Immunological Memory is Associative

Derek J. Smith
Department of Computer Science
University of New Mexico
Albuquerque, NM 87131, USA
dsmith@cs.unm.edu
Tel: (505) 277-3112
Fax: (505) 277-6927

Stephanie Forrest
Department of Computer Science
University of New Mexico
Albuquerque, NM 87131, USA
forrest@cs.unm.edu
Tel: (505) 277-3112
Fax: (505) 277-6927

Alan S. Perelson
Theoretical Division
Los Alamos National Laboratory
Los Alamos, NM 87545, USA
asp@t10.lanl.gov

Abstract

This paper argues that immunological memory is in the same class of associative memories as Kanerva's *Sparse Distributed Memory*, Albus's *Cerebellar Model Arithmetic Computer*, and Marr's *Theory of the Cerebellar Cortex*. This class of memories derives its associative and robust nature from a sparse sampling of a huge input space by recognition units (B and T cells in the immune system) and a distribution of the memory among many independent units (B and T cells in the memory population in the immune system).

Keywords: Immunological Memory, Associative Memory, Cross-Reactive Memory, Original Antigenic Sin, Sparse Distributed Memory.

1 Introduction

Cowpox vaccination, used to protect humans from smallpox, was the first known deliberate use of associative recall in the immune response (Jenner, 1798). The modern investigation of associative recall began with the observation that antibodies induced during an influenza infection often have greater affinity to prior strains of influenza than to the infecting strain—suggesting that the antibodies were generated by memory cells to prior infections (Francis, 1953; Davenport *et al.*, 1953). Gilden (1963) and Fazekas de St. Groth & Webster (1966) continued the investigation by injecting laboratory animals with one antigen and then recalling the memory of that antigen by subsequent injection of a second, related, antigen. Some researchers considered associative recall "a degeneracy in the secondary immune response" (Eisen *et al.*, 1969). However, as research continued, and associative recall was observed in many animal models with many types of antigen, it became clear that it was a general phenomenon of immunological memory (Ivanyi, 1972; Deutsch

& Bussard, 1972; Fish *et al.*, 1989; Nara & Goudsmit, 1990; Smith, 1994). Immunologists refer to associative recall as a cross-reactive secondary response, or as original antigenic sin (Fazekas de St. Groth & Webster, 1966).

The purpose of this paper is to show that immunological memory is an associative and robust memory that belongs to the class of sparse distributed memories. This class of memories derives its associative and robust nature by sparsely sampling the input space and distributing the data among many independent agents (Kanerva, 1992). Other members of this class include a model of the cerebellar cortex (Marr, 1969), the *Cerebellar Model Arithmetic Computer* (CMAC) (Albus, 1981), and *Sparse Distributed Memory* (SDM) (Kanerva, 1988). First, we present a simplified account of the immune response and immunological memory. Next, we present SDM, and then we show the correlations between immunological memory and SDM. Finally, we show how associative recall in the immune response can be both beneficial and detrimental to the fitness of an individual.

2 Immunological Memory

The immune system must recognize a large number of cells and molecules (antigens) that it has never seen before, and it must decide how to respond to them. Some antigens, such as those that make up the individual's body, must not be attacked[1]. Other antigens, such as viruses, bacteria, parasites and toxins are responded to by mixtures of the T cell response and the different types of antibody responses. The immune system remembers antigens it has seen before and when it sees them again is often capable of eliminating them before disease occurs. This memory is the basis for vaccination, and the reason why we do not get most diseases more than once. The immune response and immunological memory are complex and not fully understood. This exposition is necessarily simplified and restricted.

Recognition. The human immune system uses a large number of highly specific B and T cells to recognize antigen. An individual has the genetic material and a combinatorial based randomizing mechanism to express $>10^{10}$ distinct B cell receptors (Berek & Milstein, 1988). At any time, the immune system expresses a subset of these consisting of the order of 10^7 to 10^8 distinct B cell receptors (Köhler, 1976; Klinman *et al.*, 1976; Klinman *et al.*, 1977). The number of possible distinct antigens is difficult to calculate, but it is thought to be in the range 10^{12} to 10^{16} (Inman, 1978). B and T cell receptors are stimulated by antigen if their affinity for the antigen is above some threshold. Typically 10^{-5} to 10^{-4} of an individual's B cells are stimulated by an antigen (Edelman, 1974; Nossal & Ada, 1971; Jerne, 1974a). B cells that are stimulated by an antigen are said to be in the *ball of stimulation* of that antigen (Perelson & Oster, 1979) (Figure 1a).

