

Low-dimensional models of immune sensing and signaling

François Bourassa

Doctor of Philosophy

Department of Physics
McGill University, Montreal

June 7, 2024

A thesis submitted to McGill University in partial fulfillment of the requirements
of the degree of Doctor of Philosophy

© François Bourassa, 2024

Pour Marcelle, Alphonse, Luce, Antonin, Irène, Émilie, Louise, Benoît et Marie, mon équipe.

Contents

List of Figures	v
List of Tables	viii
List of Abbreviations	ix
Abstract	xi
Résumé	xii
Acknowledgements	xiii
Contribution to original knowledge	xvi
Contribution of Authors	xviii
Introduction	1
1 Review: chemical sensing and immunity	4
1.1 Biophysics of chemical sensing in living systems	4
1.2 Chemical sensing challenges in immunity	16
1.3 Antigen recognition mechanisms in T cells	26
1.4 Low-dimensional models in biophysics and immunity	34
2 Antigen encoding of T cell activation from cytokine dynamics	40
2.1 Introduction: the cytokine code of T cells	41
2.2 Experimental platform and data processing	45
2.3 Latent space discovery with machine learning	52
2.4 Robustness of the antigen encoding latent space	55
2.5 Cytokine time course reconstruction	61
2.6 Dynamical latent space models and parameter correlations	64
2.7 Channel capacity of antigen encoding in the latent space	82
2.8 Universality of antigen encoding	98
2.9 Summary of key insights from antigen encoding	104
3 Revisiting antagonism in T cell receptors	106
3.1 Timeliness of revisiting TCR activation models	107
3.2 Antigen encoding by altered TCRs	113

3.3	Revisiting the classical AKPR model	118
3.4	Initial AKPR model for TCRs with fewer ITAMs	125
3.5	Revising the AKPR model of TCR/TCR antagonism	133
3.6	Summary of progress on TCR antigen mixtures	151
4	Antagonism in chimeric antigen receptor T cells	152
4.1	Introduction: receptor interactions in engineered T cells	153
4.2	Experimental characterization of TCR/CAR interactions	156
4.3	Modelling TCR/CAR antagonism	160
4.4	TCR/CAR antagonism in various immunological settings	177
4.5	Outlook on cross-receptor interactions in CAR T cells	191
	Discussion and conclusion	193
	Bibliography	201
	Appendices	225
A	Examples of biological chemical sensing	225
A.1	Bacterial chemotaxis	225
A.2	Cell fate decisions in embryos	225
A.3	Olfaction	227
B	Supplementary results of universal antigen encoding	228
B.1	Filtering of large background noise in diverse immunological settings	228
B.2	Supplementary figures related to antigen encoding	231
C	Existence and uniqueness of the solution for bound receptors	236
D	Parameter estimation methods	240
D.1	Calibration of receptor and ligand abundances	240
D.2	Peptide binding time conversion	242
D.3	Parameter values in the initial AKPR model	243
D.4	Fixed parameter values	243
D.5	Overview of the parameter estimation strategy and of the generation of predictions .	244
D.6	Calculation and statistics of antagonism fold-change	246
D.7	Likelihood function and prior distribution	251
D.8	MCMC simulation details	251
D.9	Predictions in other T cell types	253
D.10	Antagonism predictions from TCR antigen libraries	255

List of Figures

1.1	Physical limit to molecule counting due to diffusion	6
1.2	Distribution optimizing information on concentration	12
1.3	Optimal immune strategies depending on pathogen statistics	18
1.4	Excitable dynamics in a toy model of immune responses	21
1.5	Bifurcations of chronic infections in the toy model	21
1.6	Schematic of the TCR structure	23
1.7	The TCR signalosome	28
1.8	Biochemical network of the classical KPR model	29
1.9	Illustration of the antigen quality discrimination problem	32
1.10	Biochemical network of the simplest AKPR model	33
2.1	Robotic platform to quantify T cell activation dynamics	44
2.2	Immunotron robotic platform	46
2.3	Time series processing steps	48
2.4	Antigen encoding according to quality in a cytokine latent space	54
2.5	Dimension of the cytokine manifold and linear reconstruction procedure	56
2.6	Supervised vs unsupervised learning	59
2.7	Impact of excluding IL-4 and IL-10	60
2.8	Effect of preprocessing on the classifier	62
2.9	Cytokine reconstruction results with linear least-squares regression	63
2.10	Cytokine reconstruction results with non-linear regression	65
2.11	Residuals of the cytokine reconstruction	66
2.12	Classes of T cell activation in antigen encoding	70
2.13	Fits of the constant velocity model	71
2.14	Graphical explanation of Kendall tau distance order accuracy metric	73
2.15	Parameter v_0 ranks antigens according to quality	74
2.16	Fits of the force model with matching	77
2.17	Comparison of the quality of fits for the two latent space models	78
2.18	Residuals of the various latent space models	79
2.19	Latent space modelling and reconstruction provide a model for cytokine dynamics .	80
2.20	Generating synthetic cytokine time courses from the latent space model	81
2.21	Procedure to estimate the channel capacity of cytokine dynamics for peptide quality	84
2.22	Empirical and fitted multivariate normal distributions of model parameters	86
2.23	Comparison of OT-1 EC50s	89
2.24	Interpolation of multivariate normal distributions in the model parameter space . .	90
2.25	Blahut-Arimoto algorithm to derive the cytokine channel capacity	92
2.26	Neural network weights interpretation	96
2.27	Reconstructed cytokine trajectories for the 6 prototypical antigen classes	97
2.28	Universality of antigen encoding across immune settings	99
2.29	Antigen encoding with different antigen-presenting settings	100

2.30	Antigen encoding with different tumor settings	101
2.31	Latent space projections of cytokines under varied molecular perturbations	103
3.1	Frequency distribution of peptide MHCs on APCs	107
3.2	Example of TCR antagonism	108
3.3	Illustration of ITAM mutations in the 6F TCR	111
3.4	Mutual information analysis of 6F(i) T cells	114
3.5	Mutual information analysis of 6F T cells	115
3.6	Latent space analysis of 6F T cells	116
3.7	Reaction network of the classical AKPR model	121
3.8	Response curves of the classical AKPR model	124
3.9	Classical AKPR model for 6F TCRs	126
3.10	Approximate pulse – L map for early studies of 6F TCRs	129
3.11	Model predictions of antagonism in 6Y and 6F T cells	131
3.12	Modeling TCR responses to antigen mixtures to quantify receptor crosstalk	134
3.13	Quantification of surface molecule abundances and affinities	136
3.14	MCMC autocorrelation for the classical AKPR model fitting	140
3.15	Biochemical network of the revised AKPR model	142
3.16	Biochemical implementation of the nonlinear inhibitory module	145
3.17	MCMC autocorrelation for the revised AKPR model fitting	148
3.18	TCR response curves of the revised AKPR model	149
3.19	MCMC parameter estimation for the revised model of 6F TCRs	150
4.1	Chimeric antigen receptor (CAR) T cell therapy	154
4.2	Differential modulation of CAR activity over a spectrum of TCR antigens	157
4.3	Modeling TCR/CAR crosstalk in CAR T cells	162
4.4	Biochemical network of the classical AKPR model extended to CAR T cells	163
4.5	Typical curves of TCR/CAR antagonism in the classical model	166
4.6	Biochemical network of the revised AKPR model extended to CAR T cells	167
4.7	Typical curves of TCR/CAR antagonism in the revised model	169
4.8	MCMC autocorrelation for TCR/CAR parameter estimation	172
4.9	Generation of TCR/CAR model predictions	173
4.10	Illustration of TCR/CAR antagonism or enhancement in the model	174
4.11	TCR/CAR crosstalk shapes CAR T cell responses in solid tumors	179
4.12	Prediction of <i>in vivo</i> survival from the cross-receptor model	182
4.13	Universality of TCR-driven antagonism of CARs	185
4.14	Antagonism of human CAR T cells by a co-transduced TCR	187
4.15	AEBS CAR T cell concept	189
4.16	Experimental proof of the AEBS concept	190
5.1	Cytokine dose response curves to OT-1 peptides	198
B.1	Cytokine time series in the human TCR experiment	229
B.2	Signal-to-noise ratio of each cytokine	230
B.3	IL-2 time series background filtering based on IFN- γ	231
B.4	Latent space model fits for more datasets	232
B.5	Latent space concentration model fits for more datasets	233

B.6	Antigen encoding with effector T cells	234
B.7	Latent space and order accuracy for different numbers of OT-1 T cells	235
C.1	The three possible arrangements of real roots for the number of bound receptors . .	238
D.1	Correcting missing EC_{50} s based on maximal activation	256

List of Tables

3.1	Definition of classical AKPR model parameters.	123
3.2	Peptide – τ map for early studies of 6F TCRs	128
3.3	Parameter values in our initial model of 6F and 6Y TCRs	130
3.4	Definition of model parameters.	144
A.1	Pairs of cell surface receptor and secreted ligand families	226
D.1	Surface molecule abundances.	241
D.2	Information on MCMC simulations.	247
D.3	Model parameter constraints.	248
D.4	Estimated model parameter values.	249

List of Abbreviations

AEBS	Antagonism-enforced braking system
Ag	Antigen or agonist
AKPR	Adaptive kinetic proofreading
APC	Antigen presenting cell
ATP	Adenosine triphosphate
B16	A melanoma cell line
BEAS-2B	A human lung epithelial cell line
BMP	Bone morphogenetic protein
CAR	Chimeric antigen receptor
CD	Cluster of differentiation (<i>e.g.</i> , CD3)
Csk	C-terminal c-Src kinase
CTLA-4	Cytotoxic T-lymphocyte associated protein 4
DNA	Deoxyribonucleic acid
E2aPBX	An acute lymphoblastic leukemia mouse cell line
ERK	Extracellular signal-regulated kinase
HER2	Human epidermal growth factor receptor 2
HHAT	Hedgehog acyltransferase
ITAM	Immunoreceptor tyrosine-based activation motif
IL	Interleukin, <i>e.g.</i> , IL-2, IL-4, etc.
IFN	Interferon, <i>e.g.</i> , IFN- γ
KPR	Kinetic proofreading
LAT	Linker for activation of T cells
Lck	Lymphocyte-specific protein tyrosine kinase
LOD	Lower limit of detection
MAPK	Mitogen-activated protein kinase
MCMC	Markov Chain Monte Carlo
MHC	Major histocompatibility complex; pMHC: peptide-MHC

MI	Mutual information
NCI	National Cancer Institute
NIH	National Institutes of Health
NICHD	Eunice Kennedy Shriver National Institute of Child Health and Human Development
NF-κB	Nuclear factor-kappa B
NY-ESO	New York esophageal squamous cell carcinoma 1
ODE	Ordinary differential equation
OT-1	Ovalbumin-specific TCR transgenic mouse line 1
OTOT	On-target, off-tumor (toxicity)
OVA	Ovalbumin
PCA	Principal component analysis
PD-1	Programmed cell death protein 1
PD-L1	Programmed cell death ligand 1
PI3K	Phosphoinositide 3-kinase
PLC-γ	Phospholipase C, gamma
SHP-1	Src homology region 2 domain-containing phosphatase-1
SI	Supplementary information (or supplemental information)
SVD	Singular value decomposition
TCR	T cell receptor
TNF	Tumor necrosis factor
ZAP-70	Zeta-associated protein 70

Abstract

T cells face high-dimensional chemical spaces. These white blood cells must sense fine differences within convoluted mixtures of similar antigens. Through complex T cell receptor (TCR) signaling, they transform these inputs into multivariate responses. Details of these mechanisms quickly become too complicated to formulate exhaustive models that are still useful. Instead, we should seek to uncover emergent physical laws regulating T cell responses and biological chemical sensing more generally. Such quantitative principles are however still lacking.

Working in close collaboration with experimental immunology labs, we developed low-dimensional theoretical models of T cell antigen sensing and receptor signaling. We studied the production of cytokines – small extracellular messenger proteins – by activated T cells. We found that high-dimensional cytokine dynamics could be represented in a two-dimensional space and parameterized with simple equations inspired from ballistic physics. We used information theory to quantify antigen encoding in this cytokine latent space, revealing a continuum of T cell responses. Building on the insight that these responses are mainly controlled by early sensing of antigen quality, we could improve adaptive kinetic proofreading models of TCR signaling to explain antagonism and enhancement effects in antigen mixtures. To account for all observed patterns when antigen qualities, quantities, and receptor phosphorylation sites are varied, we needed to simplify previous models, and then introduce a nonlinear inhibitory coupling between receptors. We could quantitatively fit the revised model to our collaborators’ data to estimate its parameters. Then, we generalized the model to T cells expressing chimeric antigen receptors (CAR) to target tumor cells in immunotherapy. We generated predictions for different receptor constructs, matching *in vitro* experiments; we could also approximately predict mouse *in vivo* treatment outcomes against tumors expressing TCR and CAR antigen mixtures. The model finally led us to design CAR T cells in which antagonism by weak TCR stimulation is optimized to protect healthy tissues from otherwise severe side effects. Our findings show that low-dimensional, physics-inspired models can be accurate enough to capture meaningful dynamics in biological sensing and signaling.

Résumé

Les lymphocytes T baignent dans des espaces chimiques à hautes dimensions. Ces globules blancs doivent détecter d'infimes différences entre des antigènes similaires et présentés en mélanges. La signalisation complexe de leurs récepteurs de cellules T (TCR) transforme ces stimuli en réponses multivariées. Un modèle exhaustif de ces mécanismes deviendrait rapidement trop labyrinthique pour être utile. Il serait préférable de tenter d'établir les lois physiques émergentes régissant la réponse des lymphocytes T et, plus généralement, la détection chimique biologique. De tels principes quantitatifs restent toutefois à formuler.

En étroite collaboration avec des laboratoires expérimentaux d'immunologie, nous avons développé des modèles à basse dimension décrivant la détection d'antigènes et la signalisation des récepteurs de lymphocytes T. Nous avons étudié la production de cytokines, de petites protéines messagères extracellulaires, par les lymphocytes T. Nous avons réduit la dynamique à haute dimension des cytokines à une représentation bidimensionnelle décrite par des équations inspirées de la physique balistique. Grâce à la théorie de l'information, nous avons quantifié la propriété d'encodage antigénique de cet espace latent des cytokines, révélant un continuum de réponses cellulaires. En constatant que ces réponses sont déterminées dès les premiers instants par la qualité des antigènes, nous avons pu améliorer les modèles de relecture cinétique adaptative des TCRs pour expliquer des effets d'antagonisme et d'amélioration de la réponse dans des mélanges d'antigènes. Pour expliquer les subtilités des effets produits par la variation de la qualité et de la quantité d'antigènes ainsi que des sites de phosphorylation des TCRs, nous avons dû simplifier les modèles préexistants, puis leur ajouter un module d'inhibition non-linéaire couplant les récepteurs. Nous avons pu estimer les paramètres du modèle amélioré en l'ajustant quantitativement aux données de nos collaborateurs. Par la suite, nous avons étendu le modèle amélioré aux lymphocytes T utilisés en immunothérapie, modifiés génétiquement pour exprimer des récepteurs antigéniques chimériques (CAR) leur permettant de cibler des tumeurs. Les prédictions du modèle correspondaient bien aux données *in vitro* pour divers types de récepteurs, et permettaient même d'anticiper les résultats de traitements par cellules CAR sur des souris ayant des tumeurs présentant des mélanges d'antigènes TCR et CAR. Le modèle nous a permis de concevoir des cellules CAR où l'antagonisme causé par une faible stimulation du TCR est optimisé pour protéger les tissus sains d'effets toxiques sévères. Nos résultats montrent que des modèles de basse dimension issus de la physique permettent de décrire avec précision les dynamiques de détection et de signalisation en biologie.

Acknowledgements

I first want to thank the Physics department staff members, Diane Koziol, Lauren Kay, Louise Decelles, Eddie Del Campo and others, for their unwavering efforts behind the scenes, even through a pandemic and a change of HR systems, among other challenges. Many thanks to Saverio Biunno, Philippe Mofavegh, and all other lab staff members for their help with setting up several physics demonstrations for school visits and public lectures through the years. Special thanks to Juan Gallego for maintaining the computational cluster on which most simulations in this thesis were performed, and for always finding a solution, within minutes, when my code would not run.

I acknowledge funding from the Fonds de recherche du Québec – Nature et technologie and the NSERC-CREATE Complex Dynamics Graduate Award Program, which made this thesis possible.

Thank you to Drs. Anmar Khadra, Martin Grant, and Nikolas Provatas for serving on my supervisory committee, and for having taught classes at McGill that proved crucial in my training.

I have been very lucky to collaborate with Drs. Gautam Reddy and Massimo Vergassola. The research we have been doing together on olfaction has not made it into this thesis, but I must thank them for their scientific input, their trust, their patience, and their help with the next steps of my career. I am also very thankful for my collaboration with Drs. Guillaume Gaud and Paul Love at the Eunice Kennedy Shriver National Institute of Child Health and Human Development (Bethesda, MD); they trusted me, a physicist, to jump into their ongoing project.

I feel blessed to collaborate closely with the lab of Dr. Naomi Taylor at the National Cancer Institute (NCI; Bethesda, MD), and in particular with Dr. Taisuke Kondo. Naomi, thank you for your openness to work with theorists, your attention to detail, and your valuable scientific and clinical perspective. Taisuke, thank you for your outstanding skills in the lab, your uncanny ability to materialize theoretical concepts *in vivo*, and your frankness. It is a rare chance for a physicist to have true collaborators offering such a window into actual therapeutical applications.

I am incredibly grateful to have been collaborating for nearly six years with the lab of Dr. Grégoire Altan-Bonnet at the NCI. Thank you to all members of the lab – Angela Lee, Emanuel Salazar-Cavazos, Madison Wahlsten and others – who have contributed to setting up cocultures and running the robot, as well as for their warm welcome every time I visited the NIH.

In Grégoire's lab, I must thank in particular Sooraj Achar, who developed the lab's robotic platform and performed so many experiments since we both started in 2018. I could not imagine a better or brighter collaborator and co-author. Sooraj, your astonishingly hard work and resilience have pushed me to carry out research diligently and thoroughly. I will miss our morning writing sessions on the TCR/CAR manuscript.

Je dois remercier tout aussi personnellement Grégoire. Grégoire, tu m'as continuellement poussé à penser plus loin tout en me témoignant ta confiance à chaque étape. Je te suis grandement reconnaissant de m'avoir ouvert les portes de ton labo pour plusieurs visites et un stage à distance (COVID-19 oblige), de m'avoir souvent conseillé et d'avoir soutenu ma candidature pour de nombreuses conférences et bourses.

Il me faut remercier encore davantage Paul François, mon directeur de recherche. Paul, je n'aurais pas pu souhaiter un meilleur superviseur que vous. Votre vision de l'interdisciplinarité entre physique et biologie a façonné la mienne. Votre optimisme et vos encouragements m'ont offert un parfait équilibre pour contrer ma propre tendance à voir tous les problèmes avant les bons côtés; merci pour votre patience, votre soutien et votre compréhension. Merci pour votre relecture de ma thèse. Rester dans votre groupe à Montréal pour mon doctorat fut l'une de mes plus judicieuses décisions jusqu'à présent.

I now want to thank all members of Paul's group through the years for their input, collegiality, and friendliness. Merci tout particulièrement à Félix Proulx-Giraldeau et Laurent Jutras-Dubé de m'avoir accueilli dans le groupe lors de ma première visite en 2017, de m'avoir montré le chemin depuis, et d'être devenus des amis. Merci à Victoria Mochulska pour son amitié, son intérêt pour le français et ses précieuses connaissances en biophysique. Merci à Christian et Nacer pour la camaraderie, les discussions scientifiques et non-scientifiques, et l'aide lors du déménagement vers l'UdeM. Thank you also to Adrien, Allen, Jimmy, Kaushik, Maya, Rabea, Yeeren, and many rotation and summer students for the discussions and memories. Thank you to Hannah Fronenberg, Rutherford 208 office colleague, for the help with outreach visits in French, the astrophysics perspective, and the tips on MCMC.

I want to thank especially Thomas Rademaker. Thomas, it was a real blast to work alongside you on the cytokines project and to visit Grégoire's lab together. You have been a model I tried to emulate since the start of my master's degree. I am grateful for our intellectual connection and our continued friendship. Thanks for welcoming Marie-Laurence and I to your new home in Australia; we hope to see you in person soon again.

Merci à mon ami Vincent pour les précieuses discussions et le contact avec le monde de la physique théorique en cosmologie. Merci à Philippe, Myriam, Noël, Pam, Lucas, Lia, Yvon et Danielle pour leur soutien et leur intérêt. Merci à Marie pour sa confiance absolue et pour les séjours prolongés chez elle, où j'ai réalisé plusieurs parties de cette thèse. Merci à Julie et Guillaume pour leur constante gentillesse et leur intérêt. Merci à Sophie et Charles-Olivier pour l'accueil toujours chaleureux à Joliette, les conseils et les fous rires. Merci à Rachel, Édouard, Charlotte, Gabriel, Annie et Arthémis d'incarner l'avenir. Merci, surtout, à mes parents, Denise et Paul, de m'avoir élevé avec tant d'amour et de dévouement, de m'avoir inculqué le sens du travail et du devoir, et de m'avoir toujours soutenu dans mes projets.

Enfin, merci par-dessus tout à mon épouse, Marie-Laurence. Merci de toujours m'appuyer dans mes projets malgré les sacrifices qu'ils impliquent. Je ne serais pas allé bien loin sans la joie et le soutien que tu m'apportes au quotidien. J'espère avoir pu te le rendre. J'ai hâte que nous poursuivions l'aventure ensemble au sud de la frontière.

Contribution to original knowledge

This thesis focuses on my original contributions to biological physics through three research publications in the field of quantitative immunology: one co-first authored and published in *Science* (*Antigen encoding*, [1]), one co-second authored and published in *Nature Immunology* (*CD3 ζ ITAMs*, [2]), and one co-first authored and under review (*TCR/CAR antagonism*, [3]). This thesis is presented in the traditional format because my colleague Thomas Rademaker included his contributions as co-first author of publication [1] in his PhD thesis [4]. Excerpts from the three immunology publications are presented as indented quotations in Computer Modern Sans font, with the source of the quote in parentheses at the end, *e.g.*, (*Antigen encoding*, [1]). The text in these excerpts is verbatim, except for a few ellipses (indicated by “[...]”) and rare minor edits (put in [brackets]), to harmonize some quotes with the rest of the thesis. Moreover, the numberings of figures, tables, sections, and references, have been adapted to this thesis. The next section details my contributions and those of my co-authors to these works. Here, I summarize how they represent original research scholarship.

Chapter 1 is a review of existing literature on chemical sensing in biology and on T cell immunology. Still, I performed two original theoretical calculations in that chapter. I applied a small noise approximation of channel capacity developed for transcriptional regulation to the different problem of concentration sensing by surface receptors (section 1.1.6), and I analyzed a modified version of an existing model of viral-immune dynamics, which leads to quite different phenomena and bifurcations compared to the original model (section 1.2.2).

In *Antigen encoding*, [1], I worked in close collaboration with the lab of Grégoire Altan-Bonnet at the National Cancer Institute (NCI, USA), in particular co-first author Sooraj Achar. They developed a unique robotic platform to record the dynamics of T cell activation. With Thomas Rademaker, we built a processing pipeline tailored to their datasets, and discovered a low-dimensional description of cytokine dynamics. I formulated an accurate mathematical model of these dynamics, and I developed numerical methods to quantify the channel capacity of cytokines. These methods and findings, relevant for immunologists and biophysicists alike, are presented in chapter 2.

In *CD3 ζ ITAMs*, [2], I collaborated with the labs of Grégoire Altan-Bonnet (NCI) and Paul E. Love (NICHHD) to dissect the ambiguous role of phosphorylation sites called ITAMs on the TCR.

I extended mathematical models of TCR activation to account for T cell receptors with altered ITAMs, and performed mutual information calculations based on their experimental data. These analyses showed that ITAMs improve the ability of T cells to distinguish antigens by mediating an inhibitory coupling between receptors. These findings are presented in chapter 3.

In *TCR/CAR antagonism*, [3], I collaborated with the labs of Grégoire Altan-Bonnet and Naomi Taylor (NCI) to discover and quantitatively explain non-linear inhibitory interactions between T cell receptors and chimeric antigen receptors (CAR) on T cells used for immunotherapy. I revisited and improved TCR activation models, extended them to CARs, and generated predictions quantitatively matching the measurements of co-first authors Sooraj Achar and Taisuke Kondo. We then used the mathematical model's insight to engineer cross-receptor interactions. These findings are presented in chapters 3 and 4, and they have important implications for the model-guided design of precision immunotherapies.

Contribution of Authors

As a general rule, this thesis focuses on my theoretical contributions to collaborative projects. I detail my own theoretical work and quote manuscripts sections that I wrote. Chapter 2 summarizes theoretical contributions from co-first author Thomas Rademaker (PhD, McGill Physics, 2021). All experimental data and methods are the work of my collaborators in the labs of Grégoire Altan-Bonnet, Naomi Taylor, and Paul Love at the National Institutes of Health, USA, in particular Sooraj Achar, Taisuke Kondo, and Guillaume Gaud.

In *Antigen encoding*, [1], which is the subject of chapter 2, as a first author, I developed the data preprocessing methods (sections 2.2.2 and B.1), assessed some aspects of the neural network’s robustness (2.4), developed the cytokine reconstruction method (2.5), the mutual information analyses (2.7). Section 2.3 summarizes co-first author Thomas Rademaker (TJR)’s theoretical contribution using neural networks to initially discover the latent space. I developed the constant velocity model and the fitting procedure jointly with TJR (2.6.1–2.6.2) while the refined model (2.6.4) is my own work. Section 2.2 summarizes experimental methods used by co-first author Sooraj R. Achar (SRA; Altan-Bonnet lab, NCI) and other authors (except theorists – TJR, PF and I) on the paper. The Kendall Tau metric in section 2.6.3 and the Earth Mover’s Distance analysis in 2.8.1 were developed jointly by SRA and I. Figures were made by SRA, TJR and I according to our contributions but edited collectively. Regarding the excerpts of [1] quoted in the chapter, I wrote all quoted excerpts of the supplementary information (SI), since they concern my theoretical contributions. Last authors Paul François (PF) and Grégoire-Altan-Bonnet (GAB) wrote the main text of that paper (a few excerpts are quoted), with edits by SRA, TJR and I.

In *CD3 ζ ITAMs*, [2], which is part of chapter 3, I was co-second author with SRA and John S. Davies (NCI). I was the only theorist on the project with PF for supervision and conceptualization. I developed the mathematical model of TCR antagonism and analyzed its predictions (section 3.3), and I performed the mutual information calculations (3.2.1) for 6Y vs 6F T cells. SRA performed the latent space analysis of 6F T cells (3.2.2) using the pipeline we had developed jointly with TJR in [1]. First author Guillaume Gaud (Love Lab, NICHD) and co-second author SRA performed the experimental work. Figures from that paper shown in the chapter were made by Sooraj Achar and I according to our respective contributions. I wrote modelling parts of the supplementary information (SI), which are quoted in the chapter.

In *TCR/CAR antagonism*, [3], covered in chapters 3 and 4, I was co-first author with Taisuke Kondo (TK; Naomi Taylor's lab, NCI) and SRA. I was also the only theorist with PF for supervision and conceptualization. I developed the revised AKPR model (section 3.5) and its extension to CAR T cells (4.3), I designed and applied the MCMC parameter estimation pipeline, and generated the model predictions for *in vitro* (4.3.4 and 4.4.4) and *in vivo* (4.4) experiments, and for the mutated peptide library (4.4.3). The AEBS CAR T cell concept (4.4.5) was proposed by the NCI first authors and corresponding authors, with inspiration drawn from the mathematical model predictions. TK and SRA performed most of the experiments – with TK focused on the cell line cultures and *in vivo* experiments, and SRA on the *in vitro* robotic assays and some *in vivo* measurements. SRA, TK, and I prepared the figures according to our respective contributions. SRA and I wrote the first draft of the main text together, with subsequent edits from all first and corresponding authors. I wrote the supplemental information (SI) sections which are quoted in this thesis, with edits from PF and GAB.

Introduction

Immunology, physics, and interdisciplinarity

The immune system is sometimes described as a “liquid brain” [5, 6]. What this catchphrase means is that the immune system is a complex, out-of-equilibrium network of cells communicating with each other to respond collectively to a multitude of stimuli. The analogy may therefore be overly suggestive – we do not imply that the immune system gives rise to a form of consciousness – but it has the merit to place the brain and immunity on an equal footing, challenging scientists to understand the latter with the same quantitative level of detail as the former. We are currently short of the mark: while neurobiology and neurophysiology have long since integrated seminal contributions from physics and mathematics, notably the 1952 Hodgkin-Huxley model of electrical action potentials in neurons [7], quantitative immunology faces a dearth of universally recognized mathematical principles. Meanwhile, the practical benefits to be reaped from quantitative immunology are significant, with applications ranging from the treatment and prevention of diseases like COVID-19 [8] to the improvement of crop resilience against pathogens [9]. In particular, the next generation of treatments in cancer immunotherapy [10, 11, 12, 13] will greatly depend on model-guided design to improve precision and specificity [14].

For physicists, the ultimate ambition would be to go one step further in the “liquid brain” analogy: to understand immunity not just as accurately as the brain, but as accurately as liquids. However, we are still far from this level of quantitative knowledge – it may not be attainable at all. Working towards it nonetheless, we should avoid the main pitfall existing in the study of any living system from the physics perspective. The risk is to evacuate what makes biological systems interesting in the first place – the fact that they are living – and treat them as just another state of matter, labelled “living matter”, on which to apply hydrodynamics, thermodynamics, and so on. In Eric Siggia’s words, we should instead “deal with problems first and tools second”, because “[b]iology is not an excuse for doing physics with the names on the variables changed” [15]. Biological physics should ultimately be oriented towards the defining features and functions of living systems. What is life, and how does it evolve? How do organisms self-assemble and develop, move and maintain themselves, make decisions and learn in their fluctuating environments? Satisfactory answers should take the form of emergent physical laws, consistent with but not reducible to fundamental microscopic laws; as noted by Schrödinger, the distinctive properties of biological systems will

require a “new type of physical law” [16, p. 80]. Working towards such laws also requires different mathematical tools and formalisms; information theory, nonlinear dynamics, and statistical learning might be especially appropriate to dissect complex living systems [17, 18, 19].

Is this style of research on immunity still part of physics? We believe so. The object of study may differ compared to traditional fields of physics, but the scientific method is the same: we seek mathematical predictions derived from theoretical principles and agreeing with experimental data. This common quantitative spirit is what constitutes physics. In William Bialek’s words: “Academic disciplines can define themselves either by their objects of study or by their style of inquiry. Physics is firmly in the second camp.” [17]. While the deductive method of physics remains, “[t]he aspects of the world which capture the interest of the physics community can and do change, not least as new phenomena become accessible to the physicists’ style of inquiry” [17]. This is currently the state of affairs in biology, particularly in immunology: there is a deluge of new quantitative data, thanks to rapid technological advances unlocking high-throughput, precise, and multiplexed measurements [20, 1, 21]. These data have no fundamental reason to be ignored by the realm of physics. In fact, biology and physics have spent centuries hand in hand: Aristotle’s *Physics* covered the fundamental causes of movement in inanimate and living entities alike [22] and Lucretius attributed human perception of odors and tastes to “atoms” of different shapes and sizes [23, IV, 615–672]. Modern physics also has a tradition of studying living systems: Helmholtz worked extensively on hearing [24], the “psychophysics” of olfaction and taste are an active research area [25], and physicists have been contributing to theoretical immunology for at least 50 years (Perelson [26]).

Research objectives: immune sensing and signaling

The methods of physics can push quantitative immunology forward in several ways. For instance, the statistical physics of immune repertoire generation [27] and the structural biophysics of antigen-receptor affinity [28] have received significant attention and clarified how adaptive immunity is organized [29]. In this thesis, we focus on another fundamental challenge: the sensing and signaling functions of T cells. These white blood cells, part of the adaptive immune system, must constantly sense their surroundings for signs of infection or damage, and produce appropriate responses through intracellular and extracellular signaling. What parameters determine T cell response dynamics? How do these cells leverage their receptor structures and repertoires to perform their functions? These processes depend on parallel chemical inputs and biochemical pathways, and produce diverse outputs. These high-dimensional attributes complicate the formulation of a concise theory of T cell sensing and signaling.

The physics style of inquiry might alleviate this difficulty. Satisfactory physical theories explain natural phenomena with parsimonious mathematical rules and few parameters. We therefore attempted, in this thesis, to find emergent quantitative descriptions of complex T cell activation processes. Our main objective has been to derive low-dimensional representations of T cell dynamics, and to predict T cell activation by diverse antigenic stimuli. We wanted to compress the high-dimensional chemical spaces of T cell inputs and outputs, to obtain a simplified description of T cell activation and reverse engineer how it depends on the quality and quantity of antigens. We also wanted to connect different time scales of immune activation by building phenomenological models of T cell receptor (TCR) signaling and use them to predict immune responses on slower time scales. Since the distinctive feature of T cell chemical sensing is to disentangle mixtures of similar antigens, we focused in particular on nonlinear effects that arise in such mixtures. We wanted to show how T cell antigen recognition differs from the concentration sensing problem commonly treated in biophysics [30] and that it can nonetheless serve as a model for other combinatorial sensing and signaling problems in biology (*e.g.*, in olfaction and developmental pathways). Then, we finally wanted to show the potential for applications unlocked by quantitative theories of T cell activation. To do so, we investigated cross-receptor interactions in T cell used for cancer immunotherapy. Throughout, we collaborated closely with immunologists (principally the labs of Grégoire Altan-Bonnet, Naomi Taylor, and Paul Love at the National Institutes of Health), and we sought quantitative agreement between model and data.

Thesis outline

We proceeded in several steps to achieve these goals. In chapter 1, we compared and contrasted concentration sensing by bacteria with antigen sensing by T cells, providing background on the immune system and seizing the opportunity to introduce nonlinear dynamics and information theory methods used in subsequent chapters. The next three chapters cover our original research contributions. In chapter 2, we found a low-dimensional compression of T cell secretion of cytokines (extracellular messenger proteins), allowing us to correlate early and late time scales of immune responses with the quality, rather than quantity, of recognized antigens. In chapter 3, we examined responses to antigen mixtures, prompting us to revise previous phenomenological models of TCR signaling. We attempted to explain, in particular, antagonism in mixtures of weak and strong TCR antigens. In chapter 4, we extended our modelling efforts to chimeric antigen receptor (CAR) T cells, to explain and harness cross-receptor antagonism. We covered specific literature review elements at the beginning of these three main chapters.

Chapter 1

Review: chemical sensing and immunity

The immune system must sense early signs of infection and damage amidst an overwhelming majority of healthy, functioning cells. In response, it must produce appropriate intracellular and intercellular signals to restore homeostasis. Immune cells consequently possess complex biochemical machineries to perform their functions. The high-dimensional nature of immune inputs, processes, and outputs poses considerable challenges towards their physics-based understanding, *i.e.*, with simple models based on quantitative principles. Fortunately, there are several examples of biological systems which in fact have underlying low-dimensional descriptions, making simple biophysical models unexpectedly efficient at capturing their properties. This thesis hopes to show that T cell sensing and signaling might be another such example.

In this chapter, we review classical biophysical principles of concentration sensing by cellular receptors (section 1.1), to contrast them with the unique challenges of T cell antigen sensing (section 1.2). We then describe the mechanisms used by T cells to perform their sensing and signaling functions, highlighting how their biochemical complexity might hide simpler principles (section 1.3). Lastly, we summarize various examples of successful low-dimensional models in biophysics (section 1.4), and identify reasons to hope T cell responses also have a low-dimensional structure (section 1.4.3). Some topics also serve as pretexts to introduce notions of information theory and nonlinear dynamics, which form the theoretical foundations of this thesis.

1.1 Biophysics of chemical sensing in living systems

1.1.1 Overview of chemical sensing in biology

Living systems stand to gain selective advantages by acquiring information about their environment and responding accordingly [31]. For this reason, across the tree of life, species have evolved mechanisms to sense chemical signals from their surroundings. These mechanisms are usually surface proteins acting as receptors which can trigger downstream intracellular pathways upon activation [32]. By chemical signals, we mean the identity, concentration, or dynamics of molecules

in the extracellular medium or on other cells (be it small compounds or large proteins). To make these concepts more concrete, we briefly describe a few examples of chemical sensing by biological systems in appendix A: animals smelling odorant molecules via their olfactory receptor neurons [33], bacteria performing chemotaxis in search of nutrient sources [34], or embryonic cells reading morphogen concentrations to differentiate [35]. Despite each having peculiarities, these examples share common features with the main topic of this thesis, sensing and signaling in T cell immunity: antagonism and other nonlinear effects in antigen mixtures, discrimination of relevant and background signals, combinatorial computations, communication between cells. We will highlight these challenges again in our review of antigen and cytokine sensing by T cells, in section 1.2 and the introductions of the following three chapters.

In all examples mentioned above, cells face the same physical limits and share a few fundamental principles. We summarize in the following subsections the canonical biophysics treatment of concentration sensing by cellular receptors, before pointing out other biologically relevant questions (subsection 1.1.8) and transitioning to the additional detection challenges solved by T cell receptors (section 1.2).

1.1.2 Berg-Purcell fundamental physical limit

Consider a single receptor of typical dimension ℓ (typically a few nanometers) faced with the task of estimating its cognate ligand concentration c in the extracellular medium. Even if it could instantaneously count all ligands within a neighborhood of volume ℓ^3 around it, its accuracy would still be limited by the intrinsic variability in the number of ligands within that small volume at a given time, due to the random walk of molecules diffusing in and out of it. This situation is illustrated in figure 1.1.

We can estimate this variability using simple probabilistic arguments. In a stationary, homogeneous state, each of the M molecules in the environment of volume V at concentration $c = M/V$ has the same probability p of being in the small volume ℓ^3 or not, by symmetry. This probability is proportional to the ratio of volumes: $p = \ell^3/V$. Hence, the random number N of ligands within the neighborhood follows a binomial distribution,

$$\mathbb{P}[N = n] = \binom{M}{n} p^n (1 - p)^{M-n} \quad (1.1)$$

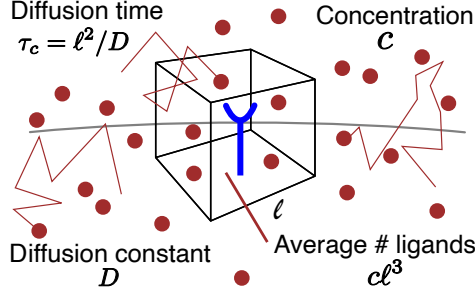


Figure 1.1: Physical limit to molecule counting due to diffusion in the neighborhood of a receptor with linear dimension ℓ . Ligands are at average concentration c , diffusing in and out of the small volume with diffusion constant D and equilibration time $\tau_c = \ell^2/D$, such that $c\ell^3$ ligands are present in the volume on average, with Poisson variance $c\ell^3$. Redrawn after figure 4.16 in [17].

Since $p = \ell^3/V \ll 1$ and $N \ll 1$, this is approximately a Poisson distribution with mean $\bar{N} = Mp$,

$$\mathbb{P}[N = n] = \frac{(Mp)^n e^{-Mp}}{n!} \quad (1.2)$$

which has variance $\sigma_n^2 = Mp$. This intrinsic noise, originating from the discrete nature and small number of ligands in the neighborhood ℓ^3 , is analogous to Poisson shot noise in detector electronics [17]. Due to shot noise, the best possible relative accuracy of this receptor's ideal concentration estimate, $\hat{c} = N/V$, is

$$\frac{\text{Var}[\hat{c}]}{\mathbb{E}[\hat{c}]} = \frac{\sigma_n}{\bar{N}} = \frac{1}{\sqrt{\bar{N}}} = \frac{1}{\sqrt{c\ell^3}} \quad (1.3)$$

where we defined $\sigma_c^2 = \text{Var}[\hat{c}]$ the variance of the concentration estimated, and used the change of variable $\bar{N} = Mp = M\ell^3/V = c\ell^3$ in the last equality. This is the first fundamental physical limit stated by Berg and Purcell [30] in the context of bacterial chemotaxis.

1.1.3 Time and receptor averaging in the Berg-Purcell limit

Given this fundamental limit on a single measurement, there are two obvious ways to improve it, which were also considered by Berg and Purcell: to repeat the measurement over time and to use several receptors. In both cases, averaging k independent measurements $\{N_1, \dots, N_k\}$ should reduce the relative error by a factor $\sqrt{k}$, since $\text{Var}\left[\frac{1}{k} \sum_{j=1}^k N_j\right] = \frac{k}{k^2} \text{Var}[N_1]$.

If the measurement is averaged over a time T , truly independent values are obtained after waiting long enough for molecules to diffuse out of the small volume and be replaced by new ones, which occurs over a characteristic diffusion time $\tau_c = \ell^2/D$, where D is the ligand diffusion constant. Therefore, there are $k = TD/\ell^2$ independent measurements in time T , reducing the relative error

to $\frac{\sigma_{\hat{c}}}{c} = \frac{1}{\sqrt{D\ell cT}}$. Taking into account the autocorrelation time of ligand diffusion introduces extra numerical factors, resulting in a more precise fundamental physical limit¹

$$\frac{\sigma_{\hat{c}}}{c} = \sqrt{\frac{4}{5D\ell cT}} \quad (1.4)$$

If a cell averages the readout of R receptors, each of size ℓ , the relative error is reduced by a factor $\sqrt{R}$, provided that receptors are sufficiently spread out to have uncorrelated neighborhoods. Otherwise, the ligands around nearby receptors becomes correlated, as one receptor can bind a ligand and prevent it from reaching the surrounding receptors [17, 36]. As the number of receptors $R \rightarrow \infty$ on a cell of dimension a , the physical limit approached is that where the sensing volume is the entire cell, $\sigma_{\hat{c}}/c = 1/\sqrt{DacT}$. Putting numbers together for a bacterium with $a = 1 \mu\text{m}$ averaging for $T = 1.5 \text{ s}$ a ligand concentration of 1 nM, we find $\sigma_{\hat{c}}/\mathbb{E}[\hat{c}] \approx 1/30$, which is sufficiently accurate for that bacterium to perform chemotaxis and decide if a concentration gradient is increasing or decreasing.

1.1.4 Berg-Purcell limit for a simple receptor

The calculations above assumed an ideal observer that could perfectly count ligands within a given volume. This served to set an absolute physical limit on concentration sensing which can at best be approached by cells that do not have access to this instantaneous count. Instead, what can be read out by signaling pathways in cells is the sequence of receptor binding and unbinding events. Considering only the available receptor-ligand kinetics, we can set more biologically realistic bounds on cellular sensing accuracy.

Berg and Purcell [30] also treated the case of such a “simple observer” in which concentration would be estimated based on the average receptor occupancy (*i.e.*, fraction of time the receptor is bound). We follow here a slightly different derivation, based on the master equation [37, 38] for the probability $p(t)$ that the receptor is bound at time t , with binding rate k_+c and unbinding rate k_- ,

$$\frac{dp(t)}{dt} = k_+c(1 - p(t)) - k_-p(t) . \quad (1.5)$$

Solving for the stationary distribution ($dp_b/dt = 0$), we find the average receptor occupancy, p_b , is

$$p_b = \frac{k_+c}{k_+c + k_-} = \frac{c}{c + K_D} \quad (1.6)$$

¹The autocorrelation time introduces a factor 4/5; in the original article, a spherical volume was considered instead of a box here, which introduced an additional factor of 3/4 π .

matching the expected result for the macroscopic chemical reaction with equilibrium constant $K_D = k_-/k_+$. Inverting the relationship, the concentration estimate for this cell is

$$\hat{c} = \frac{p_b K_D}{1 - p_b} . \quad (1.7)$$

The occupancy p_b has to be computed in the cell by averaging the receptor state $s \in \{0, 1\}$ over some time T , defining a time-averaged occupancy²

$$\hat{p}_b = \frac{1}{T} \int_0^T dt s(t) \quad (1.8)$$

which has an average equal to the true occupancy, $\mathbb{E} [\hat{p}_b] = p_b$, and variance

$$\sigma_{\hat{p}_b}^2 = \mathbb{E} [\hat{p}_b^2] - \mathbb{E} [\hat{p}_b]^2 = \frac{1}{T^2} \int_0^T \int_0^T dt dt' \mathbb{E} [s(t)s(t')] - p_b^2 \quad (1.9)$$

To find out this variance, we need the autocorrelation function of s , $G(t, t') = \mathbb{E} [s(t)s(t')]$, which we can get by solving the time-dependent master equation (1.5) of $p(t)$ with the receptor bound at the initial condition, $s(0) = 1$. Indeed, at stationarity, $G(t, t') = G(0, |t' - t|) = G(\tau)$ where $\tau = |t - t'|$, and

$$G(\tau) = \mathbb{E} [s(0)s(\tau)] = 0 \times (1 - p(\tau)) + 1 \times p_b p(\tau) = p_b p(\tau)$$

Solving eq. 1.5, which is a simple linear differential equation, with initial condition $p(0) = 1$, we find

$$G(\tau) = p_b p(\tau) = p_b^2 + p_b(1 - p_b)e^{-\kappa|\tau|}$$

where $\kappa = k_+c + k_-$ is the total reaction rate. Inserting in the integral of eq. 1.9, we find the variance on the occupancy time average,

$$\sigma_{\hat{p}_b}^2 = \frac{2p_b(1 - p_b)}{\kappa T} (1 - e^{-\kappa T}) \approx \frac{p_b(1 - p_b)}{\kappa T}$$

if the integration time $T \gg 1/\kappa$. Finally, changing variable with a Taylor approximation $\hat{c} \approx c + \frac{d\hat{c}}{dp_b}(\hat{c} - c)$, this gives us the variance $\sigma_{\hat{c}}^2$, resulting in a relative error on the concentration

²Practically, this can be computed, for instance, by a phosphorylation network with a slow degradation time T ; see *e.g.*, the ‘‘Duty fraction decoder’’ model in [39].

estimate of

$$\frac{\sigma_{\hat{c}}}{c} = \sqrt{\frac{2}{(1 - p_b)k_+cT}}. \quad (1.10)$$

This form becomes more intuitive by realizing that k_+cT is the average number of ligand-receptor encounters in time T . Encounters result in a binding event if the receptor is free, which happens a fraction $(1 - p_b)$ of the time; hence, $(1 - p_b)k_+cT = \nu$, the average number of binding events in time T . Therefore, the simple observer, based on average receptor occupancy, has a concentration sensing error bound of [30]

$$\frac{\sigma_{\hat{c}}}{c} = \sqrt{\frac{2}{\nu}}. \quad (1.11)$$

This result can be contrasted with the ideal observer (shot noise) limit, which had error $2/\sqrt{5D\ell cT}$; considering that $D\ell c$ is the diffusion-limited arrival rate k_+ of ligands on a square patch of side ℓ , the physical limit due to shot noise, eq. (1.4), is

$$\frac{\sigma_{\hat{c}}}{c} = \sqrt{\frac{2(1 - p_b)}{5}} \sqrt{\frac{2}{\nu}} < \sqrt{\frac{2}{\nu}}.$$

Hence, the ideal observer's physical limit is not quite reached by the simple observer receptor, but the $\nu^{-1/2}$ scaling of both limits is the same.

1.1.5 Energy-accuracy tradeoff in concentration sensing

The simple receptor limit above, based on the average time the receptor is bound, can be improved to approach the ideal observer shot noise limit. Using a maximum likelihood approach, Endres and Wingreen [40] have derived the optimal sensing accuracy available in principle from the entire time series of receptor occupancy. They found that signaling networks which would rely on the unbound time intervals only, rather than the bound ones, could improve the accuracy by a factor $\sqrt{2}$. The fundamental limit for simple observers therefore becomes $\frac{\sigma_{\hat{c}}}{c} = \frac{1}{\nu}$ for a system with complete access to receptor binding and unbinding events.

This maximum likelihood calculation did not, however, solve the question of which biochemical signaling scheme in cells can achieve this optimum. Trying to address that question, theoretical studies showed that only out-of-equilibrium receptor signaling schemes (*i.e.*, breaking detailed balance) can approach the theoretical limit of $\sqrt{1/\nu}$, at the expense of entropy production and energy consumption. Such energy-accuracy tradeoffs were exhibited in various models of receptors with non-equilibrium states [41, 42]. Recently, Harvey *et al.* [43] used large deviation theory and stochastic thermodynamics to derive a universal theoretical bound [43, eq. 13] synthesizing

previous results,

$$\left(\frac{\sigma_{\hat{c}}}{c}\right)^2 \geq \frac{2}{T\Sigma^\pi/4 + \nu} \quad (1.12)$$

where Σ^π is the total rate of entropy production in the receptor network [44],

$$\Sigma^\pi = \sum_{i < j} (\phi_{ij} - \phi_{ji}) \log \left(\frac{\phi_{ij}}{\phi_{ji}} \right) \quad (1.13)$$

and where $\phi_{ij} = \pi_i Q_{ij}$ is the probability flux from state i to state j when the system with transition rate matrix Q_{ij} is in its stationary (non-equilibrium) probability distribution π_i . The bound in eq. (1.12) holds until the maximum likelihood limit of $1/\nu$ is reached, beyond which energy consumption cannot improve accuracy further. As we will see, this idea that non-equilibrium reaction schemes and entropy production are needed to improve biological sensing accuracy also underlies models of T cell antigen sensing [45].

1.1.6 Information transmission by simple receptors

We now examine concentration sensing under the angle of information theory, to introduce a few notions which will be useful in chapter 2. We ask how much a cell with R_{tot} simple, independent receptors can learn about the external concentration C of a ligand, based on the number R of its receptors that are bound. This can be quantified by mutual information (MI) [46], that is, how much entropy is dissipated about C once R is known:

$$\text{MI}(C; R) = H(C) - H(C|R) . \quad (1.14)$$

This is the difference between the entropy $H(C)$ of C , and the conditional entropy left when R is known, $H(C|R)$; in the limit where C and R are continuous variables, we have differential entropies, in bits,

$$\begin{aligned} H(C) &= - \int dc f_C(c) \log_2 f_C(c) \\ H(C|R) &= - \int dr f_R(r) \int dc f_{C|R=r}(c) \log_2 f_{C|R=r}(c) , \end{aligned}$$

where $f_X(x)$ denotes the probability density function of a random variable X . Using the definition of conditional probability densities, $f_{C|R=r}(c) = f_{C,R}(c, r)/f_R(r)$, and of marginal probabilities

$f_R(r) = \int dc f_{C,R}(c, r)$, the expression for $\text{MI}(C; R)$ can be rearranged in two useful forms [17],

$$\text{MI}(C; R) = \int dc \int dr f_{C,R}(c, r) \log_2 \left(\frac{f_{C,R}(c, r)}{f_R(r) f_C(c)} \right) \quad (1.15)$$

$$= \int dc f_C(c) \int dr f_{R|C=c}(r) \log_2 \left(\frac{f_{R|C=c}(r)}{f_R(r)} \right) \quad (1.16)$$

The first form shows that MI is symmetric in R and C , and would thus also be equal to the dissipated entropy $H(R) - H(R|C)$; it also shows that $\text{MI}(R; C) = 0$ for independent variables. The second form is useful when treating the external concentration C as an input to the system and the receptor state R as the output; the distribution of inputs $f_C(c)$ is a property of the environment, while the conditional distribution $f_{R|C=c}(r)$ is the input-output mapping of the system, specifying how the output R depends on the input C , with some randomness specified by that density.

To compute $\text{MI}(C; R)$, we consider the concentration C to be a random variable that changes slowly enough to allow R to reach its stationary distribution quasi-statically. Hence, at a given concentration $C = c$, each receptor has a probability p to be bound, given by the stationary solution of the master equation 1.5: $p(c) = \frac{c}{c+K_D}$ where $K_D = k_-/k_+$. The number of bound receptors R , then, is a sum of Bernoulli random variables, which has a binomial distribution (eq. 1.1)³ with mean $R_{\text{tot}}p$ and variance $R_{\text{tot}}p(1-p)$. From the De Moivre-Laplace theorem [48], $X = R/R_{\text{tot}}$ approximately follows a normal distribution $f_{X|C=c}(x)$ with mean $\mu_X(c) = p(c)$ and variance $\sigma_X^2(c) = p(1-p)/R_{\text{tot}}$ (we neglect the continuity correction here).

In this general context, we do not have the concentration distribution f_C . Instead of computing MI for a specific choice of environment, we can compute the channel capacity of the receptors,

$$C(C; X) = \max_{f_C} \text{MI}(C; X) \quad (1.17)$$

that is, the most information R could ever transmit about C , if f_C is well tuned. In general, this maximization over a functional space cannot be performed exactly, in particular because of the term involving $f_X(x) = \int dc f_{X|C=c}(x) f_C(c)$. Here, since the relative error $\sigma_X/\mu_X \sim 1/\sqrt{R_{\text{tot}}} \ll 1$, to gain some intuition, we can make a small noise approximation [49, 50, 51, 52, 53]. We perform a saddle point approximation of the integral $\int dx f_{X|C=c}(x) \log_2(f_X(x))$ with $f_{X|C=c}(x)$ sharply peaked around its mean, $\mu_X(c)$. Then, only $f_X(\mu_X(c))$ remains, and it is related to f_C via the change of variable formula, $f_X(\mu_X(c)) = f_C(c) \left| \frac{dc}{d\mu_X} \right|$. These approximations allow us to solve

³This binomial distribution also be derived by writing a master equation for R directly; see [47, appendix B].

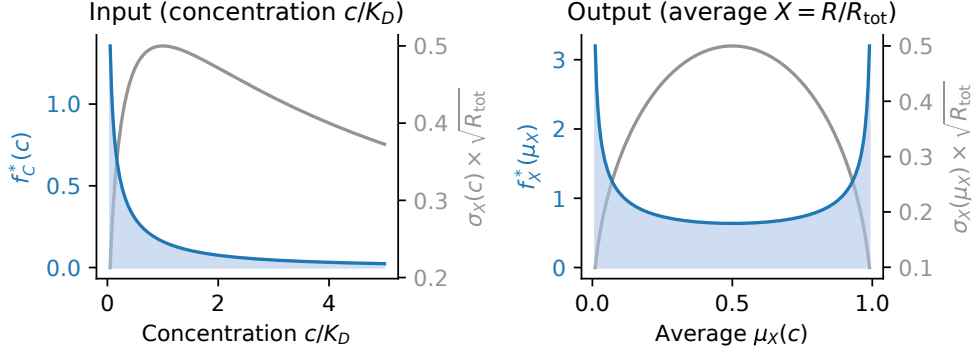


Figure 1.2: Distribution optimizing information on concentration from R_{tot} simple receptors. (Left) Optimal input concentration distribution, $f_C^*(c)$ found by a small-noise approximation in the channel capacity $C(X; R)$. Plotted against the output standard deviation depending on concentration, $\sigma_X(c)$, to illustrate that regions of small noise receive more probability weight: $f_C^* \sim \sigma_X^{-1}$. **(Right)** Optimal distribution of the output's concentration-dependent average, $\mathbb{E}[X] = \mathbb{E}[R/R_{\text{tot}}] = \mu_X(c)$, obtained by changing variable from C to μ_X : $f_X^*(\mu_X) = f_C^*(c) \left| \frac{dc}{d\mu_X(c)} \right|$. Plotted against the output standard deviation as a function of the average μ_X , to illustrate the inverse relation between the two, $f_X(\mu_X) = \frac{1}{Z} \sigma_X^{-1}$.

the variational problem, with a Lagrange multiplier Λ enforcing $f_C(c)$ normalization,

$$\frac{\delta \mathcal{L}}{\delta f_C} = 0 \quad \text{where} \quad \mathcal{L} = \text{MI}(X; C) - \Lambda \int dc f_C(c) .$$

Following steps analogous to those in Tkacik *et al.* [50], we find the optimum

$$f_C^*(c) = \frac{1}{Z} \frac{1}{\sigma_X(c)} \left| \frac{d\mu_X}{dc} \right| \quad \text{and} \quad C(C; X) = \log_2 \left(\frac{Z}{\sqrt{2\pi e}} \right) \quad (1.18)$$

where $Z = \int dc \sigma_X(c)^{-1} \left| \frac{d\mu_X}{dc} \right|$. Replacing $\mu_X(c)$ and $\sigma_X(c)$, we have, for R_{tot} receptors,

$$f_C^*(c) = \frac{1}{\pi} \frac{1}{c + K_D} \sqrt{\frac{K_D}{c}} \\ C(C; X) = \frac{1}{2} \log_2 \left(\frac{\pi R_{\text{tot}}}{2e} \right) .$$

This distribution is illustrated in figure 1.2. The important insight is that $f_C^*(c) \sim \sigma_X(c)^{-1}$: to optimize information transmission, input values which produce smaller noise on the number of bound receptors should occur more often – here, this happens at small concentrations producing very few bound receptors. Consequently, since $f_X(\mu_X) = (Z\sigma_X)^{-1}$ by a change of variable, the average output should have more weight in the region of small noise. Another important

conclusion is that, as we would have expected, the mutual information increases logarithmically with the number of receptors R_{tot} .

1.1.7 Adequacy of well-mixed approximations in chemical sensing

Features of the optimal solution – maximum probability weight in $f_C^*(c)$ at vanishing concentrations and channel capacity increase with R_{tot} – are exaggerated by the small noise approximation. To begin, the relative error σ_X/μ_X scales as $1/\mu_X$, which blows up when $c \rightarrow 0$. Noise is not really small in that region, so the small noise approximation made this regime more accurate than it really is. For more accurate channel capacity calculations, numerical methods are needed, chiefly the Blahut-Arimoto algorithm [54, 55, 56]; we will in fact use it in chapter 2.

Although we neglected shot noise contributions in this information-theoretic analysis of concentration sensing, the ideas behind it are not completely wrong. Using chemical master equations such as eq. 1.5 does assume that molecules are well-mixed and that mass-action kinetics apply. Nonetheless, Van Zon *et al.* [57], studying the example of gene promoter binding with spatially-resolved stochastic simulations, have shown that “the effect of spatial fluctuations can be described by a well-stirred, zero-dimensional, model by renormalizing the reaction rates for repressor-DNA (un)binding” [57]. These results suggest that well-mixed models of biochemical reactions can be useful approximations, especially when parameter values are phenomenologically corrected to account for underlying, microscopic physical processes. When experimental data are available, the “renormalization” of chemical rates can be achieved by fitting model parameters; this is the strategy we will employ in chapters 3 and 4.

Approximating exact physical processes by well-mixed chemical reactions constitutes a useful strategy to establish low-dimensional, interpretable, and tractable models of biological dynamics. Simplifications like these are imperative to discern, from a physics perspective, the emergent laws regulating biological systems, without getting tangled in microscopic details. Neglecting microscopic details by leveraging experimental measurements will allow us to estimate information transmission in T cells responses without having to enumerate all intrinsic physical and biological sources of noise in the system, which would be a hopeless task in our current state of understanding of immunity. These estimations are valuable despite their approximate nature, since tracking information flow is an important method to dissect biological systems from a physics perspective [17, chap. 6].

1.1.8 Beyond concentration sensing by simple receptors

The calculations above provide a biophysical background to the problem of chemical sensing by cells, specifically the determination of ligand concentration. However, it assumed receptors with only one possible cognate ligand, and with a fixed time to average their occupancy state and acquire information. Therefore, it does not address other fundamental chemical sensing challenges faced by living systems.

Decision on the fly

To begin, living organisms probably do not always need very precise concentration estimates, especially when the environmental concentration keeps changing anyways. Instead, they may simply need to distinguish two hypotheses, *e.g.*, the ligand concentration is closer to L_1 or L_2 , and wait until reaching a certain confidence level in either option. Siggia and Vergassola [58] have developed a theory of “decisions on the fly” based on statistical hypothesis testing, showing that organisms can make faster decisions using Wald’s ratio test than with fixed-time averaging. Siggia and Vergassola [58] proposed actual biochemical models of receptor signaling which implement this test. The decision on the fly strategy implies different optimal network architectures compared to the classical Berg-Purcell treatment.

Although we will not directly use it in subsequent chapters, the decision on the fly framework applies to a central problem in the immune system: antigen recognition, that is, as we will discuss in the next section, the detection of rare, strongly binding ligands amidst a sea of slightly weaker ones. Lalanne and François [59] used a similar statistical hypothesis testing framework to investigate this chemodetection problem when, additionally, the concentration of background ligands fluctuates over time. They compared the performance of a few different receptor signaling models, and found that receptor coupling by a global inhibitory feedback helps to buffer environmental fluctuations. This contrasts with Berg-Purcell-style concentration sensing, where independent receptors provide the best improvement, reducing the relative error by $\sqrt{R_{\text{tot}}}$.

Sensing of various other ligand properties

In fact, antigen recognition is one of several other facets of biological sensing beyond the classical Berg-Purcell concentration estimation. Depending on the context, cells instead need to sense ligand combinations [60], identities [61], affinities [62], or dynamical patterns [39]. Real biological

receptors do not simply function independently, in parallel, to each determine the concentration of one ligand.

Regarding combinations, receptors often bind to several similar ligands, for example the receptors for bone morphogenetic proteins (BMP) [63]. Moreover, receptors from the same family often share sub-units (*e.g.*, cytokine receptors in immune cells). This situation leads to nonlinear effects in ligand mixtures. For instance, because of competition for cytokine receptor sub-units, one cytokine can reduce a cell’s sensitivity to another cytokine [64, 65]. Similarly, in olfaction, odorant mixtures often exhibit antagonism: a reduction in the response to a mixture of molecules compared to the individual responses. This effect is due to competition for receptors between the odorants, as explained by a biophysical model of olfactory neuron signaling [66]. Ligand promiscuity (*i.e.*, similarity) and cross-receptor interactions can actually improve concentration sensing of multiple ligands [67], but it also serves other purposes beyond that, like computing ratios or logic functions of ligand combinations and addressing specific cell types based on receptor expression profiles [68]. Mathematically, these effects translate into complicated polynomial equations, requiring computational and algebraic solution techniques such as Gröbner bases [69].

Regarding ligand identities and affinities, the cross-reactivity of receptors for several promiscuous ligands seem to complicate the task of distinguishing ligands, which is important in situations where different ligands mark different biological contexts. However, thinking outside of the concentration sensing paradigm, cells have evolved various mechanisms to distinguish antigens. For example, in the case of Notch receptors and Delta ligands, the manner in which receptors clusters in response to different Delta ligand variants can trigger very distinct temporal signals in cells [61]. In the case of T cell antigen recognition, which we will review in detail below, receptors amplify small affinity differences between antigens to distinguish them [62].

Regarding the recognition of temporal patterns, some ligands have markedly different dynamics dependent on the context. During embryonic development, the Wnt morphogen displays oscillations in propagating waves of gene expression when vertebrae form [70], but static Wnt gradients establish anterior-posterior positional information [71], and Turing reaction-diffusion of Wnt leads to digit patterning [72]. Cells could therefore extract valuable contextual information from these different dynamics. Biochemical networks can in fact compute specific temporal features, such as the period of oscillating signals [39]. However, we still lack a general formalism to formulate how kinetic features of time series, *i.e.*, their “shapes”, encode information that cells can decode. Most formalisms are limited to the stochastic component of time series and require large sample

sizes to be evaluated [73]. We will propose a different formalism, based on a parametrization of low-dimensional dynamics, in section 2.7.

In summary, the Berg-Purcell treatment remains instructive about fundamental physical limits, but we need new theoretical frameworks beyond it to study biologically relevant aspects of cellular sensing. Our main topic, T cell receptor signaling, has some unique features (sensitivity, speed, specificity), as well as important applications, making it interesting in its own right from the biophysics standpoint. It is also a great model system to better understand sensing of ligand identity (according to affinity), cross-receptor interactions, and combinatorial effects in antigen mixtures. After all, contrary to chemotactic bacteria, which have other functions to perform for their own fitness, immune cells have evolved for the greater good of the organism: they are optimized solely for chemical detection, and can therefore solve more complex sensing and signaling tasks.

1.2 Chemical sensing challenges in immunity

Having reviewed the canonical treatment of chemical sensing in biophysics, we can now introduce in this context the specific biological problem examined in this thesis: antigen sensing and signaling in T cell activation. We briefly introduce biological notions about mammalian immunity, emphasizing problems of high dimensionality in cellular sensing and cellular responses, before focusing on T cell receptor (TCR) activation mechanisms.

1.2.1 Brief introduction to the immune system

Roles of the immune system

The most apparent role of immunity is to protect the host against pathogens. In the current canonical view, tracing back at least to Burnet and Fenner [74], the immune system discriminates between self and nonself, or in other words, between host constituents (“friends”) and foreign invaders (“foes”). Obviously, this challenge differs fundamentally from the concentration sensing problem above: a good immune system should sense the *identity* of antigens – by definition, molecules recognized by immune receptors – from self and nonself sources, and respond only to the latter.

To complicate matters, immunologists have unearthed an entire constellation of immune system functions beyond pathogen detection, nuancing the self/nonself paradigm. These functions include clearing and repairing damaged tissues, regulating development and metabolism [75], and maintaining beneficial gut and skin microbiomes [76, 77]. Healthy immune systems tolerate several

“foreign” entities, for example commensal microbiota, grafts, and fetuses. Conversely, autoimmune disorders arise when the immune system attacks antigens which are obviously derived from the self. Instead of self/nonself discrimination, immunity would be better defined as the system responsible for maintaining homeostasis and cohesion within the limits of the organism, which it itself delineates; it does not merely enforce pre-existing, absolute self boundaries [78].

The multifaceted role of immunity has several theoretical implications. First, since the immune system needs to repair tissues and clean up debris, it must be able to sense self antigens as well, and respond to them in these contexts. The influence of self antigens will in fact be central to chapters 3 and 4. In the latter chapter, we will also see that immunity can target mutated self, *i.e.*, cancer. Second, as we have alluded to in section 1.1.8, immune antigens are never presented alone, but rather in highly diverse, fluctuating mixtures; immune cells therefore exploit nonlinear cross-receptor and cross-ligand interactions to perform the appropriate functions in response to these antigen combinations. Third, to respond in a context-dependent manner, immune cells need to sense external chemical cues, such as small proteins like cytokines and chemokines secreted in the extracellular medium (more on them in chapter 2). They also sense costimulatory ligands presented on the surface of other cells [79]. In short, immune chemical sensing is not at all a one-dimensional concentration estimation, nor a binary decision between self and nonself. Rather, it seems that immune cells need to produce complex response within a high-dimensional continuum of possible chemical cues.

Structure of the immune system in Vertebrates

All organisms have some form of immune defense strategies, likely tuned by evolution to match the typical pathogen statistics they encounter – see figure 1.3 for a phase diagram of these possible strategies, derived from a stochastic population dynamics model by Mayer *et al.* [80]. Vertebrates, who typically live and evolve on much slower time scales than their pathogens (blue sector in the phase diagram), have an adaptive immune system, in addition to an innate one.

The innate system detects molecular patterns that are broadly conserved across bacteria or viruses, and mounts generic responses against them. It comprises several lines of defense, from the complement system – proteins combining together on pathogens to pierce holes through their cell walls – to dendritic cells and macrophages – polyvalent sentinel cells that can phagocytose bacteria and dead cells [81]. The adaptive immune system comprises T cells (which mature in the thymus) and B cells (which mature in the bone marrow). These cells have large receptor repertoires to specif-

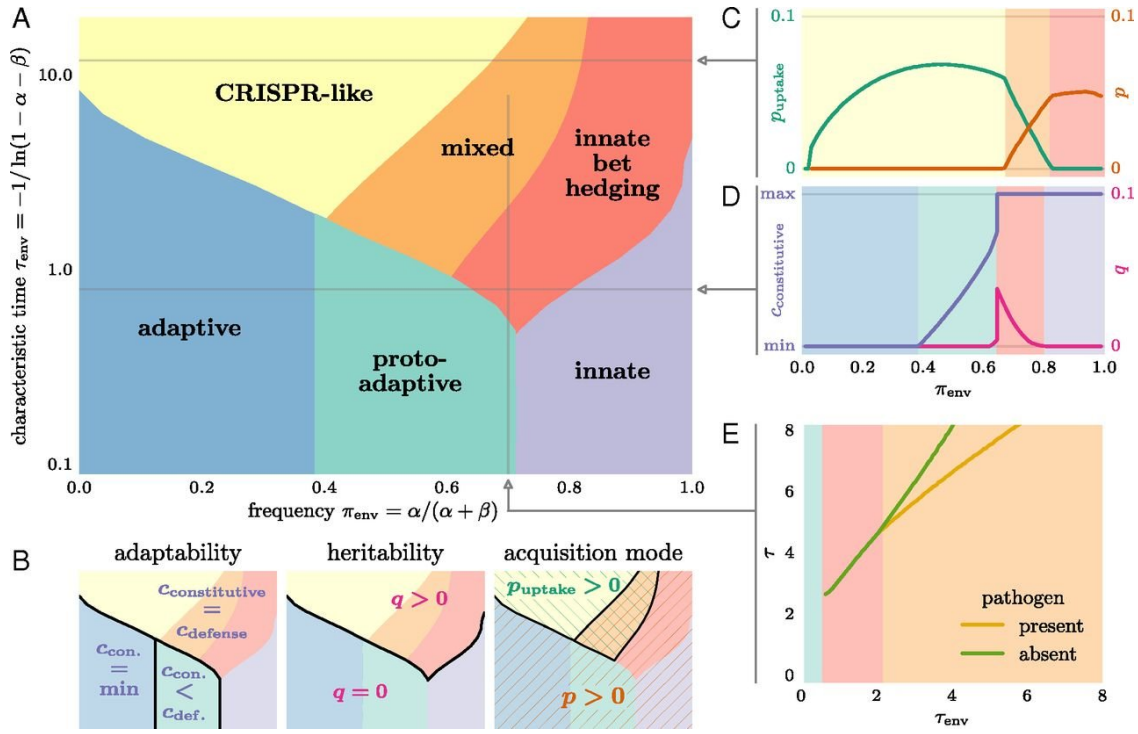


Figure 1.3: “Optimal immune strategies as a function of the frequency and characteristic time of pathogens. (A) Distinct optimal immune strategies emerge for different statistics of appearance of the pathogens. Each phase is characterized by the value of parameters indicated in B and named after a known immune system that has similar characteristics (the term ‘adaptive’ refers to the vertebrate immune system).” [80]. Probabilities α and β are probabilities of the pathogen to appear and disappear in a generation, such that π_{env} and τ_{env} defined on the axes are the frequency of a pathogen appearing, and the characteristic time it stays around (in number of generations). (B) “The different phases of immunity are defined by the values of parameters along three main axes: adaptability (constitutive cost $c_{\text{constitutive}}$), heritability ($1 - q$), and mode of acquisition (p and p_{uptake}).” [80] These three immune parameters represent the cost of the immune protection, the probability to transfer protection against a given pathogen to offsprings, and the rate at which the system mounts a defense (a system could choose to incur the cost of infection instead of the cost of defense, hence a small p could be optimal). To determine the phase of immunity, these three parameters are optimized to maximize population fitness under a given pathogen statistics α, β . (C, D) Optimal parameters depending on π_{env} for two fixed τ_{env} values. “For slowly varying environments (C), rare pathogens are best targeted by CRISPR-like uptake of protection, whereas frequent pathogens are best dealt with by spontaneous acquisition of protection, with a crossover in between where both coexist. For faster varying environments (D), the constitutive cost invested in the protection goes from negligible to maximal as the pathogen frequency increases. When it is maximal, the best strategy transitions from bet hedging ($q > 0$) to a full protection of the population ($q = 0$).” [80] (E) Correlation time (typical duration) of the protection to a pathogen in its presence or absence, for a fixed π_{env} and as a function of τ_{env} . Figure and quoted caption excerpts reproduced from Mayer *et al.* [80] with permission from PNAS.

ically recognize a vast range of possible antigens, with the defining feature that they adapt their response to each encountered pathogen, and form memories of previous encounters [29].

Importantly, the innate immune system initiates and orchestrates adaptive immune responses [82]. Innate immune cells provide much of the chemical cues giving context to T and B cells. They first recognize that damage or infection has occurred, and push adaptive cells to mount a specialized response, targeted at the relevant antigens. For instance, dendritic cells and macrophages bring antigen samples into lymph nodes to activate antigen-specific T cells, and produce inflammatory cytokines to attract and stimulate memory T cells on site [83, 84, 81]. Because of this sequence of events, adaptive immune responses *in vivo* typically span a week or more, whereas innate responses start within the first few days.

1.2.2 Nonlinear dynamics of immune responses

Immune responses exhibit complex, nonlinear dynamics in response to transient pathogenic encounters or chronic auto-immune stimulation. In nonlinear dynamics terms, an immune system is an excitable system [85]: small perturbations away from the stationary state (homeostasis) can cause large transient responses away from it, but ultimately returning to the resting point. We can illustrate this behavior with a simple, two-dimensional system of nonlinear ordinary differential equations (ODEs). We modify the model of Mayer *et al.* [86], which captured the power law scaling of T cell proliferation after immunization with an inactivated virus; here, we add proliferation of the virus (or other pathogens) and killing by immune cells⁴,

$$\begin{aligned}\frac{dV}{dt} &= rV(V - V_0)(1 - V/V_{\max}) - kVC \\ \frac{dC}{dt} &= \frac{aVC}{C_0 + V + C} - bC ,\end{aligned}$$

where V are viruses and C are immune cells. Viruses are killed at rate kVC and immune cells, as a result, proliferate with a saturable rate $\frac{aVC}{C_0 + V + C}$ [86] where C_0 is the saturation threshold. The proliferation rate a also reflects the strength at which virus antigens stimulate the immune cells. Immune cells die at rate bC . Viruses replicate at rate $rV(V - V_0)(1 - V/V_{\max})$, where $V_{\max}$ is the carrying capacity and V_0 represents a minimal inoculum for successful infection, since a single virion is unlikely to suffice [87, 88], and since several virus species (*e.g.*, enteroviruses) travel *en bloc*, hidden inside vesicles [89]. Our toy model (the cubic viral proliferation function, in

⁴Virus killing is not necessarily direct; rather, infected host cells are eliminated by cytotoxic immune cells.

particular) also draws inspiration from the 2D model of excitable auto-immune flare-ups introduced and analyzed by Lebel *et al.* [90]⁵.

If incoming viruses push this system a little bit away from the healthy state⁶ $V = C = 0$, it exhibits dynamics typical of an immune response, shown in figure 1.4: viruses proliferate, but immune cells catch up and eventually clear them before waning back to the resting state. This is what we meant by an excitable system. To understand this behavior more quantitatively, we first change to dimensionless variables and parameters. We eliminate V_0 and δ by using them as scales of cell number and time, to obtain

$$\begin{aligned}\frac{dv}{d\tau} &= \rho v(v-1)(1-v/v_m) - \kappa vc \\ \frac{dc}{d\tau} &= \frac{\alpha vc}{c_0 + v + c} - c.\end{aligned}\tag{1.19}$$

Greek parameters are dimensionless versions of the original ones: $\alpha = aV_0/b$, $\kappa = \kappa v_0/b$, etc. This model has nullclines (lines where either derivative is zero)

$$\begin{aligned}\frac{dv}{d\tau} = 0 &: v = 0 \text{ or } c = \frac{\rho}{\kappa}(v-1)(1-v/v_m) \\ \frac{dc}{d\tau} = 0 &: c = 0 \text{ or } c = (\alpha-1)v - c_0\end{aligned}\tag{1.20}$$

shown as red and blue lines in figure 1.4. The $v = 0$ and $c = 0$ nullclines ensure that cell numbers remain non-negative. This system's excitability comes from the parabolic c nullcline: as soon as a sufficient initial perturbation crosses it, the flow under the parabola has a strong $dv/d\tau > 0$ component; viruses replicate, until hitting the other branch of that nullcline, at which point $dv/d\tau$ is small again and immune cells have time to combat the infection. This parabolic nullcline sets the threshold for long excitations.

The intersections of v and c nullclines define fixed points of the system ($\frac{dv}{dt} = \frac{dc}{dt} = 0$, also called critical points, steady-states, stationary solutions). Chapters 3 and 4 will be especially concerned

⁵While excitable systems in neuroscience are typically relaxation oscillators poised just below a Hopf or SNIC bifurcation of the stable fixed point, here and in [90], we only retain the notion of a large excursion being generated by a small initial perturbation. Our model could be modified to have a bifurcation to oscillations around the healthy state, but the latter would have to move away from $(0, 0)$, for instance by adding constant cell and virus influxes. This would have made the mathematical analysis too complicated for the purpose of this chapter.

⁶A better model would have a small non-zero number of immune cells around at homeostasis, ready to respond; this could be obtained by adding a small constant influx of immune cells a_0 to dC/dt , but we are simplifying here. For the same reason, to facilitate visualization, we chose parameter values ($V_{\max}$) making cell and virus numbers very small. This model could describe viral-immune interactions in a very small region, which is rapidly saturated and requires viruses and immune cells to migrate out of it.

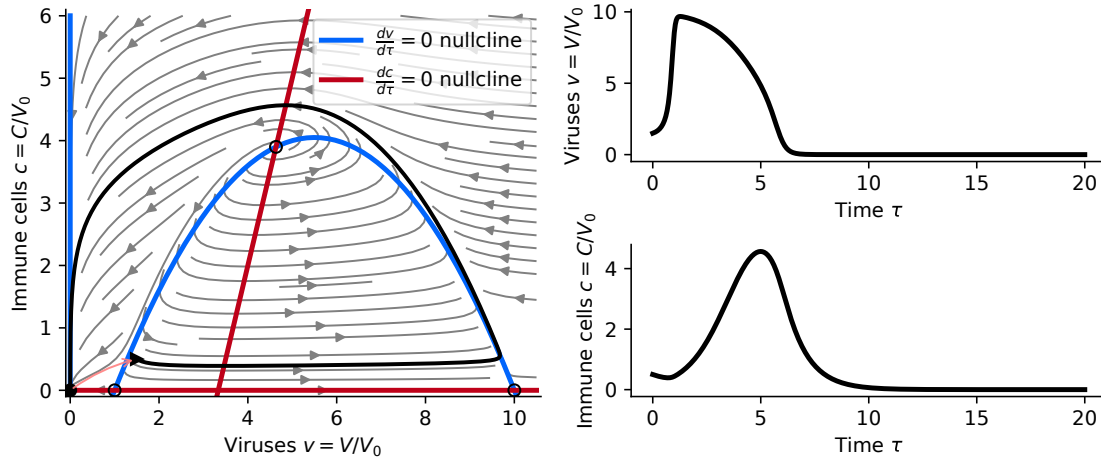


Figure 1.4: Excitable dynamics in a toy model of immune responses. (Left) Phase portrait of the model. Flow lines represent the vector field $\mathbf{f}(v, c) = (\frac{dv}{d\tau}, \frac{dc}{d\tau})$ (eq. 1.19). The black line is a solution trajectory (obtained by integrating the differential equations) after displacing the model away from equilibrium (pale red arrow). Solid colored lines are the model nullclines (eq. 1.20); at their intersections, stable fixed points are indicated by black dots, and unstable ones, by open dots. (Right) Time series of v and c corresponding to the trajectory shown on the phase portrait. Parameter values used: $\alpha = 4$, $\kappa = 1$, $\rho = 2$, $c_0 = v_m = 10$.

