Elucidation of T cell signalling models

Nick D.L. Owens, Jon Timmis, Andrew Greensted, Andy Tyrrell

www.elsevier.com/locate/yjtbi

Elucidation of T cell signalling models

Nick D.L. Owens, Jon Timmis, Andrew Greensted, Andy Tyrrell

PII: S0022-5193(09)00492-5
DOI: doi:10.1016/j.jtbi.2009.10.017
Reference: YJTBI 5745

To appear in: *Journal of Theoretical Biology*

Received date: 2 March 2009
Revised date: 8 October 2009
Accepted date: 8 October 2009

Cite this article as: Nick D.L. Owens, Jon Timmis, Andrew Greensted and Andy Tyrrell, Elucidation of T cell signalling models, *Journal of Theoretical Biology*, doi:10.1016/j.jtbi.2009.10.017

Elucidation of T cell Signalling Models

Nick D. L. Owens, Jon Timmis, Andrew Greensted and Andy Tyrrell

Department of Electronics, University of York, UK

Abstract

A potential mechanism that allows T cells to reliably discriminate pMHC ligands involves an interplay between kinetic proofreading, negative feedback and a destruction of this negative feedback. We analyse a detailed model of these mechanisms which involves the TCR, SHP1 and ERK. We discover that the behaviour of pSHP1 negative feedback is of primary importance, and particularly the influence of a kinetic proofreading base negative feedback state on pSHP1 dynamics. The CD8 co-receptor is shown to benefit from a kinetic proofreading locking mechanism and is able to overcome pSHP1 negative influences to sensitise a T cell.

Key words: T cell signalling; Kinetic Proofreading; Negative Feedback; SHP1; CD8; T cell tunability; Stochastic Modelling; Master Equation; Continuous Time Markov Chain; Numerical Analysis.

1. Introduction

The T lymphocyte is of great importance in the response of the adaptive immune system. The T cell must perform fine grain discrimination of peptide bound Major Histocompatability Complex (pMHC) molecules on antigen-presenting cells (APCs) through its T cell Receptor (TCR) [1]. The discrimination occurs between abundant self-pMHC, $99.9 - 99.99\%$ of all pMHC on an APC, and non-self-pMHC which comprises the other $0.01 - 0.1\%$ of the total pMHC expressed [2]. The ability of the T cell to respond reliably is remarkable given the bind between TCR and pMHC is low-affinity [3] and the TCR is randomly generated through somatic mutations [1]. An explanation of the T cell's behaviour can be found in the complex and dynamic signalling cascades that arise from the TCR [2]. Of particular interest is the possibility that these pathways exhibit *tunable* properties [4] which enable a T cell to alter response to a particular pMHC ligand to a greater or lesser degree.

In particular we are intersted in early TCR signalling events, those that occur immediately upon TCR pMHC engagement and before immunological synapse formation. A set of signalling events that involve kinetic proofreading, a negative influence from SH2 domain-containing phosphatase (SHP1) and a positive influence from extracellular signal regulated kinase (ERK) have been discussed in reference to tunability in [2, 5] and have been experimentally investigated in [6, 7]. This system has also received much modelling attention in [7–11]. The important quantitative and qualitative concepts that these signalling events impose on T cell discrimination are well summarised in Chan et al. [8]. Altan-Bonnet and Germain (ABG) [7] complement biological experiment with a detailed mathematical model of these signalling events. This ABG model is further discussed in [12] and analysed in [13]. The approach taken by [7] and [13] is to model, as faithfully as possible, the bio-chemical reactions associated with the TCR, SHP1 and ERK and to demonstrate their model replicates biological behaviour. To this end [7] achieves a model of 557 chemical reaction equations which are converted to 238 ordinary differential equations. Artyomov et al. [9] and Lipniacki et al. [11] both produce simplified versions of the concepts in the ABG model and perform stochastic analysis. [9, 11] argue the importance of stochastic modelling particularly when small molecule numbers can produce bistabilities in a stochastic model that do not exist in the deterministic model.

The simplified models analysed in [8, 9, 11] convey an understanding of the qualitative aspects of TCR, SHP1, ERK signalling. We take this opportunity to dissect the biologically detailed ABG model with the objective of gaining clarified understanding of the biological processes that contribute to the qualitative behaviour. We build on our previous work [14] which stochastically investigated a reduced version of the ABG model. Here we stochastically analyse the entire ABG model[1] by investigating behaviour at the single TCR level and progressively include components to compose the entire model. With this compositional analysis we achieve a greater understanding of the elements of the model without the simplifying assumptions of [11].

Section 2 provides a biological background to the model; section 3 justifies the stochastic approach and overviews the modelling methods used. The

[1]We are grateful to Gregoire Altan-Bonnet for proving us with the details of the ABG model

subsequent sections break the model and analyse sub-models. Section 4 investigates coreceptor CD8, TCR and pMHC binding; section 5 kinetic proofreading; section 6 negative feedback; section 7 the mitogen-activated protein kinase (MAPK) MAPK cascade and protection of TCR; section 8 then re-composes the entire model; finally section 9 provides discussion and conclusions.

2. Biological Background

The quality of a TCR pMHC association is well abstracted by its average lifetime [12]. It may be classified by the response illicited in the cell, taking the definitions of [2]:

- *Agonist.* Will induce all possible activation signals within a cell. A TCR-pMHC lifetime of ≈ 18 seconds [7].
- *Antagonist.* Will actively inhibit activation signals within the cell.
- *Partial agonist.* Will induce a subset of all possible activation signals within a cell.
- *Null.* Will not have any affect, activatory or inhibitory.

A range of signal strengths exists within the first three classes. It is not the case that the signals induced by a weak agonist are necessarily stronger than those induced by a partial agonist. Self-peptides fall into antagonist, partial agonist or null classes [2]. We outline the three salient features of this TCR signalling model: kinetic proofreading, negative feedback and a destruction of the negative feedback. Figure 1 depicts these features and provides reference to the appropriate section for analysis.

2.1. Kinetic Proofreading

Kinetic proofreading was first introduced to describe the accuracy of DNA replication and protein synthesis [15]. McKeithan [16] applied kinetic proofreading to T cell signalling and it is now a widely accepted model to account for ligand discrimination [6]. The process entails energy consuming steps that occur after association of pMHC to the TCR. The steps must be overcome for successful TCR signalling. With dissociation of pMHC from the TCR the steps are rapidly reversed. The result is step-like discrimination of pMHC ligands. Fast dissociating pMHC ligands fail to complete all steps, conversely

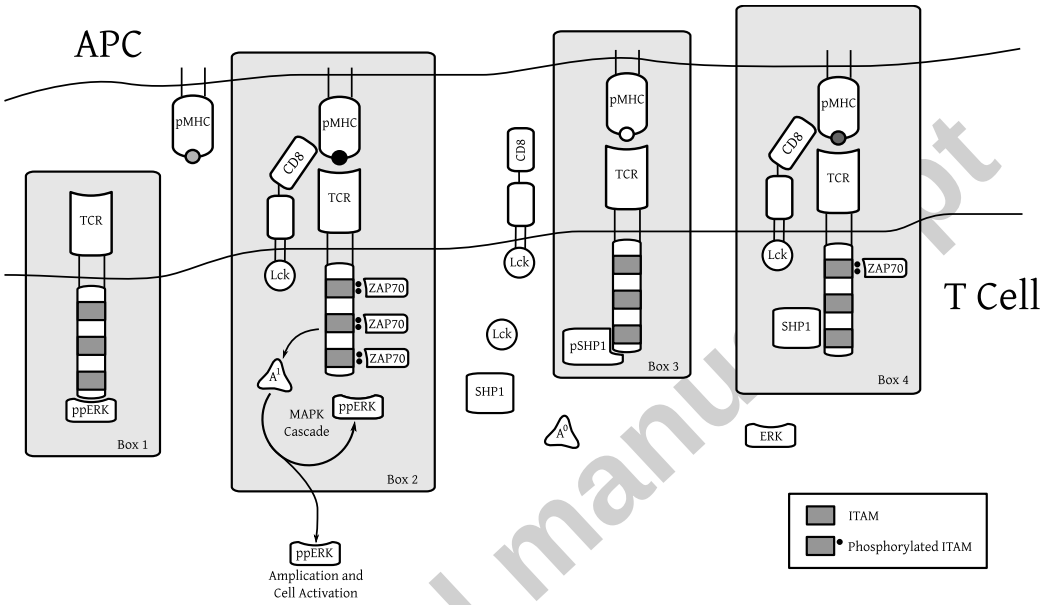

Figure 1: The TCR signalling processes. The TCR and pMHC may bind with co-receptor CD8, see *Box 2* and section 4 for analysis. Upon association of pMHC to a TCR kinetic proofreading may start which involves phosphorylation of ITAMs and binding of ZAP-70 molecules. *Box 4* depicts a partially complete proofreading process; *Box 2* depicts a fully completed proofreading process. See section 5 for kinetic proofreading analysis. A TCR internal chain with at least one ZAP-70 molecule may initiate negative feedback by binding and phosphorylating SHP1, see *Box 4*. Phosphorylated SHP1 (pSHP1) may then bind a TCR and upon further Lck action all phosphorylations will be lost, see *Box 3*. See section 6 for analysis of negative feedback. A completed proofreading process may phosphorylate an adapter protein which initiates the MAPK cascade, resulting in high levels of ppERK, see *Box 2*. ppERK carries the activation signal of the cell and may bind a TCR internal chain and protect it from the action of pSHP1, see *Box 1*. See section 7 for analysis of the MAPK cascade and TCR protection; see section 8 for simulations of the entire system.

all pMHC that bind long enough to complete all steps signal equally well. Kinetic proofreading provides a measure of the time the TCR and pMHC are associated. The steps involve phosphorylations by Leukocyte-specific protein tyrosine kinase (Lck) of Immunoreceptor tyrosine-based activation motifs (ITAMs) on the TCR's internal ζ, ϵ and δ chains. Kinetic proofreading is, however, insufficient to explain antagonism [12]. Further the model also fails in the high-density low-quality ligand case where stochastic fluctuations can allow poor quality ligands to overcome kinetic proofreading.

2.2. Negative Feedback

A negative feedback investigated experimentally in [6] may augment proofreading to explain antagonistic behaviour. The process is initiated by phosphorylation of SHP1 by Lck on the TCR internal complex. Phosphorylated SHP1 (pSHP1) may then associate to the TCR and dephosphorylate TCR internal chains. Thus the process is initiated by proofreading steps and actively inhibits proofreading and so is a true negative feedback. The inclusion of negative feedback provides an explanation for antagonism and prevents large populations of low-quality ligands stochastically overcoming proofreading. A model of proofreading with a negative feedback would suggest that the highest quality pMHC ligands would induce the largest negative feedback. However this is not the case, there is a point as ligand quality increases where the pSHP1 negative signal disappears [6].

2.3. Breaking the Negative Feedback

An explanation for the disappearance of the negative feedback is ascribed to a positive feedback through doubly phosphorylated ERK (ppERK) [7]. ppERK protects the TCR internal complexes from the action of pSHP1 by preventing pSHP1 from binding. Completion of kinetic proofreading initiates the MAPK cascade which results in the amplification of the proofreading activation signal by the production of large amounts of ppERK [7]. This acts dually to break the negative feedback and to carry the activation signal for the cell. How the activation signal determines cell fate is beyond the scope of this work.

The ppERK signal is often labelled as *digital* [7] in that it exhibits a step-like response. However a clearer description is *binary*: it is either high or low and not found in discrete steps as the term digital would imply. Moreover this signal is strictly not a positive feedback and it can be misleading to label it as one. The signal breaks the negative feedback allowing kinetic

proofreading to continue with no inhibition, but only at the rate dictated by kinetic proofreading. One would expect this to be confirmed by experiment: a T cell with SHP1 removed would exhibit no negative feedback and allow observation of ppERK behaviour in isolation. Properties traditionally associated with positive feedback such as explosive amplification are present but are facets of the feed-forward MAPK cascade. We interpet a result found in Feinerman et al. [13] to confirm this is not a positive feedback. Variation in the concentration of ppERK has no influence on the ability of a cell to appropriately signal. If ppERK was involved in a positive feedback its intracellular concentration represents a maximum bound on the positive signal. Thus variations in ppERK concentration would change the ability of the cell to signal. Interpretations of results in Lipniacki et al. [11] give further weight to this not a positive feedback argument. Changes in Lck concentration influence pSHP1 levels far more than ppERK levels, if ppERK were in positive feedback the action of Lck would be part of the feedback loop and so would influence ppERK levels. We note that the behaviour of the MAPK cascade (analysed in section 7) also contributes to these results. We shall refer to the ppERK signal as a break in the negative feedback and not a positive feedback. As discussed in Chan et al. [8] the break in the negative feedback allows the T cell to remain sensitive to good agonists and helps to define a sharp discrimination threshold.

The signalling mechanisms may be mapped to TCR-pMHC bind classification:

1. Agonist – pMHC remains associated sufficiently long for a high probability of inducing MAPK cascade despite pSHP-1 dampening. The induction of the MAPK cascade will result in a high ppERK protection and activation signal.
2. Antagonist – pMHC associates sufficiently long to produce high levels of pSHP-1 but not to induce MAPK cascade.
3. Partial Agonist (Endogenous Ligand) – Does not induce high levels of pSHP-1, but may induce some partial phosphorylation of TCR internal chains.

The spreading of the pSHP-1 and ppERK signals is of interest, it is discussed in [7] and analysed directly in [8]. pSHP1 generated by a TCR will spread to dampen surrounding TCRs. The protection signal generated by a TCR will spread to protect surrounding TCRs. A hypothesis discussed

in [8, 12] suggests that the protection signal allows endogenous ligands to "synergise" with agonist ligands and contribute to activation signalling. We discuss this suggestion in our conclusions (section 9) in light of the results of this paper and present a clarification of this argument. The model of Chan et al. [8] investigates the spreading of the negative feedback and protection on a lattice. No investigation to the necessity of space to the model is given, for our model we continue with the assumptions of [7] and analyse a non-spatial model.

3. Modelling Methods

In contrast to the deterministic approach of [7] and [13] the models of [8–11] are stochastic in nature. The case for stochastic modelling in biology in general and particularly in T cell signalling is well argued by Artyomov et al. [9]. Stochastic fluctuations become important when small molecule numbers are involved (as is the requirement in T cell signalling, sensitivity to < 10 pMHC ligands [12]). Under small molecule conditions bistablities can exist in the stochastic model which are not present in the deterministic/macroscopic model. Artyomov et al. [9] give necessary and sufficient conditions for this deterministic versus stochastic discrepancy. These conditions are fulfilled by certain reactions investigated in this paper. Lipniacki et al. [11] also demonstrate the quantitative discrepancies that occur in a simplified version of the ABG model when the deterministic formulation misses the stochastic bistabilities. Further, as noted in [17] when choosing between a stochastic and deterministic approach the stochastic model should be the default and a deterministic model should only be used with sufficient justification. To avoid any deterministic model errors we take the stochastic approach. This approach also provides a method of investigating the reactions of an individual TCR which proves extremely beneficial to understanding the many TCR simulations performed in section 8.

We take two approaches to stochastically analyse the reaction equations of the ABG model. We analyse sub-systems via a master equation/continuous time Markov chain [18]. Due to model size we are generally restricted to numerical solutions associated of a *rate* matrix, an overview of these methods is given in appendix Appendix E. As we increase the size of the model we perform simulation with the Gillespie algorithm [19].

