Emergent Networks in Immune System Shape Space

A thesis
submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Computing
in the School of Computing,
Dublin City University,

DCU

John Burns, B.Sc.
Supervisor: Prof. Heather J. Ruskin

March 1, 2005

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Ph.D. in Computing is entirely my own and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____

Candidate

ID No.: 5117336 I

Date: 2/3/2005

Acknowledgment

No goal worth attaining in life is ever likely to be easy to achieve. Neither is it likely that such an achievement is the result solely of individual merit. Rather, behind every person and their success, one nearly always finds a pattern of opportunity and support. In my case, I have been fortunate in my life to have had both, and in large measure.

In reaching the point of submitting an approved dissertation for the award of the degree of Ph.D., it is always appropriate to acknowledge the contribution of others. Pre-eminent among them is my supervisor and mentor, Professor Heather J. Ruskin. Heather's unflagging support, advice and encouragement have helped turn this work from an interesting idea into a proposition with scientific merit. Of equal importance, Heather's rigorous assessment of my work over the years has resulted in my learning something of the *scientific method*. Her preference for me to write regularly and submit for journal and conferences, in order to test and defend theories, has, I feel, helped me get my work to this point, even if I do get a little defensive sometimes!

To wife, my parents and my family, I extend my love. They have always supported me. Having someone to share the ups and downs with has been a crucial factor for me, and in no small part do I owe the success of this achievement to them. I should also mention Ruili Wang, a man of originality and insight. Ruili always advised me well in the early days of this work. In particular, we discussed many times his pyramid theory of scientific research, a novel idea I hope he will publish someday.

Finally, and perhaps rather unusually, I extend my thanks to Dublin City University as a whole. Having given me the opportunity to enter the university directly (as a mature student) in 1991, I have always regarded DCU with affection and appreciation. Thanks to John Murphy and Merve Barmann who interviewed me and accepted my application on to the B.Sc. in Computer Applications. Every journey starts with a first step: my first step on the journey to the degree of Ph.D. was taken in a cramped office,

one damp day in April 1991. John sat behind the desk, Merve sat *on* the desk, and the three of us talked about sorting algorithms [1]. An auspicious start indeed.

Thank you all, so very much.

John Burns, Dublin, 1/3/2005.

[1] I believe we came out in favour of *Quicksort*

Emergent Networks in Immune System Shape Space

John Burns
School of Computing
Dublin City University
Dublin.

Abstract

The development of a computational model is reported which facilitates the study of emergent principles of human immune system effector T cell clonotype repertoire and its distribution and differentiation. In particular, the question of systemic *self-organisation* is addressed. The model represents an extension to earlier immune system *shape space* formalism, such that each activated effector T cell clonotype and respective immunogenic viral epitope is represented as a node in a two-dimensional network space, and edges between nodes models the affinity and clearance pressure applied to the antigen presenting cell bearing the target epitope. As the model is repeatedly exposed to infection by heterologous or mutating viruses, a distinct topology of the network shape space emerges which may offer a theoretical explanation of recent biological experimental results in the field of murine (mouse) cytotoxic T cell activation, apoptosis, crossreactivity, and memory - especially with respect to repeated reinfection. In the past, most discrete computational models of immune response to viral infections have used separate real space or shape space formalisms. In this work, however, we have developed a model based on a combination of the two, with the objective of demonstrating how emergent behaviour and principles of self organisation may arise from a many-particle microscopic system. This is achieved by using a stochastic model of the lymphatic system as stimulus to a network

shape space model. The extension to the shape space formalism presented here was partially motivated by the need to address unrealistic restrictions imposed by early network models, but also by the need to propose a model by which mediation of early and protective immunity by memory T cells generated by a previous heterologous viral infection, can be explained. We propose that the variable topology of the emergent network in immune system shape space offers a mechanism by which the course of identical primary infection events across two individuals can vary in both duration and outcome.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

Almost since the invention of the software programmed computer, scientists and researchers have been designing and implementing models of biological processes. A popular early fore-runner of biological simulation was "The Game of Life" (*Life*), first presented by Gardner (1970). This application demonstrated how the 'survival' or 'death' of a point on a two-dimensional grid (or lattice) could be modelled as a function of the properties of its neighbours.

Although the early biological simulations were necessarily limited in scale due to the restricted computational resources available, they demonstrated that, in principle, computer systems could be utilised to model discrete biological processes. Following its publication, *Life* was the focus of attention because of the surprising ways in which systemic patterns emerged. In fact, *Life* is an example of a class of systems which manifest the ability to *self-organise*, in that their internal organisation increases without external influence. This class of system will be of particular interest in the work presented here, and is discussed in detail below.

As Computational power increases according to *Moore's Law* (Moore,

1965), which states that the number of components on a CPU doubles approximately every 18 months, the ability to numerically simulate systems of massive complexity has become ever more feasible. For example, Apple have recently demonstrated an application (running on 'off-the-shelf' technology), which can simultaneously model the trajectory of all known geo-stationary satellites currently in orbit around the earth.

Over the years, the use of computers in biology has emerged as a respected research area in its own right, and is now commonly referred to as *Computational Biology*, which can be defined as *the development and application of data-analytical and theoretical methods, mathematical modeling and computational simulation techniques to the study of biological, behavioral, and social systems* [1].

Numerous publications have come into existence to support this burgeoning field, including, but not restricted to *Bioinformatics* (Oxford University Press), *The Journal of Bioinformatics and of Computational Biology* (World Scientific) and *Journal of Molecular Modeling* (Springer-Verlag Heidelberg), and over the last decade, the number of papers being published in this field has risen dramatically. For example, Liebman (2001) has noted that by the middle of 2001, some 1,974 papers with *Bioinformatics* in the title were published, compared to only 12 papers in 1987.

As mentioned above, there are two complementary strategies in the field of Computational Biology.

1. *Bioinformatics*: The computational technologies for the study of how information is collected and transmitted in biological systems, starting from the molecular level, which then makes possible prediction and knowledge discovery (Bergeron, 2002).

2. *Modelling*: the design of *in machina* models of biological processes in order to aid the understanding of discrete complex biological phenomena (Chao *et al.*, 2004).

Recently, a further development in the field of computational biology

[1] NIH working definition of Bioinformatics and Computational Biology, July 17, 2000

has been the emergence of *computational immunology*. Petrovsky and Brusic (2002) have observed that the role of computational immunology is to transpose immunological problems into computational problems, to solve these problems using mathematical and computational methods and then to convert these results into biologically meaningful interpretations. This is the approach adopted in this thesis.

This research exploits the modelling process as a tool for *hypothesis testing*. Classical hypothesis testing does not set out to *prove* or *disprove* the hypothesis (Kanji, 2000), but rather, that an idea is untenable as it results in an unsatisfactorily small probability. However, in this research, such a strategy would be unsuitable. The system under study is not a naturally occurring one (such as an ant colony, or the population of a country), but is one constructed, *ab-initio*, as a *model* of a naturally occurring system (the immune system). As a model, it will be constrained, in part, by the selection of model parameters, and by extension, by the exclusion of other parameters.

Conversely, if an attempt is made to introduce *every* parameter from the naturally occurring system into the model, the model itself assumes a level of complexity approaching that of the system being modelled. If this happens, the value of modelling is reduced, or eliminated (Solomon, 2001). Thus, a key strategy in modelling large-scale complex systems, is the *selection of appropriate parameters*.

Mindful of this limitation, it is however, acceptable to pose hypotheses regarding model *fidelity*. In particular, in this work, the results of model simulations are qualitatively validated against relevant biological results, enabling further speculative questions to be raised about the system being modelled. Thus, modelling enables a range of experiments to be conducted which may be difficult or impossible in a clinical laboratory context. The focus of the last two chapters of this thesis will be to present some core (and, it is hoped, well-founded) assertions on the nature of specific immunological phenomena.

1.1 Systems of Self-Organisation

Many naturally occurring systems exhibit some form of *self-organisation*. For example, collections of stars form galaxies, which in turn are recognisable from the spiral structures they adhere to. Molecules form chemical compounds characterised by structural regularity and symmetry, and recognisable societies emerge from populations of unrelated individuals.

Traditional scientific theory sets out to explain such features by referencing the microscopic properties or laws applicable to their component parts, thus leading to fields of specialisation which although highly successful in deriving domain-specific models. This approach, often termed *Reductionism*, has been very successful, and can be illustrated by noting that over 150 years of disease research has come to identify *genes* as the source of many diseases and cellular dysfunction (see, for example, Wallis (1999)).

It is possible to approach the study of such systems in a different manner, looking instead for properties applicable to *all* such many-body problems, regardless of size or nature. If some form of self-organisation is evidently at work in the examples given above, it is reasonable to ask if such systems therefore share organising principles *in common*. It is here that modern computers prove essential, enabling the investigation of dynamic changes that occur over vast numbers of time steps and with a large numbers of initial options.

1.2 Emergent Behaviour

Emergent Behaviour can be defined as *the appearance of a property or feature not previously observed as a functional characteristic of the system* (Dimitrov, 1997). Generally, higher level properties are regarded as emergent. There are three relevant aspects to emergent behaviour.

1. *Supervenience.* The emergent properties will no longer exist if the lower level is removed.

2. *Non-aggregation.* New properties are not aggregates in that they are

not just the predictable results of summing part properties

3. *Causality.* Emergent properties are not phenomenological (either illusions or descriptive simplifications only).

This means that the higher level properties should have causal effects on the lower level ones, a property known as *downward causation* (Campbell, 1974). This implies also that the emergent properties restrict the freedom of the parts by imposing boundary conditions or constraints. In examples reported in this thesis, the principle of emergent behaviour is demonstrated in our model.

1.3 Motivation for this work

Biological systems are, by their nature, complex systems. Most are high-order and non-linear, with behaviour which is always more than the sum of their parts. The immune system is a good example of such a system. The following description is consistent with most undergraduate texts in immunology, (for example, Janeway *et al.* (1999)). The immune system consists of some 10^{12} cells (though this number may increase by several orders of magnitude depending on the context), 10 distinct cell lineages, and can defend the host against an almost infinite variety of challenges. The immune system has no central control, yet is capable of exquisite self-organisation, co-ordination and lethal retaliation. As well as being both robust and error tolerant, the immune system is adaptive in that it can dynamically alter its behaviour to suit a rapidly changing context.

Since Edward Jenner's discovery (in 1796) that cowpox (also known as *vaccinia*) induced protection against the often fatal disease of smallpox, and went on to call this procedure *vaccination*, 200 years of research has identified much at the individual immune cell-level. Furthermore, many important discoveries have helped eradicate diseases which once took a massive toll on populations. For example, in 1979, the World Health Organisation declared smallpox to be eradicated. However, very little is known about the governing and self-organising principles of the immune system as a whole. For

example, the underlying dynamics which affect the outcome of an identical strain of infection across two similar individuals are largely unresolved (Selin *et al.*, 1998).

A commonly cited shortcoming of immunological research is the typically reductionist nature it follows. To solve cell population questions, intra-cell behaviour is addressed, leading to individual cell signalling analysis, which then requires a study of internal cell chemistry, thereby leading to exploration of molecular kinetics, and, finally, to gene expression research. Therefore, the following objective is set for this thesis.

> To construct a theoretical computational model of the immune system which can provide a high-fidelity localised viewpoint of the secondary lymphatic system, and to show how patterns of emergent and self-organising behaviour may arise. This self-organisation is then expected to provide insight into how the course of an infection may differ in terms of duration, severity and outcome across individual immune systems.

Focusing on the emergence of individual variation is particularly relevant when presenting population models, because in many models of individual responses (see, eg, Nowak and May (2000)), populations are treated as ensembles of homogeneous individuals, although clearly they are not. A fuller description of the immune system may be found in texts such as Janeway *et al.* (1999).

1.4 Organisation of the Thesis

The remainder of this thesis is arranged as follows. Chapter 2 is an introduction to immunological phenomena, as well as a study of related models of immune response. The chapter is intended to give the reader sufficient grounding in adaptive-cellular immunology such that they will be able to read the remainder of the work without the need to refer to further subject matter material. The chapter also traces the development of relevant immunological models, with the objective of defining the *state of the art.*

Chapter 2 also serves to identify some of the limitations of previous work with a view to addressing some of these in later chapters. Chapter 3 introduces the basics of the model to be explored in subsequent chapters, and explores its behaviour under various infection regimes. The objective in Chapter 3 is to study how stochastic events (such as chance transitions from active cell to memory cell) may be crucial in determining the course of an infection. Chapter 3 also demonstrates the fidelity of the model at a microscopic level, in order to examine purely *localised* model dynamics.

Chapter 4 introduces an extended model which illustrates *hybridisation*: the microscopic model of Chapter 3 is coupled with a shape space model, in order to study the principle of *downward causation*. A single strain pathogen is presented to the simulation and the globalised and localised effects which result, are analysed. The material presented in Chapter 4 then sets the foundation for the proceeding chapters.

In Chapter 5, the model is extended in order to support visualisation of shape space in order that patterns of stimulation may be studied and classified.

Chapter 6 presents a theoretical network model of shape space, and shows how such a network naturally emerges and develops over time. This chapter is crucial in identifying the link between network topology and immune function. The chapter presents a study of the robustness of the shape space network, describing the model behaviour when network nodes are removed randomly and in a targeted manner.

Finally, Chapter 7 is a summary of the key findings, and presents the next steps to be taken in this research.

CHAPTER 2

Background and Related Work

2.1 Immunology Background

Common to all vertebrate immune systems is the principle of sensing of localised space for the purposes of intrusion detection (Coles *et al.*, 2002). Intrusion, in this case, is the appearance of a bacteria, viral particle or infected cell which has the property of *agonist* (triggering an immune response against it). Agonist genetic material discovered must be eliminated in order to prevent infection (or even death), of the host. Broadly speaking, the means by which the intruder gained access to the blood stream or lymphatic compartments is not of interest [1].

Sensing of the lymphatic compartments (of which there are many) for intruders, is a systematic function of immune cell (lymphocyte) recirculation. The immune system responds to challenge using one of two approaches (in some texts, known as response *arms*): (i) the humoral response, consisting of B-cell and antibody production, or (ii) the cellular response, consisting

[1] Some viruses, for example, the influenza and corona viruses, enter the host through the air passages and not through tissue damage.

of T helper cell (Th) and cytotoxic lymphocyte (CTL) cell production. Cytotoxic lymphocyte precursor cells constantly recirculate and sample their environment in the search for foreign pathogens. A precursor cell is one which is naive or undifferentiated. It is a cell which has not met with a prior agonistic challenge.

The process of sampling involves two cells binding for some small time period, during which the immune cell senses the receptors of the bound cell to determine if the bound cell is an invading pathogen (or not). If the bound cell *is* an invading pathogen, the immune cell may be stimulated to produce clones of itself in order to attack and remove other cells bearing the same genetic material. Under normal circumstances, the production of clones ceases after some fixed period of time (usually, 5-6 days), and once the infection has been cleared, most CTL cells will undergo programmed death (apoptosis).

A small percentage of the clone population will remain activated indefinitely, and this population represents effector memory. A specific class of cell (known as the antigen presenting cell or APC), has the job of engulfing pathogens in order to display the pathogen's genetic markers on their surface (hence the name "presenting"), and thus to alert any recirculating cytotoxic lymphocyte precursor cells to the infection.

2.1.1 Infection Take-up and Response

When a pathogen or antigen has been taken up by an antigen-presenting cell, such as a dendritic cell, it is degraded into one or more peptide chains within the cytosol region of the APC, and is then bound to the major histocompatible complex (MHC) class I molecule (a process known as *antigen processing*) before finally being presented on the surface of the APC as an MHC:peptide complex, a process known as *antigen presenting*.

As the immune system may be faced with an infinite number of genetically varied challengers, the final form of the MHC:peptide complex may be characterised by enormous structural variability. To reliably detect this antigenic variation, the immune system generates its own diversity in the

form of a set of T cells capable of recognising MHC:peptide sequences, by means of a T cell receptor (TCR), with variable degrees of efficacy (Buseyne and Riviere, 2001). The complete TCR set is known as the *repertoire*. A set of T cell clones possessing the same TCR is said to be of the same *clonotype*.

During T cell maturation in the thymus, genes encoding for the T cell receptor undergo several cycles of rearrangement, resulting in a mature immune repertoire capable of recognising a large range of MHC-bound non-self peptides. The affinity with which a T cell receptor binds to the MHC:peptide complex arises from the sum of the binding interactions among the Complementarity Determining Region (CDR) and the exposed peptide (Germain and Stefanova, 1999). Variation in affinity at the TCR:MHC:peptide bind site can dictate whether the pathogen challenge has the properties of agonist, partial agonist, antagonist or null compound.

All agonists (both strong and weak) will cause the T cell to begin a process which will eventually end in the death of the infected APC, and which is characterised by the onset of *clonal expansion* (whereby the T cell which successfully bound the MHC:peptide complex gives rise to a population of clones, each sharing the same phenotype). Partial agonists may not trigger T cell effector response, and thus may not result in the death of the infected APC. Antagonists inhibit the functioning of the T cell effector, and null compounds do not interact with the TCR strongly enough to cause any signal transmission. The T cell will simply sample the MHC:peptide complex and move on.

2.1.2 Biological Assumptions

In this thesis, only the cellular or cell-mediated arm of the immune response is studied. This is because most of the initial reaction to viral pathogens encountered by the immune system is dominated by the cell-mediated response (Fuller *et al.*, 2004). Furthermore, only a subset of the cellular response *entities* is modelled, which are the effector T cell and the antigen presenting cell. These two components are modelled because they play the most crucial role in determining the detection and clearance of a viral chal-

lenge (Buseyne and Riviere, 2001; Klenerman *et al.*, 2002). The action of $CD4^{+}T$ helper cells and cytokine molecules are considered implicit in this model and are not treated directly.

2.2 Related Work

Classifying such a diverse research arena as theoretical immunology is bound to be fraught with difficulties. For one thing, much work is hard to unambiguously categorise, and many authors have presented work which draws together disparate areas of research, for example, Weisbuch *et al.* (1990) (mathematical networks), Stauffer and Weisbuch (1992) (statistical physics) and Chao *et al.* (2004) (statistical computing). However, in order to present a structured time-line on relevant work over the years, models will be reviewed which may generally be classified by one of more of the following general headings.

1. *Mathematical-Network Models*: typically continuous; either ordinary differential equations (ODE) or partial differential equation (PDE), discrete or network based.

2. *Discrete Automaton Models*: discrete and usually agent-based, meaning the status of each component in the system is individually calculated.

3. *Emergent Network Models*: an intersection of graph theory and complex systems.

2.2.1 Mathematical-Network Models

One of the earliest and most influential mathematical-network models presented was that of Jerne (1974). In this work, a mathematical model was proposed which attempted to explain the means by which immunological memory could be maintained. To this end, the concept of the *idiotypic network* was introduced, in which it was hypothesised that the immune system, rather than being a set of discrete clones that respond only when triggered

by antigen, is a regulated network of molecules and cells that recognise and stimulate each other even in the absence of antigen. The theory states that because antibodies are created in part by random genetic mechanisms, they must look like novel molecules to the rest of the immune system and thus should be treated like antigens.

The novel or idiosyncratic parts of an antibody are called idiotopes. The set of idiotopes that characterises an antibody is called its *idiotype*. Due to the completeness of the repertoire, the immune system should recognise the idiotopes on its own antibodies and make antibodies against them. Jerne suggested that during an immune response antigen would directly elicit the production of a first set of antibodies Ab1 . These antibodies would then act as antigens and elicit the production of a second set of antibodies Ab2, which recognise idiotopes on Ab1 antibodies. Similarly, a third set of antibodies Ab3 could be elicited that recognised Ab2 antibodies, and so forth.

A seminal work in theoretical immunology was the paper on generalised *shape space* (Perelson and Oster, 1979). The shape space formalism was introduced as a way to represent antibody-antigen binding dynamics. In shape space, of primary interest is the clonotype repertoire distribution and its differentiation. Since it is referred to extensively in what follows, an outline of this important paper follows. For simplicity, assume that the features which govern the clonotype of the CTL receptor and APC, can be represented by N integer parameters. If the N parameters are combined into a vector, the clonotype for each CTL and APC can be considered as points within an N-dimensional Euclidean space of length L_s. Cells having the same clonotype have identical shape space vectors, and reside at the same location in shape space. Denoting CTL and APC clonotype vectors as $\mathbf{c}$ and $\mathbf{a}$ respectively, shape space develops as follows:

Surrounding each $\mathbf{c}$ is a disc of radius r (clearly, with $N = 2$, the area of this disc is πr^2). Any $\mathbf{a}$ located within this disc will be subject to a clearance pressure inversely proportional to the distance (d) between the $\mathbf{c}$ and $\mathbf{a}$ in shape space ($d = ||\mathbf{c} - \mathbf{a}||$). A strong agonist is one for which $d \to 0$. An increasingly weak agonist is one where $d \to r$, while for a null compound $d > r$. Every $\mathbf{c}$ will have a set of agonist APC clonotypes, $A_{\mathbf{c}}$,

where $|A_{\mathbf{c}}| \geq 0$. Conversely, every $\mathbf{a}$ belongs to *at least one* $A_{\mathbf{c}}$ (if this were not the case, then some $\mathbf{a}$ would remain undetected indefinitely). It is axiomatic that shape space is always completely *covered*. That is to say, with n CTL clonotypes and $N = 2$:

$$n\pi r^2 \gg L_s^2 \tag{2.1}$$

A suitable value of N (the dimensionality of the shape space) has been the subject of much discussion and debate in the years that followed the publication of this paper. The authors originally suggest $5 \leq N \leq 10$ would be sufficient, but numerous papers subsequently modelled $1 \leq N \leq 5$ (for example, Weisbuch *et al.* (1990); De Boer *et al.* (1992); Papa and Tsallis (1996)). Later criticism of shape space (Carneiro and Stewart, 1994) specifically suggested $N > 20$ as a *minimum* requirement. This is because the authors experimentally demonstrated that shape complementarity alone was insufficient to predict binding strength.

By the authors' contention, a realistic shape space would required at least 20 dimensions. Later work has sharply disagreed with this position. Papa and Tsallis (1996) introduced a hybrid real-space shape-space model (reviewed below) with $N = 1$. Indeed, the models bit-string models of Chao *et al.* (2004) do not address the question of *dimensionality*, but focus instead on the cardinality of distinct clonotypes which may be represented. The specific criticisms of Carneiro and Stewart are dealt with in Section 2.2.3.

Other authors have introduced alternatives to shape space. In their presentation of IMMSIM, Seiden and Celada (1992*b*) implemented a deterministic recognition rule in which *any* infected cell may be recognised and cleared by any lymphocyte. Other models use N-length *bit-strings* to represent the parameters which determine binding and recognition. Such models are usually discrete, agent-based and computational, as opposed to continuous mathematical models. The affinity between two cells with bit-string receptors is calculated as an extended form of *Hamming* distance. Bit-string models are not traditionally implemented as co-ordinate points in an N-dimensional Euclidean space, although in principle there is no reason

why not.

A further development of idiotypic networks and shape space was proposed by De Boer *et al.* (1992). In their B-cell idiotypic network model (based on a PDE implementation), the authors demonstrate the concept of *pattern formation* in shape space in response to perturbations (a perturbation, in these models, is generally the appearance of an infected cell, and the reaction which follows). The authors include in their model a continuous bell-shaped Gaussian *activation function*. This function controls the sensitivity of the network to perturbations. Low stimulation or over-stimulation [2] causes little or no B-cell activation. As such, this was an inclusion of biological function known to occur when B cells were presented with various concentrations of antigen *in vitro*. However, as discussed in Section 2.2.2, some of the features observed in the work are difficult to account for from a biological point of view.

The idea of pattern formation in shape space was further developed by Noest *et al.* (1997), in their continuous B-cell only idiotypic network model. This model was one of the first to include the parameter of cross-linking (as well as binding between infected cell and immune lymphocyte cell). Cross-linking is the process by which a secondary stimulatory signal (beyond the recognition of antigen) is required by the B-cell to induce clonal expansion of the B-cell. This PDE model produced two interesting pattern formation classes in shape space which the authors designated "domains" and "dots". These formation patterns emerged from a two-step infection-stimulation. Step one caused the emergence of a mixture of stable and unstable domains. The stable domains represented an uninitialised shape space, and the unstable domains were immunised-suppressed (or suppressed-immunised). In step two, when the model was stimulated a second time, the unstable domains broke into stable "dots" equivalent to memory of the infection. Although the model was implemented as a set of differential equations in a continuous shape space, the approach further developed the grey scale techniques of shape space visualisation by showing how development

[2] Over-stimulation was later proposed by Wick (1998) to explain CD4+T cell loss in a model of Human Immunodeficiency Virus (HIV).

over *time* could be shown, and demonstrating an "imprinting" technique. Previous work (De Boer *et al.*, 1992) had the limitation of once-off snapshots of shape space at the end of the simulation. In Chapter 6, an alternative approach to shape space development over time is presented, which further improves the formalism by demonstrating how an emergent network links topology and function.

The shape space formalism was extended by Weisbuch *et al.* (1990); Weisbuch (1990), such that the B cells and antibodies occupied a generalised shape space. In these models, the dimensionality of shape space was low, ($N = 1$ or $N = 2$), but the work was notable in that it brought together the formalisms of shape space and idiotypic networks, explaining how mean field theory (the affect of localised or neighbouring clonotypes within a 'field') could be used to model clonal excitation. However, the early idiotypic network models of Weisbuch and others lack any analysis of *topological* network features such as node and edge addition and deletion [3]. This is because these networks had static and symmetrical topology, and generally presented little stochastic features. In fact, the authors do not directly refer to nodes and edges (or their equivalent) at all, generally leaving it up to the reader to infer such detail. The network dynamics of these early immune networks are as follows. Each clonotype node (i) in shape space is surrounded by a local field (h_i), computed as:

$$h_i = \sum_j J_{ij} x_j \tag{2.2}$$

where J_{ij} is the affinity matrix mapping some affinity constants from each of the x_j neighbours of the i-th node. Edges are implicit in Eqn. (2.2), in that any non-zero element at the row/column intersection of J_{ij} indicates an edge from i to j. From Eqn. (2.2) it is easy to see why this class of network suffered from spatial instability. As the population of nodes in shape space increased, the next neighbours would be stimulated and increase in population, and this process would repeat, causing a system-wide cascade or

[3] Later in this thesis, a shape space network topology is developed which plays a key role in regulating immune function.

percolation (especially in higher dimensions of shape space (Fortunato *et al.*, 2003), where $N > 5$). This would then require some form of field dampening in order to maintain network stability. Weisbuch *et al.* (1990) considered an *a priori* network topology, known as a Cayley Tree [4], (Fig 2.1), which demonstrated how an idiotypic network can attain localised stability (which represents immune memory) without addressing the question of spatial instability inherent in models using pre-defined network topology.

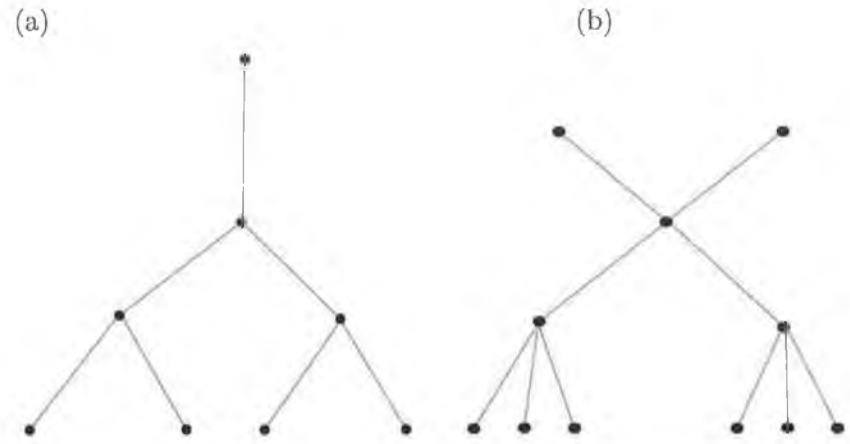


Figure 2.1: Two simple n-Cayley trees, in (a) $n = 3$ and (b) $n = 4$.

The Cayley tree approach was pursued up by Anderson *et al.* (1993), in a PDE model in which antibody dynamics were explicitly modelled. The authors experimented with the effects of increasing co-ordination number [5]. Further statistical-mechanical properties of Cayley trees were identified, such as limit cycles and localised chaotic attractors. Neumann and Weisbuch (1992) were the first to experiment with randomised idiotypic network structures, an important step in addressing the limitation of homogeneous edge configuration of earlier models. The authors found that variation in topology does indeed play a major role in the preservation or loss of localised attractors (which in turn represent memory or immunised states).

[4] An acyclical tree in which each non-leaf vertex has a constant number of edges.

[5] The exact number of edges to each node, being homogeneous in a Cayley tree, is referred to as the co-ordination number.

This, while important, still did not address previous concerns with idiotypic network models in that topology is still constructed *a priori* and biological relevance is *applied* to the model, rather than evolving *from* the model.

Further, the model does not readily address perturbations in the network, and a suppression mechanism (mentioned earlier) is needed to prevent memory loss when closely related antigens are presented to the network. Again, due to their pre-determined nature, these models have had limited success in replicating specific immune phenomena, largely because analysis reveals features which do not necessarily have a clear immunological equivalent, a point acknowledged by Weisbuch *et al.* (1990).

2.2.2 A Comment on Immune Network Models

It is interesting to note that research on network models of the immune system seems to have declined in the last five to six years. Perelson and Weisbuch (1997) concluded this was because biological experiments which can be done have already been done, and as such, no exciting new results have emerged to stimulate further research into immune networks. However, a number of additional reasons can be identified for this.

1. Idiotypic network models focused on B-cell inter-clonotype stimulation as a means to model humoral memory. It is generally accepted at this time that the phenomenon of immune memory, especially effector T cell memory is a function of long-lived primed effector T cell populations (Murali-Krishna *et al.*, 1999; Swain *et al.*, 1999; Crotty *et al.*, 2003; Liu *et al.*, 2003; Naumov *et al.*, 2003; Kim and Welsh, 2004), and not constant inter-clonotype stimulation, as had been believed for many years. In the light of this fact, one of the central threads of idiotypic network models has lost any direct biological counterpart.

2. Although such models used shape space as a means to differentiate clonotypes, they make the compromise that the affinity between clonotypes was modelled either as a binary variable, or as a three-state variable (where affinity is low, high or none). In the original work (Perelson

and Oster, 1979), affinity decreased as a function of increasing distance in space between the lymphocyte and the antigen, to a cut-off point ρ (also known as the *clonal cutoff*, or threshold).

3. The rigid *a-priori* topology of idiotypic networks tended to force researchers to find biological parallels to statistical-physical phenomena. Although spatial instability could be used to explain the process of autoimmune disease (for example), this dysfunction is actually quite rare. On the other hand, percolation was an inherent part of idiotypic networks, and again, the problem of identifying regularly occurring biological processes to match the characteristics of idiotypic networks to has remained. Another such feature is *chaotic attractors*: regions of the network where attractor nodes (or clusters of nodes) randomly appear and cycle unpredictably through the system. These networks usually have bi-directional edges, connecting B-cell idiotypes [6] to one another, and are therefore susceptible to percolation processes.

2.2.3 Validity of Shape Space

One of the criticisms that Carneiro and Stewart (1994) direct at the Continuum Shape Space theory is that the function $f(\mathbf{Ab}, \mathbf{Ag})$, where **Ab** is an antibody and **Ag** is an antigen, must be highly irregular and discontinuous. This assertion is based on work by chemists which has shown that predicting affinity and bonding between two molecules is not simply a deterministic issue of understanding the dynamics between the individual molecular constituents. However, Perelson and Oster clearly indicate that shape space does not *need* to be characterised by a uniform distribution of antibody shapes. Further criticism of the shape space paradigm is the question of the value for N.

Although Carneiro and Stewart suggest the original value of $5 \leq N \leq 10$ is too small (they suggest a value closer to $N \geq 20$), they do not, in principle, question the theoretical foundation of representing antigenic determinants by a fixed, N-sized set of parameters. The actual value for N is clearly

[6] The set of unique differentiating markers on the cell surface.

something that is system-specific and may vary. In fact, Yates *et al.* (2000) have shown how the presence of cytokine regulatory molecules crucially affects the dynamics of helper T cell populations. The presence of cytokines can up- or down- regulate function to an extent which can effectively negate or enhance the size of N.

It is therefore plausible that the actual value and parameters of N are not only dependent on the characteristics of **Ab** and **Ag**, but also of external and localised state information such as the density of cytokines. It seems reasonable to conclude that setting N to a fixed and relatively small number is sufficient to represent the antigenic determinants.

2.2.4 Discrete Automaton Models

Discrete automata models have characteristics, which in some respects make them preferable to continuous mathematical models already discussed. In particular, the immune system itself is a discrete system in which the individual behaviour of each single cell collectively determines the system-wide characteristics. Cellular Automata (CA) have been applied to numerous areas of complex physical systems modelling (Wolfram, 1982; Brass *et al.*, 1993; Bernardes and Zorzenon dos Santos, 1997; Bandini *et al.*, 2001). CA have several important characteristics which make them amenable to efficient computational implementation, including ease of representation (in the form of n-dimensional arrays). Furthermore, the CA model types are an obvious choice, given the discrete nature of the underlying computations, simplicity of rules or laws which are programmed into the CA, and the highly repetitious nature of the processing steps.

However, cellular automata possess additional fascinating properties, not least the ability to produce patterns of self-organisation of a highly complex nature which cannot be derived analytically from the rules on which the underling cellular automata is based.

As a result of this complexity, Wolfram (2001) has postulated that some form of CA must underlie many complex physical phenomena visible in nature. Furthermore, with the application of non-deterministic (*stochastic*)

cellular automata, the idea of randomness in CA site selection and update rule enforcement has yielded further insight into modelling natural stochastic systems such as molecular motion, turbulence in water flow and various biological processes, especially models of the human immune system.

Cellular automata have been the basis of many microscopic computational models of immune response (Stauffer and Pandey, 1992; Castiglione *et al.*, 1997; Mannion *et al.*, 2000; Bernaschi and Castiglione, 2001). Such microscopic models are usually, but not always, implemented on regular lattices (Wolfram, 1982; Bandini *et al.*, 2001; Wolfram, 2001), and are often described as "agent-based" in that each site of the lattice behaves independently according to the conditions found in the local neighbourhood.

Thus, agent-based models are often a good choice when patterns or structures in local space have significant effect on model fidelity. The state of the lattice over time is simply a sequence of random variables $\mathbf{x}^{(0)}, \mathbf{x}^{(1)}..., \mathbf{x}^{(t)}$ defined on a finite space $\mathcal{X}$. The sequence is a *Markov chain* (Liu, 2001) as the value of $\mathbf{x}^{(t+1)}$ is dependent on its history only through its recent past $\mathbf{x}^{(t)}$.

Several common neighbourhood specifications exist, including *von Neumann*, which includes neighbouring cells at a distance of 1 along exactly one of each of the coordinate lines (left and right in one). The *Moore*, includes the von Neumann neighbourhood as well as the the cells found on the diagonals, to yield 8 neighbours in a two dimensional lattice. Other authors choose less common specifications, for example Bernaschi and Castiglione (2001), use a 6-neighbour model arranged in the form of a diamond in order to more realistically model the number of neighbours which participate in site update evaluation.

In the model presented in later chapters, a square neighbourhood of radius $r = 1$, in two dimensions, is employed, yielding 8 nearest neighbours when the diagonals are included. Stauffer and Sahimi (1994) noted that variation in neighbourhood structure has a *quantitative* but not a *qualitative* effect on results; this finding is used to support the choice of neighbourhood size in the model developed in this thesis.

Finally, agent-based approaches have yet a further benefit which is suit-

ability for incorporation of stochastic events. Germain (2001) has proposed that stochastic events may be crucial in determining immune response and the course of infection.

The first direct computational model of the local spaces of the secondary immune organs was presented in Seiden and Celada (1992*a*,*b*), though more recently appearing as IMMSIM (Puzone *et al.*, 2002). These models were implemented using a cellular automaton to mimic the behaviour of individual immune cells within the thymus.

Each lymphocyte and antigen carried an 8-bit string which acted as the antigenic determinant. Even though the early models were limited in scope, due to the modest computational resources available at the time [7], recent models have been quite successful in reproducing the localised behaviour of the immune system. For example, recent versions of IMMSIM now incorporate humoral and cellular arms of the immune response, as well as B-cells, T cells, antigen, antigen presenting cells and stimulatory cytokines.

The work of Seiden and Celada gave rise to many further CA models for work on specific phenomena of the immune response, for example, Brass *et al.* (1993); Castiglione *et al.* (1997); Pandey (1998); Mannion *et al.* (2000, 2002); Sloot *et al.* (2002); Castiglione *et al.* (2001); Bernaschi and Castiglione (2001).

An early large scale automaton model of shape space only (as opposed to the continuum shape space models discussed in Section 2.2.1) was developed by Stauffer and Sahimi (1994), applying large computational resources for calculating update rules on hypercube lattices of dimensions 16000^2, 35^5, 7^9, and 5^{10}. The authors find that when shape space dimensionality is in the region of $5 - 10$, sharp phase transitions between the immune states of uninitialised and immunised appear, but in lower dimensions, such phase transitions are absent. This finding has been interpreted as being in agreement with the original Perelson and Oster value of N.

In contrast to the agent-based models of secondary lymphoid organs discussed above, a number of cellular automaton models of *shape space* have been proposed; for example, Stauffer and Weisbuch (1992) and more re-

[7] For example, the maximum concentration of T cells was limited to around 800.

cently, Stauffer and Proykova (2004), introduce a high-dimensional discrete computational shape space model ($5 \leq N \leq 10$) in part to address an observation made in the original shape space paper on the appropriate value of N.

The authors do not, however, directly tie model results to any biological phenomena, and neither do they comment on the improvement obtained in higher dimensions. It is interesting to note too that the authors observe no qualitative difference in results beyond $N \geq 5$, which supports one of the comments made in Perelson and Oster (1979). Bernardes and Zorzenon dos Santos (1997) develop this work further and explore the behaviour of a discrete n-dimensional hypercube shape space at the edge of chaos. The point at which a phase transition occurs in shape space which represents a transition from the uninitialised state to the immunised (or memorised) state. Significantly, these authors also conclude that there are no qualitative differences (in their model) between low ($N = 3$) dimensions and high ($N = 5$) dimensions.

More recently, a deterministic model of shape space consisting only of the B-cell repertoire (but including antigen presenting) was presented by Lagreca *et al.* (2001). Here, shape space dimension was modelled as $N = 12$ (implemented as a bit-string, where this results in a repertoire size of only 4096), and the model included re-infection. The authors demonstrate how exposure to an antigen previously seen resulted in faster elimination and higher densities of antibody generation. Hershberg *et al.* (2001) also proposed a HIV-specific model based exclusively on shape space, which is considered in some detail here. Notably, the authors do not model real spaces of the immune system. Instead, they assume infected cell and killer cell will meet with some constant probability λ. No distinction is made between virion particles (a complete virus particle with its DNA or RNA core and protein coat as it exists outside the cell) and infected $CD4^+$ T cells. In fact, they do not make any distinction between $CD4^+$ T cells and effector cytotoxic lymphocytes (CTL) cells. This lack of distinction somewhat reduces the biological fidelity of the model, because under no circumstances are $CD4^+$ T cells and CTL cells *synonymous*. This is an important consider-

ation because HIV is a rather atypical pathogen in that it specifically targets $CD4^+$ T cells (McKinney *et al.*, 2004; Palmer *et al.*, 2004), thus impairing the ability of the immune system to mount an effective CTL response.

This work also modelled shape space affinity as a lock-and-key relation, in that either full recognition of the antigen, or no recognition occurs. In recent immunological research, however, this assumption has been abandoned (Mason, 1998; Brehm *et al.*, 2002). It does seem unlikely that there is a constant probability of an immune cell [8] encountering an infected $CD4^+$ T cell, for the obvious reason that immune recirculation dynamics are known to alter once an infection has been detected (see, e.g., Janeway *et al.* (1999), pp 265). The model represents viral strain *mutation* as the probability of a specific strain diffusing across the shape space lattice. Thus, in this respect, every mutation is an escape mutation. The authors reach a similar conclusion to Nowak and May (1991) in noting that antigenic diversity is a key determinant in the outcome of HIV. Interestingly, this work on diversity presents a mechanism to combine microscopic and macroscopic dynamics.

Of the former, the dynamics of recognition and removal of infected cells is the determinant dynamic, and of the latter, the rate of diffusion (ie, mutation) of the HIV virions across the shape space is the determinant. Their findings are in broad agreement with Zorzenon dos Santos (2000), who demonstrated a stochastic cellular automaton model which reproduced the characteristic three phases of the HIV infection quite well. It is important to observe that HIV pathologically affects recirculation dynamics Kirschner *et al.* (2000), and that any model of this virus would need to explicitly include mobility as a factor. The work of Mannion *et al.* (2002) was notable in addressing this aspect.

Discrete agent-based models combining both real space and shape space formalisms are somewhat less common in the literature, the main contribution in this respect being Papa and Tsallis (1996). In this paper, the authors presented an interesting B-cell model, but one which was constrained in size (two dimensional 20×20 real space, and a shape space of one dimension,

[8] Again, immune cell meaning CTL cell, although the authors do not make this distinction explicit.

with 18 clonotypes), and which did not include any antigen presentation.

Because of the limited capacity of shape space, the authors employed a very much simplified recognition rule which triggered a recognition event if clone A was exactly $r, r-1, r+1$ units from clone B, where r is the radius of the periodic one-dimensional shape space arranged as a ring (see Fig. 2.2). This approach admitted the possibility of only 3 clones being allowed to stimulate a given clonotype, and as the shape space was limited to 18 clonotypes, the probability of *some* recognition is given by $p = 0.166$, which is an unrealistically high value (in the original work of Perelson and Oster, for example, the fraction of clonotypes which bind a randomly selected antigen was estimated conservatively at 10^{-3}). Although this work demonstrated simple structure formation in real space, the authors did not present any shape space dynamics. This may have been due to the formulaic structure adopted in shape space. From Fig 2.2 it can be seen that the ring implementation does not support stochastic evaluation of binding.

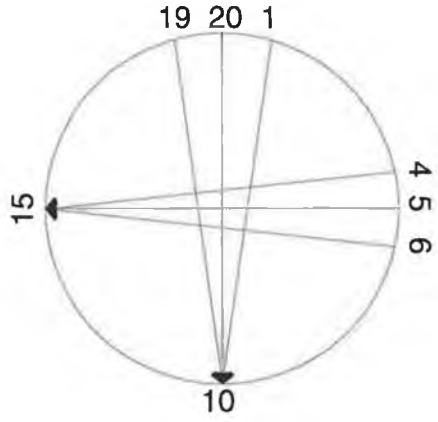


Figure 2.2: A ring based shape space, with $r = 10$. The cells at 10 and 15 are suppressed by cells exactly $r, r-1, r+1$ units away $(19, 20, 1$ and $14, 15, 16$, respectively).

Recently, discrete computational stochastic stage-structured models were presented (Chao *et al.*, 2003, 2004). Here, the authors avoid the computa-

tional overhead inherent in agent-based models, by dividing the cell life-cycle into distinct stages. All cells in a particular stage are considered identical and thus do not require specific data/memory structures to represent them. State transitions from one stage to another are controlled by probabilities which apply to *all* cell populations in a particular stage. Although this approach addressed the computational overhead of agent-based systems, it may not be *necessary* in that modest sized, off-the-shelf PC clusters can efficiently scale to simulate 10^6 agents when coded to limit memory allocation (Burns and Ruskin, 2004*a*). This model used only $40MB$ of physical memory with a 10^6 real space lattice, and taking approximately 4hrs to run on a single CPU machine.

These models have succeeded well in replicating the dynamics of effector T cell memory (in particular), although, as these models do not explicitly include individual cell kinetics, no emergent systemic behaviour is addressed. In fact agent-based computing does not have to be prohibitively expensive. In a recent paper (Burns and Ruskin, 2004*a*), an agent-based model, which was coded efficiently and reproduced cell density levels typically observed in murine (mouse) experimentation, was presented.

2.2.5 Emergent Network Models

From the perspective of theoretical modelling, the study of complex systems and emergent behaviour offers a mechanism by which well understood microscopic features of immune models may be used to predict macroscopic behaviour. For example, complex adaptive systems can usefully be represented as a network in which each entity is represented as a node, and interaction as edges between nodes. When the topology of networks as diverse as the metabolic system, protein interaction and protein domains are studied, non-random organising features become apparent (for example, Jeong *et al.* (2000); Wuchty *et al.* (2003); Barabasi *et al.* (2004)).

Principles such as scale free degree distribution, addition and deletion of nodes and preferential attachment [9] have been observed across many

[9]The process by which a new node preferentially attaches an edge to a node with a

heterogeneous network domains (for example, Albert and Barabasi (2000, 2002); Wuchty and Stadler (2003)).

As biological systems are characterised by evolution and adaptation, networks to model specific aspects of such systems must be capable of supporting edge and vertex addition and deletion, along with edge *rewiring*, and such networks are then termed *evolving* or *dynamic*. In evolving networks, non-random organising principles lead to the study of *emergent topology* which may provide insight into network-wide characteristics. In this section some important characteristics associated with complex networks are reviewed.

1. The degree $deg(k)$ of a node is the number of edges associated with that node. The average degree $\langle k \rangle$ characterises the overall graph, and the variance of $\langle k \rangle$ is captured by the degree distribution $P(k)$.

2. The shortest path l_{ij} between two nodes i and j is used to calculate the mean path length for the network:

$$\langle l \rangle = \frac{2}{N(N-1)} \sum_{i=1}^{N} l_{ij} \tag{2.3}$$

   One further description is appropriate when considering the mean path length, the concept of *small worlds*, meaning that any node on a network can be reached by any other node by using a small number of steps (the 6-degrees of freedom theory, in fact). This is a measure of how easily a network may be traversed. A network which can be crossed with relatively few steps is said to possess a small world quality. This attribute is also known as the "six degrees of freedom" attribute.

3. The clustering coefficient C_i for node i is a measure of the average connectiveness of its neighbours and is given by:

$$C_i = \frac{2n_i}{k_i(k_i-1)} \tag{2.4}$$

large number of existing connections.

were n_i is the number of links connecting the k_i neighbours of i to each other.

The mean clustering coefficient $\langle C \rangle$ of the overall network is:

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i \tag{2.5}$$

A further measure of network structure, based on $\langle C \rangle$, is the function $C(k)$, which determines the average cluster coefficient for all nodes of degree k. If $C(k)$ turns out to be *independent* of the choice of k, this may indicate that the network is either *homogeneous* (all nodes having the same degree), or else dominated numerous small, tightly knit clusters.

On the other hand, if $C(k)$ follows $C(k) \sim k^{-1}$, the network is characterised by sparsely connected within highly connected areas. An example of the two network structures is shown in Figs. 2.3 (a) and (b). The class of sparsely-connected, highly-connected networks is typical of the structure found (for example) in metabolic networks (Ravasz *et al.*, 2002).

