (\ln P_E)/\epsilon] \frac{\exp(P_R n \epsilon)}{P_R}. \quad (11)$$

The main difference between this equation and its B cell counterpart, equation (3), is that for it to be valid  $m$  must be greater than  $\bar{m}$ , i.e. the argument of the first logarithmic term must be positive.

Equation (11) is plotted in figure 1*b-d* as a function of  $n$ ,  $\epsilon$  and  $P_R$ . Its minimum, found by setting  $\partial R_0/\partial P_R = 0$ , is also at  $\bar{P}_R = 1/n\epsilon$ , implying that in minimal repertoires antibodies and T cell receptors should have similar specificity. From equation (1a), the fraction,  $\bar{f}$ , of non-deleted T cells in the minimal repertoire is the same as that of the B cells, i.e.  $\bar{f} = \exp[-1]$ . This corresponds closely to empirical data showing that 50–70% of the T cells die in the thymus (Scollay & Shortman 1985).

Substituting equation (4) into equation (11) we obtain for the minimal pre-immune T cell repertoire

$$\bar{R}_0 = -e \ln[1 + (\ln P_E)/\epsilon] n \epsilon, \quad (12)$$

which, for fixed  $\epsilon$  and  $P_E$ , is again a linear function of  $n$ , the number of self antigens. For the minimal functional T cell repertoire, equation (2), we obtain

$$\bar{R} = \bar{R}_0 \bar{f} \simeq -\ln[1 + (\ln P_E)/\epsilon] n \epsilon. \quad (13)$$

When  $\epsilon$  is sufficiently large, we can approximate  $\ln[1 + (\ln P_E)/\epsilon]$  by  $\ln(P_E)/\epsilon$  to obtain equation (6). Thus, in the limit  $\epsilon \rightarrow \infty$ , the expected size of the T cell repertoire approaches that of the B cell repertoire and becomes independent of the number of MHC molecules. This limit is probably irrelevant because MHC molecules are known to strongly influence immune reactions. Equation (13) does not have a minimum as a function of  $\epsilon$ .

#### 5. DISCUSSION

Our main result is that the diversity of the immune system principally reflects the diversity of self antigens rather than the diversity of foreign antigens. This result

has been obtained for both B and T cell repertoires when we calculated the repertoire required for having a 'sufficiently good' immune system, i.e. equations (3) and (11). Allowing for evolutionary optimization of receptor specificities, equation (5) and (12), did not affect this result.

Another recent attempt (Percus *et al.* 1993) at modelling the specificity of the immune system is based upon optimizing the probability that the immune system evolves a repertoire that is capable of recognizing all foreign antigens but not capable of recognizing any self antigen. Although the approach differs significantly from the one outlined here in that deletion of self-reactive clones is not considered, the authors also reach the conclusion that recognition of foreign antigens can be achieved by small repertoires, and that it is the requirement of not reacting with self that requires repertoire diversity.

T cell activation by peptides presented on antigen-presenting cells involves competition amongst the T cells for binding antigen. We have recently developed a model which demonstrates that, for each antigen, this competition process leads to the exclusion of all but one of the T cell clones (De Boer & Perelson 1993). This result becomes relevant for the main conclusion of this paper if we accept that the T cell repertoire is largely maintained by cross-reactivities with self antigens (Rock & Benacerraf 1984; Stutman 1986): competitive exclusion implies that each self antigen can maintain the proliferation of one T cell specificity only.

Our model of the T cell repertoire shows that to recognize antigen reliably there is a lower bound to the number of different MHC molecules an individual should have. The model also suggests that if the immune system operates with T cell receptors having optimal specificity, then the system can accommodate any number of MHC molecules even though negative selection is operating. Other models (Nowak *et al.* 1992; Celada & Seiden 1992) have suggested that, if there are too many MHC molecules, negative selection would eliminate too much of the pre-immune repertoire, and thus concluded that there is an optimum number of MHC molecules. This is also true in our model if we assume that the repertoire size has not been minimized by evolution. Thus in figure 1*d* we find an optimal value of  $\epsilon$ . However, it is not true for the 'minimal' repertoire of equation (13). This discrepancy depends on which variable(s), i.e.  $m$ ,  $P_R$ , or  $m$  and  $P_R$ , one subjects to evolutionary optimization.