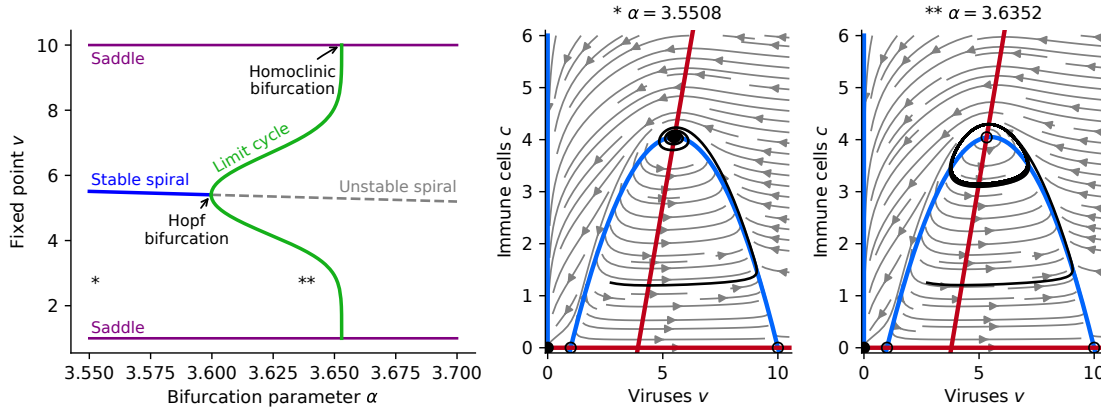


Figure 1.5: Bifurcation diagram and sample trajectories of chronic infections in a toy model of immune responses. (Left) Bifurcation diagram as a function of immune proliferation parameter α . It shows the v coordinate of fixed points as α is varied. The appearance of a stable limit cycle out of a spiral that goes from stable to unstable is a supercritical Hopf bifurcation, and the disappearance of the stable limit cycle as it collides with saddle points is a homoclinic bifurcation. The limit cycle was tracked numerically with the XPPAUT software [91]; other fixed points were evaluated analytically. (Center and right) Sample trajectories in phase space when α is set such that the chronic infection state is a stable spiral (center, point * in the bifurcation diagram) or an unstable spiral with a limit cycle (right, **). Figure 1.4 was showing a case where α is large enough to eliminate the chronic state. Other parameters are fixed as in figure 1.4.

with the fixed points of T cell receptor signaling models. The stability of fixed points – whether nearby trajectories converge towards them or move away – can be assessed by linear stability analysis. In summary, denoting $\mathbf{x} = (v, c)$, we linearize the flow $\frac{d\mathbf{x}}{dt} = \mathbf{f}(v, c)$ around a fixed point of interest. We find a linear equation $\frac{d\mathbf{x}}{dt} = D\mathbf{f} \cdot \mathbf{x}$, where $D\mathbf{f}$ is the jacobian matrix $\frac{\partial(f_v, f_c)}{\partial(v, c)}$. The eigenvalues of this matrix determine the fixed point’s stability (in short, negative real parts imply linear stability).

This method is generally covered in nonlinear dynamics textbooks [85, 92, 93, 94], so we only summarize the results for this toy model of immune activation, which are contained in figure 1.4. The fixed point $(v, c) = (0, 0)$ is always a stable node, meaning that nearby points converge to it. For biologically meaningful parameter values (*i.e.*, positive and carrying capacity $v_m > v_0 = 1$), fixed points at $(1, 0)$ and $(v_m, 0)$ are typically saddles, that is, an asymptotically unstable point but with one attracting line (stable manifold of codimension 1); however, for large α , $(1, 0)$ can be an unstable node (no attracting manifold) while for small α , $(v_m, 0)$ can become stable – a successful infection because immune cells are too slow to respond.

Such changes in existence or stability of fixed points are called bifurcations; they modify qualitatively the system’s global dynamics. The model has another fixed point with interesting bifurcations: a chronic infection state $(v^*_{\pm}, c^*_{\pm})$, solution of the quadratic equation for the intersection between the parabolic v and linear c nullclines. For small c_0 , when α is small, both solutions can be in the positive quadrant with one being stable: immune proliferation saturates too much to clear the infection. As α increases, however, the chronic state disappears (through a saddle-node bifurcation). For larger c_0 , (as in figure 1.4), the $+$ root only is in the upper quadrant. For small α , it is stable (either a node or a spiral): the immune system tolerates the weakly stimulatory pathogen. As α increases, this chronic state becomes oscillatory through a subcritical Hopf bifurcation, and eventually disappears through a homocline bifurcation, as shown in figure 1.5. Hence, even with just a simple toy model, we found a rich repertoire of possible nonlinear dynamics in the immune system.

This toy model is a first illustration of how low-dimensional models can provide insight into immune dynamics. However, it reveals that the overall immune response depends finely on initial stimuli to trigger excitable dynamics, with outcomes changing drastically for even moderate differences in the activation strength of immune cells by pathogens – encoded by α in the toy model. This suggests that we must understand particularly well the early events of immune responses. In other words, we need good models of immune sensing and receptor signaling to determine which antigenic stimuli will cross the threshold for response, and how large that response will be. How-

ever, it is not immediately clear how low-dimensional models could also be useful in this next task, since the chemical spaces in which immune cells evolve are, in principle, high-dimensional. This is particularly the case for T cells, on which we now focus.

1.2.3 High-dimensional nature of T cell activation and functions

TCR antigens for different T cell types

As previously mentioned, T cells are white blood cells which mount specific responses to pathogens. To do so, they express T cell receptors (TCRs) on their surface, with which they scan the surface of other cells to detect signs of infection or other abnormalities. TCR antigens are short peptides loaded on larger proteins called major histocompatibility complexes (MHC) on the surface of the antigen presenting cells (APCs). These peptides act as ligands for TCRs, or in other words, bind to them. This binding is specific: each T cell has a unique TCR sequence which is cognate to a small fraction of all possible antigens, while different T cells generally have different TCRs such that they collectively cover the antigenic space [95]. To recognize antigens, a TCR is made of two extracellular chains, generally $\text{TCR}\alpha$ and $\text{TCR}\beta$, and six chains with an intracellular domain: two central $\text{CD3}\zeta$, two lateral $\text{CD3}\epsilon$, one $\text{CD3}\gamma$ and one $\text{CD3}\delta$ (figure 1.6). The $\text{TCR}\alpha\beta$ chains interact with antigens and is the part that varies from cell to cell to confer its specific binding properties to a TCR. The intracellular chains transduce signals and trigger signaling pathways, in particular through phosphorylation sites called immunoreceptor tyrosine-based activation motifs (ITAMs).

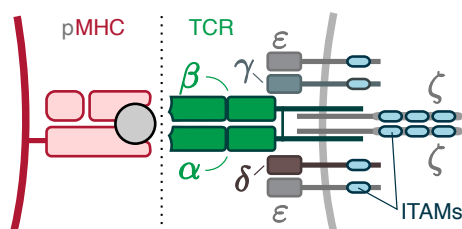


Figure 1.6: Schematic of the T cell receptor structure with its α and β chains and its six CD3 chains, shown bound to a peptide-loaded MHC. The CD4 or CD8 co-receptor is not shown.

There are two main types of T cells, depending on the co-receptor that they express: CD4^+ and CD8^+ T cells. CD4^+ T cells recognize class II peptide-MHC (pMHC) complexes, which load peptides of typically 13–17 amino acids, while CD8^+ T cells recognize class I pMHCs, with peptides of typically 8–10 amino acids [84, chap. 4]. Essentially all mammalian cells have dedicated pathways to chop up and load samples of their proteome on class I MHCs to be inspected by CD8^+

T cells. Infected or defective cells thus present altered pMHCs that can trigger $CD8^+$ T cells. Additionally, some innate immune cells like dendritic cells and macrophages are professional APCs expressing class II MHCs, on which they present peptides coming from the pathogens or dead cells they have phagocytosed, to activate $CD4^+$ T cells. Hence, the two T cell types generally have different functions: $CD8^+$ T cells become cytotoxic cells which can directly kill their targets (since their pMHCs show signs of infection or damage), while $CD4^+$ T cells generally become helper T cells which, for instance, assist B cell response [84, chap. 6]. In this thesis, we focused on $CD8^+$ T cells, except where mentioned otherwise.

The space of TCR antigens and the TCR repertoire

The number of possible peptides with length 8–10 and made out of the 20 canonical amino acids is $\sum_{n=8}^{10} 20^n \sim 10^{13}$. Even if only $\sim 3\%$ of these sequences have the optimal hydrophobicity in anchor positions to actually be presented on MHC molecules [96], the space of possible TCR antigens is still very large. In principle, a naive description of this antigenic space would require a high-dimensional vector $\mathbf{L} = (L_1, L_2, \dots, L_A)$ specifying the concentration of all $A \sim 10^{11}$ antigens possibly presented. These possibilities are compounded with the diversity of MHC alleles in humans – there are enough variants that each individual can have a unique combination of 5 or 6 alleles. These variants influence the loading of peptides and TCR-pMHC interactions in principle.

To cover this huge antigenic space, mammals evolved to have a large TCR repertoire. As mentioned above, each T cell has a single TCR sequence, but different cells have different receptors, with a wide range in how many T cells are part of the same clone, *i.e.*, share the same TCR [97]. The diversity of this repertoire is generated through a process called VDJ recombination, whereby gene segments are mixed and matched, then further modified at their junctions, to form unique TCR α and β chains in each T cell [27]. At least $\sim 10^{39}$ TCR $\alpha\beta$ sequences are theoretically possible, although not all with the same probability of being generated [98]. A human has 10^7 – 10^8 of these sequences among its 10^{11} T cells [99].

Importantly, the stochastically generated TCR $\alpha\beta$ diversity is filtered through thymic selection: T cells are negatively selected to eliminate those responding too strongly to self peptides, and positively selected to eliminate completely unresponsive TCRs [100]. Thus, the TCR repertoire of an individual is shaped to cover “gaps” between self peptides in antigenic space – in fact, targeting sequences one or few amino acid substitutions away from self peptides [95]. TCRs are cross-reactive:

each TCR can in principle respond, with varying intensity, to over 10^5 different peptides⁷ [96], resulting in adequate coverage of the antigenic space. Mixtures of antigens can therefore produce nonlinear combinatorial effects as they stimulate the same given TCR with different strengths.

Other chemical cues sensed by T cells

In addition to antigens, T cells also sense other kinds of proteins. T cells have receptors besides the TCR to detect co-stimulatory or inhibitory ligands from APCs. The CD4/CD8 co-receptor is an example, sensing the presence of MHCs of the right class; another is CD28, binding to the B7 co-stimulatory ligand. Some co-receptors, for example CTLA-4 and PD-1, are inhibitory when they bind to their surface ligands, (B7 and PD-L1, respectively). They are leveraged in cancer immunotherapy (chapter 4) [84, 101]. Together, co-stimulatory ligands provide a signature of APC type and immune context which does modulate T cell responses [79].

Additionally, T cells respond to diffusible extracellular proteins, chiefly cytokines and chemokines. We will review cytokines in more detail in chapter 2; we mention here that these are extracellular messenger proteins that can stimulate, sustain proliferation, and guide migration of T cells. There are over 30 different cytokines in humans [102] and a similar number of chemokines [103]; this multiplicity perhaps serves to multiplex information [104] or to improve robustness through redundancy [102]. Moreover, receptors for cytokines, in particular, share sub-units and therefore can sense combinatorial effects of these messenger proteins [64, 65].

High-dimensional outputs of T cells

Cytokines are produced by innate immune cells, but also by T cells themselves as part of their outputs after activation. They have the potential to create important feedback effects within T cell populations, where different clones can co-opt each other into activation [105]; on the contrary, cytokines like IL-2 activate regulatory T cells, which suppress self-reactive T cells [106]. In addition to cytokines, T cells secrete other effector molecules; in particular, cytotoxic T cells release chemicals like granzymes and perforin to lyse their target cells [107]. CD4⁺ helper T cells upregulate surface antigens such as the CD40 ligand that help activate B cells [108]. Besides secretion of proteins, activated T cells also start proliferating, growing in size, and leaving lymph nodes, stopping their migration when encountering stimulatory antigens at the site of infection [84, chap. 1]. We

⁷The converse is not true: one antigen typically does not trigger more than ~ 100 possible TCRs in an individual.

lack quantitative principles to understand how the timing and amplitude of these multiple T cell outputs are determined by the various inputs they integrate.

1.3 Antigen recognition mechanisms in T cells

We will focus our efforts on TCR antigen recognition, since pMHCs are still the most decisive input among all biochemical cues mentioned above. While the detection of co-stimulatory molecules and cytokines rely on simpler receptor binding-unbinding mechanisms like those reviewed in section 1.1.4, antigen sensing by TCRs is representative of a different chemical sensing paradigm.

1.3.1 The TCR antigen recognition problem

To state it again clearly, the function of a T cell is to recognize and respond to its cognate antigens, which constitute a small fraction of all possible peptide sequences, are typically present on APCs in far smaller numbers than self antigens, and have only small affinity differences with the latter. As we will see in chapter 3, there could be up to 10^6 self antigen copies on an APC, and only ~ 10 cognate antigens, which T cells nonetheless detect. To make the challenge seem even more difficult, T cells having undergone thymic selection do not respond strongly to self antigens (negative selection), yet they still interact weakly with them – because they need to make it through positive selection, and because these self antigens tend to be one amino acid substitution away from cognate TCR antigens [95]. Despite that, T cells are not activated by the large numbers of self antigens always present in the background. Moreover, TCRs are cross-reactive enough to recognize, in principle, several different peptides, meaning they must also be able to distinguish small differences between these ligands. Overall, this suggests that TCRs have an exquisite sensitivity to small differences between their cognate antigens and other peptides mixed with them on APCs [62], despite fluctuations of these mixtures.

1.3.2 Complexity of the TCR signalosome

Consequently, T cells have evolved an incredibly complex biochemical machinery to transduce signals from complex antigen mixtures and turn them into appropriate high-dimensional outputs. When a T cell encounters an APC, it sticks to it for a duration ranging from seconds to minutes, forming a zone of close contact between cell membranes called an immunological synapse [109, 110, 111]. Within this synapse, receptors on the T cell side and pMHCs on the APC side diffuse on their respective membranes and interact with each other [112, 113].

Upon binding of a pMHC to a TCR, several biochemical events are set in motion. The kinase Lck is brought closer to the CD3 chains to phosphorylate their ITAMs. This allows enzymes and signaling proteins to be sequentially recruited and activated (generally, by further phosphorylation) to the TCR, starting with proteins called ZAP-70, then LAT and PLC- γ , and so on [114, 84, 115]. Some smaller enzymes, like the phosphatase SHP-1, can also diffuse around and may couple individual receptors. The resulting cluster of signaling proteins is referred to as the TCR signalosome and is shown in figure 1.7. When it assembles, it triggers amplifying signals like ERK phosphorylation [116], leading to activation of key biochemical pathways like the phosphoinositide 3-kinase (PI3K) pathway [105] and thus starting the production of T cell chemical outputs described above.

The exact biophysical nature of this large aggregate of signaling proteins is unclear; it probably involves a combination of direct binding (*e.g.*, ZAP-70 [117]), clustering, and phase separation (*e.g.*, LAT [118] and Lck-CD3 ϵ [119] condensates). Antigen recognition is the problem for which T cells are made by an organism, hence they can spend the resources and energy (stored in adenosine triphosphate, ATP) needed to maintain this machinery out of equilibrium and to perform above the limits imposed by chemical detailed balance [45]. The signalosome's complexity is intimidating, making exhaustive physical models hard to interpret and analyze. In contrast, our understanding of T cell activation would be facilitated if we could find low-dimensional quantitative models of the effective processes implemented by the TCR signalosome.

1.3.3 TCR-antigen binding: the lifetime hypothesis

As a starting point for a simplified description of TCR signaling, we first examine the physical basis on which TCRs distinguish different antigens. What antigen properties are sensed by the TCR signalosome? While B cell antibodies have very strong binding affinity for their target epitopes (dissociation constant $K_D = k_{\text{off}}/k_{\text{on}} \sim 1 \text{ nM}$), TCR-pMHCs affinities are much weaker ($K_D > 1 \mu\text{M}$), with K_D s in the same range for strong and weak antigens [120]: hence, TCRs are not recognizing ligand identities based on very specific, antibody-like binding. It has been proposed that TCR-pMHC binding induces conformational changes in the TCR or forms catch bonds (where mechanical pulling on the receptors increases the lifetime) [121]. However, recent cryo-EM studies have not found evidence of such molecular structure changes in the binding region of TCRs [122]. Conformational changes in the intracellular domain of the TCR, such as unfolding of the CD3 ζ chains out of the plasma membrane [123], could explain how pMHC binding triggers downstream signals, but not how different ligands are distinguished in the first place.

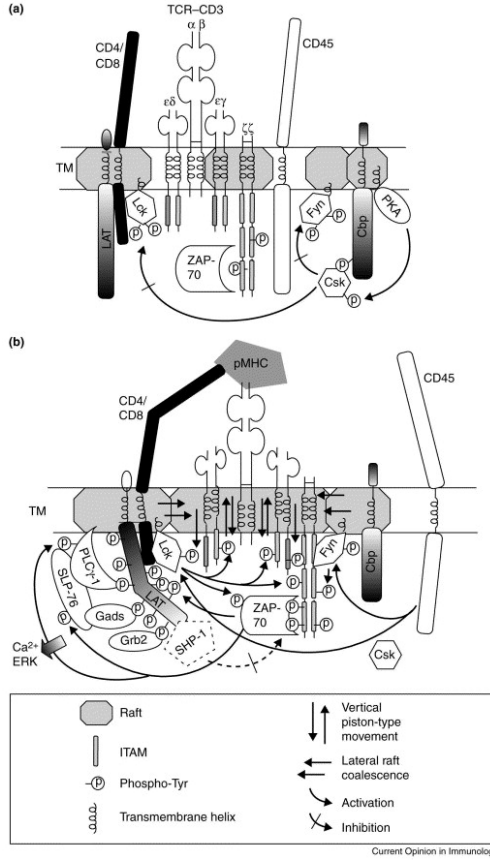


Figure 1.7: Structure of the TCR signalosome (a) Steady-state in absence of a bound pMHC. ITAMs are mostly non-phosphorylated, since Lck is inhibited by Csk and CD45. (b) Recruitment of several signaling molecules upon pMHC binding. Lck is active, CD45 is segregated away from the receptor, the receptor ITAMs are fully phosphorylated, ZAP-70 is bound to CD3 chains, leading to phosphorylation of LAT, activation of PLC-γ, and so on, culminating with Ca²⁺ and ERK signaling to activate cell effector functions. Phosphatase SHP-1 is also recruited to keep this signaling complex in balance and prevent over-activation. Figure from Werlen and Palmer [114, figure 1], reproduced with permission from Elsevier.

Instead, it seems that the dissociation rate, typically reported as a binding lifetime $\tau = 1/k_{\text{off}}$, is the property recognized by TCRs to establish differences between ligands, while the binding rates k_{on} are diffusion-limited and quite similar between antigens. TCR-pMHC binding times typically range between roughly 0.5 s (for non or weakly stimulating antigens) to 15 s (for strong antigens), with excellent correlation between the binding time and antigenicity of a peptide [120]. We will thus often use the term “antigen quality” to designate TCR binding times τ . Although recent measurements using surface plasmon resonance suggest a wider range of binding times [124], the fact remains that the TCR can sense, amplify, and respond specifically to small differences in binding times between different antigens. This idea is sometimes referred to as the “lifetime dogma”. As with any dogma in biology, other factors are likely at play [62], but pMHC binding times are central in most recent TCR modelling efforts [125, 124, 115, 113].

1.3.4 Kinetic proofreading

Thus, the TCR signalosome performs chemical sensing of antigen binding times, but the essence of the mechanisms carrying out that function remains obscured under complicated biochemistry (figure 1.7). Does it in fact boil down to simple principles? Simple bound receptors like those of

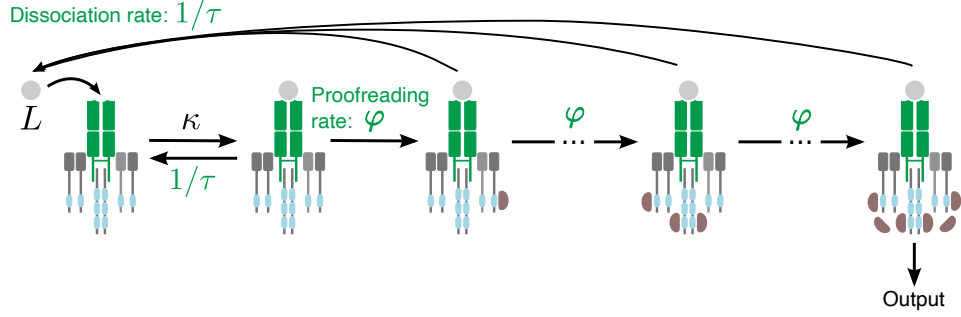


Figure 1.8: Cartoon representing the biochemical reaction network of the classical KPR model of TCR antigen recognition.

section 1.1.4 are poor sensors of antigen quality (binding times), since increasing the concentration c of a ligand can compensate its smaller τ to obtain the same receptor binding state. Indeed, the average number of bound receptors is

$$\mathbb{E}[R] = R_{\text{tot}} \frac{c}{c + K_D} = \frac{ck_{\text{on}}\tau}{ck_{\text{on}}\tau + 1}$$

such that only the product $c\tau$ appears in the equation, showing that c and τ cannot be sensed independently by this scheme.

Key inspiration came from the work of Hopfield [126] and Ninio [127] on the surprising accuracy of DNA replication. They noticed that error rates $\sim 10^{-8}$ in DNA replication are much smaller than the equilibrium prediction $\sim 10^{-4}$ based on the small free energy differences $\Delta G \approx 10k_B T$ between correct (A-T, C-G) and incorrect (A-G, C-T, etc.) base pairings. They proposed a mechanism, termed kinetic proofreading (KPR), in which an extra reaction step – the proofreading step – amplifies these small differences in dissociation rates between correct and incorrect pairs, approximately squaring the error rate down to $(e^{-\Delta G/k_B T})^2 = \left(\frac{k_{\text{off,correct}}}{k_{\text{off,incorrect}}}\right)^2 \approx 10^{-8}$.

This amplification of small differences in unbinding rates is exactly what T cells need to sense small differences in antigen quality without letting differences in antigen abundances interfere⁸. McKeithan [128] turned this idea into a low-dimensional model of TCR binding and signaling. Their key idea is that upon binding a pMHC, the TCR undergoes a sequence of N biochemical modification steps leading to an output-producing activated state, with a chance to unbind with

⁸CD4⁺ T cells may be more sensitive than CD8⁺ T cells to antigen quantity: they select B cells in germinal centers based on how many antigens they present bound on their BCRs [108], because that number depends on the affinity of their antibodies. So CD4⁺ cells, to select B cells with the best antibodies, must read antigen quantity very accurately, once antigen quality is fixed (by the high specificity of the B cells' antibody, ensuring they bind the relevant antigen only).

rate $1/\tau$ at each step. We assume that these steps are irreversible when the ligand is bound, but rapidly undone after it unbinds. This network of reactions is illustrated in figure 1.8. Consider now R_{tot} receptors interacting in an immunological synapse with L ligands presented by the APC. We treat the KPR network as a well-mixed, deterministic biochemical network, where we model C_n , the average number of TCRs in state n , for each $n \in \{0, 1, \dots, N\}$, as a function of time. Applying mass-action kinetics, and considering the fact that the number of free ligands at a given time is L minus the number of bound receptors $R_b = \sum_0^N$, the corresponding system of ODEs is

$$\begin{aligned}\frac{dC_0}{dt} &= \kappa(R_{\text{tot}} - R_b)(L - R_b) - (\varphi + \tau^{-1})C_0 \\ \frac{dC_n}{dt} &= \varphi C_{n-1} - (\varphi + \tau^{-1})C_n \quad (1 \leq n < N) \\ \frac{dC_N}{dt} &= \varphi C_{N-1} - \tau^{-1}C_N.\end{aligned}\tag{1.21}$$

We take the space here to detail the steps to solve for the fixed point solution, since similar calculations will be often used in this thesis. The steady-state occurs within minutes after the immune synapse formation [129]. Summing all equations, we have a constraint for R_b :

$$0 = \kappa(R_{\text{tot}} - R_b)(L - R_b) - \tau^{-1} \sum_{n=0}^N C_n = \kappa(R_{\text{tot}} - R_b)(L - R_b) - \tau^{-1} R_b$$

We make the approximation that L is small compared to the total number of receptors – we will instead use the exact quadratic solution in chapter 3 – so $R_b \approx \frac{\kappa R_{\text{tot}} \tau}{\kappa R_{\text{tot}} \tau + 1} L$. Then, we solve the first-order recurrence relation (coming from $\frac{dC_n}{dt} = 0$),

$$C_n = \frac{\varphi \tau}{\varphi \tau + 1} C_{n-1} = \Phi C_{n-1} \quad (1 \leq n < N)$$

where we defined $\Phi = \frac{\varphi \tau}{\varphi \tau + 1}$. Iterating, we find that $C_n = \Phi^n C_0$. We also use the last equation, $\frac{dC_N}{dt} = 0$ to find $C_N = \varphi \tau C_{N-1} = \varphi \tau \Phi^{N-1} C_0$. Inserting these expressions into $R_b = \sum_{n=0}^N C_n$, and using the geometric series for the resulting sum of Φ^n factors, we find $C_0 = (1 - \Phi) R_b = \frac{R_b}{\varphi \tau + 1}$, and thus the stationary solution for all C_n in terms of the model parameters (R_{tot} , κ , φ) and antigen properties (L , τ),

$$\begin{aligned}C_n &= R_b (1 - \Phi) \Phi^n = \frac{R_b}{\varphi \tau + 1} \left(\frac{\varphi \tau}{\varphi \tau + 1} \right)^n \quad \text{for } 0 \leq n \leq N \\ \text{where } R_b &\approx \frac{\kappa R_{\text{tot}} \tau}{\kappa R_{\text{tot}} \tau + 1} L\end{aligned}\tag{1.22}$$

In the limit where $\varphi \ll \tau^{-1}$ (slow proofreading for better accuracy) and $\kappa R_{\text{tot}} \gg \tau^{-1}$ (fast initial binding), then we have the approximate scaling of the solution with antigen quality and quantity,

$$C_N \sim L\tau^N \quad (1.23)$$

showing that the ratio of responses to two different antigens would be amplified exponentially as $C_{N,1}/C_{N,2} \sim L_1/L_2 (\tau_1/\tau_2)^N$, improving with the number of proofreading steps N .

Early experimental evidence in favor of this mechanism suggested that sequential phosphorylation of CD3 ζ ITAMs could play the role of kinetic proofreading steps [130]. More recently, events further down in the TCR signalosome assembly, such as the binding of ZAP-70, LAT, PLC γ , have also been shown to act as putative kinetic proofreading steps [131, 115]. As we will see in chapters 3 and 4, we do not need to model the biochemical particularities of each step to get predictive power. Hence, a generic KPR scheme constitutes a reasonable low-dimensional backbone for phenomenological models aiming to capture high-dimensional TCR signaling events.

In this basic model, reverse reactions have been neglected under the assumption that the forward KPR rate dominates when a pMHC is bound. An exactly zero reverse reaction flux is not possible, as it would imply infinite entropy production terms in equation 1.13, so this is only an approximation for a non-zero but negligible reverse rate. Including these nonzero reverse rates decrease the KPR performance, but it remains sufficiently high [132], as long as a (finite) breaking of detailed balance between forward and reverse rates is maintained. This non-equilibrium regime is necessary to improve the antigen quality discrimination of simple equilibrium receptors, $R \sim L\tau$. Thus, KPR costs energy to T cells, consuming a power per receptor on the order of $10^4 k_B T/s \approx 10^3 \text{ ATP/s}$, which is still only a tiny fraction (10^{-6}) of the total energetic budget of a cell [133]. This enhancement of the forward rate, as well the rapid undoing of the reaction steps following pMHC unbinding, could be explained by kinetic segregation of the CD45 phosphatase away from the bound TCR-pMHC complex, due to its large extracellular domain [134, 135].

1.3.5 Adaptive kinetic proofreading

The reason why TCR signaling amplifies small binding time differences is, really, to ensure these differences dominate changes in antigen abundance. However, the output in KPR still scales proportionally to antigen quantity L , since $C_N \sim L\tau^N$. It is possible to improve this scaling further to eliminate most of the L dependence and explain how TCRs can be sensing almost exclusively τ . The key solution is to add an inhibitory coupling between TCRs, whether it is a negative feedback

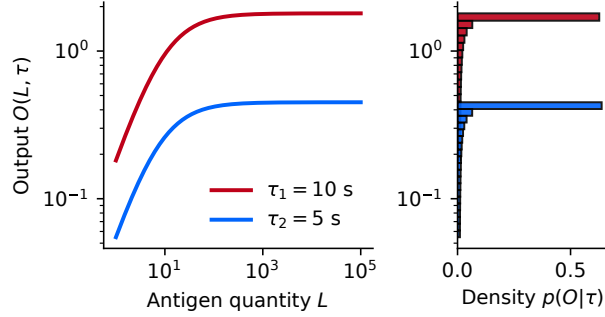


Figure 1.9: Illustration of the antigen quality discrimination problem. (Left) TCR model output curves, in the deterministic limit, depending only on antigen quality (red vs blue curves) over a wide range of antigen quantities L . (Right) Marginalization of the TCR output distribution over the range of L , to obtain $p(O|\tau)$. With a perfect model, these distributions would not overlap for different antigens. Redrawn after Lalanne and François [132, fig. 2].

or an incoherent feedforward loop. This was first shown using detailed computational models of TCR activation, where the phosphatase SHP-1 played this inhibitory role [116, 136]. The mechanism became especially clear when a much simpler, low-dimensional phenotypic model of this SHP-1 feedback was proposed [129].

In parallel, the same class of models was found *de novo* using an *in silico* evolution algorithm [132], φ -evo [137], which simulates rounds of mutation and selection on a population of biochemical networks optimized for a fitness function. For TCR activation models, this fitness function should measure how well the output reflects antigen quality τ without depending on antigen quantity L . Information theory provides a quantitative formulation: the input-output function of the model, $p(O|\tau, L)$, should maximize the mutual information $MI(O; \tau)$ between its output O and τ while the inputs are marginalized over all possible antigen quantities L :

$$\max_{p(O|\tau, L)} \int d\tau p(\tau) \int dO p(O|\tau) \log_2 \left(\frac{p(O|\tau)}{p(O)} \right)$$

where $p(O|\tau) = \int dL p(O|\tau, L)p(L|\tau) = \int dL p(O|\tau, L)p(L)$

and $p(O) = \int d\tau p(\tau)p(O|\tau)$.

This objective is illustrated in figure 1.9. The distribution $p(L)$ over which we marginalize could be, for instance, uniform over a wide range of possible quantities, $L \in [10^0, 10^5]$ approximately.

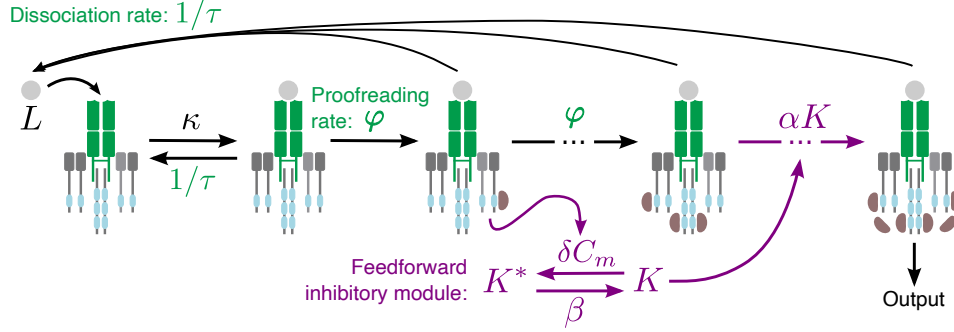


Figure 1.10: Cartoon representation of the biochemical network of the AKPR model introduced in [132]. The global inhibitory coupling between receptors is mediated by the feedforward module, where the kinase K is deactivated by intermediate complexes, thus reducing the final proofreading step rate.

KPR schemes with inhibitory coupling are called adaptive kinetic proofreading (AKPR) models, and will be at the core of chapters 3 and 4, where we will analyze them in more depth. Here, we introduce the simplest version of AKPR, found by *in silico* evolution [132], depicted in figure 1.10. The basic KPR model is modified to have a shared pool of a kinase, K , mediating the last proofreading step with rate αK . This kinase is deactivated by receptor complexes in an intermediate state C_m , with $m < N$, thus implementing an incoherent feedforward loop. Intuitively, the point of this feedback is to decrease the response to weak TCR antigens present in large quantities: these antigens produce a lot of C_m to activate the feedback, but – because of the $L\tau^m$ versus $L\tau^N$ scaling of C_m versus C_N – not as much C_N to overcome it.

Compared to the basic KPR model, only the ODEs for C_{N-1} and C_N are affected, changing term φC_{N-1} for $\alpha K C_{N-1}$, and one new equation is needed for K ,

$$\frac{dK}{dt} = \beta(K_{\text{tot}} - K) - \delta C_m K .$$

Solving for C_N at the fixed point, one factor Φ is replaced by a factor containing K ,

$$C_N = R_b \left(\frac{\alpha K_{\text{tot}} \tau}{\alpha K_{\text{tot}} \tau + 1} \right) \left(\frac{\varphi \tau}{\varphi \tau + 1} \right)^{N-1} \quad (1.24)$$

where the steady-state value of K is

$$K = K_{\text{tot}} \frac{C_{m,\text{thresh}}}{C_m + C_{m,\text{thresh}}} . \quad (1.25)$$

where $C_{m,\text{thresh}} = \beta/\delta$, typically chosen small. Hence, at small L , $K \approx K_{\text{tot}}$ and the model behaves as regular KPR, but as C_m increases at moderate L , eventually $C_m \gg C_{m,\text{thresh}}$, and then $K \sim L^{-1}\tau^{-m}$. Inserting in C_N 's expression, this cancels the L dependence and we are left with $C_N \sim \tau^{N-m}$, *i.e.*, an output independent of antigen quantity. This gives output curves $C_N(L, \tau)$ as in figure 1.9.

This version of AKPR is not completely biologically faithful, because the inhibitory coupling is likely mediated by a phosphatase like SHP-1 [138] rather than by a kinase. It is also unclear which kinase could be mediating only the last reaction step and not others. Nonetheless, it makes the adaptive part of AKPR very clear: the point of the inhibitory coupling between receptors is to cancel the L dependence in C_N . Beyond improving concentration-independent sensing of antigen quality, such inhibitory coupling between receptors causes non-linear effects (such as antagonism) in antigen mixtures [139], as we will see in chapters 3–4. Hence, the AKPR framework provides inspiration to find low-dimensional models clarifying how the TCR signalosome processes high-dimensional antigenic inputs to trigger appropriate T cell outputs.

1.4 Low-dimensional models in biophysics and immunity

1.4.1 Definition and examples of low-dimensional models

What do we mean by low-dimensional?

As outlined in section 1.2.3, biological systems evolve in high-dimensional chemical spaces: they face millions of potential different proteins and molecules, requiring large vectors of concentrations and affinities to describe them all. They also have complicated, redundant biochemical signaling networks and gene expression programs, such as the TCR signalosome and the pathways it triggers downstream. A complete pedestrian mathematical description of all these concentrations and biochemical reactions with, *e.g.*, conventional biochemical rate equations or chemical master equations, would require an inordinate amount of parameters and dynamical variables.

Thus, by low-dimensional model, we mean a mathematical description of biochemical inputs, processes, and outputs that uses far fewer parameters and variables than dictated by the naive enumeration of all molecular players or underlying microscopic dynamics. This definition is not very precise, since any model is to some extent lower-dimensional than the complete system, under standard approximations – not modelling unimportant or unrelated proteins, neglecting very small reaction rates, and so on. The kind of low-dimensional models we are most interested in are those where the possibility of dimensional reduction is not at all obvious initially, compresses

the system to very low dimensions (say, 1-10 variables) compared to the initial system, and may require some data-driven analysis to be found. This implies trying to find the “natural variables” of a biological system, which can simplify its dynamics or collapse its statistics onto universal curves at the time and complexity scale of interest [140, fig. 2]. Being able to graph the system’s dynamics or parameter values in 2D plots, as we did with our toy model of immune responses (figure 1.4), is a good sign of low dimensionality.

Examples of low-dimensional models in biophysics

To illustrate this definition, we summarize a few recent examples of successful low-dimensional models in biophysics. In a data-driven approach very similar to what we will do in chapter 2, Seyboldt *et al.* [141] have reduced the expression dynamics of four gap genes in larval *Drosophila* to a two-dimensional embedding, where each axis turns out to be the difference between two genes. They parameterized the dynamics with a polar angle in this 2D space, revealing that interactions between genes are in fact limited, and predicting the effect of various mutations. Also in *Drosophila*, but at a very different level, Berman *et al.* [142] have compressed long recordings of fly postural behaviors down to a 2D space categorizing typical behaviors, using a sequence of dimensional reduction techniques: PCA, wavelet transforms, t-SNE. They also quantified stochastic transition dynamics between these states [143].

In developmental biology, theorists have proposed low-dimensional models of cell fate decisions under the names of gene-free [144] or geometric [145, 146, 147] models. The idea is to represent cell states as attractors in an abstract 2D or 3D phase space, properly disposed to reflect possible transitions between states, and then to build phenomenological equations for the flow of cells in that space. Without explicit reference to complicated gene regulatory networks, and without fitting tons of parameters, this method can successfully predict bifurcations (changes in available states or transitions) that occur as developmental conditions are changed.

Geometric models are connected to the idea of Waddington landscapes in development and hematopoiesis, according to which cell differentiation paths are trajectories travelling down different junctions in a low-dimensional epigenetic landscapes with quasi-potential dynamics [148]. Recent studies have made this idea very concrete by deriving epigenetic landscapes from high-throughput single-cell datasets [149, 150]. A recent study has revealed a Waddington-like landscape in the memory learning process of generalized Hopfield neural networks learn memories, with successive splits taking place in that low-dimensional landscape between prototypes of the learnt classes [151]. These various examples show that effective, low-dimensional representations of

biological dynamics enable theoretical and experimental progress in different organisms and contexts.

1.4.2 Motivating low-dimensional models in biophysics

Why seek low-dimensional models?

Despite being simplifications of the underlying physical, chemical, and biological processes, low-dimensional models should still capture the important dynamics or statistics of the system, and generate quantitative, experimentally testable predictions [152]. Hence, by seeking emergent physical laws and phenomenological models of this kind, we can gain significant interpretability without sacrificing accuracy at the scale of interest. Such streamlined yet predictive theories are also valuable for engineering applications, because they provide a better intuition to propose new design concepts, together with quantitative tools to actually implement them. To take a non-biological example, classical thermodynamics laws and state variables emerging from the statistical physics of microscopic components [153] have enabled uncounted innovations in mechanical and chemical engineering. Hopefully, chapter 4 will show that low-dimensional models can have translational value in immunology too.

Why might there be low-dimensional models?

Why can we hope that simple models have anything useful to say about complicated, messy biological systems? The usual statistical physics trick of self-averaging $\sim 10^{23}$ identical particles does not usually work for biological systems, where the main challenge is the large number of *different* chemical species, each present in small quantities. There are, however, other reasons for which biological systems might have low-dimensional representations.

Complex systems with many dynamical variables tend to have slow manifolds on which dynamics converge, once fast fluctuations have dissipated [154]. This separation of time scales leads to effective low-dimensional dynamics on the slow manifolds. The effect has been well studied in molecular simulations of chemical reactions, where a key goal is to identify reaction coordinates and other collective variables distinguishing stable and transition states [155]. Similar slow manifolds may well underlie low-dimensional models of biological systems, and there is a potential evolutionary explanation for it. Husain and Murugan [156] have shown that slow modes of dynamical systems and soft modes of protein structures restrict the impact of mutations (changes in model parameter or sequence space) to these slow directions. Mutations naturally pushing the

system in these directions therefore have a higher chance of conferring a significant evolutionary advantage, such that biological systems would keep evolving in regions of parameter space maintaining these slow modes in existence.

This canalization could also explain a related phenomenon: the low-dimensional parameter space structure of biological models. It seems these models universally have a few parameter combinations that significantly impact the model’s fitness or fit accuracy, while most other directions in parameter space are irrelevant [157, 158]. This effect, called parameter sloppiness, suggests that biological functions can be captured with few effective model parameters, instead of an exhaustive list of reaction rates – the main challenge being to find these parsimonious parametrizations. At least, this challenge can be somewhat simplified by using well-mixed approximations for biochemical reaction rates, as discussed in section 1.1.7, with effective rate values compensating for the neglected microscopic processes. Parameter space compression algorithms based on sloppiness theory could also automate the task of simplifying large models of biological functions [159].

Criticality is another property that could confer a low effective dimension to biological system. Work on maximum entropy models for neuron populations, protein sequences, and flocks of birds suggests that these systems are poised near criticality, *i.e.*, phase transitions in their macroscopic order or bifurcations of their dynamics [160, 161]. This creates long-ranged correlations between individuals or molecules and presumably reduces the underlying dimensionality of these systems. Moreover, biological systems tune themselves near optimal operating points; some of their parameters are in fact slow dynamical variables regulated by various feedbacks [162]. This is the case, for instance, in neurons, where calcium activity feeds back onto the genetic regulation of ion channel levels, and can tune conductance by changing these levels [163].

We mention two more mechanisms that could contribute to low dimensionality in biological systems. One is the fact that cells have core pathways that integrate inputs from several different sources and control several outputs, such as the NF- κ B and MAPK pathways [164, 104]. Furthermore, in these pathways, interaction networks often have a “hub-and-spoke” or “bow-tie” structure where one protein is a central node interacting with many others, for instance c-di-GMP in bacteria [165] or p53 in mammalian cells [166]. These network structures with bottlenecks have the potential to create some correlations between cellular outputs. Another possible explanation, related to network structure, is the low-rank hypothesis recently proposed by Thibault *et al.* [167]; using mathematical properties of singular value decompositions (SVD), they have shown that several classes of high-dimensional dynamical systems often have interaction matrices with effective low ranks (*i.e.*, a rapid decrease in their singular value spectra). Their work gives mathematical

substance to phenomenological, low-dimensional models, but it also shows that rigorous reductions sometimes should keep a significant fraction of the original dimensionality. In the end, we think that proof is in the pudding: validation of model predictions against experimental data remains the best way to check that a low-dimensional approximation was justified.

1.4.3 Hopes for low-dimensional descriptions of immunity

In this chapter, we have formulated the theoretical challenge posed by immune sensing of high-dimensional inputs, and immune signaling leading to high-dimensional outputs, contrasting it with the classical biophysics problem of concentration sensing. We have then discussed low-dimensional models and their advantage of interpretability. Is there hope that low-dimensional models could also apply to T cell immunity? Could they reveal simpler quantitative principles underlying complex biochemical mechanisms of TCR sensing and cytokine communication? We have a few reasons to hope so.

To begin, we have already seen examples of low-dimensional models of immunity in this chapter. Our two-dimensional toy model of immune excitability could capture the global dynamics of an immune response, without detailing all cell types or signaling pathways (section 1.2.2). Moreover, the KPR and AKPR models of T cell receptor signaling clearly simplify the TCR signalosome’s complicated machinery (section 1.3.2) down to a few phenomenological ODEs, yet they still capture the same key features (*e.g.*, how the output scales with L and τ) found in detailed computational models [116].

Moreover, the general hypotheses mentioned above in section 1.4.2 to explain low-dimensionality in general biological systems most likely apply to immunity as well. There are a few likely dimensional bottlenecks in T cell signaling pathways, starting with the TCR itself: a T cell senses all antigens presented to it with its unique TCR sequence (present in many copies). Any pMHC can thus be reduced to one scalar number, its affinity or binding time τ . Since many pMHCs in the space of all possible amino acid sequences will necessarily have very similar τ s, a complex mixture of pMHCs can probably be coarse-grained to a simpler effective mixture of a few different τ s – this is why we will study binary mixtures in chapter 3. Moreover, the TCR, as the single entry point for antigenic signals, activates all main pathways for proliferation and cytokine production, such as the phosphoinositide 3-kinase (PI3K) pathway [105], so we can expect correlations to exist between all these outputs, effectively reducing their dimensionality as well. These pathways also receive inputs from cytokine receptors, such that extracellular feedbacks from cytokines might not be that complicated, merely adjusting the level of antigenic stimulation.

With the hope of finding underlying simplicity in T cell activation, we seized the unique opportunity to develop new theories offered by our collaborators' new robotic platform, which records high-dimensional T cell cytokine dynamics in response to antigen stimulation. We first focused on deriving a data-driven, low-dimensional model of these dynamics controlled by antigen quality (chapter 2). Building on this first step, we developed improved phenomenological models of early TCR antigen recognition mechanisms, to explain nonlinear effects in our collaborators' measurements of cytokine responses to antigen mixtures – in regular T cells (chapter 3) and in T cells with synthetic receptors engineered for cancer immunotherapy (chapter 4).

Chapter 2

Antigen encoding of T cell activation from cytokine dynamics

We first approached the problem of finding low-dimensional descriptions of T cell responses with a data-driven modelling strategy. Analyzing quantitative measurements of cytokine time series from the Altan-Bonnet lab (section 2.1) with small neural networks, we found a two-dimensional latent space in which cytokine trajectories separated very well according to antigen quality. We parameterized the dynamics in this latent space using ballistic-inspired equations, and found that all parameters were strongly correlated to antigen quality and to the initial rates of cytokine production. I quantified the information content of this latent space, to establish that T cells can recognize (and respond to) a continuum of antigen qualities, which can be subdivided into six non-overlapping classes (corresponding to 2.6 bits of information). The insights offered by the present chapter into the structure of T cell responses – and their strong dependence on initial antigen quality recognition – paved the way for the more mechanism-focused investigation presented in chapters 3 and 4.

This chapter contains my theoretical contributions to *Antigen encoding*, [1]:

- Time series processing and filtering (sections 2.2.2–2.2.3 and B.1), turned into a full pipeline with Thomas Rademaker and Sooraj Achar;
- Mutual information estimation in cytokine time series (2.2.4);
- Robustness analysis of the antigen encoding latent space (2.4);
- Cytokine time series reconstruction from the latent space (2.5 and 2.6.5–2.6.6);
- Developing a simple latent space model and parameter fitting procedure jointly with Thomas Rademaker (2.6.1–2.6.2), and a refined latent space model myself (2.6.4);
- Ranking antigens with a Kendall Tau metric, jointly with Sooraj Achar (2.6.3);
- Channel capacity analysis of antigen encoding in latent space (2.7);
- Earth Mover’s Distance analysis of drug perturbations (2.8.1).

The code that I wrote to produce my theoretical analyses is available on Github [168]. The neural network classifier in section 2.3 was primarily Thomas Rademaker’s work. The experiments (2.2.1) were performed by Sooraj Achar (mainly) and by Emanuel Salazar-Cavazos, Angela Lee, and Grégoire Altan-Bonnet. Most of the quoted excerpts are from modelling-related sections of the supplementary information (SI) which I wrote.

2.1 Introduction: the cytokine code of T cells

2.1.1 Complexity and multiplicity of cytokine mechanisms

Within the complicated, multifarious world of chemical signals in the immune system, cytokines play a crucial role: immune cells communicate with each other by producing and detecting these small extracellular messenger proteins. When T cells get activated by antigens recognized through their T cell receptor (TCR), they start producing a complex cocktail of cytokines to alert other cells and sustain their own activation and proliferation [105]. However, we lacked a quantitative framework to understand how TCR antigen recognition, initiated within minutes [131], can map diverse antigenic inputs to high-dimensional cytokine responses spanning days. Therefore, we decided to tap into the cytokine communication network to reverse-engineer the principles of TCR-mediated chemical sensing and information transmission through cytokines.

As mentioned in section 1.2.3, there are over 30 different cytokines in the human (and mouse) immune system, and each has been associated with various functions and idiosyncratic patterns of secretion by a whole body of biological studies [102]. Without going into the minute details of production and consumption mechanisms, a brief overview suffices to show that the cytokine code can be tremendously complex, and that finding the principles connecting T cell activation to cytokine dynamics requires quantitative, physics-inspired methods.

The spatiotemporal evolution of the concentration of a given cytokine, $c(r, t)$, during a T cell population response in some region of the body (*e.g.*, a lymph node or an infected tissue), could in principle be described by a relatively simple partial differential equation [169, 98, 4], neglecting stochastic effects,

$$\frac{\partial c}{\partial t} = \vec{\nabla} \cdot (D \vec{\nabla} c) + k_+(c, r, t) - k_-(c, r, t) , \quad (2.1)$$

where the first term is diffusion with diffusion constant D , and $k_+(c, r, t)$ and $k_-(c, r, t)$ represent the production and consumption of cytokine c . This equation may be useful in very simple scenarios: for instance, for a single activated cell surrounded by consumer cells, it establishes a typical length scale over which the produced cytokines can diffuse before being captured: $\lambda = \sqrt{\frac{D}{kn_c}}$,

where k is the consumption rate and n_c is the density of consuming cells [169]. However, in general, equation (2.1) is far from telling the complete picture of cytokine dynamics, as each term encompasses multiple biological processes going on in parallel inside every immune cell and in the extracellular medium, with no systematic way of determining *a priori* which processes are relevant or not.

Several mechanisms contribute to the production term $k_+(c, r, t)$. In each T cell, upon activation of key signaling molecules (*e.g.*, ZAP-70, LAT, PLC γ , ERK) by an antigen-bound TCR, various transcription factors (such as NFAT or NF- κ B) translocate into the cell nucleus to execute the gene expression program corresponding to T cell activation [170]. It includes the synthesis of cytokines, which are then exported (through vesicles) outside of the cell, into the extracellular medium [171]. This program also comprises the acquisition of effector functions and cell proliferation [105]; hence, the density of active, cytokine-producing cells $N(t)$ will increase over time, increasing the overall rates $k_{\pm}(c, r, t) \propto N(t)$. In addition to TCR antigen recognition, co-stimulatory signals (such as CD28 binding) and immune checkpoints (such as CTLA-4 or PD-1) modulate the overall activation [84] and thus the cytokine production rate. Moreover, antigen-presenting cells (APCs) also produce several cytokines (especially IFN- γ , IL-6, TNF), either through transcription-translation, or through recycling of previously captured protein copies [172].

Importantly, T cells also have receptors with which they detect and internalize cytokines – those produced by other cells (paracrine signaling) but also their own (autocrine signaling) [102]. The consumption of cytokines by T cells and by antigen presenting cells directly enters the degradation term $k_-(c, r, t)$, although the detailed dynamics can become complicated. Receptor sub-units are up-regulated only following T cell activation [170, 105], and the receptors for different cytokines often compete for the same receptor sub-units, such that consuming some cytokines can alter the sensitivity of a T cell to others [64, 65]. Antigen presenting cells also consume cytokines and contribute to k_- [84]. Furthermore, cytokine sensing impacts the production term k_+ as well: inflammatory cytokines such as IL-2 or IFN- γ act as additional activation signals for T cells [105]¹, while anti-inflammatory cytokines like IL-10 can bring down the activation level [173]. Complex feedback loops can thus play out in the T cell population due to cytokines: following initial TCR antigen recognition, cytokine production is turned on, but cytokine receptors are also upregulated, such that cytokine consumption increases, maybe enough to dominate the production rate, but also providing supplementary activation signals to more weakly reactive T cells. This can go on for each cytokine and with coupling effects between them.

¹This explains autocrine cytokine signaling: T cells export cytokines to the extracellular medium to bind their own surface receptors and thus sustain their activation through cytokine receptor activation pathways.

Consequently, to be mechanistically exhaustive, the overall production and degradation rates, k_+ and k_- , would have to be complicated functions of many dynamical variables, external or internal to individual cells, beyond the cytokine c itself. Such a ground-up approach is unlikely to encompass the whole complexity of cytokine dynamics while still being comprehensible and instructive about the core principles. It takes cytokine concentrations, production and degradation rates too literally, while these may not be the most informative variables or the most parsimonious description of the system. For this reason, we explored a different approach in this chapter, taking as a starting point our collaborators' experimental measurements of overall high-dimensional cytokine dynamics in a T cell and APC population, and trying to reconstruct how they were connected to the initial antigen recognition properties.

2.1.2 Connecting cytokine dynamics to antigen quality

In particular, we sought to determine whether T cell activation, and ensuing cytokine dynamics, could in fact hide a low-dimensional structure. In our physics style of inquiry, striving to keep the number of parameters to a minimum, we thought that such a structure would be determined by only a few immune parameters such as the antigen quality (strength or binding time τ) and quantity (abundance L or density on the surface of APCs) as well as the initial number of T cells, the kind of APCs used, and so on. We hypothesized that T cell responses could indeed be low-dimensional after all, since the TCR effectively acts as a dimensional bottleneck, through which T cells must sense all pMHC antigens they encounter before producing the aforementioned high-dimensional activation genetic program (figure 2.1A). We also hypothesized that antigen quality would be the main driver of the T cell response, as their primary function for defense against pathogens is to respond to foreign peptides despite their small number (at the start of an infection), without ever responding to the large amounts of self-derived peptides normally presented by APCs; in other words, to sense antigen quality without being sensitive to antigen quantity (figure 2.1B).

“Current understanding of T cell responses defines three functional classes of antigen — non-agonist, weak and strong agonists — based on their impact on thymic selection: death by neglect, positive selection, and negative selection, respectively [174]. This observation led to the concept of antigen quality, which has proved useful to predict successful eradication of viruses or tumors but has been difficult to predict from antigen sequences [175]. Estimating antigen quality usually involves measuring the antigen levels triggering 50 % of the maximal immune response (EC_{50}) [174, 175, 176]. However, antigen quality should be defined as an absolute property, as

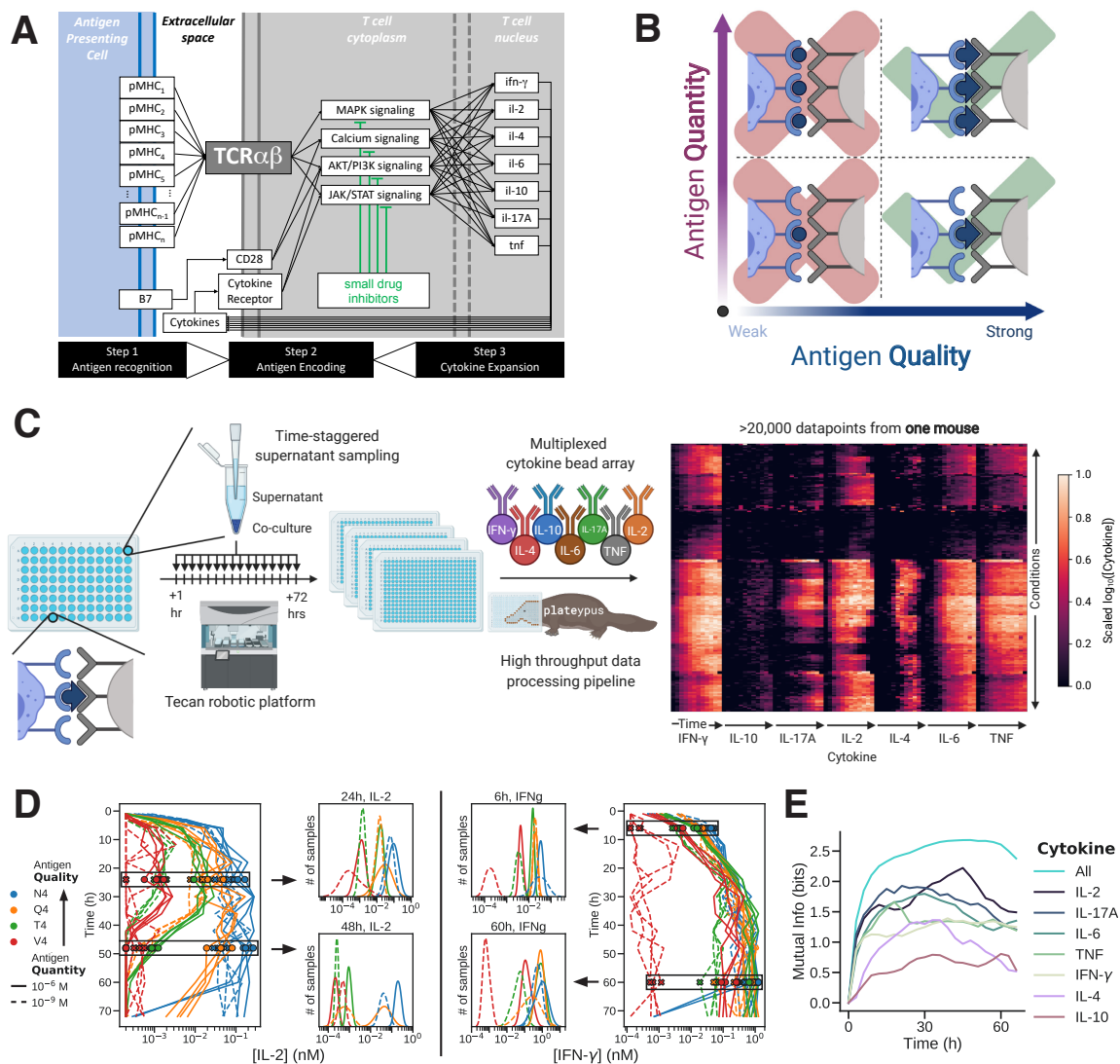


Figure 2.1: “Robotic platform to quantify the dynamics of antigen-driven activation of T cells. (A) The TCRαβ signaling machinery of T cells engages a repertoire of peptide-MHC (pMHC) ligands (antigens), triggers a signal transduction cascade and activates a functional response (*e.g.*, cytokine secretion). (B) T cells must deconvolve quality and quantity to accurately discriminate between antigens. (C) Robotic platform to multiplex immunological settings and dynamic measurements (*e.g.*, cytokines). This platform generates large datasets quantifying the inflammatory outputs generated by T cell activation *ex vivo*. We used a custom-designed pipeline to compile high-throughput bulk cytokine, surface marker and single cell data (section 2.2.1). (D) Single readouts (here IL-2 and IFN-γ) of OT-1 T cell activation at varied time points for different antigenic peptides [SIIVFEKL(V4), SIITFEKL(T4), SIIQFEKL(Q4) and SIINFEKL(N4)] (left) failed to deconvolve antigen quality and quantity, no matter the time point (center); this resulted in partial overlap between amount of secreted cytokines depending on antigen quality and complete confusion depending on antigen quantity (right). (E) Mutual information between antigen quality and secreted cytokines, combined vs individually.” (*Antigen encoding*, [1])

higher quantities of weak antigens do not functionally match low quantities of strong antigens (Fig. 2.1B) [62].” (*Antigen encoding*, [1])

Challenging the canonical thinking in terms of three antigen classes, we wanted to determine whether T cell populations can actually make finer distinctions between antigen qualities, and whether cytokine dynamics reflect this refined readout of antigen quality, independent of quantity.

To answer these questions, we needed not only static, end-point snapshots of immune activation, but time-dependent measurements:

“Although high-dimensional snapshots of immune cells are routinely acquired [177, 20, 178], they are often ill-suited to elucidate the time-dependent processes of immune responses. For instance, expression of the Programmed-Cell-Death-1 (PD-1) protein correlates with either T cell activation or exhaustion depending on the timing and context [178]. Such complexity stems from the multivariate responses triggered by multiple ligands (antigens) engaging a monogenic T cell receptor (TCR) (Fig. 2.1A). Downstream responses drive distinct patterns of activation dynamics encoding different stimuli using feed-forward processing [179, 180].” (*Antigen encoding*, [1])

We however lacked such a set of experimental measurements, and the corresponding quantitative theoretical framework, to connect antigen parameters to T cell responses over days. Therefore, with our NCI collaborators, we

“set out to robotically generate high precision, multidimensional kinetic data of T cell activation, and applied supervised machine learning approaches [181] to learn a dynamic encoding of information [182, 183]. We focused on modeling the responses of T cells to antigens on target cells.” (*Antigen encoding*, [1])

2.2 Experimental platform and data processing

2.2.1 Robotic experimental platform

To start systematically mapping out T cell responses, the Altan-Bonnet lab, with Sooraj Achar and Angela Lee in particular, built a robot-assisted platform to track the dynamics of multiple immune molecules in parallel. Figure 2.1C illustrates the main experimental steps. We provide

here a summary of this platform; more details on the experimental protocols are provided in the supplementary information of [1], section 1.

In a typical experiment, they first (manually) prepared cocultures of T cells and antigen-presenting cells that were loaded with various qualities and quantities of antigens. Some experiments also included cocultures with varying initial T cell numbers (10^5 by default) to assess the impact of this parameter as well. To load antigen peptides on the APCs, they were pulsed, that is, APCs were placed in a solution with a given concentration of peptides, which then bind to the MHC molecules [84] within a few hours of incubation. By default, the T cells came from TCR-transgenic mice expressing a unique TCR sequence called OT-1, which is specific to the ovalbumin-derived antigen peptide (OVA, amino acid sequence: SIINFEKL, called N4 for shorthand), the APCs were splenocytes (spleen cells) from B6 mice, and the peptides loaded on them were variants of the OVA peptide with well-characterized affinities [174, 176]. These variants are named A2, Y3, etc.; the letter specifies an amino acid substitution to the default sequence SIINFEKL, and the digit specifies the position of this substitution. For instance, A2 corresponds to peptide sequence SAINFEKL. Over the course of the project, we also tested a wide range of other TCR lines, APC types, and peptide repertoires.

Second, right after starting the immune reaction by mixing the APCs and T cells, our collaborators placed the cocultures in their Immunotron robotic platform (figure 2.2), which has an automated arm to handle pipetting and moving plates. The robot took care of sampling supernatant or collecting cells from the cocultures at pre-defined time points (typically one per every 2-6 hours), replacing the plates in an incubator and freezing the collected samples between sampling operations.



Figure 2.2: Photo of the Immunotron robotic platform built by the Altan-Bonnet lab, with annotations defining its main parts. (*Antigen encoding*, [1], SI)

Third, after completion of the time series, various reagents were added (manually) to the collected samples, such as cytokine capture beads – polystyrene beads coated with antibodies specific to a given cytokine – and fluorescent antibodies to tag the bead-captured cytokines or cell surface markers. Then, these samples were analyzed by flow cytometry². The flow cytometer used by our collaborators in most experiments had 5 lasers and 18 detection frequency channels, allowing for the multiplexed measurement of several cytokines or cell surface markers in parallel.

Fourth, the raw counts of the flow cytometer were sorted, converted to absolute concentration units (molar) based on calibration curves, and organized into convenient data tables by a Python pipeline, *plateypus*, developed by Sooraj Achar [184].

In the end, the datasets we received for theoretical analysis typically consisted in the supernatant concentration of 7 cytokines – IFN- γ , IL-2, IL-4, IL-6, IL-10, IL-17A, and TNF – at ~ 12 time points spanning 72 hours, for an array of cocultures prepared with varying antigen qualities, antigen quantities, and initial T cell numbers (figure 2.1D). These cytokines were chosen by our collaborators for their biological relevance and importance in the literature about T cell immunology. Some datasets were supplemented with single-cell expression levels of over 10 different surface markers, such as IL-2 receptors (IL-2R α , also called CD25) or the early activation marker CD69.

2.2.2 Time series processing

Most of the time series preprocessing steps were initially developed for my Master’s thesis [47], which used some early datasets from the Altan-Bonnet lab, for theoretical analyses completely distinct from those in the present thesis. Thomas Rademaker added preprocessing steps to correct missing data points [4], and with Sooraj Achar, we turned this algorithm in a more systematic Python pipeline [185]. Here, we provide a summary (written for the SI of [1]) of the time series smoothing and interpolation steps. The purpose of these steps was to mitigate noise in experimental measurements, rescale datasets similarly to correct for batch effects, and use interpolation to generate more time point samples and thus have more data on which to test and train machine learning and dimensionality reduction methods.

²A flow cytometer is a microfluidic instrument that can channel one cell at a time, with a high throughput, through an array of lasers and light detectors, allowing for the detection of single-cell properties. Cytokine beads are thus necessary to measure extracellular concentration: beads flow in the microfluidic channel, and the amount of fluorescence detected per bead is proportional to the amount of cytokine bound to the bead, itself proportional to the concentration in the supernatant.

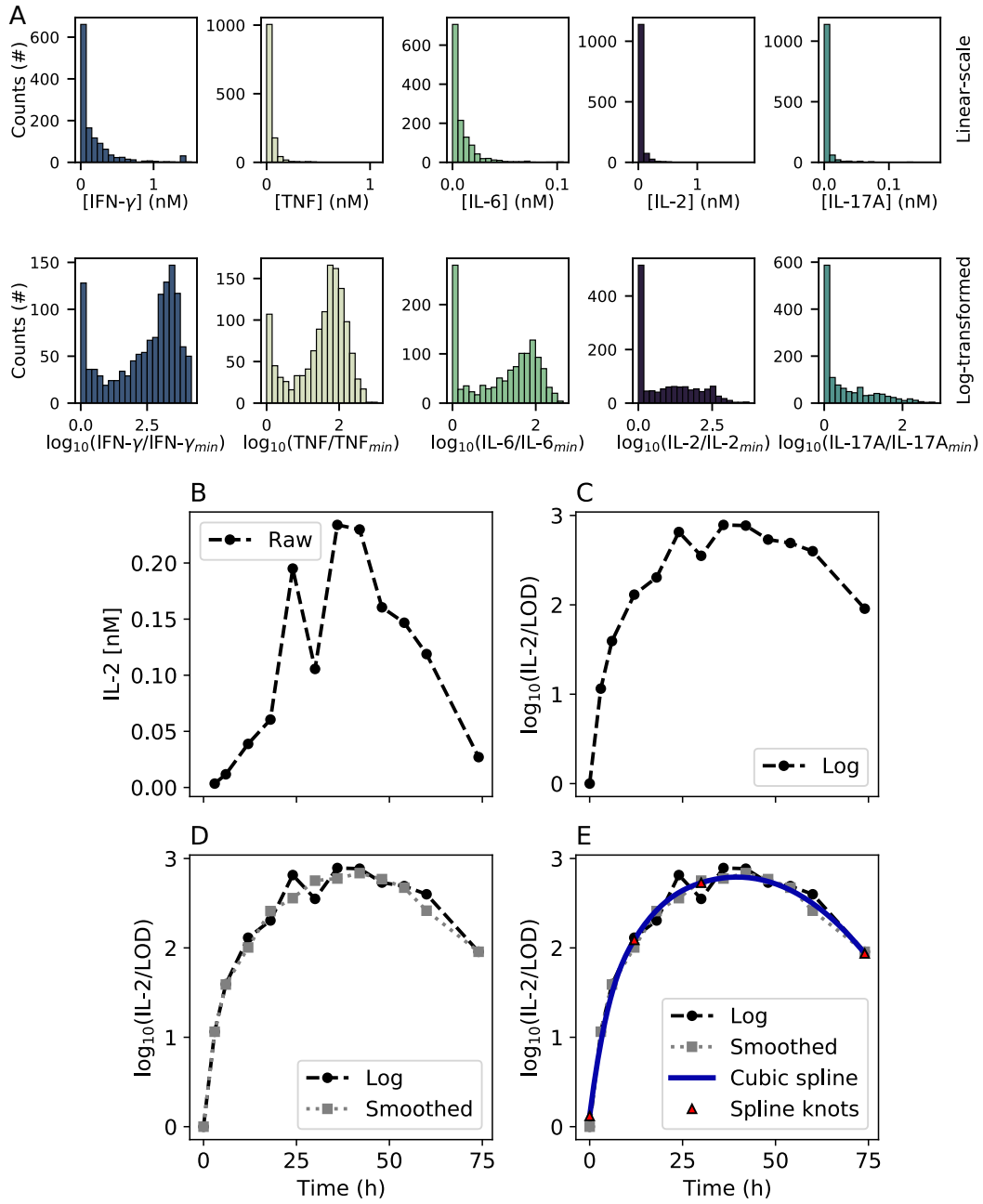


Figure 2.3: “ Time series processing steps described in section 2.2.2. (A) Histogram of cytokine concentration measurements in the training datasets, in linear scale (top row) and after log transformation (bottom row). (B) Time series for the concentration of IL-2, as measured with a calibrated cytokine bead array (here, response to peptide N4 at pulse concentration 10 nM). (C) The same time series, after taking the logarithm of the concentration normalized by the lower limit of detection for that cytokine (3.0 pM in that experiment). (D) The result of applying a moving average of width 3 on the log-transformed time series. (E) The cubic B-spline smoothing interpolation fitted to the log-transformed and smoothed time series, with relative smoothing parameter $r = 0.2$. ” (*Antigen encoding*, [1], SI)

“Cytokine concentration measurements exhibit long-tailed distributions which become more evenly distributed in logarithmic scale, as can be seen by comparing the first and second rows of figure 2.3A. Therefore, we first took the logarithm of concentrations normalized by the lower limit of detection (LOD), as shown in figure 2.3C. We then applied a moving average filter with the smallest possible window width (3 time points), to minimize edge effects on our experimental time series typically containing 12 time points. Fig. 2.3D shows an example of a time series after applying this moving average.

Because there remains visible noise after this first step of filtering, we used an interpolation algorithm that both interpolated and smoothed further the data points [186]. This algorithm (as implemented in *SciPy* [187]) fits a cubic B-spline function $c_k(t)$ to data points $y(t_i)$, $i \in \{1, 2, \dots, n\}$, automatically selecting the minimal number k of internal knots that allows c_k to satisfy the following criterion:

$$\sum_{i=1}^n [y(t_i) - c_k(t_i)]^2 \leq s \quad (2.2)$$

where s is the tolerance (smoothing parameter). Note that the curve does not have to go through all data points. We set s equal to some fraction r of the residuals between the the data points before ($y(t_i)$) and after ($z(t_i)$) applying the moving average filter,

$$s = r \sum_{i=1}^n [z(t_i) - y(t_i)]^2 \quad (2.3)$$

which ensures that s has an appropriate scale compared to the amount of noise in the data. We found that a value of $r = 0.5$ gives smooth curves $c(t)$ which still reasonably capture the essential kinetic features of the experimental data. Fig. 2.3E illustrates this last processing step [...]” (*Antigen encoding*, [1], SI)

In summary, these preprocessing steps turned noisy experimental time series of 7 cytokine concentrations $y(t)$ into smooth, time-continuous interpolations $c(t)$ of $\log_{10} \left(\frac{y(t)}{\text{LOD}} \right)$, where LOD is the lower limit of detection for that cytokine. Some datasets required additional processing steps, mainly to deal with experimental variability; these are described in appendix B.

2.2.3 Excluding IL-4 and IL-10 from inputs

Before embarking on machine learning analyses to search for a low-dimensional representation of cytokine dynamics, we sought to reduce the dataset dimensionality based on interpretable feature selection. An immediate simplification came from realizing that two cytokines, IL-4 and IL-10, were not produced at detectable levels in most experiments with mouse CD8⁺ T cells; we could therefore discard them before focusing on more informative cytokines. We motivated further this choice with information theory (next subsection), and further below in subsection 2.4.3.

2.2.4 Mutual information estimation from cytokine trajectories

We performed an information theoretic analysis of cytokine time series, to confirm that IL-4 and IL-10 were indeed the least informative cytokines, and to ensure we could hope to extract information about antigen quality from these dynamics.

We set up the calculation of mutual information between cytokines and antigen quality in the following way. We define Q to be the random variable giving the quality of an unknown peptide; here, this is a discrete variable equal to one of the peptides in the OVA family (N4, A2, Y3, etc.). We also define $\mathbf{X}$ to be a random vector of quantities derived from cytokine time series. We consider two possible choices of $\mathbf{X}$ in particular: either the concentration of one cytokine, *e.g.*, $\mathbf{X} = c_{\text{IFN-}\gamma}(t)$, or the vector of concentrations of all cytokines, $\mathbf{X} = (c_{\text{IFN-}\gamma}(t), c_{\text{IL-2}}(t), \dots)$, where the $c(t)$ are the log-transformed, smoothed and interpolated cytokine concentrations at a given time.

“We will be interested in the mutual information between Q , the input, and various such $\mathbf{X}$ vectors, the outputs – the cytokine response caused by the input peptide Q . Intuitively, mutual information equals the amount of entropy (or uncertainty) dissipated about Q when $\mathbf{X}$ is measured, since $\text{MI}(Q; \mathbf{X}) = H(Q) - H(Q|\mathbf{X})$.

The mutual information (MI) between Q and $\mathbf{X}$ is computed, in bits, as

$$\text{MI}(Q; \mathbf{X}) = \sum_q p_Q(q) \int d^k \mathbf{x} f_{\mathbf{X}|Q=q}(\mathbf{x}) \log_2 \left(\frac{f_{\mathbf{X}|Q=q}(\mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})} \right). \quad (2.4)$$

Here, $p_Q(q)$ is the probability mass function (*pmf*) of the input variable, the quality Q , while $f_{\mathbf{X}|Q=q}$ is the conditional output probability density function (*pdf*) when Q takes on one of its possible values q . The marginal *pdf* $f_{\mathbf{X}}$ derives from the other

two distributions as

$$f_{\mathbf{X}}(\mathbf{x}) = \sum_q p_Q(q) f_{\mathbf{X}|Q=q}(\mathbf{x}). \quad (2.5)$$

To accurately estimate the mutual information, we used our own Python implementation of a bin-less, distribution-free mutual information estimator proposed in [188] and adapted to the discrete input case in [189]. We take data points coming from all concentrations of a peptide $Q = q$ as different samples from the same distribution $f_{\mathbf{X}|Q=q}$." (*Antigen encoding*, [1], SI)

The central idea of this bin-less MI estimator is to compute, for each data sample, the size of the neighborhood extending up to the k th nearest-neighbor point of the same category, and then count how many data points of any category are present in this neighborhood. Intuitively, if all points within the neighborhood are from the same category, the MI will be large, while many neighboring points from other categories indicate overlapping marginal distributions $P(\mathbf{X}|Q)$ and thus lower MI. k is a hyperparameter that should remain (much) smaller than the total number of data points per category. Values of k between 3 and 6 typically work well [189].

Using this algorithm,

"we computed mutual information over time, $MI(Q; \mathbf{X}(t))$, with $\mathbf{X}(t)$ representing the vectors of cytokines or latent space variables at time t . To do so, we aggregated time points within a sliding window of 3 h as coming from the same distribution. In Fig. 2.1E, we show the information on Q over time for various $\mathbf{X}$ built from cytokines: each cytokine concentration alone, and the vector of the five cytokine concentrations. We found that cytokines IL-4 and IL-10 were not informative (justifying further why we discarded them from the latent space construction). For the five other cytokines taken jointly, the mutual information roughly peaked at $t \approx 45$ h at a value of slightly above 2.5 bits. This corresponds to $2^{2.5} > 5$ categories of peptides that can be perfectly discriminated." (*Antigen encoding*, [1], SI)

In short, we found that we could discard IL-4 and IL-10, and that high-dimensional cytokine dynamics contained significantly more information about the absolute antigen quality than any single cytokine. This meant that T cells do not rely on a single cytokine to fully encode antigen quality, and that we needed to consider the remaining five-dimensional space of cytokine dynamics to

really dissect the structure of T cell responses. To make initial progress in this direction, we utilized machine learning techniques (section 2.3), which then unlocked more traditional theoretical approaches (sections 2.6 *sq.*).

2.3 Latent space discovery with machine learning

2.3.1 Using the cumulative time integral of cytokines

Our initial attempts to decode cytokines relied on *in silico* evolution of biochemical networks [47] with the φ -evo algorithm, which simulates rounds of mutation and selection of candidate models. We found several models able to extract a coarse-grained measure of antigen quality from cytokine inputs simply by effectively taking the average of cytokine time series over a long time window. In the limit of infinitely slow averaging, this amounts to computing the cumulative time integral of cytokines. These methods revealed that information could be decoded from cytokine dynamics by relatively simple schemes; however, the degeneracy and variability of network solutions encouraged us to seek more robust decoder models.

We retained an important insight from this preliminary work: time-integrating cytokines improves their separation in terms of antigen quality. Therefore, instead of working directly with (log-transformed) cytokine concentrations, we used the cumulative time integrals of cytokines as input features in machine learning models:

$$C_i(t) = \int_0^t dt' c_i(t') = \int_0^t dt' \log_{10} \left(\frac{y_i(t')}{\text{LOD}_i} \right) \quad (2.6)$$

where $c_i(t)$ is the smoothed, log-transformed, interpolated concentration time series of cytokine i , that is, $\log_{10} \left(\frac{y_i(t')}{\text{LOD}_i} \right)$ where $y_i(t')$ are the experimental measurements (with the understanding that the time interpolation is performed on the log-transformed values).