The reaction rates supplied in the ABG model are macroscopic and require conversion to mesoscopic rates for stochastic analysis/simulation. The

conversion applies to second order reactions and amounts to a change of scale and units (from $\text{mol}^{-1}s^{-1}$ to s^{-1}). To reduce computational complexity we analyse a fraction $\epsilon \in (0, 1]$ of the cytoplasmic cell volume V. To convert a macroscopic rate d to a mesoscopic rate r we use the following:

$$r = \frac{d}{N_A V \epsilon} \tag{1}$$

$N_A = 6.022 \times 10^{23}$ is the Avogadro Number. For first order reactions the mesoscopic rate is augmented with the number of reactants; for second order reactions the rates are augmented with the number of pairs of reactants. The model contains no reactions higher than second order. The macroscopic rates of the ABG model are supplied in table A.5. If a molecule has an initial concentration $c \in \mathbb{R}$ then the initial discrete number of molecules $N \in \mathbb{Z}$ is taken to be $N = \lfloor cN_A V\epsilon \rfloor$. Three differing choices of ϵ are used in this paper, we provide justification of these choices in appendix Appendix B. The definitions of names and rates are given in appendix Appendix A and notation used in reaction equations is given in appendix Appendix F.

4. TCR, pMHC, CD8 Binding

The ABG model enumerates the complexes nascent to the CD8-TCR-pMHC ternary complex. Reactions in and out of intermediate states TCR-pMHC, CD8-TCR, and CD8-pMHC are included. Other models that include co-receptor binding such as [10] do not include CD8-TCR and CD8-MHC intermediates. We investigate the role of these intermediates, particularly the influence of TCR-CD8. The reactions are depicted in figure 2 and are as follows:

Reaction 1. *TCR pMHC CD8 Binding/Debinding*

$$\mathrm{T} + \mathrm{M} \underset{m_d}{\overset{m_a}{\rightleftharpoons}} \mathrm{B} \qquad \mathrm{T} + \mathrm{C} \underset{c_d}{\overset{c_t}{\rightleftharpoons}} \mathrm{TC} \tag{2}$$

$$\mathcal{A} + \mathrm{C} \underset{c_d}{\overset{c_a}{\rightleftharpoons}} \mathcal{A}\mathrm{C} \qquad \mathcal{A} = \{\mathrm{M}, \mathrm{C}\} \tag{3}$$

$$\mathrm{TC} + \mathrm{M} \underset{c_d \times c_s}{\overset{c_a}{\rightleftharpoons}} \mathrm{BC} \qquad \mathrm{MC} + \mathrm{T} \underset{c_d \times c_s}{\overset{c_a}{\rightleftharpoons}} \mathrm{BC} \tag{4}$$

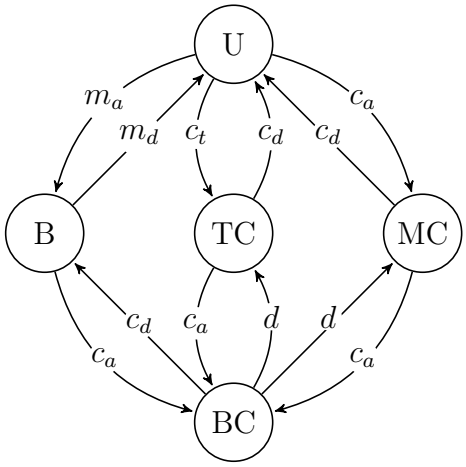

Figure 2: Diagrammatic representation of reaction 1. A single TCR, pMHC and CD8 molecule are the reactants. U gives the completely unbound state, all other states are denoted by the bound product. Arcs are labelled with the reaction rate, d denotes reaction rate $m_d c_s$.

The reactions describe the following assumptions: The CD8 co-receptor acts to hook-in pMHC to the TCR increasing the on-rate; CD8 stabilises the TCR-pMHC by a multiplicative factor $c_d < 1$; CD8 may dissociate and re-associate from with fast kinetics [20], with the exception of the state TC whose on rate c_t which is considerably smaller than c_a.

A volume restriction $\epsilon = N_T^{-1}$ is applied to scale the rates to the volume occupied by a single TCR. The stationary and transient distributions parameterised by $m_d \in [0.001, 1000]$ are given in figures 3 and 4 respectively. The distributions show that the unbound state U and the bound state B have the highest probability. The probability of being in state TC is orders of magnitude lower than being in any other state nascent to BC. If we write $\mathbb{P}_i(j)$ for the probability that given we are in state i the next state is state j, then setting $m_d = 1/18$ and all other rates as table A.5:

$$\mathbb{P}_{\mathrm{U}}(\mathrm{TC}) < \mathbb{P}_{\mathrm{U}}(\mathrm{B}) < \mathbb{P}_{\mathrm{U}}(\mathrm{MC}) \tag{5}$$

There is an order magnitude difference between the three probabilities. In the opposite direction, the probabilities for leaving state BC:

$$\mathbb{P}_{\mathrm{BC}}(\mathrm{TC}) = \mathbb{P}_{\mathrm{BC}}(\mathrm{MC}) \ll \mathbb{P}_{\mathrm{BC}}(\mathrm{B}) \tag{6}$$

The inequality will hold if $m_d < c_d/c_s = 1000$ which is the case for realistic ligands. If we write $\tau(i)$ for the holding time of state i (that is the expected

time before leaving state i given that we have entered it):

$$\tau(\mathrm{BC}) < \tau(\mathrm{TC}) = \tau(\mathrm{MC}) \ll \tau(\mathrm{U}) < \tau(\mathrm{B}) \quad \text{for } \epsilon = 1 \tag{7}$$

$$\tau(\mathrm{TC}) = \tau(\mathrm{BC}) < \tau(\mathrm{MC}) \ll \tau(\mathrm{B}) < \tau(\mathrm{U}) \quad \text{for } \epsilon = N_T^{-1} \tag{8}$$

$\tau(\mathrm{U}) \approx \tau(\mathrm{B})$ and $\tau(\mathrm{BC}) \approx \tau(\mathrm{MC})$ for $\epsilon = N_T^{-1}$ and the inequalities hold for the majority of realistic ligands. The state TC is the least likely state to enter and has one of the least holding times, which explains its low probability at equilibrium.

We recalculate the transient and stationary distributions with the removal of state TC (the new stationary distribution is labelled π^{TC}) and they are shown in figures 3 and 4 respectively. The stationary distributions and the trajectories are effectively identical. Under the assumption of approximately equal TCR, pMHC and CD8 concentrations the state TC may be removed from the model. This provides a predication that may be validated against biology, should more than $2 \times 10^{-3}\%$ of all TCRs be found bound solely to the CD8 co-receptor then this is an incorrect abstraction of biology. The incorrect abstraction may not be in structure of reactions but in choice of rate. A clear example of this is state MC which is symmetrically identical except the rate from U to MC is two orders of magnitude greater than that U to TC which gives rise to an approximately three orders of magnitude increased stationary probability in figure 3.

To understand the implications of parameter choice a general system is analysed algebraically. The rate matrix and solutions for the system are given in appendix Appendix C in equation C.3. Computing the partial derivatives $\partial\pi/\partial c_t$ further demonstrates it is only the low value of c_t in the ABG model that limits the probability of state TC. We also solve the stationary distribution equations of the simpler system with state TC removed (equation C.4) for m_d subject to the additional constraint:

$$\pi_U = \pi_B + \pi_{BC} \tag{9}$$

The solution provides the dissociation rate at there is equal probability of the TCR and pMHC being bound and unbound. This gives a quadratic in m_d:

$$\begin{aligned} m_d^2 c_d c_s + m_d(c_d^2 + c_d(c_a + c_a c_s - m_a c_s) - c_a^2) \\ -c_a^3 - c_a^2(m_a + c_d) - 2c_a c_d m_a - c_d^2 m_a = 0 \end{aligned} \tag{10}$$

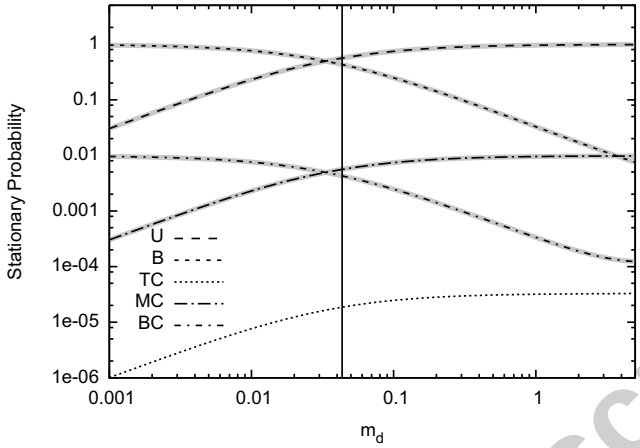

Figure 3: Stationary distributions for $m_d \in [0.001, 5]$. Black dotted lines are the distribution π, thick grey lines are the distribution π^{TC}. They are effectively identical behaviour across the range of m_d. The vertical line is at $m_d = 0.0434$ and represents the point where $\pi_U = \pi_B + \pi_{BC}$.

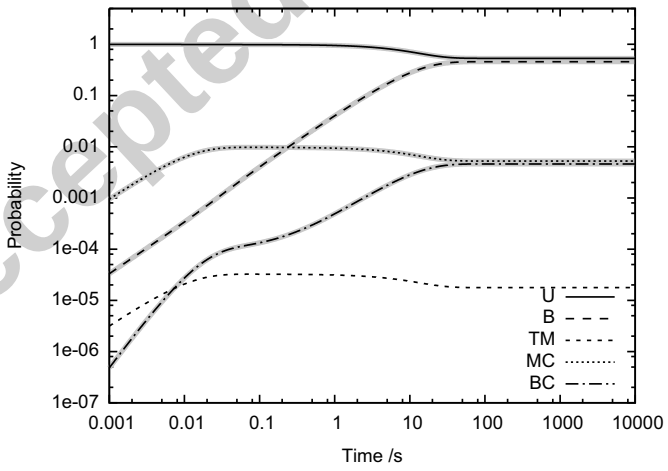

Figure 4: Transient distributions for time $t \in [10^{-3}, 10^4]$, starting in an unbound state U with $m_d = 0.05$. Black dotted lines are the distribution π, thick grey lines are the distribution π^{TC}. The probabilities follow the same trajectories.

There is a single positive solution $m_d = 0.0434$ which gives an average bind life time of 23.04 seconds, a strong agonist. The $m_d = 0.0434$ line is marked on figure 3.

The analysis thus far does not apply if the concentrations of TCR, CD8 or MHC change. The concentration of CD8 is of interest as it has been shown to be a potential parameter for T cell tuning [13, 21]. We define $v = |\mathrm{C}|/|\mathrm{T}|$ and then define new rates: $c_t' = vc_t$ the association rate of T and C; $c_a' = vc_a$ the association of C to M or B. If we recalculate transient and stationary distributions a difference is only observed at very high densities $v > 100$, so we may be confident in stationary analysis of the reduced binding system when $v < 100$.

5. Kinetic Proofreading

The kinetic proofreading in the ABG model has added complexity over standard kinetic proofreading models [16]. It is described by the following reactions:

Reaction 2. *Soluble Lck Phosphorylation:*
Normal Phosphorylation:

$$\mathrm{B}_z^q + \mathrm{L} \underset{l_d}{\overset{l_a}{\rightleftharpoons}} \mathrm{B}_z^q\mathrm{L} \xrightarrow{l_c} \mathrm{B}_z^{q+1} + \mathrm{L}, \qquad q = 0, 1;\ z = 0, 1, 2 \tag{11}$$

$$\mathrm{B}_3^m + \mathrm{L} \underset{l_d}{\overset{l_a}{\rightleftharpoons}} \mathrm{B}_3^m\mathrm{L} \xrightarrow{l_c} \mathrm{B}_3^{m+1} + \mathrm{L}, \qquad m = 0, 1, 2 \tag{12}$$

Reaction 3. *CD8 association Lck Phosphorylation:*
Normal Phosphorylation:

$$\mathrm{B}_z^q\mathrm{C} \underset{l_{cd}}{\overset{l_{ca}}{\rightleftharpoons}} \mathrm{B}_z^q\mathrm{CL} \xrightarrow{l_{cc}} \mathrm{B}_z^{q+1}\mathrm{C}, \qquad q = 0, 1;\ z = 0, 1, 2 \tag{13}$$

$$\mathrm{B}_3^m\mathrm{C} \underset{l_{cd}}{\overset{l_{ca}}{\rightleftharpoons}} \mathrm{B}_3^m\mathrm{CL} \xrightarrow{l_{cc}} \mathrm{B}_3^{m+1}\mathrm{C}, \qquad m = 0, 1, 2 \tag{14}$$

Lck associated with CD8 is implicitly part of the C complex.

Reaction 4. *ZAP-70 Binding:*

$$\mathrm{B}_z^2 + \mathrm{Z} \underset{z_d}{\overset{z_a}{\rightleftharpoons}} \mathrm{B}_{z+1}^0, \qquad z = 0, 1, 2 \tag{15}$$

ZAP-70 Debinding from TCR:

$$\mathrm{T}_z \xrightarrow{z_d} \mathrm{T}_{z-1} + \mathrm{Z}, \qquad z = 1, 2, 3 \tag{16}$$

Reaction 5. *Desphosphorylation upon TCR binding/debinding:*

$$\mathcal{Q}_z^q \xrightarrow{\delta} \mathrm{T}_z \qquad \mathrm{T}_z \xrightarrow{\alpha} \mathcal{Q}_z, \qquad z = 0, 1, 2, 3 \quad q > 1 \tag{17}$$

Here if $\mathcal{Q} = \mathrm{B}$ then $\alpha = m_a, \delta = m_d$ and if $\mathcal{Q} = \mathrm{BC}$ then $\alpha = c_a, \delta = m_d c_s$

The reactions imply the following assumptions:

- Upon association of pMHC to the TCR internal chains undergo enzymatic Lck phosphorylation. This may occur via cytosolic Lck or Lck associated with CD8. The CD8 associated Lck has a greater rate of association to the TCR $l_{ca} \gg l_a$. The phosphorylations are restricted to occur in serial in order.

- ZAP-70 may bind and protect a double phosphorylated ITAM and 3 ZAP-70 molecules may bind in all. Three phosphorylations may occur once the third and final ZAP-70 molecule has bound.

- Dissociation of TCR and pMHC causes loss of phosphorylations not protected by ZAP-70 via a fast acting phosphatase. The fast-acting phosphatase is not modelled here, the phosphorylations are immediately lost upon dissociation. ZAP-70 and its protected phosphorylations may persist between TCR binds.

- The TCR internal chain may initiate activation signalling from a state B_3^q with $q > 1$. The specific details of the activation signalling are not included until section 7.

These reactions do not include the negative effects of pSHP1 and so describe the case when the TCR has been protected by ppERK. We write $\mathbf{K}$ for the rate matrix describing the kinetic proofreading reactions for single TCR, pMHC, CD8 molecules. We calculate the expected kinetic proofreading state $\mathbb{E}_K$. We assign weights $\omega(s)$ linearly to all states $s \in \mathcal{K}$ the set of bound kinetic proofreading states:

$$\omega(\mathrm{B}_z^q) = 3z + q + 1 \tag{18}$$

m_d	$\max\{\mathbb{P}(\mathrm{B}_3^q,\ q>1)\}$
0.03	1.6×10^{-1}
0.06	0.3×10^{-1}
0.12	4.5×10^{-3}
0.23	2.6×10^{-3}
0.44	1.0×10^{-5}
0.85	4.3×10^{-7}
1.64	2.5×10^{-8}
3.16	2.5×10^{-9}

Table 1: Maximum probability of signalling state for various m_d. We can approximate the expected number of ligands needed to achieve activation via $[\max\{\mathbb{P}(\mathrm{B}^q\mathrm{Z}_3,\ q>1)\}]^{-1}$ (the coupling between TCRs at this stage only involves competition for Lck and CD8, as such we believe this approximation should be reasonable). For $m_d^{-1}\approx 30$s about 6 ligands are needed to guarantee a signalling state; for $m_d^{-1}\approx 16$s about 25 ligands are required to guarantee signalling state.

