Modeling T-cell activation via condensation with kinetic proofreading

Konstantina Foti

Minor research project, MCLS

First examiner: Dr. Florian Berger

Second Examiner: Dr. H.D. Mac Gillavry

Department of Cell Biology, Neurobiology, and Biophysics

Universiteit Utrecht

Acknowledgments

To Florian, for the guidance through a paradigm shift.
To Kyriacos, for making the transition to theory feasible with his translations.
To Amalia and Evi, for always supporting my full-of-struggles academic journey.
To Mum, for everything above and in between.

Contents

Abstract

T-cell receptor interactions with antigens on the surface of immune cells initiate a critical activation process through multiple phosphorylation steps, understood as a kinetic proofreading mechanism that enhances the specificity of the response. The activation process is associated with the formation of protein condensates, involving the linker of the receptor (LAT) and other signaling proteins. However, the contribution of these condensates to the remarkable sensitivity of the mechanism is not clear still. Here we develop a dynamical coarse-grained model to investigate the coupling of kinetic proofreading with condensate formation of the activation signal's final product. We analyze characteristics such as specificity and activation speed, comparing these with traditional kinetic proofreading properties. Our findings suggest that condensation leads to a higher specificity of the synapse's response with the same speed of activation compared to simple kinetic proofreading.

Layman Summary

This study focuses on how T-cells, a type of immune cell, recognize and respond to antigens from foreign invaders, that trigger an immune response. When T-cells encounter an antigen, they go through a series of steps called phosphorylation as part of a "kinetic proofreading" process. This ensures that T-cells respond specifically to the correct antigens, avoiding false alarms. During this activation process, certain proteins form clusters or "condensates". The exact role of these condensates has been unclear. To investigate their potential contribution to a T cell's response, we developed a mathematical model to describe how these condensates enhance the kinetic proofreading process. Our findings suggest that protein condensates help improve the accuracy of the T-cell response without slowing down the activation process, implying a more effective antigen distinction when these condensates are present. They achieve specificity through a series of biochemical steps that enhance their ability to discriminate between different molecules. One challenge has been understanding how these steps can be both fast and highly specific. Incorporating the concept of protein condensation into the kinetic proofreading model, the study reveals that these condensates act like a positive feedback loop. Once condensation starts, it promotes more condensation, similar to how positive feedback in other systems enhances its own output. This self-reinforcing loop leads to a significant increase in the production of signaling molecules, thus amplifying the initial signal. The results show that positive feedback through condensation not only increases the signal strength but also improves the system's ability to specifically respond to the correct antigens. This improved specificity is crucial for effective immune responses, as it means T-cells can quickly and accurately respond to pathogens while ignoring harmless molecules. This research bridges the gap between the biochemical understanding of T-cell activation and the biophysical processes involved. By combining kinetic proofreading with protein condensation, the study provides a more comprehensive model of how T-cells achieve both high sensitivity and specificity in their responses. This enhanced understanding is important to unravel the molecular mechanisms of highly specific immune response that are misfunctioning in cancer and autoimmune disorders.

1 Introduction: Molecular mechanisms underlying T-cell Activation

The T-cell receptor (TCR) is the key receptor on the surface of adaptive immune cells, that can recognize antigenic peptides presented by major histocompatibility complexes (pMHCs) and initiate T-cell activation and immune responses against foreign pathogens or cancerous cells (reviewed in Malissen and Bongrand, 2015). The affinity of each pMHC can be translated into differentiated phosphorylation profiles that subsequently trigger various downstream signaling events. The increased chemical similarity of foreign and self-peptides and the significantly higher concentration of the latter (Irvine et al., 2002) constitute a major bottleneck for the robustness of T-cell activation. Multiple theories have been suggested to explain the simultaneous high sensitivity and selectivity of T-cell activation: kinetic segregation, kinetic proofreading, receptor scanning, and serial triggering (reviewed in Courtney, Lo, and Weiss, 2018). It would be fair to propose that all of them, and maybe more, contribute to the essential traits of T-cell activation. In this thesis, we will focus mainly on kinetic proofreading and its correlation with another biophysical phenomenon taking place in T cells, biomolecular condensation.

Biochemistry of the immune synapse

Before we describe the biophysics of the project, it would be useful to present the rapid, spatiotemporally arranged biochemical events that take place (Figure 1.1). The receptor's architecture has been thoroughly investigated to characterize the initiation of the process. The $\alpha\beta TCR$ is comprised of the antigen-recognition α and β subunits and three CD3 signaling subunits–CD3$\epsilon\delta$, CD3$\epsilon\gamma$, and CD3$\zeta\zeta$ (Wucherpfennig et al., 2010)–that all facilitate signal transmission. Tyrosine residues on the cytosolic regions of the CD3 chains are phosphorylated by the Src family protein tyrosine kinase Lck. Subsequently, the autoinhibited cytosolic kinase ZAP70 is recruited and activated to phosphorylate multiple tyrosine residues of the linker of activation of T cells, LAT, which then serves as the docking site for the recruitment of cytosolic adaptors such as Grb2, Gads and enzymes like PLC-γ1 that interact with SLP-76 and Sos1 amongst others (reviewed in Balagopalan et al., 2015). This phosphorylation cascade is suggested to drive multi-protein condensation that modulates signaling in the formed immunological synapse (reviewed in Balagopalan et al., 2021).

Kinetic proofreading

The high sensitivity and specificity of ligand discrimination could not be explained via equilibrium thermodynamic processes. McKeithan (1995) proposed a model based on kinetic proofreading, initiating the development of multiple theoretical frameworks for T-cell activation. As a general principle, kinetic proofreading enhances ligand discrimination by inserting a delay between the initial recognition and downstream signaling events (reviewed in Boeger, 2022). Therefore, a response should occur when receptor-proximal signaling molecules undergo a series of reversible biochemical modifications that create a lag between the input and the output. Thermodynamically irreversible commitment step(s) ensure that only specific complexes trigger activation. Conclusively, the probability of signaling translating to activation depends on the pMHC

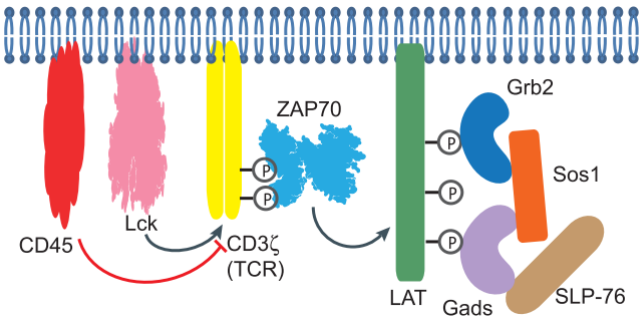


Figure 1.1: *Schematic of the components of early T-cell activation (adjusted from Su et al. (2016)). From left to right: the cytoplasmic domains of the excluded phosphatase (red arrow) CD45, the engaging kinase (black arrow) Lck, and the co-receptor CD3ζ with its phosphorylation sites (P)-via which it interacts with the kinase ZAP70. The latter phosphorylates (black arrow) LAT in multiple tyrosine residues (P), that serve as docking sites for adaptors and enzymes, like Grb2, Gads, SLP-76, and Sos1.*

dwell time to TCR and the dissociation constant of each kinetic step.