2.3.2 Neural networks reveal antigen encoding in a latent space

The preprocessed cytokine time integrals displayed some organization as a function of antigen quality (figure 2.4A), but we wanted to reduce further the dimensionality of these cytokine dynamics, which were still too complex to interpret and decipher. Therefore, to reverse-engineer the relationship between antigen quality and cytokines, we trained neural networks to perform supervised classification of antigen quality from cytokine dynamics. Thomas Rademaker initially performed this important work (I replicated it in the revision stage of the project); the details are

provided in his PhD thesis [4] and in the SI of [1]. In this subsection, we provide a summary of these findings.

We trained three-layer perceptrons to predict the antigen responsible for a given cytokine response (figure 2.4B). The input layer takes a single time point from $\mathbf{C}(t)$, the cumulative time integrals of the five log-transformed cytokines – time points from the same time series are given one at a time to the network and are classified independently. The output layer is a softmax layer predicting the probability that the input came from one of six OT-1 antigens included in the training set and spanning a wide range of affinities (from E1, a nearly null peptide, to N4, the original strong OT-1 agonist). Mathematically, to transform an input into a prediction, the network projects inputs $\mathbf{C}(t)$ (at a given time point t) to the intermediate, or “hidden”, layer $\mathbf{h}(t)$ of size n_h as

$$h_i(t) = \tanh \left(\sum_{j=1}^5 P_{ij} C_j(t) + b_i \right) \quad \forall i \in \{1, \dots, n_h\}, \quad (2.7)$$

where the P_{ij} are the learned input layer weights (from 1st to 2nd layer) and the b_i are learned biases. The last layer computes a softmax transform of projections of the hidden layer to generate predicted probabilities g_q that the input came from each antigen class q , for each of $n_q = 6$ classes in the training dataset:

$$g_q(t) = \frac{\sum_{j=1}^{n_h} W_{qj} h_j(t)}{\sum_{q'=1}^{n_q} \left(\sum_{j=1}^{n_h} W_{q'j} h_j(t) \right)} \quad \forall q \in \{1, \dots, n_q\}. \quad (2.8)$$

The W_{ij} are the learned output weights (from 2nd to 3rd layer). The neural network weights were optimized by stochastic gradient descent to minimize a cross-entropy loss function (*i.e.*, to predict the correct antigen as often as possible with as little uncertainty as possible), using the *scikit-learn* package in Python [190]. The data used for training comprised a total of 80 time-smoothed time series for six different OT-1 antigens (N4, Q4, T4, V4, G4, E1), pooled from 6 different experimental repeats, and sampled every hour for 72 hours. The neural network was then tested and cross-validated on at least 3 other OT-1 experimental repeats.

By systematically testing all possible layer sizes, we found that a two-node bottleneck in the intermediate layer ($n_h = 2$) achieved excellent classification accuracy – with over 80 % of time series correctly classified [4] – despite the dimensionality reduction. But beyond considerations about classification performance which are conventional in machine learning, we were most interested in what the latent space revealed. We projected the time series of cumulative cytokine integrals in the

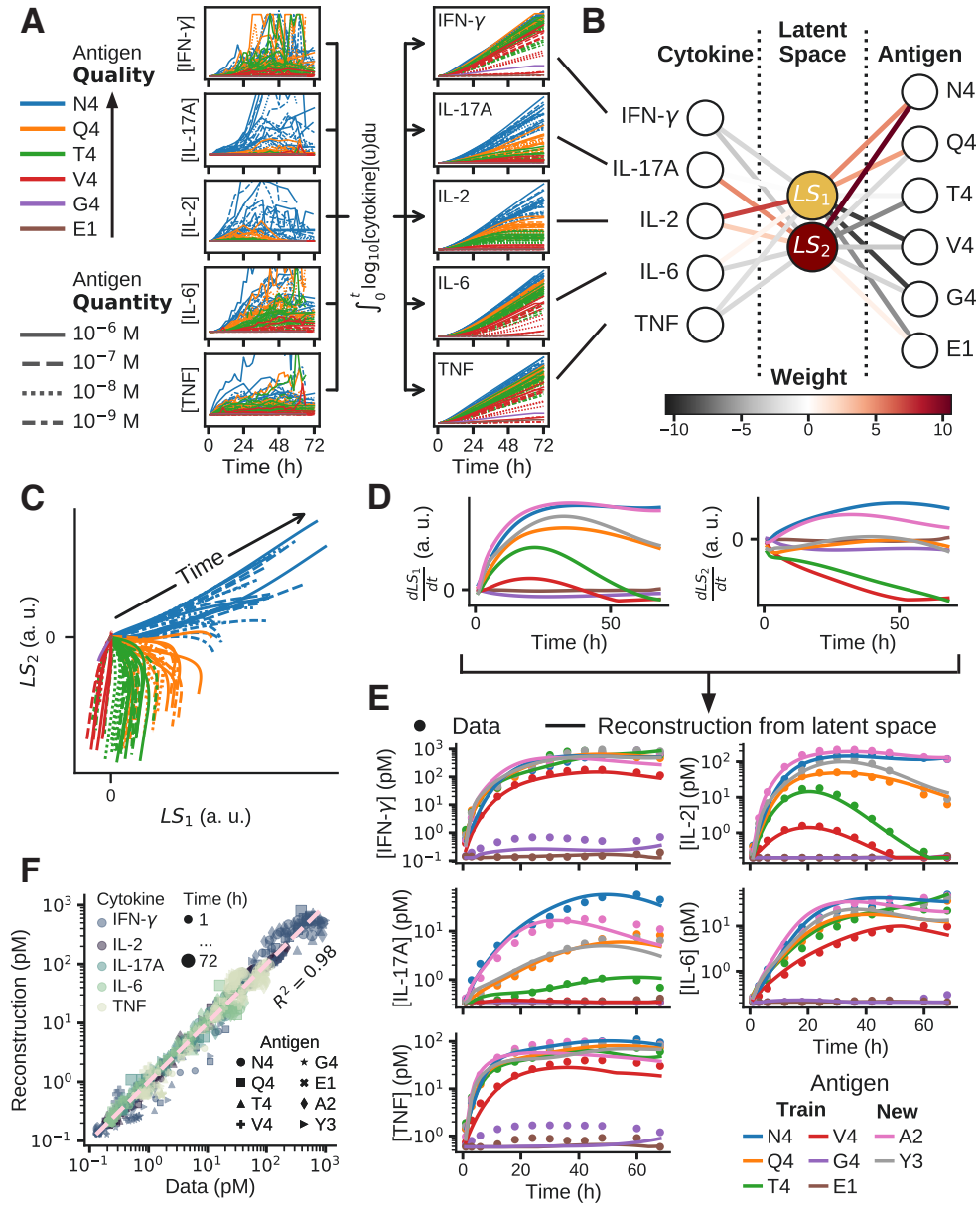


Figure 2.4: “Antigen encoding from a neural network that compresses the dynamics of T cell activation according to quality, independently of quantity. (A) Dynamics of five cytokines for different quality and quantity of OT-1 antigens (left) are pre-processed and (B) classified according to antigen sequences through a trained three-layer neural net with a 2-dimensional bottleneck (latent space LS_1 , LS_2). (C) Time courses of compressed cytokine dynamics in the latent space stratify T cell activation according to antigen sequence (quality), independently of antigen quantity [. . .]. (D) Examples of latent space trajectories for different antigen qualities. (E) Nonlinear reconstruction of cytokine time series from latent space trajectories yielded an accurate fit with experimental data (Fig. 2.11), (F) demonstrating completeness of antigen encoding [4 peptide antigens are introduced here: EIINFEKL(E1), SIIGFEKL(G4), SIYNFEKL(Y3) and SAINFEKL(A2)].” (*Antigen encoding*, [1])

latent space, neglecting the tanh activation function and the biases b_i (time points of a series are projected individually, then connected back together in the latent space). This defined two latent space coordinates, LS_1 and LS_2 ,

$$LS_i(t) = \sum_k P_{ik} C_k \quad (i \in \{1, 2\}) . \quad (2.9)$$

In these latent space coordinates,

“2d trajectories separated well according to antigen sequences and were independent of antigen quantity over 3 orders of magnitude (Fig. 2.4C). We refer to such well-separated projections in latent space as ‘antigen encoding’ to emphasize how T cell responses are classifiable according to antigen quality (as determined by antigen sequences). This encoding reliably distinguished antigens even at high dose when other markers typically saturate (Fig. 2.4D).” (*Antigen encoding*, [1])

Discovering the latent space antigen encoding property was a crucial simplification that unlocked further theoretical understanding of T cell responses. We worked to dissect this striking structure and gain a full quantitative handle on T cell cytokine dynamics. We first made sure it was robust to preprocessing and machine learning model choices (section 2.4) and that it was preserving information about cytokines (section 2.5). We then wanted to write down equations accurately describing the dynamics in latent space (section 2.6), which could be used to generate realistic cytokine time series (section 2.6.6). Then, we quantified the information capacity of T cell responses and the typical antigen classes they can distinguish (section 2.7), and we finally explored the universality of antigen encoding across immune settings (section 2.8).

2.4 Robustness of the antigen encoding latent space

To determine the robustness and reliability of the latent space found via machine learning, we explored other dimensionality reduction methods and variations in our processing choices. Thomas Rademaker performed some robustness analyses in his thesis, for instance showing how the latent space is refined as more cytokines are added [4]. The verifications described in this section, however, are distinct and are my own work.

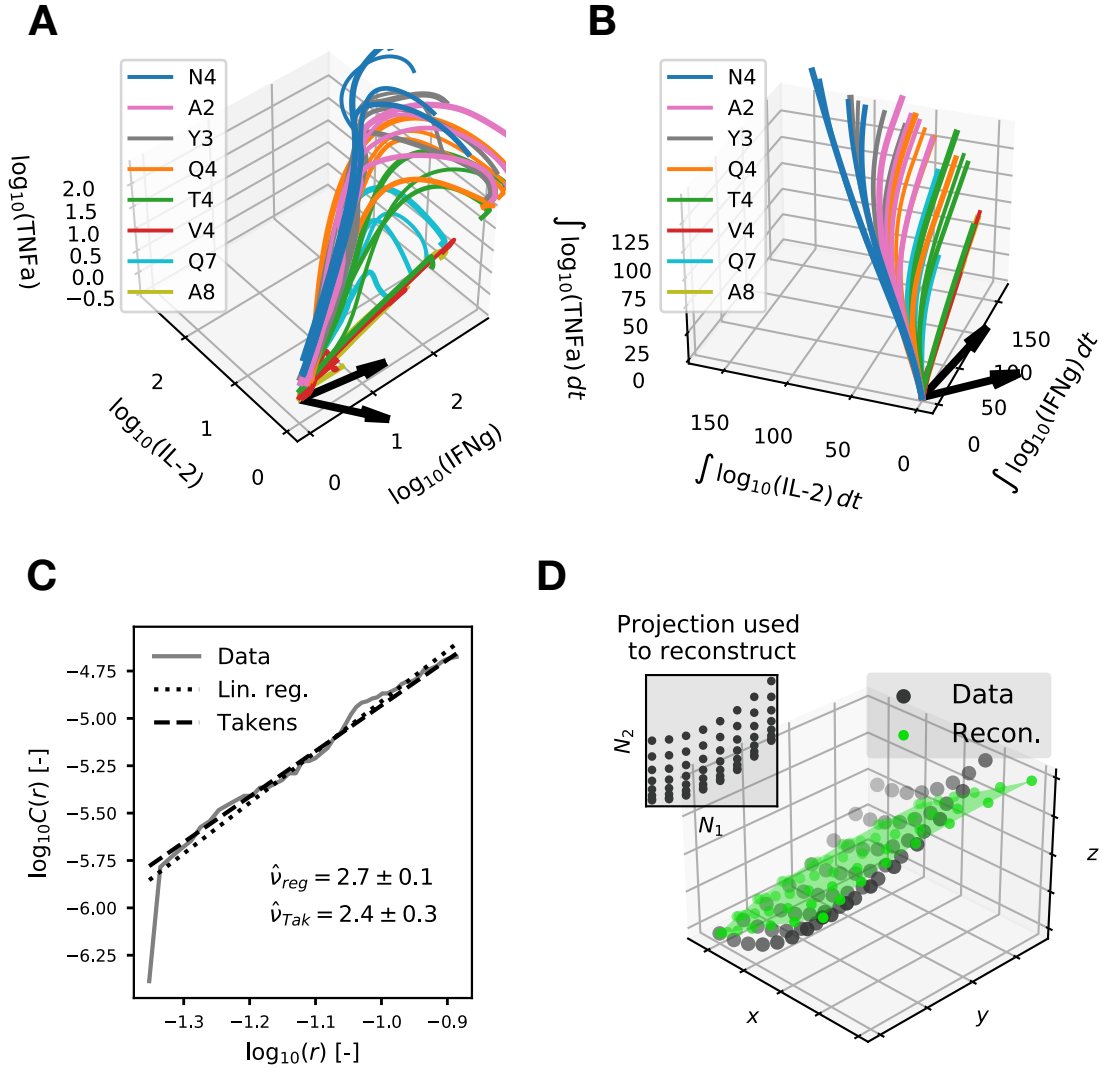


Figure 2.5: “Dimension of the cytokine manifold and linear reconstruction procedure. (A) Three-dimensional plot of cytokine concentration time courses with IFN- γ , IL-2, and TNF shown; the plots are similar when IL-6 and IL-17A are included. The arrows represent the latent space axes on which the neural network projects data. (B) Three-dimensional plot of cytokine time integrals (from 0 to t , plotted for increasing t) with IFN- γ , IL-2, and IL-17. See supplemental movie for an animated version of these two graphs. (C) Linear regression and Takens estimator of the log-log scaling of the correlation function $C(r)$ at small cytokine distances, giving estimates of the manifold’s intrinsic dimension. The line fit for the Takens estimate is illustrative: its slope is fixed by the estimator, and the intercept is found later by linear regression. (D) Linear least-squares fitting of a parabolic surface as an illustration of why the non-linear cytokine reconstruction method introduced in section 2.5.2 is necessary.” (*Antigen encoding*, [1], SI)

2.4.1 Cytokine trajectories lie on an approximately 2d manifold

To begin, we sought to determine whether the underlying dimensionality of cytokine time series really was two-dimensional. Why could cytokine dynamics be compressed by the mere linear mapping (matrix product) given in components form in equation 2.9? As shown by

“the three-dimensional plots in figure 2.5A and B, cytokine trajectories – time integrals and concentrations alike – approximately lie on a bidimensional manifold. The “true” dimensionality of the data can be quantified by estimating the Hausdorff dimension of the structure, as described in [191]. Classically, calling $C(r)$ the fraction of neighboring points within an Euclidean distance r of a point on a manifold, the Hausdorff dimension ν is obtained from the scaling of $C(r) \sim r^\nu$ with r in the $r \rightarrow 0$ limit (for instance, $C(r) \sim r^2$ for a plain square). A more refined estimate ν_T proposed by Takens removes finite size effect [191].

We computed both estimates from the correlation function of the 5D cytokine time integrals, which are shown in figure 2.5B; the scaling of $C(r)$ is shown in figure 2.5C. We obtain $\nu_H = 2.7 \pm 0.1$ and $\nu_T = 2.4 \pm 0.3$ with the linear fit and the Takens estimate, respectively. This confirms the visual intuition that the time integrals of our five cytokines are correlated on a manifold approximately of dimension $\nu = 2$, and explains why a projection onto a two-dimensional latent space was so successful.”
(*Antigen encoding*, [1], SI)

2.4.2 Comparing dimensionality reduction methods

“To further understand the structure of our data, we found it interesting to compare how unsupervised learning would perform compared to our supervised learning described above. Such a comparison is shown in Fig. 2.6. We took our training features (time integral of log cytokines) and performed a Principal Component Analysis (PCA: Fig. 2.6A-B) or ran an auto-encoder (A-E: Fig. 2.6C-D). In both cases, we indeed found a very similar structure for the first two principal components or the autoencoder latent space, separating trajectories according to ligand quality, independently of ligand quantity, in a very comparable way to our learnt latent space (Fig. 2.6E-F). But when we quantified the amount of Mutual Information between antigen quality and latent space projections (see section 2.7 for details), we found that our Antigen Encoding model derived from supervised training [...] outperforms the PCA and A-E unsupervised training methods (Fig. 2.6G). Moreover, one-layer

perceptron models trained on each latent space resulted in higher accuracy when using our classifier-derived latent space, compared to latent spaces found through unsupervised methods (Fig. 2.6H). This demonstrated that, while both supervised and unsupervised approaches can capture the general structure of data in a similar way, our supervised training allows for a better classification and ranking of antigens.” (*Antigen encoding*, [1], SI)

2.4.3 Impact of excluding IL-4 and IL-10 on the latent space

“We excluded the measurements for the cytokines IL-4 and IL-10 on the basis that their noisy time series mainly reflected the background fluorescence in the cytokine bead assay, rather than real concentrations.

We quantified the level of noise using the signal-to-noise ratio (SNR), defined here as

$$\text{SNR} = \frac{\langle c \rangle}{\sqrt{\langle (y(t) - c(t))^2 \rangle}}, \quad (2.10)$$

where $\langle c \rangle$ is the mean cytokine signal, and $\sqrt{\langle (y(t_i) - c(t_i))^2 \rangle}$ is the root-mean-squared deviation of data points $y(t_i)$ from the smoothing spline fits $c(t)$. Fig. 2.7A shows that IL-4 and IL-10 have SNRs at least twice smaller than other cytokines. Note that we computed the SNR for $y(t)$ and $c(t)$ either with (left panel) or without (right panel) the logarithmic transform described in section 2.2.2; this transformation clearly attenuated experimental noise and lead to higher SNR. We further ascertained the low SNR for IL-4 and IL-10 by dividing the range of each cytokine in quintiles and computing the SNR within each interval. Even for the top 20 % recorded concentrations, their SNR is at least twice as small as for any other cytokine (figure 2.7B).

Moreover, we checked that the latent space found when training a new classifier with IL-4 and IL-10 included in the inputs had the same structure as when training with five cytokines only (figure 2.7D). The training score was slightly higher with seven cytokines, but this could be attributed to overfitting. Indeed, the cross-validation and test scores reduced to those of the classifier using five cytokines only (figure 2.7C).” (*Antigen encoding*, [1], SI)

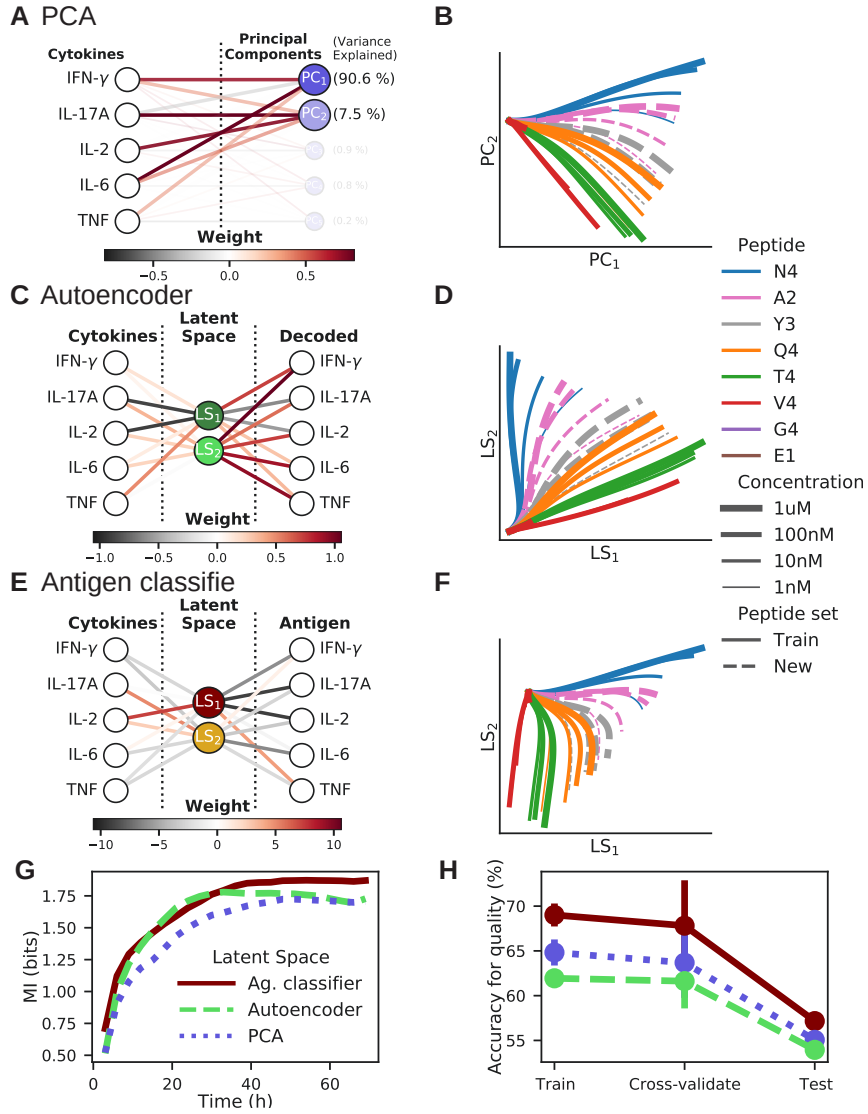


Figure 2.6: “Comparison of supervised vs. unsupervised learning procedures We compared three dimensional reduction procedures to analyze the data used in Fig. 2 of the main text, *i.e.* the time integral of the log of cytokines. **(A-B)** Principal Component Analysis (PCA) of the 5d cytokine dynamics resulted in trajectories displayed along the first two Principal Components **(C-D)** Trained auto-encoder (A-E) with a two-node bottleneck resulted in trajectories in 2d latent space **(E-F)** Neural network and latent space as used in the main text **(G)** Mutual Information as a function of time for the three learning modalities, computed on the training set **(H)** Performance of a one-layer perceptron trained for antigen classification on the latent spaces derived from the three learning modalities (same color code as in **(G)**). Note how our supervised learning procedure performed best for training, cross-validating and testing, compared to the PCA and A-E unsupervised methods.” (*Antigen encoding*, [1], SI)

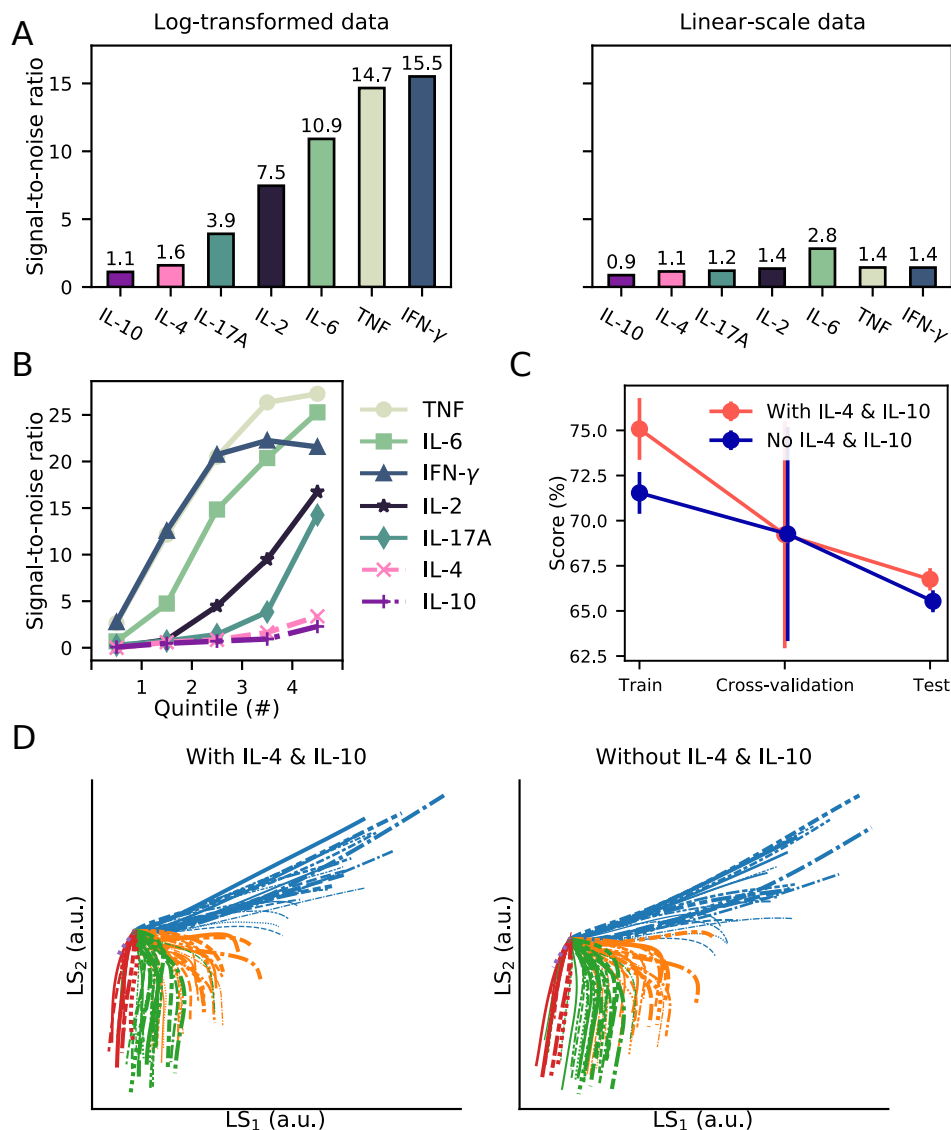


Figure 2.7: “Negligible impact of IL-4 and IL-10 for the classifier inputs. (A) Signal-to-noise ratio (SNR, eq. 2.10) of each cytokine computed across all times and conditions in the training data. The left panel shows how the log-transformation of cytokine concentrations more than doubles the SNR compared to the SNR derived from data in linear scale (right panel), for all cytokines except IL-4 and IL-10. (B) SNR analysis by quintiles. The ranges of each cytokine concentration (in log scale) were divided by quintiles and SNR were computed within each interval between quintiles. Even when considering only the top 20 % of recorded concentrations, the SNR of the 5 cytokines under consideration are more than 4 times larger than the SNR for IL-4 and IL-10, justifying why including these 2 cytokines made a negligible contribution to antigen encoding. (C) Training, cross-validation (five-fold) and testing scores for antigen quality classifiers (same structure with one 2d hidden layer). Test scores are evaluated on data from separate experiments; error bars on test scores come from 100 bootstrap replicates. (D) The latent spaces found by the classifier with or without IL-4 and IL-10 are essentially identical.” (*Antigen encoding*, [1], SI)

2.4.4 Impact of logarithmic transform, time integration, and smoothing

“Our data pre-processing consisted in a combination of logarithm, smoothing and integral of trajectories. We compared those modalities and trained a neural network on each to see how they influence the final results (Fig. 2.8). We found that the logarithmic transformation and, to a smaller extent, time integrals were crucial for the emerging structure of latent space. Smoothing further improved the accuracy in training and testing, but to a lower extent.” (*Antigen encoding*, [1], SI)

2.5 Cytokine time course reconstruction

“Our goal in this section was to map the system’s trajectories in the 2d latent space back to the actual cytokine concentration. In other words, we aimed at reconstructing the entire dataset based only on the trajectories in the latent space, to demonstrate the encompassing strength of our data compression.” (*Antigen encoding*, [1], SI)

2.5.1 Linear reconstruction

“We first explored whether standard linear reconstruction would be sufficient. Figure 2.9 compares the original and reconstructed cytokine time courses for a test dataset (one experiment, 100 k initial T cells) computed with a linear least-square regression. The training and test data are different replicates coming from the same experiment.

With this method, IL-2 is accurately reconstructed because node 1 of the neural network (LS_1) has a heavy contribution from IL-2 and thus closely matches its shape. Other cytokines, however, suffer from important reconstruction artifacts: IFN- γ and IL-6 in particular have unnatural peaks at intermediate times and then drop below their true steady-state value at later times. This is akin to the residuals expected when fitting a curved surface with a plane, as in the example of figure 2.5D.” (*Antigen encoding*, [1], SI)

2.5.2 Nonlinear reconstruction

“We thus considered increasingly complex reconstruction methods to remove the artifacts visible in figure 2.9 and better capture the curvature of the cytokine manifold.

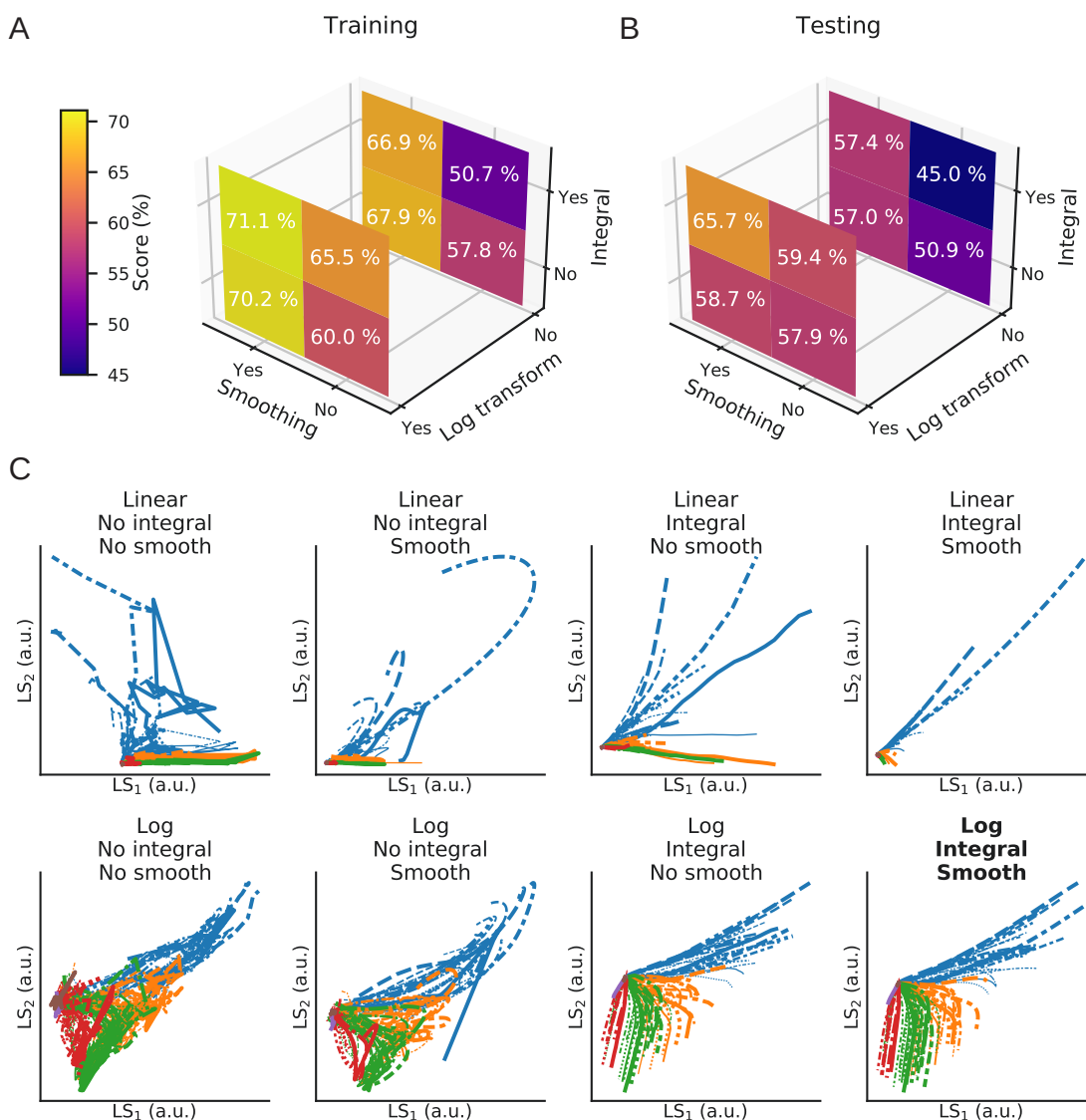


Figure 2.8: “Effect of each preprocessing step on the classifier accuracy We trained the same classifier architecture (one 2d hidden layer) on the same datasets but processed with each possible combination of log-transformation, time-integration, and spline smoothing of the data. **(A)** The training score of classifiers trained on each possible combination of preprocessing modalities, indicated by the three axes, is given by the color bar and the numerical labels in each square. **(B)** The test scores of the classifiers on previously unseen datasets, for each preprocessing modality (same color scale as A). Testing scores are always lower because of experiment-to-experiment variability. **(C)** The latent space found by the classifiers trained on each preprocessing modality (irrelevant rotations and reflections were automatically applied when necessary to allow comparison of all latent spaces in similar orientations). The last latent space (Log-Integral-Smooth) is the one used in the main text.” (*Antigen encoding*, [1], SI)

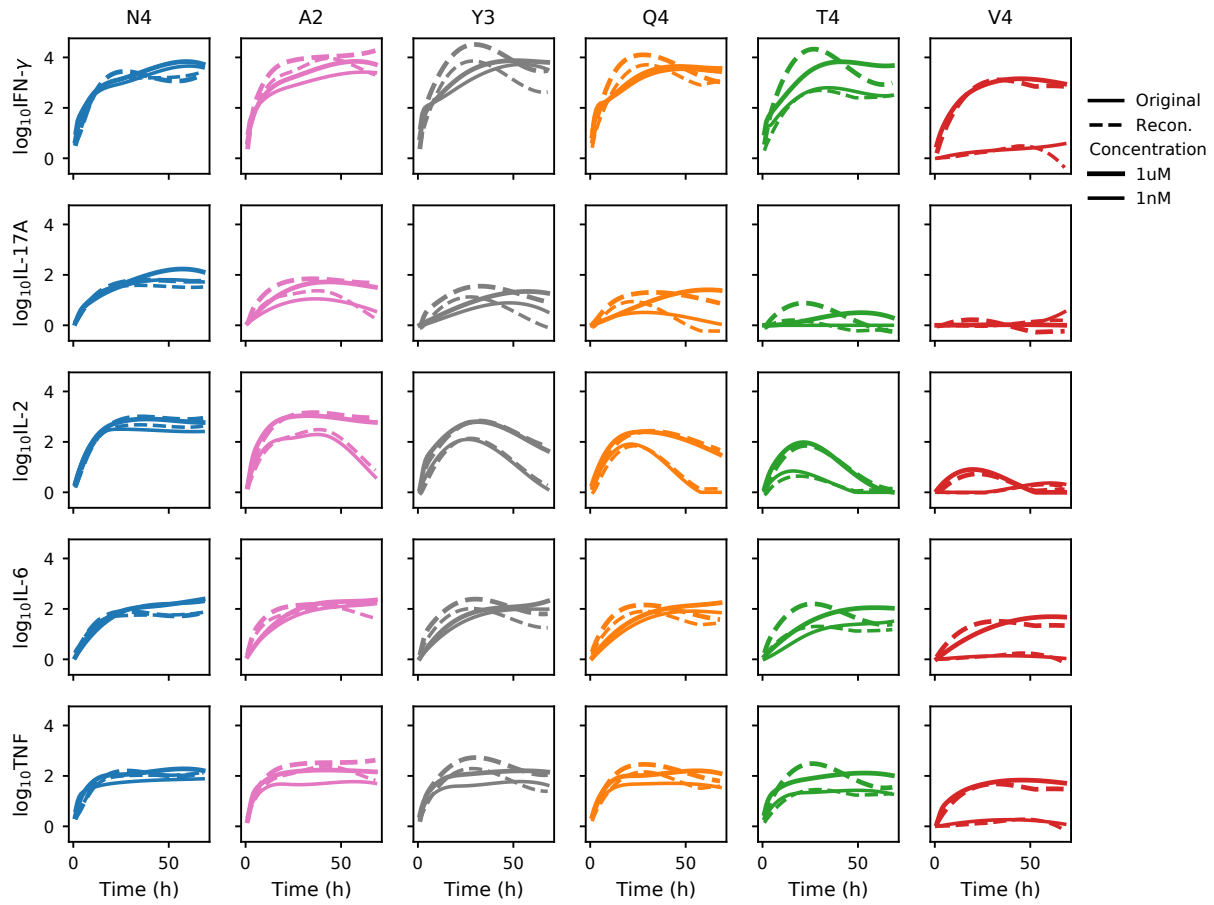


Figure 2.9: “Cytokine reconstruction results with linear least-squares regression.” (*Antigen encoding*, [1], SI)

A significant improvement resulted from including, as inputs to the reconstruction, quadratic terms in $ls_1(t)$ and $ls_2(t)$ and the hyperbolic tangent ($\tanh$) of the latent space time integrals, $\tanh(LS_1(t)/\bar{N}_1)$ and $\tanh(LS_2(t)/\bar{N}_2)$, where $\bar{N}_i$ is a normalization constant. Mathematically speaking, the $\log c_i(t)$ are reconstructed as

$$\begin{aligned} \log c_i = & Q_{i1}ls_1 + Q_{i2}ls_2 + Q_{i3}ls_1^2 + Q_{i4}ls_2^2 + Q_{i5}ls_1ls_2 \\ & + Q_{i6} \tanh(LS_1(t)/\bar{N}_1) + Q_{i7} \tanh(LS_2(t)/\bar{N}_2) \end{aligned} \quad (2.11)$$

The $\tanh$ function prevents excessive increase of the integrals' contribution at late times. For simplicity, the normalization constants $\bar{N}_i$ were chosen as the average value of $LS_i(t)$ across all times and condition in the reconstruction training data.

The coefficients Q_{ij} were fitted using least-square regression on a subset of the data, and tested on other datasets. With this accurate method, cytokine concentrations were almost perfectly reconstructed, both for the training and test data ($R^2 > 0.94$), the latter shown in figure 2.10. Fig. 2.11 supports this claim by showing the distribution of reconstruction residuals (across peptides, concentrations, and replicates) as a function of time for both methods. The artifacts produced by the linear method correspond to the large, time-dependent deviations of the mean residuals from zero. In contrast, the accurate method has residual distributions centered around zero and with noticeably smaller standard deviations and extrema, especially for IL-6, TNF, and IL-17A. This demonstrated that our projection of the cytokine dynamics onto the 2d latent space (antigen encoding) could be used to reconstruct the entire cytokine dynamics (antigen decoding)," (*Antigen encoding*, [1], SI)

and, hence, that it does not cause information loss about any cytokine in particular. Figure 2.4D-F illustrates our non-linear reconstruction method on sample latent space trajectories from the training data (D-E), and the reconstruction versus data for all training and test datasets generated with OT-1 antigens in this project (F).

2.6 Dynamical latent space models and parameter correlations

After employing a classifier neural network to discover the latent space of cytokine dynamics,

"we aimed at going beyond antigen quality prediction to build a parsimonious generative model of the cytokine trajectories from the latent space downstream of TCR

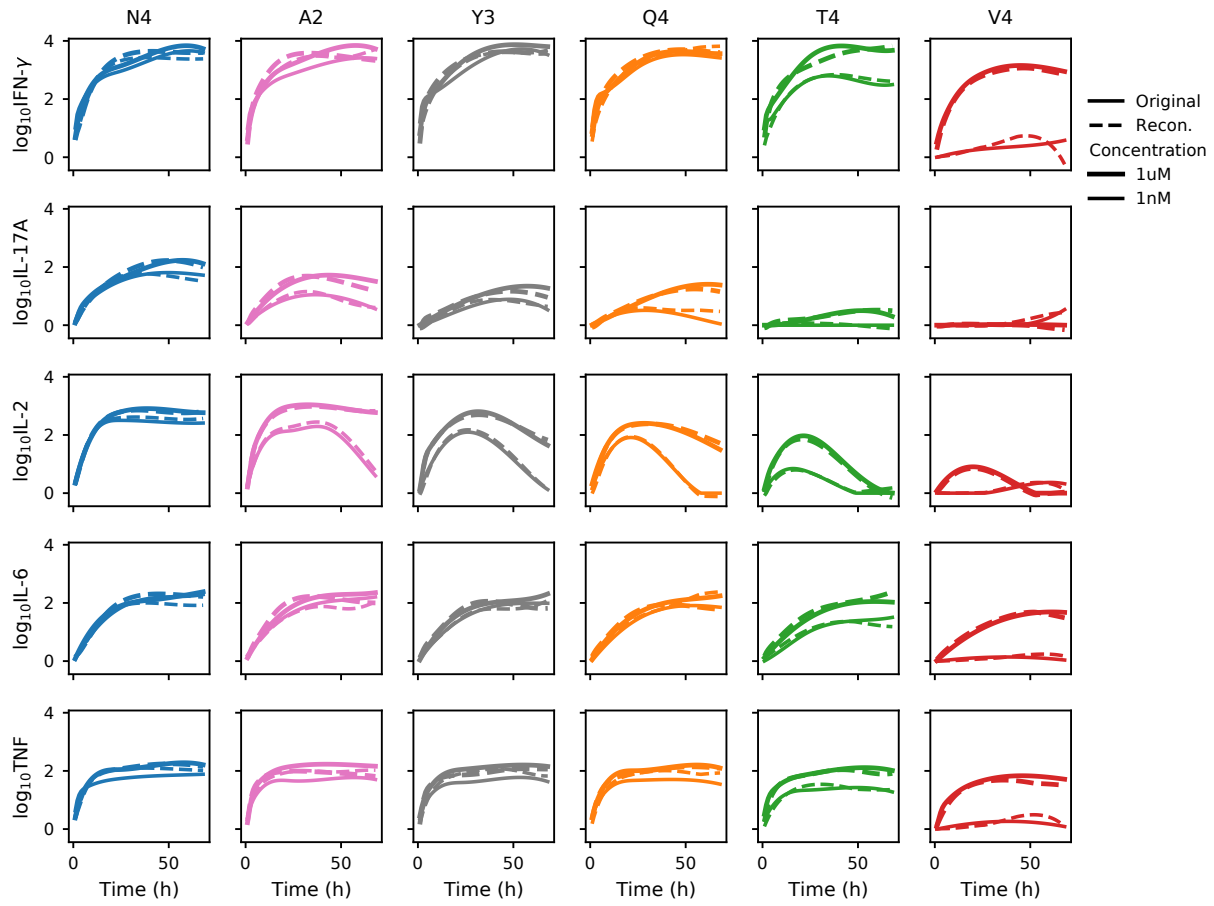


Figure 2.10: “Non-linear cytokine concentration reconstruction results using least-squares regression on quadratic and $\tanh$ -transformed time integral terms from the latent space.” (*Antigen encoding*, [1], SI)

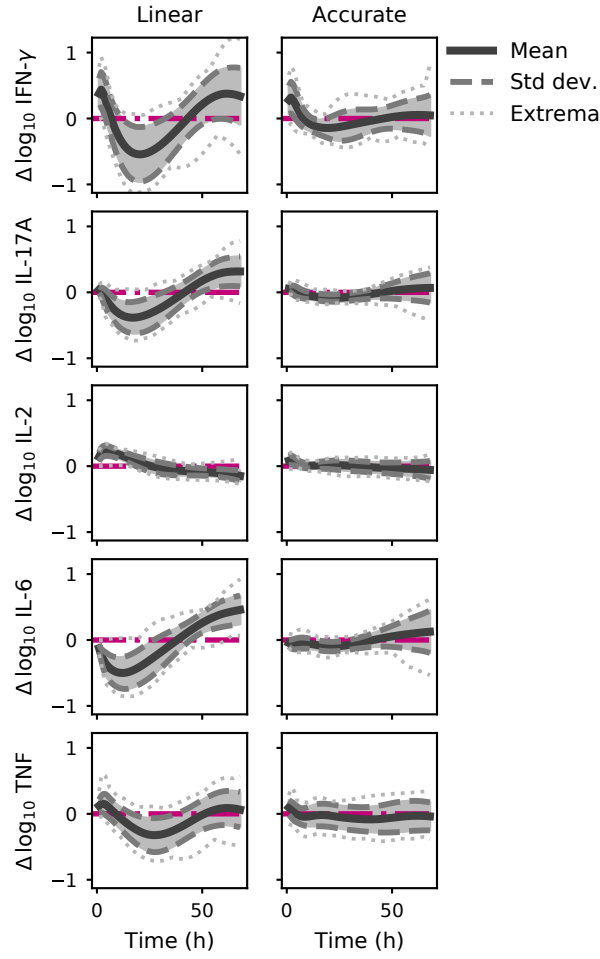


Figure 2.11: “Residuals of the cytokine reconstruction. Distribution of residuals of cytokine concentrations reconstructed from latent space (for a test dataset) as a function of time. The “accurate” non-linear method, where we include $\tanh LS_1$, $\tanh LS_2$ and quadratic terms in ls_i , gives overall lower residuals with less trend over time. The improvement is especially significant for IL-17A, TNF, and IL-6.” (*Antigen encoding*, [1], SI)

engagement [192]. Modelling of trajectories with simple equations allowed us to derive a quantitative understanding of the impact of antigenicity on T cell activation (when individual cytokines were too tangled to allow classification – see Fig. 2.5A) and, more practically, to estimate the number of antigen classes encoded by the cytokine dynamics.

Two dynamical models were derived. The first model is called the ‘velocity’ model : this is the simplest model, with a minimum number of parameters, easy to interpret and fit. Its dynamics contains two discrete phases (see below). To get a more accurate generative model of cytokine trajectories and quantify more precisely information content, we also developed a refined model, called the ‘Force model with matching’, with more parameters and matching terms to ensure a smooth transitions between the two dynamical phases. Having models of diverse complexity levels allows for a tuning between easiness of interpretation (velocity model), accuracy of reconstruction (Force model with matching) and parameter estimation (which can be challenging in models with bigger number of parameters due to parameter sloppiness [193]).” (*Antigen encoding*, [1], SI)

2.6.1 Constant velocity model

The simpler model, called “constant velocity”, was developed jointly by Thomas Rademaker and I, as well as the model parameter fitting procedure in the next subsection. We used it for simple analyses of the data, but not for the accurate parameterization of dynamics or for the channel capacity calculations below. Recall how we

“defined $LS_1(t)$ and $LS_2(t)$ as the projections of time integrals (from 0 to t) of the log-transformed cytokine concentrations, on node 1 and node 2 of the neural network. In equations,

$$LS_i = \sum_k P_{ik} C_k \quad \text{where } C_k = \int_0^t du \log_{10} c_k(u) \quad (2.12)$$

where $c_k(t)$ is the concentration of cytokine k normalized by the lower limit of detection for that cytokine (as explained in section 2.2.2) and the P_{ik} are the neural network weights from the input layer to the intermediate layer. We defined in an analogous way $ls_1(t)$ and $ls_2(t)$, the projections of the time courses of log-transformed

cytokine concentrations on node 1 and node 2. Note that $l_{S_i}(t) = \frac{dLS_i}{dt}$ ($i = 1, 2$) because the projections defined in (2.12) are linear.

The trajectories of $LS_1(t)$ and $LS_2(t)$ resemble ballistic motions. Therefore, we derived a model in analogy with the motion of a projectile or rocket with two phases: first, a constant velocity phase, and second, a free fall with air resistance. The first phase of a trajectory is thus described by $\mathbf{r}(t) = \mathbf{v}_0 t$, where $\mathbf{r}(t) = (LS_1(t), LS_2(t))$ and $\mathbf{v}_0$ is a vector with magnitude v_0 and angle θ , measured counterclockwise relative to the direction of the terminal velocity (see Fig. 2.12A). The second phase's description starts with a second-order linear differential equation for $\mathbf{r}(t)$,

$$\frac{d^2 \mathbf{r}}{dt^2} = -k \left(\frac{d\mathbf{r}}{dt} - \mathbf{v}_t \right) \quad (2.13)$$

where k is a “drag” constant (units of inverse time) and $\mathbf{v}_t = (v_{t1}, v_{t2})$ is the terminal velocity of the trajectories. Defining t_0 as the time at which the constant velocity phase ends and integrating eq. (2.13) twice with the initial conditions to match the end of the first phase at t_0 , we get the following parametric equations describing $\mathbf{r}(t) = (LS_1(t), LS_2(t))$:

$$\mathbf{r}(t) = \begin{cases} \mathbf{v}_0 t & t \leq t_0 \\ \frac{\mathbf{v}_0 - \mathbf{v}_t}{k} (1 - e^{-k(t-t_0)}) + \mathbf{v}_t(t - t_0) + \mathbf{v}_0 t_0 & t > t_0. \end{cases} \quad (2.14)$$

In total, six parameters ($v_0, \theta, t_0, \mathbf{v}_t, k$) are used for this model. Each can be given a biological interpretation. The initial (TCR-mediated) T cell activation strength is captured by v_0 . The initial orientation θ gives the ratio, at early times, of cytokines associated to the innate (node 2: IL-6 and TNF) and adaptive (node 1: IL-2 and IL-17A) responses of the immune system [...]. The time t_0 measures how long the initial cytokine production phase lasts; it may correspond to the delay before T cells switch between IL-2 production and IL-2 consumption [105]. $\mathbf{v}_t$ describes the final slope of LS_2 versus LS_1 in latent space, which relates to the ratio of innate and adaptive cytokines. This slope is remarkably conserved across all peptide conditions in an experimental repeat (see Fig. 2.13B for instance). Finally, the rate k in Eq. 2.14 is introduced to capture the decay of cytokine concentrations over long time scales. From previous models, we know that this corresponds to the rate of consumption of cytokines [105].” (*Antigen encoding*, [1], SI)

Figure 2.12A illustrates the four main parameters of this model on sample latent space trajectories.

2.6.2 Parameter fitting procedure

“Without complex curve fitting, we could determine two of the parameters for all trajectories in a given experimental repeat. First, as can be seen from the final orientation of all $(LS_1(t), LS_2(t))$ trajectories in Fig. 2.13B, the final slope $m = v_{t2}/v_{t1}$ was the same for peptides Q4 and weaker at all concentrations. We simply took the median of all final slopes found by linear regression in this LS_2 vs LS_1 space. This eliminated, *e.g.*, the component v_{t1} , derivable as $v_{t1} = v_{t2}/m$. Second, we typically fixed the time scale $k = \frac{1}{20} \text{ h}^{-1}$, because this gave reasonable fits throughout. This left four dimensionless parameters to determine: v_0/k , kt_0 , θ , and v_{t2}/k . To fit them, we numerically solved a non linear least squares problem, optimizing the cost function :

$$\begin{aligned} C &= \sum_i ||\mathbf{r}(t_i) - \mathbf{N}(t_i)||^2 + \sum_j |p_j - p_j^0| \\ &= \sum_i [(r_1(t_i) - LS_1(t_i))^2 + (r_2(t_i) - LS_2(t_i))^2] + \sum_j |p_j - p_j^0| \end{aligned} \quad (2.15)$$

where the term $\sum_j |p_j - p_j^0|$ introduces a L1 regularization. p_j^0 are “default” values for parameters corresponding to trajectories for null peptides (no constant velocity phase, no terminal velocity, minimal angle): $v_0^0 = 0$, $t_0^0 = 0$, $\theta^0 = -2\pi/3$, $v_{t2}^0 = 0$. We also set reasonable bounds on parameter values:

$$v_0 \in [0, 5k], \quad t_0 \in [0, t_{exp} + 20] \text{ hrs}, \quad \theta \in [-2\pi/3, \pi/3], \quad v_{t2} \in [0, 5k]. \quad (2.16)$$

Optimization was performed with a customized version of the `curve_fit` method in *SciPy* [187]. Minimizing this cost function for each cytokine time course (caused by one peptide at one concentration and initial T cell number) in 11 independent OT-1 datasets (one including 4 technical replicates) yielded the parameter distribution shown in Fig. 2.13C.” (*Antigen encoding*, [1], SI)

2.6.3 Parameter fits show that early kinetics control the response

“Fig. 2.13 shows how well the model fits the time course of $LS_1(t)$ and $LS_2(t)$ and the distribution of parameters v_0 , t_0 , θ and v_t for different peptides and concentrations.

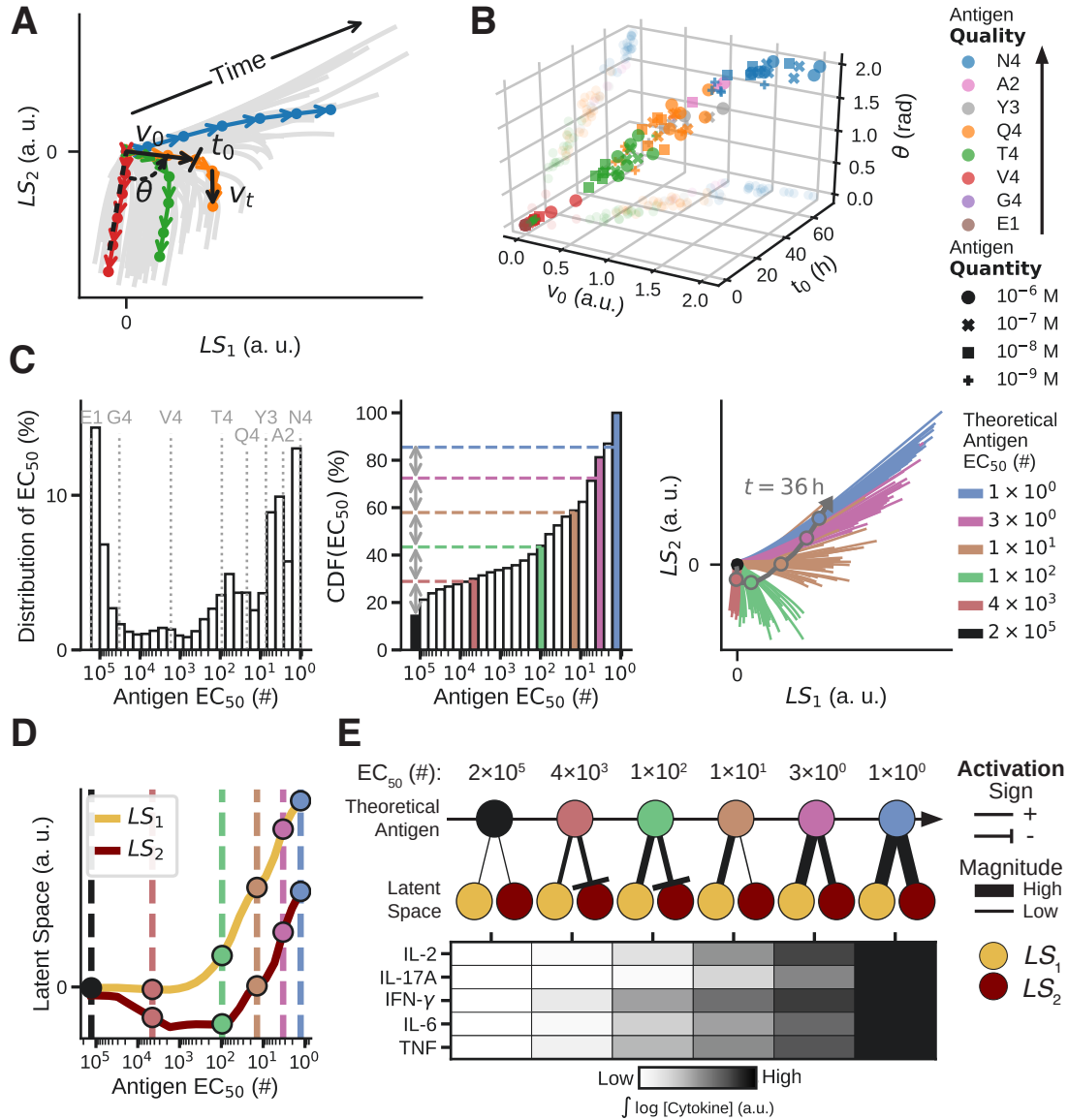


Figure 2.12: “Quantifying the classes of T cell activation using antigen encoding in latent space. (A) Cytokine trajectories in the latent space can be fitted with a 4-parameter ballistic model. (B) Although one parameter (v_t) is constant, the three others (v_0, θ, t_0) are strongly correlated. (C) Distribution of EC_{50} (left) that maximizes the mutual information extracted from model parameters (channel capacity ≈ 2.6 bits, section 2.7). By evenly sampling the cumulative distribution function (center), we find $2^{2.6} \approx 6$ classes of antigen with non-overlapping latent space trajectories (right). The arrow connects the 36-hr coordinates of trajectories of increasing antigen EC_{50} . (D) Latent space coordinates (LS_1, LS_2) at 36 hr, as a function of antigen quality as defined by EC_{50} . (E) Sketch of the biochemical mechanisms governing antigen discrimination, as suggested from the latent space coordinates in (D): each class of antigen activates positive and negative signals differentially and yields varied patterns of cytokine secretion computed from the decoder (Fig. 2.4E) and trajectories in (C).” (Antigen encoding, [1])

Constant velocity

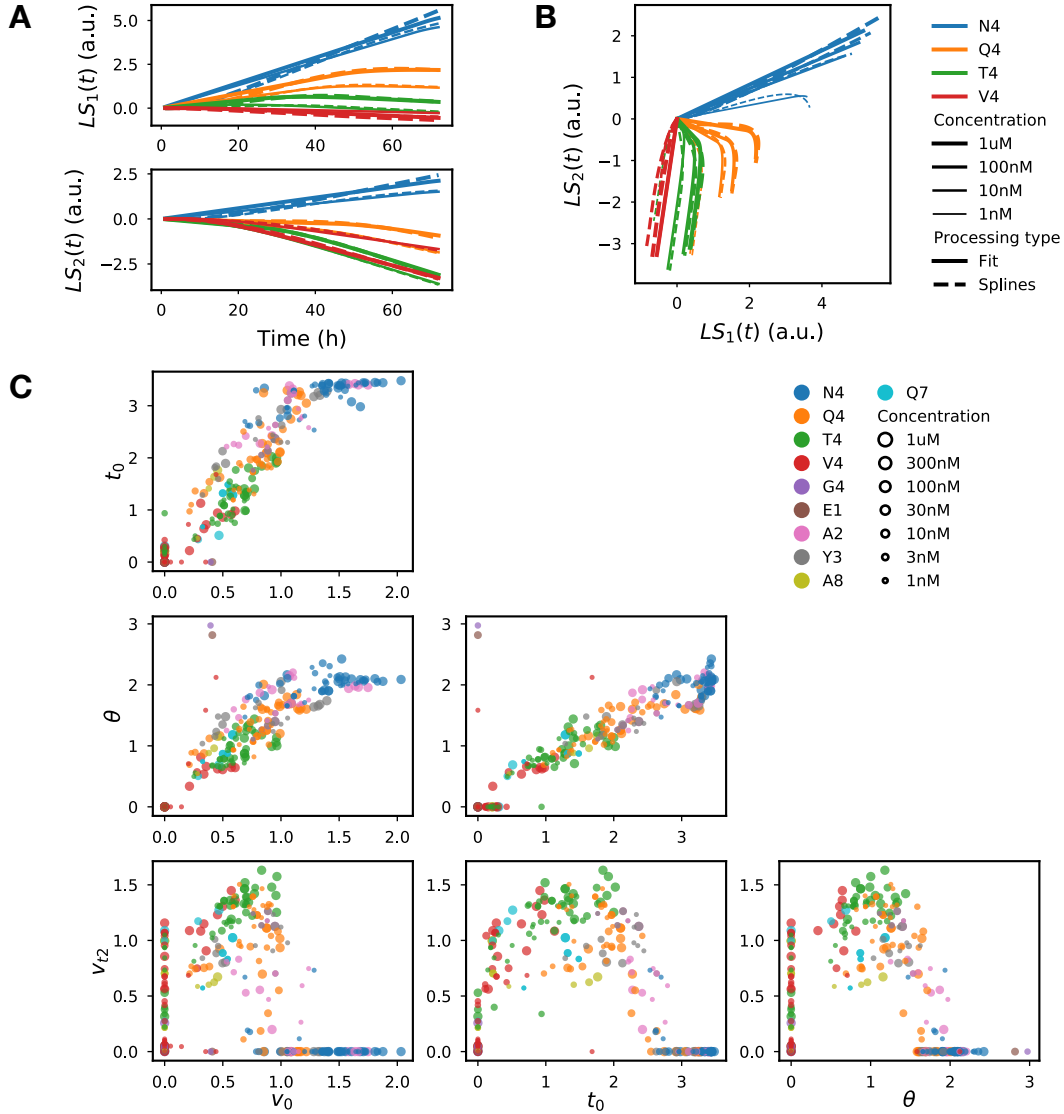


Figure 2.13: “Fits of the constant velocity model. (A) Fitted time courses of LS_1 and LS_2 from the constant velocity model (eq. 2.14) compared to data (spline interpolation) on a representative experimental repeat (OT-1 T cells, 4 peptides $\times$ 2 concentrations shown). (B) Same trajectories as (A), plotted against each other (4 concentrations shown per peptide). (C) Constant velocity model parameter values, fitted on each time series from 11 independent datasets (one including 4 technical replicates), for many peptides at different concentrations, with 100k initial T cells. ” (*Antigen encoding*, [1], SI) Two new antigens, A8 (SIINFEKA) and Q7 (SIINFEQL) are introduced in this figure; they were present in only 1-2 datasets each.

Strikingly, the first three parameters are strongly correlated, and they are controlled by antigen quality, with only minor variability caused by antigen quantity (Figures 2.12B and 2.13C). The initial T cell activation strength (v_0), which depends on the antigen, also determines the initial orientation of the trajectories, θ , and how long it takes before cytokine consumption kicks in, t_0 . This correlation defines an antigenicity axis in parameter space, with antigen quality controlling the variation along this axis (see also figs. 2.14-2.15)." (*Antigen encoding*, [1], SI)

We used the Kendall tau distance to estimate how well model parameter values are ordered on the antigenicity axis. This distance, first proposed by Kendall in 1938 [194], quantifies the number of inversions needed to reorder a list; it is explained in figure 2.14. Applying it to model parameters to check how well they reflected the ordering of antigens according to their quality, we found that parameter v_0 ranks antigens more accurately and with less variability than any single cytokine time point (figure 2.15); other parameters, correlated to v_0 , have similar accuracies. This analysis was joint work between Sooraj Achar (who prepared the figure) and I (who wrote the analysis code).

2.6.4 Refined model

"The constant velocity model was enough to perform the antigen classification and to understand most quantitative aspects of cytokine trajectories. Yet, it introduced discontinuities in derivatives at the transition between different phases at $t = 0$ and $t = t_0$. We improved this model by introducing two extra parameters to smooth out the transition between phases." (*Antigen encoding*, [1], SI)

While the constant velocity model was joint work with Thomas Rademaker, developing, integrating, and fitting the refined model on data was my own work.

Constant force model with matching

"The constant velocity model captured well the shape of trajectories in LS_1 , LS_2 space, but there was a noticeable time scale of activation of the velocity (first) phase. For instance, $LS_1(t)$ trajectories in the model for peptide N4 in Fig. 2.13A are straight lines, while the data curves have a noticeable curvature. An immediate improvement came from considering a constant acceleration a_0 in the first phase, rather than a constant velocity. Moreover, we fixed the remaining discontinuity in

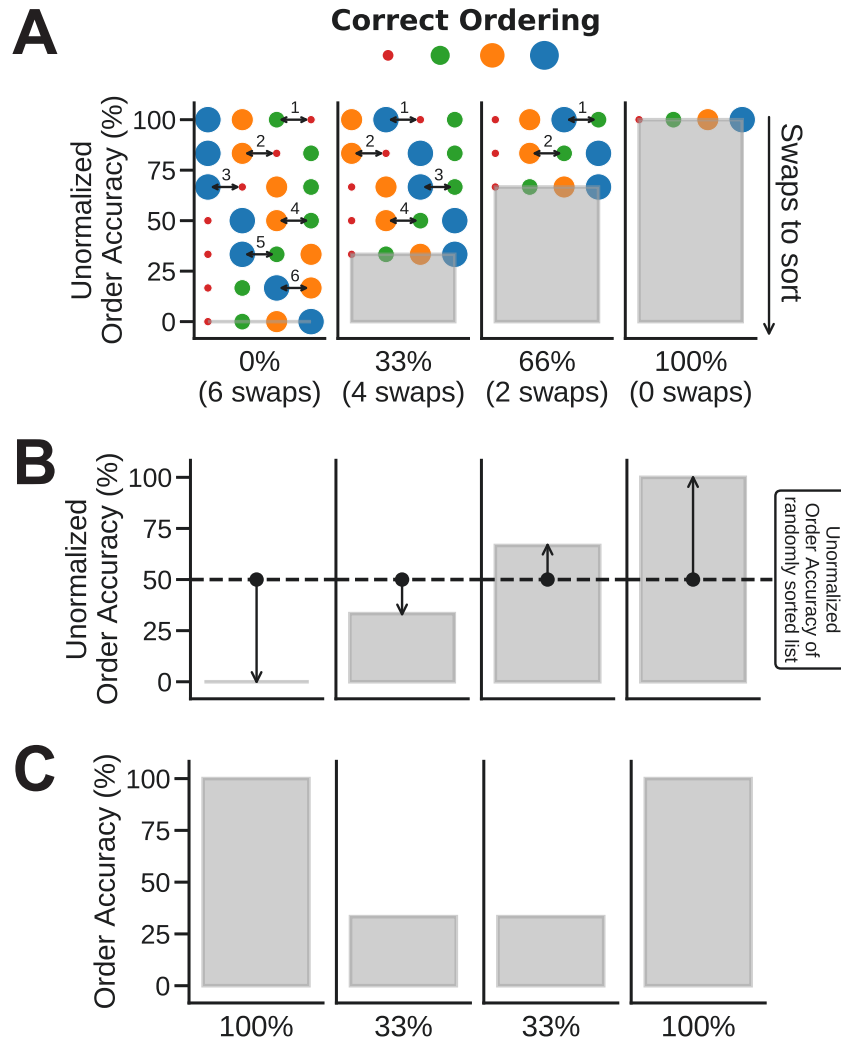


Figure 2.14: “Graphical explanation of Kendall tau distance accuracy metric. We used the Kendall tau distance metric to estimate how well a parameter or marker preserves the order amongst antigens according to quality. **(A)** This metric counts the number of swaps between neighboring elements one needs to perform to reorder a ranked list. **(B)** Our accuracy metric is initially scaled between 0 and 100%, the former for a completely reversed list – $n(n-1)/2$ swaps needed for n elements – and the latter for a well-ordered list. **(C)** The accuracy metric is then renormalized so that a randomly sorted list, which would have an **unnormalized** order accuracy of 50 % ($n(n-1)/4$ swaps), now has a **normalized** order accuracy of 0 %. We now look at the absolute, not signed, order accuracy, to allow both a reverse-ordered list (requiring $n(n-1)/2$ swaps) and a well-ordered list (requiring 0 swaps) to both formally correspond to an accuracy of 100 %. This is important because the directionality of our parameters is mostly arbitrary, so we want there to be no difference in order accuracy between a metric perfectly correlating or perfectly anti-correlating with antigenicity.” (*Antigen encoding*, [1], SI)

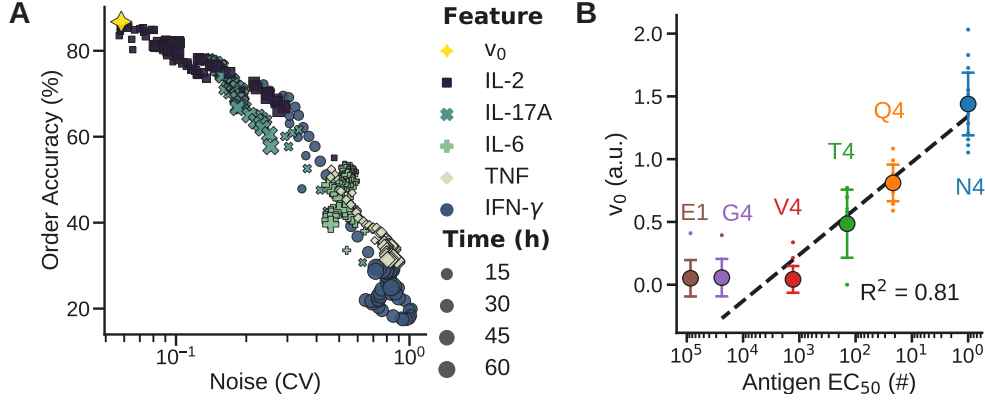


Figure 2.15: Parameter v_0 ranks antigens according to quality. (A) A single model parameter (v_0) yields a higher order accuracy (see Fig. 2.14) for antigen quality independently from quantity, more consistently (*i.e.*, with smaller coefficient of variation [CV]) than any single time point cytokine measurement. Variability of order accuracy (CV, horizontal axis) was estimated by computing order accuracy in 11 independent OT-1 datasets, one comprising 4 technical replicates. (B) The parameter v_0 , derived from the ballistic model of the cytokine dynamics in the latent space, correlates with the antigenicity of peptides as defined by their $\log_{10}(\text{EC}_{50})$ measured in classical dose-response assays [174, 176].” (*Antigen encoding*, [1], SI)

$\frac{d^2\mathbf{r}}{dt^2}$ at t_0 by analogy with boundary layer theory [94]. We matched the velocity phase and the asymptotic phase by introducing a sigmoidal term $(1 + e^{\beta(t-t_0)})^{-1}$ between the two phases in ls_1 and ls_2 :

$$\begin{aligned} ls_1(t) &= (1 - e^{-\alpha t}) \left(\frac{a_0 \cos \theta + v_{t1}}{e^{\beta(t-t_0)} + 1} - v_{t1} \right) \\ ls_2(t) &= (1 - e^{-\alpha t}) \left(\frac{(a_0 \sin \theta + v_{t2})(1 - e^{-\alpha t})}{e^{\beta(t-t_0)} + 1} - v_{t2} \right) \end{aligned} \quad (2.17)$$

We squared the bounded exponential term in the first part of the equation for ls_2 , in order to better capture its typical time courses. The expressions in equation 2.17 are analytically integrated to obtain expressions for $LS_1(t)$ and $LS_2(t)$, which can be fitted on the data as explained in section 2.6.2. Define $\tau = \alpha t$, $\tau_0 = \alpha t_0$, and

$\gamma = \beta/\alpha$. The result is

$$\begin{aligned} LS_1(t) &= \frac{a_0 \cos \theta + v_{t1}}{\alpha} \left(I(\tau, \tau_0, \gamma) - \frac{1}{\gamma} \ln(e^{-\gamma\tau} + e^{-\gamma\tau_0}) \right) - \frac{v_{t1}}{\alpha} (\tau + e^{-\tau}) + K_1 \\ LS_2(t) &= \frac{a_0 \sin \theta + v_{t2}}{\alpha} \left(2I(\tau, \tau_0, \gamma) - \frac{1}{2} I(2\tau, 2\tau_0, \frac{\gamma}{2}) - \frac{1}{\gamma} \ln(e^{-\gamma\tau} + e^{-\gamma\tau_0}) \right) \\ &\quad - \frac{v_{t2}}{\alpha} (\tau + e^{-\tau}) + K_2 \end{aligned} \quad (2.18)$$

where the K_i are chosen to get $LS_i(0) = 0$. The complicated part is the integral $I(\tau, \tau_0, \gamma)$, which is given by

$$\begin{aligned} I(\tau, \tau_0, \gamma) &= \int d\tau \frac{-e^{-\tau}}{e^{\gamma(\tau-\tau_0)} + 1} \\ &= \begin{cases} e^{-\tau} {}_2F_1\left(1, \frac{-1}{\gamma}; 1 - \frac{1}{\gamma}; -e^{\gamma(\tau-\tau_0)}\right) & \frac{1}{\gamma} \notin \mathbb{N}^+ \\ e^{-\tau} + e^{-\tau} \sum_{j=1}^{n-1} \frac{n}{n-j} (-1)^j e^{(\tau-\tau_0)\frac{j}{n}} \\ \quad + (-1)^n n e^{-\tau_0} \ln(e^{-\tau/n} + e^{-\tau_0/n}) & \frac{1}{\gamma} \in \mathbb{N}^+ \end{cases} \end{aligned} \quad (2.19)$$

where ${}_2F_1$ is Gauss' hypergeometric function. Note that, in this version of the model, θ was defined relative to the horizontal axis, and that we fitted $\tau_0 = \alpha t_0$ instead of t_0 directly, because the former is a dimensionless parameter." (*Antigen encoding*, [1], SI)

Quality of fits and parameter space

"The force model with matching (equation 2.18) is fitted to data for $LS_1(t)$ and $LS_2(t)$, as outlined in section 2.6.2. Fig. 2.16A and 2.16B show how well the model captures trajectories $LS_1(t)$ and $LS_2(t)$: the difference between data and model curves is almost unnoticeable. Fig. 2.17D also shows how this accuracy of fits extends to latent space concentration trajectories, $ls_1(t)$ and $ls_2(t)$. The squared residuals plotted in Fig. 2.17A and B are indeed roughly 10 times lower than with the constant velocity model, especially if parameters α and β are fitted separately (as opposed to fixing $\alpha = 1/20 \text{ h}^{-1}$)." (*Antigen encoding*, [1], SI)

We provide, in figures B.4 and B.5, further examples of model fits of latent space integrals and concentrations, respectively, for six randomly selected datasets (representative of all 14 OT-1 T

cell datasets). We provide a dataset-per-dataset breakdown of fit residuals in figure 2.18, clearly showing how the force model with matching captures ls_1, ls_2 dynamics better than the constant velocity model.

2.6.5 Model fits on high-dimensional cytokine dynamics

Combining the latent space model with the nonlinear decoder (eq. 2.11), we effectively obtained a dynamical model that describes the time series of all cytokine concentrations. To fit (smoothed, log-transformed) data with these expressions, we first projected the data to the latent space, where we fitted the ballistic parameter values; then, we computed the $(ls_1(t), ls_2(t))$ model curve corresponding to these parameter values, and we reconstructed the five cytokines using the nonlinear decoder (which was already trained on separate data). This gave model curves of cytokines that should closely match the original data, despite the dimensional bottleneck through which the model goes. We show an example of such fits in figure 2.19, with the goodness-of-fit quantified by a multivariate χ^2 test [195, 196] and found to be acceptable (large p -value) for all cytokines.

2.6.6 Generating synthetic cytokine time courses

Moreover, combining our latent space model with the reconstruction procedure of section 2.5 allowed us

“to generate synthetic time courses for the five cytokines we considered. The full process for synthetic data generation is illustrated in figure 2.20 and detailed below :

1. We fitted multivariate kernel density estimates (KDEs, from the *scikit-learn* module in Python [190]) in the parameter space of the force model with matching (with free α), for each peptide. More specifically, the KDEs were fitted on parameter values from similar OT-1 datasets with 100k initial T cells. Model parameters were standardized prior to fitting the KDEs and scaled back after sampling. This scaling ensured that a KDE bandwidth appropriate for all parameters could be chosen, even for those with typically smaller values, such as v_{t1} . Those KDEs (transformed back to the original parameter magnitudes) are shown on figure 2.20A. Multimodality of some of the KDEs come from different peptide concentrations (which were aggregated together) and from experimental variability.