We have approached the diversity of the immune system only from the viewpoint of the probability of recognizing pathogens. Recognition is not sufficient for antigen elimination. Effector functions and memory (i.e. immunity) are also important aspects of the adaptive immune response. Interestingly, in idiotypic network theory (Weisbuch 1990), it has been argued that the number of memories that can be encoded in an immune network is of the order of  $\sqrt{R}$ . The ability to have immunity to  $10^4$  different pathogens would then require a functional repertoire of  $10^8$  specificities, which is higher than the results reported here.

The crucial test for the evolutionary part of our hypothesis would be a comparative study determining

the repertoire size as a function of the number of self antigens for a large number of species. Equations (6) and (12) predict a linear relation.

Unfortunately, the empirical estimates for the number of self antigens are very approximate. The estimate of  $10^4 \leq n \leq 10^5$  (Cohn 1987) is based upon the size of the genome. Such an estimate could be wrong by several orders of magnitude. First, the combinatorics of forming different glycoproteins from the  $10^4$ – $10^5$  encoded proteins with the carbohydrates of an organism may give rise to an astronomically large number of different self antigens. Because immunoglobulins generically recognize conformational antigens this would be especially true for the B cell self epitopes. Second, for T cell epitopes, it has been argued that tolerance only involves a subset of the epitopes (Gammon *et al.* 1991). The authors make a distinction between the 'dominant' and the 'cryptic' self. As there might only be a limited number of dominant epitopes for each tissue of an organism, the total number of T cell self epitopes could be orders of magnitude smaller than the number of self proteins. Thus, on the basis of our theory, we could speculate that the B cell repertoire should be larger than the T cell repertoire, or that self tolerance for B cells should be incomplete.

Additionally, precise data do not exist on the size of the immune repertoire, and on the specificity of lymphocytes for a variety of species. It is known that lower vertebrates such as fish, amphibians and reptiles have a repertoire that is much smaller than that of mammals. The antibody diversity of anuran amphibians such as *Rana* and *Xenopus* is estimated between  $5 \times 10^4$  and  $5 \times 10^5$  specificities (Du Pasquier 1982). We think that this lack of data calls for a more 'comparative approach' investigating the immune systems of a large variety of species. The prediction of our theory would be that lower vertebrates have a much lower number of (dominant) self epitopes.

We thank Dr Martin A. Nowak for making his unpublished manuscript available to us. We thank Professor Niels Jerne, Professor Paulien Hogeweg, Dr Martin A. Nowak and Dr Ton Logtenberg for helpful comments on earlier versions of the manuscript. This work was done under the auspices of the U.S. Department of Energy. It was supported in part by NIH Grants AI28433 and RR06555. It was also supported by the Santa Fe Institute through their Theoretical Immunology Program and the Center for Nonlinear Studies at Los Alamos.

## REFERENCES

- Avrameas, S. 1991 Natural autoantibodies: from 'horror autotoxicus' to 'gnothi seauton'. *Immunol. Today* **12**, 154–159.
- Berek, C. & Milstein, C. 1988 The dynamics nature of the antibody repertoire. *Immunol. Rev.* **105**, 5–26.
- Bona, C. A. & Kaushik, A. K. (ed.) 1992 *Molecular immunology of self reactivity*. New York: Marcel Dekker.
- Cancro, M. M., Gerhard, W. & Klinman, N. R. 1978 The diversity of the influenza-specific primary B-cell repertoire in Balb/c mice. *J. exp. Med.* **147**, 776–786.
- Celada, F. & Seiden, P. 1992 A computer model of cellular interactions in the immune system. *Immunol. Today* **13**, 56–63.

