Integrative metamodeling of early T-cell antigen receptor (TCR) signaling accounts for nanoscale TCR activation patterns

Keywords: T cell, immunological synapse, Bayesian metamodeling, T cell activation, kinetic segregation, LCK, T cell receptor, CD45, peptide-MHC

# Abstract (<250 words)

T cells trigger highly specific and sensitive effector responses against foreign pathogens by recognizing cognate antigens presented by Antigen-presenting cells. Various microscopic processes that give rise to rapid and robust TCR activation have been proposed, but they each individually fail to capture all critical data related to TCR activation. Here, we apply Bayesian metamodeling for integrating multiple aspects of TCR triggering from fragmented input models of varying representations and scales. Specifically, we integrate models of the organization, activity, interactions, and dynamics of the signaling molecules TCR, CD45 and Lck, and of membrane shape and elasticity. Inputs include physical models, Monte-Carlo simulations, single molecule localization microscopy and published results. Our metamodel accounts for nanoscale dynamic patterns of TCR activation that could not be accounted for by the individual partial models. The scalability and flexibility of our metamodeling approach can serve to iteratively expand its predictive power for T cell activation and other biological systems.

# Significance Statement (<120 words)

T cell signaling is determined by a complex interplay of multiple microscopic processes. Metamodelling can integrate fragmented models of such processes and account for yet unexplained nanoscale activation patterns.

Introduction

T cells mount an immune response by recognizing foreign cognate antigens, presented on the surface of antigen-presenting cells (APCs). For that, the T cells form a dynamic interface with the APCs, known as the immune synapse (IS). Within this synapse signaling molecules dynamically organize to facilitate signal activation and regulation [PMID: 21179118].

Recently, multiple nanoscale patterns of these molecules have been observed, including clustering of molecules, and the segregation of receptors from glycoproteins (such as CD45). The abundant CD45 glycoproteins are phosphatases and can quench the TCR signal. Their physical separation from TCRs has been proposed and then shown to promote TCR activation (a model known as ‘kinetic segregation’ [DOI: 10.1038/ni1389]). These patterns have been shown to affect TCR signaling.

Multiple additional models have been proposed to account for the various aspects of T cell activation. 5. Such mechanisms include receptor clustering 6,7, conformational changes of receptor chains 8,9, dynamic formation of signaling complexes 10,11, cooperativity in triggering within nanoclusters 12,13, physical segregation of glycoprotein-phosphatases occurring in early tight contacts with APCs 14,15 where signalling molecules are enriched 16, effects of cell topography 17, etc.

Still, each of these models cannot capture the entirety of data related to TCR-dependent T cell activation [Ref]. The integration of such models have been proposed as a necessary step for achieving comprehensive and predictive understanding TCR activation [PMID: 21127503].

Arguably,integrating all available data sources on T-cell activation and its underlying mechanisms is impractical using existing data integration approaches. However, it is possible to integrate partial data subsets, and use these data subsets to inform partial models of narrower aspects of T-cell activation (components, subsystems, functions) [[3–16]](https://paperpile.com/c/G64qnd/wWBFZ+FVhg5+gEpPz+gFDLy+cnEhQ+XtLSQ+NsxBm+3HfWJ+zLDv2+y0mPD+NoxYt+OdYrN+7tFz5+SJJl7). Thus, we propose that Big Data integration can be divided-and-conquered if we shift the focus from direct data integration to model integration. For that, we turn to Bayesian metamodeling [[17]](https://paperpile.com/c/G64qnd/Y5bI5) (Fig. 2). Through this approach data can be collected from diverse data sources. Following data collection, the data is broken into **partial data subsets;** the partial data subsets are used as input for constructing **partial models**; each of these partial models is converted to a **probabilistic surrogate model** over some variables of the corresponding partial models (e.g. using a Bayesian network or a generative deep-learning model); the surrogate models are then coupled through imposition of **statistical coupling restraints** resulting in a a **Bayesian metamodel** - a single **joint probability distribution function (PDF)** over variables from all surrogate models. Finally, hyperparameters of partial input models are updated using their posterior estimates in the Bayesian metamodel. Critically, the output metamodel integrates all data used to inform any of the partial models.