The McKeithan model (Figure 1.2) proposes that the initial specific or non-specific ligand-receptor complex, C_0, undergoes a series of energy-consuming modifications and interactions that produce a sequence of intermediates, C_i, before the former's conversion to an active complex C_n. Dissociation of the complex translates into reversal of modifications and therefore waste of metabolic energy for the cell. Non-specific intermediate complexes are characterized by high dissociation rates to ensure that signal propagation will not be successful (McKeithan, 1995). Since the quantification of the kinetic rates of activation is still a problem to be tackled, multiple simplifications have to be made. For example, we assume that the intermediate steps occur in a particular order, are of equal rate and their dissociation rate is the same regardless of the stage of modification. Additionally, the association constant is independent of the peptides' nature, but the dissociation constant varies for foreign and self-peptides. It is important to mention that the dissociation rate is significantly smaller for the activated complex.

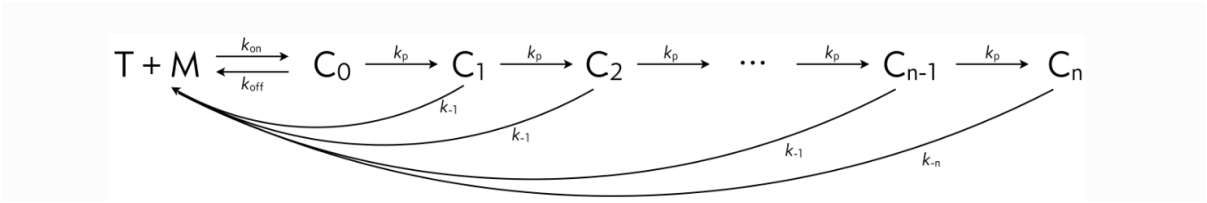


Figure 1.2: *Kinetic proofreading in T-cell activation (adjusted from McKeithan (1995)): T: TCR, M: pMHC, C_0: ligand-receptor complex, $C_1 - C_{n-1}$: intermediate complexes, C_n: active complex. The formed complexes ($C_0 - C_{n-1}$) undergo N modifications, with constant phosphorylation rate k_p, to generate the active complex, C_n. Each complex can dissociate with dissociation rate k_{-1}, leading to the unbound receptor and ligand. The dissociation rate constant is assumed to be the same in every step, thus k_{n-1} equals k_{-1}. The dissociation rate for the initial complex C_0 is k_{off} and the association rate is k_{on}.*

Embedding the complexity of biological data in the model can be quite challenging. Lo and Weiss (2021) reviewed the mechanism (Figure 1.3) in regard to the kinetics of

LAT phosphorylation. Voisinne et al. (2022) suggested that ligand discrimination is encoded in the multi-step activation of ZAP70. Initially, ZAP70 binds to CD3 and is then phosphorylated by Lck on critical tyrosine residues, promoting its active conformation. This activation step involves increased phosphorylation of two tyrosine residues via transphosphorylation within the TCR complex. According to the kinetic proofreading model, weak-affinity ligands unbind from TCR before transphosphorylation completes. There is a correlation between TCR signaling and ZAP70's ability to phosphorylate LAT, with condensation occurring 15 seconds after ligand engagement (McAfee et al., 2022), suggesting two kinetic proofreading steps (Voisinne et al., 2022). Slow phosphorylation of LAT's Y132 residue by ZAP70 is a rate-limiting step for response selectivity (Lo and Weiss, 2021), facilitating the recruitment and activation of PLC-γ1 in LAT condensates.

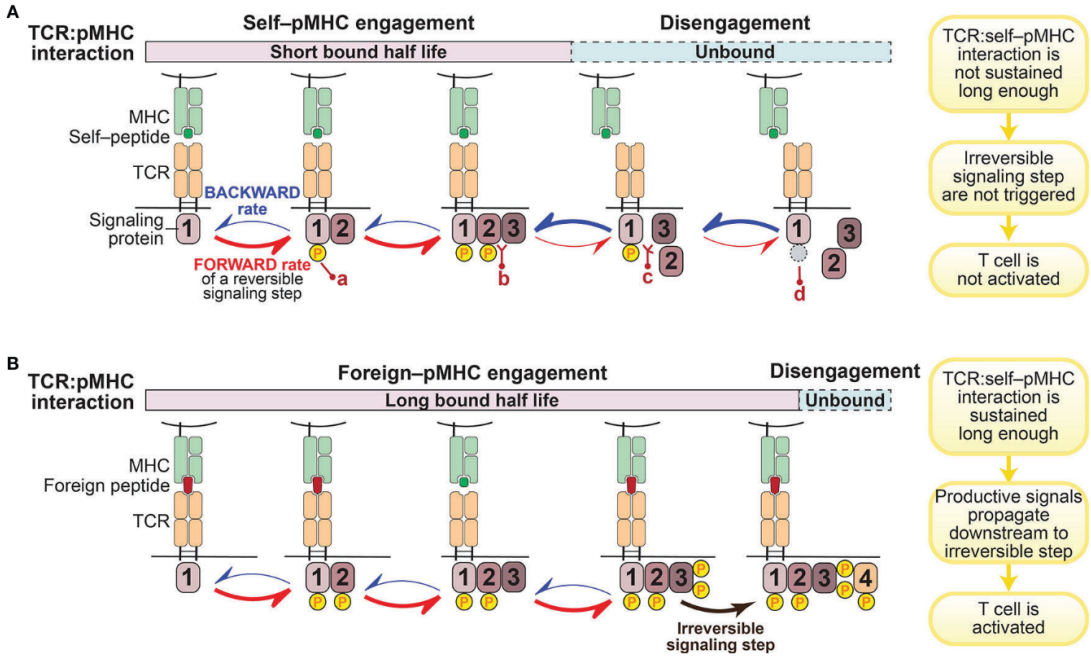


Figure 1.3: *Analytical schematic of kinetic proofreading model in T cells (adjusted from Lo and Weiss (2021)). This model proposes that the initial TCR-pMHC binding initiates a sequence of biochemical reactions that lead to T-cell activation. The reactions (yellow circle; a)/dephosphorylation (gray circle; d), or protein-protein interaction (b)/dissociation (c)) are reversible, allowing TCR disengagement to promptly reset the signaling intermediates to their initial state. (A) The interaction between TCR and self–pMHC is relatively weak with a short binding duration, preventing the signal from reaching an irreversible step before TCR dissociates. (B) The interaction between TCR and foreign pMHC is long enough to progress to a final irreversible step. T cell activation occurs only if the TCR interaction is long enough to complete all reversible kinetic proofreading steps and reach the crucial irreversible step.*

Liquid-liquid phase separation and kinase regulation

Liquid-liquid phase separation (LLPS) is governed by core biophysical principles that apply universally in cells for the membrane-less, spatiotemporal compartmentaliza-

tion of biochemical reactions. The liquid-like properties of condensates distinguish them from macromolecular assemblies and suggest that phase separation provides a unifying mechanism for segregated cellular biochemistry. As their name implies, biomolecular condensates are comprised of biomolecules that could be diverse in multiple aspects, but their behavior can be examined under classic thermodynamic principles like the maintenance of chemical equilibrium in demixed droplets (reviewed in Banani et al., 2017).

In molecular systems, phase separation occurs at a critical concentration, where the energetics of weak interactions overcome the entropy of a homogeneous mixed system and hence minimum free energy is reached at the condensed state. The achieved equilibrium eliminates diffusive flux and allows the maintenance of condensates in cells (Hyman, Weber and Jülicher, 2014). The principle of equal chemical potential in the two phases indicates that altering the local concentrations of key components could be essential. At the same time, post-translational modifications modulate the valency and solubility of vital constituents (Su et al., 2016).