Force model with matching, free α

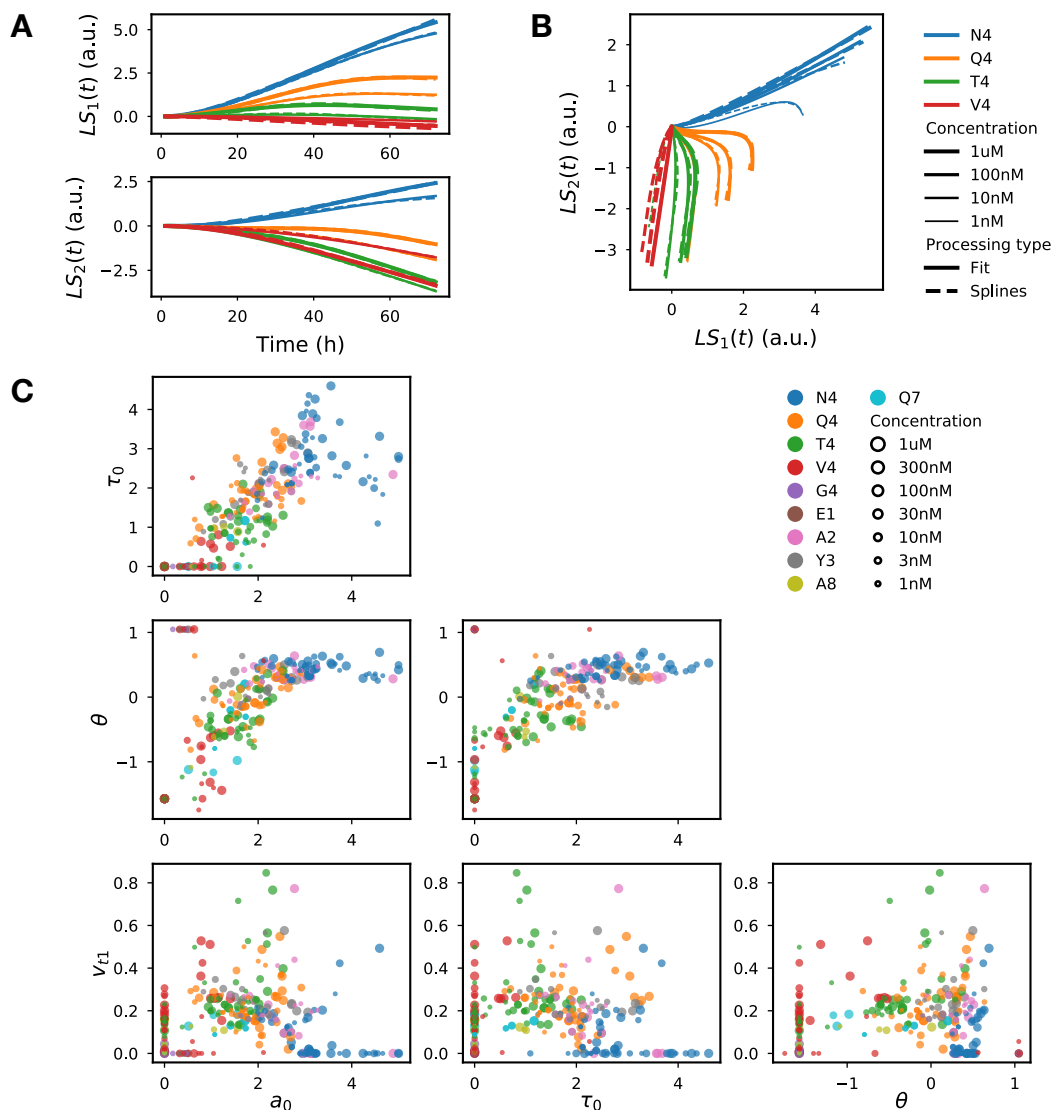


Figure 2.16: “Fits of the force model with matching. (A) Fitted time courses of LS_1 and LS_2 from the constant velocity model (eq. 2.14) compared to data (spline interpolation) on a representative experimental repeat (OT-1 T cells, 4 peptides $\times$ 2 concentrations shown). Both time scales (α and β) are adjusted to the data, but the fits for LS_1 and LS_2 are almost identical, for 100k T cells, if α is fixed to $1/20 \text{ h}^{-1}$. (B) Same trajectories as (A), plotted against each other (4 concentrations shown per peptide). (C) Value of four parameters of the force model with matching (α and β not shown but fitted as well), fitted on each time series from 11 independent datasets (one including 4 technical replicates), for many peptides at different concentrations, with 100k initial T cells. There is a trade-off between this model’s fit accuracy and the simplicity of the constant velocity model’s parameter space.” (*Antigen encoding*, [1], SI)

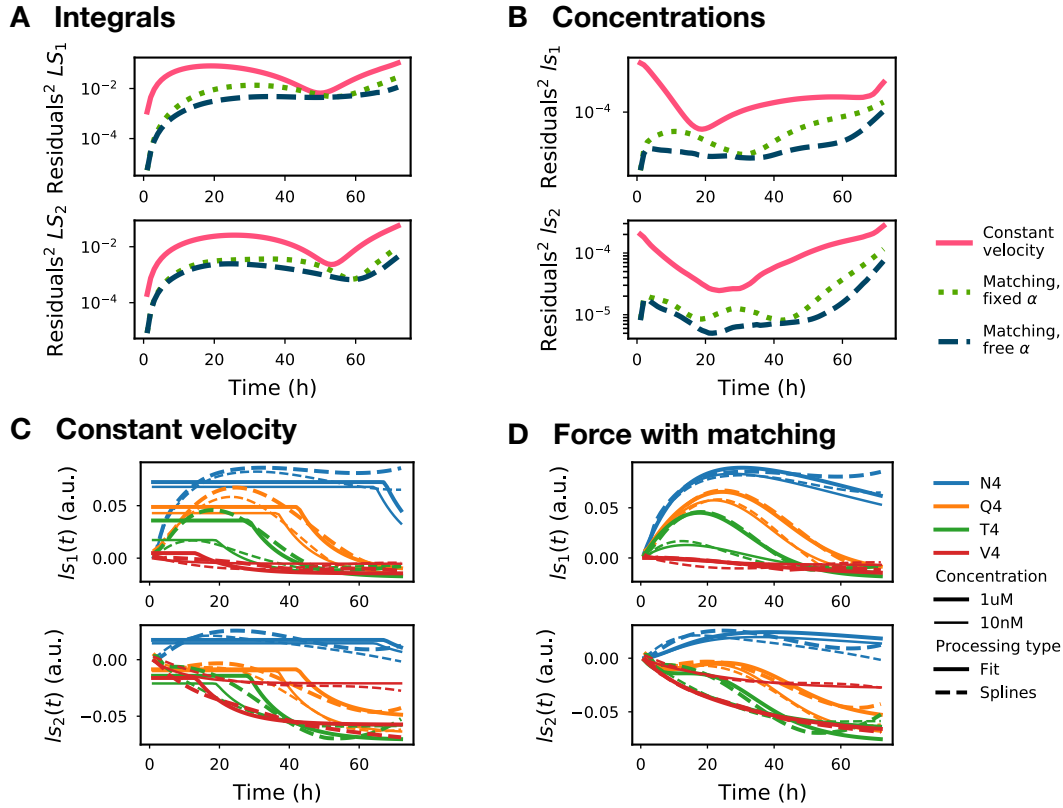


Figure 2.17: “Comparison of the quality of fits for the two latent space models. (A) Squared residuals between the data projected in latent space and fitted curves of $LS_1(t)$ and $LS_2(t)$, for four different models, averaged across 11 independent datasets (one including 4 technical replicates) containing 8 peptides, 8 concentrations, and 8 initial T cell numbers. (B) Same as (A), but for residuals on ls_1 and ls_2 (but recall that model parameters are always fitted on LS_1 and LS_2). (C, D) Comparison between cytokine concentrations projected in latent space and model trajectories of $ls_1(t)$ and $ls_2(t)$ for a representative dataset (100k initial OT-1 T cells, 4 peptides at 2 concentrations each shown). (C) Constant velocity model (eq. 2.14) (D) Force model with matching (eqs. 2.17 and 2.18). α and β are fitted separately. This model clearly describes more accurately the actual dynamics of cytokine concentrations in latent space.” (*Antigen encoding*, [1], SI)

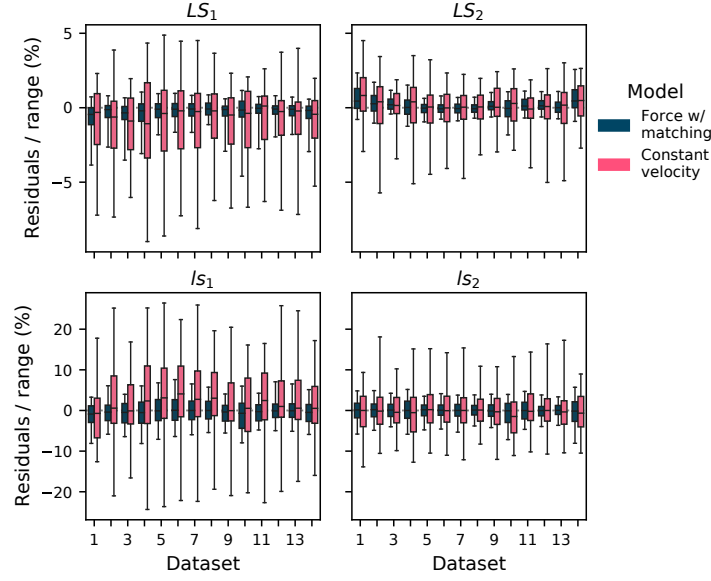


Figure 2.18: Residuals of the various latent space models : (top) for the latent space time integrals (LS_1 , LS_2); (bottom) for the latent space concentrations (ls_1 , ls_2). The residuals are normalized by the range spanned by the latent space variable ($LS_{1,2}$ or $ls_{1,2}$) within each experiment, across all times and conditions. The boxes' whiskers mark the 5th and 95th percentiles. For every dataset, zero is contained within the first and third quartiles, meaning there is no strong bias towards fits being above or below the spline-smoothed data in latent space. Residuals were calculated at every hour in every time series, for each of the 14 OT-1 T cells datasets used in figure 2.13. (*Antigen encoding*, [1], response to reviewers)

2. We sampled representative parameter values from those KDEs for each peptide. We computed the corresponding latent space concentration trajectories $ls_1(t)$, $ls_2(t)$, as well as $\tanh(LS_1(t)/\bar{N}_1)$ and $\tanh(LS_2(t)/\bar{N}_2)$ (figure 2.20B). Those are the latent space quantities that enter the accurate reconstruction algorithm described above.
3. We optimized the reconstruction coefficients on the same ensemble of datasets (equivalent results are obtained when optimizing on a separate dataset with many replicates). We finally applied this reconstruction method to the latent space trajectories computed from the sampled parameters, to obtain cytokine concentration time courses that look very similar to actual data.

Fig. 2.20C shows an example of synthetic time course for each peptide we usually have in experiments." (*Antigen encoding*, [1], SI)

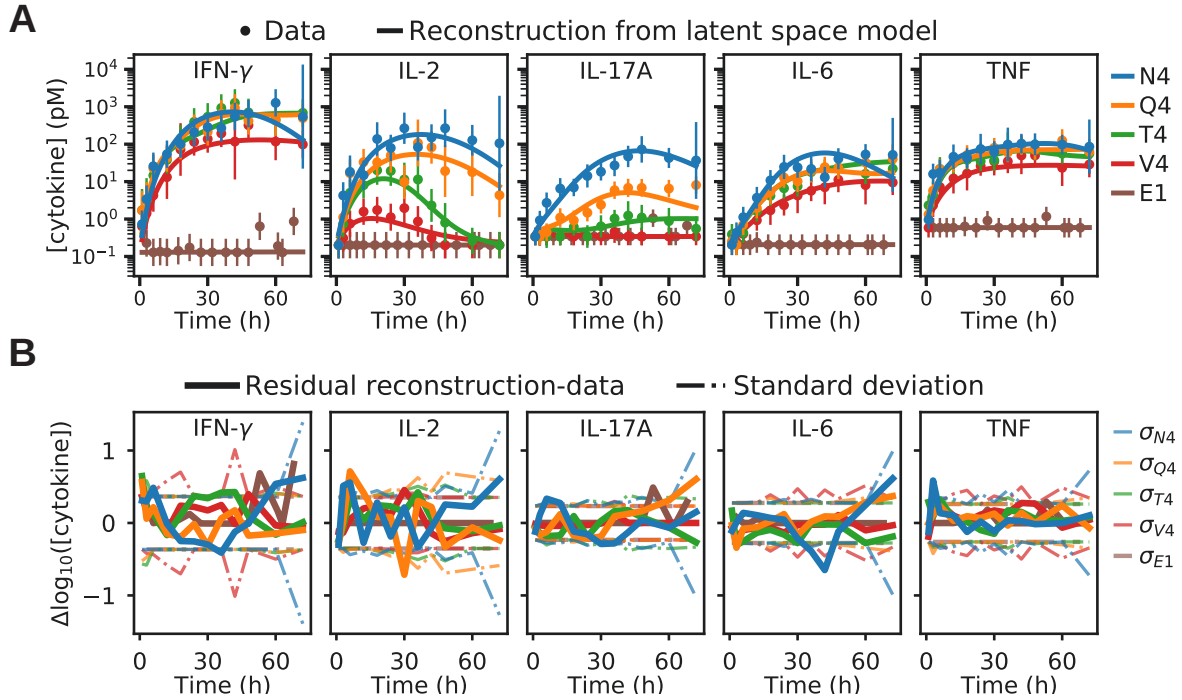


Figure 2.19: Latent space modelling and reconstruction provide a model for cytokine dynamics.

(A) Comparison, in log-scale, between experimental cytokine time courses (data) and time courses reconstructed from fits of the force model with matching in latent space, as well as (B) residuals of this model fit. Cytokine trajectories were obtained from latent space model curves using the non-linear reconstruction method, with reconstruction coefficients optimized on a separate dataset. Standard deviations of cytokine data (shown as error bars on (A) and dash-dotted lines on (B)) and covariances (not shown) were estimated from 11 OT-1 experimental replicates. Goodness of fit was assessed for each time course with a multivariate χ^2 test. For the time series shown, χ^2/ν values ranged between 1.16 (Q4, worst) and 0.56 (V4, best). Here, $\nu = 5 \times 12 - 6 = 54$ is the number of degrees of freedom in a five-dimensional time series of 12 time points fitted with 6 parameters. The corresponding p-values, which indicate the probability that a correct model would give a similar or larger χ^2 , ranged between 0.20 (Q4) and 0.998 (V4)." (*Antigen encoding*, [1], SI)

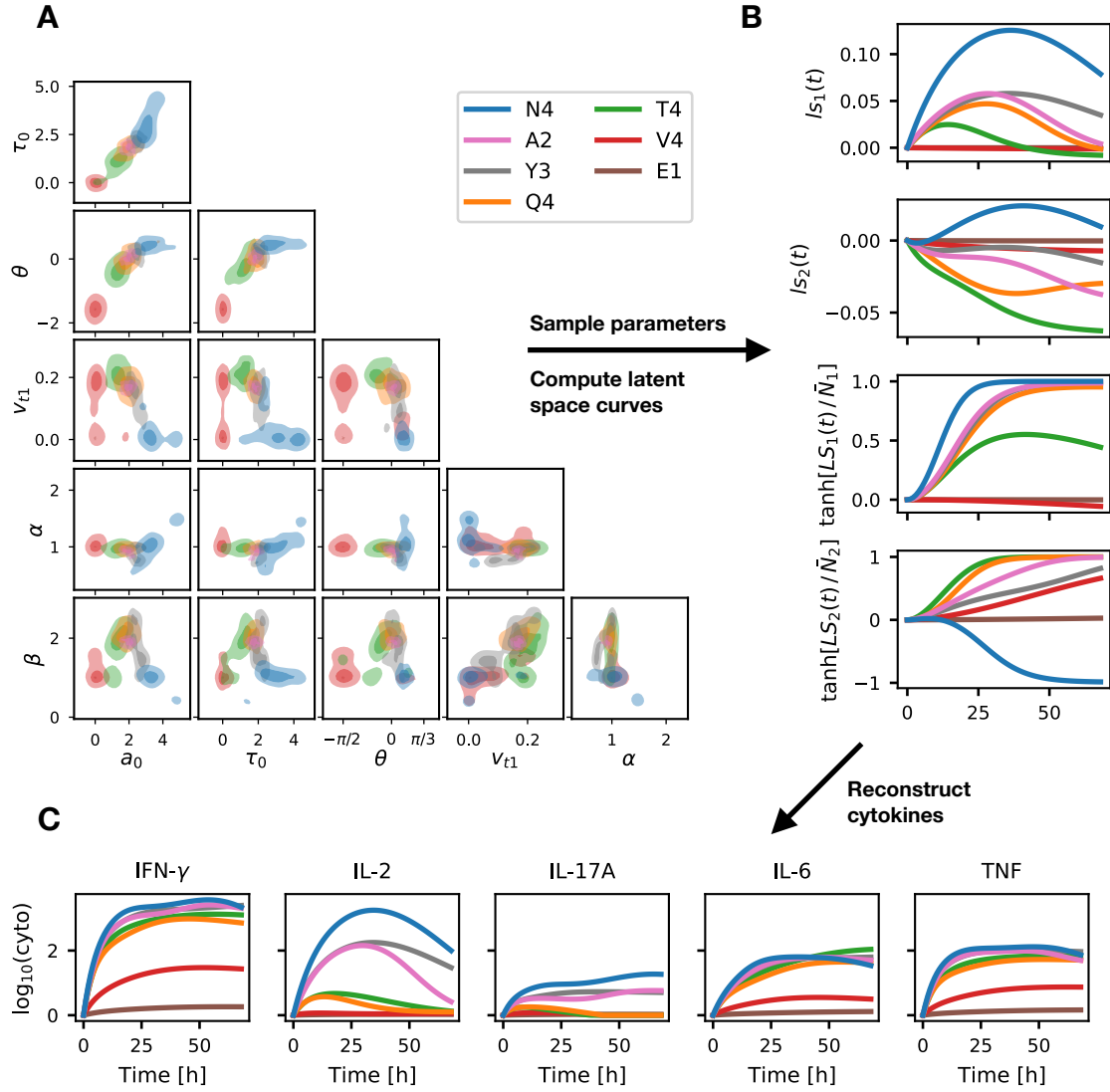


Figure 2.20: Generating synthetic cytokine time courses from the latent space model and reconstruction procedure. (A) Kernel density estimates (KDE) for each peptide over the complete parameter space of the force model with matching, fitted on a selection of 9 datasets. (B) Latent space time courses (concentrations and tanh-transformed integrals) corresponding to one set of model parameters sampled from the KDEs for each peptide. (C) Cytokine concentration time courses reconstructed from the above latent space trajectories, using reconstruction coefficients fitted on a separate experiment, as in figure 2.10. They are very similar to actual cytokine time series." (*Antigen encoding*, [1], SI)

To summarize, by building upon our initial discovery of antigen encoding in a latent space, we have developed dynamical equations which can be used to fit experimental cytokine time series or to generate realistic, synthetic time series simply by picking the value of a few interpretable parameters. Moreover, since all model parameters are mainly correlated to antigen quality, this means that we can generate and predict the cytokine time courses caused by the antigen of our choice. Hence, building on insights from a machine learning model solving the inverse problem (predicting antigenicity given cytokines), we have developed a complete model of the forward problem (predicting cytokines given antigen inputs), as it takes place in the actual biological system. We next decided to use this end-to-end model to quantify the information being transmitted by T cells via cytokines.

2.7 Channel capacity of antigen encoding in the latent space

We have already taken a first step towards quantifying the information content of cytokine dynamics in section 2.2.4, where we computed $MI(\mathbf{X}(t), Q)$ where $\mathbf{X}(t)$ was the concentration of a cytokine, or all of them, at a given time t .

“This calculation identified the most informative time points; however, it discarded information available in kinetic features. What is more, experimental variability may impact negatively the optimality of the chosen time point. Therefore, we estimated the information contained in time series with a second approach, inspired by [197]. It consisted in fitting the time courses in latent space with a model whose parameters capture the kinetic features of interest, and computing mutual information between peptide quality and those parameters, taken as $\mathbf{X}$. The distributions $f_{\mathbf{X}|Q=q}$ are then estimated in the space of model parameters (*i.e.*, figure 2.13). We found, with the same data used in figure 2.1E, that the constant velocity model parameters $\mathbf{X} = (v_0, t_0, \theta)$ contain nearly as much information (≈ 2.3 bits) about Q as cytokines do at any time point. Hence, the compression of cytokine trajectories to latent space and their model-based parameterization preserves almost all information available in cytokines. Note that this mutual information assumes an equiprobable choice of peptide strengths among a limited OVA set; the MI could be higher if more peptide categories were allowed and if p_Q were better tailored to the cytokine channel.”
(*Antigen encoding*, [1], SI)

The information theoretic quantity which answers this need of optimizing p_Q is the channel capacity between $\mathbf{X}$ and Q [198, 199, 200, 201, 55, 46]. It has been suggested that information capacity is a fundamental quantity optimized by living systems through evolution and adaptation [49, 51, 17]. We thus set out to compute information transmission in our concrete, data-driven example, to check whether this potential fundamental principle of biological physics would provide additional insight into T cell responses.

2.7.1 Channel capacity in model parameter space

“We define the channel capacity,

$$C(Q; \mathbf{X}) = \max_{p_Q} \text{MI}(Q; \mathbf{X}) \quad (2.20)$$

that is, the MI maximized over all possible Q distributions. This measure of information has two advantages: it characterizes the information capacity of the cytokine channel itself (dependence on the input distribution is eliminated), and it predicts the peptide distribution p_Q that can optimally exploit the channel.

In a nutshell, we start by fitting the parameter space for different peptides using multivariate distributions (section 2.7.2), then we interpolate distributions (section 2.7.4) to apply the Blahut-Arimoto algorithm [54, 55] and to compute channel capacity (section 2.7.5). Fig. 2.21A gives a simplified illustration of the cytokine response seen as a communication channel and of the procedure just outlined to estimate its capacity by re-sampling the peptide quality axis more densely via interpolation. In the next subsections, we give technical details on each step.” (*Antigen encoding*, [1], SI)

Treating the T cell cytokine response as a communication channel provides a more formal definition of the antigen encoding concept: T cells transmit information about antigen quality (input) into cytokine time series (output), and this information is encoded in the latent space and its dynamical model parameters. The question of how much information the latent space contains is thus well posed, allowing us to compute that quantity.

2.7.2 Fitting multivariate Gaussian distributions in parameter space

“We chose parameters a_0 , $\tau_0 = \alpha t_0$ and θ of the force model with matching to compose the vector $\mathbf{X}$ in parameter space, because the joint distribution of these

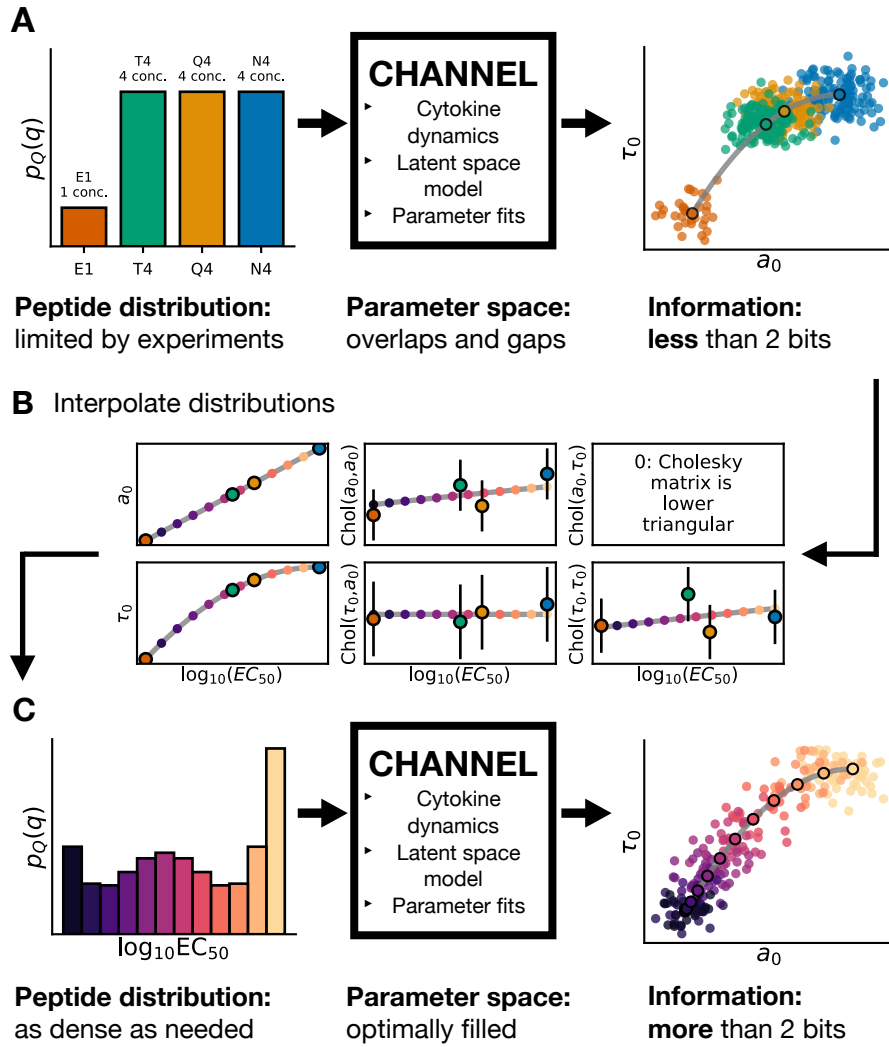


Figure 2.21: “Cartoon illustration of cytokine dynamics as an information channel for peptide quality, and of the procedure used to estimate the channel capacity. (A) Cytokine dynamics and their parameterization with a model (here, the constant force with matching model) act as a channel that maps peptides of different qualities (left) to different distributions in model parameter space (right). If we have only a few peptides from the experiment and do not maximize the MI, the channel’s output space is not optimally exploited and the information is limited by $H(Q)$. (B) We can interpolate the multivariate normal distributions in parameter space as a function of EC_{50} to obtain more closely sampled peptide qualities. More precisely, we interpolate the means and the Cholesky decomposition elements of the covariance matrix. (C) Information transmission is maximized after optimizing the input distribution p_Q over the re-sampled EC_{50} axis.” (*Antigen encoding*, [1], SI)

parameters, for each peptide, was the most closely Gaussian. We fitted the model on an experimental dataset comprising 8 OVA peptides, each at 4 log-spaced concentrations (from 1 μ M down to 1 nM), and with 9 replicates of each condition – giving a total of $K = 36$ time series per peptide. We applied the background noise filtering described in section B.1.2 and rejected 12 time series (out of $8 \times 36 = 288$) because of experimental errors. We rejected three parameter fits where $\theta > -\pi/3$ for peptide G4 as experimental outliers.

For model parameter fitting, we used prior knowledge on the correlations between a_0 and τ_0 as well as a_0 and θ , by adding terms to the cost function (eq. 2.15) of the form $c(p_i - aa_0 - b)^2$ ($p_i \in (\tau_0, \theta)$). Here, c is small (0.05), and a, b represent a linear relationship between parameters a_0 and τ_0 or θ , roughly estimated from a first model parameter fit without correlation terms. This extra regularization did not negatively impact the quality of the model fits, but attenuated the model parameters' sloppiness [193].

Then, we fitted multivariate normal distributions $\mathcal{N}(\boldsymbol{\mu}_q, \Sigma_q)$ on the sample points $\mathbf{x}_{k,q}$ in model parameter space, for each peptide q . We used the sample mean $\hat{\boldsymbol{\mu}}_q$ and the (unbiased) sample covariance matrix $\hat{\Sigma}_q^{ij}$ as estimators of the true underlying distribution for each peptide q . The variance of the mean estimator is simply

$$\text{Var}[\hat{\mu}_q^i] = \frac{\Sigma_q^{ii}}{K} \approx \frac{\hat{\Sigma}_q^{ii}}{K} . \quad (2.21)$$

The variance on the covariance matrix estimator was not needed, since we only used the Cholesky decomposition of Σ_q^{ij} , as explained in section 2.7.3 below.

Consequently, the multivariate normal distribution fitted on the sample points for peptide q is $\mathcal{N}(\hat{\boldsymbol{\mu}}_q, \hat{\Sigma}_q)$, which has the probability density function:

$$f_{\mathbf{X}|Q=q}(\mathbf{x}) = \frac{1}{(2\pi)^{N/2} |\hat{\Sigma}_q|^{1/2}} \exp \left(-\frac{1}{2} (\mathbf{x} - \hat{\boldsymbol{\mu}})^T \hat{\Sigma}_q^{-1} (\mathbf{x} - \hat{\boldsymbol{\mu}}) \right) , \quad (2.22)$$

where $N = 3$ is the number of dimensions of the parameter vector $\mathbf{X}$. Fig. 2.22 compares the original samples in parameter space to the fitted multivariate normal distributions." (*Antigen encoding*, [1], SI)

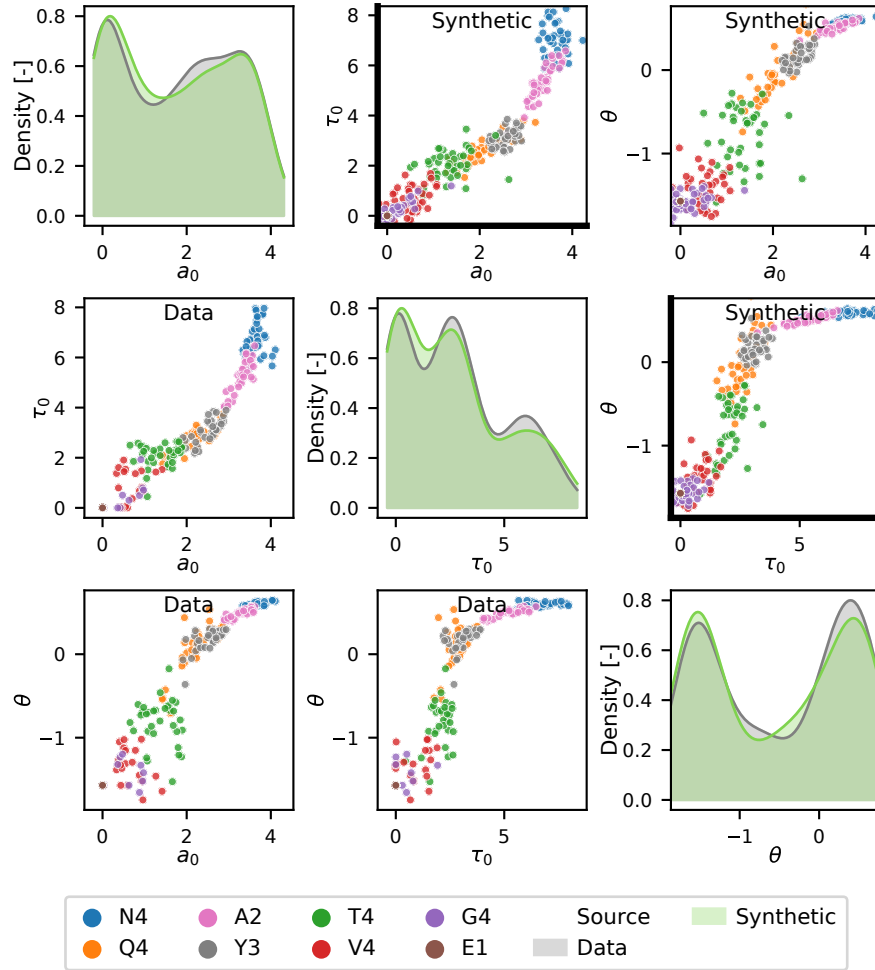


Figure 2.22: Comparison of empirical and fitted multivariate normal distributions of model parameters (“Data:” bottom half of the plot) and an equal number of points resampled from the multivariate normal *pdfs* fitted on them (“Synthetic”: upper half of the plot). Corresponding pair plots are located symmetrically across the diagonal. Graphs on the diagonal show the marginal distribution of each model parameter, all peptides combined (kernel density estimates from the scatter points), for the empirical and fitted distributions.” (*Antigen encoding*, [1], SI)

2.7.3 Cholesky decomposition estimator and its statistical variance

“We furthermore needed a statistical estimator $\hat{\Psi}$ for the Cholesky decomposition of the covariance matrix, $\Sigma = \Psi\Psi^T$. Ψ is a lower triangular matrix and can be computed straightforwardly for any positive definite matrix [196], such as Σ and $\hat{\Sigma}$. As explained below in section 2.7.4, this was necessary to interpolate between the covariance matrices fitted on the parameter distributions of peptides while ensuring their positive definiteness. We needed, in particular, the statistical variance of this estimator, to have error bars to interpolate between fitted distributions.

Olkin [202] provides an unbiased estimator $\hat{\Psi}$ for Ψ and the necessary steps to derive its variance. Their estimator is somewhat intricate to compute, and the covariance matrix reconstructed as $\hat{\Psi}\hat{\Psi}^T$ is biased. For simplicity, we chose instead

$$\hat{\Psi} = L \quad \text{where } LL^T = \hat{\Sigma} \quad (2.23)$$

that is, our estimator is simply the Cholesky decomposition of the covariance matrix estimator. It is biased for Ψ itself ($\langle \hat{\Psi} \rangle \neq \Psi$) but it does reconstruct an unbiased covariance matrix ($\langle \hat{\Psi}\hat{\Psi}^T \rangle = \langle \hat{\Sigma} \rangle = \Sigma$), which was our chief concern.” (*Antigen encoding*, [1], SI)

We can then derive the statistical variance of the Cholesky estimator. A mistake slipped into the result given by Olkin [202] for this variance. In the original article’s notation, they correctly prove that the variables u_{ii}^2 follow a χ^2 distribution with parameter $k = n - i + 1$ ($n = K - 1$ for us), so the correct conclusion is that $\langle u_{ii}^2 \rangle = n - i + 1$ and thus $\text{Var}[u_{ii}] = n - j + 1 - \langle u \rangle^2$. However, they report the variance of $\hat{\Psi}_{ii}$ as if it were rather u_{ii} that followed the χ^2 distribution, and thus they use the variance of a χ^2 distribution, $\text{Var}[u_{ii}] = 2(n - i + 1)$. This result is visibly incorrect, since it gives a variance that does not go to zero when $n \rightarrow \infty$.

Following the same derivation, but avoiding this confusion, we find

“

$$\begin{aligned} \text{Var}[\hat{\Psi}^{ii}] &= \frac{K - i - a_i^2}{K - 1} (\Psi^{ii})^2 \\ \text{Var}[\hat{\Psi}^{ij}] &= \frac{1}{K - 1} \sum_{l=j+1}^i (\Psi^{il})^2 + \frac{K - j - a_j^2}{K - 1} \Psi^{ij} \quad (i \neq j) \end{aligned} \quad (2.24)$$

where

$$a_i = \sqrt{2} \frac{\Gamma((K - i + 1)/2)}{\Gamma((K - i)/2)}. \quad (2.25)$$

and where we replace the (unknown) true Ψ by its estimator $\hat{\Psi}$. When $K \rightarrow \infty$, this variance goes to zero like $\frac{1}{K}$ since the asymptotic behaviour of the Γ function implies $a_j^2 \sim (K - j) - \frac{1}{2}$." (*Antigen encoding*, [1], SI)

2.7.4 Interpolating parameter distributions

"To compute the channel capacity without being restricted by the set of peptides available experimentally, we interpolated the multivariate normal distributions in parameter space. More specifically, each statistical estimator fitted on each experimental peptide's distribution – the $\hat{\mu}_q^i$ and the $\hat{\Sigma}_q^{ij}$ – is interpolated as a function of $Q = \log_{10} \text{EC}_{50}$, used as a measure of peptide quality. Q thus ranges from 0 (N4, reference) to 5.2 (E1). The $\log_{10} \text{EC}_{50}$ we use are averages of values reported by [174] and [176] and of our own EC_{50} assays (figure 2.23) for OVA variant antigens. Instead of interpolating matrix elements of $\hat{\Sigma}_q$ directly, we interpolated elements of the Cholesky decomposition $\hat{\Psi}_q$ and then built the interpolated covariance matrices as $\Sigma_{q'} = \Psi_{q'} \Psi_{q'}^T$. This method ensures that the interpolated $\hat{\Sigma}_{q'}$ remains positive definite and is thus a proper covariance matrix.

We used a two-step interpolation procedure for each individual distribution estimator (either $\hat{\mu}^i$ and $\hat{\Psi}^{ij}$). First, we fitted a smoothing cubic B-spline implemented in *SciPy* (the same we used in section 2.2.2). Each point on the Q axis was weighted as $w_q = 1/\sigma$ where σ is the standard deviation, or "error bar", of the statistical estimator, computed from formulas 2.21 and 2.24. The tolerance s was set equal to half the number of interpolated points (default setting). We then evaluated this smoothing spline function at the EC_{50} of the experimental peptides. Second, we fitted through those smoothed points a piecewise cubic Hermite interpolating polynomial (PCHIP) [203], as implemented in *SciPy* [187]. The latter interpolation functions were guaranteed to be monotonic between interpolated points, which prevents large non-monotonic artifacts.

This procedure yielded continuous functions $\hat{\mu}^i(q')$ and $\hat{\Sigma}^{ij}(q') = \sum_{l=1}^N \hat{\Psi}^{il}(q') \hat{\Psi}^{jl}(q')$ that constitute a multivariate normal distribution in parameter space, evaluated for any desired peptide quality $q' = \log_{10} \text{EC}_{50}$ in the range $[0, 5.2]$. These interpolations are shown in figure 2.24. Note how the average value of all three parameters increased

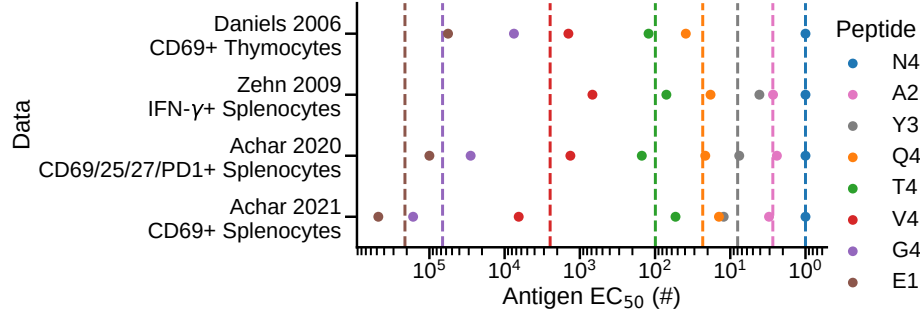


Figure 2.23: Comparison of EC_{50} of OT-1 ligands published in different papers [174, 176] with our own calibration experiments (Achar, 2020&2021). These EC_{50} were obtained with dose responses of antigen for different assays (CD69 upregulation by thymocytes, IFN- γ by naive splenocytes, and upregulation of activation markers in naive splenocytes). Dashed lines indicate the average we considered. (*Antigen encoding*, [1], response to reviewers)

almost linearly with decreasing EC_{50} in logarithmic scale, making them good readouts of antigen quality. Moreover, note how the Cholesky decomposition of the covariance increased for peptides of intermediate antigenicity (*e.g.*, T4, V4, in the 10^2 – 10^3 EC_{50} range). This variability arose because these peptides could produce detectable cytokine signals or not, depending on their concentrations.” (*Antigen encoding*, [1], SI)

We note here that this interpolation procedure amounts to writing all cytokine dynamics as a function of a single parameter: the antigen quality, measured in terms of $\log_{10} EC_{50}$. Moreover, since all dynamical parameters are related to a_0 (or v_0) through the EC_{50} , the entire T cell response over days is in fact determined by the early cytokine kinetics.

2.7.5 Blahut-Arimoto algorithm with Monte-Carlo integration

“To compute the channel capacity, we implemented the Blahut-Arimoto algorithm (figure 2.25B), which provably converges to the optimum for continuous input Q and output X random variables [54]. We discretized the input Q into $M = 25$ discrete values of $\log_{10}(EC_{50})$, because it is numerically simpler to optimize a discrete probability mass function p_Q . We coded the algorithm in C and embedded it in a Python C-API module to be executed as a Python script.

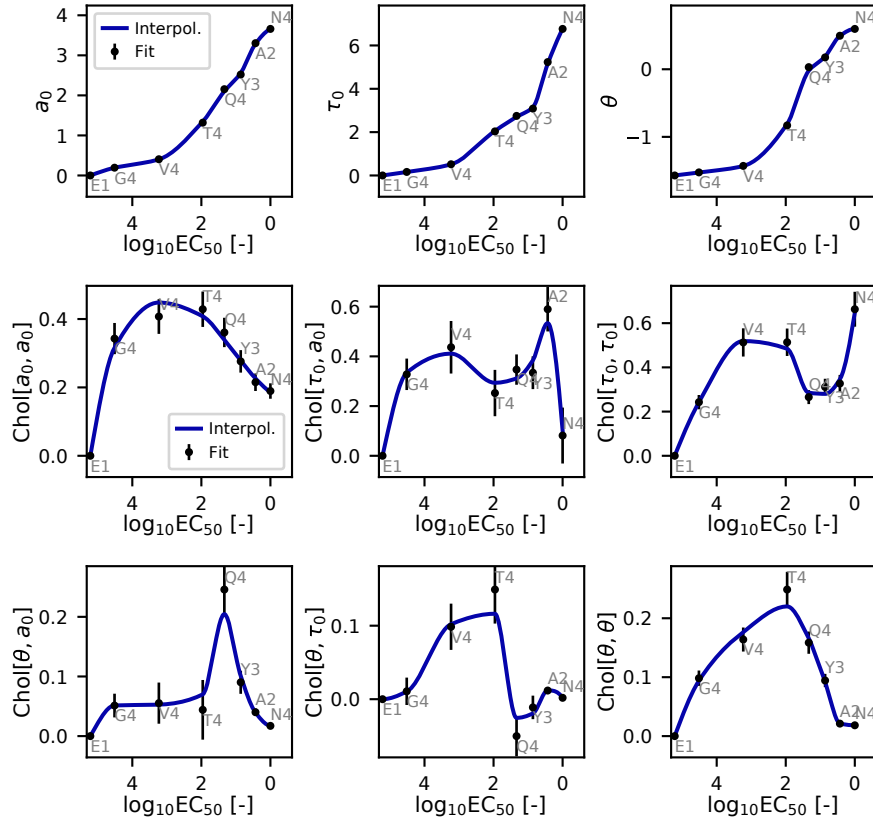


Figure 2.24: Interpolation of multivariate normal distributions in the model parameter space as a function of $\log_{10} EC_{50}$. The first row shows the interpolation of the mean value of each parameter (a_0, τ_0, θ), while the last two rows show the interpolation of the Cholesky decomposition (lower triangular matrix) of the covariance matrix. The Cholesky decomposition Ψ is used to preserve positive definiteness of the covariance matrices, reconstructed as $\Sigma = \Psi\Psi^T$. Piecewise-cubic Hermite polynomials (which are monotonous between data points) are used for the final interpolation after an initial smoothing spline interpolation. (*Antigen encoding*, [1], SI)

The crucial step in each iteration of the Blahut-Arimoto algorithm is to compute the quantities

$$c_j = \exp(b_j); b_j = \int_{\mathcal{X}} d\mathbf{x} f_{\mathbf{X}|Q=q}(\mathbf{x}) \frac{f_{\mathbf{X}|Q=q}(\mathbf{x})}{\sum_{q'} p_{q'} f_{\mathbf{X}|Q=q'}(\mathbf{x})}, \quad (2.26)$$

which serve to update the estimate of the optimal probability vector p_Q . Because the output $\mathbf{X}$ is continuous and has $N = 3$ dimensions, the sums in the original definition of the c_j are replaced by multidimensional integrals. The latter are computationally expensive and need to be evaluated often, so we resorted to Monte Carlo integration. To reduce the number of samples needed to achieve a desired accuracy σ_{c_j} on the integral estimate, we drew them from the multivariate normal distribution $f_{\mathbf{X}|Q=q}$, since it is similar to the integrand $f_{\mathbf{X}|Q=q}(\mathbf{x}) \frac{f_{\mathbf{X}|Q=q}(\mathbf{x})}{\sum_{q'} p_{q'} f_{\mathbf{X}|Q=q'}(\mathbf{x})}$, as illustrated in figure 2.25A. If the relative error (stopping criterion) resired for the channel capacity is E_C , we set $\sigma_{c_j} = E_C/M$, ensuring that the lower bound on the capacity, $C_{low} = \log_2 \left(\sum_j p_j^x c_j^x \right)$, has a relative error of E_C or less. The algorithm starts with $k = 10^5$ samples and adaptively draws more if needed (usually not necessary for $E_C = 1\%$).

Monte Carlo samples from $f_{\mathbf{X}|Q=q}$ were generated as follows. We generated uniform random numbers with the dSFMT package, which is a fast and high-quality Mersenne Twister [204]. We transformed the uniform samples to get Nk univariate $\mathcal{N}(0, 1)$ samples with the Box-Muller method [196], and transformed them again into k multivariate samples using the Cholesky decomposition Ψ of the covariance matrix Σ [205].

We validated the algorithm against various configurations of multivariate Gaussian distributions [164]. In every case, it converged to the correct value within the desired accuracy, which we typically set to 1 %." (*Antigen encoding*, [1], SI)

This compared favorably to the channel capacity estimation algorithm of [164], which has an accuracy of 4 % on the test cases.

2.7.6 Channel capacity results and final estimate for the number of antigen classes

"The Blahut-Arimoto algorithm converged to a capacity of 2.60 bits (within 1% tolerance for convergence). To assess the robustness of this result, we repeated the

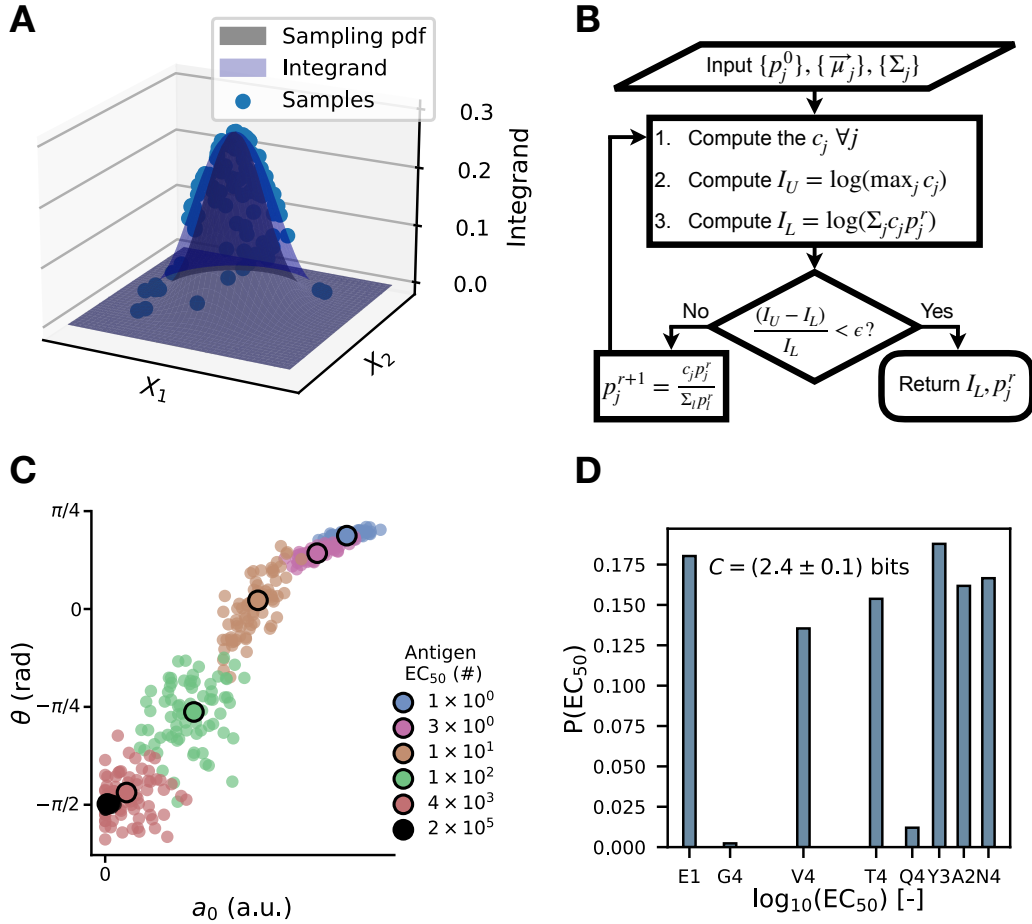


Figure 2.25: Blahut-Arimoto algorithm to derive the cytokine channel capacity from the model parameter space. (A) Cartoon illustrating the Monte Carlo integration of equation 2.26 with variance reduction by sampling from the input-output distribution, $f_{X|Q=q}(x)$. (B) Flowchart of the Blahut-Arimoto algorithm (adapted from [54]). (C) Model parameter distributions corresponding to the theoretical antigen categories determined from the channel capacity calculation, as explained in figure 2.12C. (D) Optimal peptide distribution found with the channel capacity estimation algorithm of Grabowski *et al.* [164], giving a capacity of (2.4 ± 0.1) bits (the algorithm has a relative accuracy of 4 %). As expected, the result is slightly lower than when interpolated EC_{50} are allowed $[(2.59 \pm 0.03)$ bits, subsection 2.7.6]. (Antigen encoding, [1], SI)

procedure 32 times, including the model curve fitting and multivariate distribution fitting, while perturbing randomly the regularization hyper-parameters involved in model parameter fitting. Perturbations were sampled from a normal distribution with a standard deviation of 0.05, for regularization coefficients that are on the order of 0.8. We found an average capacity of (2.59 ± 0.03) bits, where the error represents the standard deviation of the replicates, thus confirming the robustness of the calculation." (*Antigen encoding*, [1], SI)

Figure 2.12C shows the optimal distribution p_Q maximizing the mutual information $MI(\mathbf{X}, Q)$, as found by the Blahut-Arimoto algorithm. Since the capacity corresponds to $2^{2.59} \approx 6$ categories that should be completely distinguishable, we have subdivided the continuum of antigen qualities into six classes: the two limiting EC_{50} values (which include all theoretically weaker or stronger antigens as well) and four other EC_{50} values equally distributed in the cumulative probability function (CDF) of Q (figure 2.12C, center). These six prototypical antigen affinities indeed produce non-overlapping latent space responses (figure 2.12C, right), as well as non-overlapping ballistic model parameter distributions (figure 2.25C).

We note that the two probability peaks at the endpoints of the EC_{50} axis are reminiscent of the optimal distributions found for concentration sensing in chapter 1, figure 1.2. This reflects the inverse relationship imposed by channel capacity between the optimal p_Q and the variance of parameter distribution $f_{\mathbf{X}|Q=q}$. As can be seen in figures 2.22 and 2.24, there is little variability in model parameters for all-out (N4, A2) or null (None, E1) T cell responses, while there is more noise for responses in the middle, *i.e.*, for intermediate agonists (V4, T4) at the threshold of T cell activation.

“As a point of comparison, we ran the algorithm of [164] on the empirical parameter distribution for the seven available peptides. We found a channel capacity of (2.4 ± 0.1) bits and the optimal distribution shown in figure 2.25D. This slightly lower capacity is consistent with the fact that the EC_{50} axis is not optimally sampled when only empirical distributions are used, as opposed to our interpolation method.

Notice that Fig. 2.12C-D allows for a direct comparison between the EC_{50} s of the 6 antigen classes and the classical OT-1 ligands. The first category has EC_{50} larger than E1, so should correspond to self. The second category is between G4 and V4, so in the middle of the antagonistic regime. The four other categories are spread

above V4, so correspond to agonists of different strengths." (*Antigen encoding*, [1], SI)

2.7.7 Interpretation of latent space weights

We used the prototypical antigen classes to clarify how the antigen classifier of section 2.3 was exploiting the latent space to perform antigen quality classification. We first examined how $LS_1(t)$ and $LS_2(t)$ at a fixed time point ($t = 36$ h) varied as a function of antigen EC_{50} and we found that LS_1 increases monotonically with antigen quality, while LS_2 varies non-monotonically (figure 2.12D). Each prototypical antigen class lands in a distinct part of these curves, such that they each correspond to a unique pattern of LS_1 and LS_2 being null, positive or negative (figure 2.12E). This suggests that cytokine dynamics could easily be decoded for antigen quality by discrete regulatory elements, such as the cytokine receptor pathways in neighbouring T cells reading out a cytokine signal to be co-opted into response. This is consistent with findings in the context of information optimization in genetic networks, showing that non-monotonicity is important for the decoding of biological signals by binary pathways [206, figure 3].

The behaviour of each LS suggested a biological interpretation of their composition, as revealed by the neural network's set of weights connecting the inputs to the latent space (figure 2.26A): LS_1 is mostly made up of IL-2, which increases with activation, while LS_2 is a difference between the two cytokines IL-2 and IL-17A on the one hand, and cytokines IFN γ , IL-6, and TNF on the other. Since the first two are mostly associated with adaptive immunity and the last three, with innate immunity [4], we would expect the last three to fire up at low antigen qualities because innate immunity is less antigen-specific, and thus make LS_2 decrease due to their negative contribution to it, whereas adaptive immunity cytokines would have a more gradual response.

The second set of neural network weights, connecting the latent space to the output layer, closely matched the patterns seen in LS_1, LS_2 as a function of antigen quality, with weights reading out LS_1 increasing (essentially) monotonically and weights reading out LS_2 being non-monotonic (figure 2.26B). These weights decode the latent space by dividing it into sectors each attributed to one class, as seen in figure 2.26C, where we computed the softmax output $p_q(LS_1, LS_2)$ (eq. 2.8) at each point in latent space, and plotted the color of the most likely predicted antigen class among the six OVA training peptides³.

³To be precise, we plot $p_q(LS_1, LS_2)$ for all six antigen classes q at every point, on top of each other, each in a different color, but we set the opacity for each color equal to p_q : hence, only the color of the most probable peptide in a given region is non-transparent and thus shows up in that region. Boundaries are generally very sharp, except for the V4-G4 distinction.

Moreover, interpolating the network’s output weights as a function of EC_{50} and evaluating them for each of the six prototypical antigen class, we could visualize the regions of latent space most associated to each theoretical class (figure 2.26D). Unsurprisingly, they corresponded to the regions containing the prototypical latent space trajectories of figure 2.12D. Figure 2.26D also reveals how the training antigen class Q4 can in fact be subdivided roughly in two distinct classes: the training set was not completely filling the continuum of possible antigen qualities. Yet, the neural network had the ability to place intermediate new antigens like A2 and Y3 in the right regions of latent space (as seen in figure 2.4D-E), supporting further the idea that this latent space is a natural encoding of antigen quality.

2.7.8 Cytokine trajectories for prototypical antigen classes

To close the loop, we reconstructed the cytokine time series corresponding to the six prototypical antigen categories of figure 2.12C. We used the parameter covariance estimates of section 2.7.2 and the nonlinear decoder of figure 2.4D-E to reconstruct sample cytokine trajectories for each class. We show in figure 2.27, the reconstructed cytokine trajectories for archetypal antigens. We show, for each antigen, both the average cytokine time course and its standard deviation (obtained numerically by generating 100 trajectories for each antigen prototype). Consistent with our mutual information optimization, we see well-defined trajectories, especially for the stronger ligands, and larger uncertainty for weaker ligands (category 5, $EC_{50} = 4 \times 10^3$). While these trajectories exhibit a few artifacts, for some cytokines and antigens in particular, they still have the correct general shape and ordering according to antigen quality.

Generative modelling by pure machine learning usually requires large datasets to produce realistic synthetic outputs. Even then, finer details in generative modelling can easily go awry, as illustrated by symmetry or background artifacts in random face generation by generative adversarial networks [207, 208]. It is therefore expected that some artifacts remain in the synthetic cytokine time series. Owing to our biophysically-informed modelling, we were in fact able to do a lot with relatively “little” data (compared to traditional machine learning datasets). Of course, more experiments with OT-1 T cells would enable better reconstruction quality, correctly capturing all possible intrinsic and extrinsic sources of variability. This improvement in accuracy with additional data, however, was beyond the scope of our work. In the next section, we rather focused on the universal character of antigen encoding in different immune contexts.

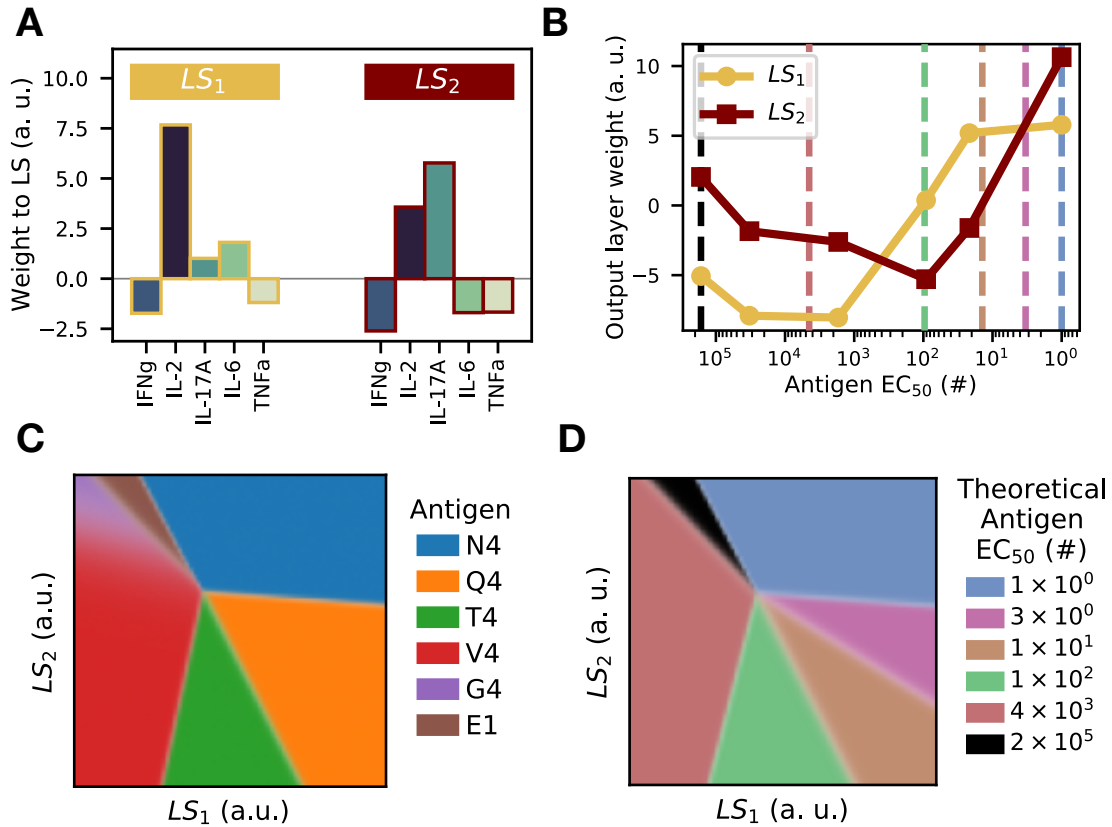


Figure 2.26: “Neural network weights to each layer and their interpretation. (A) Weights connecting the input layer to the intermediate layer give the contribution of each cytokine to each latent space variable. (B) Weights to the neural network’s prediction layer, from each latent space variable to each training antigen category. For optimal accuracy of the network, those weights have to match the monotonic (LS_1) and non-monotonic (LS_2) behaviour of the latent space variables as a function of EC_{50} , which was shown in figure 2.12D. (C) Latent space colored according to which training antigen is the most probable prediction of the neural network at each point, as determined from the weights of panel (B) and the softmax function in eq. (2.8). This reflects antigen encoding: each antigen has its own sector, corresponding to the typical location of its latent space trajectories. Notice that since the last layer before the softmax regularization is linear, those sectors are separated by straight lines; intuitively the sectors define where the latent space trajectories corresponding to each ligand can be located (D) Latent space domains for the theoretical antigen classes determined from channel capacity calculation. The weights defining this map were linearly interpolated from the weights for training antigens at the theoretical antigens’ EC_{50} s, indicated by vertical dashed lines on panel (B).” (*Antigen encoding*, [1], SI)

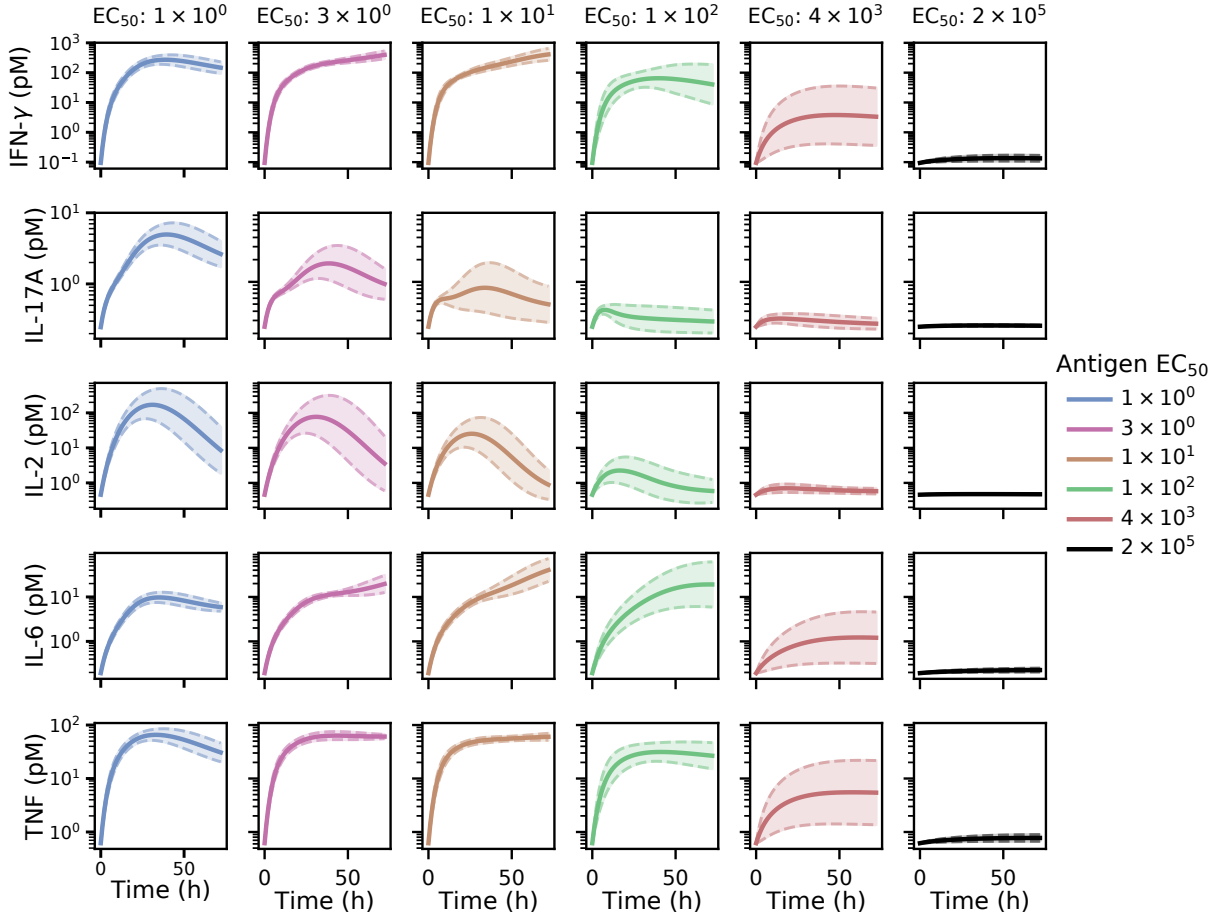


Figure 2.27: Reconstruction of cytokine trajectories for the 6 ligand categories obtained through mutual information optimization, using the decoder of Fig. 2.4D and the noise structure of Fig. 2.12C. (*Antigen encoding*, [1], response to reviewers)

2.8 Universality of antigen encoding

To support the idea that antigen encoding is a (relatively) universal feature of T cell responses, and not merely an idiosyncrasy of the OT-1 TCR and its cognate peptides, we tested how our latent space description, as trained on OT-1 data sets, could generalize to diverse immunological settings, where the number of initial T cells, the APC cell types, the TCR line, and the T cell donor species were varied. We also tested a wide panel of drug perturbations to try to break the correlation between latent space dynamical parameters. Some of these diversified datasets required custom preprocessing, which we discuss in appendix B.

2.8.1 Testing various immunological conditions

We tested antigen encoding in several different immune settings. Our collaborators performed a whole panel of robot-assisted experiments where cocultures were varied from their default preparation (OT-1 T cells and B6 splenocytes). On the T cell side, we measured cytokine responses of pre-activated (“blast”) T cells, as opposed to naive T cells; we also tested an array of initial T cell numbers, to assess how this parameter impacts cytokine trajectories. We assessed additional CD8⁺ T cell lines beyond OT-1, and even performed experiments with CD4⁺ T cells. On the APC side, we measured T cell responses in cocultures containing macrophages, dendritic cells, tumor cells, or other APC types pulsed with peptides – instead of the mixed splenocytes used for training the model. The experimental protocols are detailed in the SI of *Antigen encoding*, [1]. The results are presented below to highlight how latent space cytokine trajectories remain similarly structured by antigen quality across diverse conditions, rather than to discuss their specificities in detail.

Most figures in this subsection (2.28 – 2.31) were prepared by Sooraj Achar and generated using the latent space fitting pipeline [185] developed by us with Thomas Rademaker. I reproduce them here for completeness and to illustrate the generalizability of our approach. Analyzing the effect of T cell number on the v_0 vs θ slope (figure 2.28E) and the Earth Mover’s distance analysis of drug perturbations (figure 2.28F, right) was my own work.

To assess how our latent space model could capture antigen quality across different conditions, we

“first plotted the parameter v_0 (Fig. 2.12A) as a function of EC_{50} for both our training and test sets and recovered a log-linear correlation above a minimum threshold (Fig. 2.28A): our model of cytokine dynamics yielded a measure of antigen quality, v_0 , as faithful as an EC_{50} derived from a functional dose response. We further confirmed the predictive power of v_0 on additional immunological settings: i. two states

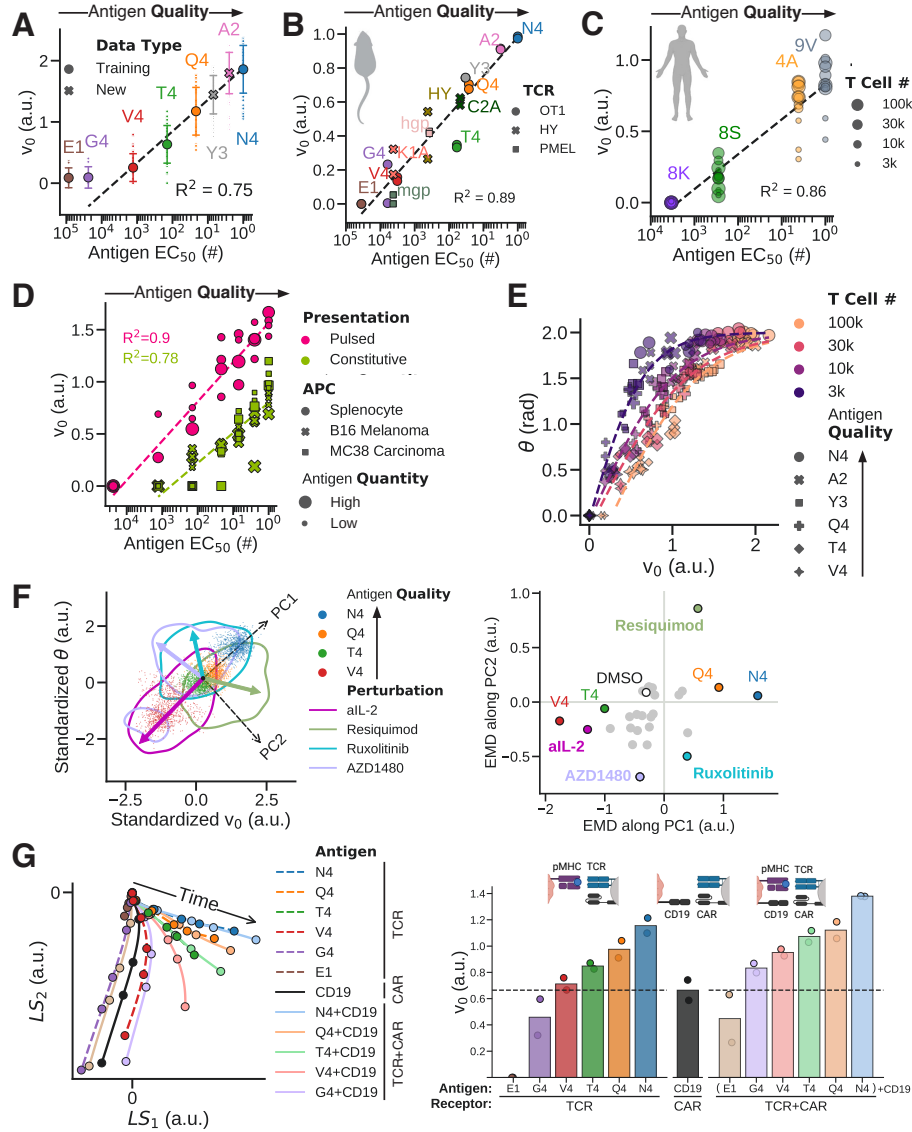


Figure 2.28: “Antigen encoding to explore diverse immunological settings. (A) Testing the ballistic model of the latent space on all OT-1 antigens yielded measurements of v_0 that correlated with their rank (as defined by the EC_{50} of a functional dose-response). (B) Antigen encoding learned from naïve OT-1 T cells yields accurate v_0 ranking with different mouse TCR transgenic T cells, (C) with human TCR transgenic T cells and (D) with OT-1 T cells responding to tumors constitutively expressing antigens. (E) Antigen encoding captures dependency on T cell numbers in (v_0, θ) space. (F) (left) Antigen encoding (v_0, θ) of T cell activation under various molecular perturbations. (right) Earth Mover’s Distance (EMD) for individual antigens. PC1 captures the variation in antigen quality ($EMD_2 = 0$). While most molecular perturbations simply downgraded the apparent antigenicity ($EMD_1 < 0$, $EMD_2 \approx 0$), JAK inhibitors (AZD1480 and Ruxolitinib) severely impacted T cell activation ($EMD_2 < 0$) and the TLR7/8 agonist Resiquimod boosted it ($EMD_2 > 0$). (G) Antigen encoding of anti-CD19 Chimeric Antigen Receptor (CAR)-T cells (left) revealed a pattern of weak activation and antagonism for cytokine responses (right).” (Antigen encoding, [1])

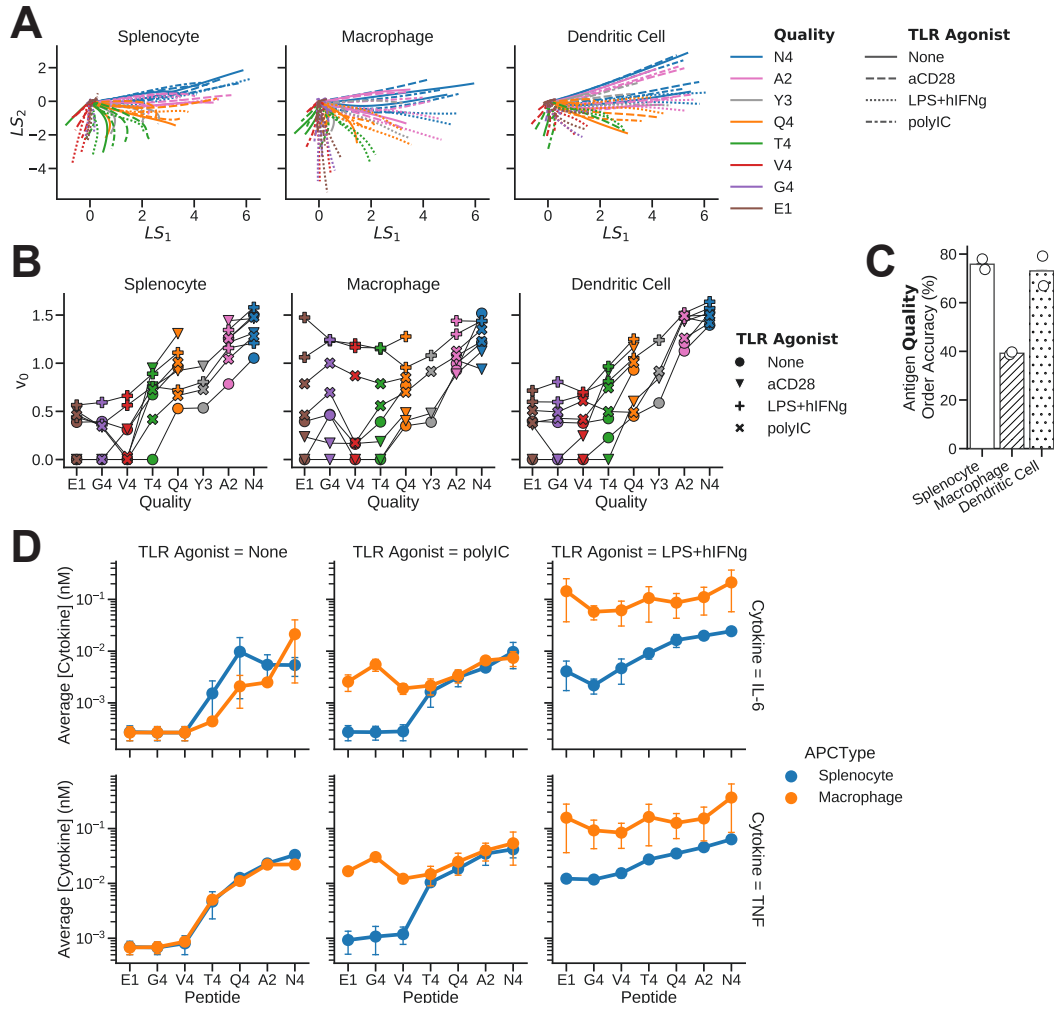


Figure 2.29: “Antigen encoding with different antigen–presenting settings. (A) We compared the latent space projection (*i.e.* antigen encoding) of the cytokine dynamics for naive OT-1 T cells, activated with peptide-pulsed splenocytes, or bone-marrow derived macrophages or bone-marrow derived dendritic cells, in the presence of varied costimulation (aCD28) or TLR agonists (LPS or poly-IC). The antigen encoding was trained with splenocytes and remained consistent across different antigen-presenting cells. **(B)** Immune velocity (v_0) as a measure of antigenicity was defined using splenocytes as antigen presenting cells, and remained essentially accurate for other antigen presentation settings – macrophages or dendritic cells with or without costimulation of inflammatory signals. **(C)** Immune velocity (v_0) has high order accuracy in splenocytes and dendritic cells regardless of costimulatory and inflammatory signals, although macrophages perform more poorly in this setting. **(D)** Examination of the input cytokines reveals that stimulation of macrophages with the polyIC and LPS TLR agonists induces a high, antigen–quality–independent secretion of the innate cytokines IL-6 and TNF, leading to a high background of activation in panel **B** and a more limited v_0 order accuracy observed in panel **C**. The data used in this figure are from two independent experiments.

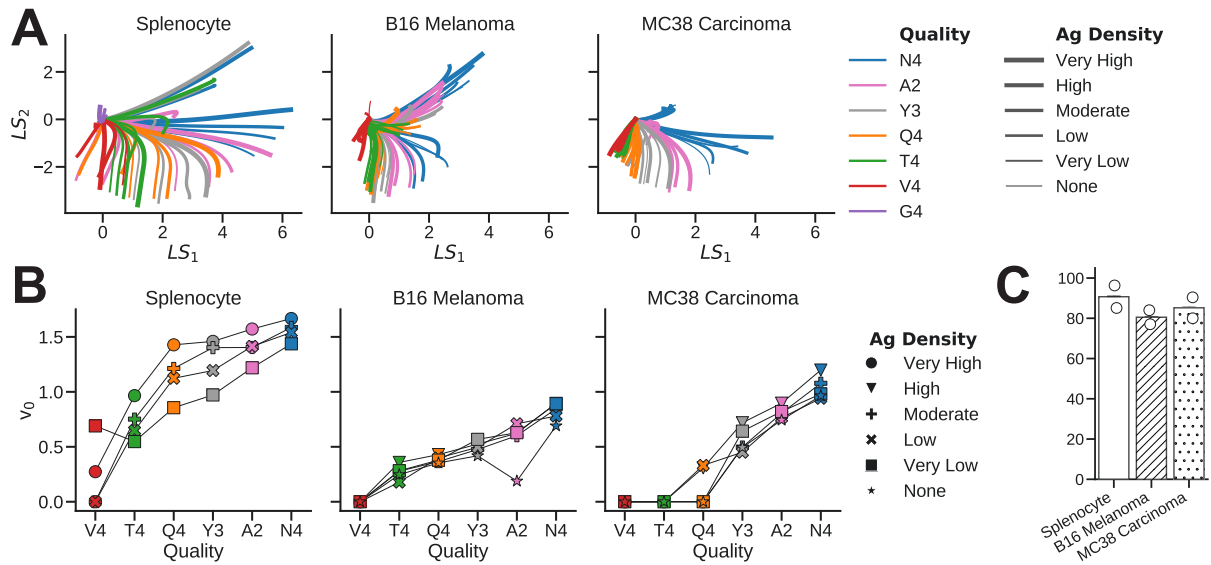


Figure 2.30: “Antigen encoding with different tumor settings. (A) We compared the latent space projections (*i.e.* antigen encoding) of the cytokine dynamics for effector OT-1 T cells, activated with peptide-pulsed splenocytes (training set) or with B16 melanoma or MC38 carcinoma tumor cells constitutively expressing OT-1 antigens. Antigen density on the tumor cells was increased using a 48hr pulsing with IFN- γ pre-activation. **(B)** Immune velocity (v_0) as a measure of antigenicity was defined using splenocytes as antigen presenting cells, and remained essentially accurate for peptides constitutively expressed and presented on the surface of B16 melanoma or MC38 carcinoma tumor cells. **(C)** Immune velocity (v_0) has high order accuracy across all tested tumors regardless of antigen density. The data used in this figure are from two independent experiments.” (*Antigen encoding*, [1], SI)

of differentiation of OT-1 T cells (naïve or pre-activated effector; Fig. B.6); ii. different antigen-presenting cells (splenocytes, bone marrow derived macrophages and dendritic cells) pulsed with peptides (Fig. 2.29); iii. different mouse TCR transgenic T cells (Fig. 2.28B); iv. human T cells expressing a TCR transgene (Fig. 2.28C); and v. different tumors constitutively expressing antigens (Fig. 2.28D, 2.30). While the latent space may rotate (angle θ , Fig. 2.12A) according to immune contexts – *e.g.*, when varying the number of T cells (Fig. 2.28E, B.7) – in all settings, the structure of the latent space was preserved and v_0 accurately ranked antigens according to their quality. Overall, our model generalized well to infer antigen quality from unseen patterns of immune activation.” (*Antigen encoding*, [1])

Drug perturbations

“We tested whether molecular perturbations (small drug inhibitors, cytokines or antibody blocking) could generate novel cytokine dynamics. For most perturbations, the strong correlation between parameters of antigen encoding remained intact, and antigen hierarchy was preserved, albeit with a downshift in apparent antigen quality. However, two types of perturbation (activation of innate signaling and inhibition of cytokine signaling) maximally broke this parameter correlation and generated distinct dynamic patterns, *i.e.*, , new classes of T cell activation with potentially new functions (Fig. 2.28F, 2.31).” (*Antigen encoding*, [1])

As explained in the figure caption, we quantified shifts in parameter distributions using the Earth Mover’s Distance (EMD). This metric is equal to the average distance over which each point of the starting distribution must be moved to create the final distribution with the minimal possible total transport work [209]. In the general case, the EMD is computed using tools from graph theory. Summarizing, we have two sets of sample points, $X_1 = \{x_{1,i}\}$ and $X_2 = \{x_{2,j}\}$, from the initial and final distributions, respectively, with weights $\{w_{1,i}\}$ and $\{w_{2,j}\}$ (typically, $w_{a,i} = \frac{1}{\text{card}(X_a)}$). First, we build a bipartite graph where nodes are samples from X_1 or X_2 , and edges connect each point in X_1 to all points in X_2 and have weights d_{ij} equal to the distance between samples $x_{1,i}$ and $x_{2,j}$. Then, we use linear programming to compute the flow f_{ij} which transports all weight from X_1 towards X_2 (f_{ij} is how much weight is transported on edge (i, j)), while minimizing the total “work” $C = \sum_{i,j} d_{ij} f_{ij}$ (weight $\times$ distance). The EMD is equal to this total minimal work, divided by the total weight in X_1 or X_2 (which is 1 if $w_{a,i} = \frac{1}{\text{card}(X_a)}$). We wrote our own Python implementation of these steps, relying on the min flow solver of the *networkx* package [210].