That is, $\omega(\mathrm{B}_0^0)=1,\omega(\mathrm{B}_0^1)=2,\ldots,\omega(\mathrm{B}_3^3)=13$. Then $\mathbb{E}_K$ is given:

$$\mathbb{E}_K=\sum_{s\in\mathcal{K}}\omega(s)\mathbb{P}(s) \tag{19}$$

Figure 5 gives the transient probability for $\mathbb{E}_K$ over a range of m_d. We summarise kinetic proofreading behaviour by calculating the maximum probability of signalling state, that is $\max\{\mathbb{P}(\mathrm{B}_3^q, q>1)\}$, the results are given in table 1.

Figure 6 gives the kinetic proofreading stationary distributions over m_d. The clumping together of probabilities of mid-kinetic proofreading states at low dissociation rates conveys the notion that all states are equally likely enroute to the final state. Due to the high concentration of ZAP-70 the states in which a new ZAP-70 molecule may bind (B_0^2, B_1^2, B_2^2) have the lowest stationary probability ($\approx 10^{-6}$).

We analyse the time taken by kinetic proofreading and particularly the influence of CD8 by calculating the expected hit times of TCR signalling state while varying v and m_d, see figure 7. CD8 density v can modulate hitting times and bring them in line with the 1-5 minute timescales discussed in [12]. An explanation as to why CD8 density is so successful at decreasing kinetic proofreading times is given in the following section 5.1.

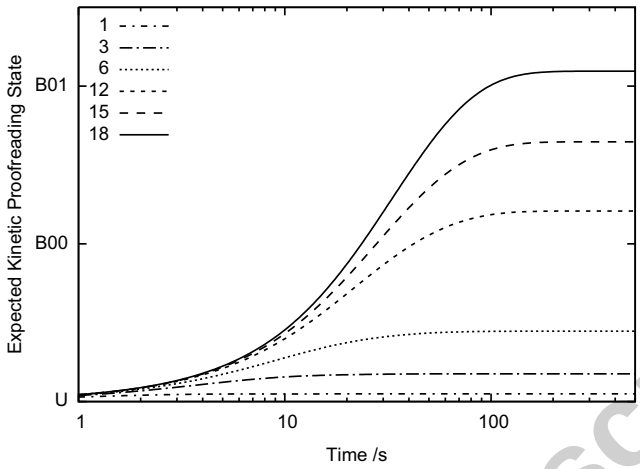

Figure 5: The transient probability of $\mathbb{E}_K$ from an initial unbound zero phosphorylation state, calculated from $\mathrm{e}^{t\mathbf{K}}$. The dissociation rates are set $m_d^{-1} = 1, 3, 6, 12, 15, 18$.

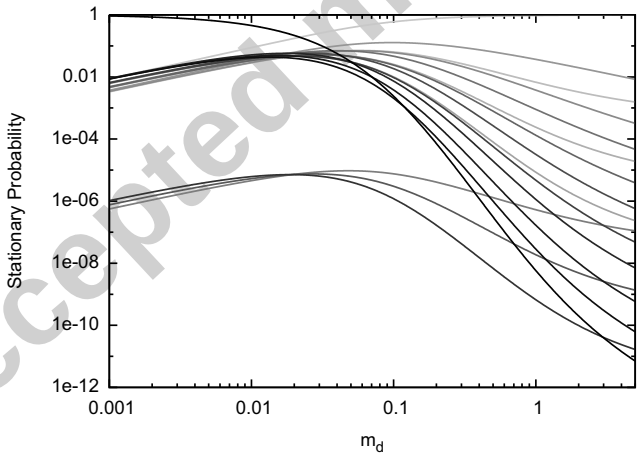

Figure 6: Kinetic proofreading stationary distribution. Lines represent U_z or B_z^q states and are coloured on a linear greyscale with lightest representing U_0 (unbound with zero ZAP-70 molecules) and the darkest B_3^3 (TCR bound to pMHC with three ZAP-70 molecules triply phosphorylated). $\mathbb{P}(\mathrm{U}_0) \approx 1$ at large m_d, and $\mathbb{P}(\mathrm{B}_3^3) \approx 1$ at small m_d. The inflexion between light and dark lines can be interpreted as the point where the latter half of the kinetic proofreading process becomes more probable.

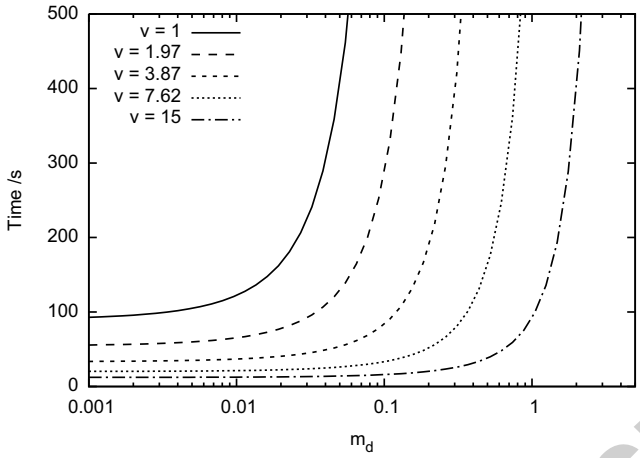

Figure 7: Expected hitting times of kinetic proofreading signalling state B_3^q, $q > 1$ varying m_d and v. Modulation of CD8 dictates the dissociation rate at which there is a steep increase in kinetic proofreading times.

5.1. Proofreading Locking Mechanism

The kinetic proofreading analysed here produces the step-like response prescribed by conventional kinetic proofreading models [16]. However the ability of ZAP-70 to "protect" TCR phosphorylations between TCR binds breaks McKeithan's model. We directly investigate this mechanism. Removing the locking system amounts to changing reaction 5 to the following:

Reaction 6. *Desphosphorylation and de-binding of ZAP-70 upon TCR de-binding:*

$$\mathcal{Q}_z^q \xrightarrow{\delta} \mathrm{T}_0^0, \qquad z = 0, 1, 2, 3 \quad q = 0, 1, 2, 3 \tag{20}$$

If $\mathcal{Q} = B$ then $\delta = m_d$ and if $\mathcal{Q} = BC$ then $\delta = m_d c_s$

Calculating non-locking kinetic proofreading transient and stationary distributions reveals a reduced probability of all later kinetic proofreading states. Further, the stationary distributions do not exhibit the inflexion of figure 6.

We investigate the differences between non-locking and locking proofreading by varying pMHC and CD8 densities. We define $u = |\mathrm{M}|/|\mathrm{T}|$ and define new rates: $m_a' = um_a$ the association rate of T and M; $c_m' = uvc_a$ the association rate of C and M; $m_c = uc_a$ the association rate of M to TC; $c_a' = vc_a$ the association rate of C to B.

We recalculate the stationary distributions for $\mathbb{E}_K$ using the new rates and vary u and v. Figure 8 gives these distributions [2] for a good agonist $m_d = 1/18$, we found the behaviour to be qualitatively similar for other values of m_d. High concentrations of CD8 and pMHC result is high $\mathbb{E}_K$. A high concentration of pMHC or CD8 increases TCR association rates and so increases the probability that the TCR and pMHC will re-associate before a ZAP-70 is lost. For $\epsilon = N_T^{-1}$ we have mesoscopic rates $m_a = 0.033$, $d_z = 0.11$. An increase in pMHC concentration of 3.33 will make re-association of the TCR and pMHC equally likely as the loss of ZAP-70. An increase in CD8 density v also contributes to the completion of kinetic proofreading steps due to efficient CD8-Lck. Thus changes CD8 density change the T cell's reactivity to a given pMHC ligand and it is the locking mechanism which allows the CD8 co-receptor to be efficacious. We confirm the ZAP-70 hypothesis in figure 9 by varying ZAP-70 dissociation rate d_z with the inclusion of the locking mechanism. Decreases in d_z result in increases in $\mathbb{E}_K$ because slow d_z increases the probability that a ZAP-70 will remain associated between TCR-pMHC associations. Due to the high abundance of cytosolic ZAP-70 any ZAP-70 molecules lost during TCR-pMHC association will be rapidly replaced. Thus the periods when the TCR is dissociated benefit the most of decreases in d_z. Slow dissociations of ZAP-70 increase the efficacy of serial triggering.

6. Negative Feedback

The negative feedback signal is carried via cytosolic SHP-1, which may bind to a TCR internal chain with at least one ZAP-70 molecule. The following describes the negative feedback process:

Reaction 7. *SHP-1 Binding, and phosphorylation to pSHP-1*

$$\mathrm{B}_z + \mathrm{S}^0 \underset{s_d}{\overset{s_a}{\rightleftharpoons}} \mathrm{B}_z\mathrm{S}^0 \xrightarrow{s_c} \mathrm{B}_z + \mathrm{S}^1 \qquad z = 1, 2, 3 \tag{21}$$

The phosphorylated pSHP-1 may load and unload from the TCR:

[2] The extremes of density here are larger than would occur naturally, they are included to demonstrate the range of behaviour, moreover the contribution to the parameters u and v could equally come from increased association rates.

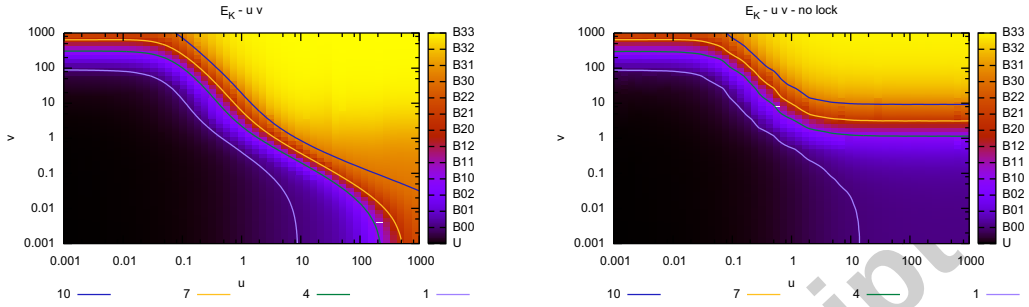

Figure 8: Stationary probabilities of $\mathbb{E}_K$ for $m_d = 1/18$, varying u and v with ZAP-70 locking left; without ZAP-70 locking right. The locking mechanism greatly increases the area of (u, v)-space with a high expected proofreading state. This is due to u and v contributing to a greater probability of TCR-pMHC re-association before a ZAP-70 is lost. Without locking and $v < 1$ results in a low expected proofreading state

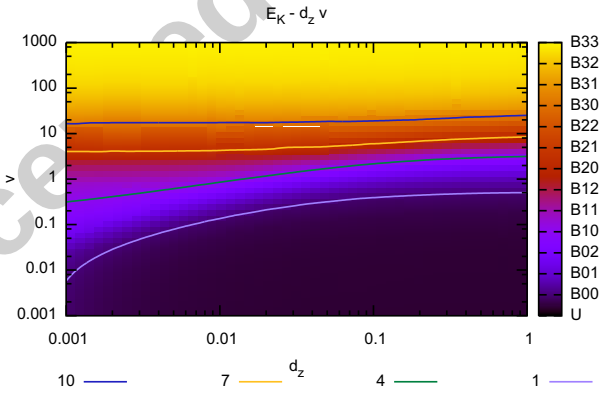

Figure 9: Stationary probabilities for $\mathbb{E}_K$ with $m_d = 1/18$ varying ZAP-70 dissociation rate d_z and CD8 density v with the locking mechanism. Decreases in d_z increase $\mathbb{E}_K$ because the probability that a ZAP-70 molecule will remain associated between TCR pMHC associations.

Reaction 8. *Binding of pSHP-1*

$$\mathcal{D} = \{\mathrm{T}_z,\, \mathrm{B}_z,\, \mathrm{B}_3^0\}, \qquad z = 0, 1, 2$$

$$\mathcal{D} + \mathrm{S}^1 \underset{s_d}{\overset{s_a}{\rightleftharpoons}} \mathcal{D}\mathrm{S}^1 \tag{22}$$

This applies for $z < 3$ and $q > 0$ cases, for $z = 3$ there is one-step dephosphorylation:

$$\mathrm{B}_3^q + \mathrm{S}^1 \xrightarrow{s_a} \mathrm{B}_3^0 + \mathrm{S}^1, \qquad q = 1, 2, 3 \tag{23}$$

Upon Lck phosphorylation of a TCR internal chain with pSHP1 loaded all phosphorylations not protected by ZAP-70 are lost:

Reaction 9. pSHP1 Desphosphorylation through Lck:

$$\mathrm{B}_z^q\mathrm{S}^1 + \mathrm{L} \underset{l_d}{\overset{l_a}{\rightleftharpoons}} \mathrm{B}_z^q\mathrm{S}^1\mathrm{L} \xrightarrow{l_c} \mathrm{B}_z^0\mathrm{S}^1 + \mathrm{L}, \qquad q = 0, 1, 2;\ z = 0, 1, 2 \tag{24}$$

$$\mathrm{B}_3^0\mathrm{S}^1 + \mathrm{L} \underset{l_d}{\overset{l_a}{\rightleftharpoons}} \mathrm{B}_3^0\mathrm{S}^1\mathrm{L} \xrightarrow{l_c} \mathrm{B}_3^0\mathrm{S}^1 + \mathrm{L} \tag{25}$$

CD8 association Lck Phosphorylation:

$$\mathrm{B}_z^q\mathrm{CS}^1 \underset{l_{cd}}{\overset{l_{ca}}{\rightleftharpoons}} \mathrm{B}_z^q\mathrm{CS}^1\mathrm{L} \xrightarrow{l_{cc}} \mathrm{B}_n^0\mathrm{CS}^1, \qquad q = 0, 1;\ n = 0, 1, 2 \tag{26}$$

$$\mathrm{B}_3^0\mathrm{CS}^1 \underset{l_{cd}}{\overset{l_{ca}}{\rightleftharpoons}} \mathrm{B}_3^0\mathrm{CS}^1\mathrm{L} \xrightarrow{l_{cc}} \mathrm{B}_3^0\mathrm{CS}^1, \tag{27}$$

The Lck associated with CD8 is implicitly part of the C complex.

The above reactions taken from [7] do not include a conversion of pSHP1 back to SHP1 and so the set of states with all SHP1 converted to pSHP1 are absorbing. To ensure that all states are irreducible and recurrent so the stationary distribution exists and is meaningful, the following reaction is added:

Reaction 10. Desphosphorylation of pSHP1.

$$\mathrm{S}^1 + \mathrm{P_S} \underset{q_d}{\overset{q_a}{\rightleftharpoons}} \mathrm{P_S}\mathrm{S}^1 \xrightarrow{q_c} \mathrm{P_S} + \mathrm{S}^0 \tag{28}$$

q_d, q_c are chosen in line with the phosphatases of the MAPK cascade; q_a is chosen similar to the association of Lck. The association rate q_a is rate-limiting and pSHP1 will persist on a timescale longer than a TCR-pMHC association. With this choice of rates we observe quantitatively identical transient behaviour over the first 200 seconds with and without reaction 10. Since a less than maximal level of pSHP1 is necessary to contain the subsequently discussed rise of $\mathbb{E}_K$, the inclusion of this reaction 10 should not significantly alter the activation threshold of the cell. Thus we feel justified in the choice of rates as they do not influence initial transient behaviour which is of importance to our conclusions (section 9). We may now calculate a potential stationary distribution for negative feedback.