The strictly regulated kinase activity in cells is thought to be orchestrated via their accumulation in condensates, amongst other reasons (Sang et al., 2022). The acceleration of biochemical reactions is related to the molecular crowding of kinases and their substrates. At the same time, there has been evidence of condensation as a response to molecular crowding (Cai et al., 2019). In their investigation of this hypothesis, Sang et al. demonstrated the modulation of phosphorylation rates in condensates due to molecular crowding *in vitro* and *in vivo*. The spontaneous and dense localization of molecules at discrete cellular sites increases reaction kinetics and thus drives cells in specific responses. Acceleration of the chemistry of selected pathway components with the concurrent exclusion of interfering ones could stabilize the phosphorylation-dependent clusters and enhance selectivity (Su et al., 2016).

Another important concept derived from polymer science employs multivalency for the natural formation of clusters via intra- and inter-molecular interactions that decrease the solubility of the constituents due to entropy-driven effects. In our case, the required multivalency of the scaffold lies in the intrinsically disordered regions of LAT, that undergo various types of homotypic and heterotypic interactions with attracted clients (Nag et al., 2009). It is important to underline that those multimolecular interactions are described as synergic, meaning that the proximity to one type of molecule increases the probability that the molecule of interest will also be close to a second type of molecule. Therefore, a molecule that is recruited to a cluster can bind directly to LAT or indirectly via interacting with another molecule that binds to LAT in a cooperative manner (Banani et al., 2017).

However, it remains unknown how the mesoscale structures of interest are maintained and regulated in composition and size. Söding et al. (2020), in an effort to characterize condensate size behavior via active regulation, described a generic mechanism. The localization-induction model states that the localization of regulatory kinase ZAP70 in the proximal area of condensation induces favorable interactions, thus pushing protein concentration in the phase separation regime (Söding et al., 2020).

The simplified biophysical depiction that sets the series of phosphorylations as the trigger of condensation could also incorporate the multivalency of the scaffold protein that

allows cooperative interactions on its flexible, intramolecular interface. The exclusion of the CD45 phosphatase from the condensates (Su et al., 2016) is an important element that constitutes the segregation model for ligand discrimination and is also assimilated into this general profile of condensation. Another remark is related to the phosphorylation rate(s), which needs to exceed a certain threshold to overcome coarsening. The capacity of this mechanistic model to incorporate biophysical fundamentals regarding the regulation of LAT condensation renders it an important tool for the investigation of the amplification of signaling in T-cell activation.

Objectives

T-cell activation has repeatedly been addressed with theoretical modeling to unravel its levels of complexity. Even though a comprehensive map of the molecular events has been constructed, little is known about the contribution of condensed signaling molecules to the discriminatory power of the mechanism. The manipulation of a simplified network of biochemical reactions (taken from Ganti et al., 2020) and the incorporation of the biophysical concept of condensation could provide more insights. Conclusively, the main goal of this thesis is to incorporate condensation into kinetic proofreading to potentially capture the amplification of the signal. For that reason, we develop two different models: model 0 for kinetic proofreading and model 1 with kinetic proofreading and a condensation process by positive feedback.

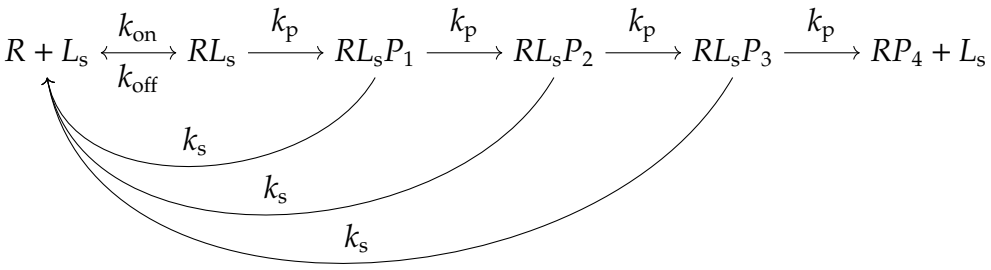
2 Methods: Two mathematical descriptions of T-cell signaling

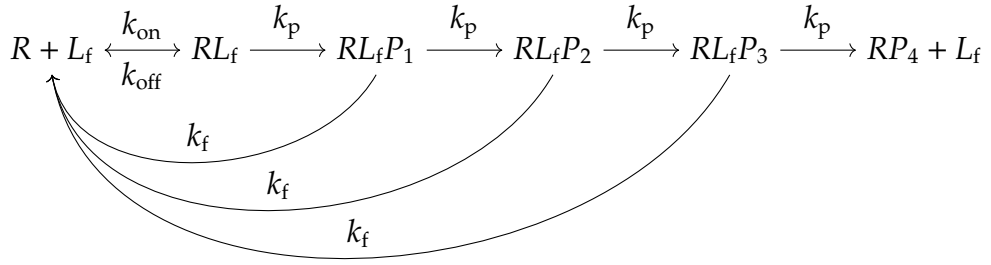
Our mathematical models aim to capture the fundamental biochemistry of T-cell activation. The biochemical reactions are described with rate equations that provide ordinary differential equations which we will solve numerically. The large number of parameters that construct and regulate the behaviour of our system could not all be incorporated in our biophysical description. Their values (Appendix) are set both on measurements and fitting because many enzymatic rates cannot be measured in cells (Ganti et al., 2020). The McKeithan (1995) model itself falsely depicts the dissociation of downstream signaling complexes directly to receptor and ligand and multiple simplified–including ours–models depend on this and many other assumptions.

2.1 Model 0: Kinetic Proofreading in T-cell activation

A simplified kinetic proofreading model that was developed by Ganti et al. (2020) to assess the minimum number of biochemical steps for kinetic proofreading in T-cell activation is utilized as the basis of our analysis. After a fixed number of steps ($N = 4$), the self ligand L_s or foreign ligand L_f remains bound to the receptor R, a non-ligand specific product RP_4 acts as an enzyme and phosphorylates a downstream substrate S, leading to the formation of the product SP. Therefore, we split the reaction network into parts: Part A for the phosphorylation cascade and part B for the signal production.

Part A: The Phosphorylation cascades towards a non-ligand-specific product can be considered as the following reaction network, here shown for both, the self ligand, L_s and the foreign ligand, L_f:





In these reactions, we treat the rates for the phosphorylation cascade for the self and the foreign ligand the same; the only difference is the unbinding rate when the complex is phosphorylated, for the self ligand k_s and for the foreign k_f, respectively. Here, k_{on} is the receptor-ligand binding rate, and k_{off} is the unbinding rate. The phosphorylation rate is k_p.

Part B: The Signal production pathway reads

$$RP_4 + S \underset{k_{off,s}}{\overset{k_{on,s}}{\rightleftharpoons}} RP_4S \xrightarrow{k_p} SP + RP_4$$

$$SP \xrightarrow{k_{off}} S \quad RP_4 \xrightarrow{k_{off}} R$$

Here $k_{on,s}$, $k_{off,s}$ refer to the binding and unbinding rate of the substrate S, respectively. We combined the chemical networks of Part A with Part B and determine a system of ODEs describing the time evolution of the model (for $N = 4$):