Response. B and T cells that are stimulated by antigen divide. The B cell receptor sometimes mutates on cell division and this can increase the affinity of its daughter cells for the antigen. Towards the end of a response, when antigen becomes scarce, higher affinity B cells have a fitness advantage over lower affinity B cells and are preferentially selected in a process similar to natural evolution (Figure 1b) (Burnet, 1959). During the replication of cells in response to antigen, some B cells change into plasma cells and secrete antibodies which eliminate the antigen. In the case

[1]Diseases such as multiple sclerosis, rheumatiod arthritis and insulin-resistant diabetes are examples of autoimmune diseases where the immune system attacks the body it usually protects.

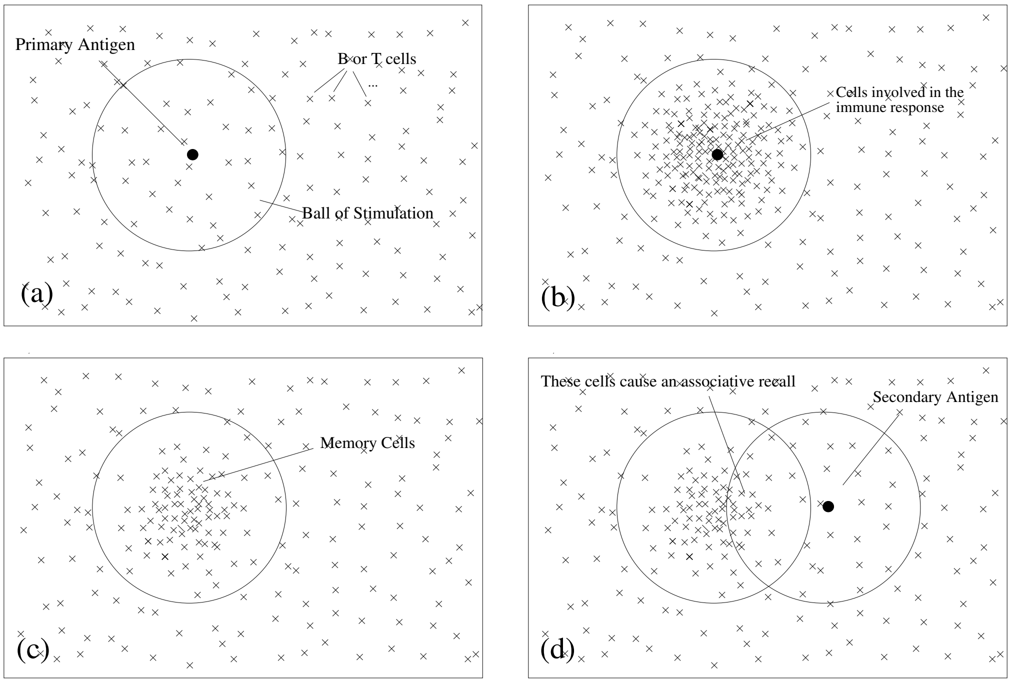

Figure 1: (a) A two dimensional illustration of (a high dimensional) sparse distribution of B or T cell receptors in a space where distance is a measure of affinity for an antigen. B and T cells within some threshold affinity bind the antigen and become activated. The region the antigen activates is called the *ball of stimulation* of the antigen. (b) Activated B cells replicate and mutate and the higher affinity ones are selected. (c) After the antigen is cleared a memory population persists. (d) A second exposure to the same antigen, or here a related antigen, restimulates the memory population inducing an associative recall.

of a viral infection, some T cells change into cytotoxic T lymphocytes (CTLs), which can kill virus-infected cells.

At the end of an immune response, when the antigen is cleared, the B cell population decreases, leaving a persistent sub-population of memory cells (Figure 1c). The mechanism(s) by which memory cells persist is not fully understood. One theory is that memory cells live for a long time (Mackay, 1993). Another is that memory cells are restimulated at some low level. A number of mechanisms for restimulation have been proposed. Jerne (1974) proposed the idiotypic network theory in which cells co-stimulate each other in a way that mimics the presence of the antigen. Another theory is that small amounts of the antigen are retained in lymph nodes (Tew & Mandel, 1979; Tew *et al.*, 1980). Another is that related environmental antigens provide cross-stimulation (Matzinger, 1994).