There is a discrepancy between pSHP1 dephosphorylation in the B_3 case compared to B_q with $q = 0, 1, 2$. An investigation is provided in appendix Appendix D and we continue with *uniform negative feedback* as described in appendix Appendix D. A volume restriction of $\epsilon = N_T^{-1}$ gives ≈ 26 SHP-1 molecules. The rate matrix for negative feedback may be written as block tridiagonal matrix:

$$\mathbf{N} = \begin{pmatrix} \mathbf{Y_0} & \mathbf{A} & & \\ \mathbf{Q} & \mathbf{Y_1} & \ddots & \\ & \ddots & \ddots & \mathbf{A} \\ & & \mathbf{Q} & \mathbf{Y_m} \end{pmatrix} \qquad \mathbf{Y_i} = \begin{pmatrix} \mathbf{K} & \mathbf{P} \\ \mathbf{D} & \mathbf{K} \end{pmatrix} \tag{29}$$

$\mathbf{Y_i}$ is the kinetic proofreading matrix with i pSHP1 molecules. $\mathbf{K}$ describes the basic kinetic proofreading reactions; $\mathbf{P}$ and $\mathbf{D}$ give the binding and debinding of the pSHP1 phosphatase respectively. $\mathbf{A}$ contains s_c from reaction 7 on appropriate diagonal positions. $\mathbf{Q}$ contains rate q_c from reaction 10 on all diagonal positions. $\mathbf{N}$ is 7420×7420 with only 50438 non-zeros (approximately 0.09% of $\mathbf{N}$), it is amenable to the analysis methods outlined in appendix Appendix E.

The influence of negative feedback will be overestimated in the stationary distribution without the inclusion of the protection from pSHP1 by ppERK (see section 7). However the TCR must generate an activation signal at least once without ppERK and so consideration of the non-protection case is necessary.

Similar to $\mathbb{E}_K$ we also calculate the expected pSHP1 level $\mathbb{E}_S$:

$$\mathbb{E}_S = \sum_{i=1}^{\max|\mathrm{S}^1|} i\mathbb{P}(|\mathrm{S}^1| = i) \tag{30}$$

In figure 10 we plot the transients of $\mathbb{E}_K$ and $\mathbb{E}_S$. The increase of $\mathbb{E}_S$ lags that of $\mathbb{E}_K$ because kinetic proofreading must reach state B_1^1 before pSHP1 can be produced. The level of pSHP1 will rise, arrest and reverse kinetic proofreading to a state $< \mathrm{B}_1^1$. With the conversion of pSHP1 back to SHP1 equilibrium is reached. We believe the existence of non-zero kinetic proofreading state that must be overcome before negative feedback is generated to be of great importance. We determine this state the *base negative feedback state*. If we recalculate the transients of $\mathbb{E}_K$ and $\mathbb{E}_S$ with the first kinetic proofreading bound state (B_0^0) as the base negative feedback state we do not observe the hump in $\mathbb{E}_K$ or $\mathbb{E}_S$. Moreover the expected kinetic proofreading state and consequently expected pSHP1 levels are very low. In section 7 we see that the existence of this base negative feedback state is of importance at the population level. The implications of this state are discussed in detail in our conclusions (section 9).

The stationary distributions of $\mathbf{N}$ and $\mathbb{E}_S$ (with the base negative feedback state reinstated) are given in figure 11. States with zero phosphorylations not protected by ZAP-70 have the greatest probability. We write $\mathbb{E}_K(m_d)$ and $\mathbb{E}_S(m_d)$ for the expectations parameterised by m_d. For extremes of realistic ligands we have $\mathbb{E}_S(0.05) \approx 15$ and $\mathbb{E}_S(1) \approx 2$. So kinetic proofreading of high quality pMHC ligands may be arrested by just over half-maximal pSHP1. One would expect confirmation of this behaviour would come from experimentation with cells with ERK removed. One could confirm that pSHP1 will rise sufficiently the dampen kinetic proofreading of all ligands including agonists. Observations of the phosphorylation states of TCR internal chains would confirm the behaviour of the ZAP-70 locking mechanism.

6.1. Positive Tuning and Negative Feedback

In section 5, particularly section 5.1, we demonstrated that CD8 and pMHC density parameters u and v may desensitise the cell. We now investigate their influence with the inclusion of negative feedback. Figure 12 gives the stationary distributions for $\mathbb{E}_K$ and $\mathbb{E}_S$. Increases in u and v initially result in increases in $\mathbb{E}_K$ and $\mathbb{E}_S$, with u being slightly more efficacious. However further increases in v result in a maximum in $\mathbb{E}_S$. This demonstrates

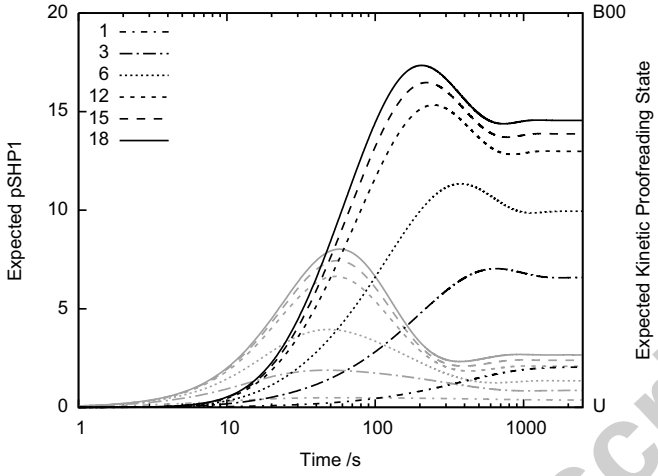

Figure 10: The transients of $\mathbb{E}_S$ (black lines, left axis) and $\mathbb{E}_K$ (grey lines, right axis) from initial unbound zero phosphorylation state U. With $m_d^{-1} = 1, 3, 6, 12, 15, 18$. Levels of pSHP1 lag the increase in kinetic proofreading state. For strong enough ligands pSHP1 overshoots its stationary value and decreases $\mathbb{E}_K$, then $\mathbb{E}_K$ falls to its stationary value. Compare the heavily reduced values of $\mathbb{E}_K$ to kinetic proofreading without negative feedback in figure 5

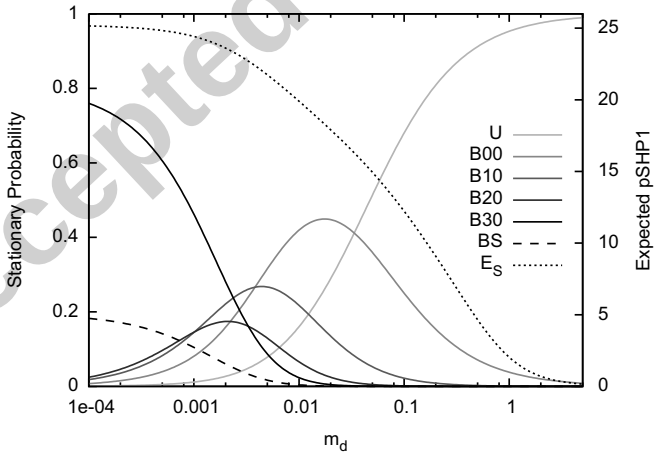

Figure 11: Stationary distribution for $\mathbf{N}$ with $m_d \in [10^{-4}, 5]$. Solid lines are selected high probability kinetic proofreading states. These are states of the form B_z^0, that is, zero non-ZAP-70-protected phosphorylations. The dashed line "BS" is total probability of TCR signalling states. The fine dashed line $\mathbb{E}_S$ is read on right axis.

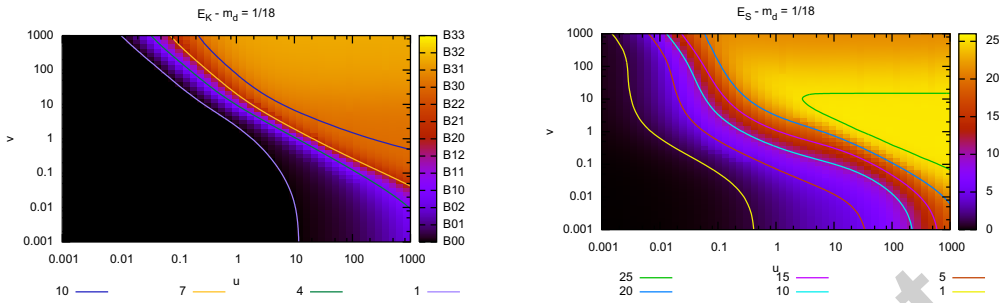

Figure 12: Stationary distributions for $\mathbb{E}_K$ (left) and $\mathbb{E}_S$ (right), varying $u, v \in [10^{-3}, 10^3]$, with $m_d = 1/18$. Contour lines $\mathbb{E}_K$ are $\omega(\mathrm{B}_0^0) = 1$, $\omega(\mathrm{B}_1^0) = 4$, $\omega(\mathrm{B}_2^0) = 7$ and $\omega(\mathrm{B}_3^0) = 10$. Contours for $\mathbb{E}_S$ are selected values of $|\mathrm{S}^1|$. Increases in u and v result in increases in both $\mathbb{E}_K$ and $\mathbb{E}_S$. However, there is a maximum in $\mathbb{E}_S$ in the v direction, which results in a decrease in $\mathbb{E}_S$ at high v.

that increases in v are able to maintain increased $\mathbb{E}_K$ despite also increasing $\mathbb{E}_S$. This also suggests that there is a v which produces maximum negative feedback. This hints towards the qualitative property to T cell tunability outlined in [21]: that a T cell is able to desensitise to a pMHC ligand μ_a whilst remaining sensitive to ligand μ_b with the dissociation rate of μ_a not necessarily faster than that of μ_b. Currently the results do not confirm or deny this behaviour, greater modelling and analysis is required.

7. MAPK Cascade and Negative feedback Destruction

The MAPK cascade may be initiated from a kinetic proofreading signalling state. The ABG model includes a generic adapter protein that connects the TCR to the MAPK cascade. The product of the MAPK cascade, ppERK, is able to bind to the TCR internal chain and protect it from the action of pSHP1. The reactions are as follows:

Reaction 11. *Phosphorylation of Adapter*

$$\mathrm{B}_3^q + \mathrm{A}^0 \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \mathrm{B}_3^q\mathrm{A}^0 \xrightarrow{k_c} \mathrm{B}_3^q + \mathrm{A}^1 \quad q > 0 \tag{31}$$

These reactions does not occur for S^1 bound to B.

Reaction 12. *Desphosphorylation of Adapter*

$$\mathrm{A}^1 + \mathrm{P_A} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \mathrm{A}^1\mathrm{P_A} \xrightarrow{k_c} \mathrm{A}^0 + \mathrm{P_A} \tag{32}$$

Reaction 13. *Phosphorylation of Raf1*

$$\mathrm{R}^0 + \mathrm{A}^1 \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \mathrm{R}^0\mathrm{A}^1 \xrightarrow{k_c} \mathrm{R}^1 + \mathrm{A}^1 \tag{33}$$

Reaction 14. *Dephosphorylation of pRaf1*

$$\mathrm{R}^1 + \mathrm{P_R} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \mathrm{R}^1\mathrm{P_R} \xrightarrow{k_c} \mathrm{R}^0 + \mathrm{P_R} \tag{34}$$

Reaction 15. *Phosphorylation of Mek/pMek*

$$\mathrm{K}^n + \mathrm{R}^1 \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \mathrm{K}^n\mathrm{R}^1 \xrightarrow{k_c} \mathrm{K}^{n+1} + \mathrm{R}^1, \quad n = 0, 1 \tag{35}$$

Reaction 16. *Dephosphorylation of pMek/ppMek*

$$\mathrm{K}^n + \mathrm{P_K} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \mathrm{K}^n\mathrm{P_K} \xrightarrow{k_c} \mathrm{K}^{n-1} + \mathrm{P_K}, \quad n = 1, 2 \tag{36}$$

Reaction 17. *Phosphorylation of ERK/pErk*

$$\mathrm{E}^n + \mathrm{K}^2 \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \mathrm{E}^n\mathrm{K}^2 \xrightarrow{k_c} \mathrm{E}^{n+1} + \mathrm{K}^2, \qquad n = 0, 1 \tag{37}$$

Reaction 18. *Dephosphorylation of pErk / ppERK*

$$\mathrm{E}^n + \mathrm{P_E} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \mathrm{E}^n\mathrm{P_E} \xrightarrow{k_c} \mathrm{E}^{n-1} + \mathrm{P_E}, \qquad n = 1, 2 \tag{38}$$

Reaction 19. *Protection of TCR by ppERK*

$$\mathcal{F} = \{\mathrm{T}, \mathrm{B}\}$$

$$\mathcal{F} + \mathrm{E}^2 \underset{e_d}{\overset{e_a}{\rightleftharpoons}} \mathcal{F}\mathrm{E}^2 \xrightarrow{e_c} \overline{\mathcal{F}} + \mathrm{E}^2 \tag{39}$$

The reaction does not occur for S^1 bound to $\mathcal{F}$

This instantiation of the MAPK cascade follows a familiar form modelled elsewhere [22]. It is necessary to understand its precise behaviour with the choice of parameters here. The $n \times n$ rate matrix that describes the MAPK cascade at $\epsilon = N_T^{-1}$ has $n = 21004075008$ with approximately 438083896320 non-zero entries, which is beyond computational means with the methods used here. Simulation offers a tractable solution. The MAPK cascade is simulated with a volume of restriction $\epsilon = 10^{-2}$; the volume of ≈ 300 TCRs. The molecule quantities are: $|\mathrm{R}| = 1000$, $|\mathrm{K}| = 4000$, $|\mathrm{E}| = 1000$, $|\mathrm{P_A}|, |\mathrm{P_R}|, |\mathrm{P_K}|, |\mathrm{P_E}| = 20$. Simulations are performed for 250 seconds and with initial $|\mathrm{A}^1| \in [1, 9]$ for 2000 repeats. A^1 is converted back to A^0 through the action of $\mathrm{P_A}$, however there is no mechanism included to convert A^0 to A^1. Thus the results give a lower bound to the behaviour of ppERK, as one would expected A^1 to be produced during the operation of the MAPK cascade. The results in figure 13 are given in terms of half-maximal ppERK response ($|\mathrm{E}^2| = 500$). A single A^1 molecule is sufficient for at least half the simulation runs to break the half-maximal E^2 barrier. With $|\mathrm{A}^1| \geq 3$ the $|\mathrm{E}^2| = 500$ barrier is broken on all runs within 25 seconds and remains broken until ≈ 220 seconds. During this period the mean $|\mathrm{E}^2|$ is at a plateau very close to the maximum for all $|\mathrm{A}^1|$ (not shown). With these parameters the MAPK cascade is sensitive, only requiring one or two initial A^1 molecules for long periods of near maximum ppERK signal. The results describe switch like behaviour of ppERK. This is potentially in line with the result of [13] that variation in ppERK concentration has little influence on the ability of the cell to signal. If the variation is small in comparison to the "on" and "off" levels of the ppERK switch then the variation is unlikely influence the designation of the switch. If small quantities of A^1 are sufficient to generate enough E^2 to be recognised as an activation signal, then a single TCR may be sufficient to activate the cell. Further if the levels of E^2 are near maximal (or at least $|\mathrm{E}^2|/|\mathrm{T}| > 1$) as the simulations suggest then it is likely that all TCRs will receive ppERK protection. This has important implications for the spreading of the protection signal, particularly whether endogenous ligands do synergise with agonist ligands and contribute to the activation signal as discussed in [12] and [10]. We will discuss these issues in detail in the concluding comments in section 9.