$$\begin{aligned}
\frac{d[R]}{dt} &= -k_{on}[R][L_f] - k_{on}[R][L_s] + k_f \sum_{i=1}^{N-1}[RL_fP_i] + k_s \sum_{i=1}^{N-1}[RL_sP_i] + k_{off}[RP_4] \\
\frac{d[L_f]}{dt} &= -k_{on}[R][L_f] + k_f \sum_{i=1}^{N-1}[RL_fP_i] + k_p[RL_fP_3] \\
\frac{d[L_s]}{dt} &= -k_{on}[R][L_s] + k_s \sum_{i=1}^{N-1}[RL_sP_i] + k_p[RL_sP_3] \\
\frac{d[RL_f]}{dt} &= k_{on}[R][L_f] - k_f[RL_f] - k_p[RL_fP_1] \\
\frac{d[RL_s]}{dt} &= k_{on}[R][L_s] - k_s[RL_f] - k_p[RL_sP_1] \\
\frac{d[RL_fP_1]}{dt} &= k_p[RL_f] - k_f[RL_fP_1] - k_p[RL_fP_1] \\
\frac{d[RL_sP_1]}{dt} &= k_p[RL_s] - k_s[RL_sP_1] - k_p[RL_sP_1] \\
\frac{d[RL_fP_2]}{dt} &= k_p[RL_fP_1] - k_f[RL_fP_2] - k_p[RL_fP_2] \\
\frac{d[RL_sP_2]}{dt} &= k_p[RL_sP_1] - k_s[RL_sP_2] - k_p[RL_sP_2] \\
\frac{d[RL_fP_3]}{dt} &= k_p[RL_fP_2] - k_f[RL_fP_3] - k_p[RL_fP_3] \\
\frac{d[RL_sP_3]}{dt} &= k_p[RL_fP_2] - k_s[RL_sP_3] - k_p[RL_sP_3] \\
\frac{d[RP_4]}{dt} &= k_p[RL_fP_3] + k_p[RL_sP_3] + k_p[RP_4S] + k_{off,s}[RP_4S] - k_{on,s}[RP_4][S] - k_{off}[RP_4] \\
\frac{d[RP_4S]}{dt} &= k_{on,s}[RP_4][S] - k_{off,s}[RP_4S] - k_p[RP_4S] \\
\frac{d[SP]}{dt} &= k_p[RP_4S] - k_{off}[SP] \\
\frac{d[S]}{dt} &= k_{off}[SP] + k_{off,s}[RP_4S] - k_{on,s}[RP_4][S]
\end{aligned}$$

2.2 Model 1: Positive Feedback for Describing Condensation

An alteration of model 0 - in which part B is modified - is designed to introduce the concept of condensation. More in detail, we consider the non-ligand specific, phosphorylated product RP_4 in two states: the free, RP_4^{free} and the condensed one RP_4^{cond}. Describing the kinetics of the biochemical network of T-cell activation in this context could provide some insights regarding the interconnection of protein condensation and acceleration of reactions in which those proteins participate. Conceptually, condensation is set to model the faster catalysis of signal transduction.

McAffee et al. (2022) reported the distinct transition in the propagation of downstream signal after TCR-triggering, that occurs via condensation of LAT. Hence, an amplification step and a thresholding mechanism seem to be responsible for signal digitization. Fluctuations in the phosphorylation rates control the initial nucleation process, the success of which is thought to be responsible for the following condensation. Thus, a positive feedback mechanism could enhance the selectivity of T-cell activation.

For simplicity, we consider a phenomenological model for condensate formation of RP_4. We first distinguish the population of RP_4 into a free state RP_4^{free} and a condensate state RP_4^{cond}. In our model, the rate of condensation is proportional to the amount of molecules already in condensate form. This, on a coarse grain level, captures the idea that free RP_4^{free} will have a higher rate of meeting and incorporating on a condensate when the total amount of condensed RP_4^{cond} is high. For the release of RP_4 from condensed to the free state, we considered simple first-order kinetics. Schematically our positive feedback model reads:

$$RP_4^{\text{free}} + RP_4^{\text{cond}} \rightarrow RP_4^{\text{cond}} + RP_4^{\text{cond}}$$
$$RP_4^{\text{cond}} \rightarrow RP_4^{\text{free}}.$$

Because the population of RP_4^{free} is produced by a more complex series of reactions, we would like to see the properties of this mechanism for an even simpler system. Therefore, we consider, C (condensed) and F (free) molecules

$$C + F \rightarrow C + C$$
$$C \rightarrow F,$$

taking into consideration the conservation of the total number $[T] = [C] + [F]$.

The rate equations describing the system are

$$\frac{d[F]}{dt} = -k_c[C][F] + k_f[C]$$
$$\frac{d[C]}{dt} = k_c[C][F] - k_f[C].$$

Using $[F] = [T] - [C]$, we decouple the equations and find the associated ODE for $[C]$,

$$\frac{d[C]}{dt} = k_c[C]([T] - [C]) - k_f[C]$$

The points at which the rate of change vanishes, define the fixed points of $[C]$,

$$0 = k_c[C]([T] - [C]) - k_f[C]$$

$$0 = -k_c[C]^2 + (k_c[T] - k_f)[C] = 0.$$

Thus one fixed point is at $[C] = 0$ and the other for

$$[C] = \frac{k_c[T] - k_f}{k_c}.$$

For $k_c[T] < k_f$ the fixed point is negative, which is not possible because $[C]$ is considered to be a concentration. Therefore, we only have the fixed point at $[C] = 0$. The transition from one to two fixed points happens when $k_c[T] = k_f$. These conditions characterize a bifurcation, where for $k_c[T] < k_f$ there is no condensation and for $k_c[T] > k_f$, we get stable condensates.

After characterizing the fixed points of our simplified system, we can modify part B in model 0, in order to incorporate the two different states of RP_4 and describe the kinetics of its condensation.

Part B(ii): We include a positive feedback in the signal production as

$$RP_4^{\text{free}} + S \underset{k_{\text{off,s}}}{\overset{k_{\text{on,s}}}{\rightleftharpoons}} RP_4^{\text{free}}S \xrightarrow{k_p} SP + RP_4^{\text{free}}$$

$$RP_4^{\text{cond}} + S \underset{k_{\text{off,s}}^{\text{cond}}}{\overset{k_{\text{on,s}}^{\text{cond}}}{\rightleftharpoons}} RP_4^{\text{cond}}S \xrightarrow{k_p^{\text{cond}}} SP + RP_4^{\text{cond}}$$

$$SP \xrightarrow{k_{\text{off}}} S \qquad RP_4^{\text{free}} \xrightarrow{k_{\text{off}}} R \qquad RP_4^{\text{cond}} \xrightarrow{k_{\text{off}}} R$$

$$RP_4^{\text{free}} \quad \underset{k_c}{\overset{k_c^{-1}}{\updownarrow}} \quad RP_4^{\text{cond}}$$

The two different states of RP_4 reflect different kinetics. Here, we introduce the modified binding and unbinding rates for the RP_4^{cond}, $k_{\text{on,s}}^{\text{cond}}$ and $k_{\text{off,s}}^{\text{cond}}$ respectively. The values of these parameters (Appendix) aim to capture the acceleration of kinetics in the condensed branch. For the same reason, there is a different phosphorylation rate k_p^{cond}. The condensation rate for the formation of RP_4^{cond} is k_c and its reverse is k_c^{-1}. Finally, there is another kinetic term associated with the condensation. The ODE of $[RP_4^{\text{cond}}]$ includes a constant (here set to 1) that accounts for spontaneous condensation. This term triggers the process of condensation, without relying on the pre-existence of condensed molecules in the system to drive nucleation.