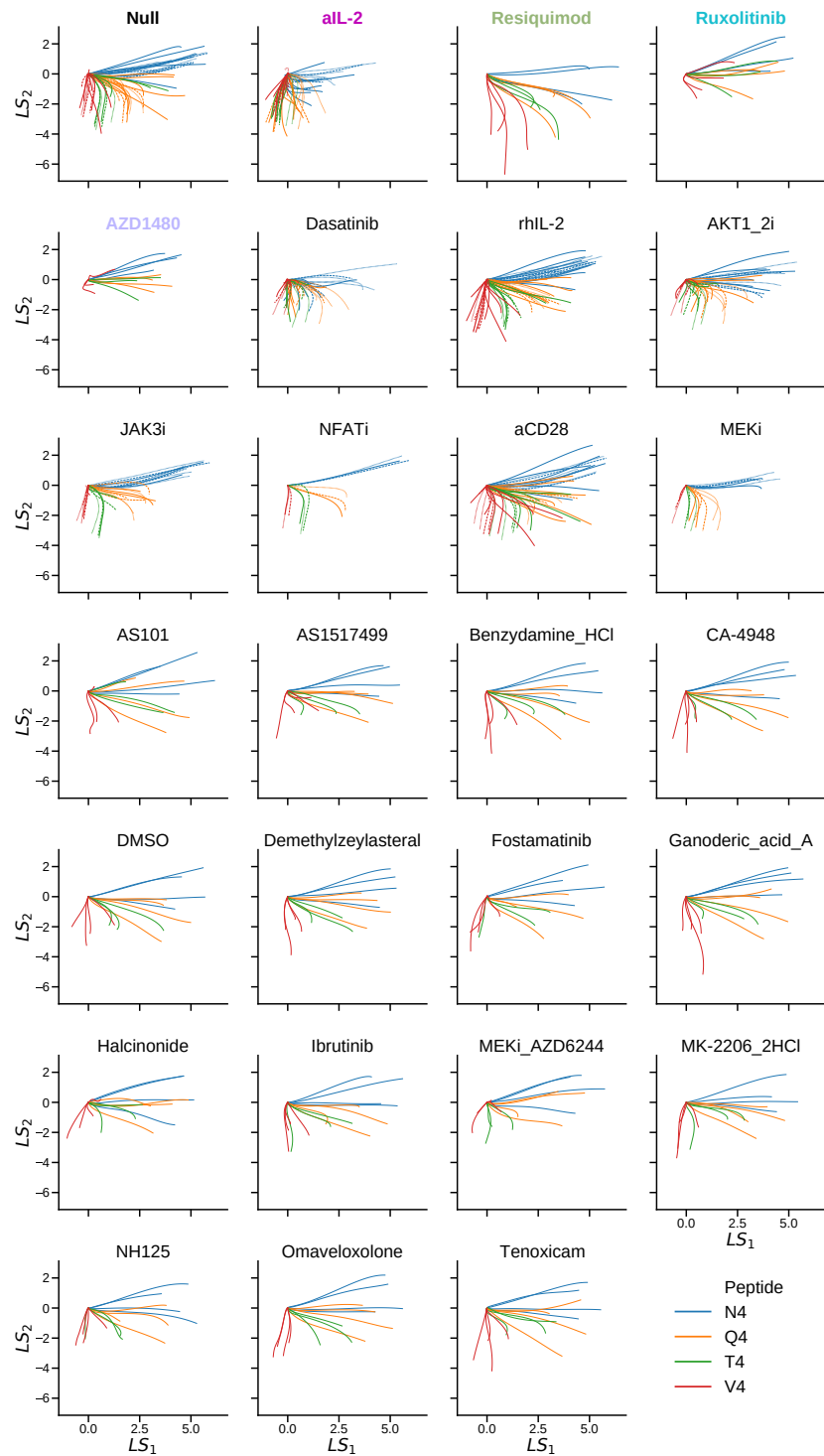


Figure 2.31: “Latent space projections (antigen encoding) for the cytokine dynamics under varied molecular perturbations. The data used in this figure are from eight independent experiments.”

To compute the EMD along each principal component (PC) direction, we first projected the initial and final distributions along the desired PC, then computed EMD between these 1D marginal distributions.

Antigen encoding with chimeric antigen receptor T cells

Given the robustness of our antigen encoding latent space, we wanted to extend its applicability to immune settings closer to clinically relevant immune therapies and to contemporary cellular engineering. We therefore

“investigated T cells genetically engineered to express a Chimeric Antigen Receptor (CAR) against ligands expressed on B cell leukemia (CD19), as used in patients with relapsed lymphoma [211]. The dynamics of cytokines released and consumed by CAR-T cells were similar to the trajectories derived from natural T cells (Fig. 2.28G-left). CAR-T cells responding to CD19 yielded cytokine dynamics similar to OT-1 TCR responding to weak agonists such as V4 (Fig. 2.28G-right), thus signaling through CAR may only sub-optimally activate T cells. We therefore tested the response of CAR-OT-1 T cells to a mixture of CAR and TCR ligands. Our latent space parametrization then revealed highly-nonlinear, hierarchical effects: weak OT-1 ligands (E1) antagonized CD19, whereas stronger OT-1 ligands dominated the response to CD19. This is consistent with the AKPR mechanism for T cell signaling with differential effects on positive and negative feedbacks (Fig. 2.12E) [62]. Complex antagonism and synergism by weak TCR ligands highlight the importance of fine tuning CAR-T cells for immunotherapeutic optimization.” (*Antigen encoding*, [1])

These preliminary observations, in the cytokine latent space, of non-linear effects of TCR/CAR antigen mixtures led us to investigate systematically cross-receptor interactions, as detailed in the next two chapters of this thesis.

2.9 Summary of key insights from antigen encoding

In summary, from the high-dimensional T cell response kinetics acquired by our collaborators with their robotic platform, we have derived a quantitative model of T cell antigen encoding (figure 2.1). Using a neural network classifier, we have discovered a two-dimensional latent space in which cytokine dynamics encode antigen quality (figure 2.4). This dimensionality reduction enabled us to

perform detailed analyses, such as deriving a mathematical model of the latent space trajectories and generating model-based cytokine time series from it. We have also rigorously quantified the amount of information transmitted by cytokine dynamics about antigen quality, finding that the continuum of T cell responses could be subdivided into $2^{2.6} = 6$ non-overlapping classes (figure 2.12). This showed how high-throughput experimental platforms and physics-inspired modelling might reveal underlying quantitative principles of immunity. Seeing how antigen encoding was robust across immune conditions (figure 2.28), we thought it could prove useful in analyzing the mechanisms of chemical sensing by T cell receptors, and ultimately to fine-tune T cell responses for immunotherapy and vaccine development [175].

We carried three main insights from antigen encoding into more mechanistic studies of TCR activation. First, T cell activation dynamics over days is low-dimensional, so we may not need to model T cell outputs as a high-dimensional set of variables; coarse-grained models may suffice. Second, the long time scales of T cell activation are essentially determined by the early activation kinetics (*i.e.*, by the v_0 or a_0 parameter); hence, models of these early events could be predictive of slower processes as well. Third, antigen quality is a continuous, quantitative property that elicits a corresponding continuum of T cell responses. Taken together, these points suggest that models of TCR-antigen interactions occurring within minutes, as introduced in chapter 1, might be used to quantitatively predict T cell activation mechanisms, even those measured on longer time scales. In the next two chapters, we have used exactly this approach to understand non-linear receptor interactions in T cells with altered TCR phosphorylation sites (chapter 3) and with synthetic receptors for immunotherapy (chapter 4).

Chapter 3

Revisiting antagonism in T cell receptors

In the previous chapter, we have essentially mapped T cell responses to a single antigen at a time, and we have treated the effect of antigen quantity as randomness in the response (*e.g.*, section 2.7). Here, we aimed to gain mechanistic understanding of TCR signaling by studying mixtures of antigens, which reveal nonlinear interactions between receptors [62]. We applied the insights gained from antigen encoding: since population-level cytokine dynamics correlate strongly with early activation kinetics, parsimonious models of T cell receptor signaling (time scale of minutes) can provide quantitative predictions of T cell responses. Aided by our collaborators' robotic platform, we systematically varied the qualities and quantities of antigens in the mixture, to better explore the high-dimensional chemical space faced by T cells in physiological settings. Moreover, T cell lines (6F-CD3 ζ TCR) from Paul Love's lab (NICHD), with altered TCR phosphorylation sites, gave us an additional experimental constraint to pinpoint receptor interactions. As we searched for a mathematical model that can account for these various phenotypes and antigen mixtures, we had to revise existing TCR activation models, which had previously been verified only for a small set of antigen mixtures.

This chapter contains my contributions to two manuscripts [2, 3] presented together here:

- Information theoretic analysis of 6F T cell responses (section 3.2.1);
- Adjusting the classical AKPR model to account for early datasets on 6F T cells (3.3 and 3.4);
- Finding this model needs to be improved to capture more extensive datasets, developing an revised model, and performing Monte Carlo simulations to estimate its parameters (3.5).

The 6F latent space analysis in section 3.2.2 was performed by Sooraj Achar using the antigen encoding pipeline [185] of section 2.2.2, which we built together with Thomas Rademaker. Experimental work was carried by collaborators, mainly Sooraj Achar and Guillaume Gaud. Quoted excerpts are from sections of the of the main text or supplementary information (SI) written by me.

3.1 Timeliness of revisiting TCR activation models

We have reviewed the mechanisms of TCR activation in chapter 1. Here, we bring out remaining questions about these mechanisms, especially when encountering antigen mixtures (as it is the case *in vivo*). We then describe the experiments performed by our collaborators to address them: the unprecedented level of robustness, accuracy, and throughput they achieve when they measure immune dynamics allowed us to push the envelope of phenomenological TCR models, seeking quantitative agreement between data and predictions to select the best model.

3.1.1 Nonlinear effects and antagonism in TCR signaling

Cells in a given tissue present on the order of 10^3 different peptide sequences on their surface MHCs, with different peptides being presented at anything between 10 and 10,000 copies per cell [212, 213, 214] (0.01% – 10% of 10^5 – 10^5 MHCs/cell). An example of the broad distribution of copy numbers for different peptides is shown in figure 3.1. Given the complexity of the TCR biochemistry and signaling pathways (section 1.3.2), it would have been most surprising that the various antigenic signals in a mixture contribute to T cell activation in a purely additive manner. Non-additive responses to mixtures of ligands are common in biology, for instance in olfactory receptor neurons [66], giving us another reason to expect them in T cells too.

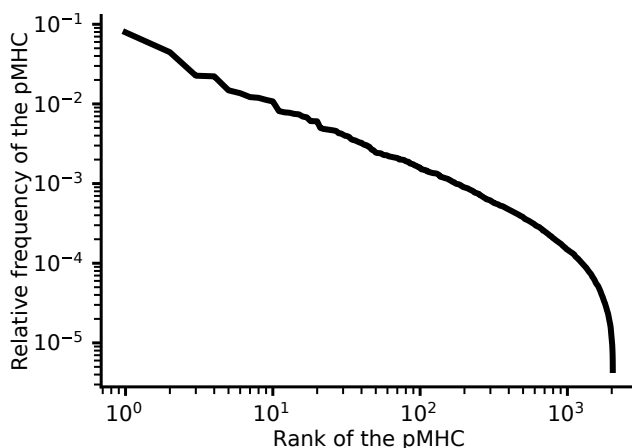


Figure 3.1: Relative frequency distribution of the class I peptide MHCs on the surface of antigen presenting cells (here, melanoma cells). Replotting the data for “background peptides” in the “DMSO” condition (*i.e.*, no treatment), averaged over three repeats, from figure 1D of [214], available publicly (Dataset_S02). We notice a Zipf-like behavior for the first 10^3 most abundant peptides, followed by a sharp cutoff. However, the abundances are approximative, based on mass spectrometry abundances; obtaining absolute cell surface abundance requires building a different calibration curve for each peptide [213], to compensate for, *e.g.*, unequal processing efficiency in the mass spectrometry pipeline.

There are in fact constructive and destructive “interferences” between TCR antigens, the most striking case of which is *antagonism*: a reduction of the T cell response due to weakly stimulating antigens. Adding weak antigens, which produce some response on their own, to a small amount of strong antigens actually reduces the overall response compared to the response elicited by the same small amount of strong antigens alone [215, 216, 217]. Figure 3.2 illustrates this effect. In this situation, the strong antigen is called an agonist and the weaker antigen causing a response reduction is called an antagonist. Since very few (< 10) agonist peptide copies are needed to trigger an immune response [110], antagonism of small agonist quantities may play a significant role in allowing or delaying T cell responses *in vivo*.

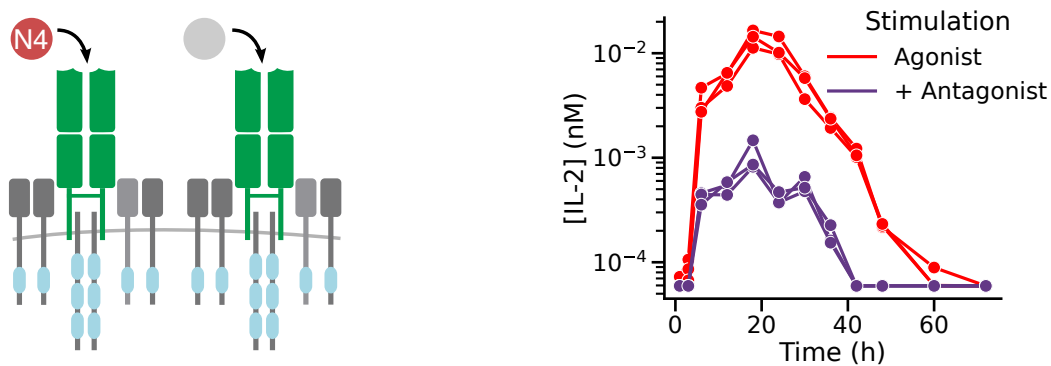


Figure 3.2: Example of TCR antagonism. (Left) TCR antagonism occurs when a T cell encounter a mixture of strong ligands (red) and weak ligands (grey) which can both bind to TCRs. (Right) IL-2 cytokine response over time of T cells (OT-1) exposed either to a strong antigen alone in a small dose (N4 at 10pM, red) or to the same amount of strong antigen plus a weak antigen (V4 at 1 μ M). The response is reduced by the addition of the antagonist peptide. This antagonist peptide, on its own and at a high dose, would produce a small amount of cytokine response as well (not shown) – it is not a completely null antigen. Hence, the response to the mixture is not at all equal to the sum of responses to each individual component. Data measured by the Altan-Bonnet lab using the Immunotron platform [2]. (Left is adapted from *TCR/CAR antagonism*, [3], figure 2; Right is adapted from *CD3 ζ ITAMs*, [2])

Antagonism was observed in T cells over 50 years ago, but its explanation has long been elusive. In the 1960s, experiments involving the sequential injection *in vivo* of a strong antigen, then a weak antigen revealed antagonism (then called “antigen competition”), but the effect was attributed primarily to “humoral” factors [218], *i.e.*, global feedbacks in the blood and lymph, such as cytokines. It was later realized – with the advent of molecular cell biology and discovery of the TCR – that cellular factors were probably more important, since mixtures delivered simultaneously still exhibited antagonism. The next proposed explanation was competition between agonist and antagonist peptides for presentation on MHCs [219]. This hypothesis was also ruled out when

antagonism was found to be TCR clone-dependent – all T cells would be antagonized if the effect came from hindered presentation of agonist peptides, but some TCRs were unaffected by the antagonists [220, 221, 222]. The next proposed explanation was rather competition for the binding of TCRs (antagonist peptides taking up all receptors) [223], but this was also ruled out by experiments showing antagonism in T cells engineered to express two TCRs, one dedicated to the agonist and the other, to the antagonist peptides, such that there was no possible competition for receptor binding [224]. Our work on antagonism in chimeric antigen receptor T cells, in the next chapter, also reveals antagonism between completely separate receptor-ligand pairs. Therefore, TCR antagonism most likely comes from intracellular receptor signaling mechanisms, where antigen mixtures change the balance between positive and negative effects triggered by TCR stimulation of different strengths [138, 116, 225].

3.1.2 Quantitative questions regarding TCR antagonism

Antagonism thus seems to occur due to inhibitory mechanisms in TCR signaling which prevent undesirable response to weakly stimulating antigens [59]: the weak antigens trigger inhibitory molecules, thus reducing the response to strong antigens, without producing enough activation on their own to compensate. In particular, the phosphatase SHP-1¹, which can dephosphorylate TCRs and prevent their signaling [226], is thought to be responsible for mediating this inhibitory feedback [138, 116, 109]. However, other inhibitory molecules are also potential candidates, such as the CD45 phosphatase [134], or the c-SRC kinase (Csk), which can restrain the activity of the Lck kinase (necessary for TCR signaling) [227]. Decisive verification of the exact molecular underpinnings of this inhibitory feedback is however difficult experimentally, due to its delicate and dynamical nature (immunoprecipitation only gives static, bulk snapshots), and because antibodies to distinguish the various phosphorylated versions of, *e.g.*, SHP-1, are lacking or unreliable (Grégoire Altan-Bonnet, private communication). Beyond any specific mechanism, François *et al.* proved mathematically that any TCR model performing absolute ligand quality recognition, independent of ligand quantity, must exhibit antagonism as a by-product [139].

Therefore, in our work, we adopted a phenomenological modelling approach where we postulate the existence of an inhibitory species, I , without specifying its precise molecular identity; it could in fact be an aggregate of several inhibitory pathways (such as SHP-1, CD45, Csk, etc.). This approach lowers the dimensionality of T cell dynamical models and enables us to address the more relevant question of how TCR antagonism is quantitatively determined by the qualities τ_l (affin-

¹SHP-1 stands for Src homology region 2 domain-containing phosphatase-1 and is also known as PTPN6, standing for Tyrosine-protein phosphatase non-receptor type 6.

ity, binding time) and quantities L_l (density, abundance) of antigens (indexed by l) in a mixture. Previous modelling studies have been limited to qualitative comparisons between predictions and data for a few different antagonist peptides presented at a high density; they could not make fine distinctions between different regulatory networks and activation functions in the inhibitory feed-back module surrounding species I [129, 132, 62]. Here, we attempted to go one step further and to compare alternative models on the basis of how accurately they match experimental patterns of antagonism.

Antagonism has generally been characterized in binary antigen mixtures [116], *i.e.*, in mixtures of two ligands with binding times τ_1, τ_2 and abundances L_1, L_2 , respectively. Two-component mixtures are interesting because most of the $\sim 10^3$ distinct peptides on an APC will have a very low affinity for the TCR; they effectively act like a background of weak peptides, reasonably well approximated by a large amount of one low-affinity peptide sequence [59]. The remaining one (or few) peptide sequences on the APC which might be cognate to a TCR can similarly be approximated by one high-affinity antigen type. Hence, receptor interactions in binary antigen mixtures give a reasonably good first approximation of the non-linear effects going on in actual physiological antigen mixtures. This is why we have also studied binary mixtures in this thesis: we have sought models able to quantitatively predict TCR antagonism for any combination of antigen mixture parameters τ_1, τ_2, L_1, L_2 , across different T cell types.

3.1.3 The role of ITAM multiplicity in TCR antigen recognition

Another open question regarding T cell activation is the role of immunoreceptor tyrosine-based activation motifs (ITAMs), which are pairs of tyrosine phosphorylation sites on the intracellular chains of the TCR [2]. As depicted by pale blue rectangles on the two TCRs of figure 3.2, one TCR has ten ITAMs (thus twenty phosphorylation sites): three on each of its two CD3 ζ chain, and one on each CD3 γ , CD3 δ , and CD3 ϵ (two copies of ϵ). The functions played by this ITAM multiplicity, unique among immune receptors, appear paradoxical. On the one hand, phosphorylation of both tyrosines of an ITAM leads to recruitment of ZAP-70, a major step towards activation of downstream TCR signaling pathways [131] (by, *e.g.*, the Lck kinase). Having multiple ITAMs per TCR may serve to amplify T cell activation signals [228, 229, 230], and T cell signaling cannot happen without at least one ITAM, yet no ITAM in particular is essential [231]. On the other hand, weak TCR stimulation leads to different states of ITAM phosphorylation compared to agonist stimulation [232, 233]. Furthermore, single-phosphorylated ITAMs seem to inhibit T cell signaling and mediate TCR antagonism [234, 235, 236], for instance by activating the SHP-1 phosphatase [138], although not all studies report this inhibitory role [237, 238].

Overall, the conflicting evidence surrounding TCR ITAMs may stem from the fact that this multiplicity supports parallel positive and inhibitory TCR signals, and that such a delicate balance evades qualitative, all-or-nothing functional descriptions, instead requiring quantitative modelling to be accounted for properly. A quantitative approach could reconcile apparently contradictory studies by avoiding all-or-nothing claims about ITAM functions². Low-dimensional, phenomenological models of T cell receptor activation are most desirable for this task, since they are more easily interpretable.

3.1.4 CD3-6F TCRs from the Love lab

To elucidate the role of CD3 ζ ITAMs, and to provide an additional phenotype in which to probe TCR antagonism quantitatively, the lab of Paul Love (NICHD) has developed OT-1 mouse lines in which the three CD3 ζ ITAMs of each ζ chain are inactivated (figure 3.3). The tyrosines (Y) in all 6 phosphorylation sites of the 3 ITAMs are substituted by a phenylalanine amino acid (F). This mutated phenotype is therefore called 6F (indicating the six $Y \rightarrow F$ substitutions), while the TCRs with the original CD3 ζ phenotype are called 6Y. The $Y \rightarrow F$ substitutions prevent ITAM phosphorylation without significantly changing the charge or structural properties of the ζ chain. The functions of CD3 ζ ITAMs can likely be inferred from the impact of their inactivation in 6F TCRs.

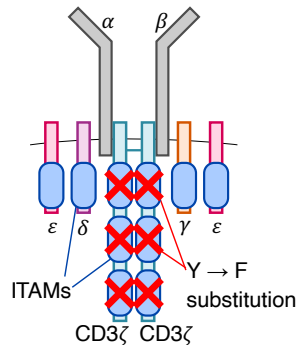


Figure 3.3: Illustration of ITAM mutations in the 6F TCR. The six red crosses indicate alteration of the six ζ ITAMs by tyrosine (Y) to phenylalanine (F) substitutions ($Y \rightarrow F$) in the two phosphorylation sites of each ITAM.

Other labs have previously used ITAM mutation or deletion in other TCR-monogenic mouse lines (e.g., P14 [239], H-Y [240], 3A9 [237]). These were however germline mutants, in which the mice

²For instance, reference [239] claims that TCR antagonism occurs “irrespective of the presence of ζ ITAMs” because they still observe *some* antagonism in CD3 ζ -deficient TCRs, while their own data in figure 6B shows that the *amplitude* of antagonism is reduced, with mutated ITAMs producing only a $\sim 2\times$ antagonism effect, compared to a $\sim 5\times$ with wild-type ITAMs.

express the altered ITAMs as soon as they are born, and go through thymic selection (growth and selection of T cells in the thymus after birth) with them. Their T cells may therefore be selected with compensatory mechanisms for the inactive or missing ζ ITAMs [241]. To circumvent this problem, the Love lab has three OT-1 mouse lines (in addition to germline 6F) where the 6F mutation is induced only after thymic selection via a Cre-Lox recombination system³. These three T cell lines are referred to as 6F(i), because they are “inducible” 6F, and they turned out to all have similar activation properties (and to the germline 6F as well), so we will not distinguish between them for modelling purposes. The 6F(i) TCR allowed us to disentangle the impact of CD3 ζ ITAMs on T cell activation and inhibition, and to probe new regimes of TCR/TCR antagonism to further decide between competing low-dimensional models of T cell activation.

3.1.5 Experimental measurements of TCR responses to antigen mixtures

Our collaborators in the Altan-Bonnet lab characterized the immune response of 6F (4-ITAM) TCRs using their robotic Immunotron platform, introduced in section 2.2.1. They compared 6F and 6Y (regular, 10-ITAM TCRs) T cell responses first to single antigens of varying strengths, and to several mixtures of two antigens. We applied our antigen encoding pipeline to process the experimental data (section 3.2), but after verifying that antigen encoding still applied to 6F T cells, we found that simply taking the average of cytokines over time proved a reasonable proxy of T cell activation levels to be compared against our models of TCR activation.

In our first study (*CD3 ζ ITAMs*, [2]), we initially tested only a high concentration of various antagonist peptides, which allowed us to develop a slightly modified version of the classical adaptive kinetic proofreading (AKPR) model in François *et al.*, 2013 [129] (section 3.3). Later on, we tested smaller concentrations of antagonists peptides. The results led us to develop a revised AKPR model, introduced in our second study (*TCR/CAR antagonism*, [3]) and presented here in section 3.5. Hence, our quantitative framework of antigen encoding, developed for single antigen responses in the previous chapter, provided the necessary insights to develop parsimonious, low-

³Cre-Lox recombination is used to replace a target gene sequence by another upon activation of the system. The Cre recombinase is an enzyme which can cut out DNA at specific sites – called Lox sequences – to insert another genetic sequence instead. For the activation of some gene or promoter to trigger Cre-Lox recombination, the gene coding for Cre is added downstream of that desired trigger promoter or gene [242]. The three Cre-Lox mouse lines from the Love lab have their genotype switch induced by administration of the drug tamoxifen, by activation of the distal Lck promoter, or by expression of the gene OX40, respectively. The dLck promoter is only active in post-selection CD4⁺ and CD8⁺ T cells, while the OX40 gene is only expressed upon T cell activation. The dLck and tamoxifen-induced 6F are used for *in vitro* experiments, to ensure all cells already have the 6F CD3 ζ at the start of the experiment; the OX40 line is used for longer *in vivo* experiments, where having the 6Y to 6F switch occur during the first few hours of the reaction does not matter too much.

dimensional models of TCR antigen recognition kinetics that could account for longer time scales of immune responses.

3.2 Antigen encoding by altered TCRs

We started by applying the latent space analysis and antigen encoding framework to 6F T cells with altered CD3 ζ ITAMs, to gain insights that we could then integrate into more mechanistic (yet still low-dimensional) models of TCR antigen recognition.

3.2.1 Information theoretic analysis

The cytokine time series data acquired by our collaborators revealed, surprisingly, that 6F(i) T cells displayed enhanced sensitivity to weak antigens (*e.g.*, G4, V4, T4), yet slightly reduced responses to strong antigens (A2, N4). This is illustrated in figure 3.4a-b, where we used the cytokine concentrations averaged over 72 hours as a readout of T cell activation. This data suggests that the six CD3 ζ ITAMs improve the TCR ability to distinguish between weak antigens – whereas 6F(i) T cells have a nearly binary response with just over 1 bit of mutual information – and that these ITAMs can also produce a small additional output upon strong antigen stimulation. Overall, the presence of additional ITAMs seemed to improve TCR antigen recognition.

To make this point more quantitative, we computed the mutual information (MI) between antigen quality Q and each cytokine produced by T cells (figure 3.4), using the MI estimator introduced in section 2.2.4. The error bars represent the statistical error on the MI estimator, computed by bootstrapping over the 12 data points available for each genotype and peptide (3 biological replicates with 4 technical replicates in each). Despite the small sample sizes, we found a statistically significant decrease in MI in 6F(i) T cells compared to 6Y in cytokines IFN- γ and TNF. The mutual information difference was more pronounced in germline, pre-activated 6F T cells, as shown in figure 3.5. The difference between 6Y and 6F was also present when we computed MI between Q and all cytokines taken together (figure 3.5b).

This information theoretic analysis of 6F T cell responses suggests that multiple ITAMs enable additional kinetic proofreading of TCR ligand binding. This was the first valuable insight from our antigen encoding analysis of 6F T cells: in AKPR models of TCR recognition, we should decrease the number of kinetic proofreading steps, N , when the number of ITAMs is decreased.

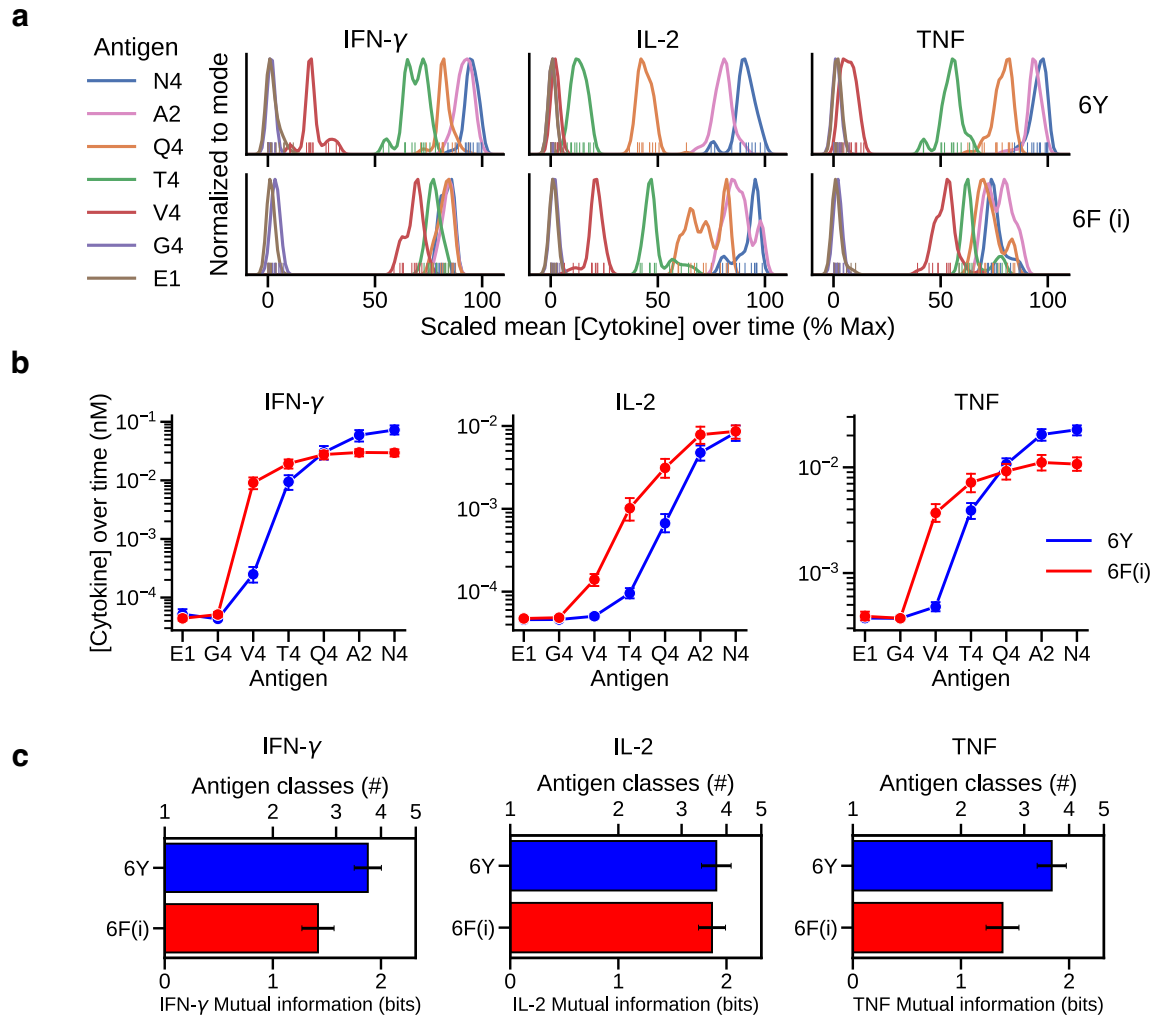


Figure 3.4: Mutual information analysis of naive 6F(i) T cells. “a, Averaged distribution of normalized cytokine secretion of 4-OH tamoxifen treated naive *OT1-6Y* or *OT1-6F(i)* (*Ert2-Cre*) $CD8^+$ T cells stimulated with APC + 1 μ M antigen of varying affinities. $n = 3$ biological replicates. The data shown are representative of two independent experiments. b, Plots of cytokine secretion for IFN γ , IL-2 or TNF averaged over the 72-h time course. $n = 3$ biological replicates. Data are shown as the mean $\pm$ s.d. and are representative of at least three independent experiments. c, Mutual information (antigen classes) between antigen quality and each secreted cytokine for each genotype. Data are expressed as the MI estimator $\pm$ s.d. $n = 3$ biological replicates representative of at least three independent experiments.” (*CD3 ζ ITAMs*, [2], adapted from figure 4)

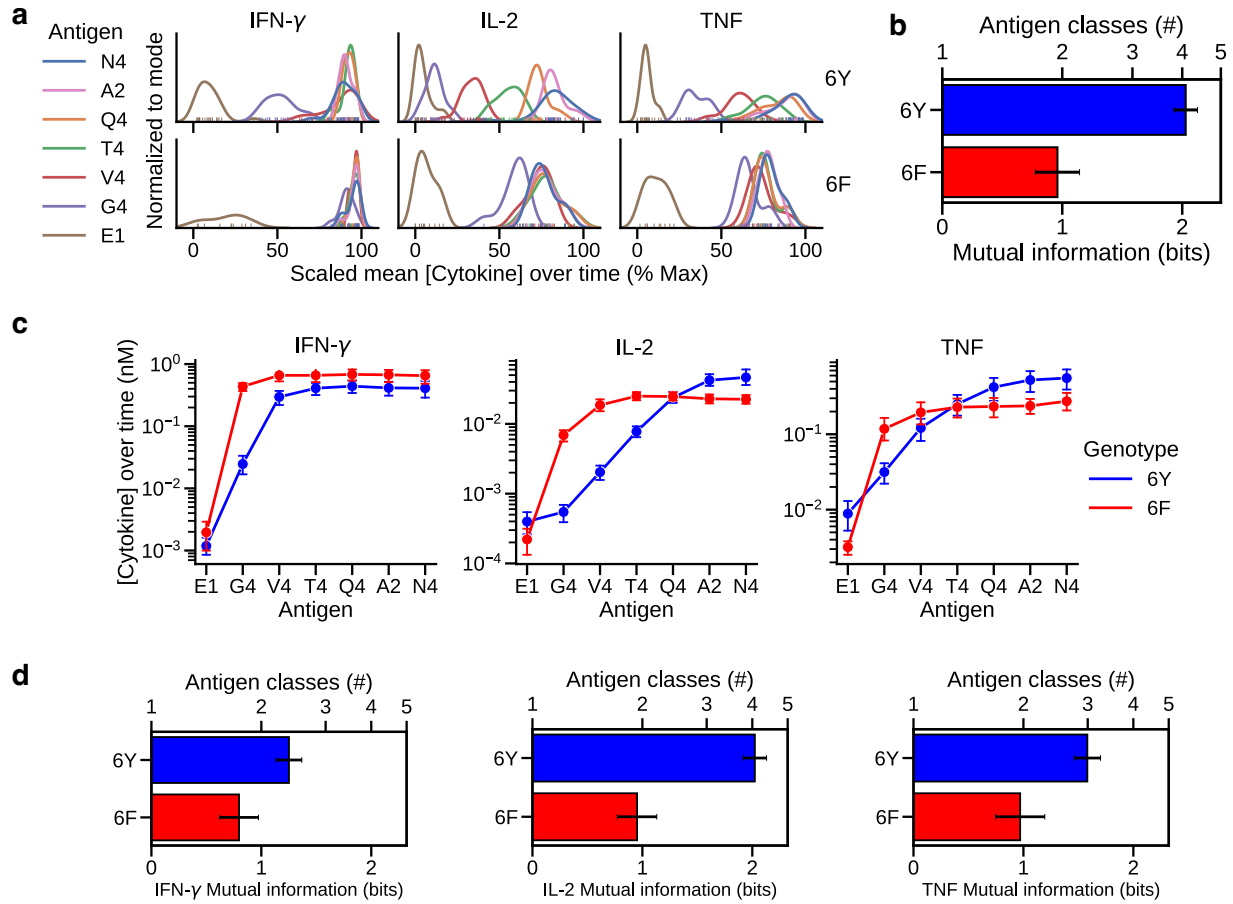


Figure 3.5: Mutual information analysis of blast, germline 6F T cells. “(a) Distribution of cytokine secretion levels of effector *OT1-6Y* or *OT1-6F* (germline) $CD8^+$ T cells pre-stimulated with 10^{-6} M N4 antigen + APCs for 6 days then re-stimulated with APC + 10^{-6} M antigens of varying affinities. (b) Mutual information (antigen classes) between antigen quality and all secreted cytokines for each genotype. While *OT1-6Y* T cells can distinguish more than four classes of antigen across the 6 peptide affinities tested, *OT1-6F* T cells only distinguish two. Data are expressed as MI estimator $\pm$ SD. (c) Summary plots of cytokine secretion showing scaled mean over time. Data are represented as mean $\pm$ SD. (d) Mutual information (antigen classes) between antigen quality and each secreted cytokine for each genotype. Data are expressed as MI estimator $\pm$ SD.” For MI estimation, we had $n = 20$ data points (technical or biological replicates) for 6Y T cells and $n = 10$ data points for 6F, because some experiments only had 6Y (10-ITAM) T cells. (*CD3 ζ ITAMs*, [2], adapted from Extended Data Figure 5)

3.2.2 Latent space analysis

Moreover, we noticed in figure 3.4 that all cytokines showed similar patterns as a function of antigen quality, suggesting that 6F T cell responses still lie on a low-dimensional manifold, just like we discovered for 6Y T cells in the previous chapter. We therefore applied our preprocessing and latent space projection method to the cytokine time series collected for 6F(i) T cells (as well as for control 6Y T cells). This analysis was performed by Sooraj Achar using the antigen encoding pipeline developed jointly with Thomas Rademaker and myself [185]; we reproduce the results in figure 3.6, because they provide important hints about CD3 ζ ITAM mechanisms.

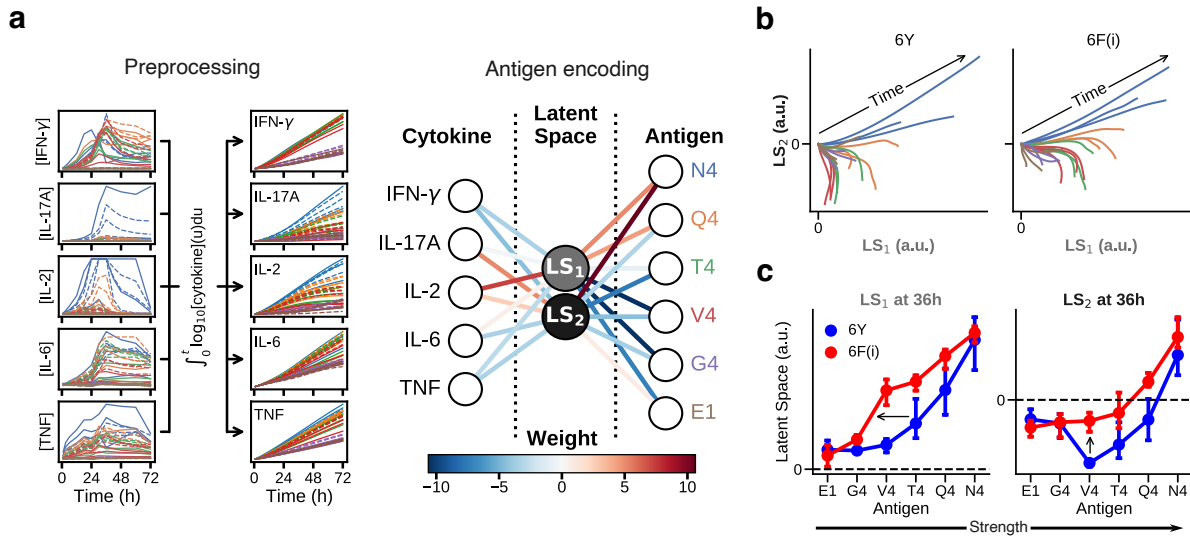


Figure 3.6: Latent space analysis of 6F T cells. “a, Cytokine secretion dynamics of naive *OT1* T cells in response to antigens of six different affinities were preprocessed (left) and then projected into a compressed two-dimensional space (LS_1, LS_2) through the application of the weight matrix obtained from a three-layer neural network trained to predict antigen quality (right) ($n = 3$ biological replicates). b, Time trajectories of LS_1 and LS_2 for each antigen in 4-OH tamoxifen treated *OT1-6Y* and *OT1-6F(i)* (*Ert2-Cre*) $CD8^+$ T cells. c, LS_1 and LS_2 values at 36 h as a function of antigen affinity. $n = 3$ biological replicates. Data are shown as the mean $\pm$ s.d. and are representative of at least three independent experiments.” (*CD3 ζ ITAMs*, [2], adapted from figure 4)

After preprocessing cytokine time series and projecting them into the latent space (figure 3.6a, using the pre-trained neural network weights of chapter 2), we observed similar latent space trajectories for 6Y (10-ITAM) and 6F(i) (4-ITAM) T cells, although with different initial angles θ for antigens of intermediate strength (figure 3.6b). We inspected more closely how LS_1 and LS_2 at a given time (here, $t = 36$ h) varied as a function of antigen quality in both TCR types (figure 3.6c). The LS_1 dimension reflected the pattern seen in individual cytokines (figures 3.4b and 3.5c), with

an increased response to weak antigens in 4-ITAM T cells. The most interesting change occurred in LS_2 : whereas it had a non-monotonic behaviour as a function of antigen quality in 10-ITAM T cells, it became monotonically increasing in 4-ITAM T cells. As mentioned in section 2.7.7, the “bell-shaped” LS_2 curve could reflect a strategy by T cells to encode information about antigen quality in a format easier to decode by other cells. The coincidence of the loss of non-monotonicity in LS_2 with a loss of information content in cytokines (figures 3.4 and 3.5) further supports the idea that CD3 ζ ITAMs help T cells encode finer distinctions between antigens in their cytokine response.

The observed (LS_1, LS_2) pattern shifts in 4-ITAM T cells further hinted at possible molecular correlates of the latent space variables. Since the increase in cytokine responses to weak antigens in 6F T cells directly translated into an increase in LS_1 , this variable most likely encodes the positive signals triggered by TCR stimulation: more activation implies larger LS_1 . Because of its non-monotonic behaviour in 6Y (10-ITAM) TCRs, LS_2 might reflect inhibitory feedbacks, such as SHP-1 activation, in TCR activation: these inhibitory signals dominate against weak antigens, for which LS_2 becomes negative in 6Y T cells. The loss of this non-monotonic LS_2 in 6F (4-ITAM) T cells would indicate a reduction in the inhibitory feedback against weak antigens. Therefore, CD3 ζ ITAMs potentially play an inhibitory function in TCR signaling, which is partially lost when these ITAMs are disabled in 6F T cells.

Our collaborators tested the hypothesis that CD3 ζ ITAMs play an inhibitory role in TCR signaling. They compared recruitment of the SHP-1 phosphatase by 6Y and 6F TCRs. Confirming our prediction, they found that 6Y TCRs activated more SHP-1, due to their 10 intact ITAMs, than 6F TCRs, with only 4 ITAMs left to recruit SHP-1 [2, figure 5]. Hence, ITAMs have a previously overlooked inhibitory function, via activation of enzymes to slow down TCR signaling. This explains their importance in lowering T cell activation due to weak antigens.

Thus, we obtained a second important insight from antigen encoding: the CD3 ζ ITAMs, in addition to contributing to kinetic proofreading, seem to support the negative feedback implemented in adaptive KPR models [116, 129, 132], which inhibits responses to weak antigens. Hence, 6F (4-ITAM) TCRs should be modelled as having a lower activation rate of the inhibitory feedback, and maybe a reduced sensitivity to the inhibitory molecule I as well – for instance because SHP-1 has fewer ITAMs on which to act.

An important prediction follows from postulating that CD3 ζ ITAMs have an inhibitory function: since 6F TCRs produce less inhibitory feedback and are less sensitive to it, they should also exhibit

reduced antagonism. We therefore set out to make this prediction quantitative by developing a low-dimensional yet predictive model of TCR activation kinetics, implementing the insights gained from antigen encoding in 6F T cells. Incidentally, having experimental data for two TCR types – 10-ITAM 6Y and 4-ITAM 6F – allowed us to better constrain our model and our understanding of how T cells parse antigen mixtures.

In section 3.3, we formally introduce our modelling framework for TCR signaling dynamics. We explain in particular (section 3.4) how we tuned the classical AKPR model for 6F T cells based on the two insights just discussed, yielding promising early predictions. Then, we explain how the classical model could not stand the test of more systematic experimental characterization, and how we ultimately revised the AKPR model, by starting from an even more parsimonious model (section 3.5).

3.3 Revisiting the classical AKPR model

To disentangle T cell responses to antigen mixtures, we used mathematical models of the early biochemical events occurring when T cell receptors are put in contact with pMHCs on the surface of antigen presenting cells. The models are all based on the idea, introduced in chapter 1, that TCRs perform kinetic proofreading (KPR) of ligand binding times [128]. Although these binding, unbinding, and phosphorylation reactions equilibrate on a time scale of minutes [116], we assumed they would be predictive of T cell activation levels experimentally measured over days in our collaborators’ Immunotron robotic platform. This assumption is based on the insight, from our antigen encoding theory (chapter 2), that cytokine dynamics upon T cell activation are determined by early antigen recognition kinetics.

In this section, we introduce in detail the adaptive KPR (AKPR) modelling framework, and explain how we extended the classical AKPR model (François *et al.*, 2013) to account for antagonism in T cells with altered ITAMs. In the next section, we explain how we directly fitted TCR model outputs to experimental data (by computing dimensionless ratios), and how we needed to propose a revised AKPR model, because the classical model failed to provide a quantitative match.

3.3.1 Modelling framework

“All models belong to the adaptive kinetic proofreading (AKPR) framework (see François and Altan-Bonnet [62], Lalanne and François [132], and François *et al.* [129]). In brief, this class of models explains the speed, sensitivity, and specificity

of immune ligand recognition, achieving so-called “absolute discrimination”, *i.e.*, discrimination between different antigenicities (ligand quality, characterized by ligand binding time τ) irrespective of ligand numbers, L .

All AKPR models comprise two components:

1. A kinetic proofreading (KPR) cascade, triggered by the formation of ligand-receptor complexes at the T cell–antigen presenting cell (APC) interface, akin to the initial KPR model proposed for T cell receptors [128]. In this cascade, a ligand-receptor complex goes through sequential activation steps (*e.g.*, phosphorylation of associated chains, recruitment of ZAP-70 and LAT). The variable C_n represents the number of complexes at step n . It is generally assumed that the last complex of the cascade, C_N , is responsible for the response. When the ligand dissociates from the receptor, it is assumed that all complexes quickly revert back to the first step of the cascade, *i.e.*, to the free, unmodified receptor. The KPR cascade is responsible for the amplification of the relative response between ligands of different qualities. However, on its own, it fails to discriminate between a high concentration of low quality ligands (large L , small τ) and a very low concentration of very antigenic ligands (small L , large τ) as explained in Altan-Bonnet and Germain [116] and Lalanne and François [59].
2. A negative feedforward or feedback interaction, slowing down the proofreading steps in the KPR cascade [116, 129, 225]. In its simplest form, the negative interaction is turned on by a given complex C_m in the KPR cascade. Depending on the model, this negative interaction can act either on the forward or backward rates of the KPR cascade [132]. This negative interaction is responsible for the adaptive part of the AKPR model, ensuring that a vast quantity of low quality ligands gives a weaker response than a small amount of high quality ligands.

It was proven mathematically [139] that all models achieving absolute discrimination must exhibit antagonism, defined as the reduction of the overall response to a mixture of ligands compared to the response when only the strongest ligands of the mixture (the “agonists”) are presented (figure 4.2). This is typically due to the fact that the less potent ligands (the “antagonists”) increase the relative intensity of the negative interactions, so that the total signal is lowered. Importantly, however, quantitative aspects of T cell antagonism (amplitude, dependence on ligand concentration, etc.)

depend on the details of the negative interactions and thereby allow us to distinguish between models.

In the following, we detail how we built upon the initial AKPR model (from François *et al.* [129]), and how we needed to modify the negative interactions to account for all experimental data collected in this study. Throughout, we modelled signaling networks in the deterministic limit as a system of ordinary differential equations (ODEs) derived assuming mass-action or Michaelis-Menten kinetics. The KPR cascade was modelled in a similar way in all versions. The full ODE systems for all models are given [...], and we used them to derive the steady-state solutions discussed below. We assumed these steady-state solutions described the average T cell population-level readouts of signals.” (*TCR/CAR antagonism*, [3], SI)

We adopt a notation where superscripts indicate receptor type and subscripts specify the meaning of a variable, *e.g.*, by indicating the proofreading step or the ligand type. For example, $C_{m,l}^\rho$ is the number of receptors of type $\rho \in \{\text{TCR}, \text{CAR}\}$ at proofreading step m and bound to ligand type $l \in \{1, 2, \dots\}$. Receptor type indices will be relevant in chapter 4 discussing TCR interactions with chimeric antigen receptors (CAR).

3.3.2 Initial AKPR model for a single TCR antigen type

“The ODEs of the François *et al.*, 2013 model for mixtures of TCR antigens were formulated in François *et al.* [129]; we reproduce them here in the notation chosen for this work. We drop the $\rho = T$ superscript, which is implied for T cells without a CAR.

$$\begin{aligned} \frac{dC_{0,l}}{dt} &= \kappa(L_l - R_{b,l}) \left(R_{\text{tot}} - \sum_{l'=1}^2 R_{b,l'} \right) + (b + \gamma I)C_{1,l} - (1/\tau_l + \varphi)C_{0,l} \\ \frac{dC_{n,l}}{dt} &= \varphi C_{n-1,l} + (b + \gamma I)C_{n+1,l} - (\varphi + 1/\tau_l + b + \gamma I)C_{n,l} \quad (1 \leq n < N) \\ \frac{dC_{N,l}}{dt} &= \varphi C_{N,l} - (1/\tau_l + b + \gamma I)C_{N,l} \end{aligned} \tag{3.1}$$

$$\frac{dI}{dt} = \beta_I \left[\left(\frac{1}{C_{m,\text{th}}} \sum_{l'=1}^2 C_{m,l'} \right) (I_{\text{tot}} - I) - I \right] \tag{3.2}$$

where $l \in \{1, 2\}$ indicates the ligand type (agonist or antagonist) bound to these TCRs, and where $R_{b,l} = \sum_{n=0}^N C_{n,l}$. The rate β_I is assumed to be fast enough for $\vec{I}$ and the C_n to reach steady state within minutes and does not enter equations (3.4)–(3.7) for that steady state.” (*TCR/CAR antagonism*, [3], SI)

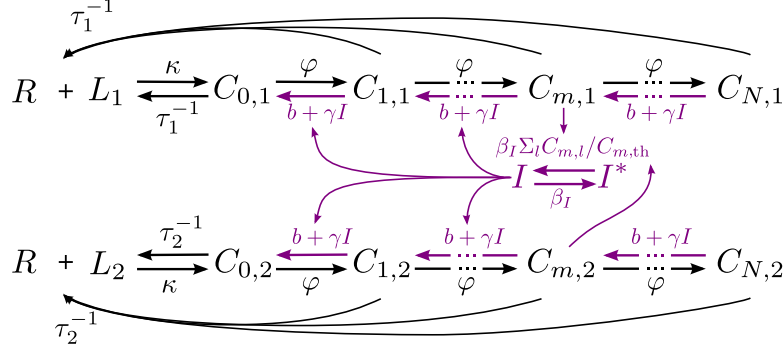


Figure 3.7: Reaction network structure of the classical AKPR model (François *et al.*, 2013) in the presence of agonist (τ_1, L_1) and antagonist (τ_2, L_2) antigens.

Figure 3.7 shows a diagram of the biochemical reaction network to which this ODE systems corresponds when there are two ligands. To summarize, the $C_{n,l}$ variables implement a kinetic proof-reading cascade, where we distinguish TCRs bound to either ligand type l , and the inhibitory feedback is implemented by a phosphatase I , which is activated by $C_{m,l}$ from its inactive form I^* out of a total pool of molecules $I_{\text{tot}} = I + I^*$.

We first consider this model in the case of a single ligand type and drop the index $l = 1$. For L ligands of binding time τ , the total number of bound receptors at steady-state is

“

$$R_b = \frac{1}{2} \left(R_{\text{tot}} + L + \frac{1}{\kappa\tau} \right) - \frac{1}{2} \sqrt{\left(R_{\text{tot}} + L + \frac{1}{\kappa\tau} \right)^2 - 4R_{\text{tot}}L} \quad (3.3)$$

and the rest of the solution for the steady-state is carried as in François *et al.* [129] (see table 3.1 for the definition of parameters). In particular, the steady-state number of complexes in proofreading step n is given by

$$C_n = R_b (a_- r_-^n + a_+ r_+^n) \quad (3.4)$$

where, for a given I , $r_{\pm}$ and $a_{\pm}$ are given by:

$$r_{\pm} = \frac{\varphi + b + \gamma I + 1/\tau}{2(b + \gamma I)} \pm \frac{\sqrt{(\varphi + b + \gamma I + 1/\tau)^2 - 4\varphi(b + \gamma I)}}{2(b + \gamma I)} \quad (3.5)$$

$$a_{\pm} = \frac{r_{\pm} - 1}{(r_{\pm}/r_{\mp})^{N+1} - 1} \quad (3.6)$$

Eqs. (3.3)–(3.6) correspond to a standard proofreading cascade with a forward rate φ and a backward rate $b + \gamma I$, where the inhibitory species I thus favors reverse reactions in the KPR cascade (see figure 3.12D). The crucial feature of the AKPR model is that I depends on a complex C_m in the cascade, so the steady-state value of I is the solution of the following implicit equation :

$$I = I_{tot} \frac{C_m}{C_{m,th} + C_m} \quad (3.7)$$

where C_m depends on I via eqs. (3.4)–(3.6), and $C_{m,th}$ is a fixed parameter setting the scale of I activation. Eq. (3.7) is solved numerically for I to close the system of equations." (*TCR/CAR antagonism*, [3], SI)

With the default parameter values used in [129], the classical AKPR model produces the steady-state response curves to a single antigen shown in figure 3.8. Increasing τ_1 , even just moderately, increases the output significantly, at all L_1 : this is kinetic proofreading. Moreover, as L_1 increases, C_N eventually stops increasing and remains flat as a function of L_1 (over three orders of magnitude for stronger antigens), due to the negative feedback implemented by I and activated by C_m . The combination of kinetic proofreading (scaling with τ_1) and adaptive feedback (independence of the output on L_1) makes it possible to imagine a response threshold on C_N (horizontal dashed line) such that intermediate antigens (*e.g.*, $\tau_1 = 3.5$ s) will never cross it, while slightly stronger antigens (*e.g.*, $\tau_1 = 5$ s) cross it at small antigen concentrations L_1 . This illustrates how AKPR models perform absolute ligand discrimination on the basis of binding time τ , with minimal effect from antigen concentration L .

3.3.3 Classical AKPR model for TCR/TCR antagonism

"This study focuses on antagonistic interactions between receptors on the surface of T cells. Hence, we revisited the initial François *et al.*, 2013 model on TCR/TCR antagonism data, to explore whether it could account for quantitative effects of varying antigen quality and density.

	Parameter	Definition
Input variables	τ_l	Binding time of ligands of type l
	L_l	Total number of ligands of type l (free or TCR-bound)
State variables	$C_{n,l}$	Number of TCRs in state n bound to ligand type l
	$R_{b,l}$	Total number of TCRs bound to ligand type l
	I	Number of active inhibitory molecules
Integer parameters	R_{tot}	Number of TCRs per T cell
	I_{tot}	Total number of inhibitory molecules (active or not)
	N	Number of proofreading steps for TCRs
	m	Proofreading state of TCRs activating the inhibitory module
Rate parameters	φ	Forward KPR rate
	κ	Binding rate of ligands (of any type l) to TCRs
	b	Basal reverse rate in the KPR cascade
	γ	Reverse KPR rate mediated by inhibitory molecules I
	$C_{m,\text{th}}$	Activation threshold of I by C_m (same for all types l)

Table 3.1: Definition of parameters in the classical AKPR model. Variables and parameters in the classical AKPR model, for TCR/TCR antagonism. Superscripts $\rho = T$, indicating the receptor type is TCR, are implicit. Note that $R_{b,l} = \sum_{n=0}^N C_{n,l}$, so it is not an independent state variable. 6Y: 10-ITAM TCR, 6F: 4-ITAM TCR (CD3 ζ ITAMs mutated).

For a mixture of two TCR antigens, we distinguish receptors bound to agonist ligands, denoted by variables $C_{n,1}$, and receptors bound to ligands of lower potency (*e.g.*, antagonists), denoted by $C_{n,2}$. Agonist and weaker ligands have binding times τ_1 and τ_2 , respectively, and quantities L_1 and L_2 .” (*TCR/CAR antagonism*, [3], SI)

We again are interested in the fixed point solution of the system of ODEs (3.2), now for $l \in \{1, 2\}$.

“The stationary solution is given by (3.4)–(3.6), with $\tau = \tau_1$ or τ_2 , respectively. However, the numbers of bound receptors to each ligand type, $R_{b,1}$ and $R_{b,2}$, are coupled via,

$$\begin{aligned}
0 &= \kappa\tau_1(L_1 - R_{b,1})(R_{\text{tot}} - R_{b,1} - R_{b,2}) - R_{b,1} \\
0 &= \kappa\tau_2(L_2 - R_{b,2})(R_{\text{tot}} - R_{b,1} - R_{b,2}) - R_{b,2} .
\end{aligned} \tag{3.8}$$

This system is solved exactly [...] by expressing $R_{b,2}$ in terms of $R_{b,1}$,

$$R_{b,2} = R_{\text{tot}} - R_{b,1} - \frac{R_{b,1}}{\kappa\tau_1(L_1 - R_{b,1})} \tag{3.9}$$

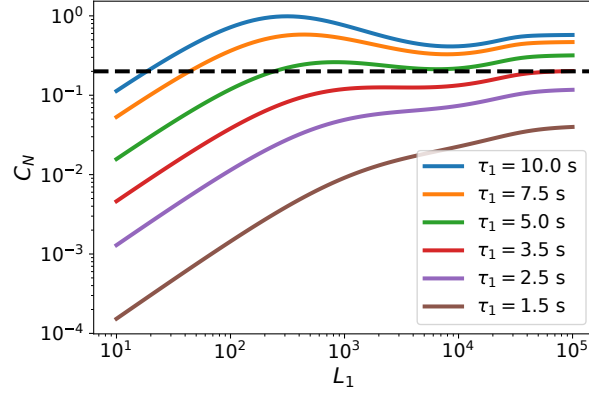


Figure 3.8: Response curves – steady-state output C_N as a function of antigen density L_1 – of the classical AKPR model [129], for various antigen qualities τ_1 . The horizontal dashed line is a hypothetical response threshold, to illustrate how intermediate antigens can be absolutely distinguished from stronger ones, at all L_1 . The steady-state solution plotted is exact and includes receptor saturation at large L_1 (contained in equation 3.3): the curves all become flat as L_1 becomes larger than the number of receptors R_{tot} .

and solving the resulting cubic equation for $R_{b,1}$,

$$p_0(R_{b,1})^3 + p_1(R_{b,1})^2 + p_2R_{b,1} + p_3 = 0 \quad (3.10)$$

where the coefficients are

$$\begin{aligned} p_0 &= \frac{\tau_1}{\tau_2} - 1 \\ p_1 &= -\left(\frac{\tau_1}{\tau_2} - 1\right) \left(R_{\text{tot}} + L_1 + \frac{1}{\kappa\tau_1}\right) - L_2 - \frac{\tau_1}{\tau_2}L_1 \\ p_2 &= \frac{\tau_1}{\tau_2}L_1^2 + \left(2\frac{\tau_1}{\tau_2} - 1\right) R_{\text{tot}}L_1 + \frac{L_1}{\kappa\tau_2} + L_1L_2 \\ p_3 &= -\frac{\tau_1}{\tau_2}R_{\text{tot}}L_1^2 \end{aligned} \quad (3.11)$$

The physically correct solution is the only root satisfying $0 \leq R_{b,1} < L_1$. It always exists and ensures $0 \leq R_{b,2} < L_2$ and $R_{b,1} + R_{b,2} < R_{\text{tot}}$ too.” (*TCR/CAR antagonism*, [3], SI)

We give a proof of these claims about the existence and uniqueness of the solution for $(R_{b,1}, R_{b,2})$ in appendix C. We derived this exact solution for $(R_{b,1}, R_{b,2})$ for the first time in [2], to avoid relying on the approximation that receptors are not saturated by ligands, which was used in previous papers

on AKPR models [129, 132] although it in fact breaks down at high ($1\mu\text{M}$) antagonist doses. Our exact solution solves this problem and remains valid in all regimes of ligand and receptor numbers.

“Lastly, the implicit solution for I in eq. (3.7) now contains a sum over ligand types

$$I = I_{tot} \frac{C_{m,1}^T + C_{m,2}^T}{C_{m,th}^T + C_{m,1}^T + C_{m,2}^T}. \quad (3.12)$$

” (*TCR/CAR antagonism*, [3], SI)

The equation for I is solved numerically, by treating $C_{m,l}$ as a function of I .

3.4 Initial AKPR model for TCRs with fewer ITAMs

In our first attempt to predict the response of 6Y and 6F TCRs to antigen mixtures, we slightly augmented the classical AKPR model to account for the effect of ITAM number, using the insights gained in section 3.2. Figure 3.9A illustrates the model variants thus defined for 6Y and 6F T cells. Then, we proceeded in two steps to extract predictions from the model. First, we tweaked model parameters and output definitions to fit the response of both TCR types to single antigens. Second, we generated model predictions for responses to antigen mixtures as a function of L_1, τ_1, τ_2 , which we compared qualitatively to experimental measurements.

3.4.1 Adapting the model to distinguish 6Y and 6F(*i*) TCRs

“We adapted the model to capture differences seen between 6Y and 6F TCRs in response to a single type of antigen at a time, such as shown in figure 3.9B. The main features we aimed to capture were that 6F(*i*) cells produce more output for weak antigens – the curve looks shifted to the left according to ligand quality– but that 6Y cells eventually catch up (for strong agonists) and produce a higher maximal response.

In our model, we hypothesized that differences in ITAM multiplicity correspond to different numbers of kinetic proofreading steps. We assigned fewer steps to 6F TCRs – which lack 6 ITAMs – than to 6Y: $N_{6Y} > N_{6F}$. Less proofreading may explain the increased response of 6F(*i*) T cells to weak agonists. Additionally, we made the last $k_{6Y} > 1$ complexes (from $C_{N_{6Y}-k+1}$ to $C_{N_{6Y}}$) contribute to the signaling output of 6Y TCRs, compared to only the last complex, $C_{N_{6F}}$, with 6F(*i*) TCRs (i.e., $k_{6F} = 1$).

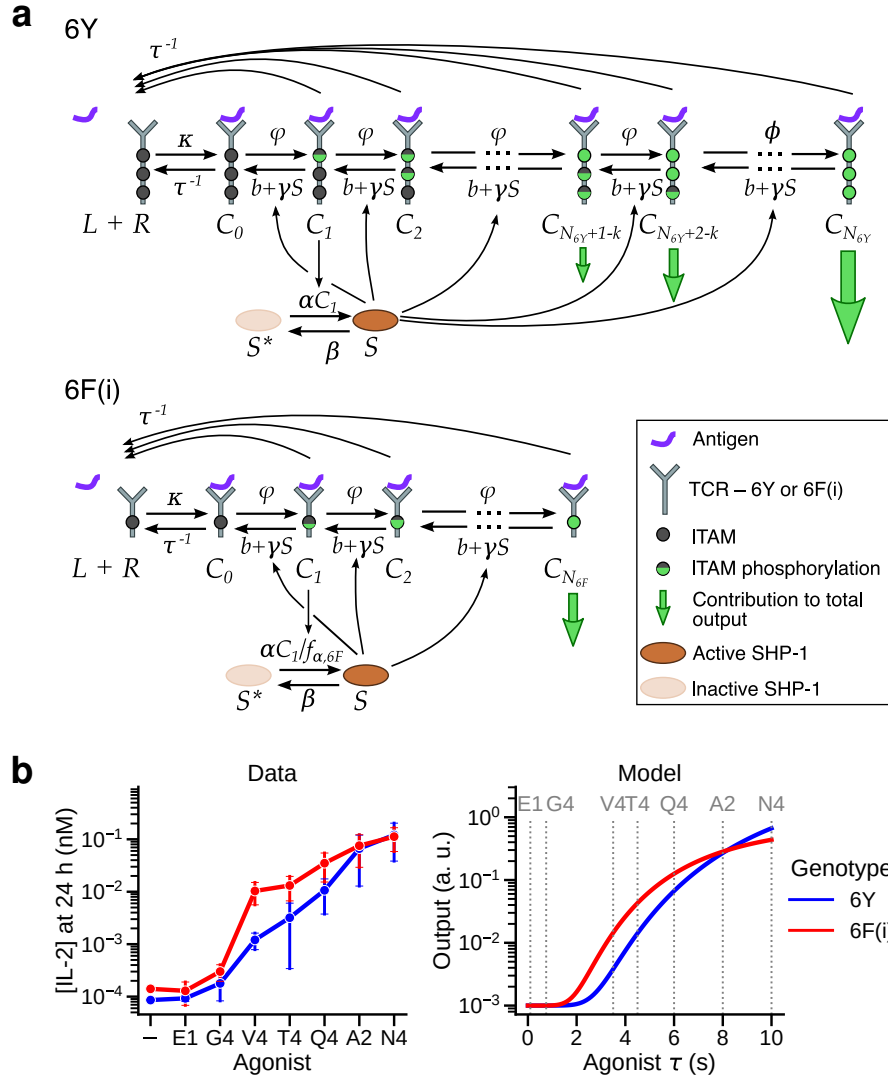


Figure 3.9: Classical AKPR model for 6F TCRs **a**, Phenomenological model of TCR activation predicting antagonism, as modified from François *et al.* [129]. On binding to a ligand (left), the TCR undergoes a series of biochemical modifications (indicated by color changes on ITAMs). Complexes can also unbind, followed by rapid dephosphorylation of ITAMs; thus, the TCR implements a KPR scheme. Complex C_m ($m = 1$) activates the negative feedback realized by the phosphatase SHP1 (S : active, S^* : inactive); SHP1 activation is reduced in $6F(i)$ by a factor $f_{\alpha,6F}$ to reflect the experimental observations reported in [2], Fig. 5. To reflect their higher ITAM multiplicity, $6Y$ TCRs can undergo N_{6Y} proofreading steps and $6F$ TCRs can undergo N_{6F} steps, with $N_{6Y} > N_{6F}$. Moreover, the last k_{6Y} complexes of $6Y$ TCRs contribute to the total output (green arrows), compared to only one for $6F(i)$ **b**, Left, model constructed to recapitulate the differences between the responses of OT1-6Y and OT1-6F(i) T cells to antigens with different affinity for the TCR. Right, experimental results as measured by IL-2 concentration at 24 h in response to different OT1 peptides.” ($CD3\zeta$ ITAMs, [2], figure 7). Error bars are standard deviations over 3 repeats with cells from 3 different mice.

This postulated effect of reduced ITAM multiplicity is consistent with the suboptimal $6F(i)$ response to antiCD3/antiCD28 crosslinking ([2], figure 1) or strong antigens (N4, figure 3.9B). The numbers we used are listed in the parameter table 3.3. Note that N_{6Y} and N_{6F} represent phenomenological TCR states and are not necessarily equal to the number of ITAMs." (*CD3 ζ ITAMs*, [2], SI)

From the steady-state solutions of the model variables $C_{n,l}$, we defined the activation output in the presence of L_1 , L_2 ligands of binding times τ_1 , τ_2 . We summed the scaled values of the last k complexes in the proofreading cascade,

$$O(\tau_1, \tau_2, L_1, L_2) = \sum_{n=N-k+1}^N \Theta_n(C_{n,1} + C_{n,2})$$

$$\text{where } \Theta_n(C_{n,1} + C_{n,2}) = \frac{(C_{n,1} + C_{n,2})^2}{\theta_n^2 + (C_{n,1} + C_{n,2})^2} \quad (3.13)$$

where, of course, $C_{n,2} = 0$ when considering the response to a single ligand type.

"The normalization constants θ_n are the values of C_n s in response to a large amount of L_1 strong agonist ligands alone ($\theta_n = C_n(\tau = 15 \text{ s}, L \rightarrow \infty)$), to ensure that Φ_n only saturates for very strong agonists. This normalization accounts phenomenologically for possible differences in the signaling potency of CD3 ζ and CD3 γ, δ, ϵ chains. It also compensates for the different scales of C_n at successive steps n in our model. Lastly, for plotting purposes, we added a small positive value (10^{-3}) to the output $O(\tau, L)$ to mimic the effect of an experimental lower limit of detection; it does not alter the predictions of the model." (*CD3 ζ ITAMs*, [2], SI)

3.4.2 Choosing binding times and ligand numbers

"To make comparisons between our model and experimental data, we had to match ligand binding times τ in the model with the binding times for OT-1 TCR binding to H-2K^b-OVA-derived antigens, and the numbers of ligands L per antigen presenting cell (APC) in the model with the number of ligands for different peptide pulse concentrations. Our model relies on phenomenological parameters so we did not attempt to match them with exact biochemical parameters; instead, we chose the following reasonable estimates. For τ , we associated evenly spaced values to the OVA-derived

antigens, which are more or less evenly spaced on a logarithmic EC_{50} axis [174, 176]. The table 3.2 below summarizes the selected values for τ .

For L , we assumed that a $1 \mu\text{M}$ pulse concentration P saturated the MHC molecules on the surface of the antigen presenting cell (APC), thus giving roughly 10^5 pMHCs per APC. Concentrations between $1 \mu\text{M}$ and 1nM are almost saturating, such that a 1nM pulse leads to 10^4 bound peptides per APC. We finally assumed that concentrations between 1nM and 10pM (those used in our experiments) are in the log-linear region of the dose response curve. To generate the continuous model prediction lines as a function of pulsing antigen concentrations P , we fitted a Michaelis-Menten calibration curve to those assumptions,

$$L = L_{\max} \frac{P}{P + P_0} \quad (3.14)$$

such that any number of ligands L could be mapped to a pulsing peptide concentration. This conversion curve is shown in figure 3.10 below.” (*CD3 ζ ITAMs*, [2], SI)

Peptide	τ (s)
N4	10
A2	8
Q4	6
T4	4.5
V4	3.5
G4	0.75
E1	0.1
None	0.001

Table 3.2: Peptide - τ map for early studies of 6F TCRs in section 3.4. It will be replaced by a more systematic conversion rule later.

3.4.3 Parameters of the model for single antigen types

For convenience of interpretation, in this version of the model, we supposed that the inhibitory molecule I was in fact SHP-1, so we called it S in this section, and we took $m = 1$ – both of these assumptions are as in the original model [129]. We also called $\beta_I/C_{m,\text{th}} = \alpha$ and $\beta_I = \beta$ the SHP-1 activation and deactivation rates, respectively.

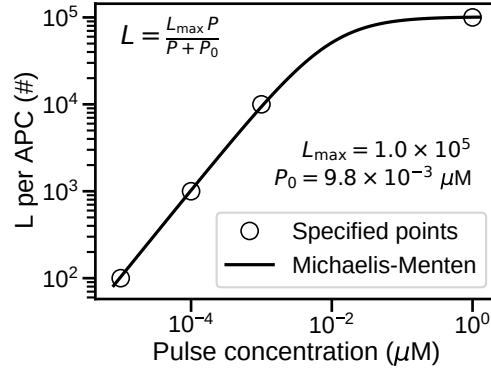


Figure 3.10: Approximate pulse concentration – L map conversion function, for section 3.4. It will be replaced by more systematic experimental characterizations later.

“After defining the model network and output, we made two modifications to the original parameter values to qualitatively match the cytokine response curves as a function of antigen quality (figure 3.9B), in response to a *single* kind of antigen (*i.e.*, $L_2 = 0$).

First, we included the important experimental observation that $6F(i)$ T cells recruit less SHP-1 by decreasing the phosphatase activation rate α by a factor $f_{\alpha,6F} = 8$, for this type of TCR only. Note that the normalization factors θ_n , however, were defined without this lowered SHP-1 rate, because those thresholds are set by the biochemistry of unaltered TCRs.

Moreover, the original model parameters were selected for pre-activated (‘blast’) T cells, while experiments here were done on naive cells; as a consequence, we had to reduce, for both TCR types, reaction rates φ , b , and γ by a factor $f_{\text{naive}} = 4$, and to increase the SHP-1 activation rate α by a factor $2f_{\text{naive}}$. Biologically, these modifications make sense: we expect pre-activated T cells to respond faster and to have less negative feedback inhibiting their response.

After making those two modifications, the model could qualitatively match data for a single antigen type (figure 3.9). The resulting model parameters are summarized in the table below. We did *not* tune parameters to fit antagonism experiments; antagonism predictions in figure 3.11f–h were generated after having selected final model parameters.” (*CD3ζ ITAMs*, [2], SI)

Parameter	Value in [129]	Value for 6Y	Value for 6F(i)	Description
R	3×10^4	Same value as original		TCRs per T cell
S_T	6×10^5	Same value as original		Total number of SHP-1 per cell
κ	10^{-4} s^{-1}	Same value as original		Ligand-receptor pair binding rate
τ	0.01 to 10 s	Same range as original		Ligand dissociation time
N	5	8	4	Number of proofreading steps
k	N/A	4	1	Last k complexes contribute to output
φ	0.09 s^{-1}	$\frac{1}{4} \times 0.09 \text{ s}^{-1}$	Same as 6Y	Phosphorylation rate
b	0.04 s^{-1}	$\frac{1}{4} \times 0.04 \text{ s}^{-1}$	Same as 6Y	Spontaneous dephosphorylation rate
γ	$1.2 \times 10^{-6} \text{ s}^{-1}$	$\frac{1}{4} \times 1.2 \times 10^{-6} \text{ s}^{-1}$	Same as 6Y	Dephosphorylation rate by SHP-1
β/α	500	$500/8 = 67.5$	500	SHP-1 deactivation-activation ratio

Table 3.3: “Model parameter values for 6Y and 6F or 6F(i) (germline or switched phenotype) naive T cells, compared to original model values (which were based on the literature and tuned for pre-activated T cells).” (*CD3 ζ ITAMs*, [2], SI)

3.4.4 Predictions of the antagonism ratio in various mixtures

To compare model predictions and cytokine data, the amount of antagonism exhibited by cytokines was quantified by the fold-change FC in cytokine output in presence of the antagonist peptide, compared to the agonist peptide alone. This defined an “antagonism ratio”,

$$FC_{\text{data}} = \frac{[\text{Cytokine}](\text{agonist} + \text{antagonist})}{[\text{Cytokine}](\text{agonist alone})} \quad (3.15)$$

“which is lower than 1 if the addition of the antagonist ligand lowers the total cytokine response, *i.e.*, if there is antagonism. From our cytokine time series, we computed the ratio at each time point, then took the geometric average of ratios across time, to have a single number characterizing each agonist-antagonist mixture.” (*CD3 ζ ITAMs*, [2], SI)

Figure 3.11a–e illustrates this process. Analogously, for the model, we computed the ratio of the steady-state outputs for the mixture compared to the agonist alone,

$$FC_{\text{model}} = \frac{O(\text{agonist} + \text{antagonist})}{O(\text{agonist alone})} \quad (3.16)$$

where O is the model output defined in equation (3.13) (for 6Y or 6F(i) T cells).

“We computed this ratio for a range of agonist binding times and agonist concentrations, in the presence of model equivalents for the antagonists we planned to used

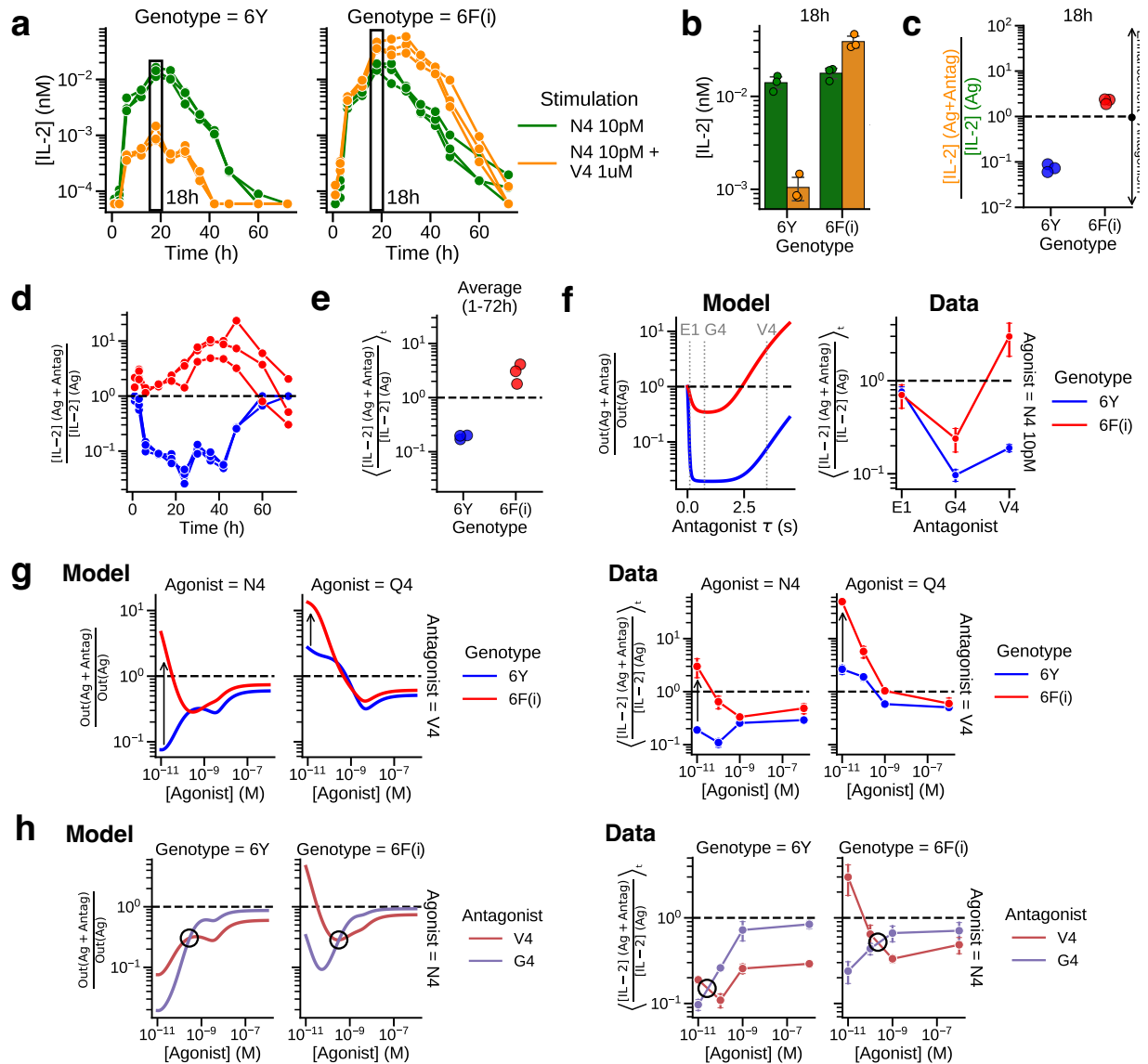


Figure 3.11: Model predictions of antagonism in 6Y and 6F T cells compared to data. Adapted from *CD3 ζ ITAMs*, [2]. “a, IL-2 secretion dynamics over 72 h for naive *OT1-6Y* or *OT1-6F(i)* (*dLck-Cre*) CD8⁺ T cells incubated with APCs pulsed with either an agonist alone (N4) or agonist plus antagonist (N4 + V4) at the indicated concentrations ($n = 3$ biological replicates). b, At 18 h, IL-2 levels were markedly lower in *OT1-6Y* T cells stimulated with both N4 and V4 compared to N4 alone, whereas V4 functioned as a co-agonist with N4 for *OT1-6F(i)* T cells. Data are shown as the mean $\pm$ s.d. c, Antagonism ratio for b. Values greater than 1 indicate enhancement of the agonist response by an antagonist ligand, while values smaller than 1 are indicative of antagonism. d, Antagonism ratio as a function of time revealed that adding V4 peptide increased IL-2 secretion by N4 peptide-stimulated *OT1-6F(i)* T cells but antagonized IL-2 secretion by N4 peptide-stimulated *OT1-6Y* T cells. e, Averaged antagonism ratio over all recorded time points.” (*CD3 ζ ITAMs*, [2], figure 6).