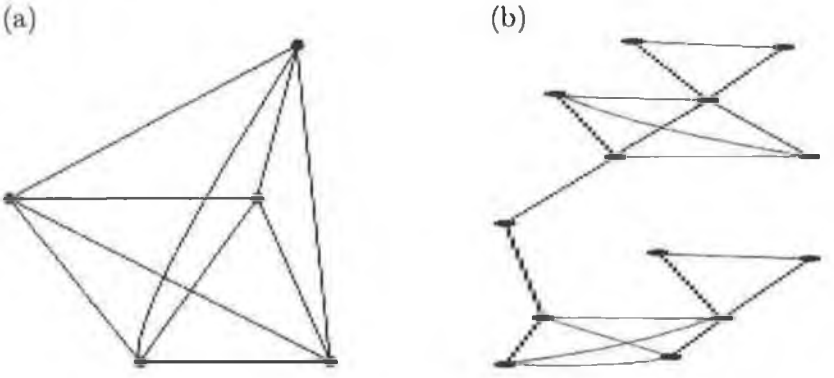


Figure 2.3: Homogeneous and hierarchical network structures, (a) and (b) respectively.

In Fig. 2.3 (a), the network is homogeneous, with $k = 4$, and $\langle C \rangle = 1$ (ie,

the most highly connected). In comparison, Fig. 2.3 (b) is more hierarchical with one node connecting two highly clustered areas.

In addition, Wuchty and Stadler (2003), note the following three *centralising metrics* associated with the nodes of evolving networks. The *centre*, the *median* and the *centroid*. The center C of a network G is the node $x \in V$ which minimises the furthest distance between itself and any other node $y \in V$:

$$C(G) = min_{x \in V}(max_{y \in V} d(x, y)) \tag{2.6}$$

The median M of G is a function which minimises the average distance to it from all other nodes:

$$M(G) = min_{x \in V}(\sum_{y \in V} d(x, y)) \tag{2.7}$$

Finally, the centroid Q of G is:

$$Q(G) = min_{x \in V}(\sum_{x \in V} d(x) - min_{x \neq y}(\sum_{y \in V} d(x, y))) \tag{2.8}$$

Large scale graphs with no (apparent) design principle have been described as *random* graphs. In such graphs, two nodes are linked together with a probability p, yielding a graph with around $pN(N-1)/2$ randomly distributed links.

Several biological networks such as protein and metabolic networks, (see, eg, Jeong *et al.* (2000, 2001); Ravasz *et al.* (2002)), have been analysed, and found to follow a power law degree distribution distribution:

$$P(k) \sim k^{-\gamma}. \tag{2.9}$$

In the work conducted by Barabasi and co-workers, the scale-free networks identified all possessed the property of $2 \leq \gamma \leq 3$. In Jeong *et al.* (2000), $\gamma = 2.2$. The apparent wide-spread existence of scale-free networks in biological systems may help to explain why such systems are usually extremely robust. Elimination of random nodes and edges (due to, for ex-

ample, mutation), has little impact on the network topology - indeed, the average diameter of the network would remain unchanged. However, when nodes possessing high cluster coefficients are eliminated, the connectedness of the network is reduced. If enough connected nodes are eliminated, the functioning of the network collapses.

In a recent paper, Albert *et al.* (2000) has compared the robustness of two types of random networks to node elimination: the Erdos-Renyi (ER) and the scale-free networks. The ER network is homogeneous, with every node having on average, the same number of edges, with probability p. On the other hand, scale-free networks are characterised by *preferential attachment.* The probability of a new node i being connected to another node l, is given by:

$$P(i) = k_l / \sum_{l}^{N} k_j \tag{2.10}$$

2.3 Summary and Conclusions

In this chapter an overview of some important research, which forms the basis of the work presented in the rest of this thesis, was presented. Several important observations regarding the limitations of previous work have been made.

In order to examine how individual immune systems may develop minor topological variations in immune system shape space which affect the clearance rates and therefore the duration and pathology of infection, the following points need to be addressed:

1. Shape space is an important formalism in modelling global characteristics of the immune system repertoire (notwithstanding the comments of Carneiro and Stewart (1994)). However, any shape space model will need to address the question of *dimensionality* (N).

2. Shape space -only models suffer from the effects of not modelling recirculation dynamics. As these dynamics are crucial in determining cell

interaction kinetics (as well as being highly non-linear), omission of recirculation dynamics may result in over-simplification of the model.

3. Pre-determined network models of immune system shape space suffer from the difficulty of applying (almost retrospectively) immune phenomena to the known statistical-physical function of the network.

4. Real space models are good at reproducing localised phenomena, but do not offer a means to capture and analyse system-wide behaviour over time.

5. Adaptive and complex network analysis of large-scale microscopic systems has yielded several methods with which to characterise and predict the behaviour of the underlying system to various stimuli.

For earlier shape space network models, the network edges are a loose notion, meaning *is capable of interacting with due to complementarity.* In Chapter 6, the network model is developed further by assuming that an edge means *has been stimulated by*, and additionally, each edge carries a flow (weight) equal to the distance (and hence, the affinity or stimulation) between the shape space entities. Furthermore, by modelling the network with *directed edges*, the problem of spatial instability is removed.

It can be concluded by saying that any model which seeks to address the *emergent behaviour* of the immune system will need to address the five observations above in order to extend knowledge in this field of research.

CHAPTER 3

Initial Microscopic Model

In the work developed in this thesis, each chapter builds on the work of the preceding chapters. As the work is concerned with developing a model of how the adaptive immune response reaches a system-wide condition (a condition which may then be used to predict disease clearance dynamics), a logical starting point is to consider the microscopic dynamics which underpin system-wide development.

Firstly, the term *microscopic* is used throughout this work to mean the localised, real space compartments of the secondary immune organs, and in particular, to mean the *lymph nodes*. There are, in fact, many lymph nodes, distributed throughout the body, but in this work only an idealised and singular node (or compartment) is considered. This is because the clinical structure of each node is remarkably similar, and because each node is designed for the same function, (to trap foreign pathogens drained from the blood, and to enable lymphocytes to mount an appropriate response) an assumption of one idealised compartment is not a crucial over-simplification.

Microscopic modelling shall be the lowest level of detail addressed in this thesis, with the objective of developing an initial model which is flexi-

ble enough for extension in later chapters, while at the same time, accurate enough to simulate a simplified immune system consisting of antigen presenting cells, cytotoxic effector T cells, one generalised lymphatic compartment, and initial, secondary and repeat infection events. A simplified lymph node structure is shown in Fig. 3.1.

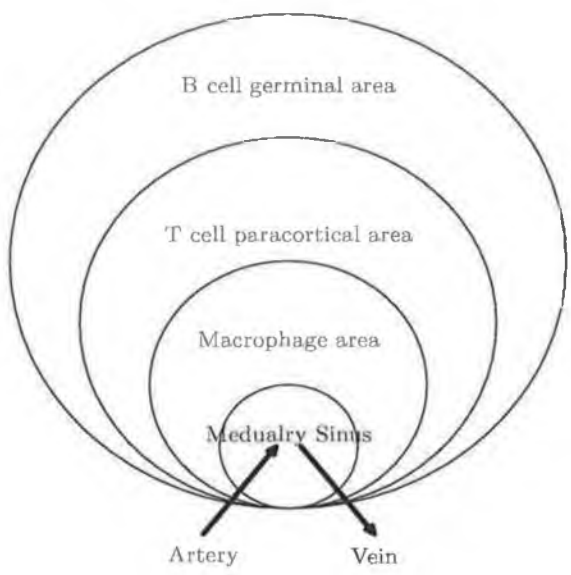


Figure 3.1: A simplified lymphatic compartment showing the different regions in which immune cells circulate.

Infected antigen presenting cells enter the lymph node as part of the blood supply to the *artery*, and progress from the *Medulary Sinus* to the *B cell germinal area*, by penetrating the consecutive walls separating each area. When the infected APCs reach the *T cell paracortical area*, they come in contact with recirculating T cells (both precursor and active), triggering an immune response if the epitope (any part of a viral protein detectable by the immune system) on the APC surface is immunogenic.

Cells leave the lymph node via the blood supply to the vein, and continue to cycle through the other lymph nodes. Kirschner *et al.* (2000), have estimated that in a one day period, lymphocytes spend a total of 30 min-

utes in the blood stream, the remainder of the time being spent recirculating through the various lymphatic compartments.

The pattern of lymphocyte recirculation in and out of the lymphatic compartment depends on the route and severity of the infection. HIV (Human Immunodeficiency Virus), a viral pathogen which mainly enters the immune system by way of infected blood supply or sexual activity, is marked by intense lymphocyte production and death within the lymphatic compartment itself (Buseyne and Riviere, 2001). By contrast, bacterial infections which enter the body through mucosal membranes or skin damage, will induce migration of activated T cell lymphocytes from the lymphatic compartment to the region where the detection arose.

In this model, the following approach is adopted. Infection notification arises when antigen presenting cells are introduced into the generalised model of the lymphatic compartment. The response which follows is based solely within the lymphatic compartment, because clearance of infected antigen presenting cells from the lymphatic compartment is a feature common to all infection dynamics, notwithstanding the fact that some responses may subsequently induce T cell migration.

Furthermore, the approach taken here also avoids a modelling process influenced *exclusively* by current experimental immune system research trends. More precisely, the study of *stochastic events*, which appear to play a crucial role in certain immune system functions (Germain, 2001), will be of particular interest. Therefore, in this, the initial microscopic model of the generalised lymphatic compartment, the *life-cycle* of the CTL cell and APC will be examined first. This is the subject of the following sections.

Although this model uses a 2-dimensional real space to model the lymphatic compartment (which is clearly a 3-dimensional structure), it was noted by Mannion *et al.* (2002) that this has little impact on overall model results, although in their model, a limited neighbourhood structure may have accounted for this.

3.1 Life-cycle of the effector T cell

The dynamics of immune response to initial infection and reinfection by the same pathogen sometime later, are considerably different. Primary response, which follows initial infection, is characterised by relatively slow precursor cell activation and population growth rates, with a consequent elongated pathogen clearance profile, typically extended over six to eight days (Busch *et al.*, 1999; Busch and Pamer, 1999).

On the other hand, secondary response (to reinfection by the same pathogen some time later) is notable for short effector activation time, high specificity of response, rapid pathogen elimination and high degree of memory cell participation (Blattman *et al.*, 2000; Barber *et al.*, 2003). In some cases, secondary response is so effective, the infection does not become symptomatic. The effectiveness of the secondary response is largely dependent on the class of effector T cell from which the response originates: the *memory* effector T cell.

Here, a seven state non-deterministic finite automata (NFA) of the effector T cell life-cycle, which is encoded as a set of states and state transitions, is developed. The objective is to study the degree to which variable infection outcome is the consequence of premature, delayed or even failed state transitions. Small variation in crucial state transition probabilities during the life-cycle is shown to induce widely variable infection outcomes.

It is worth noting that Asquith and Bangham (2003), have observed that splitting the CTL population into "memory" and "effector" subpopulations may be an over-simplification, in that the relationship between the two subpopulations is unclear. In their model of CTL fratricide (inter-CTL cell destruction) the authors implemented a continuous mathematical model of the CTL life-cycle using what they term a *continuously varying phenotype.* They noted that this approach supports the CTL effector phenotype when infected cell concentrations were high, and the CTL memory phenotype when levels were low. Although their approach is slightly at variance with the one presented in this research, there is agreement on a core assumption, that of the effectiveness of the secondary response based on memory CTLs.

The key features of this chapter are as follows: (i) an extended non-deterministic state transition model of the effector T cell life-cycle is introduced. This model successfully reproduces time-realistic effector and pathogen population dynamics during primary and secondary response, and during repeated reinfection, (ii) a crucial state transition in the effector T cell life-cycle is identified which is critical in regulating secondary response, and (iii) the model is scaled to cell density levels in the order of 10^6 CTL cells - which approaches levels typically observed in *in-vivo* murine experiments.

A note on murine experiments: Murine experiments are conducted on laboratory bred mice (non-wild type), with typical experiments involving the injection of around 5000 infected cells, with the animals being tested at regular intervals for disease progression. Often, the organs are removed some point in the experiment (usually 20-40 days after exposure), and cells levels and activation patterns are studied. In some experiments, animals are specifically bred with some genetic deficiency (for example, lacking in CD4+T cells) in order to study the consequences of such deficiencies.

3.2 Initial Model Development

The model runs in discrete 30-minute time-steps, and all entities in the model act asynchronously at each time-step (τ). The 30-minute time-step is used because most cell milestones are in the order of hours, as opposed to minutes or days. As primary response normally consists of 4 days of cell replication (clonal expansion), the cells in the model will stop dividing at $\tau = 192$. The recirculation space of the lymphatic compartment is modelled as a two dimensional stochastic cellular automata lattice of length $L = 10^3$, with periodic boundary conditions and neighbourhood radius $r = 1$ (in two-dimensions), with a maximum of 8 neighbours (the Moore neighbourhood).

An asynchronous CA (Chopard and Tomassini, 2004) is used, in which each site is selected at random for update during a time-step, but every site will be visited at each time-step, and each site can be updated at most once in any given time-step. The site to be updated is selected randomly from a uniform distribution $\sim U[0, 1]$, without replacement, with the justification

that, in biological and noisy systems, agents act at different and uncorrelated times, which precludes a global clock or synchronisation protocol (Chopard and Tomassini, 2004). Thus, an asynchronous site update strategy is preferable in approximating the true characteristics of the underlying system.

At $\tau = 0$, 5000 APCs are introduced into randomly selected sites on the lattice (following a uniform distribution), and the model executes until $\tau = 3000$ (62.5 days of elapsed time). The value of 5000 is specifically selected as it is a value typically used in murine experiments. With respect to the layout the simplified lymph node (Fig 3.1), is equivalent to the infected antigen presenting cells all appearing within the T cell paracortical area at exactly the same time. A more biologically exact simulation would model a *diffusion* of the APCs into this region over some number of time-steps. All APCs introduced into the real space display the epitope markings of the pathogen engulfed. In this model, there are no pathogen-free APC.

The CTL population grows exponentially in response to APC stimulation, with a rate which is a function of the distance between the CTL and APC in shape space (but never exceeding 0.036). Each lattice site may contain only one entity at any given time-step. The set of entities and states supported is shown in Table 3.1, which also introduces some important notation used throughout this chapter.

In contrast to other models (for example, Bezzi *et al.* (1997)) which attempt to directly simulate human immune system lymphocyte levels (the order of $\sim 10^{12}$), the model cell concentration levels are limited to realistic amounts normally observed in murine experimentation, typically of the order of 1.8×10^4 during primary response (Bousso *et al.*, 1999). The justification for this is that not *every* lymphocyte is active during a given immune challenge. Using this approach, individual behaviour of each cell can be directly modelled without resorting to approximate methods (for example, the stochastic stage structured models of Chao *et al.* (2003, 2004)).

A comment on simulation structure: Most of the simulations results presented in this thesis are *averages*, over several simulation runs. For each parameter set, some 30 simulations are usually carried out, with average cell levels shown in the figures, without reference to the sample variances. Typi-

cal variance levels are in keeping with stochastic computational simulations (see, eg, Law and Kelton (2000)), and are not treated further in this work.

3.2.1 Non-Deterministic Finite Automata

To allow the study of a *distribution* of possible outcomes (for example, immunised, chronic infected and collapse), a subset of the CTL life-cycle state transitions is identified, and the certainty of a transition from state w to state v on event e is replaced with some probability (< 1) of state transition. First, the definition of state transition relaxation is defined as follows. If X is a discrete random variable (drv) representing the transition from state w to state v, and e is some event linking wv, the relaxed state transition X_r is:

$$P(X_r|e) = 0 \leq \psi \leq 1 \tag{3.1}$$

The choice of value for ψ will naturally depend on the wv in question. A concrete example of this rule is the case where a CTL in the clonal expansion state (w), will transition to the effector state (v) when the event $\tau = 192$ occurs (e). In contrast to earlier models [1], Eqn. (3.1) implies *duality* in the presence of event e: transition on e (X_r) or not ($\bar{X}_r$). This extension results in a *non-deterministic* finite automaton (NFA) (Hopcroft and Ullman, 1979). Fig. 3.2 is a non-deterministic finite automata model of the life-cycle of the CTL (and follows notation explained in Table 3.1). E is the set of events for the model, and consists of both deterministic and non-deterministic elements. A subset of three critical non-deterministic events $S \subset E$ is defined as: $S = \{e_2, e_3, e_5, e_8\}$. Each $e_i \in S$ is defined as follows:

e_2, e_8 An infected antigen presenting cell will be destroyed by a bound cytotoxic lymphocyte cell which recognises it. Recognition is a function of the distance between the two cells in shape space.

e_3 An activated proliferating immune cell (state $ctl^{+\star}$) will normally end clonal expansion on the event $\langle e_3 : age(ctl^{+\star}) > 192 \rangle$ time-steps.

[1] in which $P(X_r|e) = 1$

e_6 Memory cells are primed from the point of a previous infection and they confer an advantage during reinfection in that they produce armed effector cells without spending time in clonal expansion. Normally the majority of $ctl^{+\dagger}$ enter $ctl^{+\bullet}$ on event $\langle e_6 : age(ctl^{+\dagger}) \geq 192 \rangle$.

e_{rpt} Repeated reinfection events, resulting in repeated doses of infected antigen presenting cells introduced into the simulation, at time-step $\tau + 300n, n = 0, 1, ..., 9$.

Each of the above events (e_n) has an associated probability ψ_n. The set $\{\psi_1, \psi_2, \psi_3, \psi_4\}$, therefore fully describes each simulation configuration of the model (all other parameters being kept constant). In the results presented in the following section, the following four experimental configurations of $\mathcal{P}$ are defined:

ψ_1 $\mathcal{P}_1 : \{0.9, 0.9, 0.9, 0.0\}$

ψ_2 $\mathcal{P}_2 : \{0.9, 0.9, 0.95, 0.0\}$

ψ_3 $\mathcal{P}_3 : \{0.9, 0.9, 0.9, 1\}$

ψ_4 $\mathcal{P}_4 : \{0.9, 0.9, 0.95, 1\}$

The first two configurations of $\mathcal{P}$ test the fidelity of the model response when confronted with a singular secondary infection event some 30 days after the initial infection. The first configuration represents a normal response and is intended to calibrate the model for near optimal conditions. For $\mathcal{P}_1$, it would be expected to see CTL production levels broadly characterised by a low, elongated peak for primary infection, followed by an increase in memory CTL. Another expected observation is APC clearance, being over $6 - 10$ days for primary response, and significantly faster during secondary response. The second configuration is an increase in ψ_3 from 0.9 to 0.95 and is intended to test the impact of a 5% decline in the number of cells which transition to the effector memory state $(ctl^{+\dagger} \rightarrow ctl^{+\oplus})$.

Some viral infections are known to cause damage or loss of the memory pool (Kim and Welsh, 2004), and the impact this has on the model is

tested. Repeated reinfection is tested with normal and depleted memory cell production ($\mathcal{P}_3$ and $\mathcal{P}_4$, respectively). Many pathogens are known to lead to acute and persistent viral infections. For example, Wherry *et al.* (2002) have found that memory function correlates to *epitope expression levels*: if excessively high levels of epitope are expressed during primary infection, the quality of memory function is reduced. Tewari *et al.* (2004) found that prolonged exposure to antigen during chronic infection tends to impair memory function, suggesting that infections which are not eliminated rapidly, but which linger, cause damage to memory efficacy. Wu *et al.* (2004) demonstrated that viral FLIP (a type of protein expressed by certain viral antigen) quickly inhibit the production of memory cells, even after normal primary response. In an important and fascinating report, Williams and Bevan (2004), observed that the formation of CTL effector memory in murine in-vivo experiments, was impaired with treatment of ampicillin 24hrs post-infection. This indicated that early termination of infection leads the generation of fewer effector memory cells. More generally, this finding may support the observation that treatment of infection with penicillin should be done conservatively and not before the infection has become chronic.

The importance of memory cell production is examined in these cases. Again, memory cell production is impaired by 5%, and the consequences of this loss on the effectiveness of infected cell clearance is studied. Section 3.3.1 examines the outcome of persistent infection in the model. Table 3.2 shows the set of state transitions possible for the NFA shown in Fig. 3.2.

3.3 Initial Results

The model is initially executed with parameter set $\mathcal{P}_1$ and $\mathcal{P}_2$ (with no repeat reinfection), and the results are shown in Fig. 3.3. In (a), the initial infection is visible at $\tau = 0$ with pathogen density $p_d = 5000$, and consequent effector response reaching a maximum value at $\tau = 300$ (6.25 days), with the concentration of effector cells, $e_d = 8.2 \times 10^3$. Antigen presenting cell density is shown by the broken line, with the solid line indicating levels of effector memory and activated cells combined. For clarity, (b),(d) show

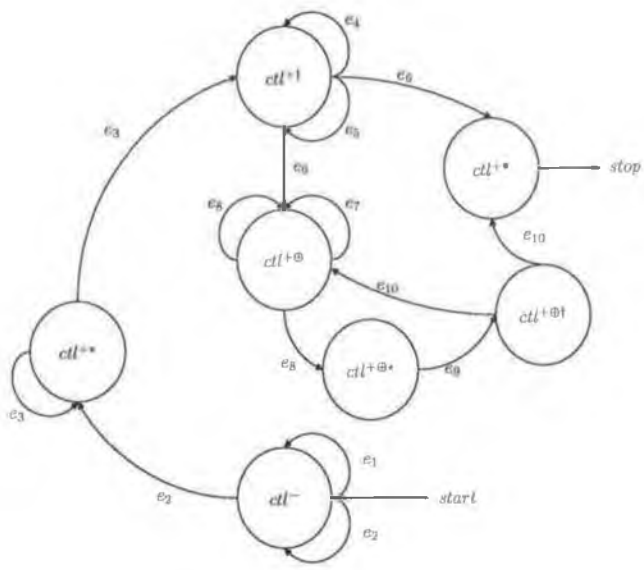


Figure 3.2: Seven-state non-deterministic finite automaton of the cytotoxic lymphocyte cell life-cycle. Transition events (e_n), carrying the same subscript are *non-deterministic*.

population levels for APC for each $\mathcal{P}$.

Fig. 3.3(b) shows the antigen presenting cell population level (only). No memory cells are present during primary response, and as such, the effector cell population is made up entirely of clones produced by stimulated precursor cells. To the right of each effector cell peak is a plateau of memory cells. The slope of the CTL density peak is extreme, indicating that the state transitions from $ctl^{+\star}$ to $ctl^{+\dagger}$ to $ctl^{+\oplus}$ (or $ctl^{+\bullet}$) occur with a high degree of certainty.

At time $\tau = 1500$ (day 31), secondary exposure to the same pathogen occurs, and the model exhibits following general behaviour: (i) the secondary immune response is preceded by a pool of committed CTL memory cells which have already been primed to respond to the re-appearing pathogen, (ii) the activated CTL density is some 10 times higher than primary response, and does not last as long, and (iii) the pathogen is reduced to half its original level much more rapidly than during primary response. With

Table 3.1: Notation and definition of model entity states

Notation	Definition
ctl^{-}	naive recirculating effector precursor
$ctl^{+\star}$	proliferating lymphocyte
$ctl^{+\bullet}$	dead activated lymphocyte (apoptosis)
$ctl^{+\oplus}$	memory effector
$ctl^{+\oplus\star}$	activated proliferating memory effector
$ctl^{+\oplus\dagger}$	armed memory effector
$ctl^{+\dagger}$	armed activated effector
apc^{+}	active infected antigen presenting cell
$apc^{+\bullet}$	dead infected antigen presenting cell

$\mathcal{P}_1$, the model exhibits efficient detection and clearance behaviour associated with a healthy immune system.

From Fig. 3.3, it can be seen the advantage in both time and infected cell clearance which is conferred on a response based largely on memory. The half life of the virus during primary response is around 3.25 days, with 90% pathogen clearance achieved at around $\tau = 480$, or 10 days of simulation time. Compared to secondary response on reinfection an infected cell half-life of $\tau \approx 60$ or 1.25 days - an efficiency of around 87% is observed. Effectively, this is because memory cells, having already been primed by a previous encounter with the specific pathogen, undergo expansion with lower death rates than during primary response, and therefore accumulate more quickly (Grayson *et al.*, 2002).

The results for $\mathcal{P}_2$ are shown in Fig. 3.3(c) and (d). Here, the probability of entering apoptosis is increased from 0.9 to 0.95. This means that the memory cell population would be around 5% of that activated effector population post-primary response. Recent work (notably Grayson *et al.* (2002)) has shown that some $\approx$ 90% of activated effector undergo apoptosis after primary response. Therefore, $\psi_3 = 0.95$ would represent an unusually high suppression of memory function.

Clearly, the reduction of memory effector production should not affect

Table 3.2: Event Transition and Definition

Notation	Definition
e_1	Recirculating CTL does not detect APC
e_2	Recirculating CTL detects APC and enters clonal expansion, or does not
e_3	Activated APC continues clonal expansion *or* ends becomes activated killer CTL
e_4	Activated killer CTL does not detect APC
e_5	Activated killer CTL detects and kills APC
e_6	Activated killer CTL enters apoptosis at $age > 192$ *or* becomes a memory CTL
e_7	Memory CTL does not detect any APC
e_8	Memory CTL detects APC but does not kill it *or* kills APC and becomes a proliferating effector
e_9	Proliferating memory CTL becomes activated killer
e_{10}	Activated memory CTL becomes a resting memory CTL *or* dies

primary response, and this is borne out by CTL density levels prior to $\tau = 1500$ (c). A normal 10-day clearance regime (d) is seen during primary response, but a less effective response during reinfection. In fact, the memory cell pool in the time range $500 \leq \tau \leq 1500$ has fallen to ≈ 500. Once reinfection occurs, the APC population is cleared some 31% more effectively than during primary response. The APC half life is $\tau = 108$, 90% clearance is achieved after reinfection at $\tau \approx 1788$ (or some 5.9 days of simulated time). However, the characteristics of $\mathcal{P}_2$ are significantly degraded compared to that observed in $\mathcal{P}_1$.

3.3.1 Persistent reinfection

Some viral pathogens are capable of persistent reinfection. That is, although population levels of infected antigen presenting cells may decline in response to clearance pressure by a specific CTL response, over time, the number of infected cells rises to chronic and sometimes acute levels. Examples of

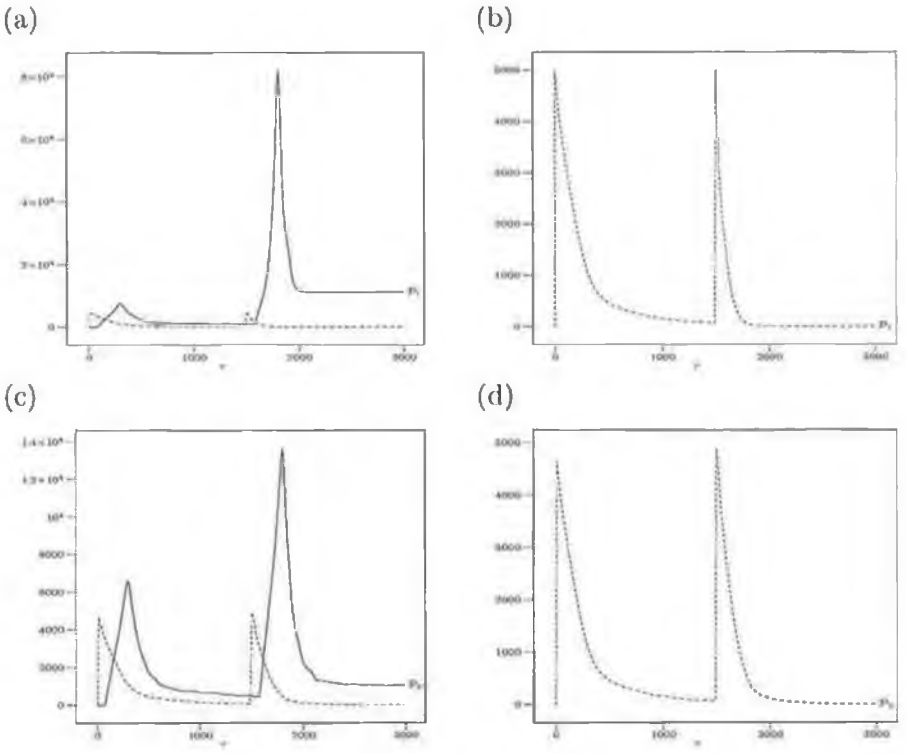


Figure 3.3: CTL and pathogen lattice density levels (a),(d) over a simulated 62.5 day period, with an initial infection at time $\tau = 0$ (x-axis), and a reinfection by the same pathogen occurring at $\tau = 1500$, for 3 values of $\mathcal{P}$. The y-axis indicates the average concentration of CTL cells over 30 simulations.

such viruses are HIV, Human T Cell Lymphotropic (HTLV), hepatitis C (HCV), hepatitis B virus, Cytomegalovirus (CMV) [2], Ebola Virus (EBV) and rubella (Kim and Welsh, 2004). Such persistent reinfection pathogens have been associated with normal immune function suppression. In this section, persistent reinfection was simulated by randomly scattering a repeat 'dose' of the pathogen into the lattice, introduced at $\tau + 300n, n = 0, 1, ..., 9$. This reinfection pattern represents a resurgence of infected cells every 6.25 days, in discrete bursts. The results of this simulation are shown in Fig. 3.4.

[2]The same group which causes Epstein-Barr Syndrome.

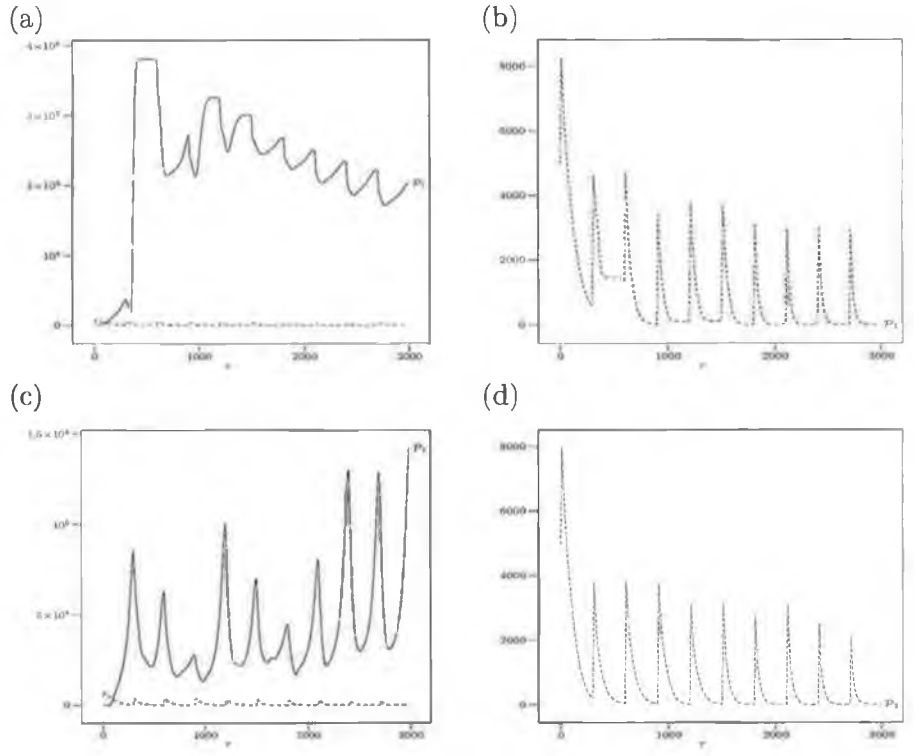


Figure 3.4: The model is exposed to repeated infection events, arising at time $\tau = 300n, n = 0, 1, ..., 9$ (x-axis), equivalent to an infection every 6 days. The y-axis indicates the average concentration of CTL cells over 30 simulations.

With respect to Fig. 3.4 (a), the response to the first reinfection is clearly strong. Some 3.8×10^5 lymphocytes are generated and the reinfection is rapidly eliminated. As further infections arise starting at $\tau = 600$, the existing memory pool never falls below 1.8×10^5, and is critical in bringing the repeated reinfections under control in time periods (b) which rarely exceed 130 time-steps (or 2.8 days of simulated time).

From (a) it can also be seen that slightly lower responses are sufficient in order to effect optimal clearance. Results from (a) and (b) support the clinical findings that the memory cell levels tends to be higher after secondary and tertiary infections (Grayson *et al.*, 2002), which in turn, supports the

clinical practice of vaccination boosting.

Finally, when the simulation is executed with diminished memory cell creation and repeatedly stressed with reinfection ($\mathcal{P}_4$), average primary and secondary response levels are similar (around 1.2×10^4). Each response is characterised by rapid expansion and reduction of effector lymphocyte clones. There are no memory cells to confer clearance advantage, and each response is initiated from low levels (around 1.2×10^2).

3.4 Chapter Summary

The approach taken in this chapter was to build an initial microscopic model of the effector T cell life-cycle, in order to study a distribution of possible simulation outcomes. The model reproduces well the time and space dynamics of initial and secondary infection. In addition, the research is novel in modelling the relationship between repeated reinfection and effector cell transition to memory or apoptosis.

The work demonstrates how repeated reinfection can be controlled only within a limited range of ψ_3. Too much memory causes the lymphatic compartment to fill-up, too little memory induces the need for clonal expansion from naive precursor cells, and a elongated APC clearance profile. When the ratio of apoptosis to memory is 'just right' ($0.88 \leq \psi_3 \leq 0.92$), antigen presenting cell levels (during repeated reinfection) are brought under control in increasingly rapid time frames. Very recent clinical work (Kim and Welsh, 2004) suggests that the immune system must periodically preferentially eliminate some memory cells which exhibit poor cross-reactivity. One of the benefits of the stochastic effector T cell life-cycle model presented here is the relative ease with which this theory could be investigated.

This chapter has put an important foundation into place. A microscopic model of an *idealised* lymphatic compartment succeeded in simulating known clinical CTL concentration levels and infected cell removal rates during primary secondary and repeat infection events. In particular, the CTL concentration levels matched those known to arise in murine laboratory experiments. Therefore, this model is a strong basis upon which to build the

crucial important thesis objective of integrating a local space model to a system-wide (or global) model. This is the subject of the next chapter.

CHAPTER 4

Repertoire Distribution and Differentiation

In the previous chapter, an agent-based model was developed which enabled detailed study of immune response to primary, secondary and repeat reinfection events. The model also supported reinfection by both the same and unrelated pathogens, and demonstrated qualitatively realistic space and time dynamics during initial, secondary and repeated reinfection events when compared to murine experimentation results. In addition, the non-deterministic state transition model of the effector T cell life-cycle reproduced well the sensitivity of response efficacy to effector cell state transition to apoptosis versus state transition to memory cell.

It is worth noting that in this work, many of the conclusions are calibrated against murine (mouse) *experiments*. This is because murine experiments contain more computationally tractable cell levels (eg, 10^6), and murine experiments tend to be more ambitious (for obvious reasons).

The features of the model from the previous chapter were purely *microscopic* in that only local (or *real*) space dynamics were studied. What was missing from this approach was a method to represent system-wide immune condition. Only when system-wide immune condition is properly modelled

can assumptions about infection duration and outcome between individuals and within a population, be made.

The concept of the *immune repertoire* is crucial in respect of system-wide immune condition, and could be defined as *the set of precursor and activated T cell clonotypes, in both the memory and non-memory pools.* The immune repertoire transcends the local spaces of the immune system (Naumov *et al.*, 2003). Furthermore, certain properties, such as clonotype activation and distribution, can be modelled directly in the repertoire, but would be impossible to represent in an entirely local space model, as such models capture only individualistic agent-based information. The state of the repertoire can be categorised by analysing the distribution of *uninitialised*, *activated* and *immunised* spaces within it. If there were a means to represent the repertoire as part of a microscopic model, this would enable both the repertoire condition to influence local space dynamics, and crucially, *for local space dynamics to influence the evolution of the repertoire.*

In this chapter the model presented in Chapter 3 has been further developed by devising an integrated shape space and real space model. This chapter addresses two important points (from Chapter 2), namely, (i) that shape space is an important formalism in representing the immune system repertoire, and (ii) that shape space -only models suffer from the effects of not modelling recirculation dynamics. The objective of this chapter is twofold:

1. To outline the method by which real space and shape space are integrated.

2. To examine the sensitivity of the primary immune response to shape space parameters.

4.1 Downward Causation

The principle of *downward causation* is central in demonstrating that a formalism is capable of emergent behaviour. Downward causation is *the condition by which all processes at the lower level of a hierarchy are restrained*

by and act in conformity to the laws of the higher level (Campbell, 1974). In this thesis, the immune repertoire is the core high-level process which influences and restrains the evolution of key immunological processes such as antigen presenting cell *recognition* and *clearance* and effector cell *clonal expansion*, all of which are lower level process in the hierarchy.

In Fig. 4.1 this concept is presented in a simplified form. The repertoire state affects how processes such as *APC Recognition* proceed, which in turn, influences *CTL Activation*. Subsequently, the CTL Activation process changes the state of the repertoire as the clonotype in question has become *activated*.

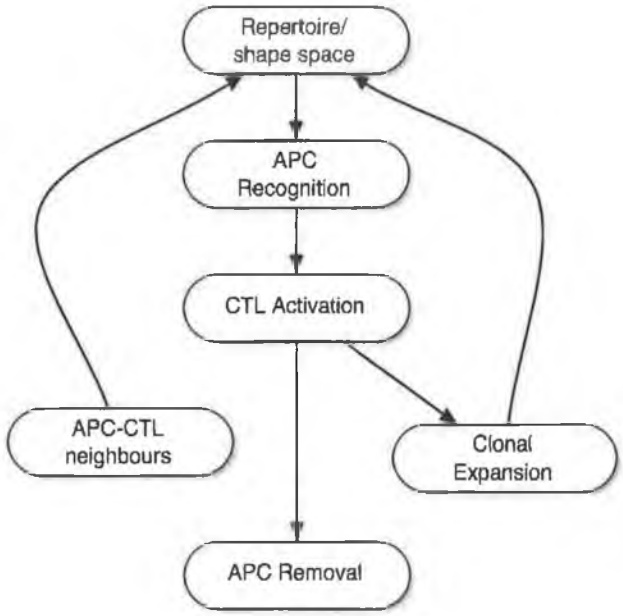


Figure 4.1: A simplified hierarchy of causation, with the repertoire acting to influence processes below and in turn, being affected *upward* by lower level processes.

The hybrid model can understood as a mapping from real space to shape space as shown in Fig. 4.2. In the lower section of the figure, the real space compartments are shown with CTL cells (white circles) and infected cells

(grey circles). When an infected cell comes into contact with a CTL, shape space is used to determine the affinity between the cells. If the affinity falls within the clonal cutoff region (denoted in the figure by the circle radius, r), then clonal expansion begins. When shape space has been stimulated by further CTL activation, the network topology emerges (top panel).

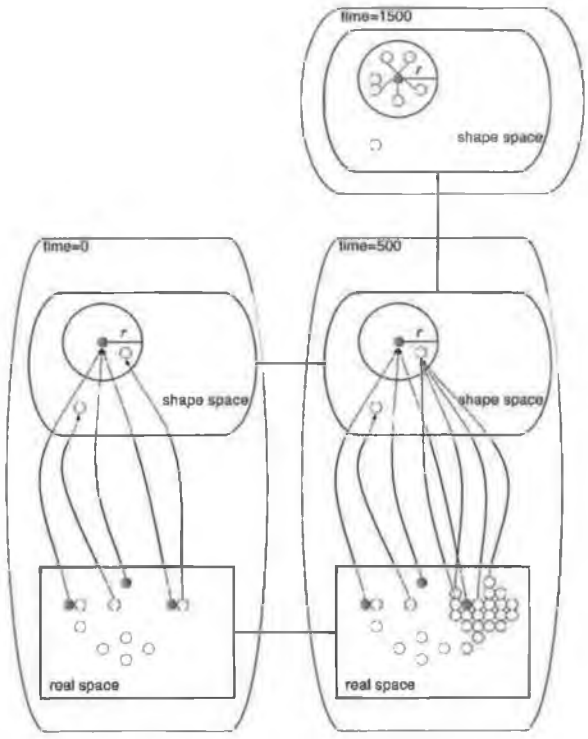


Figure 4.2: A mapping from the real space compartment to shape space, showing three time points, $(0, 500, 1500)$. At time step 1500 (top panel only), the network structure is beginning to take form in shape space.

As mentioned earlier, shape space can be used to represent the immune repertoire. In addition, shape space has the further important property of being suitable to model binding affinity. Affinity, in turn, drives the rate of clonal expansion, and in the following work, the term shape space is used to encompass the characteristics of the immune repertoire.

4.2 Shape Space Integration

The question of *how* the repertoire is integrated into the local space models of previous chapters is now addressed. Fig. 4.3 is a logical view of how shape space and real space are integrated. Each site in the real space lattice $(k_0, k_1, \ldots, k_n)$ has a pointer to a C language structure [1]. Pointers are used in order to implement fast movement about the lattice. To move occupants around, a pointer swap is all that is required.

On the right, the lattice of sites $ss_0, ss_1, \ldots, ss_n$ is the array holding shape space information (such as clonotype density levels). Information flows bi-directionally from shape space in that local space entities are assigned their clonotype or epitope co-ordinate, and local space population changes are updated in shape space. Each lattice site maintains information on the particular model entity occupying k_i at that point in time, such as:

1. *Type*: describes the current occupant of the lattice site k_i with values drawn (for example, ctl^- or apc^+) from Table 3.1
2. *Age*: age of the cell occupant, increases by 1 at each time-step
3. *Proliferation Rate*: the rate at which this cell proliferates, calculated according to the probability of binding (Eqn. (4.3)). Applies to ctl^+ cells only.
4. *Shape space coordinate*: vector of length N containing the CTL clonotype or epitope position in shape space.
5. *Activated by*: the shape space coordinates of the apc^+ which stimulated this ctl^- into proliferation.

In the particular example shown in Fig. 4.3, the antigen presenting cell has two cytotoxic lymphocyte cells as nearest neighbours. When *apc* and ctl^- are adjacent in real space, shape space affinity rules determine whether binding between the two will occur. Shape space is used to model the affinity

[1] These models are compiled and executed using gcc and under Solaris 9. A port to Fortran is not yet available

between CTL clonotype and viral epitope (and is discussed in Section 4.2.1), whereas the life-cycle of the real space effector T cells is the same as that presented in Chapter 3.

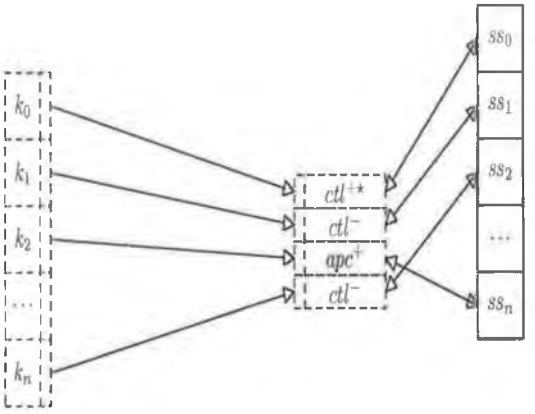


Figure 4.3: On the left, the array of lattice sites $k_0, k_1, \ldots, k_n$ represents the real space data structures in memory. Each k_i holds a pointer to a structure describing the site occupant for the current time-step τ.

Throughout this research, uppercase letters identify shape space components (so the activated CTL clonotype in shape space is denoted CTL^+) and lowercase denote real space components (an activated CTL cell in the real space compartment is therefore denoted ctl^+). The superscript is an *activation* indicator, where + implies activated, and − implies a precursor or naive cell.

4.2.1 Clonal Affinity and Shape Space

The dynamics governing affinity between antigen and lymphocyte are developed directly from the shape space formalism (the original paper is reviewed in some detail in Chapter 2). In shape space, each antigen epitope and CTL clonotype is represented as a point within the two dimensional ($N = 2$) discrete space, of size 50×50. Surrounding each CTL clonotype is a disc of radius ρ (although models such as Smith *et al.* (1998), have placed the antigen epitope at the locus of the disc of influence, without loss of generality).

Any antigen epitope-bearing APC located within the disc will be subject to clearance pressure with a force inversely proportional to the distance between the two. Shape space is used to model the density and distribution of the CTL clonotype repertoire, and to provide analytical underpinning into the critical nature of the measure of crossreactivity, which is denoted by ρ. *Crossreactivity* is the probability that a CTL clonotype stimulated by one viral epitope, ϵ_i will also react with a structurally different epitope ϵ_j, or, conversely, the number of different clonotypes which respond to the same (randomly selected) epitope.

Thus the selection of a value for ρ affects the average number of CTL clonotypes stimulated by some randomly selected epitope drawn from a uniform distribution. Some research (for example, Mason (1998)), has suggested this figure to be in the range of $50 - 111$ clonotypes, whereas others (Valitutti *et al.*, 1995; Itoh *et al.*, 1999) have suggested the range $80-200$. However, these estimates are controversial, and Borghans and De Boer (1998) have estimated a much lower level of around 5.2. In this model, this issue is treated by considering the statistical *expected value* $E(X)$. Denoting the number of shape space clonotypes as Θ_{ctl}, the length of the shape space as L_s, it is expected that the average number of clonotypes stimulated by a randomly selected epitope is:

$$E(X) = \Theta_{ctl}\pi\rho^2/L_s^2 \tag{4.1}$$

With $\rho = 10$, $\Theta_{ctl} = 100$ and $L_s = 50$, $E(X) = 12.5$, which is above the conservative estimate, but well below the rather high estimate suggested by Mason (1998). Clearly, from Eqn. (4.1), the average number of clonotypes responding to a randomly selected antigen in shape space is dependent on the crossreactivity ρ, and this dependence is further explored in Section 4.3.

For simplicity, assume that the features which govern the CTL clonotype and epitope-bearing APC binding can be represented by N parameters, taking integer values only. If the N parameters are combined into a parameter vector, each CTL clonotype and antigen epitope can be considered as points within an N-dimensional Euclidean space of length L_s. Cells of the same

clonotype have identical shape space vectors, and reside at the same location in shape space.

Shape space is a dynamic and evolving N-dimensional Euclidean space that contains one vector $\mathbf{u}$ for every immune system CTL clonotype, and one vector $\mathbf{v}$ for every pathogen epitope. Around each $\mathbf{v}$ in shape space is a disc of influence of radius ρ. Any $\mathbf{u}$ falling inside this disc of influence is subject to some pressure. That is to say, the pathogen will be removed from the real space system with some probability $P(X)$, inversely proportional to the distance between $\mathbf{v}$ and $\mathbf{u}$ in shape space. If the distance d between the two points exceeds the critical value ρ, then there is no CTL pressure on the pathogen, and no affinity or binding takes place.

A simplified shape space is shown in Fig. 4.4; in this figure, three CTL clonotypes have been activated by the epitopes $\epsilon_1, \epsilon_2, \epsilon_3$. The crossreactive cutoff (ρ) is the radius of the disc surrounding each CTL, and any epitope falling within this disc will have some clearance pressure applied against it. Shape space uses periodic boundary conditions, such that any activated CTL clonotype of less than ρ units from the space boundary will have a disc that simply 'wraps-around'.