From the chemical pathways given above, we write the system of ODEs describing the modified part B of the model as

$$
\begin{aligned}
\frac{d[RP_4^{\text{free}}]}{dt} &= k_{\text{p}}[RL_{\text{f}}P_3] + k_{\text{p}}[RL_{\text{s}}P_3] + k_{\text{off,s}}[RP_4^{\text{free}}S] - k_{\text{on,s}}[RP_4^{\text{free}}][S] \\
&\quad - k_{\text{off}}[RP_4^{\text{free}}] - k_{\text{c}}[RP_4^{\text{free}}][RP_4^{\text{cond}}] + k_{\text{c}}^{-1}[RP_4^{\text{cond}}] + k_{\text{p}}^{\text{free}}[RP_4^{\text{free}}] \\
\frac{d[RP_4^{\text{cond}}]}{dt} &= (1 + k_{\text{c}}[RP_4^{\text{cond}}])[RP_4^{\text{free}}] - k_{\text{c}}^{-1}[RP_4^{\text{cond}}] + k_{\text{off,s}}^{\text{cond}}[RP_4^{\text{cond}}S] \\
&\quad - k_{\text{on,s}}^{\text{cond}}[RP_4^{\text{cond}}] + k_{\text{p}}^{\text{cond}}[RP_4^{\text{cond}}S] - k_{\text{off}}[RP_4^{\text{cond}}] \\
\frac{d[RP_4^{\text{free}}S]}{dt} &= k_{\text{on,s}}[RP_4^{\text{free}}][S] - k_{\text{off,s}}[RP_4^{\text{free}}S] - k_{\text{p}}[RP_4^{\text{free}}S] \\
\frac{d[RP_4^{\text{cond}}S]}{dt} &= k_{\text{on,s}}^{\text{cond}}[RP_4^{\text{cond}}][S] - k_{\text{off,s}}^{\text{cond}}[RP_4^{\text{cond}}S] - k_{\text{p}}^{\text{cond}}[RP_4^{\text{cond}}S] \\
\frac{d[SP]}{dt} &= k_{\text{p}}[RP_4^{\text{free}}S] - k_{\text{off}}[SP] + k_{\text{p}}^{\text{cond}}[RP_4^{\text{cond}}S] \\
\frac{d[S]}{dt} &= k_{\text{off}}[SP] + k_{\text{off,s}}[RP_4^{\text{free}}S] + k_{\text{off,s}}^{\text{cond}}[RP_4^{\text{cond}}S] \\
&\quad - k_{\text{on,s}}[RP_4^{\text{free}}][S] - k_{\text{on,s}}^{\text{cond}}[RP_4^{\text{cond}}][S]
\end{aligned}
$$

In summary, we develop two different mathematical models that explore the biochemistry of T-cell activation, focusing on kinetic proofreading principles and the role of protein condensation. Model 0, based on Ganti et al. (2020), describes the phosphorylation cascade leading to the production of the non-ligand-specific product RP_4, split into two parts detailing the phosphorylation events and subsequent signal production. Model 1 introduces protein condensation in part B, differentiating RP_4 into free and condensed states to capture the acceleration of kinetics due to condensation.

3 Results: Condensation improves T-cell activation

The described kinetics of the phosphorylation cascade of T-cell activation are numerically solved via a ODE-solving package in Python. This computational environment also offers visualization of the models' outcomes. The values of the rate constants and the initial concentrations that are implemented in the systems of ODEs described in Methods are presented in the Appendix.

3.1 Time evolution of the biochemical networks indicate amplification through condensation

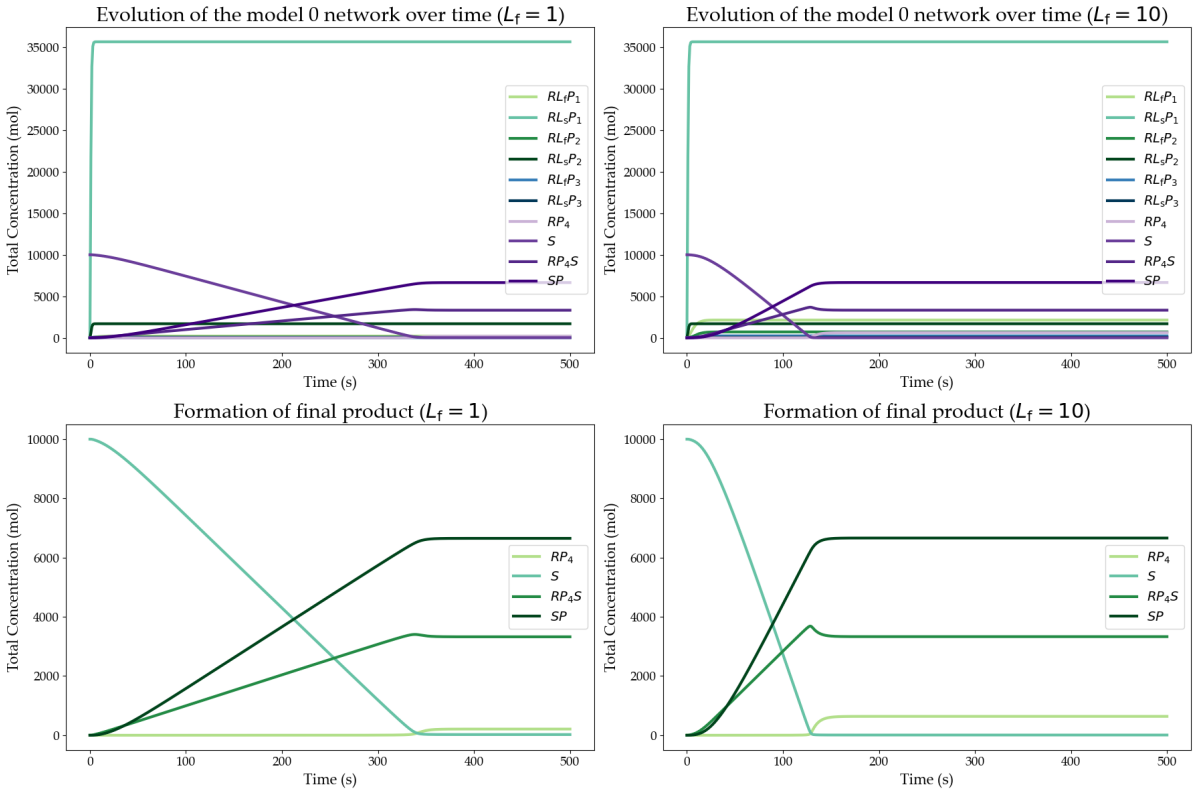


Figure 3.1: *Time evolution of the model 0 biochemical network over time (500 s) and for two different concentrations of the foreign ligand, L_f. Top left: depiction of the concentration of key components of the network (as listed in the top-left plot legend), for $L_f = 1$ mol. Top right: dynamical change of the concentration of the same components (see legend of the top-right plot) , for $L_f = 10$ mol. Bottom right: Zoom in the final products' behavior towards the steady state, for $L_f = 10$ mol. Bottom left: Zoom in the dynamical change of the final products' concentration, for $L_f = 1$ mol*

The biochemical networks of both models for T-cell activation, as described by the systems of equations (in Methods), indicate that after a transient time all concentrations reach a steady state (see Figures 3.1 and 3.2). Because of the parametrization of the system, the biologically relevant values (Appendix Tables 5.1 and 5.2) of the ligands and their initial complexes with the receptor are considerably too high and consequently not comparable with those of the intermediate complexes. For that reason, RL_f and RL_s are not depicted in the specific plots. The ongoing challenge of accurately quantifying

kinetics *in vivo* mandates numerous simplifications and assumptions - most of which are enlisted in the McKeithan (1995) model description.

As a general trend, the concentrations of all species stabilize after an initial period of significant change, reaching a steady state within approximately 150 seconds for $[L_f] = 10\,\text{mol}$ and within 300 seconds for $[L_f] = 1\,\text{mol}$, respectively. This convergence is in agreement with the typical time scale for T-cell activation (1-5 min; McAffee at al., 2022). While both figures illustrate the temporal dynamics of the network, they are characterized by different complexities. The initial set value of the foreign ligand is substantially lower than the ones of the self-peptide and the receptor (see Appendix). Interestingly, when only increased by a factor of 10 (top and bottom right of Figure 3.1), it leads to a noteworthy increase of the *SP* production, which is strictly related to the production of signal, as discussed below.