We note that currently the MAPK cascade serves to guarantee a high ppERK signal if a kinetic proofreading state is reached. We suggest as an extension to the experimental work of [13], an investigation in the variation of MAPK cascade components would be more revealing than an investigation

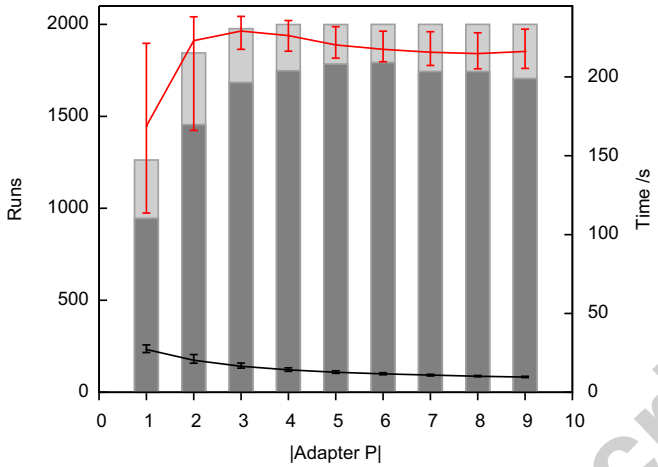

Figure 13: Half-maximal response of ppERK in MAPK cascade simulation results. Bars read from left "Runs" axis: total bar height indicates the runs in which $|E^2| > 500$ occurred; dark grey lower bar indicates the runs in which $|E^2| > 500$ occurred and then fell below 500 within 250 seconds. Lines and error bars are read on right "Time" axis: Black line (bottom) gives the median time of first $|E^2| > 500$ occurrence, the error bars are the lower and upper quartiles; red line (top) gives the median time when $|E^2|$ fell back below 500, the error bars are the quartiles. Medians and quartiles are depicted as they give a better representation of the distribution.

of ERK levels alone.

8. The Entire system

We now simulate all components investigated thus far: kinetic proofreading with negative feedback and protection from the MAPK cascade. We increase the simulation size to $\epsilon = 10^{-1}$, ≈ 3000 TCRs [3].

We specifically test the influence of increasing TCR simulation to TCR activation. The pMHC is partitioned into: M_S a class of strong antigen ($m_d = 0.055$) and M_E a class of weak antigen ($m_d = 1.0$) with $|M| = |M_S| + |M_E|$. $|M_S|$ is varied in $[1, 2000]$ and the results are given in figure 14. One or two strong ligands are as efficacious as thousands at producing at high ppERK

[3] In 1000 simulations with $\epsilon = N_T^{-1}$ with $m_d = 0.055$ no runs achieve a TCR signalling state due to the increased influence of stochastic fluctuations. However we do observe qualitative similarities between $\epsilon = N_T^{-1}$ and $\epsilon = 10^{-1}$ in figures B.21 and 17 respectively.

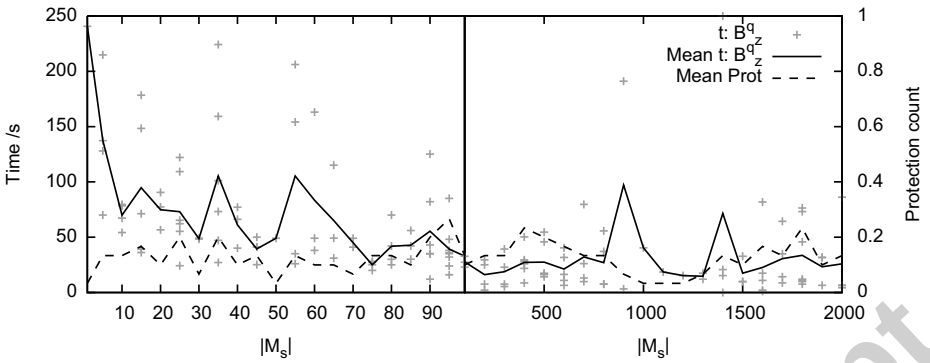

Figure 14: Simulations at $\epsilon = 10^{-1}$ with $|\mathrm{M_S}| \in [1, 2000]$ for 250 seconds and 20 repeats. Grey crosses are the times at which a signalling state is first reached, solid black line gives the mean of these times. Dashed black line is read on the right protection count axis gives the number of times a protection occurred normalised by the number of runs.

signal in $10\% - 20\%$ of all runs. Increasing $\mathrm{M_S}$ results in quicker responses, with the first signalling state being reached in the first 100 seconds.

An understanding of the activation behaviour can be gained by observing unbound pSHP1 over time. Figures 15 and 16 show 50 simulations of 250 seconds for $|\mathrm{M_S}| = 30$ and $|\mathrm{M_S}| = 3000$ respectively. The ppERK generated by a single TCR is sufficient to protect all TCRs from pSHP1 and in all cases the ppERK signal is close to maximal $\approx \epsilon N_\mathrm{E}$. When TCR protection occurs pSHP1 is unable to rebind the TCR resulting in a large spike in unbound pSHP1. For $|\mathrm{M_S}| = 30$, 5 activations occur throughout the first 250 seconds; for $|\mathrm{M_S}| = 300$ only 2 activations occur and they are within the first 50 seconds. The greater number of strong ligands produces a sharper increase in pSHP1. For $t > 50$ pSHP1 levels are so great that it is improbable that even a strong agonist will complete proofreading. Figure 16 suggests that this point occurs when $|\mathrm{S}^1| > |\mathrm{T}|$, there is at least one unbound pSHP1 molecule for every TCR. This is in agreement with section 6 that a quantity of pSHP1 far less than maximal is sufficient to arrest the kinetic proofreading of a strong agonist.

Figure 17 plots all trajectories through kinetic proofreading state of the 50 runs with $|\mathrm{M_S}| = 3000$ whose unbound pSHP1 is given in figure 16. In agreement with small-ϵ figures 10, B.21 there is a greater occurrence in late kinetic proofreading states early in the simulation runs ($t < 50$). We also note that simulations runs are generally confined to zero non-ZAP-70-protected

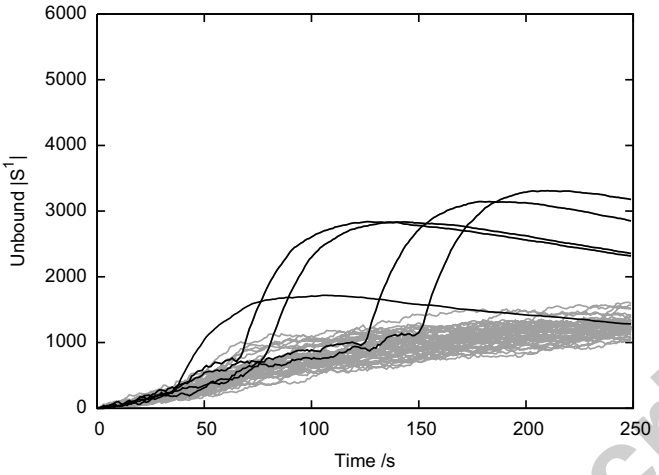

Figure 15: Simulation with $|\mathrm{M_S}| = 30$ for 250 seconds with 50 repeats plotting unbound pSHP1. Grey lines are simulations where no signalling state and no ppERK signal are achieved. Black lines are simulations where a ppERK protection of all TCR is achieved.

phosphorylation states (B_z^0) in agreement with the small-ϵ stationary distributions given in figure 11.

We compare ligands in agonist and antagonist roles. Simulations are performed varying ligand quality $m_d^{-1} \in \{1, 3, 6, 12, 16, 18\}$ for 250 seconds and 50 repeats. Agonist tests are performed with 30 pMHC at the given m_d, the other $|\mathrm{M}| - 30$ pMHC have $m_d = 1$. Antagonist tests are performed with 30 pMHC with $m_d = 1/18$ and then other $|\mathrm{M}| - 30$ pMHC are at given m_d. The results are summarised in table 2.

The antagonist tests show higher number of activations. Low quality ligands generate little pSHP1 (figure 11) allowing good ligands to signal uninhibited. This is unrelated to the synergy of [12], the very low quality ligands are "synergising" by not inhibiting the high quality ligands. As pMHC ligand quality increases levels of pSHP1 increase resulting in antagonism. Activations are seen with further increases in ligand quality, all pMHC are now good agonists and have a higher probability of completing kinetic proofreading (figure 11).

An explanation for poor quality ligands completing kinetic proofreading is due to initial $|\mathrm{S}^1| = 0$. Cells are expected to exhibit pSHP1 levels based

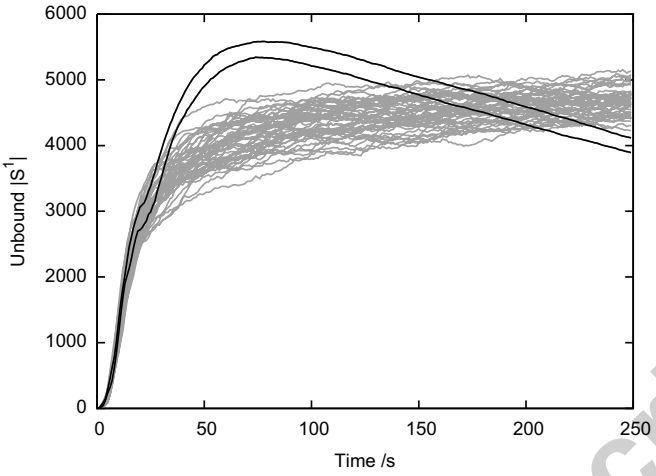

Figure 16: Simulation with $|\mathrm{M_S}| = 3000$ for 250 seconds with 50 repeats plotting unbound pSHP1. Grey lines are simulations where no signalling state and no ppERK signal are achieved. Black lines are simulations where a ppERK protection of all TCR is achieved. Two activations occur in the first 50 seconds, after this time unbound pSHP1 levels are sufficient that it is improbable that a TCR will reach a signalling state. These results are also plotted as a trajectory plot in figure 17.

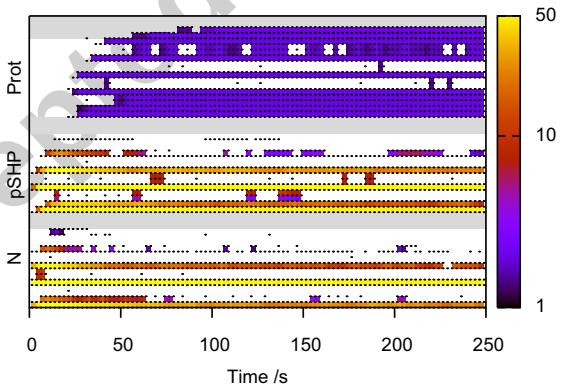

Figure 17: Trajectory plot (see Appendix Appendix G for advice on reading this plot) for the simulations given figure 16, $|\mathrm{M_S}| = 3000$ for 250 seconds with 50 repeats. Simulations are generally confined to the B_z^0 states for $z = 0, 1, 2, 3$. Two activations occur, they can be seen early within the "N" section. Note hump in later kinetic proofreading states at $t < 50$ before the rise of pSHP1 has occurred. This is in agreement with small-ϵ results in figures 10, B.21.

m_d^{-1}	Agonist	Antagonist
1	0	5
3	1	1
6	1	0
12	3	4
16	2	4
18	5	2

Table 2: Agonist and antagonist tests. The second column gives simulation runs in which a high ppERK signal occurs in the agonist tests; the third column gives simulation runs in which a high ppERK signal occurs for the antagonist tests. There is an overlap between agonist and antagonists tests, simulations with 30 pMHC with $m_d = 1/18$ and $|\mathrm{M}| - 30$ pMHC with $m_d = 1$ are performed only once and the results used for agonist and antagonist cases.

on their recent interactions [7] [4] and particularly post-thymic T cells exhibit significantly increase pSHP1 [2]. Simulations are performed varying initial pSHP1 for 1000 seconds with 50 repeats, $|\mathrm{M_S}| = 400$ agonists ligands ($m_d = 0.05$) and $|\mathrm{M_E}| = 2600$ weak pMHC ligands ($m_d = 1.0$). The results are summarised in table 3.

Increasing initial pSHP1 is able to reduce the probability of an activation signal for even strong agonist pMHC. However when observed over long enough timescale (900+ seconds) the initial pSHP1 decays to the point where stochastic fluctuations allow good agonists to produce activations. Figure 18 shows the unbound pSHP1 for initial $|\mathrm{S}^1| = 6000$.

We examine the result given in figure 12, that increases produces a maximum in $\mathbb{E}_S$ but not in $\mathbb{E}_K$. That is, that increasing the CD8 density is able to overcome increases in pSHP1. Figure 18 demonstrates zero activations in fifty with $|\mathrm{S}^1| = 6000$ and $v = 1$. Simulations are performed with $v = 2, 5$ for 1000 seconds with 50 repeats, other simulation parameters are as figure 18. The unbound pSHP1 for $v = 2, 5$ is given in figure 19. Despite initial pSHP1 increases in v is able to induce activations. Agreeing with figure 12 increases in v result in greater pSHP1 but also an increased probability of

[4] The rate at which pSHP1 is converted back to SHP1 is perhaps too great here, pSHP1 decays on the minutes rather than hours timescale, this will only mean that the affect of pSHP1 could be underestimated here.

Initial $\lvert S^1 \rvert$	Activations	Within Time /s
0	7	100
2000	6	100
3000	4	3 in 100, 1 in 900
6000	0	–
12000	1	900
8×10^5	0	–

Table 3: Simulations increasing initial $\lvert S^1 \rvert$. The second column gives simulation runs in which a high ppERK signal occurs; the third column gives an approximate timescale in which all ppERK signals occur. Since we have already performed $\lvert S^1 \rvert = 0$ simulations we do not repeat these and present the result given in figure 14, it should be noted that this result involves simulations of 250 seconds not 1000.

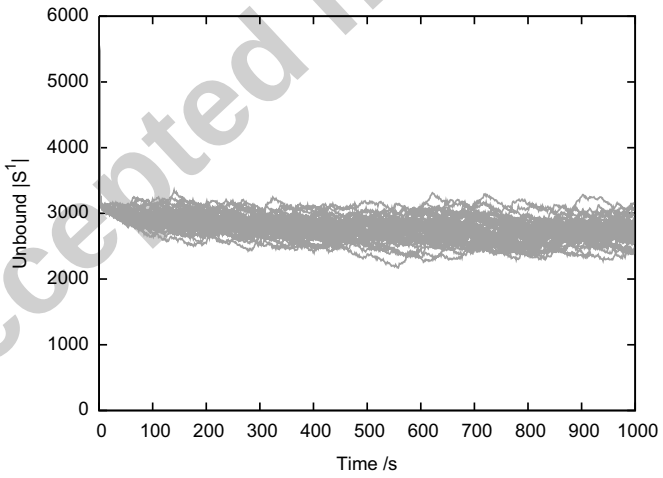

Figure 18: Simulations of $\lvert M_S \rvert = 400$, $\lvert M_E \rvert = 2600$, with initial $\lvert S^1 \rvert = 6000$ for 1000 seconds with 50 repeats. Grey lines are simulation runs in which no ppERK signal occurs. There is an immediate drop in unbound pSHP1 by $\approx \lvert T \rvert$ as it binds to all TCRs. No simulation runs achieve ppERK signalling.

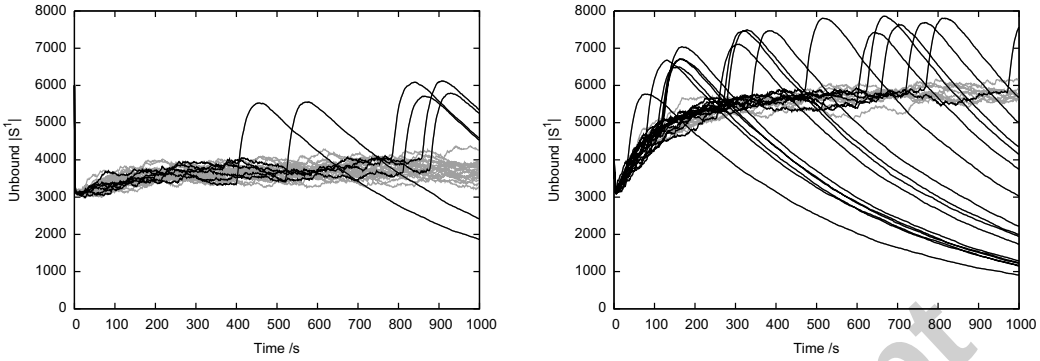

Figure 19: Simulations with $|M_S| = 400$; $|M_E| = 2600$; initial $|S^1| = 6000$; for with $v = 2$ ($|C| = 6000$) (left) and $v = 5$ ($|C| = 15000$) (right); for 1000 seconds with 25 repeats. Grey lines are simulation runs in which no ppERK signal occurs, black lines are simulation runs in which kinetic proofreading is successful. $v = 2$ results in 6 activations; $v = 5$ results in 16 activations, the most reliable activation seen thus far.

reaching later kinetic proofreading states and so increased activations.