To make informed decisions, T cells integrate multiple cues, which are sensed by specific receptors and induce signaling pathways. The detailed mechanisms by which T cells process and integrate multiple signals into accurate and reliable cellular responses remain incompletely understood. For example, in the context of pathogen recognition, T cells were shown to physically probe the surface of antigen-presenting cells (APCs) for cognate foreign antigens through their T-cell antigen receptors (TCRs). **However, the simple affinity of TCR-antigen interactions cannot explain the robustness of T-cell decision-making.** Detailed studies of TCR-dependent signaling using diverse experimental and theoretical methods have resulted in evidence for multiple possible mechanisms, which likely occur simultaneously to modulate the cell response. Such mechanisms pertain to diverse temporal and spatial scales, and depend on specific context and environment, such as **micro- to nano-scale clustering of receptors and signaling molecules,** **cooperativity,** mechano-sensing, and more. The integration of these partial mechanisms into a unified view of T-cell recognition is critically missing, leading to a lack of comprehensive and predictive understanding of such critical T-cell decision-making.

Here, we created a metamodel of T-cell recognition, based on the following models, using the following assumptions and free parameters. Currently, the first input model describes the major aspects of the kinetic segregation model, and combines them with information on the activity of LCK (TBD). If integrated correctly, they are expected to produce the observed pattern of phosphorylation and bull’s-eye spatial densities of TCRs and CD45s.

Our metamodel accounts for nanoscale dynamic patterns of TCR activation that could not be accounted for by the partial models. The scalability and flexibility of this model can serve to iteratively expand its predictive power for T cell activation and other biological systems.

Construction of the metamodel

Models are the central units of scientific theorizing and are necessary tools for an integrative understanding of heterogeneous data types. Our goal is to construct a unified model of T-cell decision-making, informed by pertinent data and prior knowledge. We create this model using Bayesian metamodeling, a novel divide-and-conquer approach for integrative whole-cell modeling. Using this formulation, we break the overall modeling problem into smaller and thus more-tractable modeling tasks. This includes: (A) collecting experimental and theoretical data using state-of-the-art imaging and complementary methods; (B) constructing multiple models describing aspects of T-cell decision-making; (C) integrating these partial models into a unified, quantitative, predictive, testable, and scalable metamodel. We then iterate through steps A-to-C to expand and improve both our experimental results, the partial models, and the unified metamodel (see Planned Activities for A-C and proof-of-concept below).

(A) Systematic and model-driven data collection across scales (Fig. 2; outer ring). We collect data on T cells and follow their decisions related to antigen recognition and differentiation. For that we will employ a broad set of imaging and other modalities, utilizing the diverse expertise of the center and committee members, as well as collaborators, as outlined in detail in Planned Activities.

(B) Constructing individual models of separate aspects of T-cell decision-making (Fig. 2; middle ring). Models, based on our data and literature, will be described using different datasets, representations, and pertaining to different parts and scales of the cell. Due to the scalable nature of our modeling approach, the list of integrated models is expected to expand iteratively over time as our scientific network expands and more experimental data becomes available.

(C) Assembling a metamodel of cellular decision-making (Fig. 2; inner ring). We will convert each partial model into a unified statistical representation, namely a probability density function (PDF). For this, we will use appropriate statistical modeling and machine learning approaches, including probabilistic graphical models and generative deep neural networks. We will then assemble the partial models in their unified statistical representations into a comprehensive metamodel of the cell, and harmonize them with each other using principles of Bayesian statistics. Importantly, the partial models can be constructed and computed independently; and can be integrated regardless of their scales by relying on statistical relations. This facilitates the sharing of data, resources, expertise and models by network members, and maximizes the modeling accuracy, precision, completeness, and scalability

The specific models

Our meta-model integrates multiple partial models. Each partial model captures specific mechanisms that contribute to signaling downstream the TCR. Specifically, we model the surface of T cells as they undergo an interaction with an antigen presenting cell (APC), forming an immune synapse (IS). Note that our focus is on the dynamic organization of molecules within the T cells. The partial models include the TCRs, glycoproteins (esp. CD45), key kinases (Lck), and the T cell membrane. We now describe each of the models and related results that serve as inputs to the metamodeling briefly, elaborate descriptions are provided in SI.