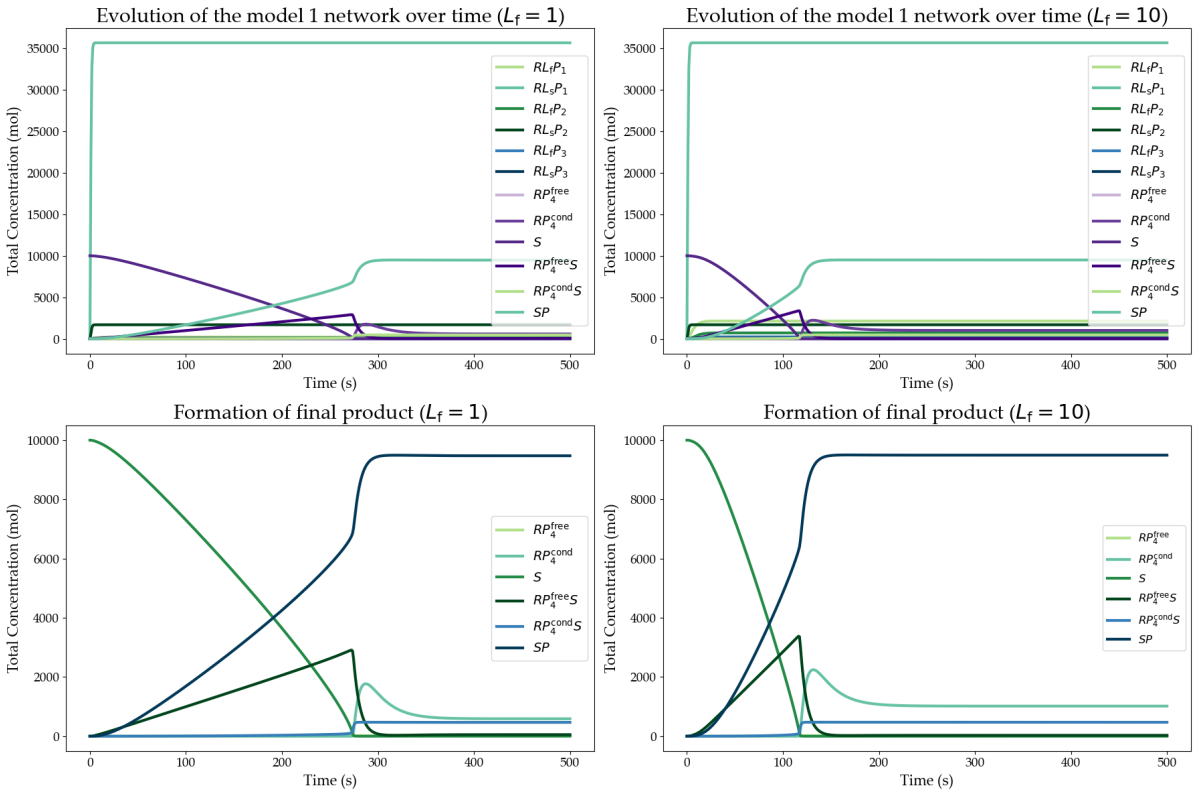


Figure 3.2: *Time evolution of the model 1 biochemical network over time (500 s) and for two different concentrations of the foreign ligand, L_f. Top left: depiction of the concentration of certain components of the network (as listed in the top-left plot legend) for $L_f = 1\,\text{mol}$. Top right: dynamical change of the concentration of the same components (see legend of the top-right plot) , for $L_f = 10\,\text{mol}$. Bottom right: Zoom in the final products' behavior towards the steady state, for $L_f = 10\,\text{mol}$. Bottom left: Zoom in the dynamical change of the final products' concentration, for $L_f = 1\,\text{mol}$*

In both models, increasing the initial amount of L_f results in reaching the steady state faster. Model 0 relies solely on the intermediates of the phosphorylation cascade for ensuring signal fidelity, resulting in moderate *SP* production. On the contrary, the substantial rise of *SP* levels (reaching 9000 mol) in model 1 indicates that this mechanism enhances signal amplification. The complex dynamics and stabilization patterns observed in Figure 3.2 (especially bottom right) are consistent with the behavior of bio-

chemical networks involving condensation reactions and underline the correlation of condensation and amplified signaling. More specifically, the presence of a pronounced $RP_4^{\text{free}}S$ peak right before the RP_4^{cond} peak and the subsequent noteworthy increase of SP hint a self-reinforcing loop that leads to amplification of the initial signal.

3.2 T-cell Activation Properties

To effectively get activated, adaptive immune cells must simultaneously exhibit high sensitivity, specificity, and speed in their responses. The quantitative analysis of the "golden triangle" (Feinerman et al., 2008) proved to not be very straightforward in our analysis and thus we proceed with our definitions and calculations of the signal, the sensitivity, and the specificity, as presented below. We analyze these quantities within the two models.

3.2.1 The activation signal indicates enhanced responsiveness of the biochemical network with condensation

For the model 0, we defined the signal as

$$\text{Signal} = [SP][RP_4^{\text{free}}],$$

implying that subsequent processing of these two products by the T cell triggers activation. Consistent with this definition, the signal for model 1 with positive feedback reads

$$\text{Signal} = [SP]([RP_4^{\text{free}}] + [RP_4^{\text{cond}}]).$$

By this definition of the signal, we aim to capture the increase in the formation of the final product SP that is linked to the absence or presence of the condensed RP_4, respectively.

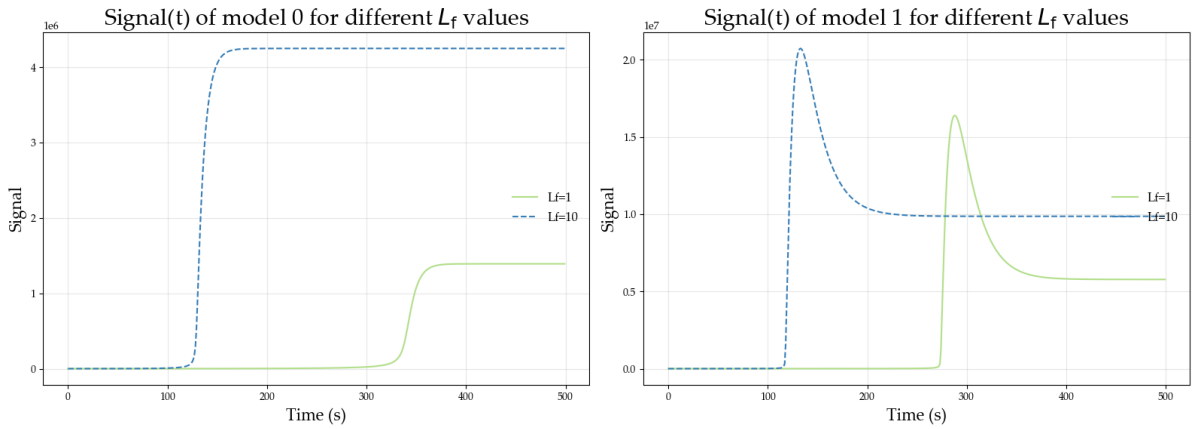


Figure 3.3: *Signal as a function of time. Left: Production of signal in model 0 for* $L_f = 1\,\text{mol}$ *and* $L_f = 10\,\text{mol}$. *Right: Production of signal in model 1 for the same values of* L_f *as in model 0.*

First, we evaluate the signal as a function of time for both models (Figure 3.3) in order to quantify how long it takes for the system to reach the steady state in each case. Conceptually, we can translate these plots into the speed of the two compared mechanisms, due to the fact that they represent how quickly and efficiently the systems respond to the presence of L_f. The steepness of the curves reflects how fast the signal is

being produced and the height of the peaks the maximum efficiency of the production. The time it takes for the signal to stabilize is also an interesting trait since it indicates the duration of the active signaling phase. The shorter time to reach the steady state implies a quicker completion of the signaling process.