9. Discussion

We have analysed a detailed biological model originally presented in [7] of TCR signalling with respect to the behaviour of cytosolic SHP1 and ERK. This signalling system has received much modelling attention in [8–11, 13] who have often focused on the macroscopic or general qualitative behaviour. However rather than make simplifying assumptions (such as those in [9, 11]) to gain understanding we have opted to retain the biological detail to discover how this detail maps to the identified general qualitative cell behaviours. We have achieved this by dissecting, analysing and then re-composing the key components of the model. In contrast to the original presentation of the ABG model [7] we have performed stochastic analysis. We believe the stochastic approach was appropriate given the small molecule numbers required to initiate the MAPK cascade (section 7) and consequently the small numbers of signalling TCRs required to generate a high ppERK activation signal (section 8). Further [9, 11] have clearly highlighted that stochastic effects leading to bistabilities are of importance in TCR signalling models such as the one here. Certainly some reactions considered fulfil the sufficiency conditions (branching, irreversibility and feedback) for a stochastic bistability given in [9]. However, it should be noted that the larger number of reactions in this

model complicates the issue. For example, the switch like behaviour and large molecule numbers of the MAPK cascade provides effective irreversibility. When observing the global cell response to the MAPK the deterministic approach could give reasonable results. However, the stochastic bistability conditions are effectively fulfilled by the reactions that initialise the MAPK cascade (reaction 11). We have branching; the reactions are reversible, but the forward rate is an order of magnitude greater than the backward rate ($N_{\mathcal{P}} \ll N_A$); the feedback in the positive direction arises from many "cascading" steps of the MAPK process (see the well known depiction of reactions in [7, 22]), there is no negative feedback. Consequently small initial molecule numbers of A^1 fulfil the bistability conditions. Further, this stochastic switch is weighted toward the "on" position as there is no negative feedback. The result is that we expect the stochastic MAPK cascade to be far more sensitive (see figure 13 compared in figure S8A in [7]). As such a smaller activation signal from the TCRs is required to produce a cell activation, the activations due to a single TCR in section 8 are confirmation.

Our results are in agreement with the qualitative behaviours discussed in [7, 8, 11] as outlined in section 2. Due to our approach direct results comparisons are difficult as we focus on differing levels of detail to the models in [7, 8, 11]. However, our results are perhaps most readily comparable with those in [7] as the underlying reactions equations are identical. Certainly we agree with the three predictions of the model given in [7]: lengthening of ppERK response time at low ligand densities (see figures 13, 14, 15); hierarchy of antagonism, that is, superior sub-agonist ligands produce greater negative feedback (see figure 11); flexibility in ligand discrimination undergoing differentiation, that is, the ability of a T cell to tune its response (see analysis involving u, v and figures 18, 19). We have refrained from stochastically running the numerical experiments in [7] and comparing the quantitative results. This task has been performed on a simpler model in [11]. We expect any quantitative differences to arise where the bistability conditions apply, as discussed in the previous paragraph this includes the sensitivity of the MAPK cascade and by a similar argument the strength of the pSHP1 negative feedback.

We now summarise the key findings of this paper. In section 4 we demonstrate that rate choices should be made in unison with model structure choices. This can reveal equivalently behaved simpler models and most importantly this provides an opportunity to validate against biology. What remains is to return to biology to discover the occurrence of the CD8-TCR

complex and adjust the model accordingly, a solution could a more complex model of CD8, TCR, pMHC binding similar to that in [21]. Section 5 investigates the ABG formulation of kinetic proofreading and the expected step-like discrimination is observed. The phosphorylation protection of ZAP-70 is shown to be of importance by aiding pMHC and CD8 density parameters u and v to sensitise the cell. This is due to the relative stability of the ZAP-70 association which allows ZAP-70 molecules to remain bound between TCR pMHC associations. Thus the kinetic proofreading state may be preserved between TCR pMHC binding events. We demonstrate that decreases in d_z can cause retention of kinetic proofreading state despite decreases in CD8 density v. The influence of this locking mechanism is largely undiscussed in [7].

The analysis of the negative feedback behaviour in section 6 demonstrates a transient hump in expected kinetic proofreading state before a rise in expected pSHP1 level. This is due to a non-zero kinetic proofreading state which much be reached before the negative feedback can be generated, we determined this the *base negative feedback state*. Examining the response to changes in pMHC and CD8 density parameters we show that u and v are able increase expected proofreading state and pSHP1 levels. However there is a maximum in $\mathbb{E}_S$ in the v direction, as discussed in section 6, this is an appropriate step towards the necessary condition for full tunability given in [21].

In section 7 the MAPK cascade is shown to act as an ultra-sensitive amplification switch, amplifying small numbers of A^1 molecules to thousands of ppERK molecules. The ppERK signal is near maximal and remains so for long periods of time and with sustained A^1 production ppERK would remain near maximal permanently. As discussed in section 7 this raises issues for the determination of the activation signal and the spreading of the protection signal. First, if a near maximal ppERK whose duration is > 200 seconds is sufficient for a cell to recognise an activation then a single TCR may be sufficient to activate the cell. Second, since numbers of near maximal ppERK are an order magnitude greater than the number of TCRs we expect the total protection of all TCRs (as is the case in the activations in section 8) [5].

[5]The TCR protection state is absorbing in this model and so all TCRs will receive protection with probability 1 over a suitable timescale. However protection is expected to persist on a timescale longer than the signalling events considered here, if a loss of protection reaction were included we would expect observe very high levels of TCR protection,

This has implications for the hypothesis presented in [12] that suggests that protected endogenous ligands may contribute towards signalling. The issue is that these ligands must still overcome kinetic proofreading. As discussed in section 2 consideration of the structure of the ppERK reactions reveal that it is not involved in a positive feedback and so ppERK protection does not aid the ability of any ligand to overcome kinetic proofreading. We discuss a classification of ligands based on the signalling induced:

- *Class I:* Are able to complete kinetic proofreading despite negative feedback and generate an activation signal. Effectively agonists.
- *Class II:* Are unable to complete kinetic proofreading in the presence of negative feedback [6], but are able to complete kinetic proofreading if protected from negative feedback. Effectively antagonists.
- *Class III:* Are unable to complete kinetic proofreading even when protected. They are strong enough achieve a kinetic proofreading state larger than the base negative feedback state, that is, they are able to generate negative feedback. Effectively antagonists/partial agonists.
- *Class IV:* Are unable to complete kinetic proofreading when protected. They are also unable to break the base negative feedback state. Effectively partial agonists/null ligands.

We have not used the agonist, antagonist etc. labels because the definitions of the classes differ from those given in section 2. Within these classes only *I & II* may "synergise" with the signalling agonists, *III & IV* will never complete kinetic proofreading. Generally we expect endogenous ligands to be members of classes *II–IV*. However a common hypothesis is that endogenous ligands are most often not antagonists [7, 12] and non-antagonist/non-agonist ligands are members of classes *III & IV*. Thus the majority of endogenous ligands will not be able to synergise due to their inability to overcome proofreading. The expected hitting time of protected kinetic proofreading signalling state

but not total

[6] Within the stochastic methods and the biological mechanisms of the models considered here all ligands will be able to complete kinetic proofreading, with or without negative feedback, if considered over a long enough time period. When we discuss inability to complete kinetic proofreading we mean that the probability a ligand will complete kinetic proofreading in a time period that is relevant to the cell is effectively nil.

(figure 7) shows a sharp increase with increases in m_d. This suggests that whilst members of class *II* can complete kinetic proofreading on a suitable timescale they may take substantially longer than members of class *I* and consequently generate substantially less A^1. It is possible that a handful of class *I* ligands can generate far more A^1 than a population of class *II* ligands. We conclude the discussions of the behaviour of the MAPK cascade with two statements. First, if a single TCR is able to generate a near maximal ppERK signal then synergising ligands can only contribute be ensuring the ppERK signal remains near maximal over a longer period of time than the original ligand could achieve. Second, if the TCRs receive total protection this is not necessarily an indicator that all TCRs require protection, simply that this guarantees that TCRs of interest are protected.

Section 8 re-composes the entire model to perform stochastic simulations on a system of many TCRs. The simulations demonstrate that a few agonists can be as efficacious as a population of thousands in producing activations. Moreover, we find that the large-ϵ results are in qualitative agreement with the small-ϵ single TCR predictions. It is elucidating to observe the dynamics of unbound pSHP1. The pSHP1 generated is dependant on TCR stimulation, which is dependant on the quantity of pMHC and distribution of dissociation rates. As a consequence in simulations with zero initial pSHP1 consisting solely of weak ligands (normally considered to be members of class *II*) little pSHP1 is generated and the weak ligands can induce activations. Increased TCR stimulation results in a sharper rise of pSHP1, the statistics of the sharp rise are related to the position of the base negative feedback state. Once the rise has occurred it becomes very unlikely that any TCR, even a good agonist, will complete kinetic proofreading. Performing simulations with ligands in agonist and antagonist roles we observe the expected agonist and antagonist behaviour. With increases initial pSHP1 we are able to prevent weak ligands from inducing activations, with further increases we are able to prevent agonist pMHC from inducing any activations. Finally we validate the small-ϵ v prediction that increased CD8 density is able to increase the expected kinetic proofreading state despite increasing the expected pSHP1.

Combining the results of the paper we discuss a hypothesis for reliable T cell activation with respect to the base negative feedback state. If the T cell is presented a population comprising $0.01 - 0.1\%$ agonists and $99.9 - 99.99\%$ endogenous ligands and if the expected time for agonists to complete kinetic proofreading is less than the expected time for endogenous ligands to reach the base negative feedback state, then the T cell will be reliably activated.

This hypothesis is depicted in figure 20 along with validating results from the model. The results state that agonist ligands with $m_d = \mu_a$ are expected to reach a kinetic proofreading signalling state before endogenous ligands with $m_d = \mu_e$ reach the base negative feedback state. The expected times given in figure 20 are calculated from the kinetic proofreading without negative feedback rate matrix $\mathbf{K}$ described in section 5. As justification for calculating without negative feedback we note that during the period in which agonist ligands generate negative feedback the endogenous ligands do not generate pSHP1. Since agonist ligands are in the vast minority and appealing to the assumption that pSHP1 is not diffusion-limited then the population of endogenous ligands "absorb" the pSHP1 generated by the agonist ligands. Further, we are assuming that the endogenous ligand comprise a sufficiently large proportion of all pMHC that the influence of agonist pSHP1 does not significantly alter their time to base negative feedback state given in figure 20.

This hypothesis requires more detailed modelling and analysis to confirm, in reality the negative feedback generated by the agonists will dampen their own proofreading as well that of the endogenous ligands. However we believe the arguments given here certainly demonstrate the point.

This hypothesis suggests that the reason why the T cell can react to only 0.01–0.1% of the total pMHC expressed [2] is because this is a very effective way of producing T cell activation. Altan-Bonnet and Germain [7] suggest that the agonist ligands can quickly overcome the negative feedback due to its "limited nature". The results in this paper suggest that the negative feedback is not limited in nature and is capable of arresting the proofreading of any realistic ligand. We wish to clarify their argument by adding that agonist ligands can induce activations by their ability to complete proofreading before weaker ligands can generate negative feedback.

We note that the hypothesis does not prevent a population of agonists inducing an activation for two reasons. First, the larger population of agonists have a higher probability of sampling a proofreading time substantially faster than the mean. Second, the negative feedback parameters currently allow agonist ligands to complete kinetic proofreading despite high pSHP1 levels over an increased time period (see table 3).

Finally we suggest biological experimentation which could test the results here. If APCs can be prepared which present endogenous ligands in the majority and agonist ligands in the minority; and if the concentration of agonist ligands can be incrementally increased then the statistics of the rise time of pSHP1 and ppERK could confirm the details of the hypothesis in

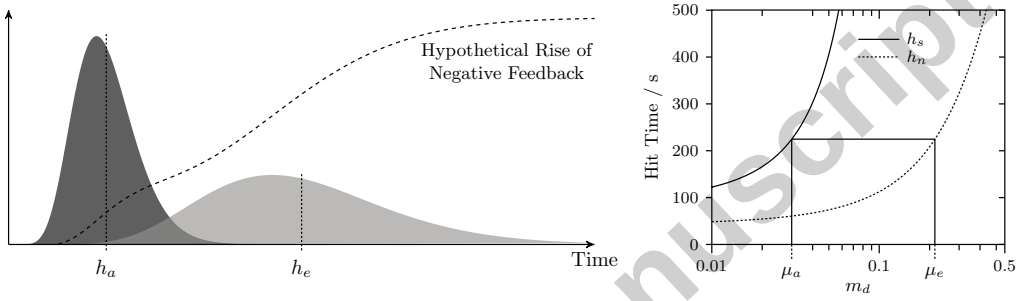

Figure 20: *Left*: Depiction of reliable activation hypothesis. The hypothetical distribution of agonist ($m_d = \mu_a$) ligands reaching signalling state is given on the left (dark grey), h_a is the expected hit time of signalling state. Hypothetical distribution of endogenous ligands ($m_d = \mu_e$) reaching base negative feedback state is given on left (light grey), h_e is expected hit time of the base state. The distributions are purely illustrative and are sum of exponential distributions, specifically Erlang Distributions (Gamma distribution with integer shape parameter) and as such the mean is greater than the mode. Dotted line gives the hypothetical rise of negative feedback, again illustrative and calculated as a function of cumulative hit time distributions, expected agonist/endogenous pSHP1 levels and assumed agonist/endogenous populations ratio. *Right:* Validation of hypothesis within the model. Solid black line (h_s) is expected hit time of signalling state for agonist ligands ($m_d = \mu_a$) (taken from figure 7), dotted line (h_n) is the expected hit time of base negative feedback state for endogenous ligands ($m_d = \mu_e$). All ligands with $m_d \leq \mu_a$ are expected to reach signalling state before ligands with $m_d > \mu_e$ reach base negative feedback state.

figure 20. Further, our analysis suggests that the pSHP1 levels needed to dampen a strong agonists are an order of magnitude smaller than the total SHP1. Consequently a variation in pSHP1 less than an order of magnitude will not necessarily inhibit a cell in producing the required pSHP1 to dampen strong agonists. However a reduction in SHP1 levels will reduce the rate at which pSHP1 can be produced and so increase the time taken to reach equilibrium. This should extend the time window in which activations can reliably occur and so cell activations will be observed over a longer time period. One could also look for the dynamics of unbound pSHP1 which would suggest its relationship to the protected TCR and particularly the spreading of the protection signal. We add that any investigations into the time taken for kinetic proofreading of varying strength ligands, particularly looking for the existence and position of a *base negative feedback state* would prove very useful in validating the hypothesis presented in figure 20.

Appendix A. Components and Rates

All components and their labels are given in table A.4, all rates are given in table A.5.

Appendix B. Justification for ϵ

In sections 4, 5 and 6 we set $\epsilon = N_T^{-1}$ with N_T the number of TCRs in the ABG model. This restricts the volume of interest to that of a single TCR and allows us to analyse a TCR in isolation. We believe this to be invaluable in understanding the reactions incorporated in the model. Certainly this will increase the influence of stochastic noise in the system. There may be quantitative discrepancies in small-ϵ results and non-obvious effects when the model is scales. However there is qualitative agreement between small and large ϵ (see figures 17 and B.21). Small-ϵ calculations are made with solely rate matrices and unlike simulation do not require many runs to view the entire distribution of behaviour. We expect the rate matrix $\mathbf{N}$ to be most susceptible to noise as it analyses varying cytosolic molecule numbers. An increase of a single pSHP1 molecule at the single TCR level corresponds to a large increase in pSHP1 in the entire cell. However we discovered that the results matched well with those obtained with a real pSHP1 concentration parameter defined similarly to v defined in section 4 (not shown).