The distance calculated between the CTL and antigen epitope is shown in Eqn. (4.2) and the probability of binding is given in Eqn. (4.3).

$$d = \|\mathbf{v} - \mathbf{u}\| = \sqrt{\sum_{i=1}^{N} (v_i - u_i)^2} \tag{4.2}$$

$$P(X) = \begin{cases} 1 & : \quad d = 0 \\ \frac{1}{d} & : \quad 0 < d \leq \rho \\ 0 & : \quad d > \rho \end{cases} \tag{4.3}$$

It is axiomatic that shape space is always completely *covered*. With n CTL clonotypes and $N = 2$:

$$n\pi r^2 \gg L_s^2 \tag{4.4}$$

It should be emphasised at this point that Eqn. (4.2) and Eqn. (4.3)

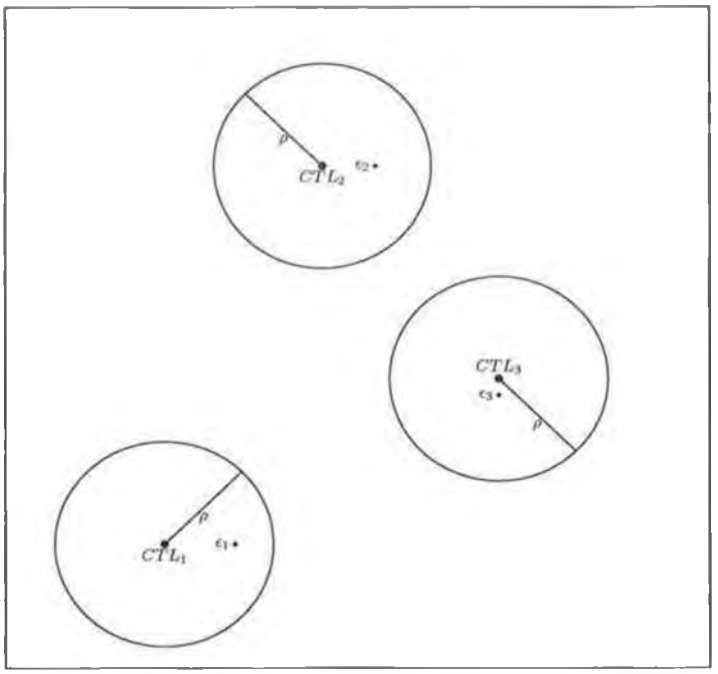


Figure 4.4: A simplified shape space in two-dimensions, with three activated CTL clonotypes, denoted by CTL_1, CTL_2 and CTL_3, with the stimulant viral epitopes denoted by $\epsilon_1, \epsilon_2, \epsilon_3$ respectively.

will be evaluated if and only if there is some *contact* between an antigen presenting cell and a cytotoxic cell within the real space model. Contact means that the two cells are no more than radius $r = 1$ from each other within the Moore neighbourhood (Fig. 4.5 (a)).

If Eqn. (4.3) does result in the activation of the CTL clonotype $P(X)$ is used as the *stimulation* rate (S) for the exponential growth during the clonal expansion phase. Therefore, the total number of activated clones which an activated CTL clonotype may give rise to at time τ is:

$$CTL_\tau = e^{0.036\tau S} \tag{4.5}$$

where the power of e depends on three parameters, τ, the duration of the expansion phase, S, the stimulation rate (dependent on the distance between

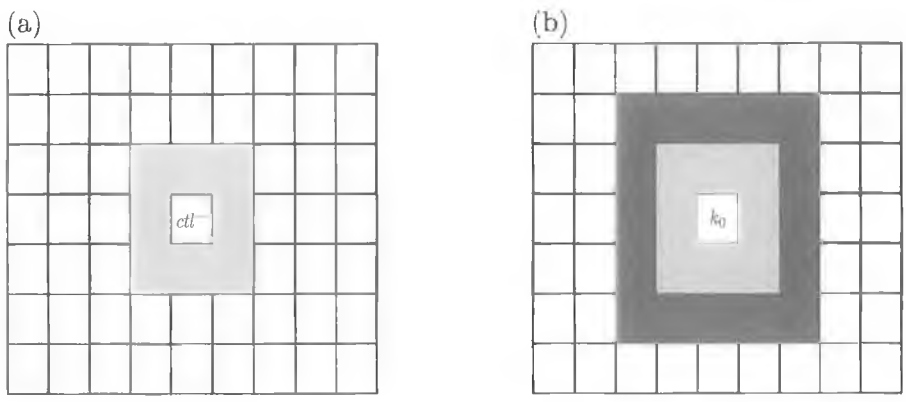


Figure 4.5: (a) A typical square lattice, with ctl_i^- at the centre and a *Moore* neighbourhood of radius $r = 1$ shaded in grey. In (b), the possible locations into which any cell k_i may move at each time step are shaded.

the antigen epitope and CTL clonotype in shape space), and a constant factor 0.036. At the end of the clonal expansion phase, when $\tau = 192$ (representing some four days of actual time), and $S = 1$, some 1000 clones have been produced. This is what would be expected in a healthy immune system (Janeway *et al.*, 1999). At the start of each simulation, shape space is characterised by two (non-zero) subpopulations, denoted CTL^- and APC^+ - representing the number of precursor cytotoxic lymphocyte and active infected antigen presenting cell *clones* respectively.

As time progresses, a further subpopulation emerges: CTL^+, representing activated cytotoxic lymphocyte clones. The CTL^+ subpopulation is recruited from the available pool of CTL^-. The process by which naïve cytotoxic lymphocytes are stimulated to become activated cytotoxic lymphocytes occurs ($CTL^- \nrightarrow CTL^+$) is described later in this section. In a healthy individual, the typical clearance rate of infected antigen presenting cells from the lymphatic system is of the order of 3-5 days, (a range commonly used in immunology texts, such as Janeway *et al.* (1999),p390).

At the end of primary response, the subpopulation of infected antigen presenting cells will tend to be eliminated ($APC^+ \rightarrow 0$). As the thymus ensures a supply of mature cytotoxic lymphocyte precursor cells into the

lymphatic system, our model follows this by ensuring that a non-zero naïve cytotoxic lymphocyte subpopulation exists at all times ($CTL^- > 0$). Finally, to support secondary immune response there is always some non-zero subpopulation of activated cytotoxic lymphocytes ($CTL^+ > 0$). These shape space subpopulation invariants are summarised in Table 4.1.

Table 4.1: Shape space subpopulation invariants.

Time	Invariant		
0	$CTL^- > 0$	$CTL^+ = 0$	$APC^+ > 0$
$0 < \tau < 475$	$0 \leq CTL^+ < CTL^-$	$CTL^- \rightarrow CTL^+$	$APC^+ > 0$
475	$CTL^- > 0$	$CTL^+ > 0$	$APC^+ \rightarrow 0$

4.2.2 Hybrid Processing

The state of the lattice over time is simply a sequence of random variables $\mathbf{x}^{(0)}, \mathbf{x}^{(1)} ..., \mathbf{x}^{(t)}$ defined on a finite space $\mathcal{X}$. The sequence is a *Markov chain* (Liu, 2001) as the value of $\mathbf{x}^{(t+1)}$ is dependent on its history only through its recent past $\mathbf{x}^{(t)}$. An update algorithm U is conditionally applied to each selected location in the lattice, depending on the occupant type - at each time step τ (U is defined in Appendix A).

Each location is sampled for update following a uniform distribution. For this reason, the update sampling step is called a *Monte Carlo* time step. At each Monte Carlo time step, U is repeatedly applied to the lattice such that the coverage degree Φ (the fraction of non-duplicate locations selected for update at each Monte Carlo time step) is in the range $0.99 \leq \Phi \leq 1$. This reduces the effect of the pseudo-random number generator as a source of errors from the simulation. Each simulation is terminated when $\tau = 475$, or after 10 days of simulation time has elapsed. This ensures that the full duration of the primary immune response is captured.

On the lattice, each immune cell k_i has two neighbourhoods: a *Moore* and an *extended Moore* (Fig. 4.5 (a) and (b), respectively), also denoted R_i and R_o, respectively. Neighbourhood sizes are $|R_i| = 8$ and $|R_o| = 16$.

With reference to Fig. 4.5 (a), any antigen presenting cell within the Moore neighbourhood is considered to be in contact with the ctl_i^-. In Fig. 4.5 (b), the possible locations into which any cell k_i may move at each time step are shaded; $|R_i|$ is shaded light grey, and the dark grey denotes $|R_o|$.

As part of the update algorithm U, each k_i recirculates within the real space, implemented as follows. First, R_i is examined in order to locate an unoccupied position into which the immune cell may move. If an empty cell is located within R_i, k_i will move into it with probability $P(inner) = 0.9$. If no space is available within R_i, R_o is searched for a free space. If a free space is located in R_o, k_i will move into it with probability $P(outer) = 0.7$.

If both R_i and R_o are occupied, then no movement of k_i will occur in this time-step. If a free location is found in R_i or R_o, the new coordinates of k_i are calculated and the cell is moved. The values chosen for $P(inner)$ and $P(outer)$ are subjective and reflect the concept of cell motion into *proximate* and *nearby* space, respectively. In the results presented here, both parameters are kept constant. Reducing $P(inner)$ would have the effect of restricting mobility and would therefore reduce the rate at which the real space lattice is sampled by the recirculating lymphocytes. This in turn would slow the clearance rate of infected antigen presenting cells.

The update algorithm U is summarised as Algorithm 1 in Appendix A. The combination of real and shape space as introduced above, results in a *hybrid* or *coupled* model. The activities within real, shape or hybrid space are defined as follows:

1. *Physical Space*

   (a) Each ctl^- and apc^+ recirculate inside the real space following the motion rules of U. The ctl^- are actively sensing the local environment for sign of infected apc^+.

   (b) Once a ctl^- and apc^+ have come into contact, the simulation transfers to shape space, in the sense that the following sequence of steps is initiated:

2. *Shape Space*

(a) The shape space distance between the two entities is calculated

(b) If recognition occurs, *clonal expansion* rate is calculated (S). If not, the ctl^- returns to its recirculation process (1a).

(c) The process of immunodominance emerges as $CTL^- \rightharpoonup CTL^+$ recruitment starts. At this point, the following steps arise in both real and shape space:

3. *Hybrid Space*

(a) The population concentration increases for each CTL clone stimulated, giving rise, after τ time steps to a concentration level $CTL_i^+ = e^{0.036\tau S}$

(b) Infected antigen presenting cells are removed from the real space system by recirculating activated effector CTL cells.

(c) At $\tau > 475$, effector cells undergo programmed death (apoptosis) and the primary immune response comes to an end.

4.3 Single-Strain Challenge

Within the real space, L_p is the length of one side of the square lattice, D_{ctl} is the density of the CTL cells and D_{apc} is the number of APC introduced into the lattice at the start of each simulation. The number of clonotypes in shape space is denoted Θ_{ctl}, while the number of epitope-bearing APC in shape space is denoted Θ_{apc}, which is the number of infection *strains*. In this simple experiment, $\Theta_{apc} = 1$. At each of the simulations, all parameters other than ρ are kept constant. In shape space, the largest value of ρ is known as ρ_{max}, the value of which is shown in Eqn. (4.6).

$$\rho_{max} = \sqrt{\frac{L_s^2}{\pi}} \tag{4.6}$$

ρ_{max} is a measure of the *maximum* crossreactivity of a given clone in shape space. Setting $\rho = \rho_{max}$ would be equivalent to having *every* CTL cross-react with every APC. Although this configuration is not explored further,

it is worth noting that it could represent a clinical condition known as auto-immune disease, where the immune system attacks both itself and invading pathogens without discretion. This disorder is, however, relatively rare (Orosz, 2001).

Of initial interest is the behaviour of the model as $\rho \rightarrow \rho_{max}$, which is expressed as the ratio of ρ to ρ_{max} as $\hat{\rho}$, and which takes values in the range $\mathcal{A} = \{0.71, 0.5, 0.25, 0.14\}$. $\mathcal{A}$ represents one possible set of values for $\hat{\rho}$ in order of decreasing crossreactivity. Any values could have been chosen for $\hat{\rho}$ that satisfy $0 \leq \hat{\rho} \leq 1$. From Eqn. (4.6), the value for ρ_{max} in our model is ≈ 28. The model parameter initial values are summarised in Table 4.2.

Table 4.2: Model parameters

Parameter	Definition	Value
D_{ctl}	CTL precursor density, real space	50000
D_{apc}	APC density, real space	5000
Θ_{ctl}	CTL clonotype density, shape space	100
Θ_{apc}	APC density, shape space	1
L_r	Length of side of real space	1000
L_s	Length of side of shape space	50
τ	Number of time steps per simulation	475
ρ_{max}	Max. shape space crossreactivity	≈ 28
$\mathcal{A}$	Range of values for $\hat{\rho}$	$\{0.71, 0.5, 0.25, 0.14\}$

4.4 Simulation Results

In the results presented in Fig. 4.6 (a) and (b), elements $\mathcal{A}_{1..4}$ are represented by $\star, \diamond, \circ, \bullet$ respectively. Fig. 4.6 (a) shows the density levels for the 10 most active shape space clonotypes when $\tau = 475$, arranged in rank order.

The values of $\mathcal{A}$ have generated two distinct clonotype activation patterns, which may be characterised as (i) sharp peak, narrow spread $(\circ, \bullet)$ and (ii) low peak, broad spread $(\star, \diamond)$. As crossreactivity decreases, more CTL cells are generated in order that immune function remains effective.

Reducing crossreactivity from 0.5 to 0.25 causes an almost 10-fold increase in the CTL produced, as well as a narrowing of clonotype response distribution from around 10 active clonotypes to around 4. This agrees with the expectation that fewer clonotype recognition events are occurring, but those that do arise are closer to exact matches. This can be thought of in terms of the disc of influence in shape space having smaller area such that those epitopes which do appear within this disc are consequently closer to the locus and therefore closer to an exact match, and in turn, cause higher stimulation because of this increased proximity.

In Fig. 4.6 (b), the corresponding clearance of infected APC from the lymphatic microscopic model is shown. Following from the assumption that shape space is *always* covered (Eqn. (4.4)), pathogen clearance from the lymphatic space is always achieved. The consequence of reducing crossreactivity from 0.5 to 0.25 is an increase in the time taken to clear APC by around 100 time steps (or 2 days of elapsed time). This is obvious in that if the probability of recognition is reduced and the rate of cell recirculation remains constant, matches will occur less frequently.

Increasing specificity results in the immune system expending more resources in terms of effector CTL in order to achieve the same end result, that of pathogen removal at the end of the primary response phase. The most successful clearance is obtained with highly crossreactive CTL ($\star$), where the APC half-life is around 10 time steps, or 5 hours of elapsed time.

This is somewhat in excess of known clinical removal rates in healthy individuals (Corbin and Harty, 2004). It is also clear that the two most successful clearance ranges ($\star, \diamond$) are characterised by a broad activation and low density levels in shape space. This is in keeping with van den Berg *et al.* (2001), who presented a mathematical model which demonstrated that a successful immune repertoire can be based on degenerate (ie, having wide-spread activation in shape space) low-affinity CTL receptors.

The results also demonstrate a further interesting property, that of *immunodominance*; this is the process whereby a small number of specific clonotypes are responsible for clearing the infected cell challenge. In Fig. 4.6 (a), the around 90% of the response is mounted by the only 1 subset of clono-

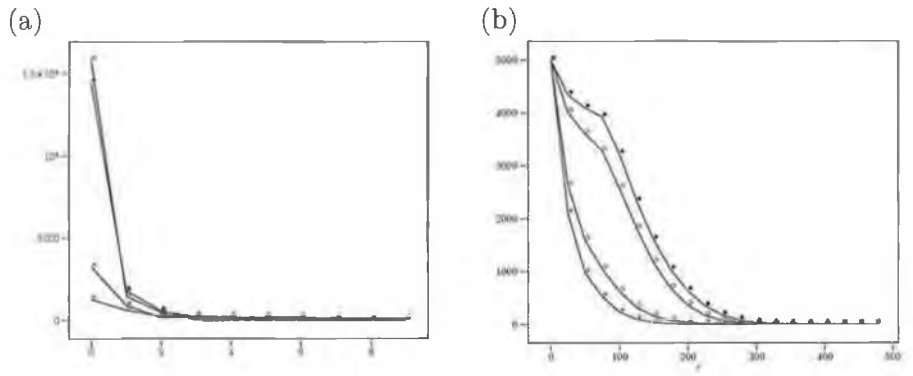


Figure 4.6: The values of parameter set $\mathcal{A}$ have resulted in two clonotype activation patterns characterised by high concentration, narrow distribution $(\circ, \bullet)$ and low concentration, broad distribution $(\star, \diamond)$. In (a), concentration levels of the ten most active clonotypes are shown (y-axis), and in (b) clearance rates of infected APC (y-axis) are shown over 500 time-steps.

types $(\circ, \bullet)$. Whereas high crossreactive response is characterised by lower immunodominance levels. An immune response which commits a single-clonotype only may have a reduced chance of dealing with an epitope mutation (such as that seen in HIV), compared to a response which is characterised by multiple clonotype activation.

4.5 Chapter Summary

The starting point for the work presented in this chapter is to assume that the immune repertoire has not been exposed to any prior pathogen, that it is uninitialised. Therefore, immunological memory or reinfection was not modelled. The objective has been to demonstrate the benefit of a model which combines two usually separate formalisms, that of immune system shape space and local (or real) space microscopic simulation.

Initial results from the model suggest that a low-affinity CTL clonotype shape space provides the most efficient APC removal. The findings are in agreement with work presented by other authors, notably van den Berg *et al.* (2001). Two distinct clonotype activation patterns were noted in the

results, observing (i) high peak, narrow distribution and (ii) low peak, broad distribution. For all other model parameters being unchanged, it was shown that reducing crossreactivity causes significantly different evolution of the shape space and correspondingly poorer APC removal.

Crossreactivity is one of the most crucial regulating parameters within the immune system. It has also been seen how immunodominance naturally emerges in shape space; the concentration and distribution of immunodominance within the CTL clonotype space is of particular interest, as a correlation between disease clearance rates and $\tilde{\rho}$ can be found.

In this chapter, the earlier models have been extended in order to incorporate shape space. The principle of *downward causation* was observed in that changes to the crossreactivity of shape space clonotypes resulted in alteration of the dynamics of the microscopic model. This is one of the key requirements for demonstrating the principle of emergent behaviour. The hybrid model presented here has another crucial characteristic namely, that of providing a means to study self-organisation.

In constructing a hybrid model, the capability to study *emergent patterns* in shape space becomes available, so that visualisation of shape space is then feasible. This is of major interest because the hallmark of many self-organising systems their tendency to exhibit visible structures or patterns. The topic of model visualisation is the subject of the next chapter.

CHAPTER 5

Pattern Formation in Shape Space

5.1 Introduction

In the previous chapter, an immune system model was developed which had the capacity to represent immune repertoire and its development. Such development includes transition from the uninitialised (or uninfected) state to the activated (or stimulated) state. As mentioned, the shape space formalism is ideal for modelling the immune repertoire, not least because, while shape space was originally proposed as a means to model binding between clonotype and epitope, it can also be refined to include clonotype *state*, such as precursor, activated, effector, memory and dead. Thus, knowing the state of clonotypes in shape space (and not just their presence or absence from a specific co-ordinate in the space) provides a fuller picture of the health of the immune repertoire, and by extension, of the immune system overall.

One limitation of the model developed thus far is that although a system-wide (or global) perspective was developed, no method was identified to study the *emergent behaviour* which may be present in the model. Emergent behaviour is the process whereby global features or structures emerge

naturally from a local system in which such features are not merely aggregates of microscopic interactions. However, emergent behaviour can be a difficult property to unambiguously identify, and therefore emergent behaviour is often characterised by the process of systemic *self-organisation* where the higher-level components of the model take on non-random spatial structures.

A widely cited example of a CA model possessing self-organisation is Conway's "Game of Life" (Gardner, 1970; Berlekamp *et al.*, 1982). A screen shot of this CA model is shown in Fig 5.1.

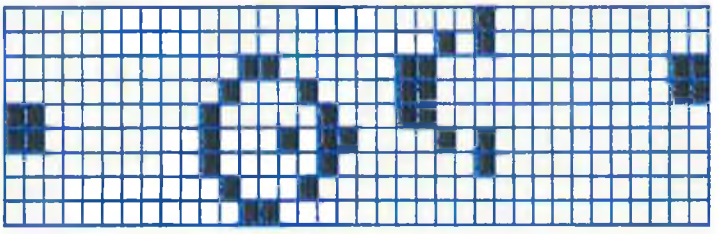


Figure 5.1: Screen shot of "The Game of Life". In the centre, two patterns of self-organisation.

Self-organisation becomes apparent through visual representation of system components and patterns of change which occur over time. Many cellular automata (CA) models implementing even simple rules will be capable of astonishingly varied pattern formation when viewed in an extended spatio-temporal context. For example, Wolfram (2001) has done much work in classifying and analysing some of the more interesting deterministic CA, where such classification typically specifies self-organising patterns, typically known as *attractors*. Very recently, some extended CA models of biological systems have been demonstrably successful in modelling two-dimensional visual characteristics of capillary development (Merks *et al.*, 2004), as well as in investigating the efficacy of current and potential therapies of non-Hodgkin's lymphoma (Ribba *et al.*, 2004). In both cases, the models displayed interesting emergent self-organising features.

In this chapter, the theme of pattern formation in shape space is devel-

oped as follows (i) a simple two dimensional visualisation of shape space is presented, and (ii) patterns and clusters which emerge in the shape space of activated CTL clonotypes, in response to single strain infections, under varying crossreactive regimes, are analysed.

Visualisation is a burgeoning field of scientific computing which seeks to provide multi-dimensional graphical representation of underlying models or data. Such representations can typically be subject to manipulation such as rotation, transformation, animation and so on. Complex data sets are transformed into visually meaningful 2-D or 3-D realisations in order to improve understanding of the underlying data, and to aid in its interpretation. Recent visualisation applications have developed enhanced virtual reality environments within which complex and often dangerous real world procedures can be practised, including those as extreme as facial surgery (Everett *et al.*, 2003). Hagen *et al.* (2003) identified a further category of visualisation known as *Intelligent Visualisation*; here, a multi agent model simplifies the monitoring and tuning of many of the parameters normally controlled by the user, such as frame-rate, rendering, library selection, load-balancing and distribution.

An excellent review of techniques in visualisation programming (using Open GL) can be found in McReynolds and Blythe (1999), which includes a section on the representation of scalar field data. A scalar field is a single-component element which can take on values only in a pre-defined range (or *scale*), and the data sampled from shape space is a form of scalar field component. In fact, the shape space scale is restricted to the set $0, 1, 2, 3$, denoting *no cell*, *precursor CTL*, *activated CTL* or *infected APC*, respectively. Typically, the scalar field data being visualised is not continuous (as the original scalar field is), but is composed of a set of discrete sampled values. The sample spacing may be regular, forming a grid, or the sample spacing might be irregular, with varied spacing between samples values.

Several previous models have presented visualisations of shape space, with the work of De Boer (De Boer *et al.*, 1992) notable in this respect. The model produced some interesting one- and two- dimensional spatial patterns in response to an analytical ODE model of B cell stimulation, and as such the

model was an example of *vector field* visualisation. As opposed to a scalar field, visualisation of a vector field (in which each point is an n-dimensional vector) often arises in applications such as computational fluid dynamics, and typically represent the flow of a gas or liquid.

The results in De Boer *et al.* (1992) were noteworthy in that they demonstrated shape space *state* (such as uninitialised, infected and suppressed) as well as cell concentration levels, by means of grey-scale shading technique. Strictly speaking, cell concentration levels in shape space was not originally included in the early research on shape space, possibly because such information can lead to visual *clutter*. Additionally, the model did not exhibit any characteristics of emergent behaviour or self-organisation, and therefore the patterns were simply visual representations of an analytical and deterministic solution. Zorzenon dos Santos and Coutinho (2001) proposed a cellular automaton model of HIV in which four distinct patterns of self-organisation were visually represented within a square lattice and modelled over a time-scale of 5,18,25 and 200 weeks.

In this chapter, a simplified visualisation of a 2-dimensional shape space is presented. This is in order that an effective representation of CTL activation patterns, especially in respect to distribution of stimulation, may be studied. Section 5.2 discusses the method by which shape space is reported and rendered as a bounded two-dimensional figure, and as such serves as an introduction to the steps involved in visualisation of shape space. Section 5.3 introduces and discusses the patterns which emerge when the model is exposed to a single-strain viral infection, under various crossreactive parameter regimes. Section 5.4 discusses the relevance of the results and suggests some potential biological implications, and concludes with comments on the next steps for model development. Use of previously established notation (introduced in Chapter 4) is continued in this chapter.

5.2 Reporting and Rendering Shape Space

In keeping with the design of most software systems, the model presented in this dissertation supports various levels of *logging* (or reporting). Al-

though there exist predefined logging services (such as the Java Logging Service `java.util.logging`), the reporting carried out by most applications is highly idiosyncratic and ad-hoc in nature, ranging from the verbose to the cryptic, the terse to the self-explanatory. In order to report useful information on the internal data structures of the model during run-time, the following five points need to be addressed.

1. Report Frequency - how often is the data collected from the simulation?
2. Report Format and Data Specification - what is the best structure for the data on the disk?
3. Data Pre-processing - what conversion does the raw data require before analysis can begin?
4. Data Analysis - what tools are required to analyse the data?
5. Data Representation - what visual methods are best to explain the data collected?

In most of the figures presented as output from the models in this thesis, each of these five points is likely to be addressed differently, depending on the data being reported and the analysis required. For example, when reporting on mean density levels of CTL cells (see, e.g., Chapter 3, Fig. 3.3), the pre-processing step (3.) consists of a C-program reading each of the input files from each simulation and calculating the mean density levels for each time-step. In contrast, the pre-processing stage for reporting activated CTL clonotypes in shape space is performed by a script executing within the R Statistical Software environment which converts raw files to internal R matrix structures. Although this model is flexible enough to execute *ad-infinitum*, the simulation is always implemented as an *event terminating simulation* (see, e.g., Law and Kelton (2000), pp 502-503), in that there is always some signal or event which as a consequence will end the simulation.

5.2.1 Output Data Analysis

The visualisation of shape space is a sequential, non-iterative activity in that the model reports the space configuration twice: on initial start-up of the simulation, and *just after* the clonal expansion phase has finished ($\tau = 192$). This is a logical point at which to sample the space, because at any time after this, apoptosis would begin and effector cells start to die. Therefore, the second sample of shape space captures the clonotype distribution at the peak of the expansion phase. The process for visualising shape space is summarised in the following six steps:

1. Simulation Starts
2. At $\tau = 0$ the initial (uninitialised) space is sampled to report the CTL^- distribution. The data is written to file `ss_init_ctl.dat`, in the format: `line 1 <x-coord>` and `line 2 <y-coord>`
3. At $\tau = 0$ the initial (uninitialised) space is sampled to report the APC^+ distribution. The data is written to file `ss_init_apc.dat`, using the file format shown above.
4. At $\tau = 192$:
   - iterate over real space lattice and write each location shape space co-ordinate to one of two files based on the cell type (either CTL^+ or APC^+), called `ss_final_ctl.dat` and `ss_final_apc.dat`, respectively, The file format is shown above.
5. Simulation Stops
6. The R statistical tool is started and the following script is executed for each raw file:
   (a) each x, y-coord in each of the data files is mapped to a row-col index for an $L_s \times L_s$ size matrix M. At the intersection of the row-col index, a number is placed indicating the cell occupant

type (CTL^-, CTL^+, APC^+), which is determined by the file being read at the time. An example of a typical output file which logs shape space co-ordinate information for each real space lymphocyte in shown in Fig. 5.2.

(b) the matrix M is then used as a parameter to the R function `points()`, with `X11()` enabled, and the image is written as an Encapsulated Postscript (.eps) file.

A sample of the raw data file structure is shown in Fig. 5.2. The first line is the x-coordinate part, and the second line the y-coordinate part. A real file would contain many more columns than the 20 shown here (but always two lines, for each dimension of shape space). There are two $\langle 06 : 05 \rangle$ pairs (at positions 2 and 14), indicating that *two* real space ctl^- are of the same shape space clonotype.

The above process of visualising shape space requires an exhaustive search of the real space lattice in order to record the x, y-coordinates of each lattice occupant. At the end of clonal expansion, there may be $\sim 10^4$ cells which share the same clonotype, but the procedure described above requires *all* duplicate clonotype data be written. This limitation does not impact on performance time very much, as shape space is only sampled twice during the simulation, at the beginning, and at the end of clonal expansion (with $\tau = 192$).

```
16 06 17 15 06 25 36 46 47 06 00 37 31 06 11 32 34 12 34 42
19 05 30 24 14 30 41 11 43 27 10 02 03 05 01 10 16 04 05 12
```

Figure 5.2: A truncated sample of a raw file containing shape space co-ordinates of the CTL^- clonotypes.

5.3 Pattern Formation

In the results presented here, the dynamics of response when the immune system is challenged by single strain pathogens is analysed. Therefore, the

shape space APC^+ distribution is confined to one single point in space (and is shown by the asterisk, $*$). At time $\tau = 0$, $apc^+ = 5000$, representing infected antigen presenting cells entering the lymphatic compartment, and in turn, triggering primary response. The parameters of the simulation are: $D_{ctl} = 50000$, $D_{apc} = 5000$, $\Theta_{ctl} = 1000$ (without replacement) and $\Theta_{apc} = 1$. The value for Θ_{ctl} is selected in order that shape space fully covered, especially for small values of $\hat{\rho}$.

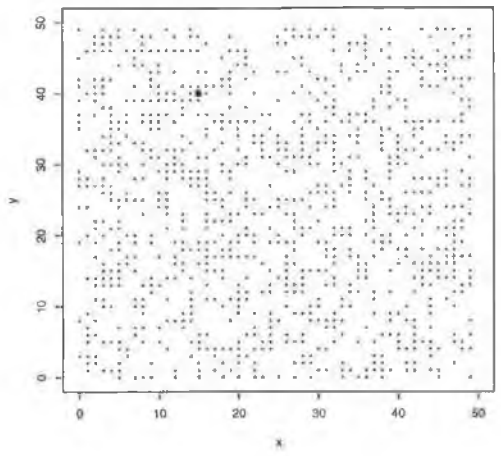


Figure 5.3: Two dimensional shape space showing the initial distribution of CTL^- and APC^+ cells (circle and asterisk, respectively) at time $\tau = 0$.

Fig. 5.3 shows a typical initial shape space at time $\tau = 0$ (for simplification, CTL^- is denoted by a smaller point compared to both the APC^+ and CTL^+). At this time, one antigen presenting cell is shown, and no recruitment from $CTL^- \rightharpoonup CTL^+$ has occurred. Biologically, this state represents the point at which an antigen has been taken up by an antigen presenting cell (such as a dendritic cell), but has not yet been detected by the recirculating cytotoxic lymphocytes in the lymphatic compartment. The shape space at this time has not been exposed to any previous infection.

Two primary conditions determine how the initial shape space distri-

bution will evolve: (i) apc^+ and ctl^- must come into contact in real space (*mobility*) and, (ii) the distance between clonotype and epitope in shape space must be within the crossreactive threshold ρ. The conditions controlling (i) are not altered during simulation executions here. With respect to (ii), although $\hat{\rho}$ may be assigned any value in the continuous range $0 \leq \hat{\rho} \leq 1$, it is here restricted to six representative values, based on the requirement of testing extreme values $(\mathcal{A}_1, \mathcal{A}_6)$ and a selection of intermediate values:

$$\mathcal{A} = \{0.928, 0.6, 0.39, 0.25, 0.164, 0.107\} \tag{5.1}$$

The effect of declining crossreactivity in shape space is analysed from four perspectives:

1. Clearance rate of the apc^+ subpopulation.
2. Density levels of CTL^+ and ctl^+.
3. Activation distribution pattern as the CTL^+ subpopulation is recruited from the CTL^- pool.
4. Efficiency of response: the measure by which the apc^+ challenge is responded to *in proportion* to the threat posed, so that immune resources are not spent unnecessarily.

Each simulation $(\mathcal{A}_i)$ is repeated 30 times, and the results are averaged. Shape space is confined to 2.5×10^3 positions while real space is modelled as a two-dimensional array with 1×10^6 positions. These small sizes represent exploratory analysis with modest computational resources and should clearly be extended. A more numerically realistic size would require a real space capable of modelling some 10^9 lymphocytes. Representing a high-order system in terms of a low-order one can result in finite size effects, with particular impact when the model exhibits first- or second- order phase transitions (Landau and Binder, 2000). Phase-transitions in immune system models would be points at which (for example, in HIV), immune collapse occurs due to destruction of helper cells, and the host is in danger of death by opportunistic infection.

In the results which follow, the consequences of finite size effects are not immediately relevant. With biological practice, most *in vitro* experiments act on numerically reduced cell populations, and the results are taken merely as indications of possible *in vivo* outcome, and are therefore highly qualitative.

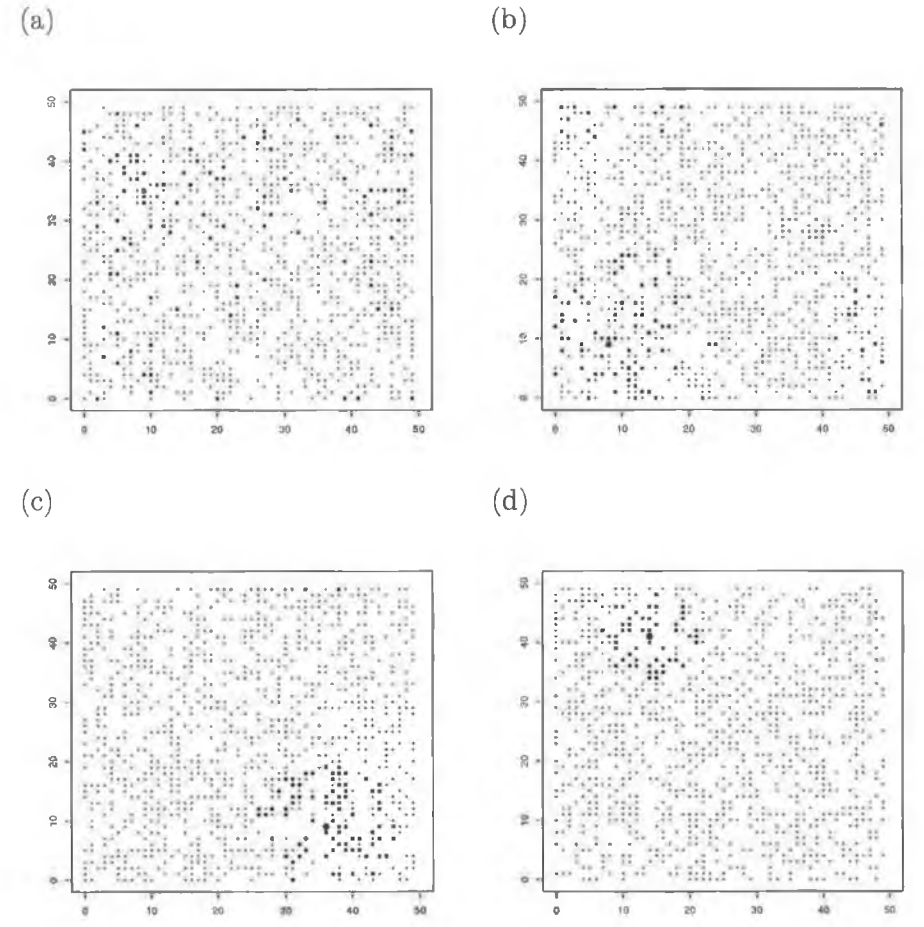


Figure 5.4: Four shape space activation patterns at time $\tau = 300$, showing the effect of declining $\hat{\rho}$ with values drawn from $\mathcal{A}_{1,..,4}$.

Fig. 5.4 shows the activated clonotype distribution pattern in shape space for each value drawn from $\mathcal{A}$, at the end of the simulation ($\tau = 300$), with

(a) to (e) representing $\mathcal{A}_1$ to $\mathcal{A}_4$, respectively [1]. For clarity, Fig. 5.5 shows the activated clonotype distribution pattern in shape space for the values in $\mathcal{A}$, (a) and (b) representing $\mathcal{A}_5$ and $\mathcal{A}_6$, respectively.

(a) (b)

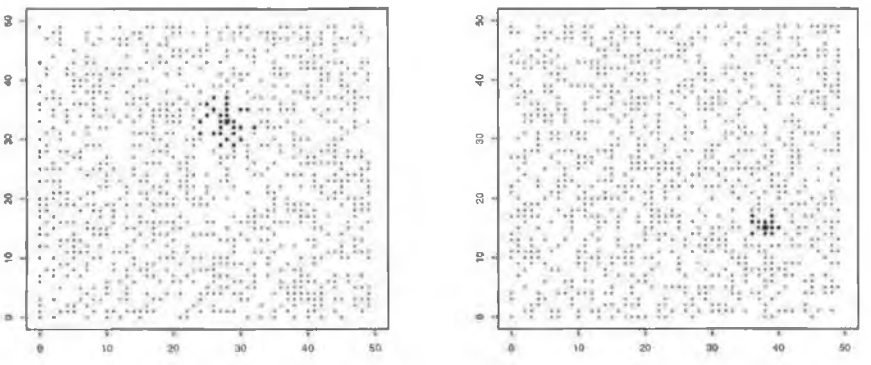


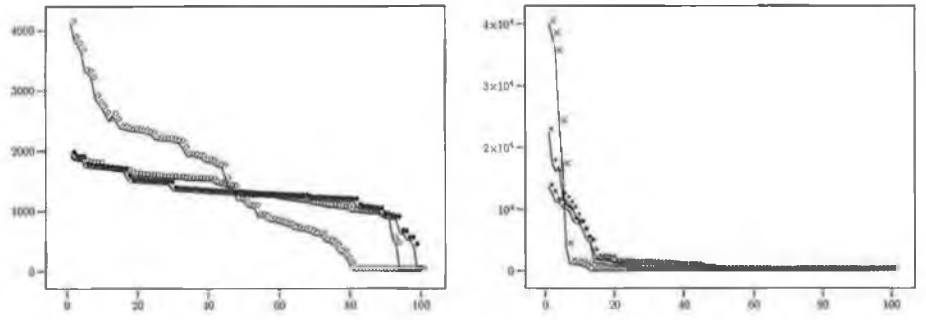
Figure 5.5: Two shape space activation patterns at time $\tau = 300$, showing the effect of declining $\hat{\rho}$ with values drawn from $\mathcal{A}_{5,6}$.

As crossreactivity declines, average distance of CTL^+ from the initial point of stimulation (APC^+) declines consistently and this is certainly in keeping with expectations. It is important to observe that the number of CTL^+ in (a) and (b) is almost identical ($\approx$ 100). At first this may seem paradoxical. After all, it is assumed that a highly crossreactive clonotype repertoire would become activated by every infected antigen presenting cell it comes into contact with. This should result in greater clonotype activation and, conversely, a reduction in crossreactivity of some 35% must surely reduce the concentration of activated immune lymphocytes. The explanation becomes clear when the coverage constraint of Eqn. (2.1) is considered. In the range $\mathcal{A}_{1,2}$ the highly mobile nature of the recirculating real space cells ensures that *sufficient* APC^+ and CTL^+ come into contact, which in turn compensates for the declining crossreactivity. This is intuitive in that

[1] As in real space, periodic boundary conditions are enforced in shape space, the results of which appear in the form of wrap-around of CTL^+ in Fig. 5.4 (a), (b) and (c)

if the sampling rate of the immune lymphocytes of their environment is high enough then, theoretically, a $\hat{\rho}$ of *just greater than* 0 would be sufficient to sustain normal clearance rates. However, realistically, such a sampling rate would be beyond the capability of the immune lymphocytes.

In the results shown in (a) to (d), the reduction of $\hat{\rho}$ is not accompanied by a commensurate reduction in activation levels. This supports the assertion that recirculation rates are high enough to ensure significant levels of lymphocyte activation, and as such, there is non-linear relationship between $\hat{\rho}$ and activation concentration. As can be seen, narrowing crossreactivity causes a sharp clustering of activated lymphocytes in the neighbourhood of the APC^+. When the activation pattern for $\mathcal{A}_6$ is examined (Fig. 5.5 (a)), it is clear that the concentration of CTL^+ has dropped significantly - to $\approx$ 12 (a decline of 88%). At this point, the recirculation rate is not sufficiently high to compensate for declining crossreactivity, and as a result, a commensurate decline in activation levels follows.



(a) (b)

Figure 5.6: Concentration of CTL clonotypes cells across the 100 most active shape space clonotypes. For clarity, (a) shows $\mathcal{A}_{1,2,3}$ and (b) shows $\mathcal{A}_{4,5,6}$. $\mathcal{A}_{1,\ldots,6}$ are denoted $\bullet, \circ, \diamond, \star, *, \times$, respectively. Only mean values are shown, from 30 simulation runs.

Fig. 5.6 shows the real space concentration levels for the 100 most active clonotypes in shape space. For clarity, the figure is shown in two parts, (a) and (b), which represent the results for the parameters $\mathcal{A}_{1,2,3}$ and $\mathcal{A}_{4,5,6}$, respectively. This figure provides insight into how the activation patterns observed in shape space are related in terms of concentration levels in real

space. There are two main CTL^+ concentration signatures: (i) low frequency and broad distribution ($\mathcal{A}_{1,2,3}$), and (ii) high frequency narrow concentration ($\mathcal{A}_{4,5,6}$). Low crossreactivity has resulted in a consistently lower overall CTL^+ activation spread at the end of clonal expansion, and conversely, higher crossreactivity has resulted in a narrower activation spread of CTL^+ far larger concentration of effector cells being produced in real space.

An immune response cannot *a priori* know the numerical strength of each pathogenic challenge (where strength is measured by the number of infected antigen presenting cells entering the lymphatic compartment at a given time). This leads to a risk and tradeoff in terms of number of lymphocyte cells produced and time spent in clonal expansion in that enough effector cells must be activated at the end of the process to effect clearance of infected antigen presenting cells, but a highly agonist (ie, stimulatory) challenge could result in over-production of effectors. Segel and Bar-Or (1999) and Segel (2000) have suggested that this balance is achieved by a form of feedback, where lymphocyte cells will look for 'scalps' (ie, dead antigen presenting cells) in order to continue to spend time in clonal expansion. In the absence of such stimulatory signals, it is suggested, the clonal expansion phase will end and effector cells will start to die. This theory goes some way to explaining how the immune system knows when to end the cell reproduction phase.

However, a different viewpoint has been proposed by Corbin and Harty (2004) in a clinical context. Murine experimentation has demonstrated that even when the initial stimulation (in the form of antigen presenting cells) are removed from the lymphatic system once primary response has begun (by manual intervention), over time, *the concentration of lymphocytes produced remains broadly comparable to the case where the stimulation is not manually removed.* This approach is the one adopted here. In this model, if the stimulation were to be removed, the lymphocytes committed to the clonal expansion phase would continue to divide and activate. This approach has also been adopted in a number of computational immune response models (Chao *et al.*, 2003, 2004).

At the end of clonal expansion, there should be a population of activated effector T cells *in proportion* to the initial threat (which in these simulations is 5000 apc^+). It is evident from Fig. 5.4 that although $\mathcal{A}_{1,2}$ resulted in the broadest [2] shape space CTL^+ activation distribution, real space concentration levels are significantly more proportional to the infected antigen challenge than the other parameter values of $\mathcal{A}$. Fig 5.7 (a) shows the concentration of infected antigen presenting cell to activated effector lymphocyte, in order to convey the increasing ineffectiveness of higher levels of effector lymphocyte production. As the crossreactivity declines, the total number of cells generated increases by some 50% when comparing $\mathcal{A}_1$ to $\mathcal{A}_6$.

Fig. 5.7 (b) shows the clearance rate of infected apc^+ from the lymphatic compartment during primary response. As crossreactivity declines, the rate at which apc^+ is cleared varies considerably. For both sets of parameter values $\mathcal{A}_1$ and $\mathcal{A}_2$, a full clearance of apc^+ is achieved at the end of primary response - with $\mathcal{A}_1$ clearing more quickly. The figure shows that as crossreactivity declines, *sensitivity* to antigen presenting cell increases, and the response is characterised by a larger perturbation from the equilibrium point (of no activated effector cells).

A further feature emerges when comparing the clearance rates of $\mathcal{A}_1$ and $\mathcal{A}_2$. In shape space, (Fig. 5.4 (a) and (b)), the activation levels are almost identical, significantly, the main difference is the pattern of clustering (a broad spread as opposed to a more focused one) which has emerged around the central stimulation point of the initial APC^+ infection. This implies that there is a qualitative difference between distribution patterns. A broad distribution with high mean distance from the centre results in greater efficiency in APC^+ clearance than does one with smaller mean distance from the centre.

Furthermore, although high crossreactivity results in large CTL^+ production (Fig. 5.6 (a)), clearance rates still degrade from $\mathcal{A}_{4,5,6}$ leading to an immune response which, although marked by high lymphocyte production, is, in effect, wasting the lymphocyte clones produced, as almost no APC^+ are cleared (Fig. 5.6 (a) - square, triangle and circle, respectively). These

[2] Measured as the mean distance of CTL^+ from the APC^+.

wasted resources represent a double threat, in that (i) valuable response time is wasted in clonal expansion for a largely ineffective activated lymphocyte pool, and (ii) the excessive ctl^+ represents wasted resources in cell generation.

(a) (b)

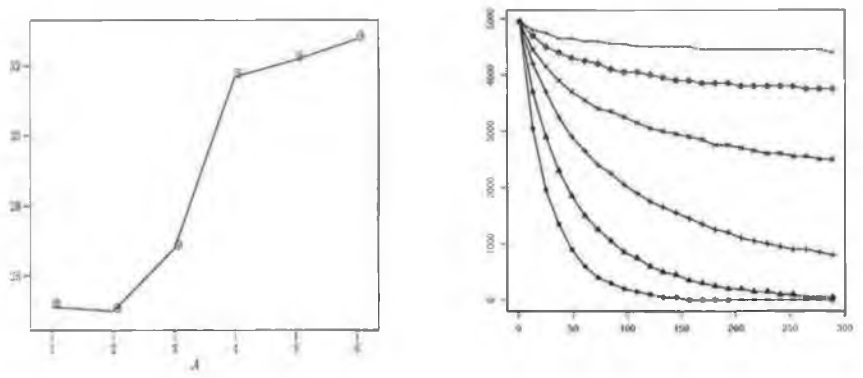


Figure 5.7: Concentration of CTL effector to APC for each of $\mathcal{A}_{1,\ldots,6}$ (a), and (b), the clearance rate of infected antigen presenting cells from the lymphatic model after 300 time steps, where $\mathcal{A}_1 = \circ$, $\mathcal{A}_2 = \triangle$, $\mathcal{A}_3 = +$, $\mathcal{A}_4 = \times$, $\mathcal{A}_5 = \diamond$, $\mathcal{A}_6 = *$.