We expected that the signal as a function of time would be in agreement with the formation of the final product (Figures 3.1, 3.2 bottom, left and right), regarding the time needed to reach the steady state in different initial concentrations of the triggering ligand. Surprisingly, there is no substantial difference in the amount of signal produced via the positive feedback mechanism for different values of L_f (Figure 3.3 right). This could reflect a digitization of the signal and the all-or-nothing response of T cells (Werterk and Xu, 2014). Therefore we conclude that the described kinetics of condensation could be realistic.

Next, we analyze the steady state of the signal as a function of the foreign ligand concentration $[L_f]$ (Figure 3.4 left). The signal increases linearly with $[L_f]$ for both models, meaning that the signal production is directly proportional to the ligand concentration. That finding was not expected because of the positive feedback mechanism. A potential explanation would be the substantial difference in the initial amount of receptor R and L_f. Moreover, we can notice that the signal increases more steeply with increasing L_f for model 1 with positive feedback compared to model 0. The former results in a higher overall signal for the same L_f values and hence enhances the network's responsiveness.

3.2.2 Condensation increases Sensitivity

Sensitivity refers to the ability of the system to correctly detect and amplify the signal, meaning it measures the system's responsiveness to the presence of the correct substrate. We can, thus, write:

$$\text{Sensitivity} = \frac{d(\text{Signal})}{d(L_f)}.$$

The linearity of the signal of both models (Figure 3.4 left) implies that the sensitivity is constant. The rate of change of the signal with respect to L_f does not vary across different values of Lf. Our observations validate that signal amplification is higher due to positive feedback, but this mechanism does not change the proportionality between L_f and the output. The parametrization of the models could again be the main reason for this result but it should be further investigated.

3.2.3 Condensation increases Signal Specificity

In our deterministic approach, we can conceptually define specificity as the ability of the system to correctly identify and produce the signal for T-cell activation, without generating false positives. It basically measures how well the system discriminates between foreign and self-ligands. In our case, we can define the specificity for the steady-state signal as

$$\text{Specificity} = \frac{\text{Signal}([L_f])}{\text{Signal}([L_f]) + \text{Signal}([Lf] = 0)}$$

Finally, the specificity (Figure 3.4 right) appears to improve with the positive feedback extension, starting higher and increasing more rapidly compared to model 0. While kinetic proofreading improves specificity with increasing [L_f], it does so at a slower rate and with a lower overall signal. We can therefore conclude that condensation enhances signal strength and accuracy and that kinetic proofreading alone might not be sufficient for highly demanding biological processes, where both high specificity and strong signals are required.

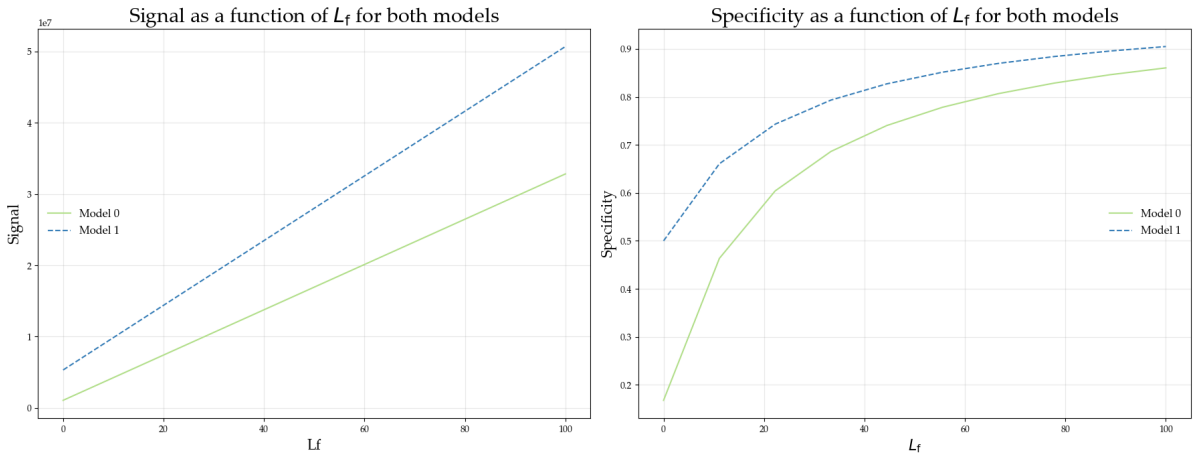


Figure 3.4: *Left: Signal as a function of* [L_f] *for model 0 and model 1. Right: Specificity of the two models for different values of* [L_f].

4 Conclusion and Perspective

An effective T-cell-mediated response hinges on the capacity to accurately and swiftly respond to an activating signal. This task is challenging due to the fact that T cells constantly interact with self peptides to ensure their survival. Therefore, T cells must effectively differentiate between agonist pMHC stimuli and the background noise of self pMHC. Molecular approaches, high-resolution imaging techniques, and biophysical tools have all significantly contributed to the identification of the main components and their interactions, hence have mapped the propagation of signaling. This quantitative element of investigation, even though critical, has not provided a conceptual comprehension of ligand discrimination nor of amplification of signaling. An attempt to capture a correlation between the former and the latter seems only rational.

Our deterministic approach aims to interconnect kinetic proofreading and protein condensation with the two aforementioned phenomena respectively. Kinetic proofreading is a mechanism that has been consistently used to characterize T-cell activation. In recent work (Ganti et al., 2020), the minimum number of kinetic proofreading steps for T cells was defined and the localization of those steps was associated with the initial receptor complex. The accumulation of critical signaling molecules in specific sites has been proposed to enhance their reaction kinetics towards certain responses (Sang et al., 2022).

Protein condensation can be translated as a positive feedback mechanism, primarily due to the way it can amplify and stabilize specific biochemical states in cells. In positive feedback, the output of a reaction enhances its own production (Mitrophanov and Groisman, 2008). Similarly, condensation can be perceived as a self-reinforcing loop in which, after the initial nucleation event, the condensed state promotes the phenomenon further. McAffee et al. (2022) reported that, once condensation has been initiated, positive feedback promotes LAT phosphorylation and characterized this as an additional layer of kinetic proofreading for antigen discrimination.

In our model, the effect of positive feedback is captured in the behavior of the condensed RP_4 and the signaling product SP. The significant increase of the latter, due to the presence of the former, further enables productive downstream signaling and functional discrimination. Here, we could quantitatively and explicitly show that incorporating positive feedback into the kinetic proofreading model notably enhances both the signal and specificity of our system (Figures 3.3 and 3.4). Our observations align with the theoretical understanding that positive feedback mechanisms can amplify biological signals and improve the accuracy of cellular responses. For T-cell mediated responses, this means a more robust and precise activation mechanism, capable of effectively distinguishing between different stimuli and responding appropriately.

Perspectives

In the McKeithan kinetic proofreading model (1995) increased specificity depends only on the specific rate constants associated with each molecular interaction, without taking into consideration the presence of other molecules. Although it is extremely appealing due to its simple mathematical formulation, it is not considered ideal when biochemical networks are associated with the condensation of molecules (referred to as aggregation in Coward, Germain, Altan-Bonnet, 2010).