Component	Label	Comments
TCR	T_z	T cell Receptor complex with z ZAP-70 molecules bound.
pMHC	M	Antigenic peptide bound to major-histocompatability complex MHC
CD8	C	Co-receptor CD8
Lck	L	Leukocyte-specific protein tyrosine kinase (Lck)
SHP1	S	SH2 domain containing tyrosine phosphatase (SHP-1). Phosphorylated pSHP-1 represented as S^1
ERK	E	Extracellular signal-regulated kinase (ERK). Single and double phosphorylated versions represented as E^1 and E^2 respectively.
ZAP-70	Z	ζ-chain associated protein kinase 70 (ZAP-70)
Adapter	A	Adapter protein that initiates the MAPK cascade. Phosphorylated form represented as A^1.
Raf1	R	Part of MAPK cascade. Phosphorylated form represented as R^1.
Mek	K	Part of the MAPK cascade. Single and double phosphorylated forms represented as K^1 and K^2 respectively
MAPKPase	P_E	MAPK Phosphatase, acts dephosphorylating ERK
MAPKKPase	P_K	MAPK Phosphatase, acts dephosphorylating Mek
MAPKKKPase	P_R	MAPK Phosphatase, acts dephosphorylating Raf1
Adapter Pase	P_A	Adapter Phosphatase, acts desphosphorylating Adapter
SHP Pase	P_S	pSHP-1 Phosphatase, acts desphosphorylating S^1
TCR-pMHC	B_z^q	TCR-pMHC complex with q non-protected phosphorylations and z ZAP-70 molecules bound.
Protection	$\overline{C_z^q}$, $\overline{T_z}$	Internal TCR complex after protection by ppERK, with q non-protected phosphorylations of the ζ-chain

Table A.4: Components in the model and their labels.

Rate	Value	Comments
m_a	1×10^4	Association rate of pMHC to TCR.
m_d	–	Dissociation rate of pMHC from TCR.
c_t	1000	Association rate of CD8 to TCR.
c_a	3×10^5	Association rate of CD8 to pMHC, TCR-pMHC complex.
c_d	100	Dissociation rate of CD8 from TCR, pMHC TCR-pMHC complexes.
c_s	0.1	Stabilising effect of CD8 to TCR-pMHC complex, multiplies dissociation rate m_d.
l_a	1×10^4	Association rate of Lck to internal TCR complex.
l_d	50	Dissociation rate of Lck from internal TCR complex.
l_c	10.4	Rate of Lck phosphorylation of the internal TCR complex.
l_{ca}	50.0	Association rate of CD8 associated Lck to ITAMs, ZAP-70 etc.
l_{cd}	40.0	Dissociation rate of CD8 associated Lck from ITAMs, ZAP-70 etc.
l_{cc}	10.4	Rate of CD8 associated Lck phosphorylation of the internal TCR complex.
z_a	1.2×10^7	ZAP-70 association to phosphorylated ITAMs.
z_d	0.11	ZAP-70 dissociation from phosphorylated ITAMs.
s_a	2.9×10^5	Association rate of SHP-1 to TCR complex.
s_d	0.13	Dissociation rate of SHP-1 from TCR complex.
s_c	35.0	Rate of phosphorylation of SHP-1.
p_a	3.2×10^5	Association rate of pSHP-1 to TCR complex.
p_d	0.05	Dissociation of pSHP-1 from complex.
a_a	5.0×10^6	Association rate of adapter to TCR complex.
e_a	1.0×10^7	Association rate of ppERK to TCR complex.
e_d	2.0	Dissociation rate of ppERK from TCR complex.
e_c	3.4	Rate of protection of TCR by ppERK.
k_a	1.2×10^7	Association rate of any component of MAPK cascade.
k_d	0.15	Dissociation rate of any component of MAPK cascade.
k_a	2.6	Rate of (de)phosphorylation of any component of MAPK cascade.
q_a	1.0×10^4	Association rate of pSHP1 phosphatase.
q_d	0.15	Dissociation rate of pSHP1 phosphatase.
q_c	2.6	Rate of (de)phosphorylation of pSHP1 phosphatase.

Table A.5: Macroscopic rates from the ABC model in [7]. All association rates are given in $\text{mol}^{-1}s^{-1}$ all other rates have units s^{-1} with the exception of c_s which is dimensionless.

Component	Quantity	Comment
V	15×10^{-15}	Cytoplasmic volume of a T cell
N_{T}	3×10^4	Quantity of TCR
N_{M}	3×10^4	Quantity of pMHC
N_{C}	3×10^4	Quantity of CD8
N_{L}	3×10^4	Quantity of Lck
N_{S}	8×10^5	Quantity of SHP1
N_{E}	10^5	Quantity of ERK
N_{Z}	1.2×10^6	Quantity of ZAP-70
N_{A}	1.5×10^5	Quantity of Adapter
N_{R}	10^5	Quantity of Raf1
N_{K}	4×10^5	Quantity of Mek
$N_{\mathcal{P}}$	2000	$\mathcal{P} = \{\mathrm{E}, \mathrm{K}, \mathrm{R}, \mathrm{A}, \mathrm{S}\}$ Quantity of all dephosphorylating phosphatases

Table A.6: Quantity of all components found in the cytoplasmic volume V, taken from [7]

In section 7 we use $\epsilon = 10^{-2}$ for simulations of the MAPK cascade and in section 8 $\epsilon = 10^{-1}$ for simulations of the entire system. We found the molecule numbers to be sufficient that stochastic fluctuations do not dominate and results to be in good agreement with other large ϵ choices, including $\epsilon = 1$.

Unlike the models in [9–11] we model all enzymatic reactions by a two stage process which should reduce the influence of stochastic fluctuations.

Appendix C. General Binding System

The binding system may be generalised and described by the following matrix:

$$\mathbf{G} = \begin{pmatrix} -\sigma_1 & k_1 & k_3 & k_5 & 0 \\ k_2 & -\sigma_2 & 0 & 0 & k_7 \\ k_4 & 0 & -\sigma_3 & 0 & k_9 \\ k_6 & 0 & 0 & -\sigma_4 & k_{11} \\ 0 & k_8 & k_{10} & k_{12} & -\sigma_5 \end{pmatrix} \tag{C.1}$$

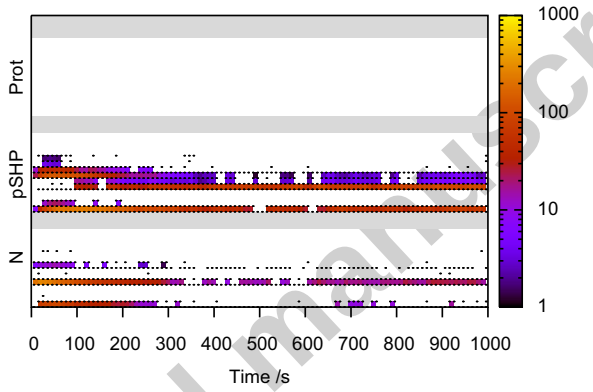

Figure B.21: Trajectories of 1000 runs of a simulation resticted to a single TCR and $m_d = 0.054$ for 1000 seconds. See appendix Appendix G for instructions on how to read this figure. No simulations reach the grey signalling region. Two orange strips in the pSHP1 bound state section indicate that the majority of simulations are in unbound state T_0S^1 or in bound state $B_0^0S^1$. Within the first 100 seconds states B_z with $z > 0$ are reached, this is during the rise time of pSHP1 maximum before the maximum is reached (see figure 11). Once the maximum pSHP1 level is reached the TCR is rarely found in a state with $z > 0$. Note the qualitative agreement with figure 17 in terms of the early hump in kinetic proofreading states and then the long term confinement to B_z^0 states.

With $\sigma_i = \sum_j g_{ij}$. The rate matrix for choice of rates here:

$$\mathbf{H} = \begin{pmatrix} -s_1 & m_a & c_t & c_a & 0 \\ m_d & -s_2 & 0 & 0 & c_a \\ c_d & 0 & -s_3 & 0 & c_a \\ c_d & 0 & 0 & -s_4 & c_a \\ 0 & c_d & m_d c_s & m_d c_s & -s_5 \end{pmatrix} \tag{C.2}$$

With $s_i = \sum_j h_{ij}$. The solution to $\pi\mathbf{H} = \mathbf{0}$ subject to $\sum_i \pi_i = 1$ is given in equation C.3.

$$\pi = \frac{\varphi}{\sum_i \varphi} \qquad \varphi = \begin{pmatrix} m_d c_d (c_d + c_a)(c_a(2c_s + 1) + 2m_d c_s + c_d) \\ c_d (c_d + c_a)(c_a^2 + c_a(m_a + c_t) + 2m_a m_d c_s + m_a c_d) \\ m_d(c_t c_d (c_d + c_a) + c_s(c_a^3 + c_a^2(m_d + c_t + m_a) + c_a(m_d c_t + m_a c_d + 2c_d c_t) + 2c_d c_t)) \\ m_d(c_a^2 c_d + c_a c_d^2) + c_s(c_a^3 m_d + c_a^2(m_d^2 + m_d(2c_d + c_t + m_a)) + c_a(m_d^2(2c_d + c_t) + m_a c_d m_d)) \\ c_a(c_d + c_a)(c_a^2 + (m_d + c_t + m_a)c_a + m_a c_d + m_d c_t) \end{pmatrix} \tag{C.3}$$

The rate matrix for the binding system with TC removed:

$$\mathbf{W} = \begin{pmatrix} -w_1 & m_a & c_a & 0 \\ m_d & -w_2 & 0 & c_a \\ c_d & 0 & -w_3 & c_a \\ 0 & c_d & m_d c_s & -w_4 \end{pmatrix} \tag{C.4}$$

With $w_i = \sum_j w_{ij}$. Solving $\pi^{\mathrm{TC}}\mathbf{W} = \mathbf{0}$ subject to $\sum_i \pi_i^{\mathrm{TC}} = 1$ gives:

$$\pi^{\mathbf{TC}} = \frac{\rho}{\sum_i \rho_i} \qquad \rho = \begin{pmatrix} m_d c_d (c_d + c_a(1 + c_s) + m_d c_s) \\ c_d(c_a^2 + c_a m_a + m_a m_d c_s + c_d m_a) \\ m_d c_a (c_d + c_s(c_a + m_a + m_d)) \\ c_a^3 + c_a^2(m_a + m_d) + c_a c_d m_a \end{pmatrix} \tag{C.5}$$

Appendix D. Negative Feedback

The action of pSHP1 on the TCR internal complex with 3 ZAP-70 molecules is a one step process, see reaction 8 (as reproduced from [7]). The action of pSHP1 on the TCR internal chains containing 0, 1, or 2 ZAP-70 molecules is a many step process. To clarify, if N and P are non-phosphorylated and phosphorylated complexes respectively, L is Lck and S is pSHP1, in the standard

case:

$$\begin{aligned}
\mathrm{N} + \mathrm{L} + \mathrm{S} &\underset{l_d}{\overset{l_a}{\rightleftharpoons}} \mathrm{NL} + \mathrm{S} \\
\mathrm{NL} + \mathrm{S} &\xrightarrow{l_c} \mathrm{P} + \mathrm{L} + \mathrm{S} \\
\mathrm{P} + \mathrm{L} + \mathrm{S} &\underset{p_d}{\overset{p_a}{\rightleftharpoons}} \mathrm{PS} + \mathrm{L} \\
\mathrm{PS} + \mathrm{L} &\underset{l_d}{\overset{l_a}{\rightleftharpoons}} \mathrm{PSL} \\
\mathrm{PSL} &\xrightarrow{l_c} \mathrm{NS} + \mathrm{L} \\
\mathrm{NS} + \mathrm{L} &\xrightarrow{p_d} \mathrm{N} + \mathrm{L} + \mathrm{S}
\end{aligned} \tag{D.1}$$

In the B_3 case:

$$\begin{aligned}
\mathrm{N} + \mathrm{L} + \mathrm{S} &\underset{l_d}{\overset{l_a}{\rightleftharpoons}} \mathrm{NL} + \mathrm{S} \\
\mathrm{NL} + \mathrm{S} &\xrightarrow{l_c} \mathrm{P} + \mathrm{L} + \mathrm{S} \\
\mathrm{P} + \mathrm{L} + \mathrm{S} &\xrightarrow{p_a} \mathrm{N} + \mathrm{S} + \mathrm{L} \\
\mathrm{N} + \mathrm{L} + \mathrm{S} &\underset{p_d}{\overset{p_a}{\rightleftharpoons}} \mathrm{NS} + \mathrm{L}
\end{aligned} \tag{D.2}$$

Rate matrix $\mathbf{R_S}$ describes the standard case:

$$\mathbf{R_S} = \begin{pmatrix}
-l_a - p_a|\mathrm{S}| & l_a & 0 & 0 & 0 & p_a|\mathrm{S}| \\
l_d & -l_d - l_c & l_c & 0 & 0 & 0 \\
0 & 0 & -p_a|\mathrm{S}| & p_a|\mathrm{S}| & 0 & 0 \\
0 & 0 & p_d & -p_d - l_a & l_a & 0 \\
0 & 0 & 0 & l_d & -l_d - l_c & l_c \\
p_d & 0 & 0 & 0 & 0 & -p_d
\end{pmatrix} \tag{D.3}$$

The stationary distribution given by $\pi_{\mathbf{R_S}}\mathbf{R_S} = 0$:

$$\pi_{\mathbf{R_S}} = \frac{\phi}{\sum \phi_i} \qquad \phi = \begin{pmatrix}
p_d p_a|\mathrm{S}|(l_c + l_d) \\
l_a p_d p_a|\mathrm{S}| \\
p_d^2(l_c + l_d) + p_d l_a l_c \\
p_d p_a|\mathrm{S}|(l_c + l_d) \\
l_a p_d p_a|\mathrm{S}| \\
(p_a|\mathrm{S}|)^2(l_c + l_d) + p_a|\mathrm{S}| l_a l_c
\end{pmatrix} \tag{D.4}$$

The rate matrix $\mathbf{R_3}$ describes the B_3 case:

$$\mathbf{R_3} = \begin{pmatrix} -l_a - p_a|\mathrm{S}| & l_a & 0 & p_a|\mathrm{S}| \\ l_d & -l_d - l_c & l_c & 0 \\ p_a|\mathrm{S}| & 0 & -p_a|\mathrm{S}| & 0 \\ p_d & 0 & 0 & -p_d \end{pmatrix} \tag{D.5}$$

Whose stationary distribution:

$$\pi_{\mathbf{R_3}} = \frac{\theta}{\sum \theta_i} \qquad \theta = \begin{pmatrix} p_d p_a|\mathrm{S}|(l_c + l_d) \\ l_a p_d p_a|\mathrm{S}| \\ l_a l_c p_d \\ (p_a|\mathrm{S}|)^2(l_c + l_d) \end{pmatrix} \tag{D.6}$$

The stationary distributions exhibit differences dependant on $|\mathrm{S}|$. For example the difference in probability of the phosphorylated state Δ is:

$$\Delta = \frac{l_a l_c p_a|\mathrm{S}|}{p_d^2(l_c + l_d) + l_a l_c(p_d + l_d) + p_a|\mathrm{S}|(p_d(l_a + l_c + l_d) + l_a l_c)} \tag{D.7}$$

$$\approx \frac{10^{-10}|\mathrm{S}|}{40 + |\mathrm{S}|} \tag{D.8}$$

The approximation is taken for CD8-associated Lck phosphorylation rates and holds for $N_T^{-1} \leq \epsilon \leq 1$. The stationary differences between the two are negligible, numerical solutions of the transient behaviour reveal larger differences however the behaviour is qualitatively identical.