Figure 3.11: (Continued) (f), “Comparison of modeled versus experimental results on the degree of antagonism in *OT1-6Y* and *OT1-6F(i)* T cells as a function of antagonist peptide affinity. Both model and experiment show that *OT1-6F(i)* T cells are less susceptible to antagonism than *OT1-6Y* T cells across a range of antagonist affinities, and that there is an optimum range of antagonist affinities needed to maximize an antagonistic response. [g], Model (left) and experimental results (right) for antagonism as a function of agonist concentration, agonist and genotype [...]. Both model and experiment show that *OT1-6F(i)* T cells are antagonized less effectively than *OT1-6Y* T cells. [h], Model (left) and experimental results (right) for antagonism as a function of agonist concentration, antagonist and genotype for different antagonists [...]. Both model and experiment show that the antagonist that produces the largest antagonistic effect changes as the agonist concentration varies. $n = 3$ biological replicates. Data are represented as the mean $\pm$ s.d. Data are representative of three independent experiments.” (*CD3 ζ ITAMs*, [2], figure 7)

in experiments (V4, G4, E1 at $1\mu\text{M}$). In particular, we predicted the antagonism effects highlighted in figure 3.11f–h. Then, we compared the model predictions to the experimental results, without further tuning of the model parameters (except G4’s τ , which was lowered from 1.5 to 0.75 s). Hence, the proper output definition and the aforementioned adjusted model parameters for *6F(i)* naive T cells were determined in advance to capture data on the response to pure antigens (figure 3.11), which came from a previous experiment.” (*CD3 ζ ITAMs*, [2], SI)

Figure 3.11 shows that the classical AKPR model, with a few small adjustments based on insights from our latent space analysis (section 3.2), provides counter-intuitive, qualitative predictions of TCR/TCR antagonism, which are verified in the experimental data. The reduction in antagonism in *6F(i)* (4-ITAM) T cells is captured by the model, and the measurements confirm our hypothesis (figure 3.11f). We did not expect the fact that the V4 peptide goes from partial agonist to antagonist as the dose of TCR agonist is increased, yet it is verified in the data (figure 3.11g). We also did not expect the crossover of which antigen is the best antagonist as agonist density is increased (figure 3.11h). Just as it happens often in other fields of physics, the results could not have been guessed from qualitative reasoning only; they required a mathematical formulation to be derived. Our results illustrate how broad insights gained from latent space cytokine dynamics can be translated into mechanistic understanding of TCR signaling, thus bridging slow and fast T cell response time scales.

However, the agreement between the classical AKPR model and experimental data started to crack as we sought a direct quantitative agreement (*i.e.*, as we tried to directly superpose data and model curves). Upon a more systematic exploration of antagonist density L_2 in particular, and after

experimentally calibrating ligand densities as a function of antigen pulse concentration (to have more accurate estimates than those hypothesized in figure 3.10), we realized that a revised model of TCR kinetics would be necessary.

3.5 Revising the AKPR model of TCR/TCR antagonism

The previous model worked qualitatively for the conditions tested in section 3.4, but we then sought a direct, quantitative agreement between model and data. Moreover, we anticipated the need to generalize our framework to synthetic receptors for cancer antigens in immunotherapy, which have very different abundances and affinities than TCR antigens. Thus, we set out to more systematically explore the space of antigen parameters and to develop a revised AKPR model.

For this purpose, our collaborators

“generated a comprehensive dataset cataloging TCR-mediated antagonism and enhancement of functional T cell responses across a large set of TCR-engaging antigen mixes (Figure 3.12A). We documented highly-variable $FC_{\text{TCR/TCR}}$ in response to different strengths and quantities of antigens, compared to the response to an agonist alone (Figure 3.12B).” (*TCR/CAR antagonism*, [3])

This $FC_{\text{TCR/TCR}}$ ratio is the fold-change in T cell activation output in the presence of a mixture of antigens, compared to the agonist antigens alone, which was already introduced in section 3.4. We discuss again its calculation from data and model outputs below (section 3.5.1). It allowed us to directly overlay model and data antagonism levels. We exploited this quantitative comparison opportunity to estimate model parameters using Markov Chain Monte Carlo (MCMC) simulations and to develop a revised AKPR model with better fits to the data than the classical model (figure 3.12C).

3.5.1 Quantitative comparison between models and data

Appendix D details the procedures by which we estimated model parameters for TCR/TCR antagonism (this chapter) and TCR/CAR antagonism (next chapter) from experimental calibration and MCMC simulations. Here, we summarize the main steps we followed to first thoroughly assess the classical AKPR model, and then develop new models based on quantitative agreement with experimental measurements.

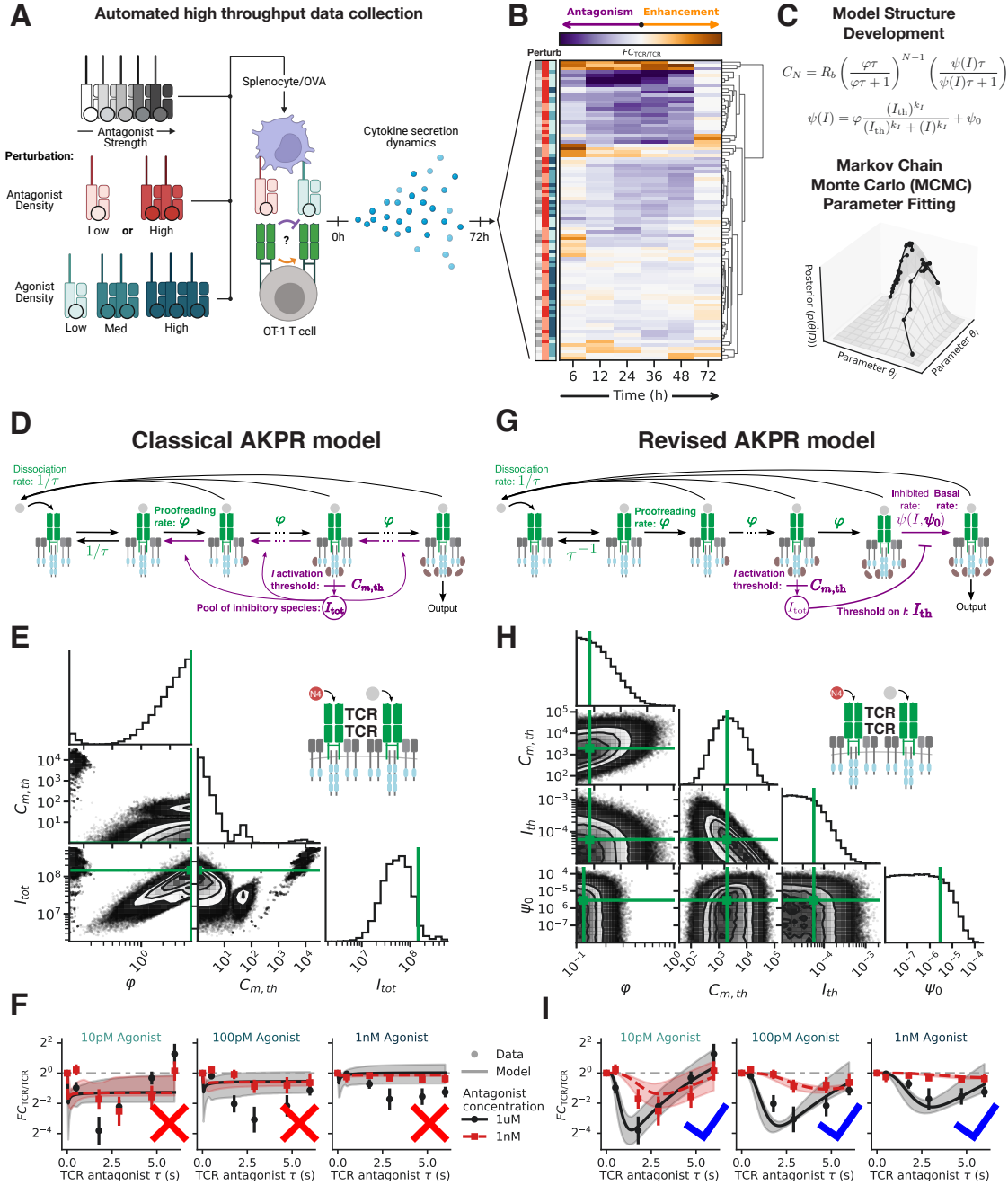


Figure 3.12: “Modeling TCR responses to mixes of antigens to quantify receptor cross-inhibition in T cells. (A) Scheme for multiplexed experimental settings to quantify TCR/TCR crosstalk. (B) Different TCR antagonist ligands modulated TCR response to the agonist ligand (OVA-N4) in both directions over time across a variety of conditions *in vitro*, as measured using fold changes (FC) of cytokine dynamics obtained from the IMMUNOtron robotic platform [1] ($n = 3$ biological replicates). (C) These data were used to develop analytically tractable mathematical models of receptor crosstalk and intracellular inhibition. Their parameters were estimated using Markov Chain Monte Carlo to fit $FC_{TCR/TCR}$ measurements. ” (*TCR/CAR antagonism*, [3], figure 2)

Figure 3.12: (Continued) “(D) Biochemical reaction network of the classical adaptive kinetic proofreading (AKPR) model used to describe TCR antagonism [129], extended to CAR and TCR receptors. Bolded parameters were fitted by MCMC. (E) Posterior parameter distributions of the classical model sampled by MCMC to fit TCR/TCR antagonism data. Crosshairs indicate the maximum a posteriori parameter estimate. (F) TCR/TCR antagonism model curves of the classical model, with the maximum a posteriori parameter estimates, compare poorly to the experimental data on which they were fitted. (G) Biochemical network of the revised AKPR model, differing from (A) by its inhibitory module, which only impacts the last KPR step. Model parameters are defined in table 3.4. (H) Posterior parameter distributions of the revised model, obtained by MCMC. (I) TCR/TCR antagonism model curves of the revised model, with the maximum a posteriori parameter estimates, fit well TCR/TCR antagonism data. Shaded areas in (F,I) are 90 % model CIs, generated by pulling 1000 parameter samples from the MCMC distributions in (E,H), while error bars are 95 % CI on the geometric mean of cytokine data from 3 experimental replicates with 6 time points each, based on a Student’s t-distribution.” (*TCR/CAR antagonism*, [3], figure 2)

Surface molecule abundances

To begin, our collaborators measured the abundance of TCRs on the surface of T cells and of MHCs on the surface of antigen presenting cells, using fluorescent antibodies and flow cytometry. Since our AKPR model describes the TCR signaling of a typical cell generalized to the whole population, we computed the geometric average across the measured single-cell abundance distributions⁴. These average abundances allowed us to fix the total number of TCRs, R_{tot} , and the total number of MHC (loaded with a peptide or not), respectively. The results for various cell types are in figure 3.13A–B and in the appendix, table D.1. Of note, while earlier modelling studies reported $R_{\text{tot}} = 30,000$ receptors per T cell [116, 129, 124], we consistently found higher receptor numbers, $R_{\text{tot}} \approx 1 \times 10^5$.

Then, to calibrate ligand abundance L_l as a function of the peptide pulse concentration, our collaborators measured the number of pMHCs on antigen presenting cells pulsed at various concentrations. We fitted these measurements with a Michaelis-Menten curve as a function of the pulse concentration $[P]$ (figure 3.13C left). We found a similar loading constant, $K_{D,\text{load}}$, for all OVA peptides (figure 3.13C, right). Hence, the ligand abundance L_l on a cell type with some total MHC abundance MHC_{tot} , for any antigen, can be determined as

$$L_l([P]_l) = \text{MHC}_{\text{tot}} \frac{[P]_l}{[P]_l + K_{D,\text{load}}} \quad (3.17)$$

⁴The geometric average is like the arithmetic average in log scale, hence it is better suited to the long-tailed distributions of surface molecule abundance across cell populations.

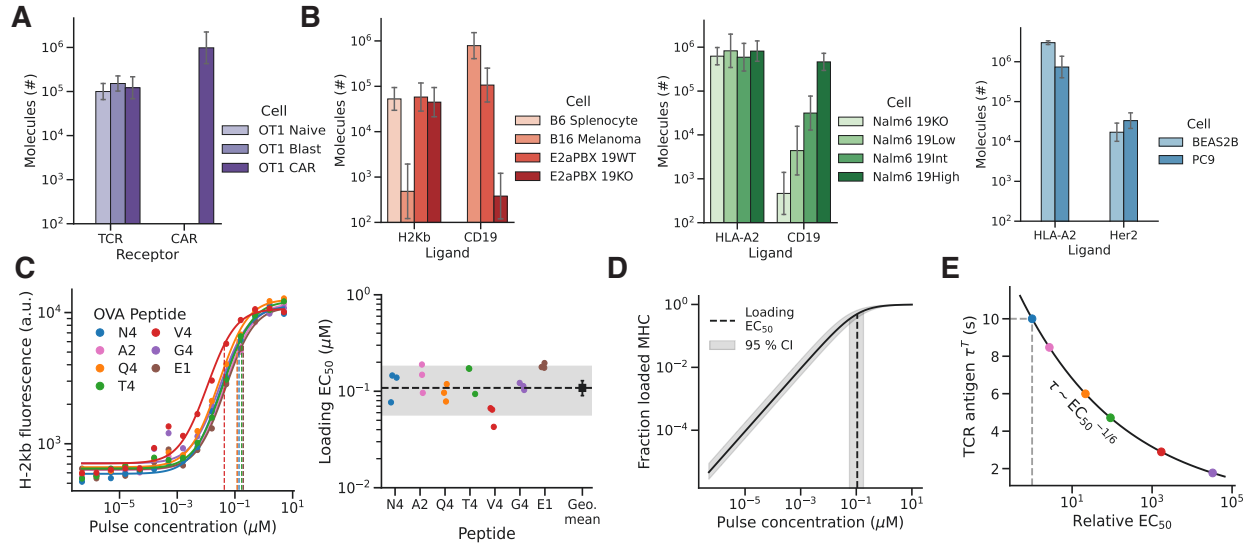


Figure 3.13: Quantification of surface molecule densities and antigen affinities. “(A) TCR and CAR receptor numbers on the surface of naive, preactivated, and CAR OT-1 T cells. (B) Major histocompatibility complex (MHC) and CAR target antigen numbers on murine antigen presenting cells (splenocytes) as well as B16 melanoma and E2aPBX leukemic cells (left), human leukemic cells (center), and human lung epithelial (BEAS2B) and adenocarcinoma (PC9) cell lines (right). (C) Calibration of OVA peptide MHC (pMHC) loading on transporter associated with antigen processing (TAP)-deficient RMA-S cells, quantified by immunofluorescence, as a function of peptide pulse concentration (representative example of three biological replicates show (left), and peptide loading EC_{50} s inferred from Hill fits on loading curves for each OVA peptide variant (right). (D) MHC loading curve built from EC_{50} s shown in (C). (E) KPR scaling law used to convert peptide affinity EC_{50} s into ligand binding times τ for the model. Error bars in panels a and b are cell population standard deviations in logarithmic scale. Shaded areas in panels (C, right) and (D, left) are 95 % CI of loading EC_{50} s across replicate and OVA peptides, and the error bar in panel (C, right) is the 95 % CI on the geometric mean, obtained by bootstrapping.” (TCR/CAR antagonism, [3], Figure S2).

We found a loading constant $\log_{10}(K_{D,\text{load}}/1\text{nM}) = 2.0^{+0.2}_{-0.3}$. Of note, this implies that a pulse concentration of 1 nM on APCs with $10^5 \text{MHC}_{\text{tot}}$ molecules corresponds to $L \approx 10^3$, which is 10 times less than what we had estimated in section 3.4. This means that we had over-estimated ligand abundances in our initial investigation of antagonism in 6Y and 6F T cells with the classical AKPR model. The wider than anticipated range of L values in the experimental data made it considerably more challenging to explain the observed antagonism patterns as the amount of antagonist peptides L_2 is decreased, and it was part of the reason why we needed to design a revised AKPR model.

Antigen quality and binding times

Moreover, we needed to assign a binding time τ to antigens based on their affinity, as quantified by their activation EC_{50} . We took the strong agonist N4 as a reference and assigned $\tau = 10$ s to it, following previous modeling studies with peptide binding times in the 0.1–10 s range [129, 124]. The binding time of other antigens was then computed by assuming EC_{50} s are defined by a KPR scaling law $O \sim L\tau^N$ passing some activation threshold $O = \Theta$, set by the reference N4. Hence,

$$\tau = \tau_{N4} \left(\frac{EC_{50, \text{pep}}}{EC_{50, N4}} \right)^{-1/N}.$$

We used $N = 6$ (since we also used this N in our revised AKPR model). This scaling calibration is shown in figure 3.13E. A different scaling of τ with EC_{50} could be (roughly) compensated for by tuning the KPR rate φ in the model; hence, this KPR assumption was sufficiently accurate for our purpose. The EC_{50} s of the very weak antigen E1 could not be determined, so we used $\tau = 0.5$ s.

Fixed and fitted model parameters

With these external parameters calibrated experimentally, we had to determine the rate parameters of the biochemical network models. We fixed some parameters to default values; for the classical AKPR model, these were the binding rate κ , the basal reverse rate b , and the I -mediated additional reverse rate γ . We set $N = 6$ for 6Y T cells; this number of KPR steps was also used in [59]. It provided a good balance between antigen discrimination, time to reach steady-state, and sufficient magnitude of C_N to avoid excessive noise in a stochastic version of the model [133, 45, 243, 244]. Other N choices (within the 4–8 range, say) could have worked too, by adjusting the φ rate and the KPR scaling law used to convert EC_{50} s to binding times. We allowed the remaining parameters to vary and fitted our models to our collaborator’s comprehensive TCR antagonism dataset. Appendix D details the different parameter fits that were performed (see especially tables D.2–D.4).

Antagonism fold-change ratio

To estimate model parameters, we fitted the model predictions to measurements of antagonism in binary antigen mixtures, quantified in terms of the antagonism FC quantity. For each mixture of agonists at concentration $[P]_1$ and antagonists at $[P]_2$, we computed the ratio of the responses to

the mixture compared to the agonist peptide (usually N4) alone:

$$FC_{\text{TCR/TCR,exp}} = \frac{\text{Out}(\text{agonist}[P]_1 + \text{antagonist}[P]_2)}{\text{Out}(\text{agonist}[P]_1)} \quad (3.18)$$

We computed this quantity for several different output readouts. All cytokines gave similar patterns of antagonism as a function of antagonist quality, simply with different amplitudes (see figure 3.12 in the next chapter). Thus, contrary to chapter 2 – where any single cytokine contained less information than the joint cytokine dynamics – here, we could use only one cytokine as a readout of T cell activation, since we were only interested in relative changes in T cell activation due to antagonism. We therefore decided to use cytokine IL-2 as the Out for model fitting purposes⁵. Since we had multiple experimental repeats of each data point p , we could also compute error bars σ_p on these measurements; we used a 95 % confidence interval based on Student's t-distribution (see section D.6).

To compute the model prediction corresponding to a mixture tested experimentally, we first obtained the antigen binding times τ_1, τ_2 and abundances L_1, L_2 corresponding to the mixture, then computed the model outputs. For TCR/TCR antagonism, we directly used the analytical steady-state solution for $C_{N,l}(\{\tau_l\}, \{L_l\})$ (*i.e.*, the numbers of bound TCRs in the final signaling state, in response to the antigen mixture $\{\tau_l\}, \{L_l\}$, eq. 3.4) as the output. Hence, we computed

$$FC_{\text{TCR/TCR,mod}} = \frac{C_{N,1}(\tau_1, \tau_2, L_1, L_2) + C_{N,2}(\tau_1, \tau_2, L_1, L_2)}{C_{N,1}(\tau_1, 0, L_1, 0)} . \quad (3.19)$$

corresponding to each $FC_{\text{TCR/TCR,exp}}$ in the fitted dataset.

MCMC parameter estimation

Then, we estimated model parameters by fitting FC measurements, using Markov Chain Monte Carlo simulations (figure 3.12C). To summarize, MCMC estimates model parameters by simulating a Markov chain, whose stationary distribution is, by construction, the posterior probability distribution of parameters θ given the data, $P(\theta \mid \text{data})$. The Markov process is built from a prior $P(\theta)$, a likelihood (or cost function) $P(\text{data} \mid \theta)$, and a proposal distribution, from which transitions in parameter space are sampled. We used the MCMC algorithm implemented in the *emcee* Python package [245], which is often used in astrophysics. It simulates an ensemble of random

⁵Specifically, we computed the FC at each experimental time point t_i , then took the geometric average of FC time points. Due to properties of the geometric average, this is equivalent to computing the geometric average of IL-2 over time, then computing FC of the averages.

walkers in parameter space and proposes transitions based on the current position of all walkers (so the moves are adapted to the local landscape). We opted for MCMC simulations, although they are more computationally costly than local optimization methods, because they provide a more complete exploration of parameter space and a Bayesian estimate of parameter uncertainty [246, 247, 248, 249, 250].

We used uniform priors with boundaries constraining parameters within biologically reasonable values (see table D.3). Our log-likelihood function was the sum of squared residuals between model and empirical $FC_{\text{TCR/TCR}}$,

$$\log P(\text{data} \mid \theta) \propto - \sum_{\text{data } p} w_p \left(\frac{\log_2 FC_{\text{model}}(\{\tau_{l,p}, L_{l,p}\}) - \log_2 FC_{\text{data},p}}{\sigma_p} \right)^2 \quad (3.20)$$

where we assigned a weight w_p to each point p ; we gave weight $w_p = 3$ to data points coming from the smallest agonist concentration (10 pM), since this was the most challenging condition to capture adequately with models, and $w_p = 1$ to other conditions. For the classical AKPR model, we performed a grid search over the index m of the complex which activates the negative feedback, performing MCMC simulation of parameters φ , $C_{m,\text{th}}$, and I_{tot} for each choice of m .

With these MCMC simulations set up for direct, quantitative comparison of antagonism between model and data, we verified more rigorously than in section 3.4 whether the classical AKPR model was able to account for TCR/TCR antagonism.

3.5.2 Shortcomings of the classical AKPR model

We performed MCMC simulations to fit the classical AKPR model (figure 3.12D) to our extensive TCR/TCR antagonism dataset, searching over the three parameters φ , $C_{m,\text{thresh}}$, and I_{tot} . We ran the simulations for 32 parallel random walkers each taking 40,000 steps – long enough to ensure proper convergence to the stationary distribution, according to autocorrelation and acceptance fraction metrics (figure 3.14). For every m value in $\{1, 2, 3, 4, 5\}$,

“MCMC optimization yielded spurious multimodal distributions of parameters (Figure 3.12E) that consistently failed to fit our data (Figure 3.12F). A careful analysis of this model failure revealed that the internal negative feedback in the classical AKPR model does not adequately vary with antigen density to match our extensive characterization, likely because it acts *en bloc* on all proofreading steps (Figure 3.12D).”
(*TCR/CAR antagonism*, [3])

Indeed, the inhibitory species I , with an activation rate linear in C_m and an impact on all KPR steps, did not have the necessary flexibility to capture two particularly challenging features of the data: first, that a small dose of 1 nM antagonists (corresponding to < 1000 ligands) can still produce significant antagonism compared to the larger dose; and second, that the antagonism curve for 1 μ M antagonists crosses the $FC = 1$ horizontal line to the left of the 1 nM curve⁶.

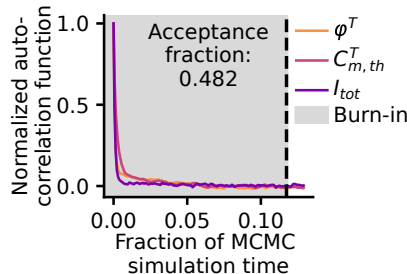


Figure 3.14: Autocorrelation function during MCMC simulations for the three rate parameters in the classical AKPR model, averaged across random walkers, computed following [245] (details in appendix D). The total simulation time is well above 50 times the autocorrelation time $\hat{\tau}$, the burn-in fraction (taken to be $20\hat{\tau}$) is small, and the acceptance fraction of proposed moves is between 0.2 and 0.5; these metrics indicate that the MCMC simulation has converged.

“The original AKPR model accounted well for antagonism as experimentally probed in earlier studies [129, 2] *i.e.*, a small quantity of TCR agonist ligands mixed with a large excess of TCR antagonist ligands.” (*TCR/CAR antagonism*, [3], SI)

However, it became unsatisfactory as we sought quantitative agreement between model and data in terms of the FC ratio, and as we varied more systematically antigen qualities and quantities L_1, L_2, τ_2 . No point in parameter space could account for all these antagonism and enhancement patterns simultaneously; even the best fit, shown in figure 3.12F, does not match at all the experimental data, merely producing a horizontal line roughly at the average antagonism FC for each agonist concentration. The classical model failed as well when we tried to extend it to antagonism in CAR T cells (next chapter). We did not attempt to fit data for 6F T cell receptors with full MCMC simulations, and instead focused on developing a refined model to understand important modulations in antagonism caused by antigen density.

⁶A model with perfect quantity-independent antigen recognition should have antagonism curves for any L_2 cross the line when $\tau_1 = \tau_2$ (as the mixture reduces to a single kind of antigen, the response to which only depends on τ , not $L = L_1 + L_2$). Even though the classical AKPR model does not have a perfectly L -independent response to antigens (figure 3.8), it cannot reproduce the crossover of the 1nM and 1 μ M antagonism curves.

3.5.3 Establishing a revised AKPR model

We needed an improved model to capture quantitatively non-linear effects in T cell responses to antigen mixtures. We remained within the hypothesis that these effects are caused by intracellular inhibitory cross-talk between receptors, since the experimental evidence reviewed in section 3.1.1 points to this cause (rather than competition for receptors or MHCs, or cytokine-mediated feedbacks). We also assumed that ODE-based models of receptor signaling kinetics can capture this cross-talk – while being prepared to explore stochastic or cell-population models, if necessary.

We focused on searching for models able to exhibit the rightwards shift of the antagonism versus τ_2 curve as L_2 is decreased, since this was an unexpected and striking feature of the data shown in figure 3.12F,I (compare the black and red sets of points). Designing more complicated models with added biochemical details did not generate this feature in a straightforward manner. We tried adding positive or negative feedbacks at various points of the kinetic proofreading cascade, based on biochemical studies of TCR crosstalk [138, 136]. We also tried larger models accounting separately for the two phosphorylation sites of each ITAM, and allowing parallel phosphorylation of these ITAMs. This led to a large space of receptor states $C_{k,l,m,n}$ giving the number of TCRs bound to ligand type l , with k ITAMs single-phosphorylated ITAMs, m doubly-phosphorylated ITAMs, and n ITAMs bound to ZAP-70, where keeping track of all possible transitions could only be done numerically. We explored several other such variations. Yet, in the end, these various models exhibited response curves to single antigens and antagonism curves in binary mixtures which were similar to effective AKPR models. Without added explanatory benefit, they were however more mathematically complicated, often requiring numerical integration of their stiff ODE systems to find steady-state solutions, which would have rendered model evaluation in MCMC simulations an order of magnitude more computationally costly.

Therefore, we stayed within the AKPR framework and adopted a different strategy. We initially sought to make the model as simple as possible, before parsimoniously adding back the ingredients necessary to capture all experimentally observed patterns of antagonism. We realized that the key solution was to introduce a non-linearity in how the inhibitory species I inhibits the KPR rate, to reduce the difference in antagonism between the 1 nM and 1 μ M antagonist doses, while keeping the $I \sim L$ scaling necessary to obtain L -independent antigen recognition over some range of L .

Inspired by the AKPR model evolved *in silico* in [132], we

“first rendered the model mathematically simpler by neglecting reverse reactions in KPR cascades, keeping only the forward rate, φ . Second, we changed the inhibitory

interaction by making the rate in the inhibited steps a Hill function of the inhibitory species I . Namely,

$$\psi(I) = \varphi \frac{(I_{th})^{k_I}}{(I_{th})^{k_I} + (I)^{k_I}} + \psi_0, \quad (3.21)$$

where I_{th} is a threshold for inhibition. This form adds flexibility in the inhibitory interaction by allowing different regimes of non-linearity and saturation as antigen densities are varied; the inhibition terms in previous AKPR models were essentially linear in I . Third, we assumed that the inhibitory species I only acts on the last f KPR steps (instead of a feedback on all steps, in the initial AKPR model). This implemented some asymmetry between proofreading steps in the cascade, allowing us to capture the antagonism patterns documented in this study." (*TCR/CAR antagonism*, [3], SI)

This change in the inhibitory coupling turns the negative feedback into an incoherent feedforward loop [251], which makes the mathematical solution simpler. Figure 3.12G illustrates this revised AKPR model, and figure 3.15 shows the details and parameters of the reaction network.

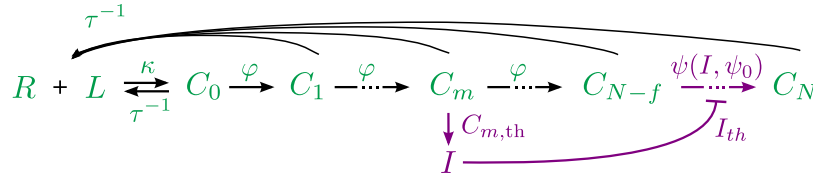


Figure 3.15: “Biochemical network corresponding to the revised AKPR model for TCR activation. Variables in green correspond to numbers of TCRs in various states, while variables, arrows and parameters in purple pertain to the inhibitory module.” (*TCR/CAR antagonism*, [3], SI)

Using mass-action kinetics for the remaining reaction rates in figure 3.15, we can derive the differential equations corresponding to this system when there are several ligand types, indexed by subscript l . We do not write the implied $\rho = T$ superscript.

“

$$\begin{aligned}
\frac{dC_{0,l}}{dt} &= \kappa(L_l - R_{b,l}) \left(R_{\text{tot}} - \sum_{l'=1}^2 R_{b,l'} \right) - (\varphi + 1/\tau_l)C_{0,l} \\
\frac{dC_{n,l}}{dt} &= \lambda_{n-1}C_{n-1,l} - (\lambda_n + 1/\tau_l)C_{n,l} \\
\frac{dI}{dt} &= \beta_I \left[\left(\frac{1}{C_{m,\text{th}}} \sum_{l'=1}^2 C_{m,l'} \right) (I_{\text{tot}} - I) - I \right]
\end{aligned} \tag{3.22}$$

where the forward rate is the same for either ligand type $l \in \{1, 2\}$ and depends on step n as

$$\lambda_n = \begin{cases} \varphi & \text{if } n < N - f \\ \psi(I) & \text{if } N - f \leq n < N \\ 0 & \text{if } n = N \end{cases} \tag{3.23}$$

For the last f steps, the regulated KPR rate $\psi(I)$ is as given in equation (3.21).”
(TCR/CAR antagonism, [3], SI)

Table 3.4 defines all parameters of the refined model of TCR/TCR antagonism, as well as for TCR/CAR antagonism (next chapter).

This revised model has mathematically simpler closed form solutions for its steady state than the classical AKPR model, due to neglecting reverse reactions. This simplicity allowed us to explore the model modifications necessary to capture the experimental data, whereby we converged to the non-linear form for the rate $\psi(I)$. Obtaining analytical solutions was also crucial to accelerate MCMC simulations, because these solutions are not computationally costly to evaluate; without them, we would have needed to integrate in time the ODE system until steady-state, for each MCMC step and each antigen mixture.

“To begin, consider a single receptor type and a single ligand type. The steady-state total number of bound receptors is still given by eq. (3.3). The numbers of receptors in each proofreading state are given by

$$C_n = \frac{R_b}{\varphi\tau + 1} \Phi^n \quad (0 \leq n < N - f) \tag{3.24}$$

$$C_n = \frac{R_b}{\psi(I)\tau + 1} \Phi^{N-f} \Phi_I^{n-N+f} \quad (N - f \leq n < N) \tag{3.25}$$

$$C_N = R_b \Phi^{N-f} \Phi_I^f \tag{3.26}$$

	Parameter	Definition
Indices	$\square^\rho$	Receptor type ($\rho \in \{\text{TCR}, \text{CAR}\}$)
	$\square_n$	Receptor state or other indication about the parameter
State variables	C_n^ρ	Number of receptors of type ρ in proofreading state n
	R_b^ρ	Total number of receptors of type ρ bound to a ligand
	I^ρ	Number of inhibitory molecules I activated by receptor type ρ
Input variables	τ^ρ	Binding time of ligands for receptor type ρ
	L^ρ	Total number of available ligands for receptor type ρ
Integer parameters	R_{tot}^ρ	Number of receptors of type ρ per T cell
	I_{tot}	Total number of inhibitory molecules (active or not)
	N^ρ	Number of proofreading steps for receptor type ρ
	m^ρ	Proofreading state of type ρ activating the inhibitory module
	f^ρ	Number of last KPR steps with forward rate reduced by I
	k_I^ρ	Hill power in $\psi^\rho(\vec{S})$
Rate parameters	φ^ρ	Forward KPR rate for states $n < f^\rho$ of receptor type ρ
	$\psi^\rho(\vec{I})$	Forward rate in the last f^ρ inhibited KPR steps
	ψ_0^ρ	Basal forward rate in the inhibited KPR steps
	κ^ρ	Binding rate of receptors and ligands of type ρ
	γ_{μ}^ρ	Strength of inhibition of receptor type ρ by I^μ .
	$C_{m,\text{th}}^\rho$	Activation threshold of I by C_m^ρ
	I_{th}^ρ	Threshold on I to inhibit KPR rate ψ
Output function parameters	τ_c^ρ	Binding time defining output activation threshold of receptor type ρ
	A^ρ	Maximum output amplitude of receptor type ρ
	$\alpha_A^{T,6F}$	Correction factor on TCR output amplitude for 6F TCR
	$\alpha_\tau^{T,6F}$	Correction factor on TCR threshold for 6F TCR
	$\alpha_A^{C,1\text{CAR}}$	Correction factor on TCR output amplitude due to 1-ITAM CAR
	$\alpha_A^{T,1\text{CAR}}$	Correction factor on CAR output amplitude for 1-ITAM CAR

Table 3.4: “Definition of mathematical model parameters, related to figures 3.12 and 4.3. Variables and parameters in the revised AKPR model, for TCR/TCR or TCR/CAR antagonism. 6Y: 10-ITAM TCR, 6F: 4-ITAM TCR (CD3 ζ ITAMs are mutated).” (*TCR/CAR antagonism*, [3], table S3)

where we defined the regular proofreading factor

$$\Phi = \frac{\varphi\tau}{\varphi\tau + 1} \quad (3.27)$$

and the inhibited proofreading factor

$$\Phi_I = \frac{\psi(I)\tau}{\psi(I)\tau + 1} . \quad (3.28)$$

Those proofreading factors play the same amplification role as $r_{\pm}$ in eq. (3.4) and are simpler to write because of the absence of backward rates. Just like in the François *et al.*, 2013 model, each complex is essentially given by the product of proofreading factors corresponding to the upstream complex in the cascades. However, contrary to the initial model, equations differ before and after $n = N - f$: the regular forward rate φ applies in the first $N - f$ steps, while the inhibited rate $\psi(I)$ applies to the last f steps. Lastly, we assumed, as in eq. (3.7), that I is activated by one complex, C_m , in the cascade:

$$I = I_{tot} \frac{C_m}{C_{m,th} + C_m} . \quad (3.29)$$

This equation is an explicit solution for I if $m < N - f$, since C_m is then known, but otherwise it has to be solved numerically when C_m depends on I via $\psi(I)$ and Φ_I ." (*TCR/CAR antagonism*, [3], SI)

Mass-action implementation of the nonlinear rate $\psi(I)$

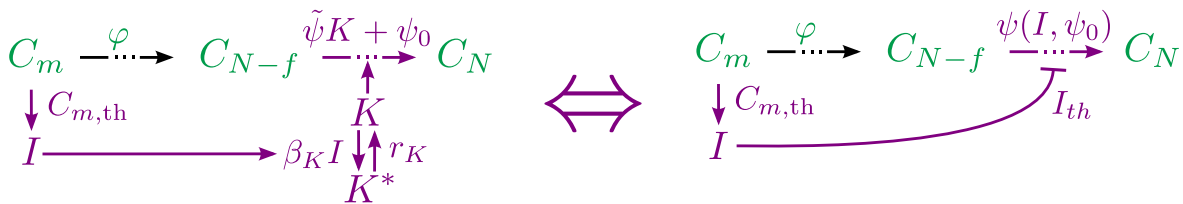


Figure 3.16: “(Left) Mass-action kinetics implementation of the nonlinear inhibitory module in our revised model, relying on a kinase K to mediate the last f proofreading steps, and deactivation of that kinase by the inhibitory molecule I . (Right) Recovering the nonlinear inhibitory module of our revised model from a quasi-static approximation on the kinase K .” (*TCR/CAR antagonism*, [3], SI)

“The nonlinear proofreading rate $\psi(I)$ can be derived as a quasi-static approximation of a biochemical model with mass-action kinetics, in which the inhibitory molecule

I deactivates the kinase K responsible for mediating the final f proofreading steps. We consider the case $k_I = 1$ for simplicity; higher exponents could be obtained by adding cooperative steps to the activation of the kinase K by I . Figure 3.16 (left) illustrates the reaction network corresponding to this implementation of the nonlinear inhibitory module in our revised model (right).

The differential equation for the time evolution of K , derived from mass-action kinetics, would read

$$\frac{dK}{dt} = r_K(K_{\text{tot}} - K) - \beta_K I K \quad (3.30)$$

where r_K is the rate at which inactive K^* is turned into active K , β_K is the rate at which I deactivates K , and $K_{\text{tot}} = K + K^*$ is the total amount of kinase in either form. Moreover, the rate of the final f proofreading steps in the main KPR cascade is, again by mass-action kinetics, $\tilde{\psi}K + \psi_0$, where we include a basal rate ψ_0 in the absence of the kinase, and where the rate $\tilde{\psi}$, to be specified, has units of $\text{time}^{-1}\text{molecule}^{-1}$. Solving equation (3.30) at steady-state, we find

$$K = K_{\text{tot}} \frac{r_K}{r_K + \beta_K I} \quad (3.31)$$

so the rate of proofreading regulated by K is

$$\tilde{\psi}K = \tilde{\psi}K_{\text{tot}} \frac{r_K/\beta_K}{r_K/\beta_K + I}$$

By taking $\tilde{\psi} = \varphi/K_{\text{tot}}$ and $I_{\text{th}} = r_K/\beta_K$, we recover the nonlinear rate $\psi(I)$ we introduced in equation (3.21) for our revised AKPR model." (*TCR/CAR antagonism*, [3], SI)

3.5.4 Results of the revised AKPR model for TCR antagonism

Model definition for binary mixtures

"For binary mixtures of TCR antigens (TCR/TCR antagonism), we distinguished as before $C_{n,1}$ and $C_{n,2}$, τ_1 and τ_2 , etc. The ODEs of our revised AKPR model in this case are presented in eq. (3.22) [with $l \in \{1, 2\}$]. At steady state, the numbers of TCRs bound to each ligand type, $R_{b,1}$ and $R_{b,2}$, are given by equations (3.9)–(3.11). Complexes $C_{n,1}$ and $C_{n,2}$ are given by eqs. (3.24)–(3.26), with τ_1 or τ_2 (including in Φ and Φ_I) and $R_{b,1}$ or $R_{b,2}$, respectively. The regulated KPR rate $\psi(I)$ is as in

eq. (3.21). The negative interaction between receptors is mediated by the inhibitory molecule I , given by (3.12), like in the François *et al.*, 2013 model. This equation is, in the revised model, an explicit solution for I if $m < N - f$, but otherwise it has to be solved numerically." (*TCR/CAR antagonism*, [3], SI)

Parameter optimization

We fitted four model parameters with MCMC: φ , $C_{m,\text{th}}$, I_{th} , and ψ_0 . Additionally, we performed a grid search over integer parameters k_I (Hill exponent in $\psi(I)$), m , and f (number of last KPR steps regulated by the inhibitory species).

"Following this update of the network structure, MCMC optimization reliably converged to unimodal distributions of parameters that fit the experimental data (Figure 3.12H). The optimized AKPR model was able to explain the observed range of antagonism/enhancement of TCR signaling across all measured mixtures of TCR antigens of varied densities and qualities (Figure 3.12I)." (*TCR/CAR antagonism*, [3])

The autocorrelation functions of parameters in the MCMC simulations attest of their convergence (figure 3.17). The best MCMC fit was obtained for $m = 4$, $k_I = 1$, and $f = 1$, meaning the data is better explained if the inhibitory coupling is a purely feedforward loop ($N - f > m$) with strong dependence on antigen quality ($C_m \sim L\tau^m$), but not a very sharp non-linearity in $\psi(I)$ ($k = 1$). Although the fits are not perfect for the 10 pM agonist, 1 nM antagonist conditions (red curve in the leftmost plot of figure 3.12I), the overall behaviour of antagonism as a function of L_1 , L_2 , and τ_2 is much better captured than with the classical model (see the side-by-side comparison in figure 3.12). Moreover, the posterior parameter distributions, $P(\theta \mid \text{data})$ shown in figure 3.12H are much smoother than for the classical model. They are not very sharply peaked, but this sort of broad parameter distribution is very frequent in biological model and has been termed "parameter sloppiness" [157, 158], as we discussed in section 1.4.

Mechanisms of the revised model

To understand better why the revised model fits adequately the experimental data, we looked at the model response curves to single antigens, shown in figure 3.18. They reveal three scaling regimes as a function of L , before receptors saturate, such that $R_b \approx \frac{\kappa R_{\text{tot}} L \tau}{1 + \kappa R_{\text{tot}} \tau} \sim L$.

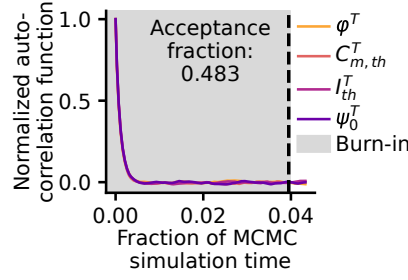


Figure 3.17: Autocorrelation function during MCMC simulations for the four rate parameters in the revised AKPR model with $m = 4$, $k_I = 1$, $f = 1$ (which yielded the best fit in the grid search over integer parameters), averaged across random walkers, computed following [245] (details in appendix D). The total simulation time is well above 50 times the autocorrelation time $\hat{\tau}$, the burn-in fraction (taken to be $20\hat{\tau}$) is less than 5 % of the total simulation duration, and the acceptance fraction of proposed moves is between 0.2 and 0.5; these metrics indicate that the MCMC simulation has converged.

1. At small L , there is a pure kinetic proofreading regime while $I \approx 0$ and receptors are unsaturated: $C_N \approx R_b \Phi^N \sim L\tau^N$. This is the region where curves are increasing linearly with L in the log-log plot.
2. Then, since the best parameter fit has a small threshold I_{th} for saturation of the regulated rate $\psi(I)$, but a large threshold $C_{m,th}$ for the activation of I by C_m , the first thing occurring as L increases is that I reaches I_{th} while $C_m \ll C_{m,th}$ still. Then, $I \approx I_{tot} C_m / C_{m,th} \sim L\tau^m$, while $\psi(I) \approx \varphi I_{th} / I$, such that $C_N \approx R_b \Phi^{N-1} \psi(I) \tau \sim L\tau^N / L\tau^m \sim \tau^{N-m}$. This is the region where C_N as a function of L is flat. Note that this regime starts at a τ -dependent L^* , when $I_{th} = I$, implying that $L^* \sim \tau^{-m}$ – hence, this happens earlier for stronger antigens.
3. Eventually, C_m reaches $C_{m,th}$, so the inhibitory molecule I saturates near I_{tot} , $\psi(I) \approx \varphi I_{th} / I_{tot}$, and thus $C_N \sim L\tau^N \times I_{th} / I_{tot}$: another KPR scaling regime but at a lower amplitude, reduced by a factor $I_{th} / I_{tot} < 1$. This is the rightmost region where response curves start increasing again (before receptors saturate). Note that this occurs at another ligand quantity $L^{**} > L^*$, but also scaling as $L^{**} \sim \tau^{-m}$. Weaker ligands thus never reach this second threshold and do not benefit from much extra activation at high doses.

The transition curves where regimes change, $I = I_{th}$ and $C_m = C_{m,th}$, are drawn as dashed and dotted lines, respectively, for a range of τ s on figure 3.18. The third regime explains the crossover between the 1 μ M and 1 nM antagonists curves in figure 3.12I: at a high dose, stronger antagonist peptides can produce some enhancement as their own output increases again, whereas at a lower dose, all antagonists produce insufficient activation, resulting in $FC < 1$. This is a subtle effect that

could not be achieved by the classical AKPR model while simultaneously fitting the antagonism data at other antigen densities.

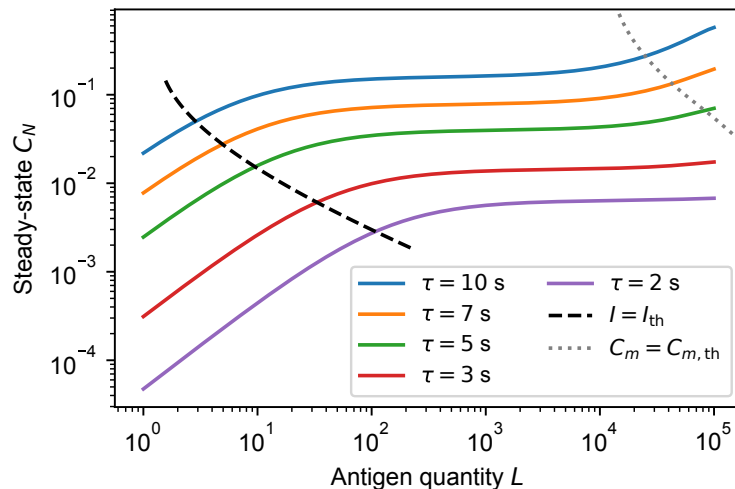


Figure 3.18: Response curves – steady-state output C_N as a function of antigen density L – of the revised AKPR model [129], for various antigen qualities τ . The steady-state solution plotted is exact and includes receptor saturation at large L . The transition between the first and second scaling regimes, when $I = I_{th}$, and the transition between the second and third regimes, when $C_m = C_{m,th}$, are shown for various τ values as dashed or dotted lines.

Revised model for 6F (4-ITAM) T cells

The classical AKPR model could not provide quantitative fits of antagonism for 6F (4-ITAM) T cells in our collaborators’ extended dataset, especially because it contained varying doses L_2 of antagonist peptides (unlike the original dataset of section 3.4). To capture differences between 6Y and 6F TCRs with the revised model, we adopted a more systematic approach. As we did for the classical model, we set $N = 4$ for 6F TCRs, instead of $N = 6$, to represent curtailed proofreading capabilities in the absence of functional CD3 ζ ITAMs. Then, instead of introducing hand-picked correction factors, we performed a separate MCMC estimation of the four fitted parameters – φ , $C_{m,th}$, I_{th} , and ψ_0 – specifically for 6F antagonism. We reasoned that any change in the 6F TCR biochemistry could be reflected by changes in these four effective kinetic rates.

Figure 3.19 shows the results of this separate MCMC parameter estimation. The simulation converged (as evidenced by the autocorrelation functions, panel B) to well-behaved parameter distributions (panel A), although exhibiting parameter sloppiness, just in as in the 6Y (10-ITAM) case (in figure 3.12H). The agreement between model and data in panel C is not perfect, but the model

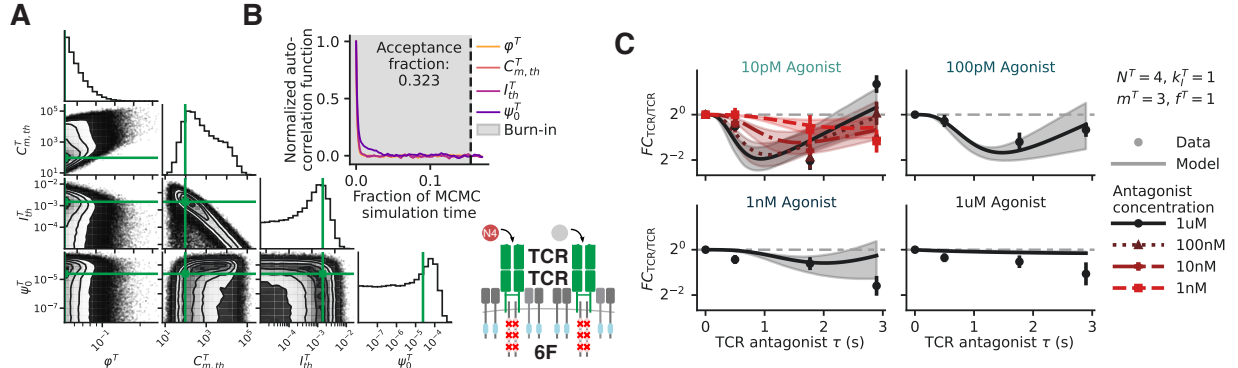


Figure 3.19: “[A] Posterior distribution of model parameters for 4-ITAM TCRs (6F), sampled by MCMC to fit TCR/TCR antagonism data.” (*TCR/CAR antagonism*, [3], figure S2) (B) Autocorrelation functions of the estimated parameters, with acceptance fraction, demonstrating convergence of the MCMC simulations. “[C] TCR/TCR antagonism curves of the revised AKPR model with maximal a posteriori parameter estimates for 4-ITAM TCRs (6F), as determined in (H), compared to the experimental data for these TCRs. Shaded areas are 90 % model CIs generated by pulling 1000 samples from the distribution in [B] while error bars are 95 % CI on the geometric mean of cytokine data, based on a Student’s t-distribution”. (*TCR/CAR antagonism*, [3], figure S2)

curves at 10 pM agonist (top left graph) vary with antagonist concentration (various line colors) similarly to the experimental data points, whereas this could not be captured by the classical model.

Comparing the optimal parameter values found for 6F (4-ITAM) and 6Y (10-ITAM) T cells, we observe that ψ_0 values are similar and poorly constrained in both receptor types; they just need to be small enough to allow the maximum antagonism amplitude observed in the data. The φ values are also similar, being $\sim 2\times$ smaller in 6F receptors, potentially to improve kinetic proofreading despite $N = 4$. This value of N could be an under-estimation of the true proofreading resolution of 6F TCRs, since N does not only represent ITAM numbers but also other downstream signaling reactions like ZAP-70 binding and LAT phosphorylation [115].

The most different parameters are $C_{m,th}$ and I_{th} ; this strongly supports our earlier findings that the two receptor types have different strengths of inhibitory signals, and that CD3 ζ ITAMs support this inhibitory function. Here, in 6F T cells, we find a smaller $C_{m,th} = 10^2$ and a larger $I_{th} = 1.5 \times 10^{-3}$. The larger I_{th} threshold indicates that 6F TCRs lacking CD3 ζ ITAMs are less sensitive to inhibitory molecules, matching our insight from the classical model. The smaller $C_{m,th}$, while it may seem counter-intuitive, makes sense in the revised model: the activation of I by 6F TCRs saturates more easily, and the second scaling regime, *i.e.*, L -independent antigen discrimination, has a reduced

range. This would contribute to their stronger responses to weak antigens and their lower mutual information for antigen quality (section 3.2.1).

3.6 Summary of progress on TCR antigen mixtures

In this chapter, we used our latent space description of T cell cytokine dynamics, developed in chapter 2, to gain a better mechanistic understanding of T cell receptor signaling, especially in the context of antigen mixtures and of TCRs with altered ITAMs. We showed how we could obtain qualitative, then quantitative, predictions of nonlinear TCR responses using parsimonious adaptive kinetic proofreading (AKPR) models, which are low-dimensional compared to the full space of biochemical receptor states (see discussion of more complicated models, p. 3.5.3). We developed a MCMC pipeline to estimate effective biochemical parameters based on quantitative agreement with antagonism FC data. This procedure allowed us to develop a revised AKPR model, and to interpret biological differences between regular (6Y) and altered (6F) TCRs.

In section 3.5, we showed that we could fit simplified models to data and interpret the fit results. However, we did not generate true predictions from the model, which could then be tested by independent experiments. In the next chapter, we attempt this final step in a new T cell system. We show how our revised AKPR model, originating from antigen encoding theory and calibrated on TCR/TCR antagonism data, can be generalized to cross-receptor interactions in T cells engineered for cancer immunotherapy.

Chapter 4

Antagonism in chimeric antigen receptor T cells

In this chapter, we generalize the revised AKPR model to understand and engineer cross-receptor interactions in chimeric antigen receptor (CAR) T cell immunotherapies. Our experimental collaborators, the laboratories of Grégoire Altan-Bonnet and Naomi Taylor (National Cancer Institute), empirically discovered that there is significant antagonism between TCRs and CARs in these T cells. To explain this effect, we built upon the results of previous chapters. We still relied on the insight from antigen encoding theory (chapter 2) that early receptor activation kinetics predict the continuum of T cell responses to antigen quality. We found that this correlation between time scales extended to *in vivo* antagonism, unfolding over weeks. Combined with latent space analyses and the Altan-Bonnet lab’s robotic platform, the revised AKPR model served as a convenient, low-dimensional framework to predict and manipulate high-dimensional immune responses.

To detail my contributions in this chapter, which are part of a submitted manuscript [3]:

- I developed a theoretical model able to account for interactions across receptor types in engineered T cells, extending the revised AKPR model (4.3.2–4.3.3);
- I obtained quantitative fits with experimental data to estimate the model parameters with Markov Chain Monte Carlo (MCMC) methods (4.3);
- I generated model predictions for various immunological settings, including *in vivo* experiments of CAR T cell treatments in mice (sections 4.4);
- In particular, the model predictions inspired a new CAR T cell design to my collaborators, wherein TCR/CAR antagonism is exploited to protect healthy tissues while maintaining treatment efficacy against tumors (4.4.5);
- I used the model as a pipeline to systematically analyze databases of TCR antigens (4.4.3).

In short, all theoretical analyses in this chapter are my work. Moreover, I wrote the supplemental information (SI) and, jointly with Sooraj Achar (Altan-Bonnet lab), the first draft of the main text.

The experiments (sections 4.2 and 4.3.1) were mainly performed by Sooraj Achar and Taisuke Kondo (Taylor lab).

4.1 Introduction: receptor interactions in engineered T cells

4.1.1 Overview of cancer immunotherapy

Cancers develop when some cells accumulate mutations conferring them an elevated growth *and* the ability to evade immune suppression [252]. Without the ability to grow unchecked, cancer cells would be cleared by immune cells in charge of maintaining homeostasis [253]. Cancer cells can acquire mutations allowing them to escape the immune system in various ways. They may over-express surface molecules that inhibit immune cells through immune checkpoints receptors [12] (such as CTLA4 [254] or PD-1 [101]). They tend to create a tumor microenvironment with physico-chemical factors such as oxygen levels, metabolites, pH, etc. which hinder immune cells [255]. To avoid T cells in particular, they may reduce their MHC levels [256]. Moreover, while cancer cell mutations often create new peptide sequences, called neoantigens, that can be presented on MHC and recognized by T cells, these peptides are not all very strong TCR agonists. In fact, immune editing shapes heterogeneous tumors by suppressing the most immunogenic cell clones while leaving weaker ones [257, 258].

Immunotherapy attempts to overcome these barriers and exploit a patient's own immune system to treat cancer cells. Several types of immunotherapy have been brought to various stages of clinical trials since the 1980s, such as administration of high cytokine doses [10], checkpoint inhibitor drugs [259, 12], cancer vaccines containing immunogenic neoantigens [11, 260], and adoptive cell transfer therapies [261]. The latter type of treatment consists in collecting T lymphocytes from a patient, activating and expanding them *ex vivo*, and re-injecting them in large numbers ($10^8 - 10^9$ cells) to attack cancer cells. Depending on the intended type of adoptive therapy, the collected cells are either selected because they have a TCR which is sensitive to tumor neoantigens (tumor infiltrating lymphocytes – TILs), or genetically engineered to express synthetic receptors that target tumor antigens. Chimeric antigen receptor (CAR) T cells are a particularly promising type of engineered T cell adoptive therapy [262, 263, 264, 265]. This is the type of immunotherapy we focused on, because our quantitative framework of T cell activation promised especially valuable insights into how natural T cell receptors interact and interfere with synthetic CARs.

When they work, immunotherapies produce a striking regression of cancer. Unfortunately, these treatments still suffer from large variability between patients, poor generalizability across cancer

types, and severe side effects like cytokine storms and toxicity in vital organs [266, 267]. Intense efforts are ongoing to overcome these hurdles, and predictive quantitative frameworks will be essential to optimize immunotherapy treatments [14].

4.1.2 Chimeric antigen receptor (CAR) T cell immunotherapy

Chimeric antigen receptor T cells offer a vast potential for model-guided bioengineering of anti-cancer treatments. In this version of adoptive transfer therapy, T cells sampled from the patient are endowed with a synthetic receptor to target surface antigens over-expressed in cancer cells, for example CD19, BCMA or CD22 (in B cell lymphomas) or HER2 (in breast cancer). After transduction of the CAR sequence and *ex vivo* expansion, the engineered cells are injected back in the patient, as illustrated in figure 4.1, left.

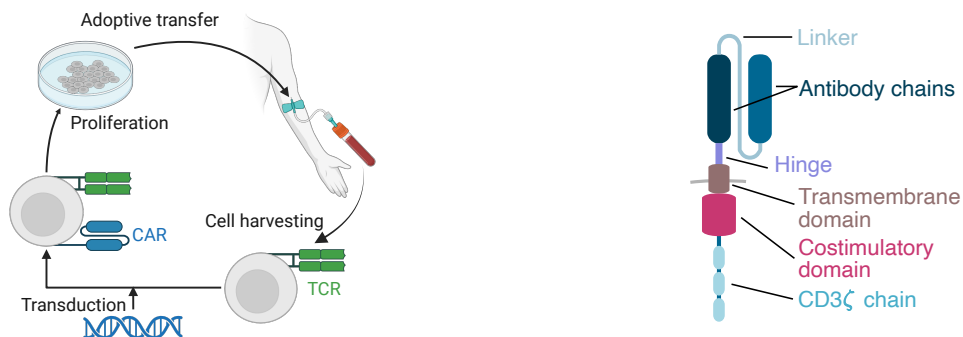


Figure 4.1: Chimeric antigen receptor (CAR) T cell therapy. (Left) Main steps of adoptive cell transfer therapy with CAR T cells. Transduction is the injection, using a viral vector, of a DNA plasmid encoding the CAR. Created with BioRender.com. (Right) Structure of a 2nd generation CAR. Due to the extracellular binding domain and the CD3 ζ chain, a CAR is a fusion between an antigen recognition domain, usually an antibody, and part of the TCR intracellular domain.

A chimeric antigen receptor (CAR) is, roughly speaking, a fusion between the extracellular part of a B cell receptor and the intracellular part of a T cell receptor. More specifically, as shown in figure 4.1, right, the extracellular part of the chimeric receptor is a binding domain, usually an antibody, specific to the target molecule, and the intracellular signaling domain is a CD3 ζ chain (of which a TCR has two, as explained in section 3.1.3). Upon binding of a CAR to its target, the CD3 ζ chain activates the same signaling pathways as TCRs (although with a potentially different balance of stimulatory and inhibitory signals). The strong binding affinity between an antibody and its target confers to CAR T cells a high activation and tumor killing potential. Moreover, a CAR also has various other connecting parts between those two main components, which can be tuned to further improve CAR function [268, 269].

4.1.3 Receptor interactions in CAR T cells

“Chimeric antigen receptor (CAR) T cell therapies have demonstrated remarkable success in the treatment of B cell malignancies [262, 263, 264, 265]. However, their application to solid tumors has been limited by potentially life-threatening on-target/off-tumor (OTOT) adverse events, caused by cytotoxicity against vital healthy tissues expressing CAR targets [266, 267, 270, 271, 272]. Recent attempts to circumvent this issue have been based on Boolean logic-gated CARs recognizing combinations of targets [273, 274, 275, 276, 277, 278], but their clinical applicability for the treatment of solid tumors remains challenging due, in part, to the limited number of tumor-specific CAR targets [279]. In contrast, T cell receptors (TCRs) recognizing mutated neoantigens on tumor cells operate under a different recognition paradigm, discriminating between tumor antigens and self-antigens on healthy tissues [258]. Nonetheless, these TCRs often elicit suboptimal tumor cytotoxicity due to the weak affinities and low densities of tumor neoantigens [280]. Combining these two tumor targeting modalities – potent CAR anti-tumor toxicity (Figure 4.2A, left) and exquisite TCR discrimination between neoantigen and self-antigen (Figure 4.2A, middle) – could significantly improve immunotherapeutic specificity.” (*TCR/CAR antagonism*, [3])

Hence, beyond molecular optimization of CAR components, there are also systems-level questions remaining about how CAR receptors interact with the rest of the TCR signalosome, and how these interactions could be exploited to fine-tune immunotherapies. Yet, we are lacking a quantitative framework to understand how CAR and TCR outputs combine in T cells. For this reason, there is no consensus on how TCRs impact CAR T cells: some labs simply knock out the TCR, while other labs report that the TCR improves persistence *in vivo* [281]. At any rate, it seemed clear to us that the TCR should have some effect, since CAR T cells clones with different TCRs have different tumor killing effectiveness and different transcriptional states [282].

Our theoretical biophysics approach, based on low-dimensional descriptions of T cell activation developed in the previous two chapters, has the potential to disentangle TCR/CAR interactions in an efficient way. Reciprocally, CAR T cells represent a practical test of our framework and its generality. Since both receptor types comprise CD3 ζ chains and their three ITAMs, CARs share signaling molecules with TCRs. They activate the same pathways, and they are most likely inhibited by the same species (SHP-1, Csk, CD45, etc.). Based on the success of adaptive kinetic proofreading in predicting antagonism in TCR antigen mixtures, we hypothesized that weak TCR

stimulation could also antagonize CAR signals by activating the inhibitory modules supported by CD3 ζ ITAMs. In fact, our first observation of universal antigen encoding in CAR T cells (chapter 2, figure 2.28) showed hints of such non-linear combination of TCR and CAR outputs.

We focused on TCR/CAR combinations, because the continuum of TCR antigen qualities, as found in chapter 2, potentially creates a corresponding continuum of CAR response modulations due to tumor neoantigens. In comparison, logic-gated chimeric receptor combinations are still confined to the CAR recognition paradigm, *i.e.*, recognition of the same antigens, merely expressed at different levels, on cancer and healthy cells. Still, a generalizable model of cross-receptor interactions could be exploited to maximize the reduction in off-tumor toxicity afforded by logic-gated CARs as well. Hence, encouraged by this generalization potential, we built upon our characterization of antigen encoding and TCR/TCR antagonism to model the nonlinear TCR/CAR interactions observed by our collaborators.

4.2 Experimental characterization of TCR/CAR interactions

4.2.1 Experimental platform

Like previous chapters, this project stemmed from a dialogue between theory and experiment. We therefore begin by summarizing the genetic engineering and immune profiling experiments performed by Taisuke Kondo, Sooraj Achar, and other authors of *TCR/CAR antagonism*, [3].

Our collaborators generated various CAR T cell lines over the course of this project. For the development and calibration of our mathematical model, they prepared CAR T cells from mouse cells expressing the 6Y OT-1 TCR – the main mouse line used in previous chapters as well – or the 6F (4-ITAM) TCR mouse line from the Love lab, discussed in chapter 3. They also generated human CAR T cells of a few other TCR types (NY-ESO, HHAT-specific TCR, etc.), using T cell samples from several donors at the NCI¹. In all cases, the CAR engineering procedure was roughly the same: cells were preactivated *in vitro*, transduced using viral vectors containing DNA plasmids for the desired CAR and/or TCR sequences, then expanded and rested for 1–2 weeks before being used for immune activation experiments.

Our collaborators also prepared or obtained at least 16 different lines of tumor cells expressing various CAR and TCR antigens. For *in vitro* experiments, different types of tumor cells (E2aPBX

¹All samples were collected and analyzed by our collaborators following NIH guidelines; the datasets I received were completely anonymous, simply containing different experimental repeats of cytokine dynamics recordings labeled ‘Donor 1’, ‘Donor 2’, etc.

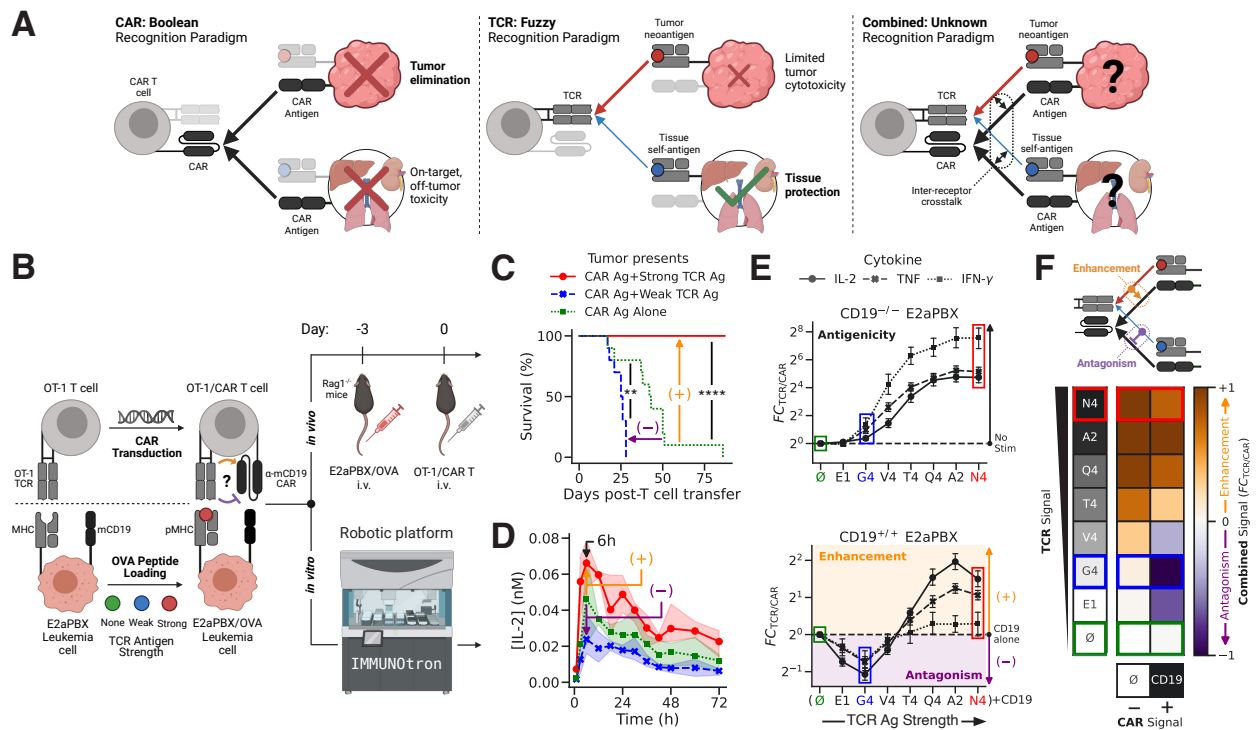


Figure 4.2: “Differential modulation of CAR activity over a spectrum of TCR antigens. (A) Outline of CAR, TCR, and combined recognition paradigms. **(B)** Experimental scheme to evaluate TCR/CAR crosstalk. Naive OT-1 T cells were transduced with a murine CD19 CAR and then stimulated with CD19⁺ E2aPBX leukemia cells expressing either no additional antigen or one of several ovalbumin peptide variants with differing antigenicities. **(C)** Survival curves for mice bearing E2aPBX/OVA leukemia cells treated with OT-1/CAR T cells ($n = 10$ mice per group). **(D)** A representative example of IMMUNOtron-acquired cytokine dynamics quantifying the divergent effects of TCR on CAR activation, as a function of the TCR antigen strength ($n = 3$ biological replicates). **(E)** TCR/CAR crosstalk for dual OT-1/CAR T cells stimulated with CD19^{-/-} (top) or CD19^{+/+} (bottom) E2aPBX leukemia cells. Y-axis represents the fold change (FC) of responses for a combination of CAR and TCR ligands compared to CAR ligand alone; an FC above 1 indicates that the strength of TCR signal enhances the overall response of the CAR, while an FC below 1 indicates antagonism of the CAR response. The degree of antagonism was evaluated by examining fold changes in the production of the cytokines IL-2, TNF and IFN- γ ($n = 6$ biological replicates). **(F)** Logic table derived from $FC_{TCR/CAR}(IL-2)$ for dual TCR/CAR stimulation of CAR T cells. Note the multivalued ($-1, 0, +1$) continuous modulation (fuzzy logic) of CAR activation (CAR Signal = CD19) as it relates to TCR signals alone (CAR Signal = $\emptyset$). Data in **(C,E)** are pooled from two independent experiments. Statistics in **(C)** were calculated using Kaplan-Meier survival estimates. Data in **(D,E)** are shown as mean $\pm$ 95 % confidence interval (CI).” (TCR/CAR antagonism, [3], figure 1)

leukemia cells, B16 melanoma cells, etc.) were pulsed with the desired set of TCR antigens, either OT-1 or NY-ESO (these cells expressed the right MHC allele to present these peptides). Some also were programmed to express different levels of the CAR target antigen (CD19, HER2) using CRISPR/Cas9 and single cell cloning systems [283]. For *in vivo* experiments, tumor cell lines had to be engineered (via lentiviral transduction) to constitutively express the desired TCR antigens, since pulsed peptides would not stay MHC-bound long enough *in vivo*.

To assess TCR/CAR interactions *in vitro*, CAR T cells were put in coculture with tumor cells expressing the CAR antigen, and pulsed with various strengths of TCR antigens (including cocultures without any TCR antigen) in the Immunotron robotic platform (figure 4.2B, bottom). Their cytokine dynamics were recorded over 72 h, providing the main datasets we used to develop our mathematical model. Cocultures containing T cells without a CAR were included in some experiments, to obtain TCR-only control data and assess TCR antigen strength on tumor APCs (*e.g.*, figure 4.2E, top). Moreover, our collaborators performed various *in vivo* experiments, where they typically injected tumor cells subcutaneously in mice, then treated them with CAR T cells a few days later and monitored survival and tumor progression over the following weeks (as outlined in figure 4.2B, top), either by direct measurement or via bioluminescence (section 4.4).

All experimental details are available in the supplemental information of *TCR/CAR antagonism*, [3]. This summary hopefully gives an idea of the extensive experimental efforts behind the datasets discussed in the next sections.

4.2.2 Evidence of antagonism (and enhancement) in CAR T cells

To characterize the impact of TCR signals on CAR T cell activation, our collaborators performed preliminary experiments *in vivo* and *in vitro*. They used mouse OT-1 T cells engineered to express a chimeric antigen receptor (CAR) specific to the surface molecule CD19, which is over-expressed by lymphoma cells (E2aPBX). For both types of experiments, the naïve expectation would be an additive model, where TCR and CAR signals simply add up to the total T cell activation. In this view, weak TCR stimulation (sufficient to produce some response on its own) should increase T cell activation and treatment efficacy compared to CAR stimulation alone (*i.e.*, in the absence of TCR stimulation). Strong TCR stimulation should increase the response even further.

In stark contrast with this linear view, the experiments revealed antagonism of the CAR by weak TCR signals. *In vivo*, the presence of the weak TCR antigen V4 on tumors reduced the CAR treatment efficacy (figure 4.2C): survival after CAR treatment of mice bearing these tumors was

significantly poorer than for tumors expressing only the CD19 CAR antigen (blue versus green lines). *In vitro*, the IL-2 concentration produced by CAR T cells responding to tumors presenting weak TCR antigen was lower, throughout time, than the IL-2 concentration produced in response to tumors without TCR antigens. In both cases, the CAR stimulation by its target molecule CD19 remains the same (all E2aPBX tumors had equal CD19 expression, see figure 3.13 and table D.1); *adding* weak TCR stimulation paradoxically *decreases* the overall T cell activation. This being said, in the case of strong TCR stimulation, the strong antigen N4 (red lines in panels C, D) enhanced T cell activation compared to CAR stimulation alone in both experiments.

To assess the extent of antagonism and enhancement regimes of CAR T cell responses, our collaborators performed further *in vitro* experiments with the whole range of OT-1 antigens (figure 4.2E, bottom). We quantified the level of antagonism or enhancement for each mixture of a TCR antigen with the CAR antigen using the FC ratio defined previously in section 3.5.1, eq. (3.18), here with the CAR antigen considered as the agonist:

$$FC_{\text{TCR/CAR}} = \frac{[\text{Cytokine}](\text{TCR Ag} + \text{CAR Ag})}{[\text{Cytokine}](\text{CAR ag alone})} . \quad (4.1)$$

A ratio $FC < 1$ indicates antagonism of the CAR by the TCR, while $FC > 1$ indicates enhancement of the CAR response. To compute FC , the experiments comprised cocultures without TCR antigen pulsed on the tumors. Moreover, to ensure that the OT-1 antigens themselves, when presented on tumors, could stimulate T cells on their own, our collaborators included, in preliminary datasets, cocultures with tumors lacking the CAR antigen (CD19 knockout, panel E, top).

The experimental data revealed, as we had hoped based on the AKPR model,

“that TCR signals transitioned from an antagonistic ($FC_{\text{TCR/CAR}} < 1$) to an enhancing ($FC_{\text{TCR/CAR}} > 1$) role as TCR signal strength was increased (Figure 4.2E, bottom). In addition, comparing the responses of OT-1/CAR T cells to CD19^{-/-} and CD19^{+/+} leukemia cells revealed maximum antagonism upon stimulation with a peptide producing very low (G4), but not zero (E1), TCR stimulation. Thus, ligands producing TCR signals at the edge of agonism (response) functioned as the best antagonists [284, 235, 116] of CAR function. Altogether, our results demonstrate that strong TCR stimulation enhances CAR T responses, while weak TCR stimulation antagonizes CAR activity, *in vitro* as well as *in vivo*. These observations are summarized in a logic table (Figure 4.2F), highlighting the multi-valued and fuzzy logic of dual TCR/CAR activation.” (*TCR/CAR antagonism*, [3])

This continuous spectrum of possible CAR T cell tuning through TCR stimulation, and the striking translation of *in vitro* antagonism to *in vivo* antagonism, sparked our interest in dissecting further TCR/CAR antagonism with our collaborators, with the hope of engineering applications at the outset. We were encouraged to attempt extending the AKPR model to TCR/CAR interactions for several reasons. The cross-receptor nature of this antagonism (acting between two separate sets of receptors and ligands) decisively rules out receptor or antigen competition as mechanisms (section 3.1.1), and instead points to an inhibitory coupling between receptors containing CD3 ITAMs, like in the AKPR model. The observed antagonism agreed with this model’s prediction for TCR/TCR antagonism at least at a qualitative level (peak antagonism is caused by weak, but not null, TCR antigens, see *e.g.*, figure 3.12I). The *in vivo* impact of antagonism indicates that the correlation between short and long time scales extends even further than we had seen in the latent space model parameters (chapter 2) and in our study of TCR/TCR antagonism *in vitro* (chapter 3). Moreover, we noticed in figure 4.2E that all cytokines produce *FC* curves of nearly identical shapes, simply with different absolute magnitudes; the *FC* ratio of a single cytokine – we chose IL-2 because of its reliability in experiments – can suffice to describe TCR/CAR antagonism. For all these reasons, it seemed feasible to capture the long-term response of CAR T cells, described by the low-dimensional *FC* metric of antagonism, with a parsimonious model of cross-receptor coupling kinetics.

4.3 Modelling TCR/CAR antagonism

“To rigorously elucidate the roles of both receptors in orchestrating the complex logic underlying dual TCR/CAR activation, we followed a two-step strategy: initially focusing on TCR recognition alone (Figure 3.12, previous chapter), then returning to TCR/CAR interactions (Figure 4.3). We relied on our IMMUNOtron platform (Figure 4.2B) that automatizes the measurement of lymphocyte activation over time, allowing the development of theoretical modeling and machine learning approaches to dissect complex immune functions [1].” (*TCR/CAR antagonism*, [3])

We presented our effort to dissect TCR/TCR antagonism in the previous chapter; it led to developing the revised AKPR model and estimating its TCR-related biochemical parameters. In the present chapter, we generalize kinetic proofreading models to TCR/CAR interactions. We first confirm that the classical AKPR model cannot account for empirical observations in CAR T cells more than it did for TCR/TCR antagonism. We then build upon the revised model, extending it to two (or more) types of receptors. Conveniently, by fitting TCR/TCR antagonism in the previous

chapter, we already have MCMC estimates of the TCR-related parameters. To ensure we could apply the same MCMC fitting pipeline to CAR-related parameters, our collaborators generated a more complete dataset of TCR/CAR antagonism data, which we describe first.