The appearance of immunodominance is evident at the end of primary response. At this time ($\tau = 300$), with the clonal expansion phase complete, the concentration and distribution of ctl^+ and CTL^+ (Fig. 5.6 (a) and (b)) represent preferentially stimulated armed effectors. In $\mathcal{A}_5$ and $\mathcal{A}_6$, around 90% of the immune response is concentrated against three CTL^+ clonotypes.

A response of this nature is deficient in several ways. In the case of a mutating pathogen (in effect, the *migration* of the clonotype from one *shape space* "position" to another within a time period of a few days of simulated time), a high concentration response targeted against one epitope or a few epitopes could potentially fail to respond in sufficient time in order to eliminate the mutated challenge. Where the pathogen is also proliferating, the time period during which the pathogen remains undetected is crucial in affecting the course of the infection. A high concentration response risks

wasting immune resources since a large fraction may be unnecessary. This is the case in delayed-type hypersensitivity reactions which can take 24-72 hours to appear (see, e.g., Janeway *et al.* (1999), pp 66).

With parameter set $\mathcal{A}_1$ and $\mathcal{A}_2$, the response pattern is different (but in agreement with expected response patterns, Fig. 5.4). The response to single pathogen strain results in over 100 distinct CTL clonotypes being activated. It was found that spreading the response across a broad range of CTL clonotypes (seen in $\mathcal{A}_1$ and $\mathcal{A}_2$) ensures that real space sampling rate is not a decisive factor in ensuring an effective response. In related work, Kirschner *et al.* (2000) have shown how a model of HIV in which altered recirculation dynamics of lymphocyte CD+4 T cells (causing them to home to the lymphatic compartments from the blood system) was able to explain the success of HIV.

5.4 Chapter Summary

In this chapter, a novel method was developed to study the emergence and dynamics of CTL clonotype diversity in an immune system simulation simulation. In particular, the model constructed has demonstrated the following properties: (i) shape space activation and its dependence on crossreactivity (ii) clearance rates are dependent on activation distribution (distance from shape space locus of the APC^+) (iii) high crossreactivity is more efficient in terms of infected cell clearance rate and lymphocyte production levels (iv) increased recirculation rates can compensate to some extent for declining crossreactivity. This model raises the following two experimentally testable hypotheses.

1. Immunisation is likely to be more effective when a spread of memory effector cells are activated. This may require several stages of stimulation of the immune response with genetically varied strains of the same viral pathogen. Oxenius *et al.* (2001) have noted that mutation has been identified in antigen epitopes which improve clonotype recognition without altering specificity. Such mutant antigens are able to

target a specific clonotype and deliver an enhanced activation signal, which in turn can lead to up to a 40-fold increase in effector function (for example, cytokine production). Such mutant antigens will be of use for boosting immune responses to specific antigens. The authors have also shown that such clonotypes respond to lower concentrations of the mutant antigen.

2. An increase in crossreactivity can contribute to effective immune response. Brehm *et al.* (2002) have shown that it is possible to stimulate such crossreactivity *in vitro*, and that the resulting clonotype activation pattern is beneficial in controlling heterologous viruses.

The model presented here leads to the following conclusions. Preferred types of activation patterns in shape space give more efficient and effective real space clearance rates than do others. As the mean distance between CTL^+ and APC^+ declines (broadly *regardless of the mean activation concentration*), the clearance of infected antigen presenting cells from the real space model becomes progressively less efficient. In agreement with the clinical findings of both Mason (1998) and van den Berg *et al.* (2001), the shape space of the model exhibits high crossreactivity which results in an immunodominance configuration which enhances antigen clearance from the real space.

Furthermore, the results here suggest that the "lock-and-key" formalism of theoretical immunology would require prohibitively vast repertoire sizes, and that effective pathogen removal is more likely to be carried out by a degenerate distribution of activated clonotypes.

A process has been presented by which shape space may be visualised, and patterns observed for clonotype activation which would not be apparent from a purely real space simulation. Visualisation is a crucial step in classifying emergent behaviour and self-organisation inherent in complex systems. The strategy in this chapter has been to challenge the immune response with a single-strain infection, and as such, it is limited in scope.

A more realistic study would involve exposing the model to multiple single-strain *and* multiple heterogeneous strains over an extended time pe-

riod in order to examine the dynamics of crossreactive memory cells and non-memory cells, and hence to develop a theory of how self-organisation can be modelled in the context of the immune system. This is the subject of the next chapter.

CHAPTER 6

Emergent Networks in Immune System Shape Space

In the previous chapter, a process for visualising shape space was presented as part of a hybrid real-space, shape-space model of a simplified immune system at end of the primary response phase. Although only a subset of immune entities was implemented, the model was shown to be sufficiently powerful to demonstrate emergent activation patterns of CTL clonotypes when subject to a single strain pathogen challenge.

The benefits of this approach were twofold: (i) visualisation of the CTL activation patterns in shape space demonstrated how CTL clonotypes clustered around the stimulant APC, following varying degrees of cluster diameter, dependent on the crossreactive ratio $\hat{\rho}$. Importantly, these clusters emerged naturally from the hybrid model, and were not pre-programmed in any sense, and (ii) the crossreactive ratio also induced varying degrees of infected cell clearance efficiency, with high crossreactive, low affinity CTL repertoire being the most successful (a finding supported by van den Berg *et al.* (2001)).

Taken together, these two points lead to an important conclusion. Activation clusters in shape space *can be used to analyse and predict the rate*

at which infected antigen presenting cells are removed from the real space compartment. In turn, this conclusion implies that the diameter of such clusters has a bearing on infection duration and classification. These can be classified as healthy (infection cleared within 6-8 days), chronic (infection lingers beyond 10 days), or acute (failure to significantly reduce infection levels).

However, the work developed in Chapter 5 lacked a realistic temporal context. Only once-off infection events were studied and only at the end of clonal expansion. Clearly, no real immune system is this naive (Welsh and Selin, 2002). Furthermore, the model presented in the previous chapter did not reflect the *dynamics* of CTL cluster creation and development. This is particularly relevant as the clusters studied in the previous chapter were all visualised once-off, and therefore, it is reasonable to ask if these clusters were dynamical structures (which undergo diameter growth, contraction) or were more static structures (and subject to little further development after initial appearance).

To address this important question requires the model to execute for a longer period of time (beyond clonal expansion), to be stimulated by more than one infection (in order to study the impact on cluster properties), and crucially, to be visualised more regularly over the simulation life-time in order to analyse cluster change. Clearly, some underlying dynamical process must exist in order for clustering to arise in the first instance, and it is the aim of this chapter to identify a theoretical framework to support the emergence and development of activated CTL clusters in shape space.

Here, the model is extended to enable the study of emergent principles of immune system CTL repertoire, incorporating the stochastic cellular automata model first presented in Chapter 3. An extension to the hybrid model presented in Chapter 4 is developed such that each activated CTL clonotype and viral epitope are represented as nodes in shape space, and edges between nodes models the affinity or clearance pressure applied to the antigen presenting cell bearing the target epitope. When the model is repeatedly exposed to infection by heterologous viruses, a distinct topology of the network space emerges.

This topology can be used to further understanding of immune system function in that nodes representing memory cells, form a crucial backbone of infection regulation, and disruption to these nodes results in an impairment of response akin to an immune system which never learns and never remembers. In turn, disruption to the memory cell nodes results in infected cell clearance reminiscent of chronic infection.

6.1 Complex Network Approach

Understanding of large scale systems of many particle interactions can often be achieved by modelling the relationship between the constituent parts in the form of a network. This approach has been successfully applied to such diverse areas as information propagation (Rosvall and Sneppen, 2003), movie-actor collaboration and protein folding, to name but a few (see Albert and Barabasi (2002) for a comprehensive review).

Recent work has characterised the dynamics of the world wide web as a directed graph (Tadic, 2001), while Albert *et al.* (2000) have have shown that the world wide web is remarkably robust when subjected to random node elimination (*robustness* is the degree of network function under node elimination). In both cases cited, global properties of the system (such as robustness) were analysed by studying the topology of web pages and web links.

In evolving networks, non-random organising principles lead to the study of *emergent topology* which may provide insight into network-wide characteristics. On the other hand, scale-free networks are characterised by the principal of *preferential attachment*. The probability of a new node i being connected to another node l, increases with *connectedness* of i:

$$P(i) = k_l / \sum_{l}^{N} k_j \tag{6.1}$$

Additionally, a network can apply a parameter (such as *cost*, *capacity*, *pressure*, *usage* and so on) and (sometimes) direction to each of the edges of the graph.

In the work conducted by Barabasi and co-workers, the scale-free networks identified all had the property of the mean degree ($\langle k \rangle$) in the range $2 \leq \langle k \rangle \leq 3$. This range for $\langle k \rangle$ turns out to be the same in all scale-free networks, For example, in Jeong *et al.* (2000), $\langle k \rangle = 2.2$, where again, $\langle k \rangle$ is the average number of node edges.

The apparent wide-spread existence of scale-free networks in biological systems may help to explain why such systems are usually extremely robust. The elimination of random nodes and edges (due to, for example, mutation), has little impact on the network topology - indeed, the average diameter of the network would remain unchanged. Mutation of a pathogen would have the effect of *shifting* the epitope co-ordinates in shape space. This shift would subsequently have an impact on network topology in that some edges would be broken, and others re-wired. A discussion of mutation is presented Chapter 7. When nodes of high degree are eliminated, the connectedness of the network is reduced. If enough connected nodes are eliminated, network function collapses.

Albert *et al.* (2000) have compared the robustness of two types of random networks to node elimination: Erdos-Renyi (ER) and scale-free. The ER network is homogeneous, with every node having on average, the same number of edges, with probability p. On the other hand, scale-free networks are characterised by the principle of preferential attachment (Eqn. (6.1)).

Fig. 6.1 shows two simple random networks with $\langle k \rangle = 2$. In (a), an Erdos-Renyi network is shown in which each node has, on average, the same degree, and in (b) a scale-free network is shown, with a small number of highly connected nodes.

Fig. 6.2 shows a histogram of the degree frequencies for each of the two topologies (from Fig. 6.1). The ER graph (a) has a wider degree and flatter distribution whereas the scale-free network has one peak occurring at $deg(1)$.

As biological systems are characterised by evolution and adaptation, networks to model specific aspects of such systems must be capable of supporting edge and node addition and deletion, along with edge *rewiring* (the process in which an edge between nodes xy is removed and added between nodes xk), and such networks are then termed *evolving* or *dynamic*, since

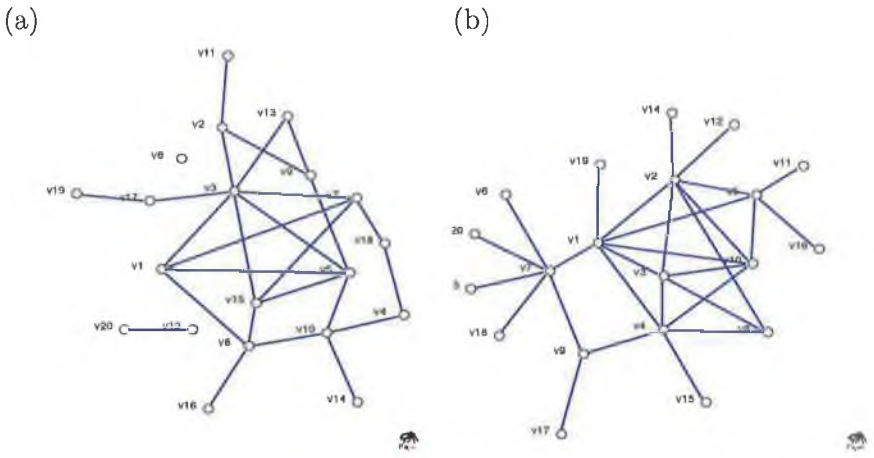


Figure 6.1: Two network topologies. The random ER graph (a) and the scale-free graph (b), with average degree $\langle k \rangle = 2$, and 20 nodes (labeled $v1, v2, ..., v20$).

these will change as the modelled system changes.

6.2 Network Applicability

Networks offer a way to categorise systems of very different origins using a single well defined analytical framework of statistical mechanics (Albert and Barabasi, 2002). If a system can be represented as a network, then analytical network metrics, such as *degree distribution*, *cluster coefficient*, along with geometric centrality measures, for example, *centre*, *median* and *centroid* (Wuchty and Stadler, 2003), can be used to study the properties of system evolution, rather than synthesising a set of metrics *de-novo*.

Due to the broad applicability of complex networks, and because the topology and growth of such networks is closely allied to underlying *function* (Albert *et al.*, 2000), the characteristics of the hybrid model outlined in the previous chapters will be represented in this chapter in network form, as nodes and edges, and with node/edge addition and deletion (that is, *network growth*). The goal of this strategy is to identify any emergent network topology which may arise, and thereby to gain insight into specific aspects

(a) (b)

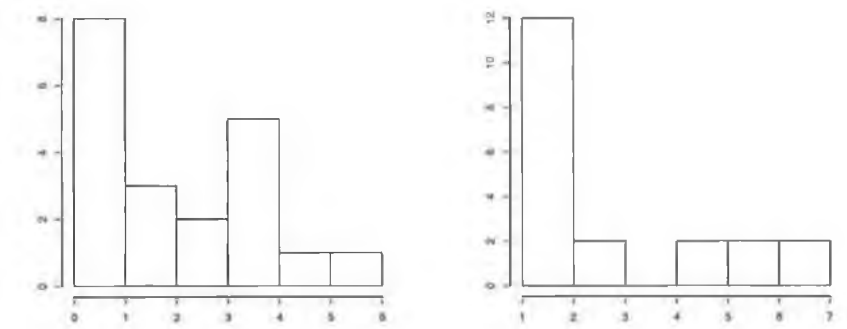


Figure 6.2: Histogram showing the two network degree frequency. The random ER graph (a) and the scale-free graph (b), with average degree $\langle k \rangle = 2$, and 20 nodes. The x-axis represents the degree of each node, and the y-axis represents the relative frequency.

of immune function, such as repertoire development, and the capacity for *recall* (where the immune system remembers an infection from a previous encounter). If an emergent network *can* be identified, this means that the topology of individual immune system networks may then be compared. In turn, this may provide a means to explain why similar [1] initial infections result in such different disease outcomes across the population.

Various network models of the immune system have already been considered in Chapter 2. To briefly reiterate the findings from this section, it was noted that much early research in network models was constrained by the following two assumptions.

1. The approach to network topology derivation was deterministic and *a priori* predefined

2. Immune memory was maintained by idiotypic networks of complementary antibody secretion, which in turn implies that memory is a function of constant B cell to B cell stimulation. Recent work has con-

[1] In terms of dose and route

tradicted this assumption (Murali-Krishna *et al.*, 1999; Swain *et al.*, 1999; Crotty *et al.*, 2003), and this is one possible reason why such models did not directly support clinical phenomena. Hagmann (1999) has pointed out that such constant stimulation could have a counter-productive effect, eventually causing the memory T cells to go into overdrive, and finally, to undergo apoptosis.

In this work, a different approach is adopted. Instead of enforcing a specific network topology, a theory for specifying exactly what constitutes a node and an edge in relation to the immune system shape space is proposed. Crucially, however, the node and edge creation and deletion dynamics *emerge spontaneously from the microscopic model interactions.*

A theoretical framework to support this approach is proposed and discussed in Section 6.2.1. With respect to Item (2.) above, this is driven by the assumption that CTL cells do not directly stimulate other complementary CTL cells to proliferate. In reality, they do not: memory cells are long-lived cells which decay only slowly over time (Liu *et al.*, 2003). This assumption also eliminates the problem of spatial instability inherent in earlier models, where spatial instability was triggered by antibody Ab1 binding to Ab2, and Ab2 is stimulated into producing antibodies which in turn binds to Ab3, and so on, causing a system-wide cascade.

6.2.1 A Theoretical Network Model

The idea of shape space is now developed using graph theory (see, e.g., Diestel (1997)) to model the relationship between infected antigen presenting cell bearing an immunogenic target epitope (denoted APC^+), and an effector activated CTL clonotype (denoted CTL^+). As discussed in Chapter 4, uppercase letters identify shape space components, so an activated CTL clonotype in shape space is denoted CTL^+.

In this approach, shape space is a graph consisting of a pair $G = (V, E)$ of sets satisfying $E \subseteq [V]^2$ (where E is the set of edges and V the set of vertices or nodes). The set of nodes V is made up of both CTL^+ and APC^+. Edges connect an APC^+ to the set of CTL^+ stimulated by its presence to

become activated. Each G is directed and weighted.

At any time $\tau_k > \tau_0$, a new node may be added in shape space with probability dependent on both apc^+ and ctl^- being neighbours in real space *and* distance between APC^+ and CTL^- in shape space is less than or equal to the crossreactive cut-off ρ. In shape space, a newly stimulated CTL clonotype node is designated CTL^+ (indicating it has been *recruited* from the CTL^- pool).

During primary infection, when a new CTL^+ node is added, an edge is added by joining the CTL^+ to the APC^+ node which stimulated its activation. Thus the relationship represented by an edge between two nodes can be understood as: the CTL^+ *was recruited from* the CTL^- pool by the presence of the APC^+, and, the CTL^+ *acts against* the stimulant APC^+:

$$CTL^- \overset{APC^+}{\prec} CTL^+ \tag{6.2}$$

and

$$CTL^+ \overset{attacks}{\prec} APC^+ \tag{6.3}$$

respectively.

Fig. 6.3 depicts the steps of Eqn. (6.2) and Eqn. (6.3), over four time steps. From left to right, top to bottom, the shape space begins in an uninitialised state, with the seven nodes all CTL^-, and no edges in network (an *unconnected* graph). In the second frame, the APC^+ denoted by the black square, is introduced into the immune system. Over the next two time steps, first one CTL^- becomes activated Eqn. (6.2), then two more. At this point, clearance pressure is applied against the APC^+, according to Eqn. (6.3).

There are two types of CTL^+, denoted α and β. An α-node is one which although emerging in response to one individual APC^+, actually affects clearance pressure against other APC^+ appearing later. A β-node is one which acts only against the APC^+ which stimulated its activation.

Fig. 6.4 shows the state of shape space having been subjected to three unrelated infections. The α-node acts against three APC^+ clonotypes (as the APC^+ is the *median* node in each subgraph, the notation m_i is used).

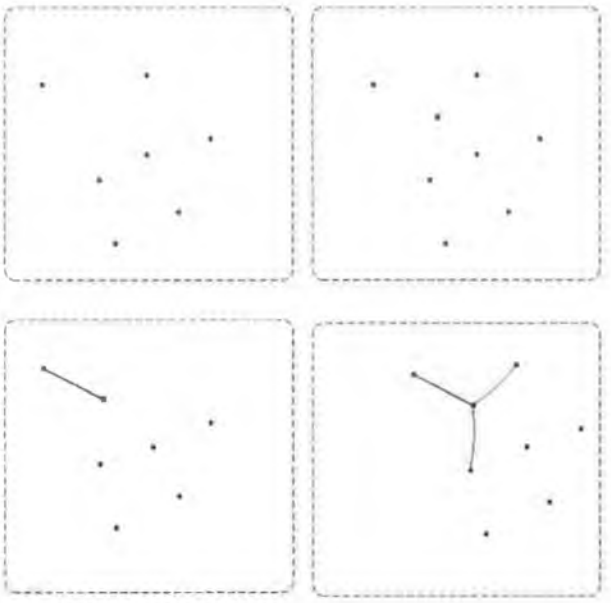


Figure 6.3: Simplified recruitment over four time steps. Left to right, top to bottom.

In so doing, it connects the otherwise unconnected subgraphs of Q, R and S.

The importance of the α-node is clear. Such nodes represent *immune memory*, not only targeting the stimulant infected cell but also other APC^+ nearby [2] in shape space (Fig. 6.4, the m_1, m_2 and m_3 nodes, respectively). Such α-nodes are unique in that the function they illustrate is a component of Eqn. (6.3) without having first been a component of Eqn. (6.2). The β nodes are actually leaf nodes (a leaf node is a node of $deg(1)$), which are non-crossreactive, but persist (beyond the end of the infection which stimulated them from the precursor pool), and do not apply pressure against any APC^+ other than the one which first stimulated it.

In shape space, clearance *pressure* (the strength of the CTL^+ response against the stimulant APC^+) acts along a *directed* and *weighted* edge. A directed edge is one which is defined by the triple:

[2] Falling within the crossreactive cut-off distance ρ.

$$\vec{E} := \{(e, x, y) | e \in E; x, y \in V; e = xy\} \tag{6.4}$$

Where e is an edge from the set E, and x and y are both nodes from the set V. The weight applied to each edge is a function of the distance between the CTL^+ and APC^+ in shape space, and is therefore an indication of the affinity between the CTL^+ and APC^+. Fig. 6.4 shows a simplified network topology after three infections (by heterologous pathogens) have arisen, with median nodes m_1, m_2, m_3. Both α and β nodes are active in infection regulation (a fact demonstrated by the edges connecting the median nodes to each). However, though originally activated by m_1, the node denoted α now affects clearance pressure against both m_2 and m_3. The α-node clearance pressure acts in three directions (for $i = 1, .., 3$):

$\vec{e}_i = (e_i, \alpha, m_i)$, with example weights marked along each edge.

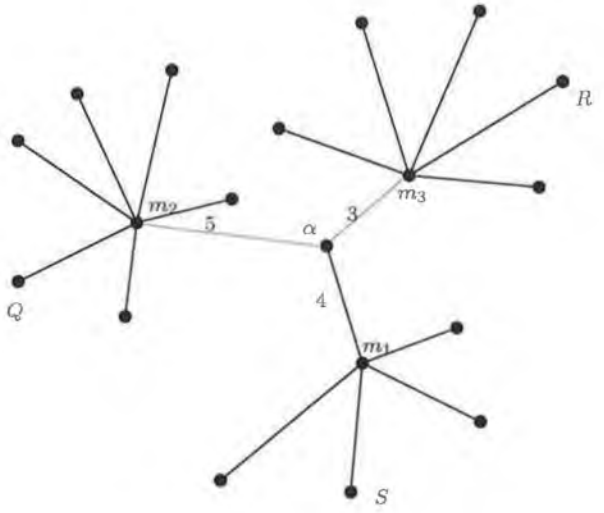


Figure 6.4: Three subgraphs with median nodes (stimulant APC^+) at m_i, connected by the α node. Leaf (β) nodes only apply clearance pressure against their stimulant APC. Weighted edges indicate the force of the clearance pressure applied to the median nodes by the α node.

6.2.2 Biological Refinement

Recent work (Selin *et al.*, 1998; Brehm *et al.*, 2002) has shown that effector CTL memory cells are capable of recognising diverse epitopes of unrelated viruses, and the authors concluded that crossreactivity between heterologous viruses (derived from a separate genetic source) may be a key factor in influencing the hierarchy of CTL responses and the shape of the memory CTL pools. Therefore, it is proposed here that shape space can be used to model both homogeneous viruses with conserved and mutated epitopes (Nowak and May, 2000), *as well as* heterologous viruses with crossreactive epitopes.

A network model of shape space emerges naturally from the real space model as follows. Each immunogenic epitope m_k and activated CTL clonotype c_j are considered as nodes in the space. As mentioned already, clearance pressure applied between two nodes c_j and m_k is represented as a *directed* edge between them. Each edge carries an implicit weight, representing the distance between the two nodes in shape space and therefore, a measure of the affinity between the nodes [3].

After initial infection, most c_j undergo apoptosis (a crucial regulator of immune system homeostasis). However, recruitment to the memory pool consumes around $5 - 10\%$ of activated CTL (Murali-Krishna *et al.*, 1999; De Boer *et al.*, 2001; Badovinac *et al.*, 2004), therefore, these nodes remain active in shape space, preserving the edge connected to the stimulatory epitope m_k.

6.3 Implementation

Having discussed the theoretical basis for the network model, in this section, the model implementation is developed. The main objective is to study the topology of the network model, especially how it emerges and evolves over time. In addition, of particular interest is the means by which network topology can represent the phenomena of *crossreactive memory*, the process by which CTLs primed by exposure to one antigen, can days or years later

[3] Conversely, a measure of the vigour which c_j mounts against, or suppresses m_k.

affect clearance pressure against a genetically unrelated antigen.

Modelling immune memory leads to a possible explanation for the widely observed clinical process that different immune responses may arise when the same stimulus is applied to two immune systems starting in the same initial state. Selin *et al.* (1998) has noted that crossreactive responses by CTL against viruses as diverse as lymphocytic choriomeningitis (LCMV) and vaccinia, have conferred beneficial immune function. Other authors (notably Brehm *et al.* (2002); Mason (1998); Borghans and De Boer (1998)) have also found in favour of crossreactive CTL responses. This feature will be of special interest in supporting the phenomena of network cluster coalescence (discussed below).

6.3.1 Model Parameters

Crossreactivity: As discussed, the number of different clonotypes which respond to the same (randomly selected) epitope is a ratio of a given clonal cutoff parameter to the maximum clonal cutoff parameter (from Eqn.(4.6), known as ρ_{max}) or $\rho/\rho_{max} = \hat{\rho}$. This crucial parameter is known as *crossreactivity*. Work by some authors (for example, Mason (1998)), has suggested this figure to be as high as $50 - 111$. In the original work of Perelson and Oster (1979), the fraction of clonotypes which bind a randomly selected antigen was estimated conservatively at 10^{-3}, so that, for a shape space of size 2.5×10^3 (used here), the number of different clonotypes which respond to the same epitope would be ~ 2.5.

In previous work (Burns and Ruskin, 2004*b*), it was found that in a model of healthy primary response, the number of clonotypes responding to a randomly selected epitope was on average 12.5 (with $\hat{\rho} = 0.35$), and this value is also used in the model presented here. This value is a mid-range figure between the low estimate of Perelson and Oster (1979) and the high estimate of Mason (1998), but supported by the findings of Blattman *et al.* (2002) who estimated a figure of around 20.

In Chapter 5 it was noted that extremely low values of $\hat{\rho}$ (for example $0.107 \leq \hat{\rho} \leq 0.164$) resulted in poor removal of antigen presenting cells

(Fig. 5.7, (b)). Conversely, as $\hat{\rho}$ approaches unity, a randomly selected epitope will be recognised by *every* clonotype in the repertoire. This is clinically never the case. T cell response to stimulation by antigen is always characterised by expansion of an *antigen specific* T cell population (Busch and Pamer, 1999), leading to an *immunodominant* population (Yewdell and Bennink, 1999). In this chapter, it is intended to study the *median* behaviour of the immune system, and thus limit the value of $\hat{\rho}$ to a value which approximates normal immune function. This value is calibrated by the following two observations:

1. A typical viral infection always leads to a *subset* of activated clonotypes (Yewdell and Bennink, 1999). This rules out $\hat{\rho} = 1$, and $\hat{\rho} \ll 0.1$. Therefore, $\hat{\rho}$ can be calibrated by examining the results from Chapters 4 and 5, where *normal* infection clearance rates are observed. Generally, normal clearance occurs in the range $0.6 \leq \hat{\rho} \leq 0.3$.

2. The number of clonotypes which respond to a single random epitope in this model supports the findings of clinical researchers. The low end of which is Borghans and De Boer (1998) (around 5.2), Blattman *et al.* (2002) (around 20) and the higher estimates of Valitutti *et al.* (1995); Itoh *et al.* (1999) (around $80 - 200$). From the models developed in previous chapters, it was noted that a median value of 12.5 emerged with $\hat{\rho} = 0.35$. It can be concluded therefore, that a healthy response is one which typically maintains $\hat{\rho} = 0.35$.

The population of recirculating precursor CTLs (D_{ctl}) is 50000, the number of clonotypes in shape space is $\Theta_{ctl} = 100$, resulting in 500 precursor cells in each group of clonotype. Selection of these values is constrained by the space/time trade-off that only 10^6 spaces are available on the lattice at the start of each simulation. Conservatively, if each stimulated CTL produced 1000 clones at the end of clonal expansion, and 500 exact matches arose, the lattice would be 50% full, and reinfection would increase this level even further. Another calibration point is that murine experiments have shown the expanded T cell repertoire results in some $\sim 10^5$ cells being produced (Kim

et al., 2002). In light of this, the selection of parameter values in this model is appropriate.

Experiment 1

Objective: To analyse the critical nature of crossreactive memory in response to epitopes from heterologous antigens. In particular, the emergent network in shape space is studied in an effort to understand principles of self-organisation.
Method: At time $\tau = 0$, 5000 antigen are placed at random locations on the real space lattice, to simulate antigen-presenting cells entering the lymphatic compartment with marker epitopes displayed on the cell surface. At this stage, the immune system will be in a pristine state without having been exposed to any previous antigen. A primary response will ensue, and a low level fraction of cells will enter the long-lived memory pool, and the remainder of the effector cells will undergo apoptosis.

At time $\tau = 1500, \tau = 3000, \tau = 4500$, further infections are introduced, this is equivalent to a viral infection reappearing every 31.25 days (by the calculation 1500/48). The choice of period for reinfection is somewhat arbitrary, and in fact does not need to be periodic at all. Further infection is included in this model because many severe viral infections are characterised by chronic and persistent phases in which the infected cells are either not fully cleared, or undergo periodic resurgence (Kim and Welsh, 2004). This is an important part of the viral pathology. Of interest here is the *development* of activated clonotype topology in shape space, and to this end, any relatively frequent stimulation which affects space transformation would also be worth considering.

At each infection point, all antigen presenting-cells carry the same epitope, but the epitope is different to the one seen at the previous infection point [4]. In this way, four heterologous infections challenging the immune system are simulated, and this enables the study of the activation network, as well to study the critical nature of crossreactive memory. The simulation

[4] In all, four distinct epitopes e_1, e_2, e_3, e_4, will be presented to the real space model.

executes for 6000 time-steps, simulating some 125 days of real time. This scenario is referred to as $\mathcal{E}_1$ in the rest of this chapter.

Assumptions

In $\mathcal{E}_1$, a case of non-proliferating antigen is assumed, and therefore the clearance rate of infected cells from the real space is a function of recognition and stimulation only (ie, *affinity*). In both experiments, CTL cells have a 8% chance of entering the memory pool once the infection has been cleared. It is assumed that shape space is completely covered (Eqn. (2.1)). Eqn. (2.1) is a realistic assumption in that an *escape* mutation does not imply that no clearance pressure is ever brought to bear on the mutated antigen.

On the contrary, in this model, escape mutation means that no active effector cells (either memory or primary response) can apply clearance pressure *at the time of mutation*. In such a case, the immune response is synthesised as a *de novo* primary response - characterised by relatively slow precursor cell activation and population growth rates, with a consequent elongated antigen clearance profile, typically extended over eight days or more. The on-going thymus generation of cytotoxic T cell precursors was replicated by injecting into the lattice some 5×10^3 precursor CTL cells (ctl^-) at the start of *each* infection event.

6.3.2 Initial Results

The output of the experiment ($\mathcal{E}_1$) is summarised in Figs. 6.5 and 6.6. The network topology of shape space is shown in Fig. 6.5, at the end of the four infections, introduced at $\tau = \{0, 1500, 3000, 4500\}$. Only CTL memory clonotypes are shown in Fig. 6.5. At the centre of each cluster is the immunogenic epitope, and each edge connected to the cluster centre is a stimulated CTL.

To monitor the state of the CTL memory pool, the lattice was sampled at $\tau = \{1500, 3000, 4500, 6000\}$, and at each sample, the position of CTL memory clonotypes and immunogenic epitopes was recorded. As such, CTL cells which had not undergone apoptosis, or become memory cells, are not

shown. Clearly, most immune challenges do not present themselves at equal-dose and equally-spaced time intervals, so the choice of infection time and sample time is somewhat contrived. However, as it serves to illustrate the underlying theory, this configuration will be retained.

After the first infection has been cleared, the CTL memory cells are arranged in a cluster formation around the immunogenic epitope (Fig. 6.5(a)). As discussed in Section 6.3.1, there is a spread of memory within a disc of radius $\hat{\rho}$ from the epitope. As this model is stochastic, the shape space activation spread is "irregular" in nature.

Real space clearance rates of the infected cells, and CTL density levels are shown in Fig. 6.6 (a) and (b), respectively, with plots superimposed to convey cell concentrations. In (b) cell levels are denoted by the symbols $\bullet, \circ, \diamond, \star$, to represent first, second, third and fourth infections, respectively. The first infection ($\bullet$) is cleared with an infected cell half-life of about 3.2 days ($\tau = 156$), with 5% remaining after 10.2 days, and this is broadly in keeping with clearance profiles expected during primary response (Bousso *et al.*, 1999).

When the second infection has been cleared, the shape space network has developed further (Fig. 6.5(b)), and two unconnected clusters emerge. Clearance rate associated with the second infection (Fig. 6.6 (a), ($\circ$)) indicates that no advantage was conferred on the immune response during elimination of the second pathogen. A normal primary response was required - and the clearance rate was almost identical to the first infection ($\bullet$), with a half life of $132 \leq \tau \leq 144$, and a 95% clearance obtained at 10.25 days.

As can be seen from Fig. 6.6 (a), both first and second infections are similar in clearance profile indicating that no previously primed memory cells participated in cell removal, and this is borne out by the two-cluster configuration in shape space in Fig. 6.5 (b).

The randomly chosen epitope location in shape space for the third infection places it within the disc of influence of the previous infection CTL, and some crossreaction between memory CTL arises. Between the two clusters, some 8 CTL have $deg(2)$, and the two clusters are fused into one. The ad-

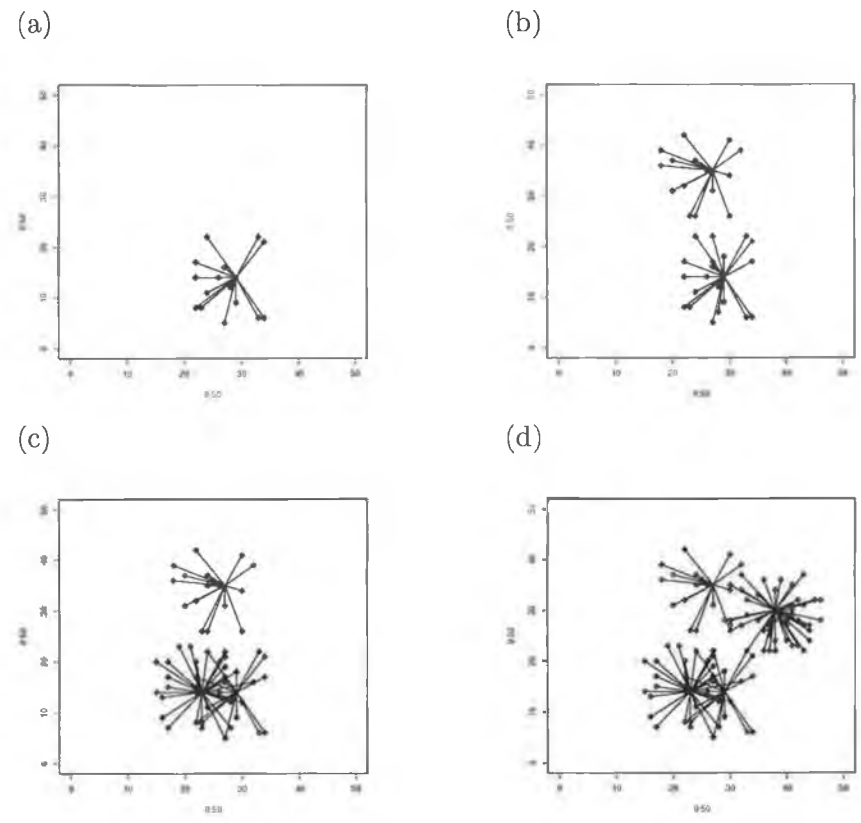


Figure 6.5: The development of a four-epitope network in shape space represented over 6000 time-steps.

vantage conferred on the immune system when memory cells respond to a challenge is clear: given that these are primed from the point of a previous infection they produce armed effector cells without spending time in clonal expansion. Having already been primed by a previous encounter with the specific pathogen, these cells undergo expansion with lower death rates than during primary response. Accumulation thus occurs more rapidly (Grayson *et al.*, 2002).

The first two infections (a) and (b) do not result in any crossreaction between memory cells. At the point of the third infection (c), there exists a pool of memory CTL clonotypes some of which are close enough in shape

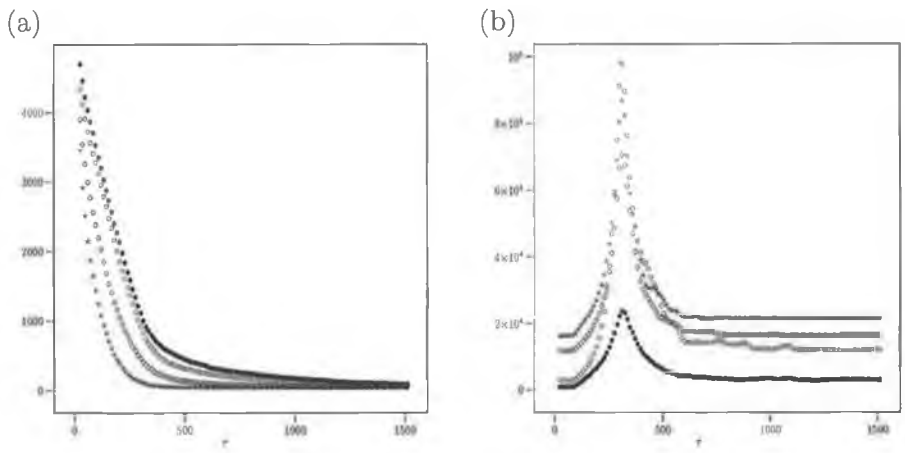


Figure 6.6: Real space density levels of antigen presenting cell (a) and CTL cell (b) respectively in the model of the lymphatic compartment over the course of each infection. The y-axes show the real space concentration levels of the respective cell types.

space to effect clearance pressure against the heterologous virus. These are non-leaf nodes with $deg(2)$. During the final infection (d), the crossreactive contribution of memory cells, specific to one virus, on suppression of the last infection results in the topology of shape space network becoming connected.

From 6.6 (a), the half life of the third infected cell population ($\diamond$) is $\sim$ 72, which is an efficiency improvement of around 50% compared to the previous two. When the final infection is analysed, an important condition has arisen in the shape space network: all small clusters have joined together and merged into one large cluster, due to the critical influence of crossreactive memory clonotypes. Only one clonotype is responsible for connecting the two clusters of Fig. 6.5 (c) into one large cluster.

The benefit of this single cluster network shape space is demonstrated by the clearance dynamics of 6.6 (a): the final infection ($\star$) is cleared so rapidly (half-life $\sim$ 36) that it would probably be asymptomatic. Analysis of the CTL density levels in real space (Fig. 6.6 (b)) explains how clearance of infected cells is so rapid. Each consecutive challenge *with the same level dose of infected cells*, is met with increasingly rapid effector population growth,

and a gradually increasing immunity.

As each infection is introduced and cleared, the average degree of the network in shape space increases. The conditions under which immunity to one virus can reduce the effects of challenge by a second virus are clear from Fig. 6.5. One or more nodes in the first cluster must also have an edge to the second cluster. Thus, damage to or suppression of these critical nodes will have a significantly greater impact than damage to leaf nodes.

Also from Fig. 6.5, the clinical phenomena by which two identical infections have a different disease outcome has a possible explanation reflected by shape space network topology. The density and distribution of α nodes will play a crucial role in determining if the ensuing immune response is primary or secondary.

In Fig. 6.7, the degree distribution of the shape space network at time $\tau = 6000$ (a), is compared with that of a random network (b) in which each the degree of each node is drawn from a uniform distribution in the range $[0, 50]$. Clearly, Fig. 6.7 (a) is not a random distribution and is typical of the density distributions observed over many trials. It has a clear mode at 1 (from Fig. 6.5 (d), the majority of nodes of $deg(1)$) and three minor modes centred around 20, 35 and 49, which represents the cluster "centres" of the immunogenic epitopes.

This non-random degree distribution clearly demonstrates some form of consistent structure emerging from a highly stochastic model of the real space, where the emergent network topology after four infections is arranged around immunogenic epitopes connected in hub-and-spoke formation to memory CTL cells, and in turn, crossreactive memory CTL clonotypes connect cluster formations to each other such that the overall network is connected at the end of the simulation.

6.4 Network Robustness

One measure of any complex network is that of *robustness*: the ability of the network to continue to function in the face of node elimination. The policy under which nodes are selected for elimination is crucial (Albert *et al.*, 2000).

(a) (b)

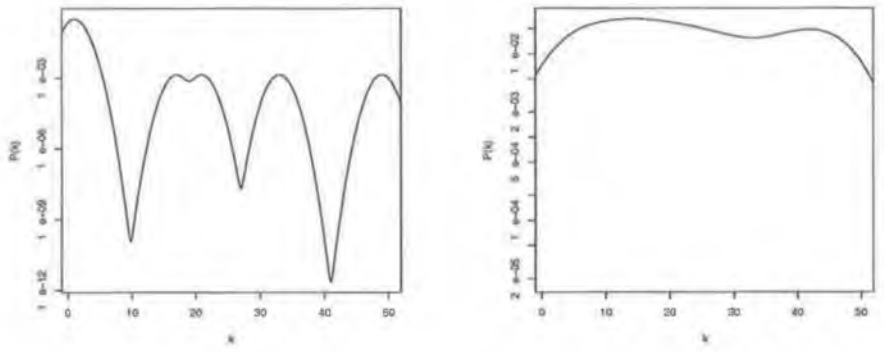


Figure 6.7: Edge degree distribution in shape space network (a) and random network (b).

If the emergent network presented in Section 6.3.2 was indeed a random network (where every new edge attaches to a given node with equal probability), then it is likely that, on average, random node elimination would impair network function gradually and without sudden loss or collapse.

Conversely, if the network was constructed following the heuristic of preferential attachment (for example), than a completely different network capability under random node elimination would be expected. As most real-world environments are noisy, a robust network should be able to tolerate *some* randomised node elimination, but not *targeted* node elimination.

Thus, it is certainly to be expected that any valid model of the immune system should exhibit the crucial real-world characteristics of robustness and fault-tolerance, just as many communications systems must be able to transmit data with a high-level of accuracy in busy or noisy network environments (Rosvall and Sneppen, 2003).

To test that the model derived in this research displays such desirable characteristics, two further tests ($\mathcal{E}_{2A}$ and $\mathcal{E}_{2B}$) are developed as follows:

Experiment 2

$\mathcal{E}_{2A}$ The probability of *any given* node not activating when it is supposed to is taken from the set $\mathcal{P} = \{0.1, 0.3, 0.5, 0.7, 0.9\}$, corresponding to each simulation run. The simulation is repeated 30 times and the results are averaged.

$\mathcal{E}_{2B}$ The probability of α nodes not activating when they are supposed to is again taken from $\mathcal{P}$, and the simulation is repeated 30 times and the results are averaged.

There is nothing particularly significant about the values selected for $\mathcal{P}$, as long as they test the extremes of node loss, and a sample in between. Other values would be equally acceptable. It is not the intention here to identify the critical values for node loss versus immune function as the primary point of interest is to determine how emergent network topology is related to immune function. Clearly, $\mathcal{E}_{2A}$ tests randomised node disruption whereas $\mathcal{E}_{2B}$ tests targeted disruption. Albert *et al.* (2000) have shown that the loss of network function can be described purely in reference to the network *topology* and the cluster coefficient C. However, of interest here is the loss of function, which is studied by measuring the clearance rate of infected antigen presenting cells from the model of lymphatic compartment.

6.4.1 Further Results

The antigen presenting cell density levels for $\mathcal{E}_{2A}$ and $\mathcal{E}_{2B}$ are presented in Figs. 6.8 (a) and (b), respectively. Plots are superimposed for each $\mathcal{P}$ to show cell levels through primary, secondary and repeat re-infection events. Parameters $1, ..., 5$ in $\mathcal{P}$ are denoted $\bullet, \circ, \diamond, \star, *$, respectively.

Two qualitative trends are noticeable in Fig. 6.8 (a). At the first infection (peak 1) immune function is not degraded until the probability of random node elimination ($p_{\alpha\beta}$) reaches $p_{\alpha\beta} = 0.9$. This makes sense when the quantitative values of the simulation are considered. From Chapter 4, Table 4.2, the density of precursor CTL (D_{ctl}) is 5000. Only when around 90% of these effectors are suppressed does the primary response fail. Naturally,

once the primary response has ended and the memory pool is initialised, randomised node loss over the next infection events (peaks 2,3,4) does not have a noticeable effect on the pattern of infected cell elimination.

After each infection event the increasing memory pool is sufficient to supply rapid effector cells in order to bring the infection under control. A qualitatively different dynamic is noticeable when α-nodes are specifically targeted for disruption (Fig. 6.8 (b)). Clearly, eliminating α-nodes should not have any effect on primary response, and this is indeed supported by the results. The clearance regime of $\mathcal{P}$ from the primary response (first peak and tail) is almost the same for each element in $\mathcal{P}$. For second and subsequent infections (peaks two, three and four), the clearance rate of infected antigen presenting cells drops by 63% when the probability of α-node disruption $0.7 < p_\alpha \leq 0.9$ (second peak, $*$).

However, there is little qualitative difference in clearance profiles in the parameter region $0.1 \leq p_\alpha \leq 0.7$ which tells us that network function continues with up to 70% of crossreactive memory absent. This is not surprising given the fact that β-cell function remains unaffected in $\mathcal{E}_{1B}$, and therefore infected cell clearance continues under loss of crossreactive memory. It is noticeable from Fig. 6.8 (b) from left to right, that clearance effectiveness improves over time. This is because the lymphatic compartment is filling up with effector memory cells which (due to their longevity) continue to target and remove infected cells over the 104 days of this simulation.

Some form of on-going purge would be required to prevent the lymphatic compartment from filling up, but this process is not modelled here as the precise mechanism has not been clinically identified (Kim and Welsh, 2004). *Interesting aside*: some clinical observations (again, Kim and Welsh (2004)) have suggested that memory clonotypes with *less* crossreactive capability are progressively lost (or deleted) in favour of the strongly crossreactive. This is logical, as over time, the lymphatic compartment would begin to fill up with memory cells retained from all the previous infections. With space at a 'premium', it would be advantageous to keep only the most crossreactive memory cells. This illustrates one of the advantages of the model in that the trade-off between increasing memory pool size and limited lymphatic

space can be studied in relation to the effect of memory recall to repeat reinfection.

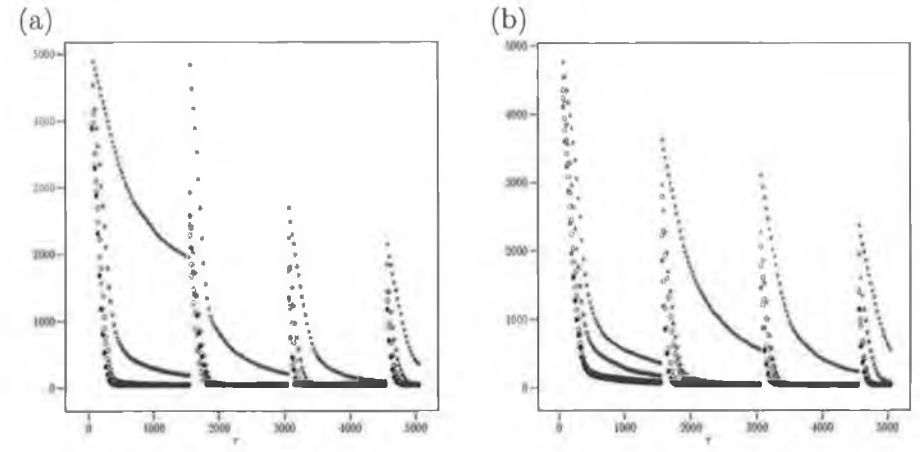


Figure 6.8: Clearance dynamics of infected antigen presentation cell over four infection events. In (a), the effect of random loss of α or β -nodes is contrasted to (b) where specifically, α-nodes are targeted for loss. The y-axes shows the real space concentration levels of infected APCs.