**Model 1** computes the spatiotemporal patterning of a population of TCR and CD45 molecules embedded in the plasma membrane of T cells during T-cell activation through the IS. The model aims to recapitulate imaging measurements of TCR and CD45 patterning when TCRs interact with peptide-MHC molecules on APCs or with molecular mimic, leading to TCR triggering [(Razvag et al. 2018; Neve-Oz et al. 2018; Chang et al. 2016)](https://paperpile.com/c/CRnYld/fjMT+H6vM+9Ity). The formation of these interfaces leads to reorganization of surface molecules, changes in the topography of the plasma membrane, and TCR activation. The tight adhesion of the two cells at the immune synapse (or the adhesion of the T cell to the mimic coverslip) causes physical segregation between the TCRs and CD45. This segregation occurs due to the smaller size of the molecular complex of the TCR-pMHC relative to the bigger size of the bulky CD45 (and other glycoproteins). Since CD45 is also a phosphatase, it has been suggested to dynamically quench the TCR signal through its dephosphorylation. Thus, the physical segregation of these molecules may promote TCR activation - a model known as ‘kinetic segregation’ [Ref Van-der Merwe+Davis].

Here, we specifically used Reaction-diffusion Markov-Chain Monte-Carlo (MCMC) simulations to capture the positions of these molecules at the tight interface. For this model, we used InterCells - our previously described computational model and simulation [(Neve-Oz et al. 2018)](https://paperpile.com/c/CRnYld/H6vM).

Briefly, our model captures the assumptions of the Kinetic Segregation model, but it does not account for TCR phosphorylation by Lck and downstream TCR spatiotemporal signaling.

**Model 2** describes the spatial patterning of a population of active Lck molecules, indicated as Lck\* [Refs Acuto, Gaus]. It computes Lck\* distribution given the density of CD45 molecules by using Monte-Carlo simulations. The free parameters for the model include the diffusion coefficient of Lck and the probability of spontaneous decay of Lck\* back to its inactive state. The input variables include mean (μ) and standard deviations (σ) of CD45 distribution. The output variables include probabilistic parameters that describe the radial distribution of steady state Lck\*.

Thus - this model captures Lck activation state at the immune synapse, but not the spatiotemporal organization of CD45 and the TCR.

**Model 3** computes the two-dimensional spatial patterning of phosphorylated TCR molecules (indicated as TCR\*) across the plasma membrane, as a function of the spatial distribution of TCR, CD45 and Lck\* molecules.

## Results

1. None of the three models captures pTCR patterning at initial contact
2. Prior information from KS and LCK models is sufficient to restore the observed pTCR pattern (roughy)
3. Metamodeling using surrogate for KS, LCK and pTCR\* models formalizes this observation, providing a quantitative probabilistic framework (including confidence intervals etc.) to couple the three models, contextualize them
4. Confronting metamodel with experimental data to infer the most probable model parameters for KS, LCK and pTCR\* models

## Discussion

T cell activation is a complex process that likely incorporates multiple microscopic processes. While the processes are inherently stochastic, there seems to be a high level of spatiotemporal organization and orchestration of the events that give rise to TCR-dependent signaling and cell activation. Here, we focused on the earliest events of TCR signaling. Specifically we attempted to integrate multiple partial models to account for previously unexplained patterns of TCR activation in early contacts. We employed a new approach to modeling, called metamodelling, that can seamlessly integrate models of various origin and spatiotemporal scales. We first showed that only the integration of multiple partial models can capture the previously observed phosphorylation pattern of TCR in the surroundings of initial tight contacts of the T cell at the IS. We then confronted the model with microscopy data and used the metamodel to infer the most probable parameters of the individual models. Such parameters include: ….

Fernandes et al have proposed that the KS model in early contacts should produce a uniform activation pattern of TCR The In model3 a CD45. Still, imaging by Razvag et al [PMID: 31825832] has shown the TCR activation is enriched at the periphery of early and tight T cells contacts with activating coverslips. Our metamodeling shows that the integration of multiple partial models was essential for capturing this patterning of TCR activation at the early contacts.

Spatiotemporal organization of molecules determines the rates of local interactions. Thus, such interactions are a common regulating mechanism for a wide range of cellular processes, including cell sensing, signaling pathways, metabolic networks, transcription, translation, and more. Spatiotemporal dependencies and stochasticisity (e.g. of molecular diffusion, interactions, etc) directly translate to system complexity. This complexity becomes daunting, esp. when one tries to gain fundamental understanding of the system working. Metamodeling has a unique and natural capability to describe complex systems in a tractable and scalable fashion. Thus, while our metamodelling approach was applied here to a concrete example of TCR activation, it could naturally grow to account for additional cellular systems.