The idiotypic network theory has been proposed as a central aspect of the associative properties of immunological memory (Farmer *et al.*, 1986; Gibert & Routen, 1994). However, in this exposition, it is but one of the possible mechanisms that maintains the memory population and is key to neither the associative nor robust properties of immunological memory.

The persistent population of memory cells is the mechanism by which the immune system remembers. If the same antigen is seen again, the memory population quickly produces large quantities of antibodies (or CTLs) and often times the antigen is cleared before it causes disease. This is called a secondary immune response. If the secondary antigen is slightly different from the primary antigen, its ball of stimulation may overlap part of the memory population raised by the primary antigen (Figure 1d). Memory cells in the overlap bind the antigen and produce antibodies and/or CTLs. Immunologists call this a cross-reactive response, people who work with associative memories would call it associative recall. The strength of the secondary immune response is approximately proportional to the number of memory cells in the ball of stimulation of the antigen. If a subset of the memory population is stimulated, by a related antigen, then the response is weaker (East *et al.*, 1980). Thus, memory appears to be distributed among the cells in the memory population. Immunological memory is robust because, even when a portion of the memory population is lost, the remaining memory cells persist to produce a response.

3 Sparse Distributed Memory (SDM)

Kanerva's SDM is a member of the class of sparse and distributed associative memories to which we show immunological memory belongs. SDM, like random access memory (the memory in a computer), is written to by providing an address and data, and read from by providing an address and getting an output. Unlike random access memory, the address space of SDM is enormous, sometimes 1,000 bits, giving $2^{1,000}$ possible addresses. SDM cannot instantiate such a large number of address-data locations so it instantiates a subset, of say 1,000,000 address-data locations. These instantiated address-data locations are called hard locations and are said to sparsely cover the input space (Kanerva, 1988).

When an address is presented to the memory hard locations that are within some threshold Hamming distance of the address are activated. This subset of activated hard locations is called the access circle of the address (Figure 2a). On a write, each bit of the input data is stored independently in a counter in each hard location in the access circle. If the ith data bit is a 1, the ith counter in each hard location is incremented by 1, if the ith data bit is a 0 the counter is decremented by 1 (Figure 2b). On a read, each bit of the output is composed independently of the other bits. The value of the ith counter of each of the hard locations in the access circle are summed. If the sum is positive the output for the ith bit is a 1, if the sum is negative the output is a 0 (Figure 2c).

The distribution of the data among many hard locations makes the memory robust to the loss of some hard locations, and it permits associative recall of the data if a read address is slightly different from a prior write address. If the access circle of the read address overlaps the access circle of the write address, then hard locations that were written to by the write address are activated by the read address, and give an associative recall of the write data (Figure 2d).

4 Correspondence between Immunological Memory and SDM

The previous two sections showed that both immunological memory and SDM are sparse and distributed, and are thus associative and robust. The correspondence between the two memories is