A detailed look at the pSHP1 reaction equations reveals that doubly-phosphorylated ITAMs are exempt from the SHP1 negative signal. This is a consequence of the precondition for Lck binding: $q < 2$ when $z = 0, 1, 2$ and $q < 3$ when $z = 3$. We compare the negative feedback as found in the ABG model in [7] and a uniform application of negative feedback with multi-step desphosphorylation and no state exempt from the action of pSHP1, we determine this *uniform negative feedback*. Calculating the transient distributions reveals the two methods are qualitatively similar, but uniform negative feedback (despite applying to a larger number of kinetic proofreading states) gives a higher stationary probability of expected kinetic proofreading states. Further we find that the for ABG negative feedback the ordering of most probable m_d is preserved in the hump but is broken as the transient moves towards stationarity. We find the multi-step form with

uniform action on all kinetic proofreading states preferable and so all results in this paper use uniform negative feedback.

Appendix E. Modelling Methods – Continuous Time Markov Chains

Definition 1. A continuous time Markov chain is a continuous time stochastic process $\{\mathbf{X}(t) : t \geq 0\}$ with state space $\mathcal{S}$ that satisfies the Markov property:

$$\begin{aligned}&\mathbb{P}(\mathbf{X}(t) = j|\mathbf{X}(s) = i, \mathbf{X}(t_{n-1}) = i_{n-1}, \ldots, \mathbf{X}(t_0) = i_0)\\ =&\mathbb{P}(\mathbf{X}(t) = j|\mathbf{X}(s) = i)\end{aligned} \tag{E.1}$$

Where $0 \leq t_0 \leq t_1 \leq \cdots \leq t_{n-1} \leq s \leq t$, and $i_0, i_1, \ldots, i_{n-1}, i, j, \in, \mathcal{S}$.

Three key properties of a CTMC are of interest: the time evolution of the chain starting from a particular state (transient analysis), the long term behaviour of the chain (its stationary distribution), and the expected time to reach of a set of states $\mathcal{C} \subseteq \mathcal{S}$ from a state h (the hitting time). Denote:

$$p_{ij}(t) = \mathbb{P}(\mathbf{X}(t) = j|\mathbf{X}(s) = i) \qquad \text{for } 0 \leq s, t \text{ and } i, j \in \mathcal{S} \tag{E.2}$$

$\mathbf{P}(t)$ is the matrix with entries $p_{ij}(t)$. The time evolution of $\mathbf{P}(t)$ is described by the following differential equation:

$$\dot{\mathbf{P}}(t) = \mathbf{G}\mathbf{P}(t) \tag{E.3}$$

$\mathbf{G}$ is the *generator* matrix or *rate* matrix of the chain and its entries are:

$$g_{ij} = \begin{cases} k_{ij} & \text{for } i \neq j, \\ -\sum_q k_{iq} & \text{for } i = j \end{cases} \tag{E.4}$$

k_{ij} describes the exponential rate of transitioning from state i to j, and k_{ii} can be thought of as the rate of leaving state i so:

$$\sum_j g_{ij} = 0 \qquad \forall i \tag{E.5}$$

The general solution to equation E.3 is given by the matrix exponential:

$$\mathbf{P}(t) = \mathrm{e}^{t\mathbf{G}} \equiv \sum_{n=0}^{\infty} \frac{(t\mathbf{G})^n}{n!} \tag{E.6}$$

Generally we calculate the matrix exponential numerically, for small dense $\mathbf{G}$ we use MATLAB's `expm(G)` function which uses a scaling and squaring algorithm with a Padé approximation as detailed in [23]. For large sparse $\mathbf{G}$ we use the methods provided in EXPOKIT [24] which uses Krylov subspace projection techniques to calculate $\mathbf{p_0}e^{t\mathbf{G}}$ directly (i.e. $e^{t\mathbf{G}}$ is not explicitly calculated) for an initial distribution $\mathbf{p_0}$.

The stationary or invariant distribution $\pi = (\pi_1, \pi_2, \ldots \pi_{|\mathcal{S}|})$ exists for a CTMC if it is irreducible and recurrent [25], which will be the case for the all the CTMCs considered in this paper. If the stationary distribution exists then it may be found by solving the following:

$$\pi\mathbf{G} = \mathbf{0} \qquad \text{with} \sum_i \pi_i = 1 \tag{E.7}$$

A numerical solution to the stationary distribution equation can often be difficult to find [26], particularly when $\mathbf{G}$ is large and sparse as is the case here. The inverse iteration method [26] is used, rather than solve equation E.7 directly solve and scale:

$$\mathbf{G}^T x_1 = x_0 \qquad \pi^T = \frac{x_1}{\mathbf{1}^T_{|\mathbf{S}|} x_1} \tag{E.8}$$

x_0 is the the column vector length $|\mathcal{S}|$ of all zeros with 1 in the $|\mathcal{S}|$ position. $\mathbf{1}_n$ is the length n column vector of ones.

The vector of expected first hit times $h^{\mathcal{C}} = (h^{\mathcal{C}}_i \mid i \in \mathcal{S})$ of states $\mathcal{C} \subseteq \mathcal{S}$, where $h^{\mathcal{C}}_i$ is the expected time to reach a $\mathcal{C}$ starting from state i can be found by solving the following equations [25]:

$$\begin{aligned} h^{\mathcal{C}}_i &= 0 && \text{for } i \in \mathcal{C} \\ -\sum_j k_{ij} h^{\mathcal{C}}_k &= 0 && \text{for } i \notin \mathcal{C} \end{aligned} \tag{E.9}$$

Appendix F. Reaction Equation Notation

To describe reaction equations more succinctly we use set notation to group together common reactions. A reaction applied to a set represents application of the reaction to every member of the set:

$$\{\text{A, B}\} + \text{C} \xrightarrow{r} \{\text{A, B}\}\text{C} \tag{F.1}$$

Represents reactions:

$$\mathrm{A} + \mathrm{B} \xrightarrow{r} \mathrm{AB}\ , \quad \mathrm{A} + \mathrm{C} \xrightarrow{r} \mathrm{AC} \tag{F.2}$$

Concatenation of terms represents a complex containing the elements of the concatenation. Generally any set manipulation can be used to reduce the reaction description:

$$\mathcal{S} = (\{\mathrm{A}\} \times \{\emptyset,\ \mathrm{B},\ \mathrm{C}\} \times \{\mathrm{D}\}) \cup \{\mathrm{GH},\ \mathrm{GJ}\} \tag{F.3}$$

Represents:

$$\mathcal{S} = \{\mathrm{AD},\ \mathrm{ABD},\ \mathrm{ACD},\ \mathrm{GH},\ \mathrm{GJ}\} \tag{F.4}$$

Then reactions may be described using $\mathcal{S}$, for example the binding of F may occur for all $i \in \mathcal{S}$:

$$\mathcal{S}{+}\mathrm{F} \xrightarrow{r} \mathcal{S}\mathrm{F} \tag{F.5}$$

Components may be supplied with indices:

$$\mathrm{A}_n + \mathrm{B} \xrightarrow{r} \mathrm{A}_n\text{-B}, \quad n = 1, 2, 3 \tag{F.6}$$

which represents:

$$\begin{aligned} \mathrm{A}_1 + \mathrm{B} &\xrightarrow{r} \mathrm{A}_1\text{-B} \\ \mathrm{A}_2 + \mathrm{B} &\xrightarrow{r} \mathrm{A}_2\text{-B} \\ \mathrm{A}_3 + \mathrm{B} &\xrightarrow{r} \mathrm{A}_3\text{-B} \end{aligned} \tag{F.7}$$

Algebraic manipulations of indices apply to all values that the index is defined, for example:

$$\begin{aligned} \mathrm{A}_n,\ n = 1, 2, 3 &\implies \mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3 \\ \mathrm{A}_{n+1},\ n = 1, 2, 3 &\implies \mathrm{A}_2, \mathrm{A}_3, \mathrm{A}_4 \end{aligned}$$

This notation can increase the possibility for ambiguity of determining reactants and products. Any ambiguities can be resolved by conservation, i.e. the same numbers of molecules on both sides of the equation. In the reactions given in this paper conservation looks to be broken in two places. First, the action of Lck causes a phosphate to explicitly appear on the right hand side of an equation. This is not unbalanced, the phosphate is implicitly included in the reactants. Second, the unbinding of MHC causes all left hand

side phosphates to disappear, this is valid due to the assumption of the model that all are removed through a highly abundant fast action phosphatase.

In all reactions if a component that may be associated to a complex and that component is not explicitly given in the reaction, then it is assumed that the reaction may proceed with and without the component associated. For example if B and D may be bound to A, then the reaction:

$$\mathrm{A} + \mathrm{C} \xrightarrow{r} \mathrm{AC} \tag{F.8}$$

includes all the following reactions:

$$\mathrm{A} + \mathrm{C} \xrightarrow{r} \mathrm{AC}, \qquad \text{neither B nor D bound} \tag{F.9}$$

$$\mathrm{AB} + \mathrm{C} \xrightarrow{r} \mathrm{ABC}, \qquad \text{B bound} \tag{F.10}$$

$$\mathrm{AD} + \mathrm{C} \xrightarrow{r} \mathrm{ADC}, \qquad \text{D bound} \tag{F.11}$$

$$\mathrm{ABD} + \mathrm{C} \xrightarrow{r} \mathrm{ABDC}, \qquad \text{B and D bound} \tag{F.12}$$

If A has p phosphorylations it is represented A^p. If no phosphorylation state is given, i.e. A, this represents that the phosphorylation state of A has no affect on the reaction.

Appendix G. Reading Trajectory Plots

This sections describes how to read the trajectory plots such as figure B.21 and 17. It is desirable to be able to visualise the results of multiple simulation runs without taking averages, indicating both common and rare behaviour. The trajectory plots go some way to do this. Time is given on the horizontal axis, and state is given on the vertical axis. Specifically the kinetic proofreading states are repeated three times on the vertical axis for the three mutually exclusive cases: N – neither pSHP1 bound nor protected; pSHP – pSHP1 bound; Prot – ppERK protected. Within each section the kinetic proofreading states are laid out in order with unbound states first: U_0^0, U_1^0, U_2^0,U_3^0, B_0^0, B_0^1, B_0^2, B_1^0, …, B_3^3. The final three signalling states are highlighted with a grey background. A dot represents that at least one simulation passed through the given state at the given time, and a coloured square represents how many runs passed through the state at the time.

Acknowledgements

This work is sponsored by EPSRC Grant Number: EP/E005187/1 and European 7th Framework Programme Project Number: FP7-ICT-2007.2.1.

References

[1] K. M. Murphy, P. Travers, M. Walport, Janeway's Immunobiology, Garland Science, 2008.

[2] R. N. Germain, I. Stefanov, THE DYNAMICS OF T CELL RECEPTOR SIGNALING: Complex Orchestration and the Key Roles of Tempo and Cooperation., Annu Rev Immunol 17 (1999) 467–522.

[3] S. Valitutti, S. Mller, M. Cella, E. Padovan, A. Lanzavecchia, Serial triggering of many T-cell receptors by a few peptide-MHC complexes., Nature 375 (6527) (1995) 148–151.

[4] Z. Grossman, W. E. Paul, Adaptive cellular interactions in the immune system: the tunable activation threshold and the significance of subthreshold responses., Proc Natl Acad Sci U S A 89 (21) (1992) 10365–10369.

[5] Z. Grossman, W. E. Paul, Autoreactivity, dynamic tuning and selectivity., Curr Opin Immunol 13 (6) (2001) 687–698.

[6] I. Stefanov, B. Hemmer, M. Vergelli, R. Martin, W. E. Biddison, R. N. Germain, TCR ligand discrimination is enforced by competing ERK positive and SHP-1 negative feedback pathways., Nat Immunol 4 (3) (2003) 248–254.

[7] G. Altan-Bonnet, R. N. Germain, Modeling T Cell Antigen Discrimination Based on Feedback Control of Digital ERK Responses., PLoS Biol 3 (11) (2005) e356.

[8] C. Chan, A. J. George, J. Stark, Cooperative enhancement of specificity in a lattice of T cell receptors., Proc Natl Acad Sci U S A 98 (10) (2001) 5758–5763.

[9] M. N. Artyomov, J. Das, M. Kardar, A. K. Chakraborty, Purely stochastic binary decisions in cell signaling models without underlying deterministic bistabilities., Proc Natl Acad Sci U S A 104 (48) (2007) 18958–18963.

[10] D. C. Wylie, J. Das, A. K. Chakraborty, Sensitivity of T cells to antigen and antagonism emerges from differential regulation of the same

molecular signaling module., Proc Natl Acad Sci U S A 104 (13) (2007) 5533–5538.

[11] T. Lipniacki, B. Hat, J. R. Faeder, W. S. Hlavacek, Stochastic effects and bistability in T cell receptor signaling., J Theor Biol 254 (1) (2008) 110–122.

[12] O. Feinerman, R. N. Germain, G. Altan-Bonnet, Quantitative challenges in understanding ligand discrimination by alphabeta T cells., Mol Immunol 45 (3) (2008) 619–631.

[13] O. Feinerman, J. Veiga, J. R. Dorfman, R. N. Germain, G. Altan-Bonnet, Variability and Robustness in T Cell Activation from Regulated Heterogeneity in Protein Levels., Science 321 (5892) (2008) 1081–1084.

[14] N. D. L. Owens, J. Timmis, A. Greensted, A. Tyrrell, Modelling the Tunability of Early T Cell Signalling Events, Proceedings of the 7th International Conference on Artificial Immune Systems (ICARIS). LNCS 5132. (2008) 12–23.

[15] J. J. Hopfield, Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity., Proc Natl Acad Sci U S A 71 (10) (1974) 4135–4139.

[16] T. W. McKeithan, Kinetic proofreading in T-cell receptor signal transduction., Proc Natl Acad Sci U S A 92 (11) (1995) 5042–5046.

[17] D. T. Gillespie, Stochastic Simulation of Chemical Kinetics., Annu Rev Phys Chem 58 (2007) 35–55.

[18] N. V. Kampen, Stochastic Processes in Physics and Chemistry, third edition Edition, North-Holland, 2007.

[19] D. T. Gillespie, Exact Stochastic Simulation of Coupled Chemical Reactions, The Journal of Physical Chemistry 81 (25) (1977) 2340–2361.

[20] J. R. Wyer, B. E. Willcox, G. F. Gao, U. C. Gerth, S. J. Davis, J. I. Bell, P. A. van der Merwe, B. K. Jakobsen, T cell receptor and coreceptor CD8 alphaalpha bind peptide-MHC independently and with distinct kinetics., Immunity 10 (2) (1999) 219–225.

[21] H. A. van den Berg, L. Wooldridge, B. Laugel, A. K. Sewell, Coreceptor CD8-driven modulation of T cell antigen receptor specificity., J Theor Biol 249 (2) (2007) 395–408.

[22] C. Y. Huang, J. E. Ferrell, Ultrasensitivity in the mitogen-activated protein kinase cascade., Proc Natl Acad Sci U S A 93 (19) (1996) 10078–10083.

[23] C. Moler, C. V. Loan, Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later, SIAM Review 45 (1) (2003) 3–49.

[24] R. B. Sidje, Expokit: A Software Package for Computing Matrix Exponentials, ACM Trans. Math. Softw. 24 (1) (1998) 130–156.

[25] J. R. Norris, Markov Chains, Cambridge Series in Statistical and Probabilistic Mathematics (No. 2), Cambridge University Press, 1998.

[26] B. Philippe, Y. Saad, W. J. Stewart, Numerical Methods in Markov Chain Modelling, Operations Research 40 (1996) 1156–1179.