Materials and Methods

The software, input files, and example output files for the present work are available at <https://github.cs.huji.ac.il/ravehb-lab/immune-synapse-metamodeling>. The metamodel was implemented using the PyMC3 package in Python{REF; ,https://docs.pymc.io/} (tested on version XX.XX; Python vXX.XX). For an outline of the approach, see Results. The specifications of the three input models to metamodeling and the technical details of their conversion to surrogate models and their coupling through metamodeling are described in detail in the SI Appendix: Supplementary Text 1.

## References

Fernandes, <https://doi.org/10.1073/pnas.1817255116>

## 6. Schamel, W. W. A. A. *et al.* Coexistence of multivalent and monovalent TCRs explains high sensitivity and wide range of response. *J. Exp. Med.* **202,** 493–503 (2005).

## 7. Lillemeier, B. F. *et al.* TCR and Lat are expressed on separate protein islands on T cell membranes and concatenate during activation. *Nat. Immunol.* **11,** 90–96 (2010).

## 8. Blanco, R., Alarcón, B. & Alarcon, B. TCR nanoclusters as the framework for transmission of conformational changes and cooperativity. *Front. Immunol.* **3,** 1–7 (2012).

## 9. Xu, C. *et al.* Regulation of T cell receptor activation by dynamic membrane binding of the CD3epsilon cytoplasmic tyrosine-based motif. *Cell* **135,** 702–713 (2008).

## 10. Sherman, E. *et al.* Hierarchical nanostructure and synergy of multimolecular signalling complexes. *Nat. Commun.* **7,** 12161 (2016).

## 11. Barr, V. A. *et al.* Development of nanoscale structure in LAT-based signaling complexes. *J. Cell Sci.* **Accepted f,** (2016).

## 12. Neve-Oz, Y., Razvag, Y., Sajman, J. & Sherman, E. Mechanisms of localized activation of the T cell antigen receptor inside clusters. *Biochim. Biophys. Acta - Mol. Cell Res.* **1853,** 810–821 (2015).

## 13. Pageon, S. V *et al.* Functional role of T-cell receptor nanoclusters in signal initiation and antigen discrimination. *Proc. Natl. Acad. Sci. U. S. A.* **113,** (2016).

## 14. Chang, V. T. *et al.* Initiation of T cell signaling by CD45 segregation at ‘close contacts’. *Nat. Immunol.* **17,** (2016).

## 15. Razvag, Y., Neve-Oz, Y., Sajman, J., Reches, M. & Sherman, E. Nanoscale kinetic segregation of TCR and CD45 in engaged microvilli facilitates early T cell activation. *Nat. Commun.* **9,** 732 (2018).

## 16. Jung, Y. *et al.* Three-dimensional localization of T-cell receptors in relation to microvilli using a combination of superresolution microscopies. *Proc. Natl. Acad. Sci. U. S. A.* **113,** (2016).

## 17. Razvag, Y., Neve-Oz, Y., Sherman, E. & Reches, M. Nanoscale Topography-Rigidity Correlation at the Surface of T Cells. *ACS Nano* **13,** 346–356 (2019).

## 

## Figures

**Fig. 1? - Intro**: biological/biophysical question - set up the challenge. Describe models (briefly) and data sources

**Fig. 2 - the approach** - combining three models, based on known literature and reasonable assumptions; learning of surrogate models

**Fig. 3 Results** (construction of model - everything before validation)

1. Building of a surrogate model (training)

**Fig. 4 - Validation**,

(A) model predictions, etc. - comparison to validation data

(B) a new integrated model of T cell molecular patterning

(C) Testable prospective prediction - what/if (try to connect to open questions in the field)

## Supporting Figures

Fig. S1A,B - Model 1 - A) Segregation of TCR and CD45 B) Conversion to surrogate

Fig. S2A,B - Model 2 - A) (CD45-dependent) Lck activity B) Conversion to surrogate

Fig. S3A,B - Model 3 - A) (Spatial patterns of) TCR activation B) Conversion to surrogate