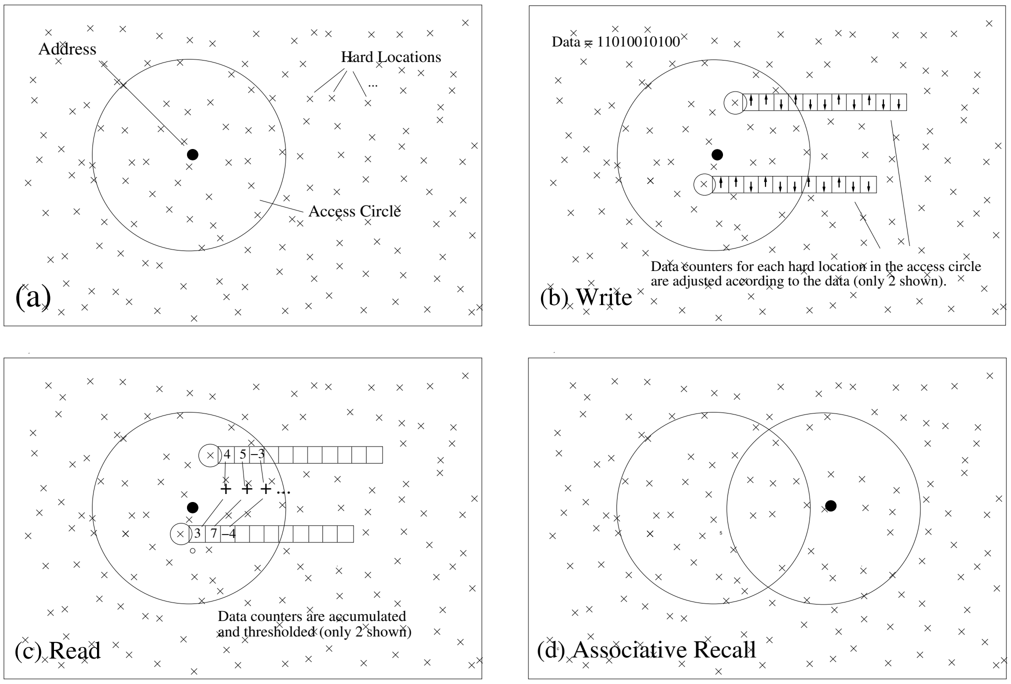

Figure 2: (a) A two dimensional illustration of hard locations sparsely and randomly distributed in a high dimensional binary input space. Distance in the space represents Hamming distance between hard locations. Hard locations within some radius of an input address are activated and form an access circle. (b) Activated hard locations adjust their associated data counters on a write and (c) accumulate and threshold them on a read. (d) The access circle of a similar address might include some hard locations from the prior write and induce an associative recall.

summarized in Table 1 and discussed below.

Both SDM and immunological memory use detectors to recognize input. In the case of SDM, hard locations recognize an address, in the case of immunological memory, B and T cells recognize an antigen. In both systems the number of possible distinct inputs is huge and due to resource limitations, the number of detectors is much smaller than the number of possible inputs—both systems sparsely cover the input space. In order to respond to all possible inputs, detectors in both systems do not require an exact match with the input, but are activated if they are within some threshold distance of the input. In the case of SDM this is Hamming distance, in the case of immunological memory it is affinity. Thus, in both systems an input activates a subset of detectors. In SDM, this subset is called the access circle of the input address, in immunological memory, it is called the ball of stimulation of the antigen.

In both systems detectors store information associated with each input. In the case of SDM the information is a bit string and is supplied exogenously. In the case of immunological memory the information is determined by mechanisms within the immune system that determine whether to respond to an antigen and if so with what class of antibody and/or CTL response.

Immunological Memory	SDM
Antigen	Address
B/T Cell	Hard Location
Ball of Stimulation	Access Circle
Affinity	Hamming Distance
Response/Tolerance	Data
Primary Response	Write and Read
Secondary Response	Read
Cross-Reactive Response	Associative Recall

Table 1: Structural and functional correspondence between immunological memory and SDM.

Both systems distribute all of the information to each activated detector. In the case of SDM the data is used to adjust counters, in the case of immunological memory a large number of memory cells are created which, in the case of B cells, have undergone genetic reconfigurations that determine the class of antibodies they will produce. Because all the information is stored in each detector, each detector can recall the information independently of the other detectors. The strength of the output, the signal, is an accumulation of the information in each activated detector. Thus, as detectors fail, the output signal degrades gracefully, and the signal strength is proportional to the number of activated detectors. Thus, the distributed nature of storing the information in both systems makes both systems robust to the failure of individual detectors.

In the case of SDM, the data is distributed to hard locations, each of which have different addresses, and in the case of immunological memory, to cells which have different receptors (although many copies of a cell with the same, or similar receptors also exist). If the activated subset of detectors of a related input (a noisy address in the case of SDM, or a mutant strain in the case of immunological memory) overlap the activated detectors of a prior input (Figures 1d and 2d), detectors from the prior input will contribute to the output. Such associative recall, and the graceful degradation of the signal as inputs differ, is due to the distribution of the data among the activated detectors.

5 Aspects of Associative Recall in the Immune Response

Prior exposure to one strain of a pathogen often protects against mutant strains. In influenza, for example, older individuals have greater immunity to new strains than younger individuals (Francis, 1953). In this case associative recall is beneficial to the individual.

Associative recall can also be a disadvantage. If the ball of stimulation of a vaccine (secondary antigen) overlaps the memory to a prior infection (primary antigen), then the vaccine may be cleared by associative recall of the prior infection and fail to make highly specific memory to the vaccine strain. If the ball of stimulation of a subsequent challenge (tertiary antigen) also overlaps prior memory then all is well (Figure 3a). However, if there is no overlap, disease can ensue and the vaccine will have "failed" (Figure 3b).