- Cohn, M. 1987 The ground rules determining any solution to the problem of the self/nonself discrimination. In *Proceedings of the EMBO Workshop on Tolerance* (ed. P. Matzinger, M. Flajnik, H. G. Rammensee, R. Rolink & B. Stockinger), pp. 3–35. Basel: Roche.
- Davis, M. M. & Bjorkman, P. K. 1988 T-cell receptor antigen genes and T-cell recognition. *Nature, Lond.* **334**, 395–402.
- De Boer, R. J. & Perelson, A. S. T cell repertoires and competitive exclusion. (Submitted.)
- Deenen, G. J., Van Balen, I. & Opstelten, D. 1990 In rat B lymphocyte genesis sixty percent is lost from the bone marrow at the transition of non dividing pre-B cell to sIgM<sup>+</sup> B lymphocyte, the state of Ig light chain expression. *Eur. J. Immunol.* **20**, 557–564.
- Du Pasquier, L. 1982 Antibody diversity in lower vertebrates – why is it so restricted? *Nature, Lond.* **296**, 311–313.
- Gammon, G., Sercarz, E. E. & Benichou, G. 1991 The dominant self and the cryptic self: shaping the autoreactive T-cell repertoire. *Immunol. Today* **12**, 193–195.
- Goodnow, C. C. 1992 Transgenic mice and the analysis of B-cell tolerance. *A. Rev. Immunol.* **10**, 489–518.
- Hartley, S. B., Crosbie, J., Brink, R., Kantor, A. B., Basten, A., Goodnow, C. C. 1991 Elimination from peripheral lymphoid tissue of self-reactive B lymphocytes recognizing membrane-bound antigens. *Nature, Lond.* **353**, 765–769.
- Kappler, J. W., Roehm, N. & Marrack, P. 1987 T cell tolerance by clonal elimination. *Cell* **49**, 273–280.
- Kappler, J. W., Staerz, U., White, J. & Marrack, P. C. 1988. Self-tolerance eliminates T cells specific for Mls-modified products of the major histocompatibility complex. *Nature, Lond.* **332**, 35–40.
- Kroemer, G. & Martinez-A., C. 1992 Mechanisms of self tolerance. *Immunol. Today* **13**, 401–404.
- MacDonald, H. R., Schneider, R., Lees, R. K., Howe, R. C., Acha-Orbea, H., Festenstein, H., Zinkernagel, R. M. & Hengartner, H. 1988 T-cell receptor  $V_{\beta}$  use predicts reactivity and tolerance to Mls<sup>d</sup>-encoded antigens. *Nature, Lond.* **332**, 40–45.
- Milon, G., Titus, R. G., Cerottini, J. C., Marchal, G. & Louis, J. A. 1986 Higher frequency of *Leishmania major*-specific L3T4<sup>+</sup> T cells in susceptible Balb/c as compared with resistant CBA mice. *J. Immunol.* **136**, 1467–1471.
- Mueller, D. L., Jenkins, M. K. & Schwartz, R. H. 1989 Clonal expansion versus functional clonal inactivation: a costimulatory signalling pathway determines the outcome of T cell antigen receptor occupancy. *A. Rev. Immunol.* **7**, 445–480.
- Nossal, G. J. V. 1983 Cellular mechanisms of immunologic tolerance. *A. Rev. Immunol.* **1**, 33–62.
- Nowak, M. A., Tarczy-Hornoch, K. & Austyn, J. M. 1992 The optimal number of MHC molecules in an individual. *Proc. natn. Acad. Sci. U.S.A.* **89**, 10896–10899.
- Percus, J. K., Percus, O. E. & Perelson, A. S. 1993 Predicting the size of the T cell receptor and antibody combining region from consideration of efficient self-nonself discrimination. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1691–1695.
- Perelson, A. S. & Oster, G. 1979 Theoretical studies on clonal selection: minimal antibody repertoire size and reliability of self non-self discrimination. *J. theor. Biol.* **81**, 645–670.
- Rock, K. L. & Benacerraf, B. 1984 Thymic T cells are driven to expand upon interaction with self-class II histocompatibility complex gene products on accessory cells. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1221–1224.
- Shaeffer, E. B., Sette, A., Johnson, D. L., Bekoff, M. C., Smith, J. A., Grey, H. M. & Buus, S. 1989 Relative contribution of “determinant selection” and “holes in the T-cell repertoire” to T-cell responses. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4649–4653.
- Schwarz, R. H. 1989 Acquisition of immunologic self-tolerance. *Cell* **57**, 1073–1081.
- Scollay, R. & Shortman, K. 1985 Cell traffic in the adult thymus. In *Recognition and regulation in cell mediated immunity* (ed. J. D. Watson & J. Marbock), pp. 3–29. New York: Marcel Dekker.
- Stutman, O. 1986 Post-thymic T-cell development. *Immunol. Rev.* **91**, 159–194.
- Weisbuch, G. 1990 A shape space approach to the dynamics of the immune system. *J. theor. Biol.* **143**, 507–522.

Received 1 February 1993; accepted 1 April 1993