A fuller picture of the immune response to randomised and targeted cell disruption emerges when considering the density levels of effector cells when challenged during primary, secondary and repeat infection. Fig. 6.9 shows the density of effector cells for $\mathcal{E}_{2A}$ and $\mathcal{E}_{2B}$ ((a) and (b), respectively). Again, plots are superimposed for each $\mathcal{P}$ to show cell levels through primary, secondary and repeat re-infection events, and parameters $1, ..., 5$ in $\mathcal{P}$ are denoted $\bullet, \circ, \diamond, \star, *$, respectively.

A comparison of the density levels for activated effector cells demonstrates the consequences of randomised and targeted node loss. Whereas Fig. 6.9 (a) shows a relatively normal primary and secondary response profile (qualitatively similar to those shown in Chapter 3, Fig. 3.3), indicating broad network function, Fig. 6.9 (b) shows that, with the exception of minor α-node loss ($\bullet$), each infection event broadly results in the *same* density profile as seen in primary infection, demonstrating that the immune system is constantly mounting a *primary response* to an infection normally controlled

by memory response.

In effect, the benefits of having been immunised (by clearing the infection at some earlier time) are lost, and the immune system is constantly having to re-acquire responses which should be already available to it. Therefore, the study of targeted and randomised cell disruption can be concluded by observing that moderate α-cell loss does not lead to failure in infected cell clearance (Fig. 6.8 (b)), but rather, it leads to an *qualitative* impairment of memory. The value conferred by prior exposure to heterologous infection is eliminated, and each repeat exposure is marked by *ab-initio* primary response dynamics.

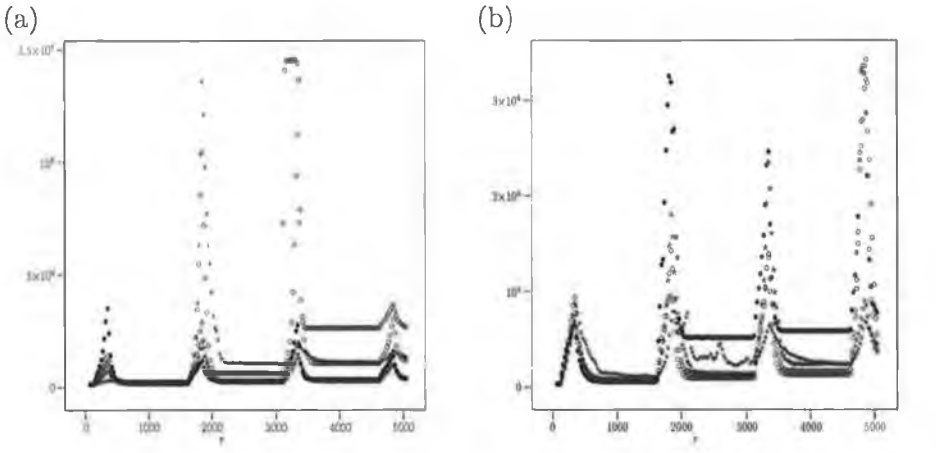


Figure 6.9: Concentration levels for activated effector cells during primary, secondary and repeat reinfection events (y-axes). In (a), density levels under randomised node elimination, and (b) shows density levels under targeted node elimination.

6.5 Discussion

Most discrete computational models of immune response to viral infections have used real space or shape space formalisms. In this work, however, a model based on a combination of the two has been presented, with the objective of demonstrating how emergent behaviour and principles of self

organisation may arise from a many-particle microscopic system. This is achieved by using a stochastic model of the lymphatic system as the stimulus to shape space differentiation and distribution.

An extension to shape space has been developed which goes beyond restrictive early network models, to demonstrate a mechanism by which early and protective immunity can be mediated by memory T cells generated by a previous heterologous viral infection. This important feature emerges in the shape space network as a memory T cell node v of $deg(v) \geq 2$. An edge between two nodes $c_j m_k$ is added whenever one exerts clearance pressure against the other. The pressure applied between two nodes c_j and m_k is represented as a directed edge between them. Each edge carries an implicit weight, representing the distance between the two nodes in shape space and therefore, providing a measure of the affinity between the nodes.

Of course, the *degree* of protective immunity offered by crossreacting memory cells is dependent on the distance between the memory T cell clonotype and the immunogenic epitope, with optimal immunity arising when reinfection is by the same antigenic epitope. In the results, increasingly effective clearance dynamics are seen as the memory pool increases, and each T cell clone has a 8% chance of becoming a long-lived memory cell.

Although it has not been modelled here, the immune system cannot continually increase the size of the memory pool. As mentioned in the *aside* earlier, Brehm *et al.* (2002) has suggested that some (as yet undefined) process probably exists to purge non-crossreactive (or at the very least, the relatively weaker crossreactive) memory cells in order that the pool does not grow beyond the real limits imposed by the restrictive spaces of the lymphatic system. However, this purging method, if it exists, is as yet unknown, but is likely to consist of progressive and selective deletion of low crossreactive memory cells. The network implication of α-nodes is that each edge connection formed from a median node to an α-node acts as a back-bone in joining disparate *subgraphs*.

6.6 Chapter Conclusions

Individuals vary considerably in their responses to viral infections, ranging from subclinical to severe. There are many factors that contribute to this variation in responsiveness, including the dose and route of infection, as well as the physiological state and genetic background of the host (Selin *et al.*, 1998). In this study, it was demonstrated that memory T cells in immune system shape space (CTL^+) which are specific to unrelated viruses may also contribute to the immune primary response to a second virus.

The presence of memory T cells has the effect of retarding the spread of infection, thus allowing time for further specific immune responses to develop. *It is expected that the effect of these cells may be the difference between clinical and subclinical infections or lethal and nonlethal infections*, and it is proposed that such clinical outcomes may be explained, at least in part, by the varying topology of the immune system shape space network of memory CTL cells.

In the results presented here, it has been proposed that α-nodes have strong biological equivalent, namely cytotoxic T lymphocyte memory cells. Such cells, having been primed by way of a previous immune challenge, require less time to respond, and, crucially, tend to be beneficially crossreactive. Thus, these cells are the source of the beneficial effects of protective memory T cell immunity. This finding is in agreement with Brehm *et al.* (2002) and others.

Disruption to α-nodes results in a significantly degraded pathogen clearance than disruption to β-nodes does. This supports the theory that during primary and secondary response, some cytotoxic clonotypes are *more important than others*.

Finally, a key objective of the investigation has been established. It has been demonstrated that a complex network in immune system shape space has emerged *naturally* from the lower level processes in the model hierarchy (that is, from the real space model of the idealised lymphatic compartment). It was shown here that this complex network demonstrates two qualitatively different node classes, corresponding to crossreactive memory effector and

non-crossreactive memory effector T cells (or α and β cells respectively). When this complex network topology was disturbed (by randomised or targeted node elimination), *immune function* qualitatively altered in response, thereby demonstrating a causal link between network topology and model function.

CHAPTER 7

Summary, Conclusions and Future Research

7.1 Introduction

Segel (1995) has observed that the defining characteristic of a complex system is that it cannot be represented by one single model. This is undoubtedly true with respect to the immune system. All one can hope to do is to shed light on some specific aspects of the system and at the same time, justify key model parameters from the biological point of view, in order to operate within accepted theoretical immunology understanding.

All research builds on the work of others; to this end, the formalisms of Stochastic Cellular Automata, Shape Space and Complex-Emergent Networks have been important components in the models developed in this work. The benefits of a combination of these formalisms have been to bring together *reductionism* and *globalism* into one model, in order to provide further insight into immunological phenomena, particularly with respect to how individuals vary in their responses to viral infections.

7.2 Summary of Findings

Most software development is characterised by the middle-out approach. The designer has the opportunity to first tackle tractable mid-level complexity before embarking on high- and low- level design issues, the difficulties and subtleties of which may only become apparent once mid-level model complexity has matured. This approach also has the benefit of helping to guide model development by avoiding unwieldy and overly-complex high-level designs and, conversely, avoiding swamping the approach in abstruse technicalities. In this research, a form of middle-out development was adopted.

In Chapter 3, a high-fidelity stochastic microscopic model of a subset of known immunological dynamics was built. A seven state non-deterministic finite automata (NFA) of the effector T cell life-cycle, which was encoded as a set of states and state transitions. When *abnormal* transitions (such as failure or delay in moving from activated effector to memory effector) were studied, the model demonstrated a relationship between repeated reinfection and effector cell transition to memory or apoptosis.

The results of Chapter 3 were validated to ensure that the model was qualitatively and quantitatively realistic. The output of the model under various parameter regimes was compared to known murine experimentation results. The findings suggest that repeated reinfection can be controlled only within a limited range of effector CTL cell transitioning to memory CTL cell. Too much memory caused the lymphatic compartment to fill-up, too little memory induced the need for clonal expansion from naive precursor cells, and an elongated APC clearance profile.

Chapter 4 devised a mechanism to introduce a macroscopic or system-wide viewpoint (shape space) into the model. Shape space was chosen to represent global properties of the immune system (such as initial state, infected, immunised), because of how readily it represents the immune *repertoire*. In turn, this enabled the principle of *downward causation* (from macroscopic to microscopic model) to be demonstrated.

In Chapter 4, when the 100 most active clonotypes in shape space were

studied after infection by a single pathogen, two distinct activation patterns emerged, characterised by (i) high cell concentration, narrow clonotype activation distribution and (ii) low concentration but broad clonotype activation distribution. For all other model parameters being unchanged, reducing crossreactivity (ρ) caused significantly different development of shape space (in terms of the clonotype activation distribution and concentration), and correspondingly reduced infected cell clearance rates.

Chapter 5 demonstrated *upward causation*, by visualising shape space and observing the emergent patterns. The results presented in this chapter lead to the conclusion that certain types of activation patterns in shape space lead to more efficient and effective real space infected cell clearance rates than do others.

It was demonstrated in Chapter 5 that as the mean distance between CTL^+ and APC^+ declines (broadly *regardless of the mean activation concentration*), the clearance of infected antigen presenting cells from the real space model becomes progressively less efficient.

In Chapter 6, a theoretical network architecture for immune system shape space was outlined, introducing α and β nodes, which corresponded to cross-reactive and non-crossreactive memory cells respectively. This network model postulated that topological clusters would emerge in response to APC^+ stimulation in shape space. The emergent network topology of the model simulation was then studied to verify that the theoretical network was present in an applied sense, and this was indeed found to be the case. When the edges were added between CTL and their stimulant APC in shape space, a network topology emerged naturally from the model.

The emergent network was tested in Chapter 6 to study *robustness*. Disruption (or suppression) of α-nodes results in a significantly degraded pathogen clearance than disruption to β-nodes did. This supported the theory that during primary and secondary response, some cytotoxic clonotypes are *more important than others*. The network implication of α-nodes is that each edge connection formed from a median node to an α-node acts as a back-bone in joining disparate *subgraphs*.

7.3 Future Research

By the very nature of such an undertaking, a Ph.D. must draw boundaries around the research topic, with the inevitable exclusion of certain specific subject matter, in order to achieve satisfactory *depth*, without unfairly sacrificing *breadth*. So it has been in this research. In each of the chapters so far, there have been several possible features which have not been incorporated into the research and which are short-term immediate targets for investigation. In this section, two key next steps in the research, and the potential benefit from their inclusion, are discussed.

7.3.1 Mutation and Shape Space

Foremost among the important next steps in this research is the incorporation of pathogen *mutation*. Mutation itself is not a straightforward process. In some pathogens, mutation is rare for example), while in others (such as the Human Immune Virus, HIV) it is both rapid and pathological. Mutation is structural change in a gene which changes the information encoded in that gene.

In shape space, a mutation would imply some form of alteration in the epitope structure such that the epitope co-ordinates in the shape change (the epitope moves). Drift and shift in shape space is distinguished as follows. A pathogen which drifts is one against which some immune response is maintained, even though the epitope structurally alters (and therefore, in shape space, it no longer occupies its original location in that it *drifts*).

When an antigen shifts, its epitope changes structurally to such an extent that the immune system is *unable* to exert clearance pressure, and the location of the new epitope in shape space is such that no CTL pressure is brought to bear. A viral epitope which does not change during the course of primary and subsequent infection is called *conserved*. These three epitope types are summarised in Fig. 7.1. The following experiment is one potential method for studying mutation.

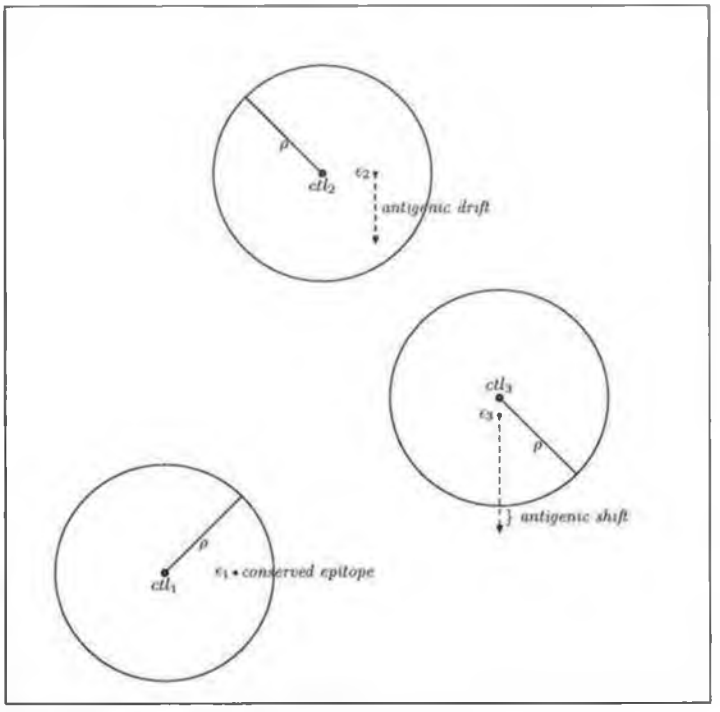


Figure 7.1: Three epitopes in shape space characterised by *conserved*, *drift* and *shift*.

Mutation Experiment

Objective: To study the contrasting dynamics of epitope conservation, drift and shift for an initially single-strain viral challenge, and to analyse the emergent shape space network.

Method: In this experiment, an initial infection of 5000 single-strain antigen presenting cells (all cells bearing the same shape space position) is introduced at time $\tau = 0$, and the model executes for 6000 time steps. Initially, each antigen would carry the same epitope, but once the primary response is over (and a memory pool has been created), there then follows second and subsequent antigen reinfection by the same strain, but at each time step, there is a probability p of epitope drift. That is, the shape space co-ordinates may change by some value, but not escaping the clearance pressure of the

stimulated cytotoxic lymphocyte effector.

To simulate antigen shift in shape space, a final parameter is added, p_{shift}, which is the probability that the epitope displayed on the antigen presenting cell surface will shift to a position in shape space which is outside the area of clearance pressure applied by the stimulated cytotoxic cells.

Fig 7.1, shows a three epitope shape space with epitope ϵ_1 is a conserved (non-mutating) epitope and the pressure exerted by ctl_1 is constant over time. ctl_2 maintains pressure against a drifting epitope, ϵ_2, but the distance between the two points never exceeds ρ, and some decreasing pressure is maintained. Finally, ctl_3 is initially able to maintain clearance pressure against ϵ_3, but due to an escape mutation, it shifts beyond ρ, and no further clearance pressure is brought to bear.

Mutation will be included as an immediate next-step of this work. From the model in this thesis, integration of mutation dynamics is a relatively straightforward task.

7.3.2 Targeted Node Elimination

As discussed in Chapter 6, Section 6.4.1, the loss of α-nodes resulted in complete memory loss to the immune system. Each subsequent reinfection resulted in almost identical primary response dynamics characteristic of continuous *re-learning*. This experiment (denoted $\mathcal{E}_{2B}$) could be further refined to study the effect of node loss for each specific node *degree*.

It is expected that a loss of nodes of high degree should disrupt response more significantly than loss of nodes of low degree. This experiment would result in a reduced cluster coefficient for the network (as more separated clusters emerge).

7.3.3 Shape Space Dimensionality, N

As noted in Chapter 2, the dimensionality of shape space (N) has been identified as a crucial characteristic of the formalism. The models presented in this thesis easily support increasing N and are not fixed at $N = 2$. Increasing N would merely require the length of the shape space coordinate

vector to be increased. Importantly, the calculation of distance (Eqn. 4.2) and hence affinity, would not change. Interestingly, increasing N does not affect any network topology assumptions, as topology is described in terms of degree distribution (and not dimension of the space). It would, however, make the network more difficult to visualise. In conclusion, increasing N is feasible as a further step in the development of these models.

7.4 Final Remarks

The models developed in this thesis offer a *viable* theory to explain how primary, secondary and continuous reinfection by heterologous pathogens may affect effector T cell function. This has been done, from the modelling perspective, by demonstrating how emergent network topology in immune system shape space affects immune function, especially when node loss or suppression is introduced.

Furthermore, it has been shown that minor topological alteration results in *downward causation* (the converse of the reductionist principle) in that the behaviour of the microscopic parts of the model are determined by the behaviour of the whole and thus determination moves downwards instead of upwards. Individual immune systems may develop minor topological variations which effect the clearance rates and therefore the duration and pathology of similar pathogen strain.

In closing, this thesis may be an important step in addressing both the limitations of early network-based models of immune response, as well as an useful initial step in understanding immune function by way of emergent network topology.

References

Albert R and Barabasi A L (2000). Topology of Evolving Networks: Local Events and Universality, *Phys. Rev. Lett.* **85(24)**, 5234–5237.

Albert R and Barabasi A L (2002). Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74**, 47–95.

Albert R, Jeong H and Barabasi A L (2000). Error and attack tolerance of complex networks, *Nature* **406**, 378–381.

Anderson R W, Neumann A U and Perelson A S (1993). A Cayley Tree Immune Network Model with Antibody Dynamics, *Bull. Math. Biol.* **55(6)**, 1091–1131.

Asquith B and Bangham C R (2003). The dynamics of T-cell fratricide: application of a robust approach to mathematical modelling in immunology, *J. Theor. Biol.* **222**, 53–69.

Badovinac V P, Porter B B and Harty J T (2004). CD8+ T cell contraction is controlled by early inflammation, *Nat. Immunol.* **5**, 809–817.

Bandini S, Mauri G and Serra R (2001). Cellular automata: From a theoretical parallel computational model to its application to complex systems, *J. Par. Comp.* **27**, 539–553.

Barabasi A L, Oltvai Z and Wuchty S (2004). Characteristics of Biological Networks, *Lect. Notes Phys.* **650**, 443–457.

Barber D L, Wherry E J and Ahmed R (2003). Cutting Edge: Rapid In Vivo Killing by Memory CD8 T Cells, *J. Immunol.* **171**, 27–31.

Bergeron B (2002). *Bioinformatics Computing*, Prentice Hall, Upper Saddle River, New Jersey.

Berlekamp E, Conway J and Guy R K (1982). *Winning Ways for Your Mathematical Plays*, ACTA Press, London.

Bernardes A T and Zorzenon dos Santos R M (1997). Immune Network at the Edge of Chaos, *J. Theor. Biol.* **186**, 173–187.

Bernaschi M and Castiglione F (2001). Design and implementation of an immune system simulator, *Computers in Biology and Medicine* **31**, 303–331.

Bezzi M, Celada F, Ruffo S and Seiden P (1997). The transition between immune and disease states in a cellular automaton model of clonal immune response, *Physica A* **245**, 145–163.

Blattman J, Antia R, Sourdive D, Wang X, Kaech S, Murali-Krishna K, Altman J and Ahmed R (2002). Estimating the precursor frequency of naive antigen-specific CD8 T cells, *J. Exp. Med.* **195**, 657–664.

Blattman J, Sourdive D, Murali-Krishna K, Ahmed R and Altman J (2000). Evolution of the T Cell Repertoire During Primary, Memory, and Recall Responses to Viral Infection, *J. Immunol.* **165**, 6081–6090.

Borghans J A and De Boer R J (1998). Crossreactivity of the T-cell receptor, *Immunology Today* **19(9)**, 428–429.

Bousso P, Levraud J P, Kourilsky P and Abastado J P (1999). The Composition of a Primary T Cell Response Is Largely Determined by the Timing of Recruitment of Individual T Cell Clones, *J. Exp. Med.* **189(10)**, 1591–1600.

Brass A, Grencis R K and Else K J (1993). A CA model for helper T cells subset polarization in chronic and acute infection, *J. Theor. Biol.* **166(2)**, 189–200.

Brehm M, Pinto A, Daniels K, Schneck J, Welsh R and Selin L (2002). T cell immunodominance and maintenance of memory regulated by unexpectedly cross-reactive pathogens, *Nat. Immunol.* **3(7)**, 627–634.

Burns J and Ruskin H J (2004*a*). A Stochastic Model of the Effector T Cell Lifecycle, *in* P.M.A. Sloot and Bastien Chopard and Alfons G. Hoekstra, ed., *Cellular Automata*, Vol. 3305 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin Heidelberg, pp. 454–463.

Burns J and Ruskin H J (2004*b*). Diversity Emergence and Dynamics During Primary Immune Response: A Shape Space, Physical Space Model, *Theor. in Biosci.* **123(2)**, 183–194.

Busch D H and Pamer E G (1999). T Cell Affinity Maturation by Selective Expansion during Infection, *J. Exp. Med.* **189**, 701–710.

Busch D, Kerksiek K and Pamer E G (1999). Processing of Listeria monocytogenes antigens and the in vivo T-cell response to bacterial infection, *Immunol. Rev.* **172**, 163–169.

Buseyne F and Riviere Y (2001). The flexibility of the TCR allows recognition of a large set of naturally occurring epitope variants by HIV-specific cytotoxic T lymphocytes, *Int. Immunol.* **13**, 941–950.

Campbell D (1974). 'Downward causation' in Hierarchically Organized Biological Systems, *in* F Ayala and T Dobzhansky, eds, *Studies in the Philosophy of Biology*, Macmillan Press, pp. 179–186.

Carneiro J and Stewart J J (1994). Rethinking "Shape Space": Evidence from Simulated Docking Suggests that Steric Complementarity is not limiting for Antibody-Antigen Recognition and Idiotypic Interactions, *J. Theor. Biol.* **169**, 391–402.

Castiglione F, Bernaschi M and Succi S (1997). Simulating the Immune Response on a Distributed Parallel Computer, *Int. J. Mod. Phys.* **8**, 527.

Castiglione F, Motta S and Nicosia G (2001). Pattern recognition by primary and secondary response of an Artificial Immune System, *Theor. in Biosci.* **120**, 93–106.

Chao D L, Davenport M P, Forrest S and Perelson A S (2003). Stochastic Stage-structured Modeling of the Adaptive Immune System, *in Proceedings of the IEEE Computer Society Conference on Bioinformatics*, IEEE Computer Society, p. 124.

Chao D L, Davenport M P, Forrest S and Perelson A S (2004). A stochastic model of cytotoxic T cell responses, *J. Theor. Biol.* **228**, 227–240.

Chopard B and Tomassini M (2004). Randomized Computation with Cellular Automata, *in* P.M.A. Sloot and Bastien Chopard and Alfons G. Hoekstra, ed., *Cellular Automata*, Vol. 3305 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin Heidelberg, pp. 71–80.

Coles R M, Mueller S N, Heath W R, Carbone F R and Brooks A G (2002). Progression of Armed CTL from Draining Lymph Node to Spleen Shortly After Localized Infection with Herpes Simplex Virus 1, *J. Immunol.* **168**, 834–838.

Corbin G A and Harty J T (2004). Duration of Infection and Antigen Display Have Minimal Influence on the Kinetics of the CD4+ T Cell Response to Listeria monocytogenes Infection, *J. Immunol.* **173**, 5679–5687.

Crotty S, Felgner P, Davies H, Glidewell J, Villarreal L and Ahmed R (2003). Long-term B cell memory in humans after smallpox vaccination, *J. Immunol.* **171(10)**, 4969–4973.

De Boer R J, Oprea M, Kaja R A, Murali-Krishna, Ahmed R and Perelson A S (2001). Recruitment Times, Proliferation, and Apoptosis Rates during the $CD8^+$ T-Cell Response to Lymphocytic Choriomeningitis Virus, *J. Virol.* **75(22)**, 10663–10669.

De Boer R, Segel L and Perelson A (1992). Pattern formation in one- and two-dimensional shape-space models of the immune system, *J. Theor. Biol.* **155(3)**, 295–333.

Diestel R (1997). *Graph Theory*, Springer-Verlag.

Dimitrov V (1997), Key Notions and Concepts in the Paradigm of Complexity. Unpublished.
URL: *http://www.zulenet.com/VladimirDimitrov/pages/notions.html*

Everett P C, Seldin E B, Troulis M, Kaban L B and Kikinis R (2003). A 3-D System for Planning and Simulating Minimally-Invasive Distraction Osteogenesis of the Facial Skeleton, *in* M Hamza, ed., *Biomedical Engineering*, Proceedings of the IASTED International Conference, ACTA Press, pp. 90–95.

Fortunato S, Stauffer D and Coniglio A (2003). Percolation in high dimensions is not understood, *Physica A* **334**, 307–311.

Fuller M J, Khanolkar A, Tebo A E and Zajac A J (2004). Maintenance, Loss, and Resurgence of T Cell Responses During Acute, Protracted, and Chronic Viral Infections, *J. Immunol.* **172**, 4204–4214.

Gardner M (1970). Mathematical games: The fantastic combinations of John Conways new solitary game "life", *Scientific American* **223**, 120–123.

Germain R N (2001). The Art of the Probable: System Control in the Adaptive Immune System, *Science* **293**(5528), 240–245.
URL: *http://www.sciencemag.org/cgi/content/abstract/293/5528/240*

Germain R N and Stefanova I (1999). The Dynamics of T Cell Receptor Signalling: Complex Orchestration and the Key Roles of Tempo and Cooperation, *Annu. Rev. Immunol.* **17**, 467–522.

Grayson J, Harrington L, Lanier J, Wherry E and Ahmed R (2002). Differential Sensitivity of Naive and Memory CD8+ T cells to Apoptosis in Vivo, *J. Immunol.* **169(7)**, 3760–3770.

Hagen H, Ebert A, van Lengen R H and Scheuermann G (2003). Scientific Visualization: methods and applications, *in Proceedings of the 19th spring conference on Computer graphics*, ACM Press, pp. 23–33.

Hagmann M (1999). Memory T Cells Don't Need Practice, *Science* **286(5443)**, 1267–1270.

Hershberg U, Louzoun Y, Atlan H and Solomon S (2001). HIV time: winning the war while, loosing all the battles, *Physica A* **289**, 178–190.

Hopcroft J E and Ullman J D (1979). *Introduction to Automata Theory, Languages and Computation*, Addison Wesley.

Itoh Y, Hemmer B, Martin R and Germain R N (1999). Serial TCR engagement and down-modulation by peptide:MHC molecule ligands: Relationship to the quality of individual TCR signalling events, *J. Immunol.* **162**, 2073:2080.

Janeway C, Travers P, Walport M and Capra J (1999). *Immunobiology. The Immune System in Health and Disease*, Churchill-Livingston.

Jeong H, Mason S, Barabasi A L and Oltvai Z (2001). Lethality and centrality in protein networks, *Nature* **411**, 41–42.

Jeong H, Tombor B, Albert R, Oltvai Z and Barabasi A L (2000). The large-scale organization of metabolic networks, *Nature* **407**, 651–654.

Jerne N (1974). Towards a network theory of the immune system, *Ann. Immunol.* **125 C**, 373–389.

Kanji G K (2000). *100 Statistical Tests*, Sage, London, Thousand Oaks, New Delhi.

Kim S, Brehm M A, Welsh R and Selin L K (2002). Dynamics of Memory T Cell Proliferation Under Conditions of Heterologous Immunity and Bystander Stimulation, *J. Immunol.* **169**, 90–98.

Kim S and Welsh R (2004). Comprehensive early and lasting loss of memory CD8 T cells and functional memory during acute and persistent viral infections, *J. Immunol.* **172(5)**, 3139–3150.

Kirschner D, Webb G and Cloyd M (2000). Model of HIV-1 Disease Progression Based on Virus-Induced Lymph Node Homing and Homing-Induced Apoptosis of $CD4^+$ Lymphocytes, *J. Acqui. Immune Defic. Syndr.* **24**, 352–362.

Klenerman P, Wu Y and Phillips R (2002). HIV: current opinion in escapology, *Current Opinion in Microbiology* **5**, 408–413.

Lagreca M, de Almeida and Zorzenon dos Santos R (2001). A dynamical model for the immune repertoire, *Physica A* **289**, 191–207.

Landau D P and Binder K (2000). *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press.

Law A M and Kelton W D (2000). *Simulation Modeling and Analysis*, McGraw Hill, New York.

Liebman M (2001). From Bioinformatics to Biomedical Informatics, *Genome Technology* **11**, 1–2.

Liu H, Andreansky S, Diaz G, Turner S J, Wodarz D and Doherty P C (2003). Quantitative Analysis of Long-Term Virus-Specific $CD8^+$-T-Cell Memory in Mice Challenged with Unrelated Pathogens, *J. Virol.* **77(14)**, 7756–7763.

Liu J S (2001). *Monte Carlo Strategies in Scientific Computing*, Springer.

Mannion R, Ruskin H and Pandey R (2000). Effect of Mutation on Helper T-cells and Viral Population: A Computer Simulation Model for HIV, *Theor. in Biosci.* **199(2)**, 145–155.

Mannion R, Ruskin H and Pandey R (2002). Effects of Viral Mutation on Cellular Dynamics in a Monte Carlo Simulation of HIV Immune Response Model in Three Dimensions, *Theor. in Biosci.* **121(2)**, 237–245.

Mason D (1998). A very high level of crossreactivity is an essential feature of the T-cell receptor, *Immunology Today* **19(9)**, 395–404.

McKinney D M, Skvoretz R, Livingston B D, Wilson C C, Anders M, Chesnut R W, Sette A, Essex M, Novitsky V and Newman M J (2004). Recognition of Variant HIV-1 Epitopes from Diverse Viral Subtypes by Vaccine-Induced CTL1, *J. Immunol.* **173**, 1941:1950.

McReynolds T and Blythe D (1999). Advanced Graphics Programming Techniques Using OpenGL, *ACM SIGGRAPH Workshop on Advanced Graphics* **65(1-2)**, 3–19.

Merks R M, Newman S A and Glazier J A (2004). Cell-Oriented Modeling of *In Vitro* Capillary Development, *in* P.M.A. Sloot and Bastien Chopard and Alfons G. Hoekstra, ed., *Cellular Automata*, Vol. 3305 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin Heidelberg, pp. 424–434.

Moore G E (1965). Cramming more components onto integrated circuits, *Electronics* **38(8)**, 1–4.

Murali-Krishna K, Lau L L, Sambhara S, Lemonnier F, Altman J and Ahmed R (1999). Persistence of Memory CD8 T Cells in MHC Class I-Deficient Mice, *Science* **286**, 1377–1381.

Naumov Y N, Naumova E N, Hogan K T, Selin L K and Gorski J (2003). A Fractal Clonotype Distribution in the CD8+ Memory T Cell Repertoire Could Optimize Potential for Immune Responses, *J. Immunol.* **170**, 3994–4001.

Neumann A and Weisbuch G (1992). Dynamics and Topology of Idiotypic Networks, *Bull. Math. Biol.* **54**, 699–726.

Noest A, Takumi K and De Boer R (1997). Pattern formation in B-cell immune networks: Domains and dots in shape-space, *Physica D* **105**, 285–306.

Nowak M A and May R M (1991). Mathematical Biology of HIV Infections: Antigenic Variation and Diversity Threshold, *Math. Biosci.* **106**, 1–21.

Nowak M A and May R M (2000). Oxford Univ. Press, Oxford New York, chapter 14, pp. 149–181.

Orosz C G (2001). An Introduction to Immuno-ecology and Immuno-informatics, *in* L A Segal and I R Cohen, eds, *Design Principles for the Immune System and Other Distributed Autonomous Systems*, Santa Fe Institute Studies in the Science of Complexity, Oxford Univ. Press, pp. 125–149.

Oxenius A, Price D A, Dawson S J, Tun T, Easterbrook P J, Phillips R E and Sewell A K (2001). Cross-staining of cytotoxic T lymphocyte populations with peptide-MHC class I multimers of natural HIV-1 variant antigens, *AIDS* **15**, 121–122.

Palmer B E, Boritz E and Wilson C C (2004). Effects of Sustained HIV-1 Plasma Viremia on HIV-1 Gag-Specific CD4+ T Cell Maturation and Function, *J. Immunol.* **172**, 3337:3347.

Pandey R (1998). A Stochastic Cellular Automata Approach to Cellular Dynamics for HIV: Effect of Viral Mutation, *Theor. in Biosci.* **117**, 32.

Papa A R and Tsallis C (1996). A local-field-type model for immunological systems: time evolution in real and shape space, *Physica A* **223**, 85–101.

Perelson A S and Oster G F (1979). Theoretical Studies of Clonal Selection: Minimal Antibody Repertoire Size and Reliability of Self-Non-Self Discrimination, *J. Theor. Biol.* **81(4)**, 645–70.

Perelson A S and Weisbuch G (1997). Immunolgy for physicists, *Rev. Mod. Phys.* **69(4)**, 1219–1267.

Petrovsky N and Brusic V (2002). Computational immunology: The coming of age, *Immunology and Cell Biology* **80**, 284–254.

Puzone R, Kohler B, Seiden P and Celada F (2002). IMMSIM, a flexible model for in machina experiments on immune system responses, *Future Gener. Comput. Syst.* **18**(7), 961–972.

Ravasz E, Somera A, Mongru D, Oltvai Z and Barabasi A L (2002). Hierarchical Organization of Modularity in Metabolic Networks, *Science* **297**, 1551–1555.

Ribba B, Alarcon T, Marron K, Maini P and Agur Z (2004). The Use of Hybrid Cellular Automaton Models for Improving Cancer Therapy, *in* P.M.A. Sloot and Bastien Chopard and Alfons G. Hoekstra, ed., *Cellular Automata*, Vol. 3305 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin Heidelberg, pp. 444–453.

Rosvall M and Sneppen K (2003). Modeling Dynamics of Information Networks, *Phys. Rev. Lett.* **91(17)**, 178701(4).

Segel L A (1995). Grappling with Complexity, *Complexity* **1(2)**, 18–25.

Segel L A (2000). Diffuse Feedback from Diffuse Information in Complex Systems, *Comlexity* **5(6)**, 39–46.

Segel L A and Bar-Or R L (1999). On the Role of Feedback in Promoting Conflicting Goals of the Adaptive Immune System, *J. Immunol.* **163**, 1342–1349.

Seiden P and Celada F (1992*a*). A Computer Model of Cellular Interactions in the Immune System, *Immunology Today* **13(2)**, 56–62.

Seiden P and Celada F (1992*b*). A model for simulating cognate recognition and response in the immune system, *J. Theor. Biol.* **158**, 329–357.

Selin L K, Varga S M, Wong I C and Welsh R M (1998). Protective Heterologous Antiviral Immunity and Enhanced Immunopathogenesis Mediated by Memory T Cell Populations, *J. Exp. Med.* **188**, 1705–1715.

Sloot P, Chen F and Boucher C (2002). Cellular Automata Model of Drug Therapy for HIV Infection, *in* S Bandini, B Cjopard and M Tomassini, eds,

ACRI 2002, Vol. 2493 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin Heidelburg, pp. 282–293.

Smith D, Forrest S, Ackley D and Perelson A (1998). Using lazy evaluation to simulate realistic-size repertoires in models of the immune system, *Bull. Math. Biol.* **60(4)**, 647–658.

Solomon S (2001). Private Communication.

Stauffer D and Pandey R B (1992). Immunologically motivated simulation of cellular automata, *Computers in Physics* **6**, 404.

Stauffer D and Proykova A (2004). Cellular automata simulation of medication-induced autoimmune disease, *Physica A* **331**, 545–551.

Stauffer D and Sahimi M (1994). High Dimensional Simulation of Simple Immunlogical Models, *J. Theor. Biol.* **166**, 289–297.

Stauffer D and Weisbuch G (1992). High-dimensional simulation of the shape-space model for the immune system, *Physica A* **180: 1-2**, 42–52.

Swain S L, Hu H and Huston G (1999). Class II-Independent Generation oc CD4 Memory T Cells from Effectors, *Science* **286**, 1381–1383.

Tadic B (2001). Dynamics of directed graphs: the world-wide Web, *Physica A* **293**, 273–284.

Tewari K, Sacha J, Gao X and Suresh M (2004). Effect of Chronic Viral Infection on Epitope Selection, Cytokine Production, and Surface Phenotype of CD8 T Cells and the Role of IFN-γ Receptor in Immune Regulation, *J. Immunol.* **172**, 1491:1500.

Valitutti S, Muller S, Cella M, Padovan E and Lanzavecchia A (1995). Serial triggering of many T-cell receptors by a few peptideMHC complexes, *Nature* **375**, 148:151.

van den Berg H, Rand D and N.J.Burroughs (2001). A Reliable and Safe T Cell Repertoire based on Low-affinity T Cell Receptors, *J. Theor. Biol.* **209**, 465–486.

Wallis G (1999). *The Genetic Basis of Human Disease*, The Biomchemical Society.

Weisbuch G (1990). A shape space approach to the dynamics of the immune system, *J. Theor. Biol.* **143(4)**, 507–522.

Weisbuch G, De Boer R and Perelson A (1990). Localized memories in idiotypic networks, *J. Theor. Biol.* **146**, 483–499.

Welsh R and Selin L (2002). No one is naive: the significance of heterologous T-cell immunity, *Nature Rev. Immunol.* **2(6)**, 417–426.

Wherry E J, McElhaugh M J and Eisenlohr L C (2002). Generation of CD8 T Cell Memory in Response to Low, High, and Excessive Levels of Epitope, *J. Immunol.* **168**, 4455:4461.

Wick D (1998). The Disappearing CD4+T Cells in HIV Infection: a Case of Over-stimulation?, *J. Theor. Biol.* **197**, 507–516.

Williams M A and Bevan M J (2004). Shortening the Infectious Period Does Not Alter Expansion of CD8 T Cells but Diminishes Their Capacity to Differentiate into Memory Cells, *J. Immunol.* **173**, 6694:6702.

Wolfram S (1982). Cellular Automata as Simple Self-Organizing Systems, *Nature* **1**, 1.

Wolfram S (2001). *A New Kind of Science*, Wolfram Media.

Wu Z, Roberts M, Porter M, Walker F, Wherry E J, Kelly J, Gadina M, Silva E M, DosReis G A, Lopes M F, O'Shea J, Leonard W J, Ahmed R and Siegel R M (2004). Viral FLIP Impairs Survival of Activated T Cells and Generation of CD8 T Cell Memory, *J. Immunol.* **172**, 6313:6323.

Wuchty S, Ravasz E and Barabasi A L (2003). The Architecture of Biological Networks, *Complex Systems Science in Biomedicine* .

Wuchty S and Stadler P F (2003). Centers of complex networks, *J. Theor. Biol.* **223**, 45–53.

Yates A, Bergmann C, Hemmen J V, Stark J and Callard R (2000). Cytokine-modulated Regualtion of Helper T Cell Populations, *J. Theor. Biol.* **206**, 539–560.

Yewdell J W and Bennink J R (1999). Immunodominance in Major Histocompatibility Complex Class I-Restricted T Lymphocyte Responses, *Annu. Rev. Immunol.* **17**, 51–88.

Zorzenon dos Santos R M (2000). On the Dynamics of the Evolution of HIV Infection. *ArXiv Condensed Matter e-prints* .

Zorzenon dos Santos R M and Coutinho S (2001). Dynamics of HIV Infection: A Cellular Automata Approaca, *Phys. Rev. Lett.* **87**, 168102–1–4.

Appendices

APPENDIX A

Update Algorithm U

Algorithm 1 Update algorithm U, applied at each time step τ.

repeat
  Randomly select a non-empty cell c_i
  if c_i *is a* ctl_i^- **then**
    for $w = 0$ to $|R_i|$ **do**
      if cell c_w *is a* apc_w^+ *and* $(|c_w - c_i| \leq \rho$ {in shape space}) **then**
        $apc_w^+ := apc_w^{+\bullet}$ {mark c_w as dead}
        $ctl_i^- := ctl_i^{+\star}$ {c_i enters the clonal expansion phase}
        break {out of the for-loop}
      end if
    end for
  else if c_i *is a* $ctl_i^{+\star}$ **then**
    for $w = 0$ to $|R_i|$ **do**
      if cell $c_w = \emptyset$ **then**
        $ctl_w^{+\star} := ctl_i^{+\star}$ {copy c_i into the free cell c_w (clonal expansion)}
        break {out of the for-loop}
      end if
    end for
    for $w = 0$ to $|R_o|$ *and* no free space in R_i **do**
      if cell $c_w = \emptyset$ **then**
        $ctl_w^{+\star} := ctl_i^{+\star}$ {copy c_i into the free cell c_w (clonal expansion)}
        break {out of the for-loop}
      end if
    end for
  end if
  if cell c_w in R_i or $R_o = \emptyset$ **then**
    $c_i \mapsto c_w$ {recirculate the cell c_i to c_w}
  end if
  if age $c_i = 300$ *and* c_i *is a* ctl_i^+ **then**
    $ctl_i^+ := ctl_i^{+\bullet}$ {activated cells switch off at $\tau = 300$ (apoptosis)}
  end if
until $\Phi = 0.99$ {repeat until coverage reaches 99%}
$\tau := \tau + 1$ {increment the clock}

APPENDIX B

Papers Published

Network Topology in Immune System Shape Space

John Burns and Heather J. Ruskin

School of Computing,
Dublin City University,
Dublin 9, Ireland
{jburns,hruskin}@computing.dcu.ie
http://www.dcu.ie/computing/msc/index.html

Abstract. We consider the emergent network topology of an immune system shape space at the end of primary response. We extend the formalism of shape space in order to model the relationship between activated immune lymphocytes and stimulant antigen presentation cells by way of a graph consisting of a pair $G = (V, E)$ of sets. The vertex set V is the set of activated genotypes, while the edge set E connects such activated immune lymphocytes and stimulant antigen presentation cell in shape space. This paper shows how shape space graph edge weighting can be viewed, from the biological perspective, as the vigour with which an activated cytotoxic immune cell suppresses the infected antigen presentation cell which stimulated it. In this research, we also identify critical vertices (called α-vertices). These α-vertices act as *bridging* vertices in that they join subgraphs of unrelated immune response. As a consequence of this, such α-vertices ideally model immune cytotoxic lymphocyte *memory* cells. By representing memory cells as highly connected vertices, we show how such cells play a significant role in the elimination of pathogenic agents.

1 Introduction

In this paper we present results from recent work carried out to model the emergence of shape diversity within the immune system. Previously [1], we introduced a new process by which two formalisms, usually separately addressed, may be integrated. These formalisms are known as shape space and physical space. We highlighted a means by which localised dynamics effect global (or shape space) condition, and how global condition in turn may feed information down to local physical space. This approach is now further refined by treating shape space as a self-organising, dynamic network in 2-dimensional space. The system is considered to be exposed to a set of genetically varied pathogens in order to simulate normal human immune experience over a fixed period of time. We then study the cytotoxic lymphocyte activation patterns which emerge naturally in shape space. The results presented here show that, at the end of primary response, a network of activated cytotoxic lymphocytes and pathogen challengers emerges in shape space.

M. Bubak et al. (Eds.): ICCS 2004, LNCS 3038, pp. 1094–1101, 2004.

The main contribution of this work is as follows: We present a means to model the genotype (or shape) space of the immune system as set of connected, directed and weighted subgraphs. These subgraphs, will, by way of the emergence of critical (or α-) vertices, merge over time. We show that disruption to α-vertex formation degrades immune response more severely than does the disruption of other (which we call β-) vertices. Disruption is likely to occur whenever a viral mutation is a factor, for example, Human Immune Virus (HIV) or Influenza. The means by which such graphs grow, and how *rewiring* of vertices improves response over time, is also investigated. This work demonstrates that edge weighting can be viewed, from the biological perspective, as the vigour with which an activated cytotoxic immune cell suppresses the infected antigen presentation cell which stimulated it. In shape space, this weighting is the distance (d) from α- and β- vertices to the stimulant pathogen (effectively, it is the length of the edge).

2 The Model

In this section we first review some important features of our previous work, and introduce new detail. For a full exposition of both shape and physical space, the reader is directed to [4]. The shape space formalism was introduced by [5] as a way to represent antibody-antigen binding dynamics. Further research refined this model, notably [6] and [7]. The features of cytotoxic lymphocyte (CTL) cells and antigen presentation (APC) cells which govern the dynamics of cell binding (known as the *antigenic determinant*), may be represented by N parameters. If the N parameters are combined into a vector, the antigenic determinant for each APC and each CTL can be considered as points within an N-dimensional Euclidean space of length L_{ss}. Some notational conventions are observed in the work which follows (where upper-case letters refer to shape space, and lower case, to physical space), and this convention is summarised as follows:

1. (CTL^+, ctl^+): activated cytotoxic lymphocyte cells which are ready to attack and remove infected antigen presentation cells. These cells are often referred to as *armed effectors*. The recirculation patterns of ctl^+ are different from ctl^-, in that ctl^+ will leave the lymphatic compartment and migrate to the location of infection detection. Alteration of recirculation patterns is a common feature of cellular immune response in healthy [8] and diseased [9] immune systems.
2. (APC^+, apc^+): infected antigen presentation cells (typically, dendritic cells) which, having engulfed a virus particle, has gone on to present characteristic viral peptide fragments on its surface.

At the start of each simulation, shape space is characterised by two (non-zero) subpopulations: CTL^- and APC^+, representing the number of precursor cytotoxic lymphocyte and active infected antigen presentation cell *genotypes* respectively. A further subpopulation (called CTL^+), arises once an APC^+ is detected by a CTL^-. CTL^+ are activated cytotoxic lymphocyte genotypes. The CTL^+

subpopulation level increases each time another detection occurs, such that the total CTL^- approaches 0. Denoting CTL and APC genotype vectors as $\mathbf{c}$ and $\mathbf{a}$ respectively, we further develop shape space as follows: Surrounding each $\mathbf{c}$ is a disc of radius ρ [1]. Any $\mathbf{a}$ located within this disc will be subject to a clearance pressure inversely proportional to the distance (d) between the $\mathbf{c}$ and $\mathbf{a}$ in shape space ($d = ||\mathbf{c} - \mathbf{a}||$). Our approach is to place into shape space an increasingly diverse set of antigen challenges and test the varying immune response.