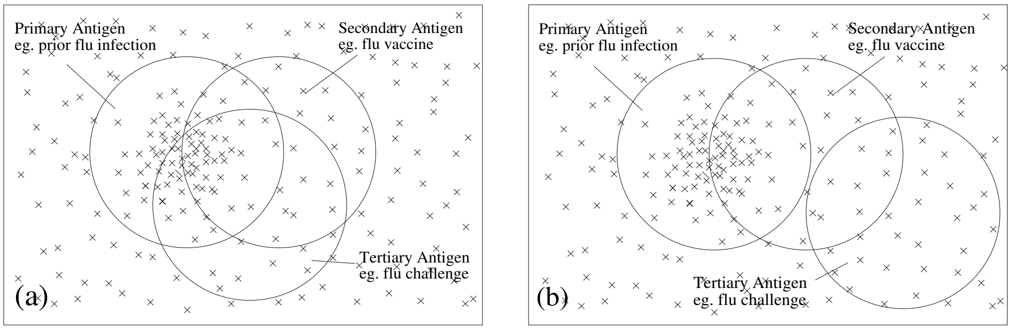

Figure 3: Associative recall of prior infection (primary antigen) can cause a vaccine (secondary antigen) to fail. The prior infection might divert the vaccine and (a) also react with the challenge strain (tertiary antigen), or (b) be of no use against the challenge strain.

A potentially confusing situation for any associative memory is when two similar inputs require different outputs. In the case of immunological memory this would occur when an antigen requires one type of response and a related antigen a different type of response. Such a situation has been hypothesized in the case of some malaria infections, where prior exposure, to a related environmental antigen, is thought to divert the response (Good *et al.*, 1993).

6 Summary

We have shown the correspondence between B and T cells in the immune system and hard locations in a SDM. In particular B and T cells perform a sparse coverage of all possible antigens in the same way that hard locations perform a sparse coverage of all possible addresses in a SDM. Also, data are distributed among many independent B and T cells in immunological memory as they are among many independent hard locations in a SDM. Immunological memory is thus a member of the family of sparse distributed memories, and its associative and robust properties are due precisely to its sparse and distributed nature.

Acknowledgments

DJS gratefully acknowledges the support of the Santa Fe Institute, the University of New Mexico Computer Science Department AI fellowship and Digital Equipment Fellowship. DJS also thanks David Ackley, Will Attfield, Ron Hightower, Terry Jones, Pentti Kanerva, Linda Cicarella, Ron Moore, Francesca Shrady and Paul Stanford for helpful comments on this work. Portions of this work were performed under the auspices of the U.S. Department of Energy. This work was supported by NIH (AI28433, RR06555), the Joseph P. and Jeanne M. Sullivan Foundation, ONR (N00014-95-1-0364), and NSF (IRI-9157644). The authors gratefully acknowledge the ongoing support of the Santa Fe Institute.

REFERENCES

Albus, J. S. (1981). *Brains, Behavior, and Robotics*. Byte Books, Peterborough, NH.

Berek, C. & Milstein, C. (1988). The dynamic nature of the antibody repertoire. *Immunol. Rev.* **105**, 5–26.

Burnet, F. M. (1959). *The Clonal Selection Theory of Immunity*. Cambridge Univerity Press, London.

Davenport, F. M., Hennessy, A. V. & Francis, T. (1953). Epidemiologic and immunologic significance of age distribution of antibody to antigenic variants of influenza virus. *J. Exp. Med.* **98**, 641–656.

Deutsch, S. & Bussard, A. E. (1972). Original antigenic sin at the cellular level. i. antibodies produced by individual cells against cross-reacting haptens. *Eur. J. Immunol.* **2**(4), 374–378.

East, I. J., Todd, P. E. & Leach, S. J. (1980). Original antigenic sin: Experiments with a defined antigen. *Mol. Immunol.* **17**(12), 1539–1544.

Edelman, G. M. (1974). In: *Cellular Selection and Regulation in the Immune System* (Edelman, G. M., ed.) page 1. Raven Press, New York.

Eisen, H. N., Little, J. R., Steiner, L. A., Simms, E. S. & Gray, W. (1969). Degeneracy in the secondary immune response: Stimulation of antibody formation by cross-reacting antigens. *Israel J. Med. Sci.* **5**(3), 338–351.

Farmer, J. D., Packard, N. H. & Perelson, A. S. (1986). The immune system, adaptation, and machine learning. *Physica D* **22**, 187–204.

Fazekas de St. Groth, S. & Webster, R. G. (1966). Disquisitions of original antigenic sin. ii. proof in lower creatures. *J. Exp. Med.* **124**(3), 347–361.