Our simplistic description of the biochemical network of T-cell activation, which mainly relies on McKeithan's framework, fails to capture the importance of multiple interactions and a series of modifications that certain components undergo. Rule-based modeling (Hlavacek et al., 2006) incorporates the necessary complexity of signaling cascades in rules that connect individual elements in elaborate ways. Multiple packages, like PySB, have been developed to illustrate multi-protein systems with higher transparency and accuracy and have already been implemented in our biological context of interest (Ganti et al., 2020). Because these models require detailed quantitative data that may not be available in many cases, alternative approaches based on logical (aka Boolean) modeling have also been used to model T-cell differentiation. Logical models do not require detailed quantitative measurements but rather allow the development of complex qualitative networks and several have been developed to investigate T-cell activation (Wang et al. 2012). There are now numerous modeling approaches and platforms available and a challenge for the future will be the coordination of these diverse methodologies into one consistent model of T-cell activation.

5 Appendix: Rate Constants and Initial Conditions

Table 1: *Rate constants and initial concentrations for model 0 presented in Fig. 3.1*

Species	**Total Concentration (number of molecules)**
$[R]_T$	30000
$[L_f]_T$	1 or 10
$[L_s]_T$	10000
$[S]_T$	10000
Rate Constant	**Value**
k_{on}	0.005 s^{-1}
k_f	0.2 s^{-1}
k_s	2.0 s^{-1}
k_p	0.1 s^{-1}
k_{off}	0.05 s^{-1}
$k_{on,s}$	0.1 (mol * s^{-1})
$k_{off,s}$	0.05 s^{-1}

The numerical values presented in Table 1 are used for the numerical integration of the ODEs describing model 0 (see subsection 2.1, page 11). Integrating the system of equations with the provided set of values results in the time evolution of the network as captured in Figure 3.1.

The numerical values listed in Table 2 are utilized for the numerical integration of the ODEs representing model 1 (refer to subsection 2.2, page 14). Using these values to integrate the system of equations yields the time evolution of the network, as depicted in Figure 3.2.

Table 2: *Rate constants and initial concentrations for model 1 presented in Fig. 3.2*

Species	**Total Concentration (number of molecules)**
$[R]_T$	30000
$[L_f]_T$	1 or 10
$[L_s]_T$	10000
$[S]_T$	10000
Rate Constant	**Value**
k_{on}	0.005 s^{-1}
k_f	0.2 s^{-1}
k_s	2.0 s^{-1}
k_p	0.1 s^{-1}
k_{off}	0.05 s^{-1}
$k_{on,s}$	0.1 (mol * s^{-1})
$k_{off,s}$	0.05 s^{-1}
$k_{on,s}^{cond}$	0.2 s^{-1}
$k_{off,s}^{cond}$	0.01 s^{-1}
k_p^{free}	0.1 s^{-1}
k_p^{cond}	1.0 s^{-1}
k_c	0.1 s^{-1}
k_c^{-1}	100 s^{-1}

6 References

Balagopalan, L., Raychaudhuri, K., Samelson, L. E. (2021). Microclusters as T Cell Signaling Hubs: Structure, Kinetics, and Regulation. Frontiers in Cell and Developmental Biology, 8, 608530.

Balagopalan, L., Kortum, R. L., Coussens, N. P., Barr, V. A., Samelson, L. E. (2015). The linker for activation of T Cells (LAT) signaling hub: From signaling complexes to microclusters. JBC, 290 (44), 26422–26429.

Banani, S. F., Lee, H. O., Hyman, A. A., Rosen, M. K. (2017). Biomolecular condensates: Organizers of cellular biochemistry. Nature Reviews Molecular Cell Biology, 18(5), 285–298.

Boeger, H. (2022). Kinetic Proofreading. Annu. Rev. Biochem., 91, 423–47.

Cai, D., Feliciano, D., Dong, P. et al. (2019). Phase separation of YAP reorganizes genome topology for long-term YAP target gene expression. Nat Cell Biol, 21, 1578–1589.

Courtney, A.H., Lo, W.L., Weiss, A. (2018). TCR Signaling: Mechanisms of Initiation and Propagation. Trends in Biochemical Sciences, 43(2), 108–123.

Ganti, R.S., Lo, W.-L., McAffee, D.B., Groves, J.T., Weiss, A., Chakraborty, A.K. (2020). How the T cell signaling network processes information to discriminate between self and agonist ligands. PNAS, 117(42), 26020-26030.

Hyman, A.A., Weber, C.A., Jülicher, F. (2014). Liquid-liquid phase separation in biology. Annu Rev Cell Dev Biol.,30, 39-58.

Hlavacek, W.S., Faeder, J.R., Blinov, M.L., Posner, R.G., Hucka, M., Fontana, W. (2006). Rules for modeling signal-transduction systems. Sci STKE2006:re6.

Irvine, D., Purbhoo, M., Krogsgaard, M., M., Davis. (2002). Direct observation of ligand recognition by T cells. Nature, 419, 845-849.

Lo, W.L., Weiss, A. (2021). Adapting T Cell Receptor Ligand Discrimination Capability via LAT. Frontiers in Immunology, 12, 673196.

Malissen, B., Bongrand, P. (2015). Early T cell activation: Integrating biochemical, structural, and biophysical cues. Annual Review of Immunology, 33, 539–561.

McAffee, D.B., O'Dair, M.K, Lin, J.J., Low-Nam, S.T., Wilhelm, K.B., Kim, S., Morita, S., Groves, J.T. (2022). Discrete LAT condensates encode antigen information from single pMHC:TCR binding events. Nat Commun. 13(1):7446.

McKeithan, T.W. (1995). Kinetic proofreading in T-cell receptor signal transduction. Proc Natl Acad Sci USA, 92, 5042-5046.

Mitrophanov, A.Y., Groisman, E.A. (2008). Positive feedback in cellular control systems. Bioessays, 30(6), 542-55.

Nag, A., Monine, M.I., Faeder, J.R., Goldstein, B. (2009). Aggregation of membrane proteins by cytosolic cross-linkers: Theory and simulation of the LAT-Grb2-SOS1 system. Biophys. J., 96, 2604–2623 .

Sang, D., Shu, T., Pantoja, C. F., Ibáñez de Opakua, A., Zweckstetter, M., Holt, L. J. (2022). Condensed-phase signaling can expand kinase specificity and respond to macromolecular crowding. Molecular Cell, 82(19), 3693-3711.

Söding, J., Zwicker, D., Sohrabi-Jahromi, S., Boehning, M., Kirschbaum, J. (2020). Mechanisms for Active Regulation of Biomolecular Condensates. Trends in Cell Biology, 30(1), 4–14.

Su, X., Ditlev, J. A., Hui, E., Xing, W., Banjade, S., Okrut, J., King, D. S., Taunton, J., Rosen, M. K., Vale, R. D. (2016). Phase separation of signaling molecules promotes T cell receptor signal transduction. Science, 352(6285), 595–599.

Voisinne, G., Locard-Paulet, M., Froment, C., Maturin, E., Menoita, M. G., Girard, L., Mellado, V., Burlet-Schiltz, O., Malissen, B., Gonzalez de Peredo, A., Roncagalli, R. (2022). Kinetic proofreading through the multi-step activation of the ZAP70 kinase underlies early T cell ligand discrimination. Nature Immunology, 23(9), 1355–1364.

Wang, R.S., Saadatpour, A., Albert, R. (2012). Boolean modeling in systems biology: an overview of methodology and applications. Physical biology. 9:055001.

Wertek, F., Xu, C. (2014). Digital response in T cells: to be or not to be. Cell Res.24(3):265-6.

Wucherpfennig, K.W., Gagnon, E., Call, M.J., Huseby, E.S., Call, M.E. (2010). Structural biology of the T-cell receptor: insights into receptor assembly, ligand recognition, and initiation of signaling. Cold Spring Harb. Perspect. Biol., 2, a005140.