Shape space may be further explored using graph theory [10] to model the relationship between APC^+ and CTL^+. In this approach, shape space is a graph consisting of a pair $G = (V, E)$ of sets satisfying $E \subseteq [V]^2$. The set of vertices V is made up of both CTL^+ and APC^+. Edges connect an APC^+ to the set of CTL^+ stimulated by its presence to become activated. Each G is directed and weighted. An initial set of vertices is introduced at time τ_0 based on model startup parameters. At any time $\tau_k > \tau_0$, a new vertex may be added in shape space with probability dependent both on an apc^+ and ctl^- being neighbours in physical space ($P(N)$) *and* the distance (d) between APC^+ and CTL^- in shape space is less than or equal to some threshold $\hat{\rho}$ ($P(d \leq \hat{\rho})$). The outcome of both events are independent of each other, so the probability of a new vertex being added is: $P(newvertex) = P(N)P(d \leq \hat{\rho})$ for any given $\{APC^+, CTL^-\}$ conjugate. A newly added vertex is designated CTL^+ (indicating it has been *recruited* from the CTL^- pool). Whenever a new vertex is added, an edge is added by joining the new vertex to the vertex which stimulated its activation. Thus the relationship represented by an edge between two vertices can be understood as: the CTL^+ *was recruited from* the CTL^- pool by the presence of the APC^+, and, the CTL^+ *acts against* the stimulant APC^+:

$$CTL^- \overset{APC^+}{\prec} CTL^+ \tag{1}$$

and

$$CTL^+ \overset{attacks}{\rightharpoonup} APC^+ \tag{2}$$

respectively.

A new edge is added whenever a new vertex is, but also, whenever a new α-vertex appears. An α-vertex is one which although emerges in response to one individual APC^+ actually effects pressure on other APC^+. As shown in Fig. 1, the α-vertex acts against three APC^+ genotypes (as the APC^+ is the *median* vertex in each subgraph, we use the notation m_i). In so doing, it connects the otherwise unconnected subgraphs of Q, R and S. The importance of the α-vertex is clear: α-vertices are promiscuous, not only targeting the stimulant infected cell but also other APC^+ nearby in shape space (Fig. 1, the m_1, m_2 and m_3 vertices, respectively). The emergence of such vertices marks a diversity threshold of immune response (α_{crit}) which once reached, favours a full and healthy clearance of infected cells from the lymphatic compartment. Such α-vertices are unique in that they participate in Eqn.(2) without having first participated in Eqn.(1). Clearance pressure acts along directed edges which are defined by the triple:

[1] Clearly, with $N = 2$, the area of this disc is $\pi\rho^2$.

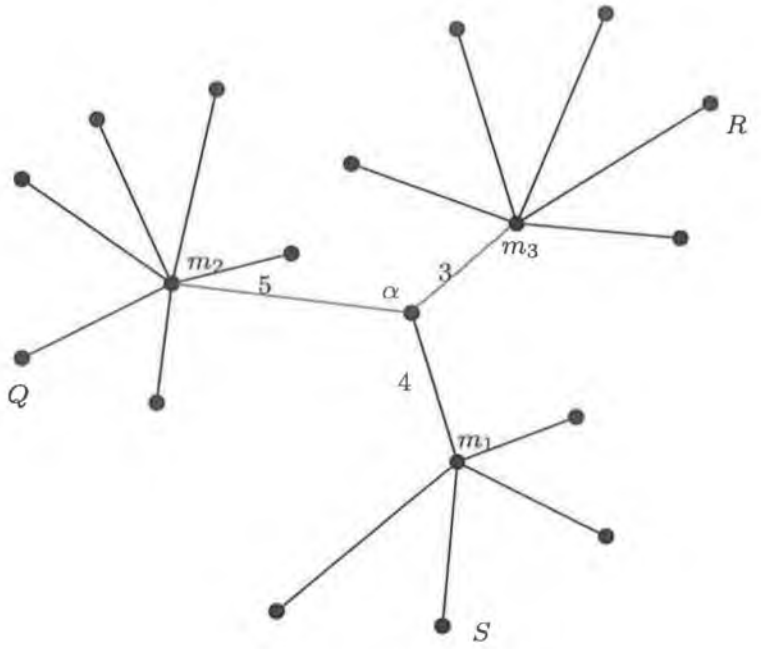


Fig. 1. Three subgraphs with median vertices (stimulant APC^+) at m_i. Leaves (β-vertices) represent activated cytotoxic lymphocyte genotypes (CTL^+). The α-vertex is also a CTL^+, but is one which though originally activated by m_1 now affects clearance pressure against both m_2 and m_3.

$$\vec{E} := \{(e, x, y) | e \in E; x, y \in V; e = xy\}$$

With respect to Fig. 1, the α-vertex clearance pressure acts in three directions (for $i = 1, .., 3$): $\vec{e}_i = (e_i, \alpha, m_i)$.

Rewiring of edges happens whenever a subpopulation of APC^+ disappears, to be replaced by a later APC^+. From the biological point of view, the disappearance of a subpopulation of APC^+ happens whenever the viral genotype challenge is completely eradicated. Rewiring can be viewed as a secondary immune response to some further antigenic challenge. Consistent with [11], we refer the APC^+ as the median vertex (m_i) in a subgraph S. Rewiring is a three step process as follows:

(i) The median vertex m_i is deleted from the graph and the vertices connected to it become disconnected.
(ii) A new median vertex m_j ($i \neq j$) is introduced, representing a new viral infection in the system.
(iii) Each leaf [2] (or β-vertex) rewires to the new median m_j. Depending on the location of m_j in shape space, there is a probability that not all leaves will reconnect to the median vertex (due to $d > \hat{\rho}$) and such leaves will remain unconnected until a *more* central vertex appears. In biological terms, the disappearance of a stimulus APC^+ usually results in a gradual decline of the effector response, to some small, non-zero level which remains as a form of immune *memory*.

[2] A leaf is a vertex of $deg(1)$.

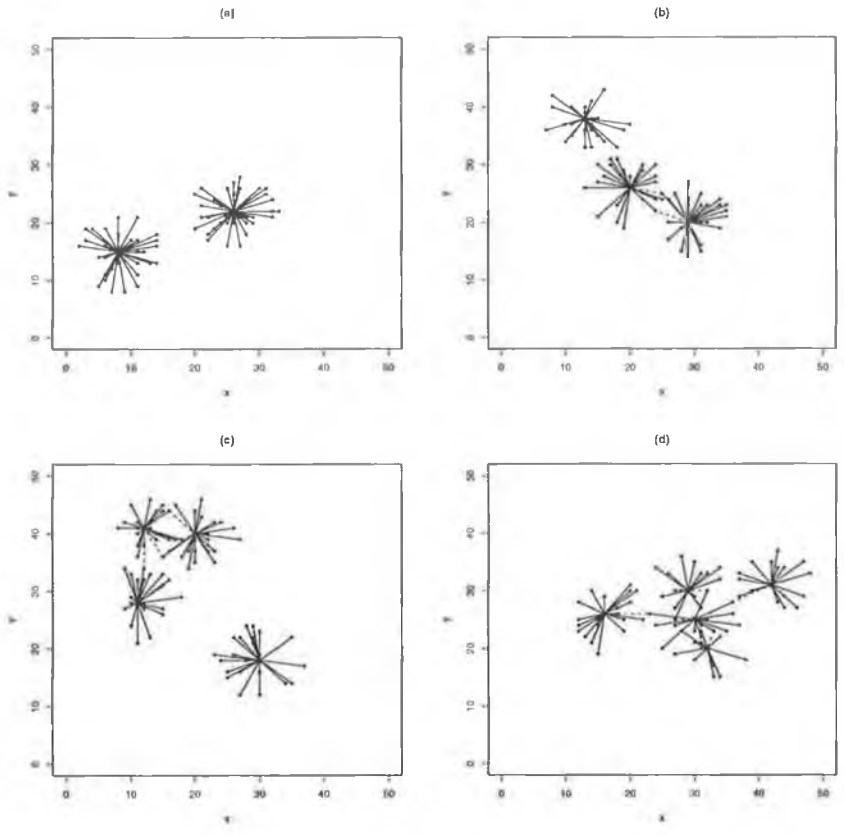


Fig. 2. Four representations of immune shape space at the end of exposure to five antigen genotypes ($\mathcal{R}_{1,..,4}$). The emergence of the α-indexes is indicated by the dashed edges appearing first when the system has been exposed to three antigens (b), with $\mathcal{R}_4$. At exposure to five antigens (d), with $\mathcal{R}_4$, the number of α-vertices is five, and all five of the subgraphs are linked to form one.

3 Results

Fig. 2 shows the state of shape space at the end of primary response. Progressive exposure to varied antigens (values drawn from $\mathcal{R}$) is shown from (a) to (d). In (d), the model has been exposed to 5 different and unrelated antigenic challenges, in much the same way that a maturing immune system would be at the end of $3-4$ years of development. The only parameter varied during simulation execution is $\mathcal{R}$. In (a) and (b), the central (or median) vertex is connected by a set of edges to leaves which appear in response to the prior appearance of the median vertices. Biologically, the infected antigen presentation cell, once recognised, triggers a process (known as clonal expansion) which eventually results in the immune system applying clearance pressure against the infected cell (and all cells presenting the same viral genetic material). In (b), when the system has been exposed to two infections, the subgraphs remain unconnected, indicating that no clearance pressure is applied cross-reactively, and there are no α-vertices. At the point of third antigen challenge, (c), there emerges two α-vertices (acting on the lower two subgraphs). The edges connecting the α-vertices are shown as

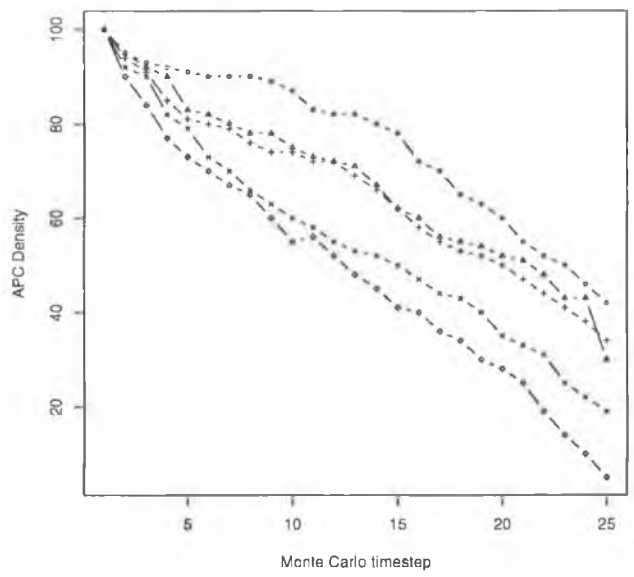


Fig. 3. Clearance rates of apc^+ for $\mathcal{R}_4$, under five different regimes of α-vertex disruption: 1 to 5 (respectively $\diamond$ $\times$ $+$ $\triangle$ and $\circ$), sampled 25 times during model execution and representing some 6 days of real-world time. Selective knock-out (or disruption due to pathogen mutation) of an α-vertex reduces the efficacy of infected cell detection and clearance. When all five vertices have been disrupted (shown by the circle), only $\approx$ 60% of infected cell clearance takes place. In all cases, the β-vertex response remained healthy. In viral pathogens known to mutate slowly (influenza) or quickly (HIV), immune memory (our α-vertices) can often be less effective (as the viral target *drifts* in shape space). The figure shown here offers some insight into consequences of memory cell disruption.

a broken line. At this point, the third infection to appear has caused an activation pattern which results in cross-reactive pressure applied against it. This pressure (by way of the α-vertex) is significant, because it does not emerge over the normal time period of clonal expansion (usually some $3-5$ days). Rather, the pressure is applied *instantaneously* as the α-vertices exist *a priori* and are therefore primed in advance. When four infections have been experienced by the system (Fig. 2 (c)), three of the subgraphs have merged into one, by way of some 4 α-vertices. When the fifth unrelated infection is encountered by the system, a complex network of subgraphs, connected by some 5 α-vertices, has coalesced into one graph. Once the subgraphs merge, the dynamical immune process of up- and down regulation may be explained as follows:

(i) The appearance of m_1 stimulates the development of both α- and β- vertices in R.
(ii) In turn, α-vertex acts to reduce (or down-regulate) m_2
(iii) The down-regulation of m_2 causes a down regulation in β-vertices of Q (as their source of stimulation declines)

And this process may be summarised thus:

$$\uparrow m_1 \Rightarrow \uparrow \{\alpha, \beta \in R\} \Rightarrow \downarrow m_2 \Rightarrow \downarrow \{\beta \in Q\}$$

Once $\alpha_{crit} = 5$ has been reached, the removal of any further infections which may arise is achieved by way of the rapid appearance of α-vertices, and rapid increase in edge density of the connected, weighted graph G in shape space. In Fig. 3, the relative importance of the α-vertices over the β-vertices is shown. This figure shows the model clearance rate of APC^+ from the lymphatic compartment during exposure to five infections. However, in each case, we have explored the effects of disrupting α-vertices. Disruption will occur whenever an APC^+ *drifts* from its original shape space coordinates $(\Delta x, \Delta y)$. Drift is likely to arise whenever an APC^+ mutates (for example, in the case of HIV). Disruption to one or two α-vertices does not seriously degrade clearance, but results in an average reduction of efficiency of $\sim 15\%$. This is intuitive: the importance of the α-vertex lies not in one individual but in the cumulative effect of all. Further disruption results in progressively worse clearance ability. Disruption of all five α-vertices reduces infected cell clearance by around 35%. We would not expect a total immune failure even under full disruption, as there still remains a healthy and effective β-vertex response. Each simulation is repeated 30 times, and the results are averaged. The results shown here were obtained from five separate model simulation runs (for $\mathcal{R}_4$), at each stage, a further α-vertex was suppressed and the edges connecting subgraphs consequentially did not emerge.

4 Discussion and Conclusions

In this research we have provided an outline of how the immune system shape space may be usefully extended to model the process by which infectious agents may be targeted by cells which have been primed in response to a previous and unrelated infection. Using an approach based on graph theory, we identified two qualitatively different vertex types: α and β. Although both vertex types form part of the cellular effector response, we have shown that an effective immune response depends largely on successful α-vertex activation for efficacious response, and only to a lesser extent, on β-vertices. We have seen how disruption to α-vertex activation results in a suppressed response characteristic of chronic infection. In the results presented we have proposed that α-vertices have strong biological equivalent: cytotoxic lymphocyte memory cells. Such cells, having been primed by way of previous immune challenge, require less time to respond, and, crucially, tend to be beneficially cross-reactive. This finding supports [12] and others. Disruption to α-vertices results in a significantly degraded pathogen clearance than disruption to β-vertices does. This supports the theory that during primary and secondary response, some cytotoxic genotypes are *more important than others* [12]. The network implication of α-vertices is that each edge connection formed from a median vertex to an α-vertex acts as a back-bone in joining disparate *subgraphs*. As these subgraphs connect, two related further questions arise (i) up and down regulation of competing CTL^+ *across subgraphs*

and (ii) the emergence of a balanced graph[3]. We have already shown how up- and down- regulation may be explained by extending shape space as an evolving graph (or network).

References

1. Burns, J., Ruskin, H.: A Model of Immune Suppression and Repertoire Evolution. In: Sloot, P.M.A., Gorbachev, Y.E., (eds.): Lecture Notes in Computer Science, Vol. 2660. Springer-Verlag, Berlin Heidelberg New York (2003) 75-85
2. Albert, R., Barabasi, A.: Topology of Evolving Networks: Local Events and Universality. Phys. Rev. Lett. 85(24) (2000) 5234-5237
3. Buseyne, F., and Riviere, Y.: The flexibility of the TCR allows recognition of a large set of naturally occurring epitope variants by HIV-specific cytotoxic T lymphocytes Int. Immunol. (13) (1999) 941-950
4. Burns, J., Ruskin, H.: Viral Strain Diversity and Immune Response - a Computational Model. In: Hamza, M.H., (ed): Proceedings of the IASTED International Conference, Biomedical Engineering. ACTA Press (2003) 60-65
5. Perelson, A.S., Oster, G.F.: Theoretical Studies of Clonal Selection: Minimal Antibody Repertoire Size and Reliability of Self-Non-Self Discrimination. J. Theor. Biol. 81(4) (1979) 645-670
6. de Boer, R.J., Segel, L.A., Perelson, A.S.: Pattern formation in one- and two-dimensional shape-space models of the immune system. J. Theor. Biol. 155(3) (1992) 295-333
7. Smith, D., Forrest, S.: Deriving shape space parameters from immunological data. J. Theor. Biol. 189 (1997) 141-150.
8. Janeway, C.A., Travers, P., Walport, M., Capra, J.D.: Immunobiology. The Immune System in Health and Disease. 4th edn. Churchill-Livingston (1999)
9. Wick, D.: The Disappearing CD4+T Cells in HIV Infection: a Case of Overstimulation?. J. Theor. Biol. 197 (1998) 507-516
10. Diestel, R.: Graph Theory. 2nd edn. Spinger (1997)
11. Wuchty, S., Stadler, P.: Centers of complex networks. J. Theor. Biol. 223 (2003) 45-53
12. Brehm, M.A., Pinto, A.K., Daniels, K.A., Schneck, J.P., Welsh, R.M., Selin, L.K.: T cell immunodominance and maintenance of memory regulated by unexpectedly cross-reactive pathogens Nat. Immunol. 3(7) (2002) 627-634

[3] A graph G is *balanced* if the maximum ratio of edges to vertices, taken over all subgraphs of G, occurs at G itself

A Stochastic Model of the Effector T Cell Lifecycle

John Burns[1,2] and Heather J. Ruskin[1]

[1] School of Computing, Dublin City University, Dublin 9, Ireland
[2] Department of Computing, Institute of Technology Tallaght, Dublin 24, Ireland
jburns@computing.dcu.ie

Abstract. The dynamics of immune response to initial infection and reinfection by the same pathogen sometime later, are considerably different. Primary response, which follows initial infection, is characterised by relatively slow precursor cell activation and population growth rates, with a consequent elongated pathogen clearance profile, typically extended over six days or more. On the other hand, secondary response (to reinfection by the same pathogen some time later) is notable for short effector activation time, high specificity of response, rapid pathogen elimination and high degree of memory cell participation. In this paper, we present a seven state non-deterministic finite automata (NFA) of the effector T cell lifecycle, which is encoded as a set of states and state transitions. Our objective is to study the degree to which variable infection outcome is dependent on the accumulation of *chance* events. Such chance events may be represented as the consequence of premature, delayed or even failed state transitions. We show how small variation in crucial state transitions probabilities during the lifecycle can induce widely variable infection outcomes. This model is implemented as a spatially extended, concurrent two-dimensional stochastic cellular automata, executing on a MPI-based Linux cluster.

1 Introduction

Cellular Automata (CA) have been applied to numerous areas of complex physical systems modelling [1]. CA have several important characteristics which make them amenable to efficient computational implementation, including ease of representing (in the form of n-dimensional arrays), discrete nature of the underlying computations, simplicity of rules or laws which are programmed into the CA, and the highly repetitious nature of the processing steps. However, cellular automata posses additional fascinating properties, for example, patterns of self-organisation of a complexity which cannot be derived numerically from the rules on which the underling cellular automata is based. As a result of this complexity, [2] has postulated that some form of CA must underlie many complex physical phenomena visible in nature. Furthermore, with the application of non-deterministic (*stochastic*) cellular automata, the idea of randomness in CA site

P.M.A. Sloot, B. Chopard, and A.G. Hoekstra (Eds.): ACRI 2004, LNCS 3305, pp. 454–463, 2004.

selection and update rule enforcement has yielded further insight into modelling stochastic natural systems such as molecular motion, turbulence in water flow and various biological processes, especially models of the human immune system [3,4,5,6,7]. In this paper we present an approach that seeks to avoid a computational modelling process exclusively influenced by current experimental immune system research trends. We propose a *relaxation* of the deterministic assumptions inherent in earlier work [8], and explore the dynamics of a more stochastic system. Stochastic events appear to play a crucial role in certain immune system functions [9]. The contribution of this work is as follows: (i) an extended non-deterministic state transition model of the effector T cell lifecycle is introduced. This model successfully reproduces time-realistic effector and pathogen population dynamics during primary and secondary response, and during repeated reinfection, (ii) we identify three stages in the effector T cell lifecycle model which are critical in regulating the course of primary and secondary response, and (iii) the model is implemented as a spatially extended two-dimensional cellular automata lattice executing concurrently on a MPI-based Linux cluster. This allows us to scale the model to cell density levels in the order of 10^6 CTL cells - which approaches levels typically observed in *in-vivo* murine experiments. This work is arranged as follows: section 2 is a brief overview of some key features of the adaptive immune response which we model, and serves as an introduction to some specific terminology. Section 2 is intended for readers who may be unfamiliar with general principles of immunology. Section 3 discusses the model structure and explains the motivation and implementation of the underlying non-deterministic cellular automata. Section 4 presents results of the simulation, and in particular, some interesting features which emerge. Finally, section 5 is a discussion of the results, and an outline of further enhancements.

2 Adaptive Immune Response

Common to all immune systems is the principal of sensing of localised space for the purposes of intrusion detection. Intrusion, in this case, is the appearance of a bacteria, viral particle or infected cell which may be classified as *non-self*. Any non-self genetic material discovered must be eliminated in order to prevent infection (or even death), of the host. Broadly speaking, the means by which the non-self intruder gained access to the blood stream or lymphatic compartments is not of interest[1]. There are a great variety in the pathogen challenge and immune response course (not all of which are a concern here). One such scenario arises as follows: when a viral particle has been taken up by an antigen-presenting cell (APC), such as a dendritic cell, it is degraded into one or more peptide chains within the cytosol region of the APC, and is then bound to the major histocompatible complex (MHC) class I molecule (a process known as *antigen processing*) before finally being presented on the surface of the APC as

[1] Some viruses, for example, the influenza and corona viruses, enter the host through the air passages and not through tissue damage.

an MHC:peptide complex, a process known as *antigen presenting*. APC will recirculate through the lymphatic system in order to alert the immune system to an infection. Sensing of the lymphatic compartments (of which there are many) for antigen-presenting cells, is a systematic function of immune cell (lymphocyte) recirculation. Cytotoxic lymphocyte (CTL) precursor cells constantly recirculate and sample their environment in the search for foreign pathogens. The process of sampling involves two cells binding for some small time period, during which the immune cell senses the receptor of the bound cell to determine if the bound cell is an invading pathogen (or not). If the bound cell *is* an invading pathogen, the immune cell may be stimulated to produce clones of itself in order to attack and remove other cells bearing the same genetic material. Under normal circumstances, the production of clones ceases after some fixed period of time, and once the infection has been cleared, most CTL cells will undergo programmed death (apoptosis). A small subset of the clone population will remain activated indefinitely, and this population represents effector memory. In the presented here, we do not model free antigen, but only antigen epitopes which have been bound to the surface of an antigen presenting cell.

3 The Model

Our model runs in discrete 30-minute timesteps, and all entities in the model act asynchronously at each timestep (τ). As primary response normally consists of 4 days of cell replication (clonal expansion), the cells in our model will stop dividing at $\tau = 192$. The recirculation space of the lymphatic compartment is modelled as a two dimensional stochastic cellular automata lattice of length $L = 10^3$, with periodic boundary conditions and neighbourhood radius $r = 1$ (in two-dimensions), with a maximum of 8 neighbours. Each site is selected at random for update during the timestep. Not every site is will be visited at each timestep, but each site can be updated at most once in any given timestep. At $\tau = 0$ some 5000 antigen entities are introduced into randomly selected sites on the lattice (following a uniform distribution), and the model executes until $\tau = 3000$ (62.5 days of elapsed time). The CTL population grows exponentially in response to APC stimulation, with a clonal expansion rate which is a function of the *affinity* between the CTL and APC. The dynamics of affinity are modelled using shape space [10,11]. The stimulation rate never exceeding 0.036, which yields a population of daughter clones of ~ 1000 after 4.5 days of clonal expansion. Each lattice site may contain only one entity at any given timestep. The set of entities and states supported is shown in Table 1, which also introduces some important notation used throughout this paper.

3.1 Non-deterministic Finite Automata

To allow the study of a *distribution* of possible outcomes, we identify a subset of the CTL lifecycle state transitions, and replace the certainty of a transition from state w to state v on event e with some probability (< 1) of state transition.

Let us start by defining what is meant by state transition relaxation: If X is a discrete random variable (drv) representing the transition from state w to state v, and e is some event linking wv, the relaxed state transition X_r is:

$$P(X_r|e) = 0 \leq \psi \leq 1 \tag{1}$$

The choice of value for ψ will naturally depend on the wv in question. In contrast to earlier models [2], Eq. (1) implies *duality* in the presence of event e: transition on e (X_r) or not ($\bar{X}_r$). This extension results in a *non-deterministic* finite automaton (NFA) [12]. Fig. (1) is a non-deterministic finite automata model of the lifecycle of the CTL (and follows notation explained in Table 1). E is the set of events the model, and consists of both deterministic and non-deterministic elements. We define a subset of three critical non-deterministic events $S \subset E$ as: $S = \{e_{2,8}, e_3, e_5\}$. Each $e_i \in S$ is defined as follows:

e_2, e_8 An infected antigen presenting cell will be destroyed by a bound cytotoxic lymphocyte cell which recognises it. Recognition is a function of the distance between the two cells in shape space.

e_3 An activated proliferating immune cell (state $ctl^{+\star}$) will normally end clonal expansion on the event $\langle e_3 : age(ctl^{+\star}) > 192 \rangle$.

e_6 The fraction of effector T cells entering the long-lived memory pool. Normally the majority of activated effector cells undergo programmed cell death (apoptosis) at the end of primary response. However, recruitment to the memory pool consumes around 5 – 10% of activated CTL [13,14, 15], thus, a further stochastic transition occurs on e_6, with of $ctl^{+\dagger}$ enter $ctl^{+\bullet}$ on event $\langle e_6 : age(ctl^{+\dagger}) \geq 192 \rangle$.

e_{rpt} Repeated reinfection events, resulting in repeated doses of infected antigen presenting cells introduced into the simulation, at timestep $\tau + 300n, n = 0, 1, ..., 9$.

Each of the above events (e_n) has an associated probability ψ_n. The set $\{\psi_1, \psi_2, \psi_3, \psi_4\}$, therefore fully describes each simulation configuration of the model (all other parameters being kept constant). In the results presented in the following section, we define the following four experimental onfigurations of $\mathcal{P}$:

1. $\mathcal{P}_1 : \{0.9, 0.9, 0.9, 0.0\}$
2. $\mathcal{P}_2 : \{0.9, 0.9, 0.95, 0.0\}$
3. $\mathcal{P}_3 : \{0.9, 0.9, 0.9, 1\}$
4. $\mathcal{P}_4 : \{0.9, 0.9, 0.95, 1\}$

The first two configurations of $\mathcal{P}$ test the fidelity of the model response when confronted with a singular secondary infection event some 30 days after the initial infection. The first configuration represents a normal response and is intended to calibrate the model for near optimal conditions. For $\mathcal{P}_1$, we would expect to see CTL production levels broadly characterised by low, elongated peak for primary infection, followed by an increase in memory CTL. Another expected observation

[2] in that $P(X_r|e) = 1$

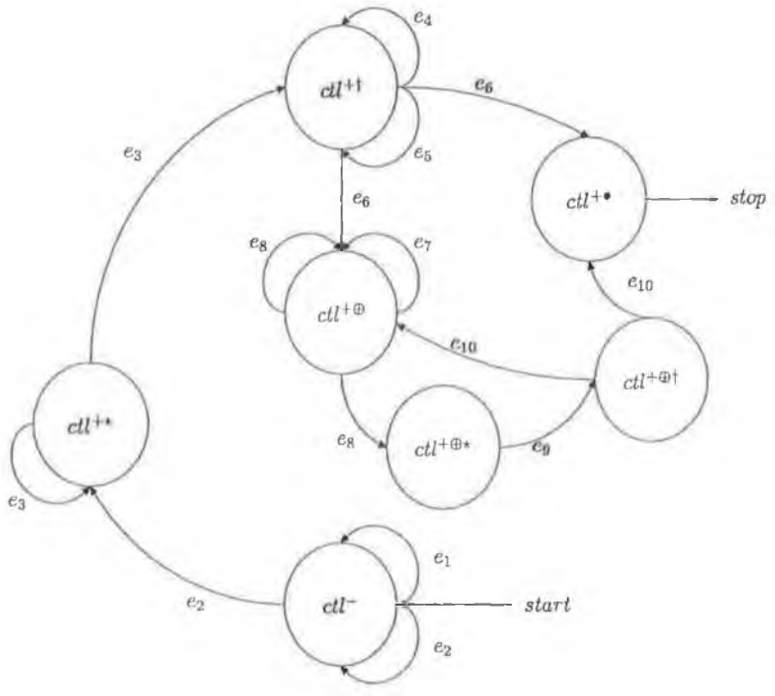


Fig. 1. A seven-state non-deterministic finite automata of the cytotoxic lymphocyte cell lifecycle. Transition events (e_n), which carry the same label, are *non-deterministic*.

is APC clearance: over some $6-10$ days for primary response, and significantly faster during secondary response. The second configuration is an increase in ψ_3 from 0.9 to 0.95 and is intended to test the impact of a 5% decline in the number of cells which transition to the effector memory state ($ctl^{+\dagger} \rightarrow ctl^{+\oplus}$). Some viral infections are known to case damage or loss of the memory pool [16], and we test to see the impact this has on our model. We test repeated reinfection with normal and depleted memory cell production ($\mathcal{P}_3$ and $\mathcal{P}_4$, respectively). Many pathogens are known to lead to acute and persistent viral infections, and we test the importance of memory cell production in these cases. Again we deplete the memory production by 5% and study the consequences of this loss. Section 4.1 examines the results of persistent infection in our model.

4 Results

The model is initially executed with parameter set $\mathcal{P}_1$ and $\mathcal{P}_2$ (with no repeat reinfection), and the results are shown in Fig. 2. In (a), the initial infection is visible at $\tau = 0$ with pathogen density $p_d = 5000$, (the broken line) and consequent effector response reaching a maximum value at $\tau = 300$, with $e_d = 8.2 \times 10^3$. Fig. 2(b) shows the antigen presenting cell population level (only). No memory cells are present during primary response, and as such, the effector cell population is made up entirely of clones produced by stimulated precursor cells. To the right of each effector cell peak is a plateau of memory cells. The slope of the CTL density peak is extreme, indicating that the state transitions

Table 1. Notation and definition of model entity states

Notation	Definition
ctl^{-}	naive recirculating effector precursor
$ctl^{+\star}$	proliferating lymphocyte
$ctl^{+\bullet}$	dead activated lymphocyte (apoptosis)
$ctl^{+\oplus}$	activated memory effector
$ctl^{+\oplus\star}$	activated proliferating memory effector
$ctl^{+\oplus\dagger}$	activated memory effector
$ctl^{+\dagger}$	armed activated effector
apc^{+}	active infected antigen presenting cell
$apc^{+\bullet}$	dead infected antigen presenting cell

from $ctl^{+\star}$ to $ctl^{+\dagger}$ to $ctl^{+\oplus}$ (or $ctl^{+\bullet}$) occurring with a high degree of certainty. At time $\tau = 1500$ (day 31), secondary exposure to the same pathogen occurs, and the model exhibits following general behaviour: (i) the secondary immune response is preceeded by a pool of committed CTL memory cells which have already been primed to respond to the re-appearing pathogen, (ii) the activated CTL density is some 10 times higher than primary response, and does not last as long, and (iii) the pathogen is reduced to half its original level much more rapidly than during primary response. With $\mathcal{P}_1$, the model exhibits efficient detection and clearance behaviour associated with a healthy immune system. From Fig. 2, it can be seen the advantage in both time and infected cell clearance which is conferred on a response based largely on memory: the half life of the virus during primary response is around 3.25 days, with 90% pathogen clearance achieved at around $\tau = 480$, or 10 days of simulation time. Compared to secondary response on reinfection we see an infected cell half life of $\tau \approx 60$ or 1.25 days - an efficiency of around 87%. Effectively, this is because memory cells, having already been primed by a previous encounter with the specific pathogen, undergo expansion with lower death rates than during primary response: they therefore accumulate more quickly [17]. The results for $\mathcal{P}_2$ are shown in Fig. 2(c) and (d). Here, the probability of entering apoptosis is increased from 0.9 to 0.95. This means that the memory cell population would be around 5% of that activated effector population post-primary response. Recent work (notably [17]) has shown that some $\approx$ 90% of activated effector undergo apoptosis after primary response. Therefore, $\psi_3 = 0.95$ would represent an unusually high suppression of memory function. Clearly, the reduction of memory effector production should not effect primary response, and this is borne out by CTL density levels prior to $\tau = 1500$ (c). We see a normal 10-day clearance regime (d) during primary response, but a less effective response during reinfection: in fact, the memory cell pool in the time range $500 \leq \tau \leq 1500$ has fallen to ≈ 500. Once reinfection occurs, the APC

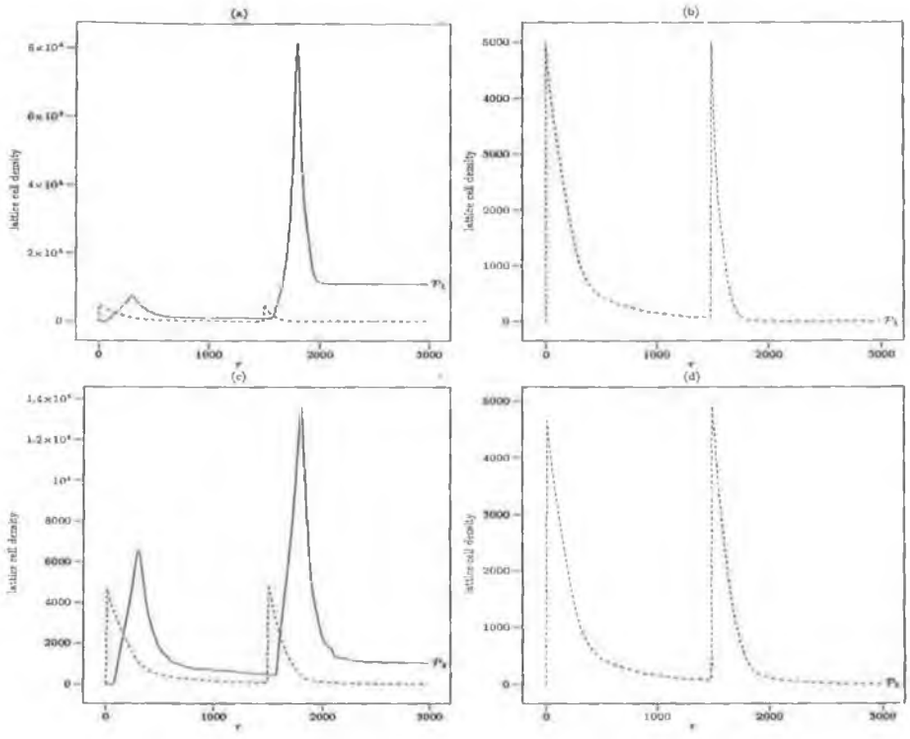


Fig. 2. CTL and pathogen lattice density levels (a),(c) over a simulated 62.5 day period, with an initial infection at time $\tau = 0$ and a reinfection by the same pathogen occurring at $\tau = 1500$, for 3 values of $\mathcal{P}$. Antigen presenting cell (APC) density is shown by the broken line, with the solid line indicating levels of effector memory and activated cells combined. For clarity, (b),(d) show population levels for APC for each $\mathcal{P}$.

population is cleared some 31% more effectively than during primary response. The APC half life is $\tau = 108$, 90% clearance is achieved after reinfection at $\tau \approx 1788$ (or some 5.9 days of simulated time). However, the characteristics of $\mathcal{P}_2$ are significantly degraded compared to that observed in $\mathcal{P}_1$.

4.1 Persistent Reinfection

Some viral pathogens are capable of persistent reinfection, in that, although population levels of infected antigen presenting cells may decline in response to clearance pressure by a specific CTL response, over time, the number of infected cells rises to chronic and sometimes acute levels. Examples of such viruses are HIV, HTLV, hepatitis C (HCV), hepatitis B virus, CMV EBV and rubella [16]. Such persistent reinfection pathogens have been associated with normal immune function suppression. In this section, we simulate persistent reinfection by randomly scattering a repeat 'dose' of the pathogen, introduced at $\tau + 300n, n = 0, 1, ..., 9$. This reinfection pattern is a represents a resurgence of infected cells every 6.25 days, in discrete bursts. The results of this simulation are shown in Fig. 3.

With respect to Fig. 3 (a), the response to the first reinfection is clearly strong: some 3.8×10^5 lymphocytes are generated and the reinfection is rapidly

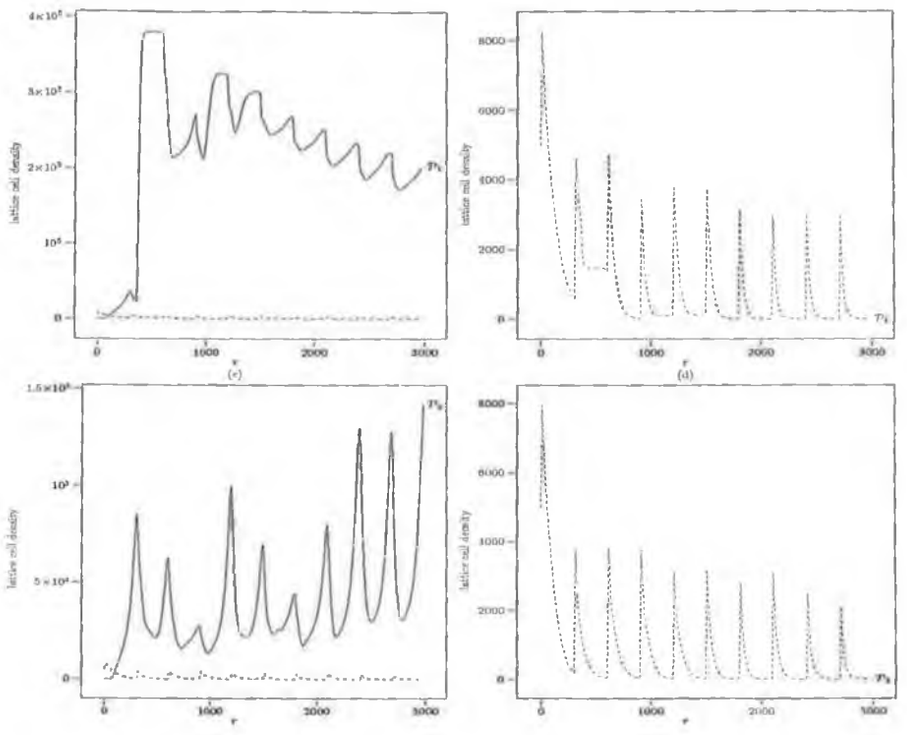


Fig. 3. The model is exposed to repeated infection events, arising at time $\tau = 300n, n = 0, 1, ..., 9$, equivalent to an infection every 6 days.

eliminated. As further infections arise starting at $\tau = 600$, the existing memory pool never falls below 1.8×10^5, and is critical in bringing the repeated reinfections under control in time periods (b) which rarely exceed 130 timesteps (or 2.8 days of simulated time). We also see from (a) that slightly lower responses are sufficient in order to effect optimal clearance. Results from (a) and (b) support the clinical findings that the memory cell levels tends to be higher after secondary and tertiary infections [17], which in turn, supports the clinical practice of vaccination boosting. Finally, when the simulation is executed with diminished memory cell creation and repeatedly stressed with reinfection ($\mathcal{P}_4$), average primary and secondary response levels are similar (around 1.2×10^4). Each response is characterised by rapid expansion and reduction of effector lymphocyte clones. There are no memory cells to confer clearance advantage, and each response is initiated from low levels (around 1.2×10^2).

5 Discussion and Conclusions

The approach taken in this research was to construct a stochastic model of the effector T cell lifecycle, in order to study a distribution of possible simulation outcomes. We have shown how the model reproduces well the time and space dynamics of initial and secondary infection. In addition, we believe the research is valuable in modelling the relationship between repeated reinfection and effector

cell transition to memory or apoptosis. We have demonstrated how repeated reinfection can be controlled only within a limited range of ψ_3: too much memory causes the lymphatic compartment to fill-up, too little memory induces the need for clonal expansion from naive precursor cells, and a elongated APC clearance profile. When the ratio of apoptosis to memory is 'just right' ($0.88 \leq \psi_3 \leq 0.92$), antigen presenting cell levels (during repeated reinfection) are brought under control in increasingly rapid time frames. The next steps in this research are to test the homeostasis of our model: where does the model break down, and what insight does this provide. Very recent clinical work [16] suggests that the immune system must periodically preferentially eliminate some memory cells which exhibit poor cross-reactivity. One of the benefits of the the stochastic effector T cell lifecycle model presented here is the relative ease with which this theory could be investigated. The benefits of selective memory cells reduction may form the basis of further work with this model.

References

1. Wolfram, S.: Cellular Automata as Simple Self-Organizing Systems. Nature **1** (1982) 1
2. Wolfram, S.: A New Kind of Science. Wolfram Media (2001)
3. Stauffer, D., Pandey, R.B.: Immunologically motivated simulation of cellular automata. Computers in Physics **6** (1992) 404
4. Castiglione, F., Bernaschi, M., Succi, S.: Simulating the Immune Response on a Distributed Parallel Computer. Int. J. Mod. Phys. **8** (1997) 527
5. Mannion, R., Ruskin, H., Pandey, R.: Effect of Mutation on Helper T-cells and Viral Population: A Computer Simulation Model for HIV. Theor. in Biosci. **199(2)** (2000) 145–155
6. dos Santos, R.M.Z., Coutinho, S.: Dynamics of HIV Infection: A Cellular Automata Approach. Phys. Rev. Lett. **87** (2001) 168102
7. Bernaschi, M., Castiglione, F.: Design and implementation of an immune system simulator. Computers in Biology and Medicine **31** (2001) 303–331
8. Burns, J., Ruskin, H.J.: Network Topology in Immune System Shape Space. In Sloot, P., Gorbachev, Y., eds.: Computational Science - ICCS 2004. Volume 3038 of Lecture Notes in Computer Science., Berlin Heidelberg, Springer-Verlag (2004) 1094–1101
9. Germain, R.N.: The Art of the Probable: System Control in the Adaptive Immune System. Science **293** (2001) 240–245
10. Perelson, A.S., Oster, G.F.: Theoretical Studies of Clonal Selection: Minimal Antibody Repertoire Size and Reliability of Self-Non-Self Discrimination. J.Theor. Biol. **81(4)** (1979) 645–70
11. Burns, J., Ruskin, H.J.: Diversity Emergence and Dynamics During Primary Immune Response: A Shape Space, Physical Space Model. Theor. in Biosci. **123(2)** (2004) 183–194
12. Hopcroft, J.E., Ullman, J.D.: Introduction to Automata Theory, Languages and Computation. Addison Wesley (1979)
13. Murali-Krishna, K., Lau, L.L., Sambhara, S., Lemonnier, F., Altman, J., Ahmed, R.: Persistence of Memory CD8 T Cells in MHC Class I-Deficient Mice. J. Exp. Med. **286** (1999) 1377–1381

14. De Boer, R.J., Oprea, M., Kaja, R.A., Murali-Krishna, Ahmed, R., Perelson, A.S.: Recruitment Times, Proliferation, and Apoptosis Rates during the CD8$^+$ T-Cell Response to Lymphocytic Choriomeningitis Virus. J. Virol. **75(22)** (2001) 10663–10669
15. Badovinac, V.P., Porter, B.B., Harty, J.T.: CD8+ T cell contraction is controlled by early inflammation. Nat. Immunol. **5** (2004) 809–817
16. Kim, S., Walsh, R.: Comprehensive early and lasting loss of memory CD8 T cells and functional memory during acute and persistent viral infections. J. Immunol. **172(5)** (2004) 3139–3150
17. Grayson, J., Harrington, L., Lanier, J., Wherry, E., Ahmed, R.: Differential Sensitivity of Naive and Memory CD8+ T cells to Apoptosis in Vivo. J. Immunol. **169(7)** (2002) 3760–3770

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Theory in Biosciences 123 (2004) 181–193
Theory in Biosciences
www.elsevier.de/thbio

Diversity emergence and dynamics during primary immune response: a shape space, physical space model

John Burns*, Heather J. Ruskin

School of Computing, Dublin City University, Dublin 9, Ireland

Received 12 February 2004; accepted 3 May 2004

Abstract

A computational model of the dynamics of diversity among T-cell receptors and MHC:peptide complex molecules is presented. We propose a method by which individual immune systems may evolve effcient or ineffcient states as a result of T-cell receptor cross-reactivity as well as genetic variation among pathogens. By combining shape space and physical space models, valuable insight is obtained into how immune system-wide state is, in large part, determined by localised space dynamics. In the model, system-wide state also informs local dynamics, especially in the lymphatic system during primary immune response. The process by which similar initial infection conditions across individuals may result in highly variable end states (a phenomenon observed in the clinical context) is modelled. Our results show that activity alone is not a good indicator of infection suppression or removal. In this work, we postulate that successful viral clearance is characterised by broad T-cell receptor activation (in shape space), and results in low average concentration levels of activated cytotoxic lymphocyte cells.

Keywords: Shape space; Cross-reactivity; Repertoire diversity

*Corresponding author.
E-mail addresses: jburns@computing.dcu.ie (J. Burns), hruskin@computing.dcu.ie (H.J. Ruskin).