4.3.1 Summary of the complete *in vitro* TCR/CAR antagonism data

To disentangle TCR/CAR interactions in a range of immunological conditions and help us constrain our modelling efforts, our collaborators

“performed extensive *IMMUNOtron*-based measurements of CAR responses in the context of concomitant engagement of the TCR under conditions wherein TCR antigen strength and TCR antigen density were modulated (Figure 4.3A). Furthermore, both TCR and CAR signaling strength were altered by changing the number of immunoreceptor tyrosine-based activation motifs (ITAMs) on these receptors (Figure 4.3A). The resulting data sets (Figure 4.3B) were then used to expand and validate our optimized model of TCR responsiveness to include TCR/CAR crosstalk, (Figure 4.3C) via the inhibitory module uncovered for TCR/TCR crosstalk (Figure 3.12H), together with additional parameters specific to TCR/CAR crosstalk (Figure 4.3C, right and Figure 4.3D).” (*TCR/CAR antagonism*, [3])

Knowing in advance which biological parameters could be varied experimentally provided some modelling guidelines. A satisfactory model of TCR/CAR interactions should have parameter handles to account for varying TCR antigen quality, TCR and CAR antigen densities, and TCR and CAR ITAM numbers. Models from the AKPR framework naturally fit this description: binding times τ , ligand abundances L , and kinetic proofreading steps (N, m, f) , respectively, can be tuned to reflect changes in the varied biological parameters. Further effects of ITAM numbers can also be described by phenomenological changes in biochemical rates (as we did for 6F TCRs in figure 3.19).

4.3.2 The classical AKPR model cannot capture TCR/CAR antagonism

We first attempted to model cross-receptor interactions in CAR T cells using the classical AKPR model, to confirm the necessity of a revised model.

“We extended the phenotypic TCR activation model proposed in François *et al.* [129] by modelling the CAR as a second type of receptor, analogous to the TCR (*i.e.*, with

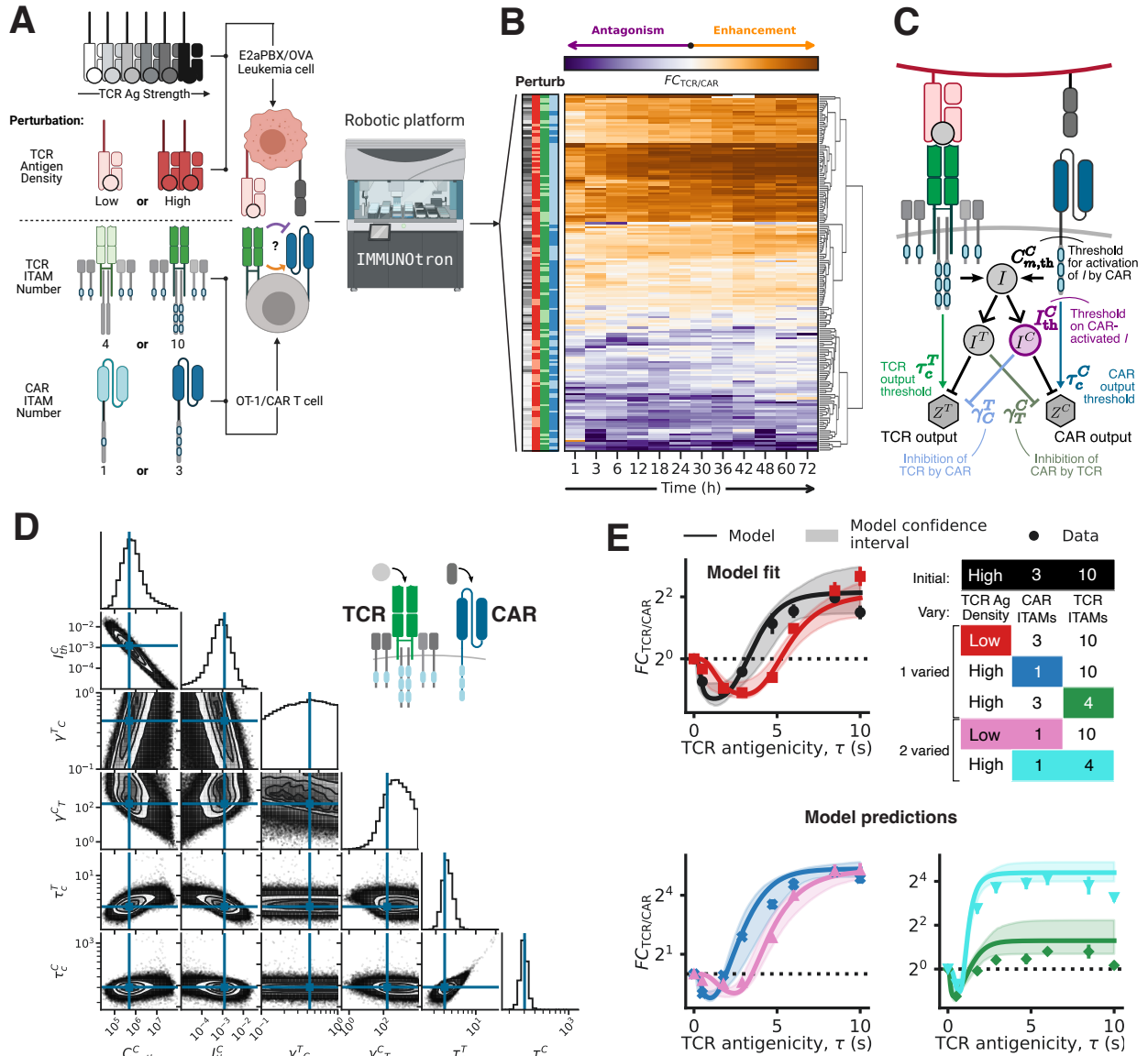


Figure 4.3: “Modeling TCR/CAR crosstalk in CAR T cells. (A) Scheme for multiplexed experimental settings to quantify TCR/CAR crosstalk. (B) Different TCR signal strengths modulated CAR activity in both directions over time across a variety of conditions in vitro, as measured using fold changes (FC) of cytokine dynamics obtained from the IMMUNOtron robotic platform ($n = 3 - 6$ biological replicates). (C) Structure of the model for TCR/CAR crosstalk. Both CARs and TCRs activate a shared inhibitory module, I , that reduces the rate of the kinetic proofreading cascade leading to formation of activated TCR (Z^T) and CAR (Z^C) outputs. The model accounts for the variable density and quality of TCR and CAR antigens, as well as the number of ITAMs of either receptor type. (D) Posterior distribution of the revised AKPR model parameters related to CAR and TCR/CAR interactions, sampled by Markov Chain Monte Carlo (MCMC) to fit TCR/CAR antagonism data for 3-ITAM CARs and 10-ITAM TCRs, for a high dose ($1 \mu M$) or low dose ($1 nM$) of TCR antigen.” (TCR/CAR antagonism, [3], figure 3)

Figure 4.3: (Continued) “(E) Best model fit with confidence intervals (top left). The model was then used to predict antagonism for other CAR and TCR constructs with altered ITAM numbers (bottom) The color of each curve refers to the color-coded condition presented in panel (top right). Data in (B,E) are pooled from two independent experiments and are shown as mean $\pm$ 95 % CI on the mean. Model curves shown in (E) are the best parameter fit and a 90 % CI obtained from MCMC parameter samples.” (*TCR/CAR antagonism*, [3], figure 3)

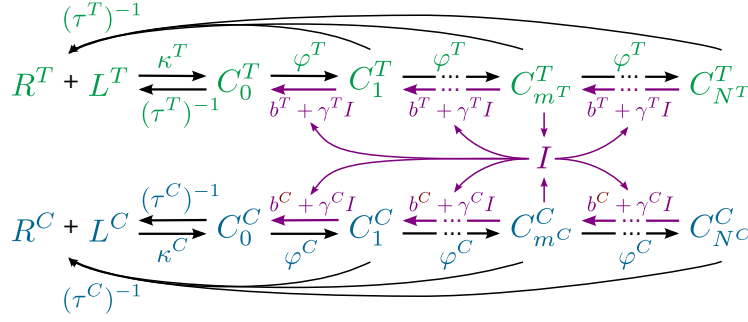


Figure 4.4: “Biochemical network corresponding to the François 2013 model, extended to TCR and CAR receptors. Blue: CAR variables and parameters, Green: TCR variables and parameters, Purple: shared inhibitory coupling, which controls the reverse reactions.” (*TCR/CAR antagonism*, [3], SI)

a KPR cascade upon receptor-ligand binding), but with its own pool of antigens. We used a superscript $\rho \in \{\text{TCR}, \text{CAR}\}$ on proofreading complexes and parameters to indicate the type of receptors considered. We then generalized the equations from François *et al.* [129] described [in section 3.3]. The two types of receptors are only coupled via the shared pool of inhibitory molecules I , which catalyzes with rate $\gamma^\rho I$ all reverse KPR reactions of each receptor type. Rate b^ρ represents the basal reverse rate. Other variables and parameters are defined as in table 3.4. Figure 4.4 illustrates this biochemical reaction network” (*TCR/CAR antagonism*, [3], SI)

and the ODEs can be derived from it by applying the law of mass-action.

“Conservation laws for the total number of ligands and receptors of each type completely determine the numbers of bound and free ligands, given the numbers of receptor in each kinetic proofreading state, $C_n^\rho(t)$. These, along with the number of active inhibitory molecules $I(t)$, are the only dynamical variables we need to consider. Writing the ODEs corresponding to the reaction network ($\rho \in \{\text{TCR}, \text{CAR}\}$), we

have

$$\begin{aligned}
\frac{dC_0^\rho}{dt} &= \kappa^\rho (R_{\text{tot}}^\rho - R_b^\rho)(L^\rho - R_b^\rho) + (b^\rho + \gamma^\rho I)C_1^\rho - (1/\tau^\rho + \varphi^\rho)C_0^\rho \\
\frac{dC_n^\rho}{dt} &= \varphi^\rho C_{n-1}^\rho + (b^\rho + \gamma^\rho I)C_{n+1}^\rho - (\varphi^\rho + 1/\tau^\rho + b^\rho + \gamma^\rho I)C_n^\rho \quad (1 \leq n < N^\rho) \\
\frac{dC_{N^\rho}^\rho}{dt} &= \varphi^\rho C_{N^\rho-1}^\rho - (1/\tau^\rho + b^\rho + \gamma^\rho I)C_{N^\rho}^\rho
\end{aligned} \tag{4.2}$$

where $R_b^\rho = \sum_{n=0}^N C_n^\rho$. The inhibitory coupling I , whose activation is catalyzed by receptors in state C_m^ρ , obeys

$$\frac{dI}{dt} = \beta_I \left[\left(\sum_{\mu} \frac{C_{m,\text{th}}^\mu}{C_{m,\text{th}}^\mu} \right) (I_{\text{tot}} - I) - I \right]. \tag{4.3}$$

" (TCR/CAR antagonism, [3], SI)

We are still mainly interested in the steady-state solution of this model, which is obtained by following steps analogous to the TCR-only model. Equation (3.3) for the total number of bound receptors applies separately to each receptor type, since they each have their own ligands. Therefore,

$$R_b^\rho = \frac{1}{2} \left(R_{\text{tot}}^\rho + L^\rho + \frac{1}{\kappa\tau^\rho} \right) - \frac{1}{2} \sqrt{\left(R_{\text{tot}}^\rho + L^\rho + \frac{1}{\kappa\tau^\rho} \right)^2 - 4R_{\text{tot}}^\rho L^\rho} \tag{4.4}$$

where $\rho \in \{\text{TCR}, \text{CAR}\}$.

Then, the steady-state number of receptors in each state can be calculated as in the TCR-only case, by solving a second-order recurrent relation for each receptor type separately.

"The number of receptors in each proofreading step is given by

$$C_n^\rho = R_b^\rho (a_-^\rho (r_-^\rho)^n + a_+^\rho (r_+^\rho)^n) \tag{4.5}$$

with

$$r_\pm^\rho = \frac{\varphi^\rho + b^\rho + \gamma^\rho I + 1/\tau^\rho}{2(b^\rho + \gamma^\rho I)} \pm \frac{\sqrt{(\varphi^\rho + b^\rho + \gamma^\rho I + 1/\tau^\rho)^2 - 4\varphi^\rho(b^\rho + \gamma^\rho I)}}{2(b^\rho + \gamma^\rho I)} \tag{4.6}$$

$$a_\pm^\rho = \frac{r_\pm^\rho - 1}{(r_\pm^\rho/r_\mp^\rho)^{N^\rho+1} - 1} \tag{4.7}$$

The level of the inhibitory species I at steady state is given by a similar implicit equation, where the C_m^μ depend on I ,

$$I = I_{tot} \frac{\sum_{\mu} C_{m,\mu}^\mu / C_{m,\text{th}}^\mu}{1 + \sum_{\mu} C_{m,\mu}^\mu / C_{m,\text{th}}^\mu} \quad (4.8)$$

Again $C_{m,\text{th}}^\mu$ defines a threshold of activation of the shared inhibitory variable by the complex $C_{m,\mu}^\mu$, $\mu \in \{\text{TCR}, \text{CAR}\}$. Those thresholds can be different since the receptors are not identical. Equation (4.8) for the level of inhibitory molecules is solved numerically to close the system of equations." (*TCR/CAR antagonism*, [3], SI)

Lastly, we defined TCR and CAR outputs, Z^T and Z^C , by applying Hill activation functions on C_{NT}^T and C_{NC}^C ; we explain this choice in more detail below, in section 4.3.3.

Despite our attempts, the classical model could clearly not match the *in vitro* TCR/CAR antagonism data collected by our collaborators and introduced in figure 4.3B. Figure 4.5 illustrates why: in this model, the maximal strength of antagonism scales inversely with the density of TCR antigens, while the data shows similar FC amplitudes at $1\mu\text{M}$ and 1 nM concentrations. Hence, as with the TCR/TCR data (section 3.5.2), the problem in the original model was manifest when varying TCR antigen densities. We did not attempt to further fit the classical model to TCR/CAR antagonism data. This analysis confirmed to us

"that the original model needed more flexibility in the implementation of the inhibitory interactions. Indeed, in the François *et al.*, 2013 model, all proofreading steps are negatively regulated identically. This assumption limits the model too much and fails to account for the TCR/CAR crosstalk and the TCR/TCR antagonism at low antagonist concentrations." (*TCR/CAR antagonism*, [3], SI)

The need to relax the symmetry in the inhibitory feedback in the classical model prompted us to turn to the revised AKPR model.

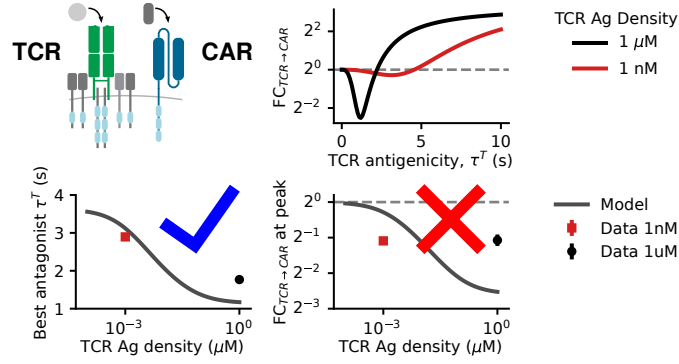


Figure 4.5: “Typical model curves (parameters not fitted) of TCR/CAR antagonism, scaling of the peak antagonism position with TCR antigen density, and the antagonism peak amplitude in the François *et al.*, 2013, compared to peak antagonism position and amplitude in experimental data. For the selected parameters, the model and data agree qualitatively in terms of how peak antagonism shifts as a function of L^T (blue check mark), but not in terms of the amplitude of this peak (red cross).” (*TCR/CAR antagonism*, [3], SI)

4.3.3 Revised AKPR model for TCR/CAR antagonism

Definition of the model

Equipped with our collaborators’ detailed dataset of antagonism as a function of antigen qualities and quantities, we were able to extend our revised TCR activation model to CAR T cells. As we did above when extending the classical model, we modelled the the CAR as a second type of KPR receptor with its own pool of antigens, and used superscripts $\rho \in \{\text{TCR}, \text{CAR}\}$ to indicate receptor type.

We generalized the equations for the revised TCR model to both receptor types. Coupling receptor types via the inhibitory species I was not entirely straightforward, since TCRs and CARs, sharing only parts of their intracellular domains, may not recruit signaling proteins in the same way (but we lack a complete empirical characterization of these details). To afford enough flexibility to fit the experimental antagonism data and potential asymmetries between CAR and TCR, we settled on the following structure for the inhibitory module.

“We applied the inhibitory effect only on the last f^ρ KPR steps of each cascade ($\rho \in \{\text{TCR}, \text{CAR}\}$). We also allowed more asymmetry in the inhibition by CAR and TCR by introducing separate forms of I activated by the TCR, I^T , or the CAR, I^C (this accounts for the recruitment of negative regulators to each receptor, based

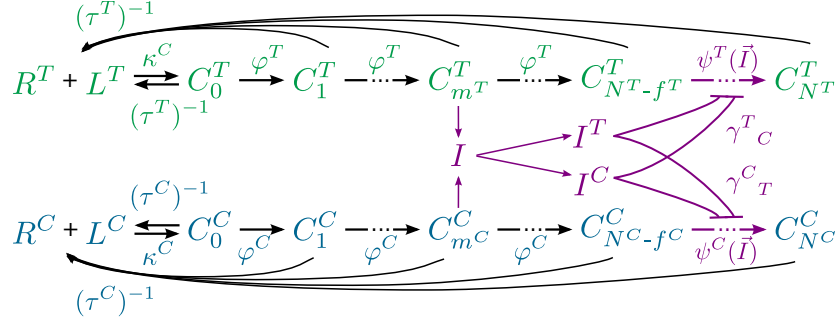


Figure 4.6: “Biochemical network corresponding to the revised AKPR model for TCR/CAR antagonism. Blue: CAR variables and parameters, Green: TCR variables and parameters, Purple: shared inhibitory coupling, which inhibits the last f^T and f^C steps of TCR and CAR proofreading, respectively.” (*TCR/CAR antagonism*, [3], SI)

on their state of activation [224, 138]). We generalized the inhibited rate in the proofreading cascade of each type ρ :

$$\psi^\rho(\vec{I}) = \varphi^\rho \frac{(I_{\text{th}}^\rho)^{k_I^\rho}}{(I_{\text{th}}^\rho)^{k_I^\rho} + (\sum_\mu \gamma_{\mu}^\rho I^\mu)^{k_I^\rho}} + \psi_0^\rho. \quad (4.9)$$

The vectorial dependence ($\vec{I}$) indicates that the forward rates depend on both I^μ . We also introduced a matrix γ_{μ}^ρ that encodes how each form I^μ inhibits the proofreading rate of receptor type ρ . This matrix has no reason to be symmetrical since receptors are different.

Figure 4.6 illustrates the biochemical reaction network of the model,” (*TCR/CAR antagonism*, [3], SI)

and the corresponding ODEs are

“

$$\begin{aligned} \frac{dC_0^\rho}{dt} &= \kappa^\rho (L^\rho - R_b^\rho) (R_{\text{tot}}^\rho - R_b^\rho) - (\varphi^\rho + 1/\tau^\rho) C_0^\rho \\ \frac{dC_n^\rho}{dt} &= \lambda_{n-1}^\rho C_{n-1}^\rho - (\lambda_n^\rho + 1/\tau^\rho) C_n^\rho \\ \frac{dI^\rho}{dt} &= \beta_I \left[\left(I_{\text{tot}} - \sum_\mu I^\mu \right) \frac{C_{m,\text{th}}^\rho}{C_{m,\text{th}}^\rho} - I^\rho \right] \end{aligned} \quad (4.10)$$

where the proofreading rate from C_n^ρ to C_{n+1}^ρ , λ_n^ρ , depends on n as in (3.23), that is,

$$\lambda_n^\rho = \begin{cases} \varphi^\rho & \text{if } n < N^\rho - f^\rho \\ \psi^\rho(\vec{I}) & \text{if } N^\rho - f^\rho \leq n < N^\rho \\ 0 & \text{if } n = N^\rho \end{cases} \quad (4.11)$$

The regulated rate $\psi^\rho(\vec{I})$ is defined in equation (4.9). The rate β_I is assumed to be fast enough for $\vec{I}$ and the C_n^ρ to reach steady state within minutes and does not enter equations (4.13)–(4.17) for that steady state." (*TCR/CAR antagonism*, [3], SI)

For reference, all parameters of the revised AKPR model for several receptor types are defined in table 3.4, in the previous chapter. Some parameters are highlighted in the condensed depiction of the revised model, figure 4.3C: these are the parameters dictating TCR/CAR interactions, through the inhibitory module and through the combination of their output signals, which had to be estimated by MCMC (see below).

Again, we solved this system of equations at steady-state. The total number of bound receptors of each kind is given by equation (4.4), as in the classical model. The rest of the solution can be obtained by iterating the recurrence relation resulting from setting $\frac{dC_n^\rho}{dt} = 0$,

$$C_n^\rho = \frac{\lambda_{n-1}^\rho}{(\lambda_n^\rho + 1/\tau^\rho)} C_{n-1}^\rho. \quad (4.12)$$

"The numbers of receptors in each proofreading state are, for each receptor type ρ ,

$$C_n^\rho = \frac{R_b^\rho}{\varphi^\rho \tau^\rho + 1} (\Phi^\rho)^n \quad (0 \leq n < N^\rho - f^\rho) \quad (4.13)$$

$$C_n^\rho = \frac{R_b^\rho}{\psi^\rho(\vec{I}) \tau^\rho + 1} (\Phi^\rho)^{N^\rho - f^\rho} (\Phi_{\vec{I}}^\rho)^{n - N^\rho + f^\rho} \quad (N^\rho - f^\rho \leq n < N^\rho) \quad (4.14)$$

$$C_{N^\rho}^\rho = R_b^\rho (\Phi^\rho)^{N^\rho - f^\rho} (\Phi_{\vec{I}}^\rho)^{f^\rho} \quad (4.15)$$

where we again defined the regular and inhibited proofreading factors

$$\begin{aligned} \Phi^\rho &= \left(\frac{\varphi^\rho \tau^\rho}{\varphi^\rho \tau^\rho + 1} \right) \\ \Phi_{\vec{I}}^\rho &= \left(\frac{\psi^\rho(\vec{I}) \tau^\rho}{\psi^\rho(\vec{I}) \tau^\rho + 1} \right) \end{aligned} \quad (4.16)$$

Each form I^ρ is activated out of a total pool of inhibitory molecules I_{tot} by the complexes $C_{m^\rho}^\rho$, such that

$$I^\rho = I_{\text{tot}} \frac{C_{m^\rho}^\rho / C_{m,\text{th}}^\rho}{1 + \sum_\mu (C_{m^\mu}^\mu / C_{m,\text{th}}^\mu)} . \quad (4.17)$$

Equation (4.17) for I^ρ is explicit if $m^\rho < N^\rho - f^\rho$ for all types ρ . When one receptor type, say $\rho = T$, satisfied this condition but not the other, we solved numerically equation (4.14) for the missing C_m^C (using Brent's method [285] in SciPy [187]), then determined $\vec{I} = (I^T, I^C)$. When both receptor types had implicit C_m , we solved numerically the 2D system of equations for $\vec{I}$ (using the 'hybr' method of Scipy's `optimize.root` function).

Notice that this model could readily be extended to more than two types of receptors. Figure 4.7 shows how typical behavior of this model as a function of L^ρ and τ^ρ promises to match experimentally observed features." (*TCR/CAR antagonism*, [3], SI)

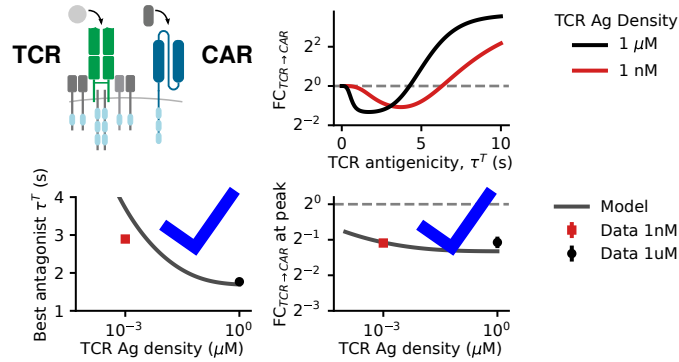


Figure 4.7: “In the revised AKPR model, typical model curves (parameters not fitted) of TCR/CAR antagonism, scaling of the peak antagonism position with TCR antigen density, and the antagonism peak amplitude, compared to peak antagonism position and amplitude in experimental data. The blue check marks highlight the qualitative agreement between model and data.” (*TCR/CAR antagonism*, [3], SI)

Definition of CAR and TCR outputs

To complete the extension of our receptor kinetic proofreading model to CAR T cells,

“we needed to combine TCR and CAR signals into a net T cell activation quantity. We assumed that the last complexes in the proofreading cascade, $C_{N\rho}^\rho$, activate downstream pathways with potentially different efficacies, since receptors are different. We assumed Hill activation functions of degree 2 and amplitudes A^ρ ; we defined TCR, CAR, and total outputs as

$$\begin{aligned} Z^\rho &= A^\rho \frac{(C_{N\rho}^\rho)^2}{(C_{N\rho}^\rho)^2 + (\Theta^\rho)^2} \quad (\rho \in \{\text{TCR}, \text{CAR}\}) \\ Z^{\text{tot}} &= Z^T + Z^C \end{aligned} \quad (4.18)$$

We set thresholds Θ^ρ based on the $C_{N\rho}^\rho$ produced by a large amount of ligands having some critical binding time τ_c^ρ for receptor type ρ only,

$$\Theta^\rho = C_{N\rho}^\rho (\tau_c^\rho, L^\rho \rightarrow \infty, L^{\mu \neq \rho} = 0) \quad (4.19)$$

We considered the critical times τ_c^T and τ_c^C as model parameters to be estimated. We normalized amplitudes $A^\rho = 1$ for regular TCRs and 3-ITAM CARs, but adjusted them based on separate data (section D.9) for 1-ITAM CARs and/or 6F TCRs (TCRs with inactivated CD3 ζ ITAMs, leaving 4 ITAMs [2]), to reflect curtailed proofreading capabilities by these altered receptors.” (*TCR/CAR antagonism*, [3], SI)

4.3.4 Fit and predictions with the revised AKPR model

Parameter estimation

With the revised AKPR model (equations 4.15–4.17) promising to have the correct scalings when extended to CAR T cells (figure 4.7), we proceeded to optimize its unknown parameters and assess how accurately it could capture the extensive TCR/CAR dataset (figure 4.3B) generated by our collaborators.

“We first fixed the parameters intrinsic for TCR activation – estimated from the MCMC optimization applied to the TCR/TCR dataset (Figure 3.12H) – and then used a new round of MCMC (Figure 4.3D) on the TCR/CAR interaction datasets collected for CARs bearing 3 ITAMs and wild-type TCRs (10 ITAMs) following stimulation with high and low TCR antigen doses (Figures 4.3E top and S2). Second, we validated our model and its MCMC-optimized parameters by generating predictions for the remaining dual TCR/CAR stimulation datasets generated with TCRs and/or CARs

bearing reduced numbers of ITAMs (Figure 4.3E, bottom).” (*TCR/CAR antagonism*, [3])

We determined model parameters similarly to our procedure for TCR/TCR antagonism (section 3.5.1). All details on the parameter estimation procedure, and on the generation of model predictions for altered TCR and CAR ITAMs, are provided in appendix D. In summary, we estimated ligand and receptor abundances from experimental calibrations, as explained in section 3.5.1 (and appendix D.1); these molecule numbers are shown in figure 3.13 and table D.1. We reused TCR parameters fitted on TCR/TCR antagonism, as appropriate for the TCR type (10-ITAM 6Y or 4-ITAM 6F) of the CAR T cells. We determined CAR binding and proofreading rates (κ^C , φ^C , τ^C for CD19) based on a previous study comparing relative TCR and CAR parameters [286]. We set the number of KPR steps of a CAR to $N^C = 3$, since it has only one CD3 ζ chain and must therefore have even less KPR steps than a 6F TCR, for which we have used $N^T = 4$ previously. We also normalized parameters $\gamma_{\rho}^{\rho} = 1$ and $I_{\text{tot}} = 1$, since only the ratios of I_{th}^{ρ} to I_{tot} and of γ_{ν}^{ρ} to γ_{ρ}^{ρ} matter in the steady-state solutions². These fixed parameters are summarized in table D.3.

The remaining parameters were determined by Markov Chain Monte Carlo simulations; these are the six parameters annotated in figure 4.3C, controlling TCR/CAR interactions in the inhibitory module and the normalization of receptor states $C_{N\rho}^{\rho}$ into outputs Z^{ρ} . Boundaries of the *a priori* distribution of parameters are given in table D.3. We performed a grid search over integer CAR parameters, running a MCMC simulation for each combination of m^C (step controlling the feedback), k_I^C (exponent in $\psi^C(\vec{I})$), and f^C (number of KPR steps regulated by I in CARs). We fitted experimental data for 6Y (10-ITAM) TCRs at either TCR antigen density (1 μM and 1 nM); we kept other receptor types for model validation. Table D.2, in the appendix, summarizes the details of these MCMC runs.

As in our TCR/TCR simulations, we used a least-squares cost function of the residuals between experimental and model antagonism values, measured by the FC ratio. For the TCR/CAR model, this ratio was computed from the receptor outputs Z^T and Z^C ,

$$FC_{\text{TCR/CAR}} = \frac{Z^{\text{tot}}(L^T, \tau^T, \text{CAR Ag})}{Z^C(\text{CAR Ag only})} \quad (4.20)$$

where $Z^{\text{tot}} = Z^T + Z^C$, that is, considering the CAR antigen as the agonist (setting the denominator) and the TCR antigen as the antagonist (in the mixture at the numerator).

²The strength of negative interactions are thus controlled by I_{th}^{ρ} and γ_{ν}^{ρ} , $\rho \neq \nu$.

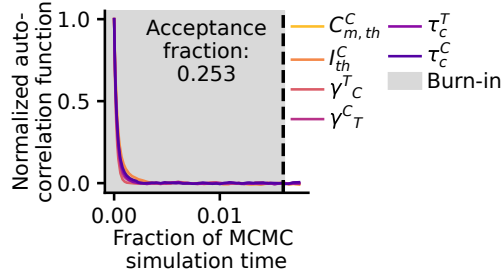


Figure 4.8: Autocorrelation function during MCMC simulations for the six estimated parameters in the revised AKPR model extension to TCR/CAR interactions. The autocorrelation estimate is averaged across random walkers, computed following [245] (details in appendix D). The total simulation time is well above 50 times the autocorrelation time $\hat{\tau}$, the burn-in fraction (taken to be $20\hat{\tau}$) is less than 5 % of the total simulation duration, and the acceptance fraction of proposed moves is between 0.2 and 0.5; these metrics indicate that the MCMC simulation has converged.

We ran MCMC simulations long enough to ensure convergence to the parameters’ *a posteriori* distribution (figure 4.8). The maximum *a posteriori* estimates we obtained from MCMC are provided in table D.4 (they are also indicated by crosshairs in figure 4.3D). Several m^C , k_I^C , f^C values in the grid search provided equivalent fit qualities; we therefore selected $m^C = 2$, $k_I = 1$, $f_C = 1$ by analogy with TCR/TCR fits which also tended to give $k = 1$, $f = 1$, and m as large as possible. Moreover, these k , f values make the inhibitory module as simple as possible and there was no biological reason to make it more complicated. The $m = 2$ value also helped to differentiate 3-ITAM CARs from 1-ITAM CARs, as explained below.

After estimating parameters using TCR/CAR antagonism data for 10-ITAM (6Y) TCRs, 3-ITAM CARs, we validated the model by generating predictions for different receptor types. We set $N^C = 1$, $m^C = 1$ for 1-ITAM CARs, since their single ITAM implies few kinetic proofreading steps. To obtain accurate predictions for antigen mixtures, we needed to compute correction factors for the activation functions of these altered receptors, using independent measurements of responses to single TCR or CAR antigens alone. These correction factors account for the different absolute magnitude of 6F TCR or 1-ITAM CAR responses, and for the different impact of 1-ITAM and 3-ITAM CARs on the TCR response³. These calculations are illustrated in figure 4.9 and detailed in appendix section D.9; values of the correction factors are provided in table D.3. It is important to emphasize that the single-antigen (TCR or CAR) data used to compute these factors is distinct

³These effects cannot be captured by the other biochemical since response amplitudes were normalized to 1 for 6Y TCR, 3-ITAM CAR T cells.

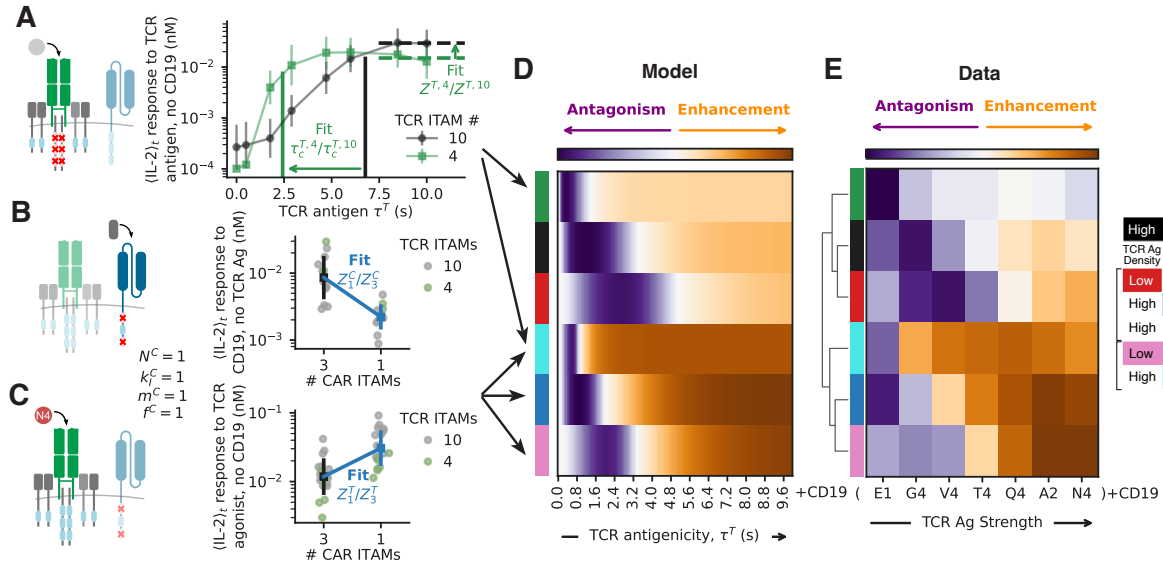


Figure 4.9: Generation of TCR/CAR model predictions. “(A) Correction for 4-ITAM TCR output amplitude and activation threshold, based on experimental IL-2 response to TCR antigens in the absence of CAR antigens. (B) Correction for 1-ITAM CAR output amplitude, based on the responses of 3- vs 1-ITAM CARs to CD19. (C) Correction for the effect of 1-ITAM CARs on the TCR output amplitude, based on the response of 1- vs 3-ITAM CAR T cells to TCR antigens in the absence of CAR antigen. (D) Model predictions of antagonism and enhancement in mixtures of TCR and CAR antigens, for various TCR and CAR constructs and TCR antigen doses, using the maximal a posteriori parameter estimates and correction factors determined previously on separate data. (E) Experimental data for conditions shown in [figure 4.3E]. Cluster ordering was performed on (D) and applied to (E). Error bars in panels (A), (B), and (C) are standard deviations across time points in two experimental replicates with 2–3 biological replicates each.” (*TCR/CAR antagonism*, [3], figure S2)

from the measurements of responses to TCR and CAR antigen mixtures. We could hence compute true model predictions, beyond mere parameter fitting, for these mixtures.

Fit quality and interpretation of the model

The revised AKPR model provided an excellent fit of $FC_{\text{TCR/CAR}}$ as a function of TCR antigen quality and quantity (figure 4.3E). It captures both the antagonism ($FC < 1$) and enhancement ($FC > 1$) regimes as TCR antigen quality is varied, and correctly captures the rightwards shift of the antagonism curve when TCR antigen density is decreased (red line compared to black line). The fit is particularly accurate for antagonism: despite the small error bars on the data (which come from having multiple time points and experimental replicates averaged), the fit goes through all data points with $FC < 1$. Enhancement is reasonably well captured, with some discrepan-

cies. The fact that the strong agonist N4 enhances more at a 1 nM dose (rightmost point in red) than at 1 μ M may be due to second-order effects, like receptor internalization [287] or antigen consumption [105]. We neglected such mechanisms, instead aiming for a simple cross-receptor inhibition mechanism able encompass the wide range of immunological conditions tested by our collaborators.

To understand how the model produces non-monotonic FC curves, we plotted, in figure 4.10, the TCR and CAR outputs as a function of L^T for a few representative TCR antigen qualities mixed with the CAR antigen. The peak antagonism occurs at the crossing of the inhibited Z^C and the start of agonism in Z^T , as seen in our preliminary experimental data (figure 4.2E). Hence, patterns of TCR/CAR cross-receptor interactions rest on a delicate balance between positive and negative signals from the TCR. In this balance, enhancement is almost entirely provided by the TCR output Z^T becoming larger than the initial Z^C output. The experimental observations can thus be explained without synergistic cross-receptor coupling in the model, which would provide direct enhancement of Z^C itself (as opposed to Z^{tot} only via Z^T) by strong TCR stimulation⁴.

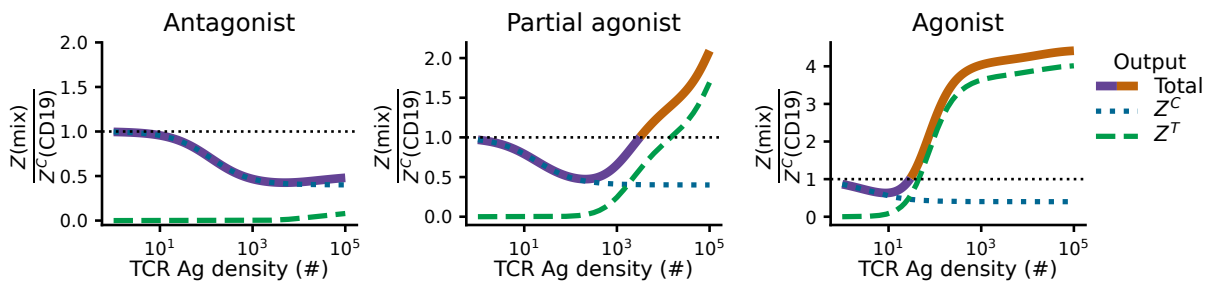


Figure 4.10: “Illustration of antagonism or enhancement of the CAR signal by a TCR antagonist (left, $\tau = 2$ s), partial agonist (center, $\tau = 4$ s), or agonist (right, $\tau = 10$ s) as a function of antigen density.” (*TCR/CAR antagonism*, [3], figure S2)

Underlying this fit, the MCMC simulations converged to smooth *a posteriori* parameter distributions, shown in figure 4.3D. The distributions reveal parameter sloppiness, as in the TCR/TCR case. Some sloppy directions correspond to understandable parameter correlations. For instance, there is a correlation between $C_{m,\text{th}}^C$ and I_{th}^C , the thresholds for activation of I by the CAR and for inhibition of the CAR by I , respectively (see top left pairwise plot in panel D)⁵. This is because in the scaling regime where $I^C \gg I_{\text{th}}^C$ and $C_m^C \ll C_{m,\text{th}}^C$, only the products $I_{\text{th}}^C C_{m,\text{th}}^C$ and $\gamma_T^C C_{m,\text{th}}^C$

⁴Future experimental analyses could reveal that the CAR output itself is increased by strong TCR stimulation, in which case the model would need to be improved to account for this true receptor synergy; positive interactions were not necessary in the present study.

⁵This was also the case for the equivalent TCR parameters, I_{th}^T and $C_{m,\text{th}}^T$, as can be seen in figure 3.12H

are important in $\psi(\vec{I})$:

$$C_m^C \ll C_{m,\text{th}}^C \Rightarrow I^C \approx \frac{C_m^C}{C_{m,\text{th}}^C}$$

$$I^C \gg I_{\text{th}}^C \Rightarrow \psi^C(\vec{I}) \approx \varphi^C \frac{I_{\text{th}}^C C_{m,\text{th}}^C}{C_{m,\text{th}}^C \gamma_T^C I^T + C_m^C}$$

showing that there are only two products of the three parameters which appear in the steady-state solution.

The optimal parameter values found by MCMC provide relevant biological observations. Most importantly, the large value of the parameter giving the strength of inhibition of the CAR by the TCR, $\gamma_T^C \approx 10^2$ (dimensionless, compared to $\gamma_T^T = 1$), indicates that the CAR must be very sensitive to I^T , the inhibitory species activated by the TCR. Parameter γ_C^T , quantifying the strength of inhibition of the TCR by the CAR, is poorly constrained, but has to be < 1 to prevent the CAR from inhibiting the TCR too much, which would be inconsistent with the observed enhancement by TCR antigens with $\tau > 5$ s. This strong asymmetry in the inhibitory module is necessary to account for the antagonism data, and stems from the necessity for TCRs to inhibit $\sim 10^6$ CARs without self-inhibiting more than what was observed in TCR/TCR antagonism data (figure 3.13). This is especially challenging to model, because CAR antigens bind very strongly to the antibody portion of the receptor (large τ^C), yet weak TCR stimulation can nonetheless antagonize that activation with weak τ^T , and enhance it with slightly larger τ^T or L^T . Hence, receptor asymmetry is a crucial component of the revised AKPR model extension to TCR/CAR cross-interactions, and it was missing from the classical model. Biochemically, there are various potential mechanisms: steric effects (*e.g.*, distance of ITAMs from the membrane) could make TCR-bound SHP-1 molecules more potent at dephosphorylating CARs than other TCRs, the I^T species could in reality be a different, potent enzyme activated only by the TCR, and so on. Tracking down these biochemical details was beyond the scope of the current project; our goal was to develop a phenomenological, low-dimensional model, and follow up on detailed mechanistic hypotheses in future work. We will nonetheless provide some biochemical characterization in section 4.4.4.

We made further observations about parameter values: we found that ψ_0 , the basal proofreading rate in the last steps regulated by I , has to be small but non-zero, to avoid complete shutdown of the CAR by the TCR. We noticed that CARs also have a small threshold for inhibition by I , $I_{\text{th}}^C \sim 10^{-3}$ (dimensionless compared to $I_{\text{tot}} = 1$), and a large threshold to activate I , $C_{m,\text{th}}^C \sim 5 \times 10^5$ receptors (compared to typically $R^C \sim 10^6$ receptors total on a CAR T cell). This implies the same three

scaling regimes explained for TCRs in section 3.5.4. Lastly, the thresholds τ^T and τ^C are the best-determined parameters (bottom right graph of figure 4.3D), because they contribute to setting the point at which TCR antigens transition from antagonism ($FC < 1$ while $Z^T \approx 0$) to enhancement ($FC > 1$ because Z^T increases). This transition point, however, also varies as a function of antigen quantity L^T , as shown by the “partial agonist” condition in figure 4.10. Hence, capturing this transition across immunological parameters was still a considerable modelling challenge, not solved only by fitting the τ^T parameter. This constrained our design of the revised AKPR model.

Prediction accuracy for altered receptors

Comparing the heat maps in figure 4.9D–E shows how the revised AKPR model qualitatively captures all variations in the TCR/CAR antagonism and enhancement patterns as receptor ITAM numbers change. Comparing model curves and data in figure 4.3E, bottom reveals that the model predictions are accurate at a quantitative level. The main discrepancy is observed for conditions with 4-ITAM (6F) TCRs and can in fact be attributed to a batch effect in the main experiment involving these receptors: the CAR T cells with 4 ITAMs had an unusually high activity in the absence of TCR antigens, thus reducing the extent of enhancement possible (whereby the model curve is higher than data points)⁶.

“Our revised AKPR model correctly predicted a decrease in antagonism when the number of ITAMs on either the CAR or TCR was decreased – from 3 to 1 and 10 to 4, respectively [2]. Crucially, the model also accounted for the observation that a low dose of partial agonist can antagonize as efficiently as a larger dose of weaker ligands, and that the peak of antagonism shifts to higher quality TCR antigens upon a reduction in TCR antigen density (Figure 4.3E and Figure 4.10), suggesting that a negative ‘background’ effect occurs as soon as the TCR is engaged, with stronger ligands required to overcome this effect at low densities. Together, our data and model predictions suggest that TCR-mediated antagonism as well as enhancement of CAR signaling can take place in multiple immune settings, in particular in environments characterized by low levels of antigen presentation such as solid tumors.” (*TCR/CAR antagonism*, [3])

⁶We could have fitted another correction factor, $\alpha_A^{C,6F}$, accounting for the different amplitude of Z^C in the presence of 4-ITAM versus 10-ITAM TCRs, but contrary to the other correction factors that we estimated separately for predictions, this one seemed to be an experimental rather than truly biological effect, hence we did not attempt to correct it.

4.4 TCR/CAR antagonism in various immunological settings

Given the success of the revised AKPR model, we attempted to generalize it to immunological settings that were increasingly relevant physiologically and clinically: solid tumors *in vitro* and *in vivo* (section 4.4.1), human CAR T cells *in vitro* (section 4.4.4), and ultimately in mice with human-derived tissues (section 4.4.5). We also show how the model can serve as a pipeline to optimize TCR/CAR interactions to protect healthy tissues via antagonism of the CAR, yet maintain tumor clearance via enhancement by the TCR (section 4.4.3). We did not aim for quantitative accuracy in all these contexts, but the model served as a motivation for the exhaustive battery of experiments carried out by our collaborators. We describe briefly the model predictions and the experimental results for each immunological setting tested by our collaborators.

4.4.1 Generalizing model predictions to other tumor cell types

To generalize the model to various contexts, we first needed a systematic way to account for different tumor cell types. Fortunately, we could account for these contexts by changing TCR and CAR antigen densities, based on calibrations with fluorescent antibodies and flow cytometry, in the same way as for the default (E2aPBX) tumors.

“MCMC parameter estimation simulations were done with data from experiments where the antigen presenting cells were either B6 splenocytes (for TCR/TCR antagonism) or CD19-expressing or CD19-KO E2aPBX tumor cells (for CAR/TCR antagonism). The effect of changing tumor cell type for further model predictions (figures 4.11–4.15) was straightforwardly reflected in the model by changing ligand numbers L^T and L^C according to surface molecule abundances we measured for these other cell types (figure 3.13 and table D.1).

We took into account differences in cell surface area when setting L^ρ from surface expression measurements. We divided total surface abundance by the average surface area of the tumor cells relative to E2aPBX, for which the model parameters were estimated. This was particularly important for B16 melanoma cells, which are large cells [288], $\sim 5\times$ larger surface area than E2aPBX based on our observations (data not shown). Nalm6 tumor cells are also moderately large cells [289], $2.5\times$ larger surface area than E2aPBX based on our observations.” (*TCR/CAR antagonism*, [3], SI)

We accounted for surface area differences because T cells only come in contact (forming an immunological synapse) with some fraction of the entire tumor cell surface, and thus only have access to some fraction of the total CAR antigen and pMHC numbers. The AKPR model's biochemical parameters were fitted to work with E2aPBX tumor cells, so we only needed to account for differences in size relative to that type. For the PC9 lung cancer cell line, used below in section 4.4.5, we did not have a clear estimate of tumor cell size, hence we did not correct them for a potentially different surface area.

4.4.2 Tumors with low TCR antigen density tested *in vivo*

"We therefore tested the functional significance of TCR-mediated antagonism in a melanoma model (B16-CD19) expressing substantially lower basal MHC levels and higher basal CD19 levels compared to E2aPBX leukemia cells (Figure 4.11A, left and 3.13B). Our mathematical model predicted that despite this lower level of MHC expression, TCRs would maintain their ability to tune CAR signals (Figure 4.11A, right). To test this prediction, we injected B16-CD19 melanoma cells, engineered to express strong (N4), weak (V4), or no (Null) OVA peptide [290], into the flanks of mice prior to treatment with OT-1/CAR T cells (Figure 4.11B, left). We evaluated immunotherapeutic activity by long-term measurements of tumor growth and high-dimensional immunophenotyping of tumor-infiltrating T cells (Figure 4.11B, right). There was no significant difference in the unperturbed growth of the three melanoma lines *in vivo* (B16-CD19/Null, B16-CD19/V4, and B16-CD19/N4; Figure 4.11C, left), and, as expected, the adoptive transfer of OT-1 T cells without a CAR limited tumor growth commensurably with the strength of the TCR antigen (null<weak<strong) (Figure 4.11C, center). Notably though, upon transfer of OT-1/CAR T cells, tumor rejection was markedly antagonized by weak TCR stimulation and enhanced by strong TCR stimulation (weak<null<strong) (Figure 4.11C, right)." (*TCR/CAR antagonism*, [3])

Hence, the predictions of our revised model of cross-receptor interactions, with parameter values estimated to fit *in vitro* experimental data acquired over 72 h, seemed to generalize at least qualitatively to *in vivo* measurements spanning weeks. We pushed this idea of correlation between time scales to the limit and tried to quantitatively predict *in vivo* outcomes, in terms of post-treatment survival times, based on the *in vitro* model output. We supposed that the survival time of tumor-bearing mice after CAR treatment, T_s , is some linear function of the receptor model total output,

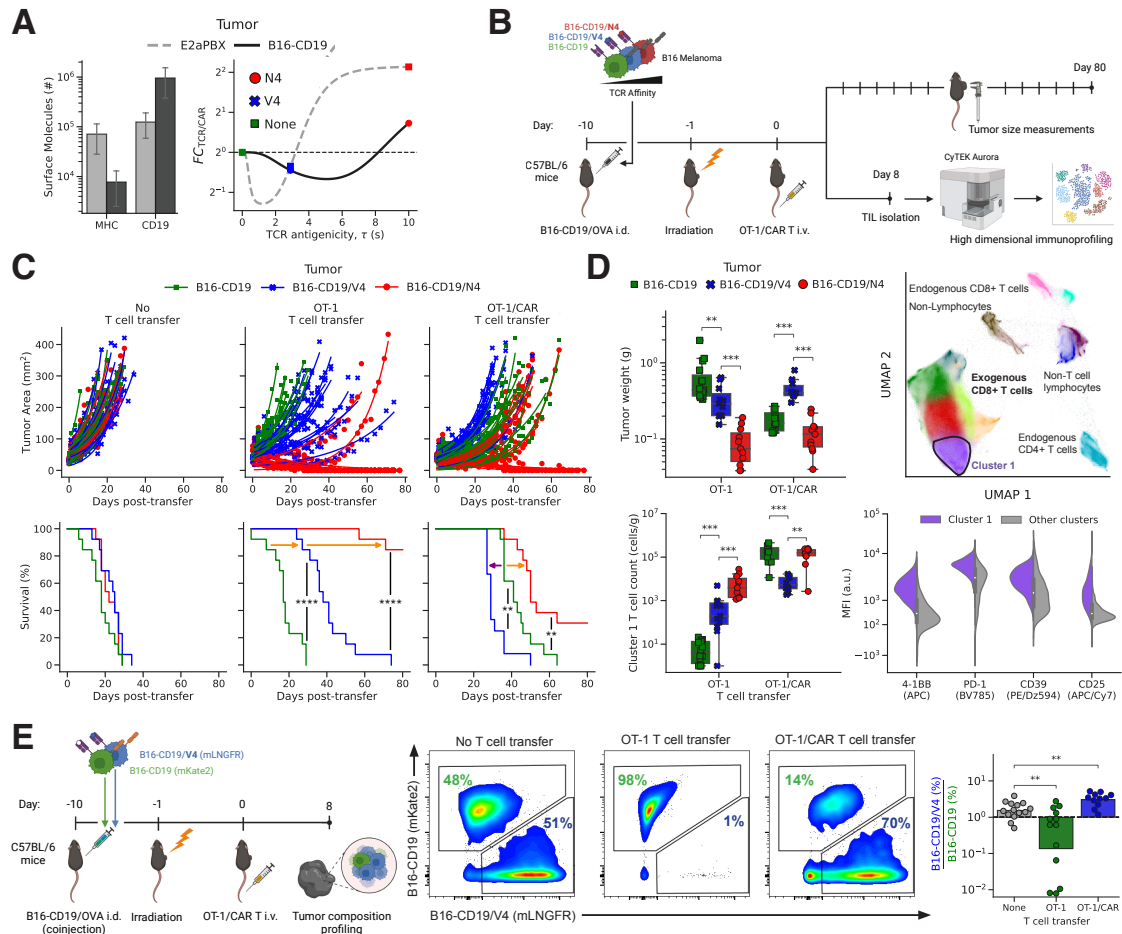


Figure 4.11: “TCR/CAR crosstalk shapes CAR T cell responses in solid tumors. (A) Baseline densities of MHC (TCR ligand) and murine CD19 (CAR ligand) in murine E2aPBX leukemia and murine CD19-transduced B16 melanoma cell lines (left) and model predictions for TCR/CAR crosstalk against B16-CD19 expressing weak (V4) and strong (N4) affinity OVA antigens (right) are presented. **(B)** Experimental scheme to evaluate TCR/CAR crosstalk in OT-1/CAR T cells in response to dual antigen B16-CD19/OVA melanoma cells. **(C)** Tumor growth (top) and survival curves (bottom) for treated groups show TCR-mediated antagonism of CAR T cell responses to melanoma cells ($n = 9\text{--}13$ mice per group). **(D)** Tumor weights at Day 8 after OT-1/CAR T-cell transfer into mice bearing melanoma cells expressing no TCR antigen, weak (V4, blue) or strong (N4, red) TCR antigens (top left). UMAP plot of 3,514,537 tumor-infiltrating leukocytes colored by cluster membership (16 clusters, top right). Leukocyte Cluster 1 frequency (bottom left) anti-correlated with tumor weight. Its phenotype (bottom right) revealed a tumor-reactive population of CD8⁺ T-cells ($n = 10\text{--}15$ mice per group). **(E)** Experimental scheme (left), representative FACS plots (center), and overall fold changes in cell frequencies (right) show specific killing of B16-CD19 melanoma cells over B16-CD19/V4 melanoma cells by OT-1/CAR T cells, due to TCR-mediated antagonism of the CAR T response. All mice were initially injected with 1:1 ratios of B16-CD19 or B16-CD19/V4 melanoma cells ($n = 11\text{--}14$ mice per group).” (*TCR/CAR antagonism*, [3], figure 4)

Figure 4.11: (Continued) “Data in (C,E) are pooled from three independent experiments. Data in (D) are pooled from two independent experiments. (** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$) Statistics in (C) were calculated using Kaplan-Meier survival estimates and (D, E) were calculated using two-tailed unpaired Student’s t-test with Holm-Bonferroni error correction applied. Data in (A) are shown as means $\pm$ standard deviation.” (TCR/CAR antagonism, [3], figure 4)

Z^{tot} :

$$T_s = mZ^{\text{tot}}(L^T, L^C, \tau^T, \tau^C) + T_0, \quad (4.21)$$

where L^T , L^C , τ^T are the ligand numbers and binding times of the CAR and TCR antigens on the tumor cells, determined from experimental calibrations, as explained earlier (section 3.5.1 and appendix D)

We computed the AKPR model output Z^{tot} for each condition in figure 4.11C: tumors with N4, V4, or no TCR antigen, and either regular T cell treatment (equivalent to $L^C = 0$ as the CAR antigen can’t be detected), 3-ITAM CAR T cell treatment, and 1-ITAM CAR T cell treatment (data in figure 4.12A). We estimated parameters m and T_0 by linear regression on all single antigen conditions – either TCR antigen only (regular T cell treatment) or CAR antigen only (tumors without TCR antigen). This fit is shown in figure 4.12B, top. Then, from the estimated m , T_0 and the previously computed Z^{tot} , we predicted the average survival time T_s for all antigen mixture conditions – 1- or 3-ITAM CAR T cell treatments of tumors expressing N4 or V4 (figure 4.12B, bottom and C). This strategy tested the ability of the model to correctly predict cross-receptor interactions, given a calibration of single antigen responses, similarly to our first attempt at modelling 6F TCRs (section 3.4), but now seeking quantitative predictions of *in vivo* results.

The linear regression of T_s as a function of Z^{tot} had a sufficient accuracy ($R^2 = 0.75$) to make predictions with a confidence interval on the same scale as real biological variability in survival time (± 20 days). The OT-1 treatment of V4-expressing tumors was under-estimated in terms of model output, potentially because the model’s TCR parameters were fitted on data from naive T cells in chapter 2, while preactivated T cells, more sensitive to weak antigens, were used here. Nonetheless, the predicted survival times lined up surprisingly well with the observations (panel B, bottom and C), especially for conditions with antagonism (V4-expressing tumors), but not for the 1-ITAM CAR treatment of N4-expressing tumors, for which the predicted survival was over-estimated. This is because the revised AKPR model was fitted to accurately capture *in vitro* results, where 1-ITAM CARs did produce larger absolute outputs than 3-ITAM CARs (blue and turquoise

lines in figure 4.3E, bottom). Hence, our model was especially accurate when predicting TCR/CAR antagonism, as we had observed even for the *in vitro* fits (figure 4.3E, top).

We performed the same survival prediction exercise for mice injected with E2aPBX lymphoma cells, with corresponding ligand abundances in the revised model. However, the mouse-to-mouse variability was greater than for the B16 melanoma experiments, and the linear regression was less accurate ($R^2 < 0.3$), leading to poor confidence on the model predictions (prediction interval on the order of ± 40 days). Still, the fact that a linear regression between Z^{tot} and T_s was possible for B16 tumors, and to some extent for E2aPBX tumors, was highly surprising. Together, the B16 and E2aPBX results indicate that modelling early receptor interaction kinetics can predict the extent of *in vivo* immune responses qualitatively (antagonism or enhancement). In some cases, we can generate quantitative predictions, but on such long time scales, other biological processes start having a significant impact after the initial activation. In particular, the short mouse survival after 1-ITAM CAR treatments of N4 tumors compared to the very strong enhancement seen *in vitro*, which is the main outlier compared to model predictions in figure 4.12C, could be attributed to rapid exhaustion of the over-stimulated CAR T cells [292], or to upregulation of other anti-inflammatory feedbacks like regulatory T cells [273] or IL-10 production [173].

Using the B16 tumor line, we wanted to ensure that CAR treatment outcomes have a real biological connection to the single-cell receptor-ligand interactions we modelled. Our collaborators verified experimentally that, even *in vivo*, TCR/CAR antagonism occurs because of interactions between individual cells and their receptors, rather than through some systemic feedback, *e.g.*, cytokines.

“We therefore tested whether the observed TCR-mediated CAR antagonism would limit the specific killing of tumor cells in heterogeneous tumor populations. We co-injected mice with a 1:1 ratio of melanoma cells expressing CD19 but no peptide and melanoma cells expressing both CD19 and the weakly antigenic V4 peptide. We then treated the mice with OT-1 or OT-1/CAR T cells, and analyzed tumor cell composition eight days after T cell adoptive transfer (Figure 4.11E, left). As expected, treatment of mice with OT-1 T cells without a CAR resulted in the preferential elimination of tumor cells expressing both the CAR ligand and the TCR antagonist V4 over tumor cells expressing the CAR ligand alone. In contrast, treatment with OT-1/CAR T cells resulted in significantly higher killing of melanoma cells expressing the CAR ligand alone than the melanoma cells expressing the CAR ligand alongside the TCR antagonist V4 (Figure 4.11E, center and right). These results demonstrate

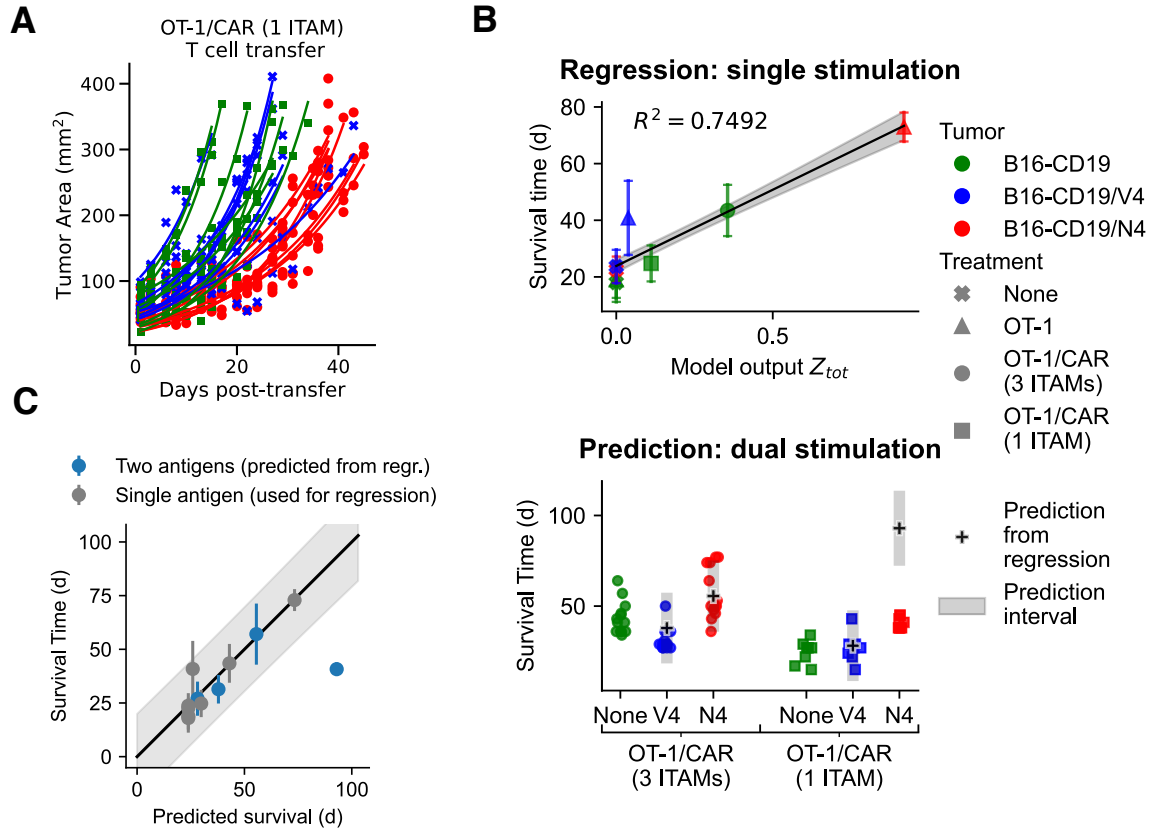


Figure 4.12: Prediction of *in vivo* survival from the cross-receptor model. (A) Tumor growth for mice groups treated by 1-ITAM CAR, OT-1 TCR T cells, showing reduced TCR-mediated antagonism of CAR T cell responses ($n = 9\text{--}13$ mice per group). Legend as in (B). (B) (Top) Linear regression of experimental survival time as a function of model output for each tumor and treatment condition where T cells have stimulation by only one antigen (either TCR or CAR). Data shown as average $\pm$ standard deviation across mice of each group. The shaded area is the 95 % confidence interval on the linear fit based on Student's t-distribution [291, eq. 8.44]. (Bottom) Prediction of survival time for tumor conditions with dual receptor stimulation (TCR and CAR), using the fitted linear regression and the pre-computed model output Z^{th} . The shaded area for each prediction is the 95 % prediction interval based on Student's t-distribution [291, eq. 8.45]. (C) Comparison of real survival times and predictions from the model output and linear regression (fitted on the grey points, which are single stimulation conditions). Data shown as average $\pm$ standard deviation across mice of each group. The shaded area is the 95 % prediction interval. The outlier point is 1-ITAM CAR treatment of tumors expressing the CAR antigen and the strong TCR agonist N4. The model (backed up by the *in vitro* data) predicts strong activation and thus good treatment efficacy, but the observed survival is poor; as explained in the text, T cell death and exhaustion or anti-inflammatory feedbacks might explain the discrepancy.

the potential for TCR-mediated antagonism to limit the anti-tumor cytotoxicity of CAR T cells on a granular, cell-by-cell basis, shaping tumor composition and impairing immunotherapeutic efficacy. Conversely, this finding raised the possibility that triggering TCR-mediated CAR antagonism could be beneficial when T cells interact with healthy cells, as it would potentially allow for a reduction in severe adverse CAR events by limiting on-target/off-tumor toxicity against CAR antigens on healthy tissues [267, 293, 294].” (*TCR/CAR antagonism*, [3])

From the theoretical perspective, this finding explains why receptor signaling models successfully capture cytokine dynamics and even mouse survival: these responses are fundamentally determined by cell-cell interactions, and slower global processes follow suit. In particular, within a fixed immunological context (APC type, *in vivo* or *in vitro*, etc.), our TCR/CAR cross-receptor inhibition model can establish accurate relative comparisons (such as the FC ratio) between the responses to different antigen mixtures. In future applications, the model could thus serve to anticipate tumor immune editing by CAR T cells, based on the unique TCR neoantigens presented by different cells in heterogeneous tumors. It could also shed some light on variability in treatment outcomes, by quantifying the possible “background effect” of antagonistic TCR neoantigens unique to each patient. But here, we have followed another promising direction: we showed how the revised AKPR model can be used as an *in silico* tool to screen antigen databases and engineer receptor combinations that help maximizing CAR T cell accuracy.

4.4.3 Prevalence of TCR/CAR antagonism in a mutated peptide library

“To quantitatively assess how frequently TCR-mediated antagonism might shape CAR T cell responses in tumors as well as healthy tissues, we leveraged our mathematical model of TCR/CAR crosstalk (Figure 4.3). In particular, we estimated, for a given TCR agonist (p^{Ag}), the prevalence of antagonists among altered peptide ligands (p^{APL}) generated by single amino acid mutations of p^{Ag} . This is particularly significant if p^{Ag} is a tumor neoantigen, *i.e.*, a mutated self-antigen, because its set of p^{APL} would likely contain its potential self-counterpart (Figure 4.13A, left and center). We hypothesized that moderately strong agonists would have the highest chance of being derived from their self-counterparts that are antagonistic to begin with (Figure 4.13A, right). To test this, we used a comprehensive mutagenesis dataset [258] of three distinct human antigens associated with seven TCR clones (appendix section D.10). We input the measured antigen strengths (EC_{50} s) for all TCR-antigen combinations

(Figures 4.13B–D) into our mathematical model (Figure 4.13E) to derive predictions for $FC_{\text{TCR/CAR}}$ (Figures 4.13F–G). We then binned p^{APL} into three categories, based on their predicted $FC_{\text{TCR/CAR}}$: 1) antagonists ($FC < 0.5$); 2) null peptides ($0.5 < FC < 2$), and 3) enhancers ($FC > 2$) (Figure 4.13H) and ordered TCRs by their agonist antigen EC_{50} (Figure 4.13I). We found that between 17 and 80 % of p^{APL} across all three antigens are potentially antagonistic, with p^{Ag} agonists of intermediate strength producing the highest percentage of p^{APL} antagonists (Figures 4.13J–K). These data confirm our hypothesis that intermediate strength neoantigens (p^{Ag}) have the highest probability of being derived from an antagonistic self-antigen (p^{APL}).” (*TCR/CAR antagonism*, [3])

This analysis illustrates how our model of CAR T cell chemical sensing can, from a single measure of TCR antigen affinity, predict the level of cross-receptor inhibition or enhancement. This creates an efficient pipeline to screen T cell receptor affinity databases (in terms of antigen EC_{50} s) in search of TCR-antigen pairs that can optimally enhance tumor elimination, or antagonize CAR T cell toxicity against normal cells. In fact, our analysis suggests that both improvements can be realized with a single receptor, since neoantigens (p^{Ag}) that are good agonists of a given TCR frequently have self counterparts (p^{APL}) that can antagonize CAR T cells.

“These predictions illustrate the ubiquitous, yet variable, nature of TCR antagonism and support the possibility of harnessing antagonism to engineer a transfer of antigen discrimination – self vs. tumor neoantigen – from TCR to CAR.” (*TCR/CAR antagonism*, [3])

4.4.4 Human CAR T cells

Before going after a specific preclinical application of TCR/CAR antagonism, with our collaborators, we

“explored whether TCR-mediated CAR antagonism could be induced in a pre-clinical setting by transducing a neoantigen-specific TCR into human CAR T cells. We engineered human $CD8^+$ T cells to express a human TCR specific for a cancer-testis antigen, and, consistent with our predictions, found that most variants of the wild-type antigen acted as weak agonists. According to our mathematical model, we

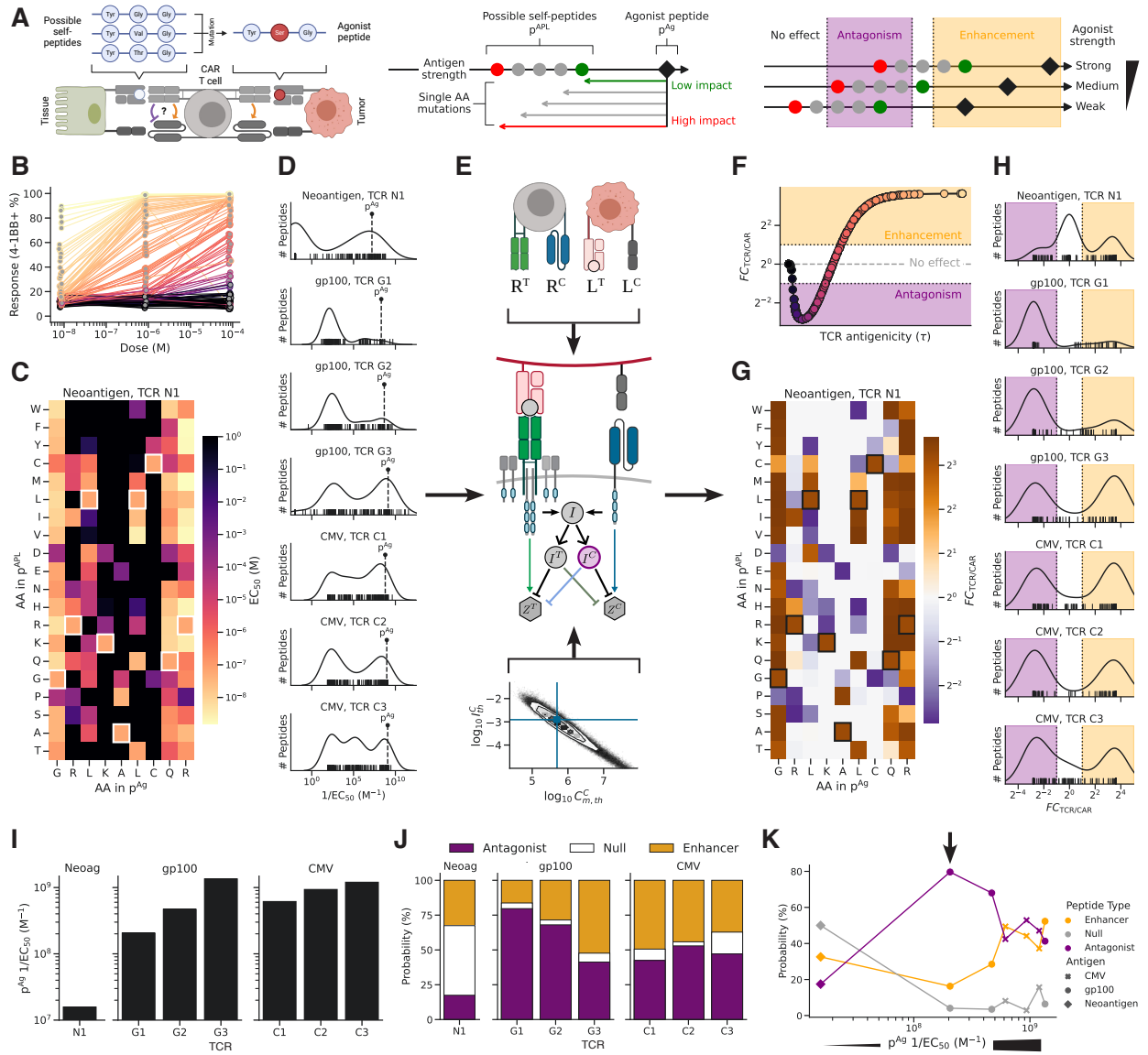


Figure 4.13: “Universality of TCR-driven antagonism of CARs. (A) Theoretical scheme to evaluate the probability of tumor neoantigens to have an antagonistic self counterpart. **(B)** Dose response curves for 7 TCR responding to 1204 altered peptide ligands (from Łuksza *et al.* [258]). **(C)** Representative EC₅₀ for all single amino acid mutations of the neoantigen. **(D)** Distributions of EC₅₀s across all seven cloned TCRs. Dashed line represents the EC₅₀ of the agonist antigen p^{Ag} used to isolate the corresponding TCR. **(E)** Schematic outlining the input parameters for the antagonism prediction pipeline: ligand/receptor data for T cell/APC system of interest (top) and previously determined MCMC parameter data (bottom), were passed to the model (middle), to estimate the magnitude of the TCR/CAR crosstalk. **(F)** $FC_{\text{TCR/CAR}}$ as a function of TCR antigenicity (1/EC₅₀s) for antigens analyzed in (C). **(G)** Representative predicted $FC_{\text{TCR/CAR}}$ for all single amino acid mutations of the neoantigen activating TCR N1. **(H)** Distributions of TCR/CAR interaction fold changes (FC) across all seven characterized TCRs.” (*TCR/CAR antagonism*, [3], figure 5) (See next page for continued caption)

Figure 4.13: (Continued) “(I) EC_{50} s for p^{Ag} for each TCR. (J) Percentage of peptide variants p^{APL} predicted to antagonize, enhance, or have no effect on CAR activation for each p^{Ag} /TCR pair. (K) Average percentage of p^{APL} predicted to antagonize, enhance, or have no effect on CAR T cell activation as a function of the strength of p^{Ag} .” (*TCR/CAR antagonism*, [3], figure 5)

predicted that these variant antigens would produce antagonism in a dual TCR/CAR system, and confirmed this experimentally (Figure 4.14A–B), thus generalizing our findings from murine to human cells (Figures 4.2 and 3.12).” (*TCR/CAR antagonism*, [3])

“To get a semi-quantitative match between PLC γ immunophenotyping quantification and model curves in figure 4.14, we had to take into account the limited dynamic range of Western blots. We noticed that the TCR antigen 9C produced no visible band, while 9V produced a weak band. We therefore added a constant background value Z_0 to the total model output Z^{tot} , equal to $Z^T \left(\frac{1}{2}(\tau_{9C}^T + \tau_{9V}^C) \right)$ (i.e. output of an antigen halfway between 9C – no visible output – and 9V – some visible output). Moreover, we took into account the very strong CAR signal in this proximal readout (compared to cytokines), by dividing the threshold Θ^C on C_{NC}^C in Z^C by a factor of 20. These adjustments do not alter the qualitative prediction from our model that lower agonist quantities are antagonized more strongly – see, for instance, the decreasing antagonism with increasing agonist concentration in figures 3.12I and 4.3E.” (*TCR/CAR antagonism*, [3], SI)

“We further quantified TCR/CAR crosstalk in this human system by evaluating microscopic colocalization and downstream signaling responses (Figures 4.14C–D). In particular, we measured the activation of the phospholipase C gamma subunit 1 (PLC- γ 1), a critical downstream regulator [295, 268] of both TCR and CAR stimulation. Our model predicted and our experiments confirmed that decreased CAR antigen density, a common mode of immunotherapeutic escape in tumors [296], resulted in increased antagonism (Figure 4.14D). Taken together, our modeling and experimental data demonstrate that TCR-mediated CAR antagonism can be universally induced in human CAR T cells by co-expressing engineered TCRs targeting a weak TCR agonist/antagonist on tumor cells.” (*TCR/CAR antagonism*, [3])

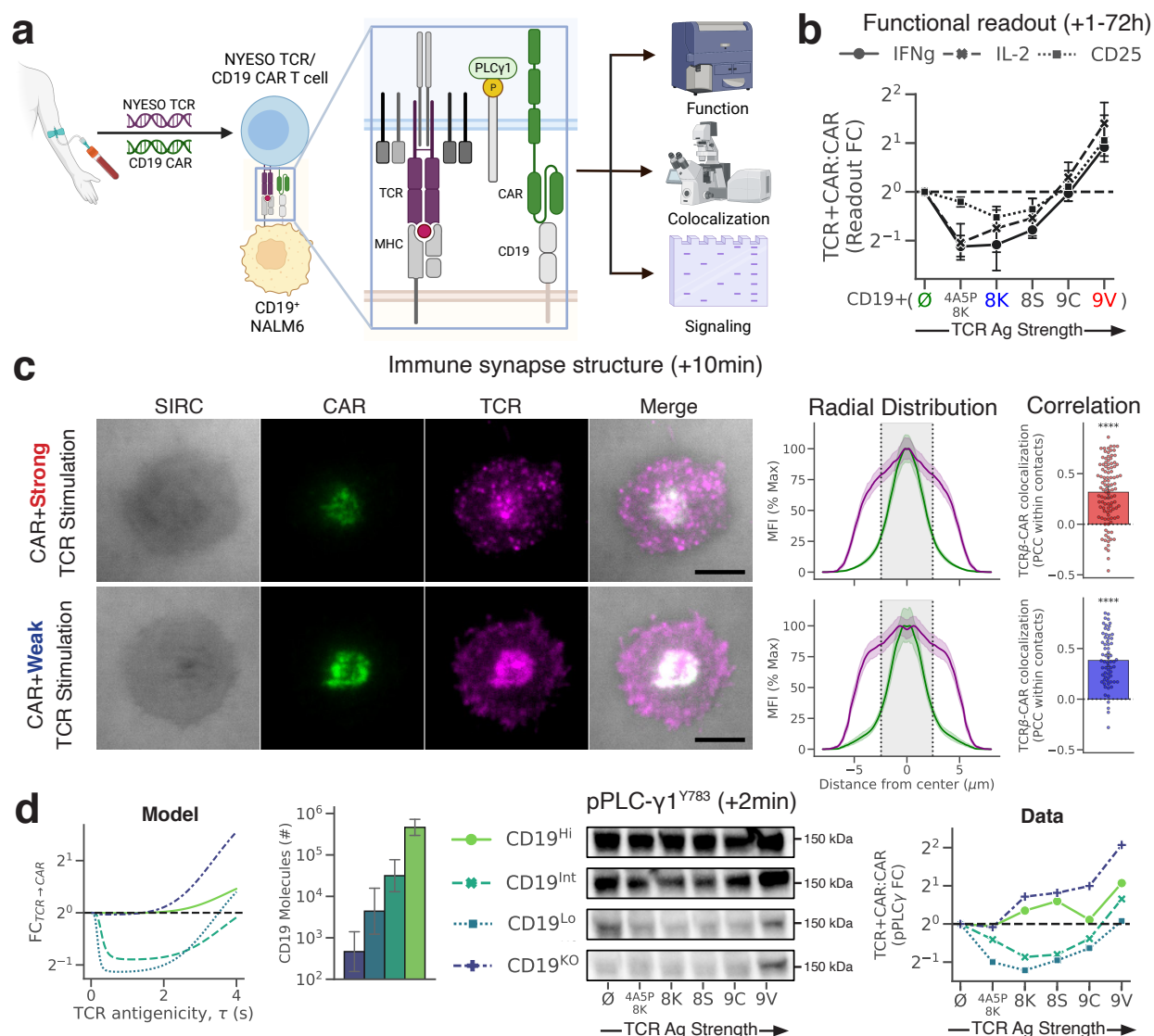


Figure 4.14: “Human CAR T cell activity can be both antagonized and enhanced by a co-transduced TCR. (A) Schematic of the generation of dual NYESO TCR/CD19 CAR T cells and interactions with targets on leukemic cells. (B) Overall TCR/CAR crosstalk quantification across IFN- γ and IL-2 functional outputs demonstrates the functional significance of TCR/CAR antagonism in the human NYESO TCR/CD19 CAR system ($n = 4$ biological replicates). (C) Representative TIRF images (left) and radial distributions (center) showing the accumulation of TCRs and CARs in the center of the immune synapse when co-engaged with their respective ligands ($n = 2$ independent experiments, with 64-108 cells per stimulation). Pixel-by-pixel correlation of receptors in TIRF images revealing a moderate level of CAR and TCR colocalization (right). Grey region represents the location of the CAR mask used during correlations.” (TCR/CAR antagonism, [3], figure S5) (See next page for continued caption)

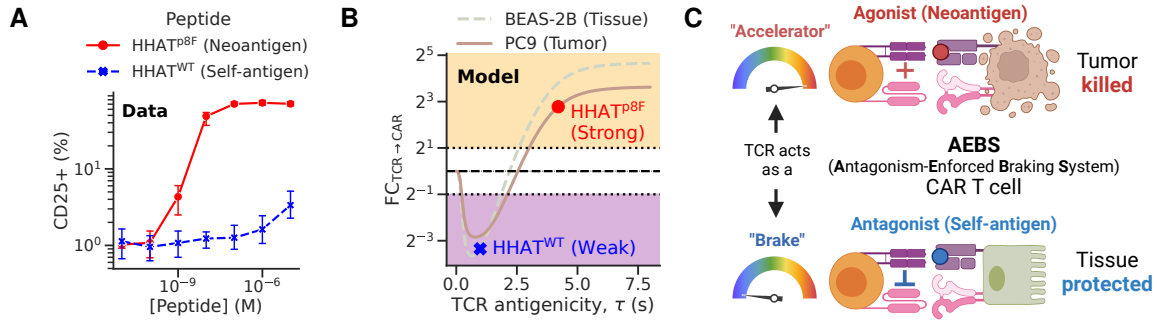
Figure 4.14: (Continued) “(D) Four HLA-A*02:01⁺ NALM6 leukemia cell lines with different CD19 densities (center left) and NYESO peptides were used to stimulate NYESO TCR/CD19 CAR T cells for 2 minutes. Representative immunoblots (center right) and quantification (right) of phosphorylated PLC- γ 1 in T cells activated via their TCR and CAR with different peptide variants (strengths indicated) and CD19 levels, respectively. Experimental measurements are consistent with mathematical model predictions (left). Data are representative of two independent experiments. Data in (B, C) are shown as means $\pm$ 95% CI. Data in (D) are shown as means $\pm$ standard deviations. Biological replicates were performed using T cells from different blood donors.” (*TCR/CAR antagonism*, [3], figure S5)

4.4.5 Application of the model to TCR/CAR engineering

The previous two subsections have shown, through model predictions verified experimentally, that TCR self antigen and neoantigen pairs can antagonize and enhance CAR responses (4.4.3), respectively, and that human CAR T cells also exhibit these cross-receptor effects (4.4.4). From these findings emerged a way to combine the recognition paradigms of CARs – strong affinity for antigens expressed more by tumors – and TCRs – amplification of small affinity differences between tumor neoantigens and healthy cell self antigens – to sharpen CAR T cell precision. Our objective was to minimize the toxicity of CAR T cells against healthy tissues expressing non-zero levels of CAR antigens. This “on-target, off-tumor” (OTOT) toxicity creates severe side effects in patients, sometimes leading to organ failures, cytokine storms, and death [266, 267]. This toxicity, as mentioned in section 4.1.3, represents one of the major barriers limiting the generalization of CAR treatments.