Fish, S., Zenowich, E., Fleming, M. & Manser, T. (1989). Molecular analysis of original antigenic sin. i. clonal selection, somatic mutation, and isotype switching during a memory b cell response. *J. Exp. Med.* **170**(4), 1191–1209.

Francis, T. (1953). Influenza, the newe acquayantance. *Annals of Internal Medicine* **39**(2), 203–221.

Gibert, C. J. & Routen, T. W. (1994). Associative memory in an immune-based system. In: *Proceedings of the Twelfth National Conference on Artificial Intelligence* pp. 852–857, Cambridge, MA. AAAI Press/The MIT Press.

Gilden, R. V. (1963). Antibody responses after successive injections of related antigen. *Immunology* **6**(30), 30–36.

Good, M. F., Zevering, Y., Currier, J. & Bilsborough, J. (1993). 'original antigenic sin', t cell memory, and malaria sporozoite immunity: an hypothesis for immune evasion. *Parasite Immunology* **15**(4), 187–193.

Inman, J. K. (1978). In: *Theoretical Immunology* (Bell, G., Perelson, A. & Pimberly, G., eds) page 243. Marcel Decker, Inc., New York.

Ivanyi, J. (1972). Recall of antibody synthesis to the primary antigen following successive immunization with heterologous albumins. a two-cell theory of the original antigenic sin. *Eur. J. Immunol.* **2**(4), 354–359.

Jenner, E. (1798). *An Inquiry into the Causes and Effects of the Variolae Vaccinae*. Low, London.

Jerne, N. K. (1974). In: *Cellular Selection and Regulation in the Immune System* (Edelman, G. M., ed.) page 39. Raven Press, New York.

Jerne, N. K. (1974). Towards a network theory of the immune system. *Annals of Immunology (Institute Pasteur)* **125C**, 373.

Kanerva, P. (1988). *Sparse Distributed Memory*. MIT Press, Cambridge, MA.

Kanerva, P. (1992). Sparse distributed memory and related models. In: *Associative Neural Memories: Theory and Implementation* (Hassoun, H. M., ed.) chapter 3. Oxford University Press.

Klinman, N. R., Press, J. L., Sigal, N. H. & Gerhart, P. J. (1976). In: *The Generation of Antibody Diversity* (Cunningham, A. J., ed.) page 127. Academic Press, New York.

Klinman, N. R., Sigal, N. H., Metcalf, E. S., Gerhart, P. J. & Pierce, S. K. (1977). *Cold Spring Harbor Symp. Quant. Biol.* **41**, 165.

Köhler, G. (1976). Frequency of precursor cells against the enzyme beta-galactosidase: an estimate of the balb/c strain antibody repertoire. *Eur. J. Immunol.* **6**(5), 340–347.

Mackay, C. R. (1993). Immunological memory. *Adv. Immunol.* **53**, 217–265.

Marr, D. (1969). A theory of cerebellar cortex. *J. Physiology* **202**, 437–470.

Matzinger, P. (1994). Immunological memories are made of this? *Nature* **369**(6482), 605–6.

Nara, P. L. & Goudsmit, J. (1990). Clonal dominance of the neutralizing response to the hiv-1 v3 epitope; evidence for original antigenic sin during vaccination and infection in animals including humans. In: *Vaccines (Cold Spring Harbor Eighth Annual Meeting), Vol. 91, Modern Approaches to New Vaccines including Prevention of AIDS* (et al, R. M. C., ed.). Cold Spring Harbor, NY.

Nossal, C. J. V. & Ada, G. L. (1971). *Antigens, Lymphoid Cells and The Immune Response*. Academic Press, New York.

Perelson, A. S. & Oster, G. F. (1979). Theoretical studies of clonal selection: Minimal antibody repertoire size and reliability of self- non-self discrimination. *J. Theoret. Biol.* **81**, 645–670.

Smith, D. J. (1994). A literature review of original antigenic sin. Technical Report TR 94–10, University of New Mexico, Albuquerque, NM.

Tew, J. G. & Mandel, T. E. (1979). Prolonged antigen half–life in the lymphoid follicules of antigen–specifically immunized mice. *Immunology* **37**, 69–76.

Tew, J. G., Phipps, P. R. & Mandel, T. E. (1980). The maintenance and regulation of the humoral immune response. persisting antigen and the role of follicular antigen–binding dendritic cells. *Immunol. Rev.* **53**, 175–211.