1431-7613/$ - see front matter © 2004 Elsevier GmbH. All rights reserved.
doi:10.1016/j.thbio.2004.05.001

Introduction

When a viral particle has been taken up by an antigen-presenting cell (APC), such as a dendritic cell, it is degraded into one or more peptide chains within the cytosol region of the APC, and is then bound to the major histocompatible complex (MHC) class I molecule (a process known as *antigen processing*) before finally being presented on the surface of the APC as an MHC:peptide complex, a process known as *antigen presentation*. As the immune system may be faced with an infinite number of genetically varied challengers, the final form of the MHC:peptide complex may be characterised by enormous structural variability. To reliably detect this antigenic variation, the immune system generates its own diversity in the form of a set of T-cells capable of recognising MHC:peptide sequences, by means of a T-cell receptor (TCR), with variable degrees of efficacy (Buseyne and Riviere, 1999). The complete TCR set is known as the *immune repertoire*. During T-cell maturation in the thymus, genes encoding for the T-cell receptor undergo several cycles of rearrangement, resulting in a mature immune repertoire capable of recognising a large range of MHC-bound non-self peptides. In this paper, we study the emergence and dynamics of two main forms of immune diversity: (i) non-self peptides bound to MHC class I molecules, and (ii) variation in the CDR (complementarity-determining region) of the T-cell receptor. The affinity with which a T-cell receptor binds to the MHC:peptide complex arises from the sum of the binding interactions among the CDR and the exposed peptide (Germain and Stefanova, 1999). The authors further indicate that the variation in affinity at the TCR:MHC:peptide bind site can dictate whether the pathogen challenge has the properties of agonist, partial agonist, antagonist or null compound. All agonists (both strong and weak) will cause the T-cell to begin a process which will eventually end in the death of the infected APC, and which is characterised by the onset of *clonal expansion* (whereby the T-cell which successfully bound the MHC:peptide complex gives rise to a population of clones, each bearing the same TCR). Partial agonists may not trigger T-cell effector response, and thus may not result in the death of the infected APC. Antagonists inhibit the functioning of the T-cell effector, and null compounds do not interact with the TCR strongly enough to cause any signal transmission: the T-cell will simply sample the MHC:peptide complex and move on. The main contribution of this work is as follows: we present a computational model of the first 6 days of primary immune response. Our model simulates the physical space of the lymphatic system in which naive cytotoxic lymphocyte (CTL) precursor cells constantly recirculate and sample their environment in the search for MHC:peptide-bearing APC. Diversity among TCR and MHC:peptide is modelled and analysed by means of an extension to the shape space formalism and integration to a physical space model. A parameter is introduced to model affinity threshold, with variation in this parameter (denoted ρ) strongly correlated to APC clearance rates. The results here show how an immune repertoire configuration may be characterised as *efficient* or *inefficient*, and how a highly cross-reactive TCR repertoire (with consequentially low specificity) affects efficient antigen removal when presented with MHC:peptide complexes. The characteristics of every cytotoxic lymphocyte and

every antigen-presenting cell is determined by the TCR and MHC:peptide complex on the surface of each, respectively. For clarity of language, we assume that APC is synonymous with MHC:*peptide complex presented on the surface of the APC*, and any further reference to APC phenotype refers to the manifestation of specific structural and binding properties of the peptide complex on the cell surface. This paper is organised as follows: the next section introduces the model, and explains the underlying details of both physical and shape space. The third section presents the results and explores the role of key model parameters. Finally, the last section is a discussion of the implications of the work.[1]

The model

The shape space formalism was introduced by Perelson and Oster (1979) as a way to represent antibody–antigen binding dynamics. Further work refined this model, notably de Boer et al. (1992) and Smith et al. (1997). In this paper, we extend the formalism to model CTL and APC diversity and immunodominance, as well as to show how strong and weak agonists can be modelled and analysed. In shape space, we are primarily interested in phenotype repertoire distribution and its differentiation. For simplicity, assume that the features which govern the binding of the CTL receptor and APC, can be represented by N integer parameters. If the N parameters are combined into a vector, the *phenotype* for each CTL and APC can be considered as points within an N-dimensional Euclidean space of length L_s. Cells of the same phenotype have equal shape space vectors, and reside at the same location in shape space. Denoting CTL and APC phenotype vectors as $\mathbf{c}$ and $\mathbf{a}$, respectively, we further develop shape space as follows: Surrounding each c is a disc of radius r.[2] Any $\mathbf{a}$ located within this disc will be subject to a clearance pressure inversely proportional to the distance (d) between the $\mathbf{c}$ and $\mathbf{a}$ in shape space ($d = ||\mathbf{c} - \mathbf{a}||$). A strong agonist is one for which $d \to 0$, an increasingly weak agonist is one where $d \to r$, while for a null compound $d > r$. Every $\mathbf{c}$ will have a set of agonist APC phenotypes, $A_{\mathbf{c}}$, where $|A_c| \geq 0$. Conversely, every $\mathbf{a}$ belongs to *at least one* $A_{\mathbf{c}}$ (if this were not the case, then some $\mathbf{a}$ would remain undetected indefinitely). It is axiomatic that shape space is always completely *covered*. That is to say, with n CTL phenotypes and $N = 2$:

$$n\pi r^2 \gg L_s^2. \tag{1}$$

With the same dimensionality of (1), this leads to $r_{\max}$, the fixed upper-bound of r:

$$r_{\max} = \sqrt{\frac{L_s^2}{\pi}}. \tag{2}$$

[1] A copy of the source code is available for download at http://www.computing.dcu.ie/~jburns/

[2] Clearly, with $N = 2$, the area of this disc is πr^2.

Table 1. Shape space subpopulation invariants.

Time	Invariant		
0	$CTL^- > 0$	$CTL^+ = 0$	$APC^+ > 0$
$0 < \tau < 300$	$0 \leq CTL^+ < CTL^-$	$CTL^- \rightarrow CTL^+$	$APC^+ > 0$
300	$CTL^- > 0$	$CTL^+ > 0$	$APC^+ \rightarrow 0$

Cross-reactivity, the probability that a CTL will bind to an APC, is denoted by ρ, and calculated as $\rho = r/r_{\max}$. Cross-reactivity is a system-wide parameter in shape space with the following effect: increasing cross-reactivity results in decreasing specificity, in turn this causes lower stimulation to the CTL during clonal expansion. Below the affinity threshold ($d \leq r$), the stimulation rate S_r, used during the clonal expansion phase is calculated as

$$S_r = \begin{cases} 1 & : \quad d = 0, \\ \frac{1}{d} & : \quad 0 > d \leq r. \end{cases} \tag{3}$$

At the start of each simulation, shape space is characterised by two (non-zero) subpopulations: CTL^- and APC^+ - representing the number of precursor cytotoxic lymphocyte and active infected antigen-presenting cell phenotypes, respectively. As time progresses, a further subpopulation emerges: CTL^+, representing activated cytotoxic lymphocyte phenotypes. The CTL^+ subpopulation is recruited from the available pool of CTL^-. The process by which naive cytotoxic lymphocytes are stimulated to become activated cytotoxic lymphocytes occurs ($CTL^- \rightarrow CTL^+$) is described later in this section. In a healthy individual, the typical clearance rate of infected antigen-presenting cells from the lymphatic system is of the order of 3–5 days. At the end of primary response, the subpopulation of infected antigen-presenting cells will tend to be eliminated ($APC^+ \rightarrow 0$). As the thymus ensures a supply of mature cytotoxic lymphocyte precursor cells into the lymphatic system, our model follows this by ensuring that a non-zero naive cytotoxic lymphocyte subpopulation exists at all times ($CTL^- > 0$). Finally, to support secondary immune response,[3] there is always some non-zero subpopulation of activated cytotoxic lymphocytes ($CTL^+ > 0$). These shape space subpopulation invariants are summarised in Table 1.

The phenomenon whereby a preferred subset of the general *CTL* population is stimulated to proliferate into armed effectors is known as *immunodominance* (Yewdell and Bennink, 1999). A stronger *immunogenic* viral peptide stimulates a stronger CTL response, which in turn down-regulates the infected antigen-presenting population level. In our model, immunodominance arises as a consequence of competition between CTL responses for antigenic stimulation. In shape space, immunodominance is restricted to the CTL^+ subpopulation only. The concentration

[3]And, as a consequence, immunological memory.

and distribution within this subpopulation (discussed in Section 3) is of particular interest, as we see correlation between disease clearance rates and ρ. A CTL^+ distribution pattern characterised by low concentration but broad activation, proves to be significantly advantageous in obtaining healthy clearance rates of infected antigen-presenting cells. The dynamics of the lymphatic compartment (referred to here as *physical space*) are modelled by way of a square cellular automaton lattice (Wolfram, 2001) of finite length, observing periodic boundary conditions. Within physical space, cytotoxic lymphocytes constantly recirculate, sampling their adjacent neighbours in order to detect the presence of infected antigen-presenting cells. The state of the lattice over time is simply a sequence of random variables $\mathbf{x}^{(0)}$, $\mathbf{x}^{(1)}$ ..., $\mathbf{x}^{(t)}$ defined on a finite space χ. The sequence is a *Markov chain* (Liu, 2001) as the value of $\mathbf{x}^{(t+1)}$ is dependent on its history only through its recent past $\mathbf{x}^{(t)}$. We define an update algorithm U which is conditionally applied to each selected location in the lattice depending on the location occupant type–at each time step τ. Each location is sampled for update following a uniform distribution. For this reason, the update sampling step is called a *Monte Carlo* time step. Each location occupant can be one of three mutually exclusive types: (i) naive cytotoxic lymphocyte precursor, (ii) activated cytotoxic lymphocyte, and (iii) active infected antigen-presenting cell. Following shape space notation (Table 1), these entities will be referred to as ctl^-, ctl^+ and apc^+, respectively (for clarity, lattice sites are here denoted with one subscript). An instance of these entities at location i determines the subscript (for example, ctl_i^- is the cytotoxic lymphocyte precursor at position i on the lattice). Additionally, we define two further substates which have relevance in physical space: proliferation ($\star$) and *death* ($\bullet$). The set of state transitions for a lattice cell at position i is as follows:

(1) $ctl_i^- \rightarrow ctl_i^{+\star}$
(2) $ctl_i^{+\star} \rightarrow ctl_i^+$
(3) $ctl_i^+ \rightarrow ctl_i^{+\bullet}$
(4) $apc_i^+ \rightarrow apc_i^{+\bullet}$

At each Monte Carlo time step, U is repeatedly applied to the lattice such that the coverage degree Φ (the fraction of non-duplicate locations selected for update at each Monte Carlo time step) is in the range $0.99 \leq \Phi \leq 1$. This reduces the effect of the pseudo-random number generator as a source of errors from the simulation. Each simulation is terminated when $\tau = 300$, or 6.25 days of simulation time has elapsed. This ensures that the full duration of the primary immune response is captured. On the lattice, each immune cell c_i has two neighbourhoods: an *inner* and an *outer*, denoted $\mathscr{R}_i$ and $\mathscr{R}_o$, respectively, with $|\mathscr{R}_i| = 8$ and $|\mathscr{R}_o| = 16$ (with radius 1 and 2, respectively). As part of the update algorithm U, each c_i recirculates within the physical space, implemented as follows: first, $\mathscr{R}_i$ is examined in order to locate an unoccupied position into which the immune cell may move. If an empty cell is located within $\mathscr{R}_i$, c_i will move into it with probability $P(inner) = 0.9$. If no space is available within $\mathscr{R}_i$, $\mathscr{R}_o$ is searched for a free space. If a free space is located in $\mathscr{R}_o$, c_i will move into it with probability $P(outer) = 0.7$. If both $\mathscr{R}_i$ and $\mathscr{R}_o$ are occupied,

then no movement of c_i will occur in this timestep. If a free location is found in $\mathscr{R}_i$ or $\mathscr{R}_o$, the new coordinates of c_i are calculated and the cell is moved. The values chosen for $P(inner)$ and $P(outer)$ are subjective and reflect the concept of cell motion into *proximate* and *nearby* space, respectively. In the results presented here, both parameters are kept constant. U is summarised as Algorithm 1.

Algorithm 1 Update algorithm U, applied at each time step τ.

Repeat
  Randomly select a non-empty cell c_i
  if c_i *is a* ctl_i^- **then**
    for $w = 0$ to $|R_i|$ **do**
      if cell c_w is a apc_w^+ and ($|c_w - c_i| \leq \rho$ {in shape space}) **then**
        $apc_w^+ := apc_w^{+\bullet}$ {mark c_w as dead}
        $ctl_i^- := ctl_i^{+\star}$ {c_i enters the clonal expansion phase}
        break {out of the for-loop}
      end if
    end for
  else if c_i is a $ctl_i^{+\star}$ **then**
    for $w = 0$ to $|R_i|$ *and* no free space in R_i **do**
      if cell $c_w = \emptyset$ **then**
        $ctl_w^{+\star} := ctl_i^{+\star}$ {copy c_i into the free cell c_w (clonal expansion)}
        break {out of the for-loop}
      end if
    end for
    for $w = 0$ to $|R_o|$ and no free space in R_i **do**
      if cell $c_w = \emptyset$ **then**
        $ctl_w^{+\star} := ctl_i^{+\star}$ {copy c_i into the free cell c_w (clonal expansion)}
        break {out of the for-loop}
      end if
    end for
  end if
  if cell c_w in R_i or $R_o = \emptyset$ **then**
    $c_i \mapsto c_w$ {recirculate the cell c_i to c_w}
  end if
  if age $c_i = 300$ and c_i is a ctl_i^+ **then**
    $ctl_i^+ := ctl_i^{+\bullet}$ {activated cells switch off at $\tau = 300$ (apoptosis)}
  end if
until $\Phi = 0.99$ {repeat until coverage reaches 99%}
$\tau := \tau + 1$ {increment the clock}

The combination of physical and shape space as introduced above, provides a *hybrid* model. Each process acts within physical, shape or hybrid space and each is defined as follows:

(1) *Physical space*
   (a) Each ctl^- and apc^+ recirculate inside the physical space following the motion rules of U. The ctl^- are actively sensing the local environment for sign of infected apc^+.
   (b) Once a ctl^- and apc^+ have come into contact, the simulation transfers to shape space, in the sense that the following sequence of steps is initiated.
(2) *Shape space*
   (a) The shape space distance between the two entities is calculated
   (b) If recognition occurs, *clonal expansion* rate is calculated (S_r). If not, the ctl^- returns to its recirculation process (1a).
   (c) Immunodominance emerges as $CTL^- \rightarrow CTL^+$ recruitment starts. At this point, the following occurs in both physical and shape space.
(3) *Hybrid space*
   (a) The population concentration increases for each CTL phenotype stimulated, giving rise, after τ time steps to a concentration level $CTL_i^+ = e^{0.036\tau S_r}$.
   (b) Infected antigen-presenting cells are removed from the physical space system by recirculating activated effector CTL cells.
   (c) At $\tau > 300$, effector cells undergo programmed death (apoptosis) and the primary immune response comes to an end.

Results

In the results presented here, we analyse the dynamics of response when the immune system is challenged by single strain pathogens. Therefore, the shape space APC^+ distribution is confined to one single point in space (and is shown by the asterisk, $*$). A multiple strain model was previously presented (Burns and Ruskin, 2003b). At time $\tau = 0$, $apc^+ = 100$, representing infected antigen-presenting cells entering the physical space, and in turn, triggering primary response. In shape space, CTL^- are distributed across 10^3 locations ensuring that the coverage requirement of Eq. (1) is always met. At the beginning of each simulation, shape space is configured by uniformly distributing the CTL^- and APC^+ subpopulations. Fig. 1 shows a typical initial shape space at time $\tau = 0$ (for simplification, we show the CTL^- in all cases, as a smaller point compared to both the APC^+ and CTL^+). Two primary conditions determine how the initial shape space distribution will evolve: (i) apc^+ and ctl^- must come into contact in physical space (*mobility*) and, (ii) shape space distance $d \leq r$. The conditions controlling (i) are not altered during simulation executions here. With respect to (ii), although ρ may be assigned any value in the continuous range $0 \leq \rho \leq 1$, it is here restricted to six representative values, in the set:

$$\mathscr{R} = \{0.928,\ 0.6,\ 0.39,\ 0.25,\ 0.164,\ 0.107\}. \tag{4}$$

Setting $\rho = 1$ would represent *maximal* recognition and binding behaviour such that *every* APC would bind to every CTL. This case would rarely arise, as such cross-reactive CTL would represent a threat to immune health and may give rise to

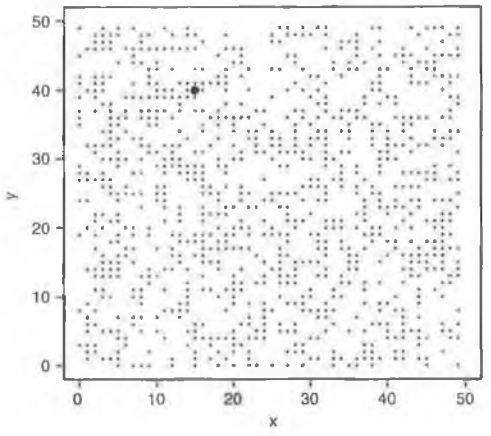


Fig. 1. Two-dimensional shape space showing the initial distribution of CTL^- and APC^+ cells (circle and asterisk, respectively) at time $\tau = 0$. At this time, one antigen-presenting cell is shown, and no recruitment from $CTL^- \rightarrow CTL^+$ has occurred. Biologically, this state represents the point at which an antigen has been taken up by an antigen-presenting cell (such as a dendritic cell), but has not yet been detected by the recirculating cytotoxic lymphocytes in the physical space.

autoimmune disease and is therefore excluded from the simulation. Likewise, $\rho = 0$ implies an immune system in which no cross-reactivity exists, and again the case is excluded from the simulation. We are interested in analysing the effect in shape space and physical space of declining cross-reactivity, from four perspectives:

(1) Clearance rate of the apc^+ subpopulation.
(2) Density levels of CTL^+ and ctl^+.
(3) Activation distribution pattern as the CTL^+ subpopulation is recruited from the CTL^- pool.
(4) Efficiency of response: the measure by which the apc^+ challenge is responded to *in proportion* to the threat posed, and that immune resources are not spent unnecessarily.

Each simulation ($\mathscr{R}_i$) is repeated 30 times, and the results are averaged. Shape space is confined to 2.5×10^3 positions while physical space is modelled as a two-dimensional array with 1×10^4 positions. These small sizes represent exploratory analysis with modest computational resources (single CPU) and should clearly be extended. A more numerically realistic size would require a physical space capable of modelling some 10^9 lymphocytes. Representing a high-order system in terms of a low-order one can result in finite size effects, with particular impact when the model exhibits first- or second- order phase transitions (Landau and Binder, 2000). In the results which follow, we believe the consequences of finite size effects are not immediately relevant. With biological practice, most in vitro experiments act on numerically reduced cell populations, and the results are taken merely as indications of possible in vivo outcome. Fig. 2 shows the activation distribution pattern in shape space for each value drawn from $\mathscr{R}$, at the end of the simulation ($\tau = 300$), with (a) to

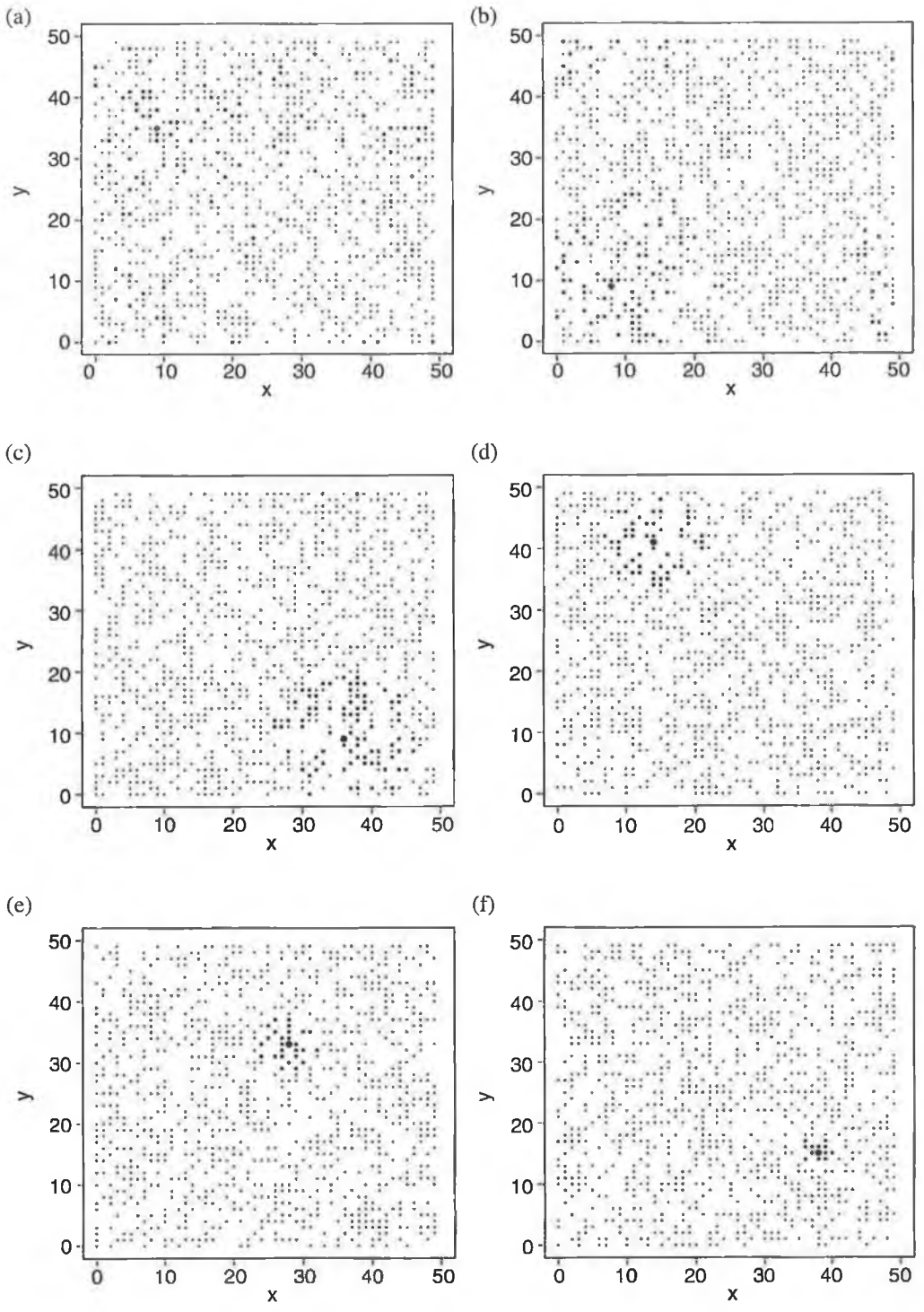


Fig. 2. Six shape space activation patterns at time $\tau = 300$, showing the effect of declining ρ with values drawn from $\mathscr{R}_{1,..,6}$. In (a) $\mathscr{R}_1 = 0.928$, $APC^+ = 0$ (indicating full eradication from the physical space), represented as a diamond ($\diamond$) to indicate the location it once occupied. In (b) $\mathscr{R}_2 = 0.6$, (c) $\mathscr{R}_3 = 0.39$, (d) $\mathscr{R}_2 = 0.25$, (e) $\mathscr{R}_2 = 0.164$, and (f) $\mathscr{R}_2 = 0.107$. In (b) – (f), $APC^+ > 0$, represented by an asterisk ($*$) indicating incomplete physical space eradication ($apc^+ > 0$).

(f) representing $\mathscr{R}_1$ to $\mathscr{R}_6$, respectively (as in physical space, periodic boundary conditions are enforced in shape space). As cross-reactivity declines, average distance of CTL^+ from the initial point of stimulation (APC^+) declines consistently and this is certainly in keeping with expectations. It is important to observe that the concentration of CTL^+ in (a) and (b) is almost identical (≈ 100). At first this may seem paradoxical – after all – if one assumes that a highly cross-reactive T-cell repertoire would become activated by every infected antigen-presenting cell it comes into contact with – this should result in greater T repertoire activation and, conversely, a reduction in cross-reactivity of some 35% must surely reduce the concentration of activated immune lymphocytes. The explanation becomes clear when the coverage constraint of Eq. (1) is considered: in the range $\mathscr{R}_{1,2}$ the highly mobile nature of the recirculating physical space cells ensures that *sufficient* APC^+ and CTL^+ come into contact, which in turn compensates for the declining cross-reactivity. This is intuitive: if the sampling rate of the immune lymphocytes of their environment is high enough then, theoretically, a ρ of *just greater than* 0 would be sufficient to sustain normal clearance rates. However, realistically, such a sampling rate would be beyond the capability of the immune lymphocytes. In the results shown in 2(a)–(e), the reduction in absolute value of ρ is not accompanied by a commensurate reduction in activation levels. This supports the assertion that recirculation rates are high enough to ensure significant levels of lymphocyte activation, and as such, there is non-linear relationship between ρ and activation concentration. As can be seen in 2(c)–(f), the narrowing cross-reactivity causes a well-defined clustering of activated lymphocytes in the neighbourhood of the APC^+. When the activation pattern for $\mathscr{R}_6$ is examined (Fig. 2(f)), it is clear that the concentration of CTL^+ has dropped significantly – to ≈ 12 (a decline of 88%). At this point, the recirculation rate is not sufficiently high enough to compensate for declining cross-reactivity, and as a result, we observe a commensurate decline in activation levels. Next, the model is analysed from the hybrid view point, specifically by examining the concentration levels of ctl^+ in physical space by way of of shape space CTL^+ (phenotypes). These data are shown in Fig. 3(a), which provides insight into how the activation patterns observed in shape space are related in terms of concentration. Notice in 3(a) that there are two main CTL^+ patterns: (i) low level — representing $\mathscr{R}_{1,2,3}$, and (ii) high level responses — representing $\mathscr{R}_{4,5,6}$. Low cross-reactivity has resulted in a consistently lower overall CTL^+ subpopulation at the end of clonal expansion, and conversely, higher cross-reactivity has resulted in greater immune resources being generated. An immune response cannot *a priori* know the numerical strength of each pathogenic challenge. This leads to a risk and tradeoff in terms of resources allocated and time spent in clonal expansion:enough effector cells must be activated at the end of the process to effect clearance of apc^+, but a highly agonist challenge could result in over-production of effectors. Heuristically, at the end of clonal expansion, we should see a population of ctl^+ *in proportion* to the initial apc^+ threat. We thus consider the question of immune response in proportion to threat. Observable from Fig. 2 is that although $v_{1,2}$ resulted in the broadest[4] shape

[4]Again, measured as the mean distance of CTL^+ from the APC^+.

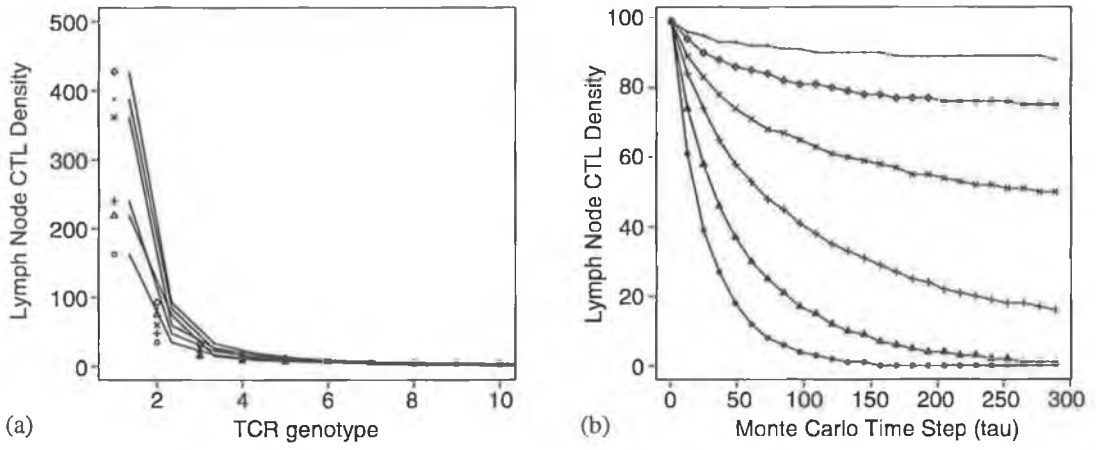


Fig. 3. (a) Concentration and distribution of ctl^+ across the 10 most active shape space phenotypes, and (b) the clearance rate for apc^+ removal from the physical space at the end of the primary response (some 6.25 days). In both figures, $\mathscr{R}_1 = \circ$, $\mathscr{R}_2 = \triangle$, $\mathscr{R}_3 = +$, $\mathscr{R}_4 = \times$, $\mathscr{R}_5 = \diamond$, $\mathscr{R}_6 = *$.

space CTL^+ activation distribution, physical space concentration levels are signifficantly more in proportion to the infected antigen challenge ($APC^+ = 10^2$) than the other parameter values of $\mathscr{R}$. In fact, the immunodominant concentration level of the most active CTL for $\mathscr{R}_1$ is 163 compared to 388 for $\mathscr{R}_5$ and 442 for $\mathscr{R}_6$. Thus, low cross-reactive configurations expend less immune resources when mounting a response than do the highly cross-reactive configurations. Consequently, the clonal expansion phase for $\mathscr{R}_{1,2}$ is shorter than for $\mathscr{R}_{4,5,6}$. In the case of a rapidly proliferating viral challenge (such as influenza) this shorter clonal expansion phase may save valuable time. Additionally, lower levels of ctl^+ reduces the risk of autoimmune disease or, for that matter, of allergic reaction. With respect to Fig. 3(b), an interesting picture emerges. This figure shows the clearance rate of infected apc^+ from the physical space over the duration of the primary response. As cross-reactivity declines, the rate at which apc^+ is cleared varies considerably: for both $\mathscr{R}_1$ and $\mathscr{R}_2$ a full clearance of apc^+ is achieved at the end of primary response – with $\mathscr{R}_1$ clearing more quickly. A further notable feature arises when comparing the clearance rates of $\mathscr{R}_1$ and $\mathscr{R}_2$: in shape space, (Fig. 2(a) and (b)), the activation levels are almost identical, significantly, the main difference is the pattern of clustering (a broad spread as opposed to a more focused one) which has emerged around the central stimulation point of the initial APC^+ infection. This implies that there is a qualitative difference between distribution patterns: a broad distribution with high mean distance from the centre results in greater efficiency in APC^+ clearance than does one with smaller mean distance from the centre. Further, we see that although high cross-reactivity results in large CTL^+ production (Fig. 3(a)), clearance rates still degrade from $\mathscr{R}_{4,5,6}$ leading to an immune response which, although marked by high lymphocyte production, is, in effect, wasting the lymphocyte clones produced, as almost no APC^+ are cleared (Fig. 3(a) – square, triangle and circle, respectively).

These wasted resources represent a double threat, in that (i) valuable response time is wasted in clonal expansion for a largely ineffective activated lymphocyte pool, and (ii) the excessive ctl^+ represents wasted resources in cell generation.

The appearance of immunodominance is evident at the end of primary response. At this time ($\tau = 300$), with the clonal expansion phase complete, the concentration and distribution of ctl^+ and CTL^+ (Fig. 3(a)) represents preferentially stimulated armed effectors. In $\mathscr{R}_4$ and $\mathscr{R}_6$, 90% of the immune response is concentrated against one APC^+ phenotype. A response of this nature suffers from a number of deficiencies: in the case of mutating pathogen (in effect, the *migration* of the phenotype from one shape space position to another), a high concentration response targeted against only one or a few MHC:peptide complexes could potentially fail to respond in sufficient time in order to eliminate the mutated challenge. Where the pathogen is also proliferating rapidly, such as in the case of HIV (Ramratnam et al., 1999), the time period during which the it remains undetected may be crucial in effecting the course of the infection. A high concentration response risks wasting immune resources a large fraction of which may be unnecessary. This reaction type may even result in tissue damage as a result of cell-mediated cytotoxicity or graft rejection. With $\mathscr{R}_1$ and $\mathscr{R}_2$ the response pattern is different (with findings in agreement with expected response patterns shown in Fig. 2): the response to single pathogen strain results in over 100 distinct CTL phenotypes (only the 10 most active are shown) being activated. Spreading the response across a broad range of CTL phenotypes ensures that physical space sampling rate is not a decisive factor in ensuring an effective response.

Discussion and Conclusions

In this paper we have presented a novel method to study the emergence and dynamics of diversity in a computational immune simulation. Our model integrates two formalisms, shape space and physical space, which are usually separately addressed. The combination allows us to show how localised dynamics effect global condition, and how global condition in turn feeds information back down to local physical space. We believe this model provides a basis for further experiment. In particular, our model permits investigation of the following properties: (i) shape space activation depends on cross-reactivity (ii) clearance rates are dependent on activation distribution (distance from shape space locus of the APC^+) (iii) high cross-reactivity is more effcient in clearance and effort spent (iv) increased recirculation rates can compensate to some extent for declining cross-reactivity. Notwithstanding the fact that shape space initially attracted considerable criticism from Carneiro and Stewart (1994) (a response to which may be found in Burns and Ruskin (2003a)), we suggest two experimentally testable hypotheses:

(1) Immunisation is likely to be more effective when a spread of memory effector cells are activated. This may require several stages of stimulation of the immune response with genetically varied strains of the same viral pathogen. Oxenius et al.

(2001) has identified mutations in peptide-MHC antigens that improve T-cell recognition without altering specificity. These mutant antigens are able to target a specific T-cell and deliver an enhanced activation signal. This, in turn, can lead to up to a 40-fold increase in effector function (for example, cytokine production). Such mutant antigens will be of use for boosting immune responses to specific T-cell antigens. It has also been shown that T-cells respond to lower concentrations of the mutant antigen.

(2) An increase in cross-reactivity can contribute to effective immune response. Brehm et al. (2002) have shown that it is possible to stimulate such cross-reactivity in vitro, and that the resulting T-cell activation pattern is beneficial in controlling heterologour[5] viruses.

The model presented here leads to the following conclusions: preferred types of activation patterns in shape space lead to more efficient and effective physical space clearance rates than do others. As the mean distance between CTL^+ and APC^+ declines (broadly *regardless of the mean activation concentration*), the clearance of antigenic from the physical space model becomes progressively less efficient. In agreement with the clinical findings of both Mason (1998) and van den Berg et al. (2001), the shape space of our model exhibits high cross-reactivity which results in an immunodominance configuration which enhances antigen clearance from the physical space.

References

Brehm, M., Pinto, A., Daniels, K., Schneck, J., Welsh, R., Selin, L., 2002. Nat. Immunology 3 (7), 627.
Burns, J., Ruskin, H.J., 2003a. In: Sloot, P., Gorbachev, Y. (Eds.), Computational Science – ICCS 2003. Lecture Notes in Computer Science, Vol. 2660. Springer, Berlin, pp. 75–85.
Burns, J., Ruskin, H.J., 2003b. In: Hamza, M. (Ed.), Biomedical Engineering. Proceedings of the IASTED International Conference, ACTA Press, pp. 60–65.
Buseyne, F., Riviere, Y., 1999. International Immunology 13, 941.
Carneiro, J., Stewart, J.J., 1994. Journal of Theoretical Biology 169, 391.
de Boer, R., Segel, L., Perelson, A., 1992. Journal of Theoretical Biology 155 (3), 295.
Germain, R.N., Stefanova, I., 1999. Annual Review of Immunology 17, 467.
Landau, D.P., Binder, K., 2000. A Guide to Monte Carlo Simulations in Statistical Physics. Cambridge University Press, Cambridge.
Liu, J.S., 2001. Monte Carlo Strategies in Scientific Computing. Springer, Berlin.
Mason, D., 1998. Immunology Today 19 (9), 395.
Oxenius, A., Price, D.A., Dawson, S.J., Tun, T., Easterbrook, P.J., Phillips, R.E., Sewell, A.K., 2001. AIDS 15, 121.
Perelson, A.S., Oster, G.F., 1979. Journal of Theoretical Biology 81 (4), 645.
Ramratnam, B., Bonhoeffer, S., Binley, J., Hurley, A., Zhang, L., Mittler, J., Markowitz, M., Moore, J., Perelson, A., Ho, D.D., 1999. Lancet 354 (9192), 1782.
Smith, D., Forrest, S., Hightower, R., Perelson, A., 1997. Journal of Theoretical Biology 189, 141.
van den Berg, H., Rand, D., Burroughs, N.J., 2001. Journal of Theoretical Biology 209, 465.
Wolfram, S., 2001. A New Kind of Science. Wolfram Media.
Yewdell, J.W., Bennink, J.R., 1999. Annual Review of Immunology 17, 51.

[5]Immunologically related but not identical.

A Model of Immune Suppression and Repertoire Evolution

John Burns and Heather J. Ruskin

Dublin City University, Dublin 9,
Ireland
{jburns, hruskin}@computing.dcu.ie
http://www.dcu.ie/computing/msc/index.html

Abstract. We present a stochastic cellular automata model that allows us to study both localized and generalized aspects of the immune system (IS). We show how critical values for T Cell Receptor (TCR) affinity and cross-reactivity (ρ) can determine the course of a viral infection. The model presented here offers insight into the widely varying pathology of infectious agents across individuals. Additionally, our model points to ways in which auto-immune disease can occur. We show that by integrating models of physical space and shape space we can analyze immune repertoire evolution and distribution over various time periods ranging from a few days up to three years.

1 Introduction

One of the questions that has long exercised immunological researchers is why different people (or, more precisely, their immune systems) respond differently to the same viral or antigen challenge. This question is exemplified by the pathology of the Human Immune Virus (HIV). In particular, the progression of the HIV infection from inital exposure to the onset of full blown AIDS (Acquired Immune Deficiency Syndrome) is known to occur over time (t), with range $2 \leq t \leq 20$ years [1]. Computer models that successfully reproduce the behaviour of the HIV pathogen are extremely good at replicating the localized intra- and inter- cellular behaviour over a certain period of time and space. In particular, the affect of viral mutation on the immune response has been demonstrated by [2] and [10] while the work of [4] and [5] have shown, at least in outline, how bit-string models can represent the state of the immune repertoire. However, to our knowledge, there are no computer models that successfully address the issue of why it is that different immune systems respond differently to the same viral exposure. It follows from this that in addition to modelling intra-cellular localized interaction, there is a challenge to find a way to express the generalized distinction that exists between the immune systems of different people. In this paper we postulate that the course of an infection is crucially dependent on the density and distribution of the immune repertoire, and furthermore, that T cell receptor cross-reactivity is a crucial factor in determining the success or otherwise of the immune response to pathogen challenge.

P.M.A. Sloot et al. (Eds.): ICCS 2003, LNCS 2657, pp. 75–84, 2003.

2 Model

The synthesis of two distinct approaches is the key to our model. Firstly, we model the physical space of the secondary immune organs (such as the lymph nodes or spleen) by way of two-dimensional stochastic cellular automaton [6] with periodic boundary conditions. This physical space model is the location in which host invading intra-cellular pathogens, having been engulfed by antigen presentation (APC) cells, are presented on the cell surface to naive recirculating cytotoxic T lymphocyte (CTL) cells. Upon recognition of a specific antigen, naive lymphocytes stop recirculating and undergo cell division. The process of cell division is limited to clones of the original lymphocyte that recognized the viral pathogen (this original lymphocyte is known as a *CTL-precursor*), and all cloned daughters inherit the specific T cell receptor that was successful in this recognition process. The daughters of the CTL-precursor cells will normally divide at a rate of 2-4 every 24 hours and this process continues for some 3-5 days [9]. This process is known as *clonal expansion*, and forms the basis of our physical space model. In our model, we assume that all antigen presentation cells carry the major histocompatibility complex (MHC) class I molecule which is required by CTL-precursors in order to produce armed effector CTL cells at the end of the clonal expansion process

The second component to our model is an implementation and extension of the shape space formalism originally presented by [7]. We utilize shape space to model the density and distribution of the T Cell Receptor (TCR) repertoire, and to lend analytical insight into the critical nature of the measure of cross-reactivity, which we denote as ρ. In our model, shape space is a dynamic and evolving N-dimensional Euclidean space that contains one vector $\mathbf{u}$ for every immune system CTL-p genotype, and one vector $\mathbf{v}$ for every pathogen genotype. Around each $\mathbf{v}$ in shape space is a disc of influence of radius ρ. Any $\mathbf{u}$ falling inside this disc of influence is subject to some pressure. That is to say, the pathogen will be removed from the physical space system with some probability $P(X)$, inversely proportional to the distance between $\mathbf{v}$ and $\mathbf{u}$ in shape space. If the distance d between the two points exceeds the critcal value ρ, then there is no CTL pressure on the pathogen, and no affinity or binding takes place. The process is summarized in equation (1) and equation (2)

$$d = \|\mathbf{v} - \mathbf{u}\| = \sqrt{\sum_{i=1}^{N} (v_i - u_i)^2} \tag{1}$$

$$P(X) = \begin{cases} 1 & : \quad d = 0 \\ \frac{1}{d} & : \quad 0 < d \leq \rho \\ 0 & : \quad d > \rho \end{cases} \tag{2}$$

It should be emphasized at this point that equation (1) and equation (2) will be evaluated if and only if there is some contact between an APC and a CTL-p within the physical space model. Contact in this case means that within the physical space lattice the cells representing the CTL-p and the APC are adjacent

- or are nearest neighbours. If equation (2) does result in the removal of the APC from the physical space, we use $P(X)$ as the *stimulation* rate (S) for the exponential growth during the clonal expansion phase ($S \equiv P$). Therefore, the total number of activated CTL daughter clones which the ith CTL-p gives rise to, (T^i_{CTL}) is given by:

$$T^i_{CTL} = e^{0.036\tau S} \tag{3}$$

where the power of e depends on three parameters, τ, the duration of the expansion phase, S, the stimulation rate (dependent on the distance between the APC and CTL-p in shape space), and a constant factor 0.036. At the end of the clonal expansion phase, when $\tau \approx 192$ (representing some four days of actual time) and where $S = 1$, we see some 1000 clones have been produced. This is what we would expect in a healthy immune system [9]. Clearly, the effect on T^i_{CTL} when S decline is noticeable, for example, for a case with weak stimulation, $S = 0.3$, $T^i_{CTL} \approx 8$.

Our physical space model is implemented as a discrete two-dimensional array of C-language pointers to data structures. Each member of the array contains a pointer to it's own structure (even if the structure is logically empty). In turn, each structure contains a set of information which is summarized in Table 1. At each time step, we randomly select locations on the array to update. We ensure full coverage of the array by using a member of the structure to indicate that this array element has been updated. When 99% of the array has been visited, we reset the visited flag on each array member, and increment the clock counter by one, thus indicating we have moved on to time $\tau+1$. Our shape space model is implemented as a one-dimensional array (of length 2). As both APC and CTL-p cells have a representation in shape space, this value is again carried by each of the structures in the physical space array. In Table 1, the location of a given APC or CTL-p in shape space is referred to as *sscoord*. We implement motion on the lattice merely by swapping pointer references (thus, no expensive in-memory copy activities are required). For example, assume an infected APC cell is located at position L_i, and at position L_{i+1} there is an unoccupied cell. We simply swap the pointers stored in L_i and L_{i+1} with each other, and the move is complete. For further treatment of physical space models see [3]. For other models of shape space, see [12].

2.1 Model Parameters

Although the model presented here is capable of supporting most of the known entities of the immune system, we restrict our inital study of the immune state to antigen presentation cells (APC), cytotoxic lymphocyte precursor (CTL-p) cells, and activated cytotoxic lymphocyte (CTL) cells. We justify the exclusion of other entities (such as Th cells, B cells and cytokines) by the fact that it is through activation and differentiation of CTL-p cells that the immune repertoire is known to evolve when the immune system is faced with an intra-cellular pathogenic challenge, such as the Human Immune Virus (HIV). Within our physical space model, we define L_{ps} as the length of one side of the square lattice, D_{ctl-p} as

Table 1. Data structure maintained at each physical space array member

Data type	Parameter	Definition
unsigned char	act	Cell type
unsigned char	state	Cell state
unsigned char	sscoord[]	Location of cell genotype in shape space
unsigned int	ticks	The age of this cell
unsigned int	max-ticks	The maximum age of this cell
double	prolif-rate	The rate at which this cell proliferates (0..1)
unsigned int	last-updated	Timestep when cell was last visited
unsigned int	updated-count	cell visited flag
unsigned int	id	Unique ID for this cell

the density of the CTL-p cells, D_{apc} as the density of the APC's. During each of the model simulation runs, we keep all parameters other than ρ constant. The largest value of ρ is known as ρ_{max}, the derivation of which is shown in equation (4).

$$\rho_{max} = \sqrt{\frac{L_{ps}^2}{\pi}} \tag{4}$$

We define ρ_{max} as a measure of the *maximum* cross-reactivity of a given genotype in shape space. Setting $\rho = \rho_{max}$ would be equivalent to having *every* CTL-p cross-react with every APC. Although we do not explore this configuration further, it is worth noting that this configuration could represent a clinical condition known as auto-immune disease (wherein the immune system attacks both itself and invading pathogens without discretion). Auto-immune disease is a relatively rare condition ([13]) and will be examined in further work. As we are initially interested in observing the behaviour of the model as $\rho \to \rho_{max}$, we express the ratio of ρ to ρ_{max} as $\hat{\rho}$, which takes values in the range $\mathcal{A} = \{0.5, 0.29, 0.10\}$. $\mathcal{A}$ represents one possible set of values for $\hat{\rho}$ in order of decreasing cross-reactivity. We could have chosen any set of values for $\hat{\rho}$ that follow $0 \leq \hat{\rho} \leq 1$. From equation (4), the value for ρ_{max} in our model is ≈ 28. The model parameter initial values are summarized in Table 2.

3 Shape Space

The shape space idea, as an underpinning of theoretical immunology, was introduced by [7] in 1979, but has attracted some criticism, notably from [16]. It is instructional at this point to review the basis of this important formalism and to address the comments of [16] . We will also justify our utilization of shape space while highlighting potential shortcomings.

In shape space, the antigenic determinant for a given antibody or antigen is an N-dimensional vector consisting of values representing such parameters as geometric configuration (at the molecular level), electric charge and other

Table 2. Table of initial model parameters

Parameter	Definition	Initial Value
D_{ctl-p}	Density of CTL-precursors	2000
θ	Number of genotype CTL in shape space	2000
D_{apc}	Density of APC	100
L_{ps}	Length of one side of physical space square lattice	100
L_{ss}	Length of one side of shape space square lattice	50
τ	Number of time steps per simulation run	300
ρ_{max}	Max. cross-reactivity of shape space genotype	≈ 28
$\mathcal{A}$	Range of $\hat{\rho}$	$\{0.5, 0.29, 0.1\}$

complex chemical characteristics that are not postulated in [7] . According to this approach, not all elements of the N-dimensional vector play an equal role in determining complementarity or match. Given an antibody vector **Ab** and an antigen vector **Ag**, if **Ab** = **Ag** then the two entities will bind absolutely. If we now assume that the N-dimensional vector represents a location in an N-dimensional Euclidean space, then

$$\|\mathbf{Ab} - \mathbf{Ag}\| = 0 \tag{5}$$

means that **Ab** and **Ag** are *coincident* in the space (that is, they share the exact same coordinates within the space), and this space is known as *shape space.* Therefore, we can say that shape space is the space of all possible vectors each one representing a unique set of antigenic determinants (or shapes). One of the criticisms that [16] level at the above theory is that the function $f(\mathbf{Ab}, \mathbf{Ag})$, which, according to [7], will determine the 'distance' (and hence, affinity) between **Ab** and **Ag**, must be highly irregular and discontinuous. This assertion is based on work by chemists which has shown that predicting affinity and bonding between two molecules is not simply a deterministic issue of understanding the dynamics between the individual molecular constituents. However, [7] clearly comment towards the end of their paper that shape space does not *need* to be uniform (that is, there does not have to be a uniform distribution of **Ab** vectors within Euclidean space). Therefore, f does not need to be either continuous or regular. In fact, [7] offer a model for representing the probability of detecting antigens in a non-uniformly distributed shape space.

Further criticism of the shape space paradigm is the question of the value for N. Although [16] insist the original value of $5 \leq N \leq 10$ is too small (they suggest a value closer to $N = 20$), they do not, in principle, question the theoretical foundation of representing antigenic determinants by a fixed, N-sized set of parameters. The actual value for N is clearly something that is system specific and may vary. In fact, [17] have shown how the presence of cytokine regulatory molecules crucially affects the dynamics of helper T cell populations. It is therefore plausible that the actual value and parameters of N are not only dependent on the characteristics of **Ab** and **Ag**, but also of external and local-

ized state information such as the density of cytokines. It seems reasonable to conclude that setting N to a fixed and relatively small number is sufficient to represent the parameters of antigenic determination.

Finally, to conclude this section, we address the specific utilization of the shape space paradigm within the model presented in this paper. Firstly, we do not require or assume a continuous **Ab** distribution across our shape space. The probability that an APC will fall into a region for which equation (1) is not defined (and hence, not be under any CTL pressure) is given by

$$P(\bar{X}) = 1 - \frac{\sum_{i=1}^{\theta} \pi\rho^2}{L_{ss}^2} \tag{6}$$

In our model we assume that shape space is populated by a uniform distribution of genotype CTL cells. It can be seen from equation (6) that ρ is a critical parameter. Additionally, in a healthy immune system, equation (6) would always result in a value close to zero. Secondly, we extend the original shape space formalism by introducing a stimulation rate which is inversely proportional to the distance as calculated using equation (1). This partially addresses the deficiencies in other shape space -based models, most notably that of [12]. Thirdly, with respect to the value for N, although we set $N = 2$, this is by no means a hard parameter of our model. As N represents the dimensionality of the shape space, we should, in principal, set N to a value such that $L_{ss}^N \to \infty$ (as the number of antigenic determinants that might be presented to the immune system is, essentially, infinite). For practical computational purposes, such a space would require resources beyond current computing boundaries available today. As the research presented here studies the affect of *specific* pathogen detection and removal, our space needs only to accomodate antigenic variation for a specific strain, and thus $N = 2, L_{ss}^N = 2.5 \times 10^3$ is sufficient for our initial purposes.

4 Results

Fig. 1 shows the state of three immune TCR shape spaces at time $\tau = 300$ (where $\tau = 1$ models 30 minutes of elapsed time), and Fig. 2 shows the denstity of APC in the lymph system over time $0 \leq \tau \leq 300$ for each of the three immune system configurations. The only parameter altered across each of the three configurations is the ratio of $\hat{\rho}$, the values of which are drawn from $\mathcal{A}$ (and are represented in the figures by the diamond, the circle and the square, respectively). We notice, in Fig. 1, that the shape space has evolved into three quite different states. The first state can be classified as *healthy* in the sense that the clearance rate of APC from the lymph system is broadly in keeping with what is known to be the case clinically [9] (and represented here as the diamond in Fig. 2). We can see that the healthy shape space state can be characterized as a low affinity, low density configuration. This is in agreement with [8]. As the immune cross-reactivity declines (represented by the circle, Fig. 1 and 2) we see an apparently paradoxical condition in the shape space realization. That is, as

promiscuity diminishes, it takes greater specificity (and hence, greater stimulation) in order to clear APC. In fact, we can see from Fig. 2 that a 90% clearance from the lymph system takes some $\tau = 300$ time steps. We conclude that the immune system configuration, with reduced cross-reactivity, is chronically unable to clear all APC and is operating at a sub-optimal level. Finally, when we test the immune configuration at the lowest level of cross-reactivity (represented by the square) - we see a very strong and highly specific four receptor response in shape space. Having reduced promiscuity to its lowest level, the effect on APC clearance is conspicuous. From Fig. 2 we can see only a 10% APC removal from the lymph system after some 6 days. It is clear that were this the immune system of a real person, it would be at the point of collapse, and hence could represent the onset of full-blown AIDS (wherein the patient dies by way of opportunistic infection).

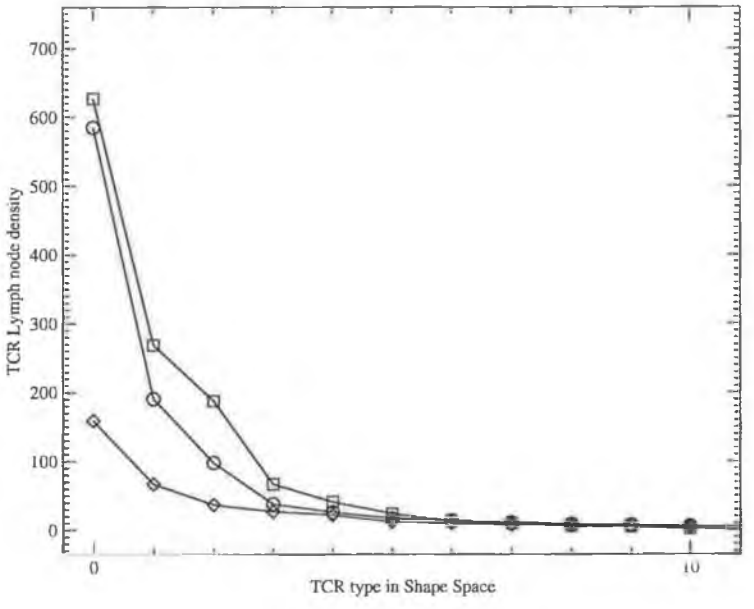


Fig. 1. Immune repertoire density and distribution for 3 values of $\hat{\rho}$ drawn from $\mathcal{A}$, with $\tau = 192$. Shown here are the density levels for the ten most dominant TCR types in shape space. The healthy state of the system (diamond) is a low-affinity, low average density configuration. As affinity increases and promiscuity declines, a more specific and active repertoire evolves. The state represented by the square is the least healthy, and represents an immune system at the point of failure.

5 Discussion

Our starting point for the experiments presented here is to assume that the immune repertoire has not been exposed to any prior pathogen. Hence, we do not model immunological memory or reinfection. Our objective has been to study how insight (into different IS evolutionary states) can be gained by integrating localized and generalized models of immune response. Informally, our theoretical

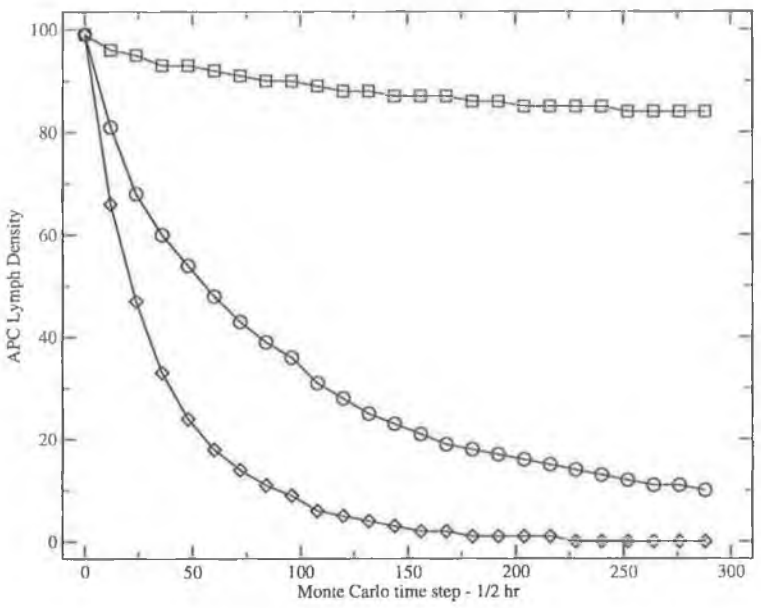


Fig. 2. Clearance rate of APC's from the lymph node system over about 6 days ($\tau = 300$) for various values of $\hat{\rho}$ drawn from $\mathcal{A}$. The three distinct states which have resulted: healthy, chronic and fatal (diamond, circle and square, respectively). These states show the affect of decreasing $\hat{\rho}$ on APC clearance.

expectation is that by studying the evolution of shape space, we can demonstrate how specific configurations of the repertoire occur. Following from this, we can also show one possible mechanism for the development of immunodominance [1]. With reference to Fig. 1, we see that immunodominance becomes more pronounced as cross-reactivity declines. An immune system with high levels of immunodominance gives rise to an inefficient detection and clearance regime. Therefore, when we examine the difference in infection outcome (Fig. 2) we notice that a more strain-specific response results in poorer short-term disease detection and clearance. Although we have not modelled pathogenic mutation in this model, we can hypothesize that the repertoire configurations that have evolved might play a role in determining the dynamics of a virus mutation and immune clearance. We feel the model presented here is a sound basis for exploring such issues in the future.

6 Conclusion

A model of the immune system is presented which utilizes a new extension to the shape space formalism, and addresses - at least in part - some of the former criticisms. We argue for the value of a shape space formalism for modelling some aspects of the immune repertoire. We have presented initial results from the model that suggest that a low-affinity T cell receptor (TCR) space provides the

[1] Immunodominance is the process by which a small number of specific TCR's are reponsible for clearing an antigen or virus

most efficient APC removal. Our findings are in agreement with recent work presented by [8]. We have shown that, for all other model parameters being unchanged, reducing TCR promiscuity causes significantly different evolution of the shape space and correspondingly poorer APC removal. We hypothesize that ρ (the cross-reactivity measure) is one of the most crucial general parameters within the immune system. We have demonstrated, for one configuration of the model, results which are similar in signature to those which characterize the onset of AIDS. A major factor underlying the evolution of an individual's immune system (and its overall health) appears to be the role played by cross-reactivity, ρ. For $\hat{\rho} < \hat{\rho}_{crit}$ (where $\hat{\rho}_{crit}$ is crudely estimated to be in the range $0.29 \leq \hat{\rho}_{crit} \leq 0.1$), the immune system is essentially undermined and faces almost immediate collapse. For more viable values of $\hat{\rho}$ ($0.5 < \hat{\rho}_{crit} < 0.29$) the residual antigen cells are maintained at a limited levels in every case, with less than 100% effective clearance.

References

1. Nowak M., McMichael A. J.: How HIV defeats the Immune System. Scientific American 273 (1995) 58–65
2. Mannion, R., Ruskin, H., Pandey, R.B.: Effect of Mutation on Helper T-cells and Viral Population: A Computer Simulation Model for HIV Theor. in Biosci. 119/2 (2000) 145–155
3. Seiden, P., Celada, F.: A model for simulating cognate recognition and response in the immune system, J.Theor. Biol. 158 (1992) 329–357
4. Lagreca, M.C., de Almeida, Zorzenon dos Santos, R.M.C.: A dynamical model for the immune repertoire. Physica A 289 (2001) 191–207
5. Castiglione, F., Motta, S., Nicosia G.: Pattern recognition by primary and secondary response of an Artificial Immune System. Theory Biosci. 120 (2001) 93–106
6. Wolfram, S.: Cellular Automata as Simple Self-Organizing Systems. Nature (1982)
7. Perelson, A. Oster, G.: Theoretical Studies of Clonal Selection: Minimal Antibody Repertoire Size and Reliability of Self-Non-self Discrimination, J. Theor. Biol. 81 (1979) 645–667
8. van den Berg, H.A., Rand D.A., Burroughs, N.J.: A Reliable and Safe T Cell Repertoire based on Low-affinity T Cell Receptors. J. Theor. Biol 209 (2001) 465–486
9. Janeway, C.A., Travers, P., Walport, M., Capra, J.D.: Immunobiology. The Immune System in Health and Disease. Churchill-Livingston 4th Edition (1999)
10. Zorzenon dos Santos, R.M., Coutinho, S.C.: Dynamics of HIV Infection: A Cellular Automata Approach. Phys. Rev. Lett. 87 (2001) 168102
11. Nowak, M., May, R.: Mathematical Biology of HIV Infections: Antigenic Variation and Diversity Threshold. Mathematical Biosci. 106 (1991) 1–21
12. Hershberg, U., Louzoun, Y., Atlan, H., Solomon, S.: HIV time: winning the war while, loosing all the battles. Physica A 289 (2001) 178–190
13. Orosz, C.: An Introduction to Immuno-ecology and Immuno-informatics. Design Principles for Immune System and Other Distributed Autonomous Systems. Oxford University Press (2001)
14. Monteiro, L.H.A., Goncalves, C.H.O., Piqueira, J.R.C.: A condition for Successful Escape of a Mutant after Primary HIV Infection. J. Theor. Biol. 203 (2000) 399–406

15. Nowak, M., May, R.: Virus Dynamics. Mathematical Principles of Immunology and Virology. Oxford University Press (2000)
16. Carneiro, J. and Stewart, J.J.: Rethinking "Shape Space": Evidence from Simulated Docking Suggests that Steric Complementarity is not limiting for Antibody-Antigen Recognition and Idiotypic Interactions. J. Theor. Biol. 169 (1994) 391–402
17. Yates, A, Bergmann, C., Van Hemmen, J.L., Stark, J., Callard, R.: Cytokine-modulated Regualtion of Helper T Cell Populations J. Theor. Biol. 206 (2000) 539–560

Viral Strain Diversity and Immune Response - a Computational Model

John Burns
Dublin City University
Dublin 9
Ireland
email: jburns@computing.dcu.ie

Heather J. Ruskin
Dublin City University
Dublin 9
Ireland
email: hruskin@computing.dcu.ie

Abstract

We present a hybrid stochastic cellular automaton computational model that allows us to study both localized and generalized aspects of the immune system. We introduce a new approach to model the dynamics of viral strain diversity, detection and clearance. We see that as strain diversity decreases, the effectiveness of infected cell removal from the lymph system declines. We build on previous work to show how critical values for T Cell Receptor affinity and cross-reactivity can also affect the course of a viral infection. The model presented here may offer some insight into the widely varying pathology of infectious agents across individuals.

KEY WORDS
Viral Strain Diversity, Cross-reactivity, Cellular Automata, Shape Space

1 Introduction

Many pathogens encountered by the immune system exhibit some form of strain diversity. That is to say, once the pathogen has been taken up by an antigen presentation cell (APC), the MHC class II/peptide complex presented on the surface of the the APC may exhibit some form of structural variation across the genotype. These variations may emerge over a long time period, typically several years in the case of the influenza virus, contrasted to just a few days in the case of the Human Immune Virus (HIV) [1]. Strain diversity within the influenza virus genotype is well understood, and world-wide strain variations are monitored closely. However, there has been little attention paid to modelling the dynamics between viral strain diversity and T-Cell Receptor (TCR) cross-reactivity. Previously, we presented work on the dynamics of APC detection and clearance under various TCR cross-reactive regimes ([2]). In this paper, we now extend our previous work to show how TCR cross-reactivity and viral strain diversity can be modelled by way of an extension to the shape space formalism first introduced by [3]. Our model supports the findings of both [4] and [5] as well as providing further evidence against the theory of lock-and-key with respect to viral and antigen detection. Our findings identify three critical ranges of viral strain diversity, representing healthy, chronic infected and fatal immune system states respectively.

2 Model

The synthesis of two distinct approaches is the key to our model. Firstly, we model the physical space of the secondary immune organs (such as the lymph nodes or spleen) by way of two-dimensional stochastic cellular automaton [10] with periodic boundary conditions. This physical space model is the location in which host invading intracellular pathogens, which, having been engulfed by antigen presentation (APC) cells, are presented on the cell surface to naive recirculating cytotoxic T lymphocyte (CTL) cells. Once the infected APC is detected, the process of clonal expansion [2], [11] ensues.

The second component to our model is an implementation and extension of the shape space formalism originally presented by [3]. We utilize shape space to model the density and distribution of both TCR and APC genotypes. Shape space offers analytical insight into the critical nature of the measure of cross-reactivity, which we denote as ρ. In our model, shape space is a dynamic and evolving N-dimensional Euclidean space that contains one vector $\mathbf{u}$ for every immune system CTL-p genotype, and one vector $\mathbf{v}$ for every pathogen genotype. Around each $\mathbf{v}$ in shape space is a disc of influence of radius ρ. Any $\mathbf{u}$ falling inside this disc of influence is subject to some pressure. That is to say, the pathogen will be removed from the physical space system with some probability $P(X)$, inversely proportional to the distance between $\mathbf{v}$ and $\mathbf{u}$ in shape space. If the distance d between the two points exceeds the critical value ρ, then there is no CTL pressure on the pathogen, and no affinity or binding takes place. The process is summarized in equation (1) and equation (2)

$$d = \|\mathbf{v} - \mathbf{u}\| = \sqrt{\sum_{i=1}^{N} (v_i - u_i)^2} \qquad (1)$$

$$P(X) = \begin{cases} 1 & : \quad d = 0 \\ \frac{1}{d} & : \quad 0 > d \leq \rho \\ 0 & : \quad d > \rho \end{cases} \tag{2}$$

It should be emphasized at this point that equation (1) and equation (2) will be evaluated if and only if there is some contact between an APC and a CTL-p within the physical space model. Contact in this case means that within the physical space lattice the cells representing the CTL-p and the APC are adjacent - or are nearest neighbours. If equation (2) does result in the removal of the APC from the physical space, we use $P(X)$ as the *stimulation* rate (S) for the exponential growth during the clonal expansion phase ($S \equiv P$). Therefore, the total number of activated CTL daughter clones which the ith CTL-p gives rise to, (T^i_{CTL}) is given by:

$$T^i_{CTL} = e^{0.036\tau S} \tag{3}$$

where the power of e depends on three parameters, τ, the duration of the expansion phase, S, the stimulation rate (dependent on the distance between the APC and CTL-p in shape space), and a constant factor 0.036. At the end of the clonal expansion phase, when $\tau \approx 192$ (representing some four days of actual time) and where $S = 1$, we see some 1000 clones have been produced. This is what we would expect in a healthy immune system [11]. Clearly, the effect on T^i_{CTL} when S declines is noticeable: for example, for a case with weak stimulation, $S = 0.3$, $T^i_{CTL} \approx 8$.

Our physical space model is implemented as a discrete two-dimensional array of C-language pointers to data structures. Each member of the array contains a pointer to it's own structure (even if the structure is logically empty). In turn, each structure contains a set of information which is summarized in Table 1. At each time step, we randomly select locations on the array to update. We ensure full coverage of the array by using a member of the structure to indicate that this array element has been updated. When 99% of the array has been visited, we reset the visited flag on each array member, and increment the clock counter by one, thus indicating we have moved on to time $\tau + 1$. Our shape space model is implemented as a one-dimensional array (of length 2). As both APC and CTL-p cells have a representation in shape space, this value is again carried by each of the structures in the physical space array. In Table 1, the location of a given APC or CTL-p in shape space is referred to as *sscoord*. We implement motion on the lattice merely by swapping pointer references (thus, no expensive in-memory copy activities are required). For example, assume an infected APC cell is located at position L_i, and at position L_{i+1} there is an unoccupied cell. We simply swap the pointers stored in L_i and L_{i+1} with each other, and the move is complete. For further treatment of physical space models see [2], [6], [7] and [8] For other models of shape space, see [2] and [14].

Table 1. Data structure maintained at each physical space array member

Data type	Parameter	Definition
unsigned char	act	Cell type
unsigned char	state	Cell state
unsigned char	sscoord[]	Location of cell genotype in shape space
unsigned int	ticks	The age of this cell
unsigned int	max-ticks	The maximum age of this cell
double	prolif-rate	The rate at which this cell proliferates
unsigned int	last-updated	Time step when cell was last visited
unsigned int	updated-count	cell visited flag
unsigned int	id	Unique ID for this cell

2.1 Model Parameters

Although the model presented here is capable of supporting most of the known entities of the immune system, we restrict our inital study of the immune state to antigen presentation cells (APC), cytotoxic lymphocyte precursor (CTL-p) cells, and activated cytotoxic lymphocyte (CTL) cells. We justify the exclusion of other entities (such as Th cells, B cells and cytokines) by the fact that it is through activation and differentiation of CTL-p cells that the immune repertoire is known to evolve when the immune system is faced with an intra-pathogen challenge, such as the Human Immune Virus (HIV). Within our physical space model, we define L_{ps} as the length of one side of the square lattice, D_{ctl-p} as the density of the CTL-p cells, D_{apc} as the density of the APC's. During each of the model simulation runs, we keep all parameters other than ρ constant. The largest value of ρ is known as ρ_{max}, the derivation of which is shown in equation (4).

$$\rho_{max} = \sqrt{\frac{L^2_{ps}}{\pi}} \tag{4}$$

We define ρ_{max} as measure of the *maximum* cross-reactivity of a given genotype in shape space. Setting $\rho = \rho_{max}$ would be equivalent to having *every* CTL-p cross-react with every APC. Although we do not explore this configuration further, it is worth noting that this configuration could represent one of two possible clinical conditions -

1. Auto-immune disease - wherein the immune system attacks both itself and invading pathogens without discretion. Auto-immune disease is a relatively rare condition ([15]) and will be examined in further work.

2. Negative selection - a condition under which any naive T-cells which express affinity to self-cells are deleted

Table 2. Table of initial model parameters

Parameter	Definition	Initial Value
D_{ctl-p}	Density of CTL-p	2000
θ	CTL in shape space	2000
D_{apc}	Density of APC	100
L_{ps}	Physical space length of one side	100
L_{ss}	Shape space length of one side	50
τ	Number of time steps per simulation	300
ρ_{max}	Maximum cross-reactivity	≈ 28
n	Sequence of cross-reactivity	2,3,...,21,22
$\mathcal{A}$	Range of $\hat{\rho}$	$\frac{n}{\rho_{max}}$
$\mathcal{M}$	Strain Diversity Density-Distribution	1,100;2,50; 4,25;10,10; 50,2;100,1

from the repertoire prior to entering the blood recirculation system

From equation (4), the value for ρ_{max} in our model is ≈ 28. Table 2 summarizes additional model parameters. As we are initially interested in observing the behaviour of the model as $\rho \rightarrow \rho_{max}$, we express the ratio of ρ to ρ_{max} as $\hat{\rho}$, which takes values from the set of $\mathcal{A} = \frac{n}{\rho max}, n = 2, 3, 4...21, 22$. $\mathcal{A}$ represents one possible set of values for $\hat{\rho}$ in order of increasing cross-reactivity. We could have chosen any set of values for $\hat{\rho}$ that follow $0 \leq \hat{\rho} \leq 1$.

Strain diversity is represented in our shape space as follows. For each of the separate six models, we place into shape space the following *per-model* (comma separated) density-distributions: $\mathcal{M}_1 = 1, 100$; $\mathcal{M}_2 = 2, 50$; $\mathcal{M}_3 = 4, 25$; $\mathcal{M}_4 = 10, 10$; $\mathcal{M}_5 = 50, 2$; $\mathcal{M}_6 = 100, 1$. In all models, the total number of infected APC's initially introduced into the system is the same (100), and decreases monotonically with time, (as we do not model viral expansion or reinfection), and it is the density-distribution relationship within shape space which is altered across models. In this way, we can exclude quantitative difference in APC density as a factor in the disease progression and outcome. We consider the area enclosed by shape space to be the set of all possible strain types for a given viral genotype.

3 Results

The results presented here show model output for two distinct parameter sets, these are: (i) variation in APC density and distribution across shape space for $\mathcal{M}_i \in \mathcal{M}$, and (ii) variation in the measure of cross-reactivity in shape space with $\hat{\rho} \in \mathcal{A}$. In both Fig. 1 and Fig. 2 we show the effect of decreasing the distribution of APC antigenic determinants in shape space. Studying this parameter yields insight into how the immune system responds to strain diversity (antigenic determinant diversity). For example, every year when the Influenza pathogen re-appears it periodically exhibits antigenic variation when compared to previous strain genotypes. It is this diversity which enables the pathogen to avoid immune detection with varying degrees of success. In the simulations conducted here, we examine the effectiveness of increasing cross-reactivity when presented with increasingly strain-specific challenges. Fig. 1 shows the probability density function (PDF) for the six values of $\mathcal{M}$ as shown in Table 2. What is immediately clear is that as the antigenic diversity narrows and becomes increasingly more specific, the CTL probability denisty function changes from unimodal to right-side bimodal, and, simultaneously, the mean shifts to the left. Effectively, we see that decreasing antigenic diversity causes lower average immune response along with increasing variance. The consequence of this response profile becomes clear when Fig. 2 is considered. Here, we see the APC density remaining in the lymph system at $\tau = 300$ (which represents ≈ 6.25 days of elapsed time). Plot (a), (b), (c) and (d), representing APC distribution in shape space of 100, 50, 25 and 10, respectively have broadly similar means (16.52, 16.86, 17.38 and 18.71) and broadly similar minima and maxima. Thus, for plots (a) to (d), it is clear that the APC clearance rates are only *somewhat* inhibited by narrowing antigenic diversity. However, when antigenic diversity declines to the value of two[1], as shown in plot (e), a clear point of instability is reached: the average amount of APC left in the lymph system jumps to 25.43, and in plot (f), to 32.14. It is interesting to see how the exponential curve of the first four plots has been replaced by an irregular distribution which cannot easily be classified. The biological explanation may now be as follows: As antigenic diversity narrows, there are fewer recombination sites on the APC to which the TCR can bind. When the number of recombination sites reaches one, we have the traditional lock-and-key method of pathogen recognition [11]. However, our model suggests that if the immune system followed a strict lock-and-key pathogen recognition process, an inefficient APC clearance regime would appear. In fact, the results of the simulation presented here suggest that any APC which provided only one region for TCR binding, would present the immune system with a serious challenge. We can see from the curve of plot (f) that even for high levels of cross-reactivity (such as $\hat{\rho} = 14$), only 75% of the APC is cleared. Contrast this with plot (a), where over 97% of APC are cleared, we can see that in order for the immune system to safely remove pathogen using the lock-and-key method, a cross-reactivity level of $\hat{\rho} \approx 21$ would be required. Problematically, such a high value for $\hat{\rho}$ would run the risk of inducing autoimmune disease, because the CTL-precursors would bind with such promiscuity that self-cells could be attacked. To further develop this point, we now divide the immune response into three categories: Healthy, Chronic Infected and Fatal - denoted as δ_h, δ_c and δ_m, respectively (where δ represents and APC clearance regime). Broadly speaking we assume that a healthy immune system is one which removes at least 95% of pathogen infected

[1] that is, there are only two distinct APC genotypes in shape space

Table 3. Table of model clearance rates (δ) for $\mathcal{M}_{1..6}$

Model $\mathcal{M}$	δ_h	δ_c	δ_m
$\mathcal{M}_1$	[12,21]	[3,11]	[1,2]
$\mathcal{M}_2$	[12,21]	[3,11]	[1,2]
$\mathcal{M}_3$	[13,21]	[3,12]	[1,2]
$\mathcal{M}_4$	[13,21]	[3,12]	[1,2]
$\mathcal{M}_5$	[15,21]	[5,14]	[1,4]
$\mathcal{M}_6$	[16,21]	[7,15]	[1,6]

APC. Likewise, a chronic infected response is one which removes at least 50% of infected APC cells. Finally, an immune response which cannot clear at least 50% of infected APC is in the process of collapse, resulting in the death of the host. Table 3 shows the range of values for δ_h, δ_c and δ_m, for each model $\mathcal{M}_i \in \mathcal{M}$, and relates directly to the plots shown in Fig. 2. Referring to Table 3, as antigenic diversity narrows, a critical point occurs in model $\mathcal{M}_5$ wherein the healthy region δ_h is squeezed into five (and subsequently, four) points of high $\bar{\rho}$. Thus, models $\mathcal{M}_5$ and $\mathcal{M}_6$ provide some evidence against the historically accepted lock-and-key nature of TCR - APC binding: under this approach, the immune system would need to maintain such high levels of $\bar{\rho}$ that some form of autoimmune disease would be likely to develop. Finally, Fig. 3 shows the mean CTL and APC density values plotted against each of the models. Clearly, we see that (apart from the third data point), as antigenic diversity narrows, overall immune activity (CTL density) decreases. Similarly, infected APC density will increase as CTL activity declines. Again, the first four data points, representing $\mathcal{M}_{1..4}$, show only slightly retarded activity levels. Once the antigenic diversity narrows to two (then one), CTL activity declines markedly. Again, APC density shows small increases for $\mathcal{M}_{1..4}$, but shows large increases for $\mathcal{M}_{5..6}$.

4 Discussion

Our starting point for the experiments presented here is to assume that the immune repertoire has not been exposed to any prior pathogen. Hence, we do not model immunological memory or reinfection. Our objective has been to study how antigenic diversity and cross-reactivity together play a role in determining the course of viral infection. Additionally, the results of our simulation cast some doubt on the notion of lock-and-key viral recognition. Based on the results above we can explore possibilities for how and why the course of infection differs among people. Also, we are in a position to understand some of the underlying mechanisms of autoimmune disease. Given an initial viral exposure there are now four potential disease progression paths:

(i) The disease has some antigenic diversity, the immune system exhibits normal cross-reactivity: full clearance

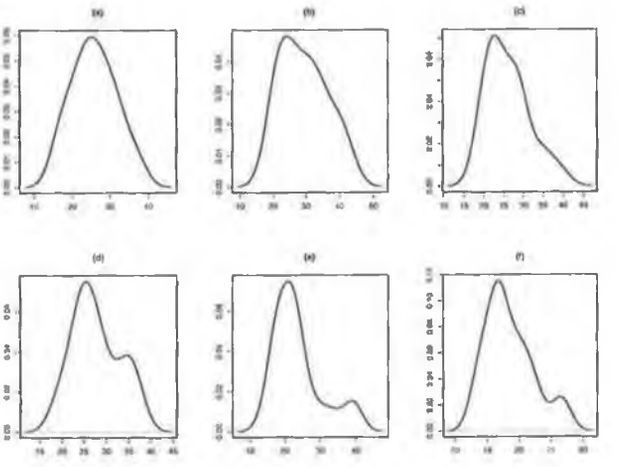


Figure 1. Shape space CTL probability denisty function (PDF) of $\mathcal{M}_{1..6}$, plot (a) to (f) respectively. Two distinct PDF's emerge: unimodal and bimodal. The unimodal distribution gives way to bimodal as $\mathcal{M}_1 \rightarrow \mathcal{M}_6$

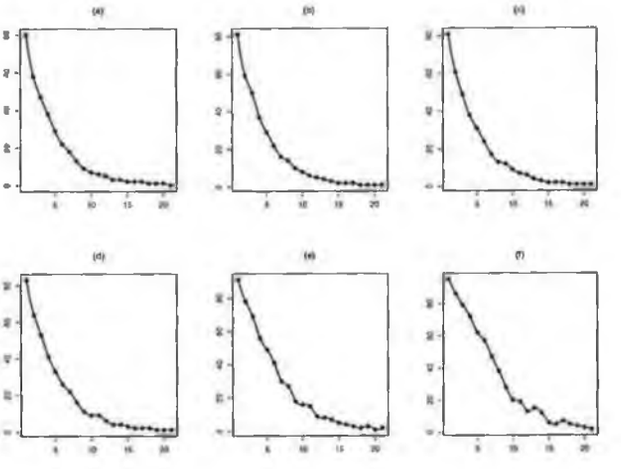


Figure 2. Lymph node APC density level ($y - axis$) for values of $\bar{\rho}$ ($x - axis$) drawn from $\mathcal{A}$ for each value in $\mathcal{M}_{1..6}$, plot (a) to (f) respectively.

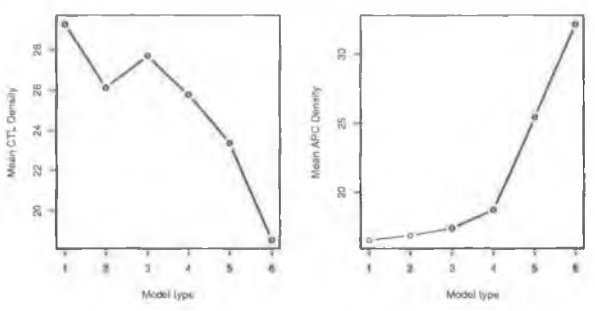


Figure 3. Mean CTL and APC density respectively. The effect of narrowing antigenic diversity can be seen. CTL density declines rapidly once the threshold diversity level of two ($\mathcal{M}_5$) is reached. Likewise the APC density increases as CTL pressure declines.

is achieved in the usual time frame of 3-5 days.

(ii) The disease is restricted in antigenic diversity, the immune system exhibits normal cross-reactivity: poor clearance leads to a chronic infected state

(iii) The disease is restricted in antigen diversity, the immune system exhibits high levels of cross-reactivity: full clearance is achieved with the risk of autoimmune disease occuring, if negative selection has not already deleted self-reactive CTL precursors from the cell pool.

(iv) The disease has some antigenic diversity, the immune system exhibits low cross-reactivity: poor clearance will result in severely chronic infection of the immune system, leading to fatality.

Given the relative rarity of autoimmune disease, we assume that excessively high levels of cross-reactivity are normally unlikely to occur. Therefore, (iii) above will be a rare disease outcome. However, the model presented here is rich enough to support all four potential disease outcomes listed above. For example, Fig. 2, plot (a) at $10 \leq \hat{\rho} \leq 15$ illustrates path 1. above, while plot (e) at $5 \leq \hat{\rho} \leq 14$ illustrates a possible path for 2. above.

5 Conclusion and Future Work

We have presented initial results from the model that suggest that a narrowing strain diversity will negatively affect the ability of the immune system to achieve healthy APC clearance, even under a highly cross-reactive regime. Conversely, we have shown how a narrowing antigenic diversity in shape space must give rise to extremely high levels of cross-reactivity in order for APC clearance to be maintained at healthy levels. In turn, we have seen that lock-and-key binding - which is the most narrow form of antigenic diversity - requires such high levels of cross-reactivity that this, in turn, may induce autoimmune disease or some form of clonal deletion to self-reactive T cell precursors. As autoimmune disease is rare in the clinical environment, we conclude that lock-and-key binding must in turn be rare.

We have assumed that the time signature for the emergence of strain diversity has been instantaneous. That is to say, each of the models $\mathcal{M}_{1..6}$, introduce their respective strain diversity levels at time $\tau = 0$. For the majority of viral pathogens encountered by the immune system, this is clearly an unrealistic starting point. However, we feel that the data presented from the above simulations is a usefule step in showing that a clear relationship exists between strain diversity, cross-reactivity and infected cell clearance rates from the secondary immune organs. Future work will include modelling the emergence of strain diversity over a more realistic time scale, as well as modelling the process of Th1 and Th2 differentiation and up and down regulation.

References

[1] Nowak M., McMichael A. J.: How HIV defeats the Immune System, Scientific American, 273, 1995, 58-65.

[2] Burns J., Ruskin H. J.: A model of Immune Suppression and Repertoire Evolution (Accepted) *Proc. of International Conference on Computational Science*, St. Petersburg, Russia. To be published in Lecture Notes in Computer Science (LNCS), 2003, Springer-Verlag.

[3] Perelson, A. Oster, G.: Theoretical Studies of Clonal Selection: Minimal Antibody Repertoire Size and Reliability of Self-Non-self Discrimination, J. Theor. Biol., 81, 1979, 645-667.

[4] Mason, D.: A very high level of crossreactivity is an essential feature of the T-cell receptor, Immunology Today, 19/9, 1998, 395-404.

[5] van den Berg, H.A., Rand D.A., Burroughs, N.J.: A Reliable and Safe T Cell Repertoire based on Low-affinity T Cell Receptors, J. Theor. Biol., 209, 2001, 465-486.

[6] Mannion, R., Ruskin, H., Pandey, R.B.: Effect of Mutation on Helper T-cells and Viral Population: A Computer Simulation Model for HIV, Theor. in Biosci., 119, 2000, 2, 1 45.

[7] Seiden, P., Celada, F.: A model for simulating cognate recognition and response in the immune system, J.Theor. Biol., 158, 1992, 329-357.

[8] Lagreca, M.C., de Almeida, Zorzenon dos Santos, R.M.C.: A dynamical model for the immune repertoire, Physica A 289, 2001, 191-207.

[9] Castiglione, F., Motta, S., Nicosia G.: Pattern recognition by primary and secondary response of an Artificial Immune System, Theory Biosci., 120, 2001, 93-106.

[10] Wolfram, S.: Cellular Automata as Simple Self-Organizing Systems, Nature 1982.

[11] Janeway, C.A., Travers, P., Walport, M., Capra, J.D.: Immunobiology. The Immune System in Health and Disease. Churchill-Livingston, 1999, 4th Edition.

[12] Zorzenon dos Santos, R.M., Coutinho, S.C.: Dynamics of HIV Infection: A Cellular Automata Approach. Phys. Rev. Lett., 87, 2001, 168102.

[13] Nowak, M., May, R.: Mathematical Biology of HIV Infections: Antigenic Variation and Diversity Threshold. Mathematical Biosciences, 106, 1991, 1-21.

[14] Hershberg, U., Louzoun, Y., Atlan, H., Solomon, S.: HIV time: winning the war while, loosing all the battles. Physica A, 289, 2001, 178-190.

[15] Orosz, C.: An Introduction to Immuno-ecology and Immuno-informatics. Design Principles for Immune System and Other Distributed Autonomous Systems. Oxford University Press, 2001.

[16] Monteiro, L.H.A., Goncalves, C.H.O., Piqueira, J.R.C.: A condition for Successful Escape of a Mutant after Primary HIV Infection. J. Theor. Biol., 203, 2000, 399-406.

[17] Nowak, M., May, R.: Virus Dynamics. Mathematical Principles of Immunology and Virology. Oxford University Press, 2000.

[18] Cameiro, J. and Stewart, J.J.: Rethinking "Shape Space": Evidence from Simulated Docking Suggests that Steric Complementarity is not limiting for Antibody-Antigen Recognition and Idiotypic Interactions. J. Theor. Biol., 169, 1994, 391-402.

[19] Yates, A, Bergmann, C., Van Hemmen, J.L., Stark, J., Callard, R.: Cytokine-modulated Regulation of Helper T Cell Populations J. Theor. Biol., 206, 2000, 539-560.

[20] Wick, D.: The Disappearing CD4T Cells in HIV Infection: a Case of Over-stimulation?, J. Theor. Biol., 197, 1998, 507-516.

Moreover, the structure and the spectral properties of the hydrogen molecule confined in a harmonic oscillator potential are studied [2]. The bond length and the vibronic transitions (the intensities and the number of lines) depend in a specific way on the strength of the confining potential. In particular, due to the confinement, the absorption and the emission vibronic bands are blue shifted.

Acknowledgements: Supported by the Polish KBN, project No. 5 PO3B 119 21

[1] D. Bielińska-Wąż, J. Karwowski, G. H. F. Diercksen, J. Phys. B **34**, 1987 (2001).

[2] D. Bielińska-Wąż, G. H. F. Diercksen and M. Klobukowski, Chem. Phys. Letters **349**,215 (2001).

Phase Diagram of a Simple Model of Water: A CVM and Monte Carlo Analysis

P. Bruscolini[1], A. Pelizzola[1], L. Casetti [2]

(1) Istituto Nazionale per la Fisica della Materia & Dipartimento di Fisica, Politecnico di Torino c.so Duca degli Abruzzi 24 I-10129 Torino (TO), Italy
(2) Istituto Nazionale per la Fisica della Materia, Unit'à di Ricerca di Firenze v. G. Sansone 1 I-50019 Sesto Fiorentino (FI), Italy

In a recent paper Patrykiejew and coworkers[1] proposed a two-dimensional model where water is described as a network forming fluid living on a triangular lattice. Fluid "particles" have three bonding arms which must lie on the lattice directions and can form bonds to other particles. A bond is weakened by the presence of a third particle on the same triangle, mimicking the perturbing effect that a water molecule may exert on an existing hydrogen bond. Thanks to its simplicity, this model can be relevant also as a starting point to study the thermodynamics of the solvation of nonpolar solutes in water.

We have studied the phase diagram of this model both with a Cluster Variation Method approach and with Monte Carlo simulations, correcting some relevant errors reported in the original paper. The results show an interesting phase diagram upon variation of the bond-weakening factor.

[1] A. Patrykiejew, O. Pizio and S. Sokolowski, Phys. Rev. Lett. **83**, 3442 (1999).

A Monte Carlo Model of Immune System T-Cell Receptor Cross-Reactivity During Primary Response

J. Burns and H. J. Ruskin

School of Computer Applications, Dublin City University, Dublin 9.
jburns@computing.dcu.ie

We present a unique Monte Carlo based cellular automata model that allows us to study aspects of the immune system by combining two distinct formalisms - (i) Physical Space and (ii) Shape Space. The motivation for combining these two formalisms comes from the observation that both local change and global condition inform the immune response to a given stimulus. One common feature of the stimuli under investigation is that they effect an alteration in the immune

CP661, *Modeling of Complex Systems: Seventh Granada Lectures*,
edited by P. L. Garrido and J. Marro

repertoire density and distribution. The shape-space formalism supports classification of the immune repertoire density and distribution, as well as classification of T-cell receptor/antigen presentation cell affinity. The objective of this paper is to examine the sensitivity of the primary immune response (during clonal expansion) to cross-reactivity of T-cell receptors (ρ). The T-cell receptors and antigen presentation cells are located at specific points within a two-dimensional shape-space, and affinity is measured by the Euclidean distance between T-cell receptor and antigen presentation cell. In order to drive our shape-space, we utilize an enhanced physical-space model to represent one lymph node. Our enhancements include - (i) realistic dynamics within the lymph node compartment accounting for cells entering and leaving via the bloodstream, (ii) Monte Carlo time steps based on the fastest aging entity, thus providing a clinically-realistic time signature, and (iii) realistic cell density levels within the lymph node compartment. As a result of these enhancements our model closely exhibits known clinical patterns during immune system primary response.

[1] Monteiro, L.H.A. et al. J. Theoretical Biol. **203**, 399-406 (2000)
[2] Hershberg, U. et al. Physica A **289**, 178-190 (2001)
[3] Janeway, C.A. et al. Immunobiology The Immune System in Health and Disease. Churchill-Livingston (1999), 4th Edition
[4] Perelson, A. and Oster, G. J. Theoretical Biol. **81**, 645-667 (1979)

Noise–Induced Nonlinear Instabilities

O. Carrillo[1,2], M. Ibañes[1] and J. M. Sancho[1]

(1) Departament d'Estructura i Constituents de la Matèria, Universitat de Barcelona, Diagonal 647, E–08028 Barcelona, Spain.
(2) e-mail: oliver@ecm.ub.es

External fluctuations can have counterintuitive ordering effects on a system. An example of these ordering effects is the appearance of a nonequilibrium transition from a disordered phase to an ordered one by increasing the intensity of the external fluctuations. Nowadays two possible physical mechanisms responsible for noise-induced ordering phase transitions have been observed: a short time *linear* instability in the disordered phase sustained by the spatial coupling and the appearance of new stable states in a nonequilibrium effective potential. Our aim here is to present analytical and numerical evidence of a short time *nonlinear* instability mechanism responsible for the appearance of second and first-order nonequilibrium phase transitions induced by an increase of the intensity of the external fluctuations. This nonlinear instability mechanism has not yet been studied in the literature, although it was conjectured in Ref. [3].

We study two models defined by the following stochastic partial differential equations:

$$\frac{\partial \phi}{\partial t} = -\phi^3(1+\phi^2) + D\nabla^2\phi + \phi^2\eta(\vec{x},t) + \xi(\vec{x},t), \tag{1}$$

$$\frac{\partial \phi}{\partial t} = -\phi(1+\phi^4) + D\nabla^2\phi + \phi^2\eta(\vec{x},t) + \xi(\vec{x},t). \tag{2}$$

Both noises ($\xi(\vec{x},t)$, internal and $\eta(\vec{x},t)$, external) are gaussian with zero mean and are δ–correlated in time. $\xi(\vec{x},t)$ is also δ–correlated in space, and $\eta(\vec{x},t)$ has a space correlation function $c(\vec{x}-\vec{x}')$.

CP661, *Modeling of Complex Systems: Seventh Granada Lectures*,
edited by P. L. Garrido and J. Marro

Glossary of Biological Terms

A

antibody A molecular secretion from a B cell which attaches to, and neutralises, the specific class of antigen which stimulated the B cell production of the antibody. As antibodies bear the same genetic receptors as the B cells which produced them they are used as a shorthand for B cell, p. 9.

antigen Genetically foreign cell or bacteria which the immune system treats as hostile, p. 10.

APC Antigen Presenting Cell. A cell, the function of which is to engulf foreign genetic material and to present genetic markers (or peptides) of the material on its surface to advertise potential infection, p. 10.

apoptosis Programmed death of a lymphocyte cell. The cell switches off in response to some internal event (such as reaching a specific age), or external event, such as no further stimulation by infected antigen presenting cells, p. 10.

B

B-cell General term for a cell originating in the bone marrow, and part of the humoral response. B-cells secrete antibodies in response to stimulation by foreign genetic material, p. 9.

C

Cayley Tree An acyclical tree in which each non-leaf vertex has a constant number of edges, p. 17.

CD4$^+$T A T-helper cell, the function of which is to up or down regulate immune activity (such as cytotoxic lymphocyte cell production) by secreting stimulatory molecules called cytokines, p. 12.

clonotype A group of T cells of the same lineage all bearing the same receptor. In genetic terms, equivalent to the *phenotype*, p. 11.

CTL Cytotoxic T Lymphocyte: A T-cell of the family CD8$^+$, which, when stimulated, becomes an 'armed effector' cell capable of killing infected antigen presenting cells by binding to the APC and injecting poision, p. 10.

cytokine Small, soluable proteins, secreted by one cell, but which can alter the properties of the cell itself or another cell. The commonest cytokine released by T cells is the interleukin (IL) class, which can surpress or enhance cytotoxic T cell efficacy, p. 12.

E

epitope General term for any part of a viral protein detectable by the immune system, p. 33.

I

idiotype The novel or idiosyncratic parts of an antibody are called idiotopes. The set of idiotopes that characterises an antibody is

called its *idiotype*. Idiotype is functionally equivalent to clonotype, p. 13.

L

lymph nodes Secondary immune organs located throughout the body, although mainly in the chest and neck, into which lymphocytes drain from the blood supply in order to sense antigen presenting cells for signs of infection, p. 32.

M

MHC Major Histocompatible Complex. The set of human immune molecules which clasp foreign antigen peptide fragments to the surface of antigen-presenting cells, p. 10.

microscopic The local space of the secondary immune organs, typically the lymph nodes, p. 21.

P

pathogen General term for a virus or bacteria which the immune system treats as hostile, p. 10.

peptide A chain of genetic material taken from a foreign antigen and displayed on the surface of the antigen presenting cell to notify lymphocytes of an immune challenge, p. 10.

precursor A cell which is not yet committed to a course of action and one which may, depending on external stimuls (such as recognition of an infected antigen presenting cell), become a cytotoxic lymphocyte, p. 10.

R

recirculation The process by which a lymphocyte travels the bloodstream, draining through some of the lymphatic compartments. Recir-

culation is a crucial process for the detection of infected antigen presenting cells, p. 9.

repertoire The set of all T-cell clonotypes, including activated, resting and naive, p. 11.

rewiring The process in which an edge between nodes xy is removed and added between nodes xk, p. 86.

robustness The degree of network function under node elimination, p. 85.

S

scalar field A scalar value is a single component that can assume one of a range of values, an example of which is *age*, p. 67.

T

TCR T Cell Receptor. A structure on the T cell surface with a fixed and variable region, the function of which is to bind to MHC:peptide fragments presented on the surface of cells, p. 11.

thymus A primary immune organ responsible for the continuous production and supply of T-cells (hence the T), p. 11.

V

vector field Vector fields have an n-component vector (usually 2 or 3 components) at each point of the space sampled for visualisation, p. 68.

virion A complete virus particle with its DNA or RNA core and protein coat as it exists outside the cell, p. 23.