Our design strategy was to endow CAR T cells with a TCR which is strongly activated by a neoantigen on tumor cells, while the corresponding (non-mutated) self antigen on healthy tissues maximally antagonizes CAR responses. This combination transfers the TCR’s specificity in distinguishing tumors from healthy cells to the CAR T cell, while preserving the CAR’s strong stimulation by over-expressed tumor surface molecules (*e.g.*, CD19). In other words, the TCR and CAR act as a NOT gate for a self antigen plus CAR antigen mixture, yet as an OR gate for the neoantigen plus CAR antigen combination. Against tumors, the CAR triggers the T cell response unimpeded, whereas the TCR acts, via antagonism, as a brake against healthy cells (figure 4.15A). Consequently, we termed this design an “Antagonism-Enforced Braking System” (AEBS) for CAR T cells.

“Accordingly, we paired a CAR with a transduced TCR that binds strongly to a mutated neoantigen on tumor cells but only weakly to its native self-antigen on healthy tissues. We manufactured these dual TCR/CAR T cells by transducing three com-



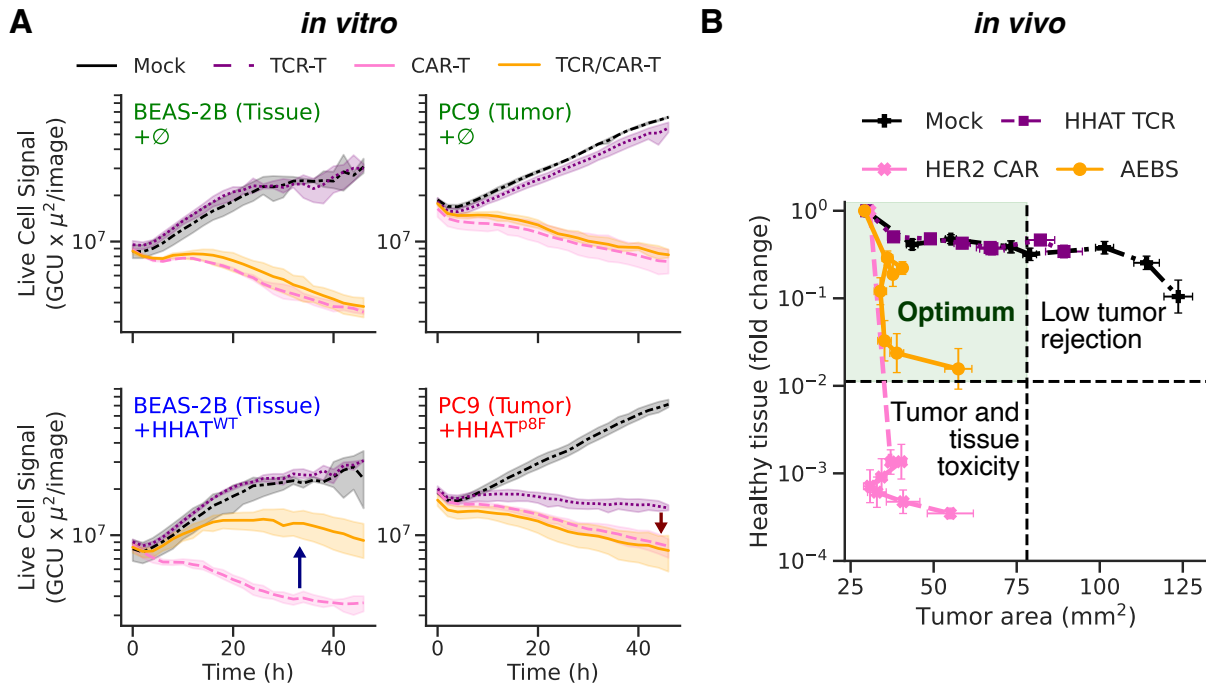


Figure 4.16: Experimental proof of the AEBS concept. (A) “Cell growth curves [...] for GFP-tagged target cells, monitored as green calibrated units (GCU), show lower cytotoxicity of dual TCR/CAR T cells against healthy tissue compared to conventional CAR T cells (upward arrow) and increased cytotoxicity against tumors compared to TCR T cells (downward arrow, $n = 5$ biological replicates).” (*TCR/CAR antagonism*, [3], figure 6) (B) “[T]wo-dimensional trajectories of healthy tissue abundance and tumor growth over time [...] demonstrate the minimal on-target/off-tumor toxicity of AEBS CAR T cells against healthy tissue compared to HER2 CAR T cells and increased anti-tumor toxicity compared to HHAT TCR T cells ($n = 20 - 29$ mice per group). Data in (B) are pooled from four sets of experiments with individual human blood donors. (**** $p < 0.0001$) Statistics in (B) were calculated using two-tailed unpaired Student’s t-test with Holm-Bonferroni error correction applied. Data in (C) are shown as means $\pm 95\%$ CI.” (*TCR/CAR antagonism*, [3], figure 7)

effector molecules by dual TCR/CAR T cells as compared to conventional CAR T cells, and anti-tumor cytotoxicity was maintained (Figure 4.16A, right). Thus, the TCR in our dual TCR/CAR T cells operates as an ‘Antagonism-Enforced Braking System (AEBS)’, enhancing CAR-mediated anti-tumor activity while simultaneously antagonizing CAR activity against healthy tissues (Figure 4.15C), thereby preventing OTOT toxicity.” (*TCR/CAR antagonism*, [3])

These *in vitro* findings confirmed the prediction from our TCR/CAR interaction model that weakly stimulating self counterparts of TCR neoantigens are the best antagonists of CAR T cell responses against healthy tissues. These findings led our collaborators to test AEBS CAR T cells *in vivo*. They developed an experimental proof-of-concept with “humanized” mice: mice without their own immune system, in which they engrafted human lung cells (BEAS-2B) presenting the HHAT^{WT} antagonist and human lung cancer cells (PC9) presenting the HHAT^{p8F} neoantigen and the CAR antigen (HER2). They then treated these mice with either regular human CAR T cells, regular human T cells with the receptor for HHAT^{p8F}, or human AEBS T cells with dual CAR and neoantigen-specific TCR. The healthy human cells were engineered to be bioluminescent (expressing luciferase), so they could be tracked by imaging in visible light; the cancer cells were injected subcutaneously and tracked by tumor size measurement.

The results of this preclinical experiment are shown in figure 4.16B. Regular CAR T cells exhibited on-tumor and off-tumor toxicity: they slowed down tumor growth, but also depleted healthy lung cells, because the latter also present the CAR antigen (HER2), albeit at a lower concentration (pink line). Regular T cells spared healthy tissues, but were not efficient at killing tumor cells (purple line), presumably because stimulation of the TCR by the HHAT^{p8F} peptide is not strong enough to fully activate T cell cytotoxicity functions. Remarkably, AEBS CAR T cells struck an optimal balance: they retained the cytotoxic ability of regular CAR T cells against tumors, but TCR-mediated antagonism significantly reduced their toxicity against healthy tissues (orange line).

4.5 Outlook on cross-receptor interactions in CAR T cells

After using the cytokine latent space in chapter 2 to zoom into biophysical TCR signaling models in chapter 3, we have built upon this mechanistic foundation to travel back up to even higher system complexity and longer time scales. We thereby proposed a way to combine TCR and CAR recognition paradigms to reduce CAR T cell toxicity. Our ability to quantitatively fit *in vitro* T cell responses and to qualitatively predict *in vivo* responses confirmed our initial insights from antigen

encoding: early kinetics in the TCR recognition of an antigen quality continuum determine, in good part, immune responses on longer time scales.

The understanding of dual TCR/CAR chemical sensing gathered from our revised AKPR model guided our collaborators to a promising preclinical concept to improve CAR T cell safety. In future developments of the AEBS CAR T cell concept, our cross-receptor interaction model could be used as a pipeline to scan databases of TCR antigen affinities. To fully improve CAR T cell treatments, it will be crucial to identify TCRs which strike the optimal balance of antagonism from self antigens and enhancement from neoantigen counterparts. More than a barcode for T cell identity, the TCR may well underlie the different functional outcomes between CAR T cell clones [282] because of its affinity landscape for self and neoantigens on tumors and healthy tissues. Our framework could also generalize to other receptor combinations in the future.

Discussion and conclusion

Discussion

Summary

T cell sensing and signaling pose theoretical challenges left unanswered by the usual biophysical treatment of concentration sensing (section 1.1). T cells have to sense small ligand affinity differences among antigen mixtures. Their complex biochemical signaling machinery obfuscates potentially simpler quantitative principles at play – if any. As such, systems immunology needs systematic frameworks to rationalize and manipulate high-dimensional T cell responses [1].

Motivated by the theoretical and practical rewards that such a framework would bring, we attempted to develop low-dimensional models of high-dimensional T cell activation processes. Our collaborators’ robotic platform afforded us a rare opportunity to reverse-engineer T cell response parameters from robust and precise cytokine dynamics measurements. We compressed cytokine data, with little loss, to a universal two-dimensional latent space in which trajectories separated mainly according to antigen quality – a property we termed “antigen encoding” (chapter 2). Ballistic physics inspired us with an accurate model of these latent space dynamics. Model parameters describing early activation kinetics (a, θ, t_0) correlated with each other and predictive of the overall response dynamics. These parameters were primarily determined by antigen quality, with minor effects from antigen quantity; they contained 2.6 bits of information about the continuum of antigen affinities that can stimulate T cells. Most molecular or drug perturbations did not break this correlation and merely shifted responses up or down the antigenicity axis (PC1, figure 2.28F). Thus, we have obtained a connection in low dimension between the antigenic space sensed by T cells and their high-dimensional cytokine outputs, based on the effective antigen quality sensed by the TCR.

These results indicated to us – as hinted by the toy model of excitable immune dynamics in section 1.2.2 – that early antigen recognition events set the tone for the rest of a T cell response. We therefore investigated phenomenological models of cross-inhibition in TCR signaling to explain non-additive effects in cytokine responses to antigen mixtures measured in our collaborators’ robotic platform. While it may have seemed over-ambitious to model effects seen in 72-hour, non-stationary cytokine dynamics with a steady-state model of receptor activation, we managed to set

up direct quantitative fits by focusing on relative response magnitudes, simplifying data and model outputs down to a one-dimensional antagonism ratio, $FC_{\text{TCR}/\text{TCR}}$ (chapter 3) or $FC_{\text{TCR}/\text{CAR}}$ (chapter 4). Performing MCMC simulations to estimate model parameters allowed us to revise the classical AKPR model, simplifying it further before adding back key ingredients, namely a nonlinear function in the cross-receptor inhibitory coupling (in rate $\psi(I)$) and an asymmetry between TCR and CAR susceptibility to this inhibition (encoded in parameters γ_C^T and γ_T^C).

We validated the model by predicting antagonism and enhancement for various altered receptor ITAMs. We even extended our predictions as far as *in vivo* time scales, with reasonable agreement but more variability, likely due to slow feedback processes (*e.g.*, tumor growth and exhaustion) modifying the outcome of initial receptor-ligand interactions. Still, our revised AKPR model represents a drastic dimensional reduction of the TCR signalosome, the latter requiring in principle hundreds of equations just to describe its known biochemical interactions [116]. Our results reinforce the idea that TCR recognition of antigen quality boils down to kinetic proofreading steps plus an inhibitory coupling between receptors; our collaborators’ precise datasets (capturing changes less than two-fold in cytokine levels) provided enough detail to discriminate between different versions of this inhibitory interaction. Our quantitative understanding of an emergent simplicity in TCR antigen recognition inspired us to develop a new CAR T cell design where antagonism is optimized, with assistance from the model, to protect healthy cells against off-tumor toxicity

Limitations, future work, remaining challenges

Although our results are encouraging, we are far from having fully elucidated the quantitative principles of TCR sensing and signaling. To moderate our claims, we now highlight a few limitations of our work, and suggest potential solutions that would strengthen our results if carried out successfully. We also highlight important aspects of T cell activation that our theoretical work has not addressed.

The full cytokine code might multiplex parameters beyond antigen quality

In chapter 2, our cytokine latent space was built from datasets including seven cytokines (quickly reduced to five, as two had no signal). These cytokines were selected based on their anticipated biological relevance, and spanned enough dimensions already to complicate interpretation, but they did not account for the full range of chemical signals emitted by T cells. The communication channel components that we overlooked might be used by T cells to transmit information about other immune parameters: antigen quantity, antigen presenting cell (APC) types, cell numbers, etc.

To address this concern, our collaborators have upgraded their robotic platform and flow cytometry equipment to acquire panels of ~ 20 cytokines and single-cell surface marker expression in parallel (by multiplexing bead sizes and color channels). Preliminary analyses of these larger datasets still show that antigen quality is the main parameter driving the response, down to the single-cell level.

It will however be interesting to attempt extracting other immune parameters from these larger datasets. There may be other latent spaces, orthogonal to the one we found for antigen quality, in which new information is encoded. For instance, T cells also sense combinations of co-stimulatory ligands on APCs and modulate cytokine response levels and dynamics as a result – differently in CAR and TCR T cells, too [79]. Hence, the rotations observed in our 2D latent space as APC types are changed (figures B.6–2.30) could reflect large changes in other, intersecting latent spaces. These spaces may be even more important with $CD4^+$ T cells, which depend more on antigen quantity to perform their helper roles in, *e.g.*, B cell germinal centers [299].

Overfitting, systematic model search, and positive feedback

When seeking a quantitative agreement between data and models with free parameters, overfitting is always a concern. We took various measures to avoid that issue: testing neural network robustness against various preprocessing choices and comparing with other dimensional reduction methods (chap. 2), estimating parameters of the classical AKPR model based on single antigen responses of 6Y versus 6F T cells to predict antagonism (chap. 3), cross-validating the revised AKPR model with predictions compared to separate datasets of T cells with altered receptors (chap. 4), and so on. These precautions ensure that our broad conclusions – about how the continuous spectrum of antigen qualities drives distinct cytokine responses and produces antagonism effects in mixtures – are not merely fitting anomalies.

However, we cannot completely exclude the possibility that some details of our models have been overfitted or poorly constrained by the experimental data. In chapter 2, the channel capacity we calculated is based on one large experiment, combining several technical replicates with cells from several mice; however, combining datasets from different days would likely have reduced the channel capacity because of batch effects and of experiment-to-experiment variability. In fact, part of this variability may be a feature rather than a bug of the immune system, a system primed to respond very sensitively to all kinds of challenges⁷. An important future theoretical task will be to anticipate these biological variations – or at least estimate their full statistics – based on the broad,

⁷Credits to William Bialek for proposing this idea in a discussion I was lucky enough to have with him.

long-tailed distributions of TCR signaling molecule basal levels in T cells of a population [136]. Acquiring enough samples to do so remains a technical challenge, even with robots.

In chapter 3, the revised AKPR model fits are not perfect; in particular, they do not capture the rightwards shift of the antagonism curve as the antagonist quantity L_2 is decreased; the antagonism amplitude still decreases too much as L_2 is lowered (red line, 10 pM agonist, figure 3.12I). In chapter 4, on the contrary, the extension of the revised model to CAR T cells seems underdetermined, since several values of integer parameters k , m , f yielded similar fits. This might be because the Hill activation function we are fitting for the outputs Z^T and Z^C could turn several reasonable choices of C_N^ρ receptor outputs into the right overall FC antagonism curve, with the fitted thresholds, τ_c^T and τ_c^C , setting the exact crossover point.

These details do not detract from the revised AKPR model’s ability to achieve quantitative accuracy in different immune contexts: it still is a useful prediction tool. However, we could be missing key mechanisms in our models. One likely candidate is a positive feedback loop, thought to be mediated by phosphorylated ERK downstream of the TCRs, preventing receptor deactivation once the output crosses a certain threshold [116]. Striving for simple models, later studies found this positive feedback unnecessary to capture antagonism qualitatively in the classical AKPR model [129, 132]. With our collaborators’ robotic platform, we may now have reached the level of quantitative precision where this positive feedback becomes noticeable.

We have, in fact, tried to include such an additional interaction in several versions of AKPR, without significantly improving the model’s behavior (discussion in section 3.5.3). One issue is that we have been manually searching for revised models, which is certainly not the most efficient or systematic approach. Algorithms exist to evolve or simplify biochemical networks performing a given function, or to compress large models into simpler ones [137, 159]. They were however not easily applicable to our problem, since we do not have a qualitative biological function that we want the model to perform, but rather a relatively small (by statistical standards; large by immunological standards) dataset to fit. Instead, we could perform a systematic search in model space where we include or remove a large set of possible network interactions (positive and negative feedbacks, etc.), similar to the methods used in [64, 300]. Alternatively, following the approach in [301], we could apply MCMC to a large network including many possible reactions, with strong exponential priors on these reaction parameters, to enforce regularization and thus keep only the essential interactions.

To properly identify an eventual positive feedback from antigen mixture datasets, we should fit, in addition to antagonism, not only enhancement but true *synergy* effects, *i.e.*, a combined output larger than the sum of either single receptor stimulation output. This could be quantified by a second dimensionless ratio,

$$FC_{\text{synergy}} = \frac{\text{Out}(\text{TCR} + \text{CAR})}{\text{Out}(\text{TCR}) + \text{Out}(\text{CAR})} ,$$

larger than 1 for true synergy, and below 1 for sub-additive responses (which are different from antagonism, defined as a response below the agonist-only output). We did not compute this second FC quantity in our thesis, because we did not have the TCR-only condition in all experimental datasets, but future experiments should include it. Lastly, if traditional biochemical networks in the AKPR framework ultimately fail to capture the range of antagonism and enhancement seen in datasets, the most practical solution would be to resort to machine learning techniques to approximate inhibitory and positive feedback modules, using recent tools like automatic differentiation and universal differential equations [302]. This would not be the most interpretable solution, but would at least enable a complete quantitative characterization of TCR sensing and signaling.

Explicitly connecting time scales to capture dose response curves

In chapter 3, it was interesting to note that we could recover absolute discrimination (output depending only on quality τ while flat as a function of quantity L) in the model curves for single antigens, after fitting the model on antagonism in binary antigen mixtures. In practice, we observed the converse statement of the “spandrel” theorem proven in [139]. However, these model output curves (C_N versus L) in figure 3.18 do not quite capture the shape of experimental dose response curves, which have very sharp increases as a function of antigen quantity before reaching a plateau resembling absolute discrimination; see figure 5.1 below.

This sharpness could partly be due to using blast T cells in the experiment (pre-activated and expanded), as opposed to naive cells. Still, we should properly predict these sharp response curves, which will require improvements to our framework. The cause may be that we neglected positive feedback loops, as mentioned in the previous section, which could sharpen the response around the threshold. Otherwise, the problem might be situated downstream of TCR recognition. One solution would be to use the AKPR model in its stochastic version, imagine a population of T cells sensing antigens, and compute the fraction of T cells reaching some activation threshold before starting to produce cytokines. A more phenomenological solution would be to apply a sharp Hill

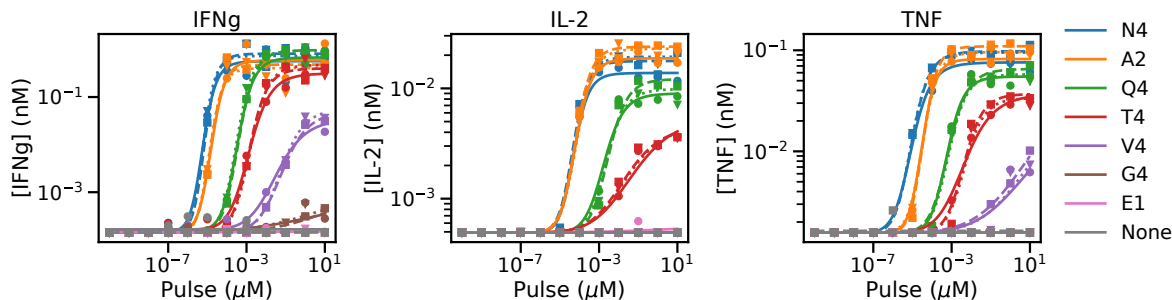


Figure 5.1: Cytokine dose response curves to OT-1 peptides. Data from Sooraj Achar, generated for *TCR/CAR antagonism*, [3].

activation function on C_N – similar to the output function we used in Z^T and Z^C for TCR/CAR antagonism, but this time fitting the Hill thresholds directly on dose response data.

The next step to capture T cell responses in absolute values, for direct data-model comparisons, would be to find a low-dimensional model of the T cell signaling pathways downstream of TCR signaling. Such an end-to-end model of T cell responses would explicitly connect experimental readouts and receptor outputs. We would no longer be restricted to fitting relative fold-changes. We have relied on an empirical correlation between short (TCR sensing of antigen quality) and long time scales (cytokine production and consumption), but we have not explained it entirely. A mathematical model of this correlation would refine our ability to predict and manipulate it. In fact, in preliminary stages of this work, we had attempted to achieve this goal using the IL-2 model of Voisinne *et al.* [105], finding quasi-static approximations to reduce its dimensionality without changing its predictions (not shown in this thesis). However, that model relied only on antigen decay to explain IL-2 decrease after 24 h, leading to a scaling of cytokine time integrals with the initial antigen quantity – very different from what we saw in chapter 2. Further work will be required in that direction.

Conclusion

Thesis objectives and potential impacts

In summary, there are still many remaining challenges regarding T cell sensing and signaling. Our work on antigen encoding and revised AKPR models was nonetheless a step in the right direction, towards a better quantitative understanding of this uniquely complex and sensitive biological system. We managed to reduce the dimensionality of T cell cytokine outputs without losing information about the antigen sensing performed by the TCR. We set up a MCMC method to fit

phenomenological models of TCR signaling to robot-acquired datasets, comparing the model and data through a dimensionless scalar ratio (the antagonism fold-change FC). The model predictions were accurate and interpretable enough to guide the design of improved CAR T cells.

By seeking quantitative agreement between theory and experiment, by using computational physics methods like MCMC simulations and the Blahut-Arimoto algorithm, we have attempted to bring models of immune sensing and signaling to the same level of accuracy as other fields of physics concerned with complex systems and using analogous techniques, for instance astrophysics [303]. We are still far from the lofty goal of unearthing immunity’s fundamental laws, but latent spaces of immune dynamics may allow the establishment of deeper theoretical connections with the notions of potential and fitness landscapes, which are increasingly common in biological physics: for development [147], for mutations in protein sequences [156], for microbes in evolving populations [304], and for the functions of microbial communities [305]. Moreover, on the applied side, our collaboration with the Altan-Bonnet and Taylor labs hopefully shows that even phenomenological, low-dimensional models can boost the engineering of precise immunotherapies.

Epistemological remarks

Optimistic and pessimistic views of low-dimensional models

Why do we find low-dimensional representations in immunity, and in biological physics more generally? The pessimistic view would say it is due to limitations of human cognitive capabilities: we cannot comprehend overly complicated models, we need low-dimensional descriptions to feel satisfied about our understanding of something, so we come up with such simple theories even if they are wrong. Such models however have little to do with the reality of the natural world and would be at best, as James W. Black put it in his 1988 Nobel lecture, “accurate descriptions of our pathetic thinking about nature” [306] (a phrase also used by Gunawardena [152]).

At the opposite end of the spectrum, the optimistic view would be some version of realism: the biological world really has a low-dimensional underlying structure that we are uncovering in mathematical terms. This conception would be similar to Galileo’s statement, in his criticism of Ptolemy’s epicycles, that simpler mathematical theories are more satisfactory because they reveal the true structure and essence of the natural world [307, p. 508–510], [308, p. 245]. This realism also shows through Albert Einstein’s phrase saying that “[t]he eternal mystery of the world is its comprehensibility” [309], and through Eugene Wigner’s statement that “[t]he miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve” [310]. Although they are both agnostic

to why that is the case, their position imply that we actually “comprehend” the world as it is, that we uncover “laws of physics” truthful to reality when we find a successful mathematical theory, including low-dimensional models.

A Kantian view as middle ground

A Kantian view can perhaps offer a middle ground between the pessimistic and optimistic views of why low-dimensional models can describe biological systems. In the *Critique of Pure Reason* [311, 312], Kant establishes an important distinction between a phenomenon and the “thing in itself” (*Ding an sich*) that underlies it. The “thing in itself” is outside the realm of human knowledge, but it is manifested as a phenomenon through our faculty of sensibility (*Sinnlichkeit*), and interpreted through our faculty of understanding (*Verstand*). These faculties have *a priori* forms – space and time for the faculty of sensibility, twelve categories such as causality and existence for understanding – which shape all phenomena apprehended through their prism: “the object (as an object of the senses) conforms to the constitution of our faculty of intuition” [311, B xvii]. Hence, the regularities and mathematical laws that we find in natural phenomena really are in these phenomena, because the latter are pre-conditioned, filtered by the *a priori* forms of our faculties; in Kant’s words, “we ourselves bring into the appearances that order and regularity in them that we call *nature*” [311, A 126]. In formulating his theory of knowledge, Kant aimed to rationalize the possibility of deductive natural sciences like physics (in his words, of synthetic *a priori* judgments) within the world of phenomena – which is ontologically different from the “thing in itself”, but not any less real.

In this Kantian view, we think that low-dimensional models may be part of the *a priori* form of our understanding. The things in themselves, beyond the realm of scientific knowledge, could be high-dimensional, but we can only perceive, directly or through complicated machines, low-dimensional projections of them in phenomena. They are not as immediately obvious as time and space (the *a priori* forms of sensibility), but we can re-discover these low-dimensional structures, latent in our perceptions and inherent to our understanding, through careful mathematical analysis of datasets and mechanistic models. Thus, like in the pessimistic view, low-dimensional models are a reflection of our perceptive and cognitive capabilities; yet, like in the optimistic view, they are real, part of the highest level of reality that human beings have access to – phenomena. The precise status that low dimensionality could have in Kant’s theory of knowledge (a category? sub-category? a schema?) remains to be clarified, but revisiting the *Critique of Pure Reason* could be valuable to interpret and guide contemporary developments of theory in biological physics.

Bibliography

- [1] Sooraj R. Achar *et al.* “Universal antigen encoding of T cell activation from high-dimensional cytokine dynamics”. *Science*, vol. 376, no. 6595 (May 2022), 880–884.
- [2] Guillaume Gaud *et al.* “CD3 ζ ITAMs enable ligand discrimination and antagonism by inhibiting TCR signaling in response to low-affinity peptides”. *Nature Immunology*, vol. 24, no. 12 (Dec. 2023), 2121–2134.
- [3] Taisuke Kondo *et al.* “Engineering TCR-controlled fuzzy logic into CAR T cells enhances therapeutic specificity”. *[submitted]* (2024).
- [4] Thomas J. Rademaker. “Discovering Biophysical Principles in Latent Space Representations of Immune Recognition”. Ph.D. Canada – Quebec, CA: McGill University (Canada), 2021.
- [5] Jordi Piñero and Ricard Solé. “Statistical physics of liquid brains”. *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 374, no. 1774 (June 2019), 20180376.
- [6] Grégoire Altan-Bonnet. *Le système immunitaire comme cerveau liquide ?* Conférence publique. Institut d’études scientifiques de Cargèse, Aug. 2022.
- [7] A. L. Hodgkin and A. F. Huxley. “A quantitative description of membrane current and its application to conduction and excitation in nerve”. *The Journal of Physiology*, vol. 117, no. 4 (Aug. 1952), 500–544.
- [8] Matthew Zirui Tay *et al.* “The trinity of COVID-19: immunity, inflammation and intervention”. *Nature Reviews Immunology*, vol. 20, no. 6 (June 2020), 363–374.
- [9] Jonathan D. G. Jones and Jeffery L. Dangl. “The plant immune system”. *Nature*, vol. 444, no. 7117 (Nov. 2006), 323–329.
- [10] Steven A. Rosenberg *et al.* “Observations on the Systemic Administration of Autologous Lymphokine-Activated Killer Cells and Recombinant Interleukin-2 to Patients with Metastatic Cancer”. *New England Journal of Medicine*, vol. 313, no. 23 (Dec. 1985), 1485–1492.
- [11] J. C. Castle *et al.* “Exploiting the Mutanome for Tumor Vaccination”. *Cancer Research*, vol. 72, no. 5 (Mar. 2012), 1081–1091.
- [12] Elizabeth I Buchbinder and Anupam Desai. “CTLA-4 and PD-1 Pathways: Similarities, Differences, and Implications of Their Inhibition”. *American journal of clinical oncology*, vol. 39, no. 1 (Feb. 2016), 98–106.

- [13] Carl H. June *et al.* “CAR T cell immunotherapy for human cancer”. *Science*, vol. 359, no. 6382 (Mar. 2018), 1361–1365.
- [14] Emanuel Salazar-Cavazos and Grégoire Altan-Bonnet. “Engineering time-controlled immunotherapy”. *Science*, vol. 378, no. 6625 (Dec. 2022), 1175–1176.
- [15] Farooq Ahmed. “Profile of Eric D. Siggia”. *Proceedings of the National Academy of Sciences*, vol. 109, no. 15 (Apr. 2012), 5551–5552.
- [16] Erwin Schrödinger. *What is Life?* Canto Classics. Cambridge: Cambridge University Press, 2012.
- [17] William S. Bialek. *Biophysics : searching for principles*. Princeton, NJ: Princeton University Press, 2012.
- [18] Gautam Reddy *et al.* “Sector search strategies for odor trail tracking”. *Proceedings of the National Academy of Sciences*, vol. 119, no. 1 (Jan. 2022), e2107431118.
- [19] Massimo Vergassola *et al.* “‘Infotaxis’ as a strategy for searching without gradients”. *Nature*, vol. 445, no. 7126 (Jan. 2007), 406–409.
- [20] Sri Krishna *et al.* “Stem-like CD8 T cells mediate response of adoptive cell immunotherapy against human cancer”. *Science*, vol. 370, no. 6522 (Dec. 2020), 1328–1334.
- [21] Matthew J. Wither *et al.* “Antigen perception in T cells by long-term Erk and NFAT signaling dynamics”. *Proceedings of the National Academy of Sciences*, vol. 120, no. 52 (Dec. 2023), e2308366120.
- [22] Aristote. *Physique*. Paris: Garnier-Flammarion, 2000.
- [23] Lucrèce. *De la nature / De rerum natura*. GF Flammarion. 1997.
- [24] Hermann L. F. Helmholtz. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Trans. by Alexander J. Ellis. 3rd ed. Cambridge University Press, 2009.
- [25] Richard L. Doty and David G. Laing. “Psychophysical Measurement of Human Olfactory Function”. *Handbook of Olfaction and Gustation*. Hoboken, NJ: Wiley, 2015.
- [26] Alan Perelson, ed. *Theoretical Immunology, Part One*. Santa Fe: CRC Press, 1978.
- [27] Zachary Sethna *et al.* “Population variability in the generation and selection of T-cell repertoires”. *PLOS Computational Biology*, vol. 16, no. 12 (Dec. 2020). Ed. by Benny Chain, e1008394.
- [28] Martina Milighetti *et al.* “Predicting T Cell Receptor Antigen Specificity From Structural Features Derived From Homology Models of Receptor-Peptide-Major Histocompatibility Complexes”. *Frontiers in Physiology*, vol. 12 (Sept. 2021), 730908.

- [29] Andreas Mayer *et al.* “How a well-adapted immune system is organized”. *Proceedings of the National Academy of Sciences*, vol. 112, no. 19 (May 2015), 5950–5955.
- [30] H C Berg and E M Purcell. “Physics of chemoreception”. *Biophysical journal*, vol. 20, no. 2 (Nov. 1977), 193–219.
- [31] BingKan Xue *et al.* “Environment-to-phenotype mapping and adaptation strategies in varying environments”. *Proceedings of the National Academy of Sciences*, vol. 116, no. 28 (July 2019), 13847.
- [32] Pieter Rein ten Wolde *et al.* “Fundamental Limits to Cellular Sensing”. *Journal of Statistical Physics*, vol. 162, no. 5 (Mar. 2016), 1395–1424.
- [33] Gautam Reddy *et al.* “Olfactory Sensing and Navigation in Turbulent Environments”. *Annual Review of Condensed Matter Physics*, vol. 13, no. 1 (Mar. 2022), 191–213.
- [34] Antonio Celani and Massimo Vergassola. “Bacterial strategies for chemotaxis response”. *Proceedings of the National Academy of Sciences*, vol. 107, no. 4 (Jan. 2010), 1391–1396.
- [35] Jonathan Desponds *et al.* “A mechanism for hunchback promoters to readout morphogenetic positional information in less than a minute”. *eLife*, vol. 9 (July 2020). Ed. by Michael B Eisen, e49758.
- [36] William Bialek and Sima Setayeshgar. “Cooperativity, Sensitivity, and Noise in Biochemical Signaling”. *Physical Review Letters*, vol. 100, no. 25 (June 2008), 258101.
- [37] Nico van Kampen. *Stochastic Processes in Physics and Chemistry*. 3rd ed. Amsterdam: North-Holland, 2007.
- [38] C. W. (Crispin W.) Gardiner 1942-. *Stochastic methods : a handbook for the natural and social sciences*. 4th ed. Springer complexity. Berlin: Springer, 2009.
- [39] Jackson O’Brien and Arvind Murugan. “Temporal Pattern Recognition through Analog Molecular Computation”. *ACS Synthetic Biology*, vol. 8, no. 4 (Apr. 2019), 826–832.
- [40] Robert G. Endres and Ned S. Wingreen. “Maximum Likelihood and the Single Receptor”. *Physical Review Letters*, vol. 103, no. 15 (Oct. 2009), 158101.
- [41] Pankaj Mehta and David J. Schwab. “Energetic costs of cellular computation”. *Proceedings of the National Academy of Sciences*, vol. 109, no. 44 (Oct. 2012), 17978.
- [42] Monica Skoge *et al.* “Chemical Sensing by Nonequilibrium Cooperative Receptors”. *Physical Review Letters*, vol. 110, no. 24 (June 2013), 248102.
- [43] Sarah E. Harvey *et al.* “Universal energy-accuracy tradeoffs in nonequilibrium cellular sensing”. *Physical Review E*, vol. 108, no. 1 (July 2023), 014403.

- [44] J. L. Lebowitz and H. Spohn. “A Gallavotti Cohen-Type Symmetry in the Large Deviation Functional for Stochastic Dynamics”. *JOURNAL OF STATISTICAL PHYSICS*, vol. 95, no. 1/2 (1999), 333–366.
- [45] Felix Wong *et al.* “Energy-speed-accuracy relation in complex networks for biological discrimination”. *Physical Review E*, vol. 98, no. 1 (July 2018), 012420.
- [46] Thomas Cover and Joy Thomas. *Elements of Information Theory*. 2nd ed. Hoboken, NJ: Wiley, 2006.
- [47] François X. P. Bourassa. “Decoding Cytokine Dynamics With Biochemical Networks”. PhD thesis. McGill, June 2020.
- [48] Athanasios Papoulis and S. Unnikrishna Pillai. *Probability, random variables, and stochastic processes*. 4th ed. McGraw-Hill series in electrical and computer engineering. Boston: McGraw-Hill, 2002.
- [49] Gašper Tkačik *et al.* “Information flow and optimization in transcriptional regulation”. *Proceedings of the National Academy of Sciences*, vol. 105, no. 34 (Aug. 2008), 12265–12270.
- [50] Gasper Tkacik *et al.* “Information capacity of genetic regulatory elements”. *Physical Review E*, vol. 78, no. 1 Pt 1 (July 2008), 011910–011910.
- [51] Gašper Tkačik *et al.* “Optimizing information flow in small genetic networks.” *Physical Review E*, vol. 80, no. 3 (2009), 031920.
- [52] Aleksandra Walczak *et al.* “Optimizing information flow in small genetic networks. II. Feed-forward interactions”. *Physical Review E*, vol. 81, no. 4 (Apr. 2010), 041905.
- [53] Julien O. Dubuis *et al.* “Positional information, in bits”. *Proceedings of the National Academy of Sciences*, vol. 110, no. 41 (Oct. 2013), 16301.
- [54] R. Blahut. “Computation of channel capacity and rate-distortion functions”. *IEEE Transactions on Information Theory*, vol. 18, no. 4 (1972), 460–473.
- [55] Manuel Razo-Mejia *et al.* “First-principles prediction of the information processing capacity of a simple genetic circuit”. *Physical Review E*, vol. 102, no. 2 (Aug. 2020), 022404.
- [56] S. Arimoto. “An algorithm for computing the capacity of arbitrary discrete memoryless channels”. *IEEE Transactions on Information Theory*, vol. 18, no. 1 (Jan. 1972), 14–20.
- [57] Jeroen S. Van Zon *et al.* “Diffusion of Transcription Factors Can Drastically Enhance the Noise in Gene Expression”. *Biophysical Journal*, vol. 91, no. 12 (Dec. 2006), 4350–4367.
- [58] Eric D. Siggia and Massimo Vergassola. “Decisions on the fly in cellular sensory systems”. *Proceedings of the National Academy of Sciences*, vol. 110, no. 39 (Sept. 2013), E3704.

- [59] Jean-Benoît Lalanne and Paul François. “Chemodetection in fluctuating environments: Receptor coupling, buffering, and antagonism”. *Proceedings of the National Academy of Sciences*, vol. 112, no. 6 (Feb. 2015), 1898.
- [60] Yaron E. Antebi *et al.* “Combinatorial Signal Perception in the BMP Pathway”. *Cell*, vol. 170, no. 6 (2017), 1184–1196.
- [61] Pulin Li and Michael B. Elowitz. “Communication codes in developmental signaling pathways”. *Development*, vol. 146, no. 12 (June 2019), dev170977.
- [62] Paul François and Grégoire Altan-Bonnet. “The Case for Absolute Ligand Discrimination: Modeling Information Processing and Decision by Immune T Cells”. *Journal of Statistical Physics*, vol. 162, no. 5 (Mar. 2016), 1130–1152.
- [63] Heidi E. Klumpe *et al.* “The context-dependent, combinatorial logic of BMP signaling”. *Cell Systems*, vol. 13, no. 5 (May 2022), 388–407.e10.
- [64] Jesse W. Cotari *et al.* “Cell-to-Cell Variability Analysis Dissects the Plasticity of Signaling of Common γ Chain Cytokines in T Cells”. *Science Signaling*, vol. 6, no. 266 (Mar. 2013), ra17.
- [65] Léa Sta *et al.* “Tuning of cytokine signaling through imbalanced abundances of receptors and kinases”. *bioRxiv* (Oct. 2022).
- [66] Gautam Reddy *et al.* “Antagonism in olfactory receptor neurons and its implications for the perception of odor mixtures”. *eLife*, vol. 7 (Apr. 2018). Ed. by Fred Rieke, e34958.
- [67] Martín Carballo-Pacheco *et al.* “Receptor crosstalk improves concentration sensing of multiple ligands”. *Physical Review E*, vol. 99, no. 2 (Feb. 2019), 022423.
- [68] Christina J. Su *et al.* “Ligand-receptor promiscuity enables cellular addressing”. *Cell Systems*, vol. 13, no. 5 (May 2022), 408–425.e12.
- [69] Léa Sta *et al.* “Algebraic Study of Receptor-Ligand Systems: A Dose-Response Analysis”. *SIAM Journal on Applied Mathematics* (July 2023), S105–S150.
- [70] Katharina F. Sonnen *et al.* “Modulation of Phase Shift between Wnt and Notch Signaling Oscillations Controls Mesoderm Segmentation”. *Cell*, vol. 172, no. 5 (Feb. 2018), 1079–1090.e12.
- [71] Christian P. Petersen and Peter W. Reddien. “Wnt Signaling and the Polarity of the Primary Body Axis”. *Cell*, vol. 139, no. 6 (Dec. 2009), 1056–1068.
- [72] J. Raspopovic *et al.* “Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients”. *Science*, vol. 345, no. 6196 (Aug. 2014), 566.

- [73] Sarah Anhala Cepeda-Humerez *et al.* “Estimating information in time-varying signals”. *PLOS Computational Biology*, vol. 15, no. 9 (Sept. 2019), e1007290.
- [74] F.M. Burnet and F. Fenner. *The Production of Antibodies*. Influenza: A Survey of the Last 50 Years in the Light of Modern Work on the Virus of Epidemic Influenza. Macmillan, 1949.
- [75] Lucille C. Rankin and David Artis. “Beyond Host Defense: Emerging Functions of the Immune System in Regulating Complex Tissue Physiology”. *Cell*, vol. 173, no. 3 (Apr. 2018), 554–567.
- [76] John F. Brooks and Lora V. Hooper. “Interactions among microbes, the immune system, and the circadian clock”. *Seminars in Immunopathology*, vol. 42, no. 6 (Dec. 2020), 697–708.
- [77] Lora V. Hooper *et al.* “Interactions Between the Microbiota and the Immune System”. *Science*, vol. 336, no. 6086 (June 2012), 1268–1273.
- [78] Thomas Pradeu. *Philosophy of Immunology*. Elements in the Philosophy of Biology. Cambridge: Cambridge University Press, 2020.
- [79] Ashna Patel *et al.* “Regulation of temporal cytokine production by co-stimulation receptors in TCR-T cells is lost in CAR-T cells”. *bioRxiv* (Jan. 2024), 2024.02.21.581341.
- [80] Andreas Mayer *et al.* “Diversity of immune strategies explained by adaptation to pathogen statistics”. *Proceedings of the National Academy of Sciences*, vol. 113, no. 31 (Aug. 2016), 8630–8635.
- [81] Lauren Sompayrac. *How the Immune System Works*. 4th ed. New York: Wiley, 2012.
- [82] Akiko Iwasaki and Ruslan Medzhitov. “Control of adaptive immunity by the innate immune system”. *Nature Immunology*, vol. 16 (Mar. 2015), 343.
- [83] Nu Zhang and Michael J. Bevan. “CD8+ T Cells: Foot Soldiers of the Immune System”. *Immunity*, vol. 35, no. 2 (Aug. 2011), 161–168.
- [84] Kenneth Murphy and Casey Weaver. *Janeway’s immunobiology*. Ninth edition. New York, NY, USA: Garland Science, Taylor & Francis Group, LLC, 2017.
- [85] Steven H. Strogatz. *Nonlinear Dynamics and Chaos*. 0th ed. CRC Press, May 2018.
- [86] Andreas Mayer *et al.* “Regulation of T cell expansion by antigen presentation dynamics”. *Proceedings of the National Academy of Sciences*, vol. 116, no. 13 (Mar. 2019), 5914.
- [87] Frank S. Heldt *et al.* “Single-cell analysis and stochastic modelling unveil large cell-to-cell variability in influenza A virus infection”. *Nature Communications*, vol. 6, no. 1 (Nov. 2015), 8938.

- [88] Rafael Sanjuán. “Collective properties of viral infectivity”. *Current Opinion in Virology*, vol. 33 (Dec. 2018), 1–6.
- [89] Ying-Han Chen *et al.* “Phosphatidylserine Vesicles Enable Efficient En Bloc Transmission of Enteroviruses”. *Cell*, vol. 160, no. 4 (Feb. 2015), 619–630.
- [90] Yael Lebel *et al.* “Excitable dynamics of flares and relapses in autoimmune diseases”. *iScience*, vol. 26, no. 11 (Nov. 2023), 108084.
- [91] Bard Ermentrout. *Simulating, Analyzing, and Animating Dynamical Systems: A Guide to XPPAUT for Researchers and Students*. Society for Industrial and Applied Mathematics, Jan. 2002.
- [92] Morris W. Hirsch *et al.* *Differential equations, dynamical systems, and an introduction to chaos*. 3rd ed. Waltham, MA: Academic Press, 2013.
- [93] John. Guckenheimer and Philip 1945- Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. Applied mathematical sciences ; volume 42. New York: Springer-Verlag, 1983.
- [94] Carl M. 1943- Bender and Steven A. Orszag. *Advanced mathematical methods for scientists and engineers I : asymptotic methods and perturbation theory*. New York, NY: Springer, 1999.
- [95] Andreas Mayer *et al.* “How different are self and nonself?” *arXiv* (2022).
- [96] Don Mason. “A very high level of crossreactivity is an essential feature of the T-cell receptor”. *Immunology Today*, vol. 19, no. 9 (Sept. 1998), 395–404.
- [97] Qian Qi *et al.* “Diversity and clonal selection in the human T-cell repertoire”. *Proceedings of the National Academy of Sciences*, vol. 111, no. 36 (Sept. 2014), 13139.
- [98] Grégoire Altan-Bonnet *et al.* “Quantitative immunology for physicists”. *Physics Reports*, vol. 849 (Mar. 2020), 1–83.
- [99] Xiaoping Sun *et al.* “Longitudinal analysis reveals age-related changes in the T cell receptor repertoire of human T cell subsets”. *Journal of Clinical Investigation*, vol. 132, no. 17 (Sept. 2022), e158122.
- [100] Andrej Košmrlj *et al.* “Thymic Selection of T-Cell Receptors as an Extreme Value Problem”. *Physical Review Letters*, vol. 103, no. 6 (Aug. 2009), 068103.
- [101] Matthew Krummel and James Allison. “CD28 and CTLA-4 have opposing effects on the response of T cells to stimulation.” *The Journal of Experimental Medicine*, vol. 182, no. 2 (Aug. 1995), 459.

- [102] Grégoire Altan-Bonnet and Ratnadeep Mukherjee. “Cytokine-mediated communication: a quantitative appraisal of immune complexity”. *Nature Reviews Immunology* (Feb. 2019).
- [103] Catherine E. Hughes and Robert J. B. Nibbs. “A guide to chemokines and their receptors”. *The FEBS Journal*, vol. 285, no. 16 (Aug. 2018), 2944–2971.
- [104] Giorgos Minas *et al.* “Multiplexing information flow through dynamic signalling systems”. *PLOS Computational Biology*, vol. 16, no. 8 (Aug. 2020). Ed. by Attila Csikász-Nagy, e1008076.
- [105] Guillaume Voisinne *et al.* “T cells integrate Local and Global cues to discriminate between structurally similar antigens”. *Cell reports*, vol. 11, no. 8 (May 2015), 1208–1219.
- [106] Harikesh S. Wong *et al.* “A local regulatory T cell feedback circuit maintains immune homeostasis by pruning self-activated T cells”. *Cell*, vol. 184, no. 15 (July 2021), 3981–3997.e22.
- [107] Jane C Stinchcombe and Gillian M Griffiths. “The role of the secretory immunological synapse in killing by CD8+ CTL”. *Seminars in Immunology*, vol. 15, no. 6 (Dec. 2003), 301–305.
- [108] Gabriel D. Victora and Michel C. Nussenzweig. “Germinal Centers”. *Annual Review of Immunology*, vol. 40, no. 1 (Apr. 2022), 413–442.
- [109] Kyeong-Hee Lee *et al.* “The Immunological Synapse Balances T Cell Receptor Signaling and Degradation”. *Science*, vol. 302, no. 5648 (Nov. 2003), 1218–1222.
- [110] Darrell J. Irvine *et al.* “Direct observation of ligand recognition by T cells”. *Nature*, vol. 419, no. 6909 (Oct. 2002), 845–849.
- [111] Michael L. Dustin and Jay T. Groves. “Receptor Signaling Clusters in the Immune Synapse”. *Annual Review of Biophysics*, vol. 41, no. 1 (June 2012), 543–556.
- [112] Daniel Coombs *et al.* “Activated TCRs remain marked for internalization after dissociation from pMHC”. *Nature Immunology*, vol. 3, no. 10 (Oct. 2002), 926–931.
- [113] Alexander S. Moffett *et al.* “Comparing kinetic proofreading and kinetic segregation for T cell receptor activation”. *bioRxiv* (June 2023).
- [114] Guy Werlen and Ed Palmer. “The T-cell receptor signalosome: a dynamic structure with expanding complexity”. *Current Opinion in Immunology*, vol. 14, no. 3 (June 2002), 299–305.
- [115] Raman S. Ganti *et al.* “How the T cell signaling network processes information to discriminate between self and agonist ligands”. *Proceedings of the National Academy of Sciences*, vol. 117, no. 42 (Oct. 2020), 26020–26030.

- [116] Grégoire Altan-Bonnet and Ronald N Germain. “Modeling T Cell Antigen Discrimination Based on Feedback Control of Digital ERK Responses”. *PLOS Biology*, vol. 3, no. 11 (Oct. 2005), e356.
- [117] Jesse Goyette *et al.* “Regulated unbinding of ZAP70 at the T cell receptor by kinetic avidity”. *EMBO Journal*, vol. Submitted (Apr. 2020).
- [118] Darren B. McAfee *et al.* “Discrete LAT condensates encode antigen information from single pMHC:TCR binding events”. *Nature Communications*, vol. 13, no. 1 (Dec. 2022), 7446.
- [119] Hui Chen *et al.* “Self-programmed dynamics of T cell receptor condensation”. *Proceedings of the National Academy of Sciences*, vol. 120, no. 28 (July 2023), e2217301120.
- [120] Gilbert J Kersh *et al.* “High- and Low-Potency Ligands with Similar Affinities for the TCR”. *Immunity*, vol. 9, no. 6 (Dec. 1998), 817–826.
- [121] Baoyu Liu *et al.* “Accumulation of Dynamic Catch Bonds between TCR and Agonist Peptide-MHC Triggers T Cell Signaling”. *Cell*, vol. 157, no. 2 (Apr. 2014), 357–368.
- [122] De Dong *et al.* “Structural basis of assembly of the human T cell receptor–CD3 complex”. *Nature* (Aug. 2019).
- [123] Lunyi Li *et al.* “Ionic CD3-Lck interaction regulates the initiation of T-cell receptor signaling”. *Proceedings of the National Academy of Sciences*, vol. 114, no. 29 (July 2017), E5891–E5899.
- [124] Johannes Pettmann *et al.* “The discriminatory power of the T cell receptor”. *eLife*, vol. 10 (May 2021), e67092.
- [125] Melissa Lever *et al.* “Phenotypic models of T cell activation”. *Nature Reviews Immunology*, vol. 14, no. 9 (Sept. 2014), 619–629.
- [126] J. J. Hopfield. “Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity”. *Proceedings of the National Academy of Sciences*, vol. 71, no. 10 (Oct. 1974), 4135.
- [127] Jacques Ninio. “Kinetic amplification of enzyme discrimination”. *Biochimie*, vol. 57, no. 5 (July 1975), 587–595.
- [128] T W McKeithan. “Kinetic proofreading in T-cell receptor signal transduction”. *Proceedings of the National Academy of Sciences*, vol. 92, no. 11 (May 1995), 5042.
- [129] Paul François *et al.* “Phenotypic model for early T-cell activation displaying sensitivity, specificity, and antagonism”. *Proceedings of the National Academy of Sciences*, vol. 110, no. 10 (Mar. 2013), E888.

- [130] Ellen Neumeister Kersh *et al.* “Fidelity of T Cell Activation Through Multistep T Cell Receptor ζ Phosphorylation”. *Science*, vol. 281, no. 5376 (July 1998), 572.
- [131] Guillaume Voisinne *et al.* “Kinetic proofreading through the multi-step activation of the ZAP70 kinase underlies early T cell ligand discrimination”. *Nature Immunology* (Aug. 2022).
- [132] Jean-Benoît Lalanne and Paul François. “Principles of Adaptive Sorting Revealed by In Silico Evolution”. *Physical Review Letters* (2013).
- [133] Wenping Cui and Pankaj Mehta. “Identifying feasible operating regimes for early T-cell recognition: The speed, energy, accuracy trade-off in kinetic proofreading and adaptive sorting”. *PLOS ONE*, vol. 13, no. 8 (Aug. 2018). Ed. by Jordi Garcia-Ojalvo, e0202331.
- [134] Kenneth G. Johnson *et al.* “A supramolecular basis for CD45 tyrosine phosphatase regulation in sustained T cell activation”. *Proceedings of the National Academy of Sciences*, vol. 97, no. 18 (Aug. 2000), 10138–10143.
- [135] Simon J Davis and P Anton van der Merwe. “The kinetic-segregation model: TCR triggering and beyond”. *Nature Immunology*, vol. 7, no. 8 (Aug. 2006), 803–809.
- [136] Ofer Feinerman *et al.* “Variability and robustness in T cell activation from regulated heterogeneity in protein levels”. *Science*, vol. 321, no. 5892 (Aug. 2008), 1081–1084.
- [137] Adrien Henry *et al.* “ ϕ -evo: A program to evolve phenotypic models of biological networks”. *PLOS Computational Biology*, vol. 14, no. 6 (June 2018), e1006244.
- [138] Irena Štefanová *et al.* “TCR ligand discrimination is enforced by competing ERK positive and SHP-1 negative feedback pathways”. *Nature Immunology*, vol. 4, no. 3 (Mar. 2003), 248–254.
- [139] Paul François *et al.* “Phenotypic spandrel: absolute discrimination and ligand antagonism”. *Physical Biology*, vol. 13, no. 6 (2016), 066011.
- [140] Rob Phillips. “Theory in Biology: Figure 1 or Figure 7?” *Trends in Cell Biology*, vol. 25, no. 12 (Dec. 2015), 723–729.
- [141] Rabea Seyboldt *et al.* “Latent space of a small genetic network: Geometry of dynamics and information”. *Proceedings of the National Academy of Sciences*, vol. 119, no. 26 (June 2022), e2113651119.
- [142] Gordon J. Berman *et al.* “Mapping the stereotyped behaviour of freely moving fruit flies”. *Journal of The Royal Society Interface*, vol. 11, no. 99 (Oct. 2014), 20140672.
- [143] Gordon J. Berman *et al.* “Predictability and hierarchy in *Drosophila* behavior”. *Proceedings of the National Academy of Sciences*, vol. 113, no. 42 (Oct. 2016), 11943–11948.

- [144] Francis Corson and Eric D Siggia. “Gene-free methodology for cell fate dynamics during development”. *eLife*, vol. 6 (Dec. 2017), e30743.
- [145] Francis Corson and Eric D Siggia. “Geometry, epistasis, and developmental patterning”. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 15 (Apr. 2012), 5568–5575.
- [146] Laurent Jutras-Dubé *et al.* “Geometric models for robust encoding of dynamical information into embryonic patterns”. *eLife*, vol. 9 (Aug. 2020). Ed. by Sandeep Krishna *et al.*, e55778.
- [147] David A. Rand *et al.* “Geometry of gene regulatory dynamics”. *Proceedings of the National Academy of Sciences*, vol. 118, no. 38 (Sept. 2021), e2109729118.
- [148] C. H. Waddington. *The strategy of the genes; a discussion of some aspects of theoretical biology*. London: Allen & Unwin, 1957.
- [149] Jin Wang *et al.* “Quantifying the Waddington landscape and biological paths for development and differentiation”. *Proceedings of the National Academy of Sciences*, vol. 108, no. 20 (May 2011), 8257–8262.
- [150] Simon L. Freedman *et al.* “A dynamical systems treatment of transcriptomic trajectories in hematopoiesis”. *Development*, vol. 150, no. 11 (June 2023), dev201280.
- [151] Nacer Eddine Boukacem *et al.* “A Waddington landscape for prototype learning in generalized Hopfield networks”. *arXiv* (Dec. 2023).
- [152] Jeremy Gunawardena. “Models in biology: ‘accurate descriptions of our pathetic thinking’”. *BMC Biology*, vol. 12, no. 1 (Apr. 2014), 29.
- [153] Frederick Reif. *Fundamentals of statistical and thermal physics*. McGraw-Hill series in fundamentals of physics. New York: McGraw-Hill, 1965.
- [154] Jesper Tegner *et al.* “A perspective on bridging scales and design of models using low-dimensional manifolds and data-driven model inference”. *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences*, vol. 374, no. 2080 (Nov. 2016), 20160144.
- [155] Ao Ma and Aaron R. Dinner. “Automatic Method for Identifying Reaction Coordinates in Complex Systems”. *The Journal of Physical Chemistry B*, vol. 109, no. 14 (Apr. 2005), 6769–6779.
- [156] Kabir Husain and Arvind Murugan. “Physical Constraints on Epistasis”. *Molecular Biology and Evolution*, vol. 37, no. 10 (Oct. 2020), 2865–2874.
- [157] Kevin S. Brown and James P. Sethna. “Statistical mechanical approaches to models with many poorly known parameters”. *Physical Review E*, vol. 68, no. 2 (Aug. 2003), 021904.

- [158] Ryan N Gutenkunst *et al.* “Universally Sloppy Parameter Sensitivities in Systems Biology Models”. *PLoS Computational Biology*, vol. 3, no. 10 (Oct. 2007). Ed. by Adam P Arkin, e189.
- [159] Félix Proulx-Giraldeau *et al.* “Untangling the hairball: fitness based asymptotic reduction of biological networks”. *Biophysical Journal*, vol. 113, no. 8 (Oct. 2017), 1893–1906.
- [160] Thierry Mora and William Bialek. “Are Biological Systems Poised at Criticality?” *Journal of Statistical Physics*, vol. 144, no. 2 (July 2011), 268–302.
- [161] William Bialek *et al.* “Statistical mechanics for natural flocks of birds”. *Proceedings of the National Academy of Sciences*, vol. 109, no. 13 (Mar. 2012), 4786.
- [162] William Bialek. “Perspectives on theory at the interface of physics and biology”. *Reports on Progress in Physics*, vol. 81, no. 1 (2018), 012601.
- [163] Jean-Marc Goaillard and Eve Marder. “Ion Channel Degeneracy, Variability, and Covariation in Neuron and Circuit Resilience”. *Annual Review of Neuroscience*, vol. 44, no. 1 (July 2021), 335–357.
- [164] Frederic Grabowski *et al.* “Limits to the rate of information transmission through the MAPK pathway”. *Journal of The Royal Society Interface*, vol. 16, no. 152 (Mar. 2019), 20180792–20180792.
- [165] Jinyuan Yan *et al.* “Bow-tie signaling in c-di-GMP: Machine learning in a simple biochemical network”. *PLOS Computational Biology*, vol. 13, no. 8 (Aug. 2017), e1005677.
- [166] L Collavin *et al.* “p53-family proteins and their regulators: hubs and spokes in tumor suppression”. *Cell Death & Differentiation*, vol. 17, no. 6 (June 2010), 901–911.
- [167] Vincent Thibeault *et al.* “The low-rank hypothesis of complex systems”. *Nature Physics* (Jan. 2024).
- [168] François X. P. Bourassa and Thomas J. Rademaker. *Antigen encoding theory*. 2022.
- [169] Alon Oyler-Yaniv *et al.* “A Tunable Diffusion-Consumption Mechanism of Cytokine Propagation Enables Plasticity in Cell-to-Cell Communication in the Immune System”. *Immunity*, vol. 46, no. 4 (Apr. 2017), 609–620.
- [170] Karen E Tkach *et al.* “T cells translate individual, quantal activation into collective, analog cytokine responses via time-integrated feedbacks”. *eLife*, vol. 3 (Apr. 2014). Ed. by Ruslan Medzhitov, e01944.
- [171] Amanda C. Stanley and Paige Lacy. “Pathways for Cytokine Secretion”. *Physiology*, vol. 25, no. 4 (Aug. 2010), 218–229.

- [172] Anthony P. Manderson *et al.* “Subcompartments of the macrophage recycling endosome direct the differential secretion of IL-6 and TNF α ”. *The Journal of Cell Biology*, vol. 178, no. 1 (July 2007), 57–69.
- [173] Kathryn Trandem *et al.* “Highly activated cytotoxic CD8 T cells express protective IL-10 at the peak of coronavirus-induced encephalitis”. *Journal of immunology (Baltimore, Md. : 1950)*, vol. 186, no. 6 (Mar. 2011), 3642–3652.
- [174] Mark A. Daniels *et al.* “Thymic selection threshold defined by compartmentalization of Ras/MAPK signalling”. *Nature*, vol. 444, no. 7120 (Dec. 2006), 724–729.
- [175] Marta Łuksza *et al.* “A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy”. *Nature*, vol. 551, no. 7681 (Nov. 2017), 517–520.
- [176] Dietmar Zehn *et al.* “Complete but curtailed T-cell response to very low-affinity antigen”. *Nature*, vol. 458, no. 7235 (Mar. 2009), 211–214.
- [177] Paul Gueguen *et al.* “Contribution of resident and circulating precursors to tumor-infiltrating CD8+ T cell populations in lung cancer”. *Science Immunology*, vol. 6, no. 55 (Jan. 2021), eabd5778.
- [178] Meromit Singer *et al.* “A Distinct Gene Module for Dysfunction Uncoupled from Activation in Tumor-Infiltrating T Cells”. *Cell*, vol. 166, no. 6 (Sept. 2016), 1500–1511.e9.
- [179] Shai S. Shen-Orr *et al.* “Network motifs in the transcriptional regulation network of *Escherichia coli*”. *Nature Genetics*, vol. 31, no. 1 (May 2002), 64–68.
- [180] Naftali Tishby *et al.* “The information bottleneck method”. *arXiv* (Apr. 2000).
- [181] Bryan C. Daniels and Ilya Nemenman. “Automated adaptive inference of phenomenological dynamical models”. *Nature Communications*, vol. 6, no. 1 (Aug. 2015), 8133.
- [182] Marcelo Behar *et al.* “The Dynamics of Signaling as a Pharmacological Target”. *Cell*, vol. 155, no. 2 (Oct. 2013), 448–461.
- [183] Nagarajan Nandagopal *et al.* “Dynamic Ligand Discrimination in the Notch Signaling Pathway”. *Cell*, vol. 172, no. 4 (Feb. 2018), 869–880.e19.
- [184] Sooraj R. Achar and Grégoire Altan-Bonnet. *plateypus*. 2021.
- [185] Thomas J. Rademaker *et al.* *Antigen encoding pipeline*. 2022.
- [186] P. Dierckx. “An algorithm for smoothing, differentiation and integration of experimental data using spline functions”. *Journal of Computational and Applied Mathematics*, vol. 1, no. 3 (Sept. 1975), 165–184.

- [187] Pauli Virtanen *et al.* “SciPy 1.0: fundamental algorithms for scientific computing in Python”. *Nature Methods*, vol. 17, no. 3 (Mar. 2020), 261–272.
- [188] Alexander Kraskov *et al.* “Estimating mutual information”. *Physical Review E*, vol. 69, no. 6 (June 2004), 066138.
- [189] Brian C. Ross. “Mutual Information between Discrete and Continuous Data Sets”. *PLOS ONE*, vol. 9, no. 2 (Feb. 2014), e87357.
- [190] Fabian Pedregosa *et al.* “Scikit-learn: Machine Learning in Python”. *Journal of Machine Learning Research*, vol. 12, no. 85 (2011), 2825–2830.
- [191] James Theiler. “Estimating fractal dimension”. *Journal of the Optical Society of America A*, vol. 7, no. 6 (June 1990), 1055–1073.
- [192] Ronald N. Germain and Irena Stefanová. “THE DYNAMICS OF T CELL RECEPTOR SIGNALING: Complex Orchestration and the Key Roles of Tempo and Cooperation”. *Annual Review of Immunology*, vol. 17, no. 1 (Apr. 1999), 467–522.
- [193] Henry H. Mattingly *et al.* “Maximizing the information learned from finite data selects a simple model”. *Proceedings of the National Academy of Sciences*, vol. 115, no. 8 (2018), 1760–1765.
- [194] M. G. Kendall. “A New Measure of Rank Correlation”. *Biometrika*, vol. 30, no. 1/2 (June 1938), 81.
- [195] Eric D. Feigelson and G. Jogesh Babu. *Modern Statistical Methods for Astronomy: With R Applications*. 1st ed. Cambridge University Press, July 2012.
- [196] William H. Press *et al.* *Numerical Recipes in C: The Art of Scientific Computing*. 2nd ed. New York, NY: Cambridge University Press, 1992.
- [197] Karsten Borgwardt *et al.* “Class Prediction from Time Series Gene Expression Profiles Using Dynamical Systems Kernels”. *Pacific Symposium on Biocomputing 2006 : Maui, Hawaii, 3-7 January 2006*. Singapore: World Scientific, 2006.
- [198] Ying Tang *et al.* “Quantifying information accumulation encoded in the dynamics of biochemical signaling”. *Nature Communications*, vol. 12, no. 1 (Feb. 2021), 1272.
- [199] Raymond Cheong *et al.* “Information Transduction Capacity of Noisy Biochemical Signaling Networks”. *Science*, vol. 334, no. 6054 (Oct. 2011), 354.
- [200] Jangir Selimkhanov *et al.* “Accurate information transmission through dynamic biochemical signaling networks”. *Science*, vol. 346, no. 6215 (Dec. 2014), 1370.
- [201] Alejandro A. Granados *et al.* “Distributed and dynamic intracellular organization of extracellular information”. *Proceedings of the National Academy of Sciences* (May 2018).

- [202] Ingram Olkin. “Estimating a Cholesky Decomposition”. *Linear Algebra and its Applications*, vol. 67 (1985), 201–205.
- [203] F. N. Fritsch and R. E. Carlson. “Monotone Piecewise Cubic Interpolation”. *SIAM Journal on Numerical Analysis*, vol. 17, no. 2 (Apr. 1980), 238–246.
- [204] Mutsuo Saito and Makoto Matsumoto. “A PRNG Specialized in Double Precision Floating Point Numbers Using an Affine Transition”. *Monte Carlo and Quasi-Monte Carlo Methods 2008*. Ed. by Pierre L’ Ecuyer and Art B. Owen. Berlin, Heidelberg: Springer, 2009, 589–602.
- [205] Luc Devroye. *Non-Uniform Random Variate Generation*. 1st ed. New York, NY: Springer-Verlag, 1986.
- [206] Marianne Bauer *et al.* “Trading bits in the readout from a genetic network”. *arXiv*, vol. <http://arxiv.org/abs/> (Dec. 2020).
- [207] Tero Karras *et al.* “Analyzing and Improving the Image Quality of StyleGAN”. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA: IEEE, June 2020, 8107–8116.
- [208] Phillip Wang. *Random Face Generator (This Person Does Not Exist)*. 2023.
- [209] Darya Y. Orlova *et al.* “Earth Mover’s Distance (EMD): A True Metric for Comparing Biomarker Expression Levels in Cell Populations”. *PLOS ONE*, vol. 11, no. 3 (Mar. 2016), e0151859.
- [210] Aric A. Hagberg *et al.* “Exploring Network Structure, Dynamics, and Function using NetworkX”. *Proceedings of the 7th Python in Science Conference*. Ed. by Gaël Varoquaux *et al.* Pasadena, CA USA, 2008, 11–15.
- [211] Wendell A. Lim and Carl H. June. “The Principles of Engineering Immune Cells to Treat Cancer”. *Cell*, vol. 168, no. 4 (Feb. 2017), 724–740.
- [212] Donald F. Hunt *et al.* “Characterization of Peptides Bound to the Class I MHC Molecule HLA-A2.1 by Mass Spectrometry”. *Science*, vol. 255, no. 5049 (Mar. 1992), 1261–1263.
- [213] Lauren E. Stopfer *et al.* “Multiplexed relative and absolute quantitative immunopeptidomics reveals MHC I repertoire alterations induced by CDK4/6 inhibition”. *Nature Communications*, vol. 11, no. 1 (June 2020), 2760.
- [214] Lauren E. Stopfer *et al.* “Absolute quantification of tumor antigens using embedded MHC-I isotopologue calibrants”. *Proceedings of the National Academy of Sciences*, vol. 118, no. 37 (Sept. 2021), e2111173118.
- [215] Paul Klenerman *et al.* “Cytotoxic T-cell activity antagonized by naturally occurring HIV-1 Gag variants”. *Nature*, vol. 369, no. 6479 (June 1994), 403–407.

- [216] Antonio Bertoletti *et al.* “Natural variants of cytotoxic epitopes are T-cell receptor antagonists for antiviral cytotoxic T cells”. *Nature*, vol. 369, no. 6479 (June 1994), 407–410.
- [217] Kristin A. Hogquist *et al.* “T cell receptor antagonist peptides induce positive selection”. *Cell*, vol. 76, no. 1 (Jan. 1994), 17–27.
- [218] Jevrosima Radovich and David W. Talmage. “Antigenic Competition: Cellular or Humoral”. *Science*, vol. 158, no. 3800 (Oct. 1967), 512–514.
- [219] Luciano Adorini *et al.* “In vivo competition between self peptides and foreign antigens in T-cell activation”. *Nature*, vol. 334, no. 6183 (Aug. 1988), 623–625.
- [220] Brian D. Evavold and Paul M. Allen. “Separation of IL-4 Production from Th Cell Proliferation by an Altered T Cell Receptor Ligand”. *Science*, vol. 252, no. 5010 (May 1991), 1308–1310.
- [221] L Racioppi *et al.* “Peptide-major histocompatibility complex class II complexes with mixed agonist/antagonist properties provide evidence for ligand-related differences in T cell receptor-dependent intracellular signaling.” *The Journal of experimental medicine*, vol. 177, no. 4 (Apr. 1993), 1047–1060.
- [222] S C Jameson *et al.* “Clone-specific T cell receptor antagonists of major histocompatibility complex class I-restricted cytotoxic T cells.” *The Journal of experimental medicine*, vol. 177, no. 6 (June 1993), 1541–1550.
- [223] M.Teresa De Magistris *et al.* “Antigen analog-major histocompatibility complexes act as antagonists of the T cell receptor”. *Cell*, vol. 68, no. 4 (Feb. 1992), 625–634.
- [224] Bonnie N Dittel *et al.* “Cross-Antagonism of a T Cell Clone Expressing Two Distinct T Cell Receptors”. *Immunity*, vol. 11, no. 3 (Sept. 1999), 289–298.
- [225] Dennis C. Wylie *et al.* “Sensitivity of T cells to antigen and antagonism emerges from differential regulation of the same molecular signaling module”. *Proceedings of the National Academy of Sciences*, vol. 104, no. 13 (Mar. 2007), 5533–5538.
- [226] David R. Plas *et al.* “Direct Regulation of ZAP-70 by SHP-1 in T Cell Antigen Receptor Signaling”. *Science*, vol. 272, no. 5265 (May 1996), 1173–1176.
- [227] Boryana N Manz *et al.* “Small molecule inhibition of Csk alters affinity recognition by T cells”. *eLife*, vol. 4 (Aug. 2015), e08088.
- [228] Nicolai S. C. Van Oers *et al.* “Regulation of TCR Signal Transduction in Murine Thymocytes by Multiple TCR ζ -Chain Signaling Motifs”. *The Journal of Immunology*, vol. 160, no. 1 (Jan. 1998), 163–170.

- [229] Himadri Mukhopadhyay *et al.* “Systems Model of T Cell Receptor Proximal Signaling Reveals Emergent Ultrasensitivity”. *PLoS Computational Biology*, vol. 9, no. 3 (Mar. 2013). Ed. by Jason M. Haugh, e1003004.
- [230] John R. James. “Tuning ITAM multiplicity on T cell receptors can control potency and selectivity to ligand density”. *Science Signaling*, vol. 11, no. 531 (May 2018), ean1088.
- [231] Lisa A. Pitcher and Nicolai S.C. van Oers. “T-cell receptor signal transmission: who gives an ITAM?” *Trends in Immunology*, vol. 24, no. 10 (Oct. 2003), 554–560.
- [232] Joaquín Madrenas *et al.* “ ζ Phosphorylation Without ZAP-70 Activation Induced by TCR Antagonists or Partial Agonists”. *Science*, vol. 267, no. 5197 (Jan. 1995), 515–518.
- [233] Nicolai S. C. van Oers *et al.* “The 21- and 23-kD forms of TCR ζ are generated by specific ITAM phosphorylations”. *Nature Immunology*, vol. 1, no. 4 (Oct. 2000), 322–328.
- [234] Joanne Sloan-Lancaster *et al.* “Partial T cell signaling: Altered phospho- ζ and lack of zap70 recruitment in APL-induced T cell anergy”. *Cell*, vol. 79, no. 5 (Dec. 1994), 913–922.
- [235] C Reis E Sousa *et al.* “Partial signaling by CD8+ T cells in response to antagonist ligands.” *The Journal of experimental medicine*, vol. 184, no. 1 (July 1996), 149–157.
- [236] Ellen N. Kersh *et al.* “Partially Phosphorylated T Cell Receptor ζ Molecules Can Inhibit T Cell Activation”. *The Journal of Experimental Medicine*, vol. 190, no. 11 (Dec. 1999), 1627–1636.
- [237] Haiyan Liu and Dario A. A. Vignali. “Differential CD3 ζ Phosphorylation Is Not Required for the Induction of T Cell Antagonism by Altered Peptide Ligands”. *The Journal of Immunology*, vol. 163, no. 2 (July 1999), 599–602.
- [238] Lisa A. Pitcher *et al.* “T Cell Antagonism is Functionally Uncoupled from the 21- and 23-kDa Tyrosine-Phosphorylated TCR ζ Subunits”. *The Journal of Immunology*, vol. 171, no. 2 (July 2003), 845–852.
- [239] Laurence Ardouin *et al.* “Crippling of CD3- ζ ITAMs Does Not Impair T Cell Receptor Signaling”. *Immunity*, vol. 10, no. 4 (Apr. 1999), 409–420.
- [240] Elizabeth W. Shores *et al.* “T Cell Development in Mice Lacking All T Cell Receptor ζ Family Members (ζ , η , and Fc ϵ RI γ)”. *The Journal of Experimental Medicine*, vol. 187, no. 7 (Apr. 1998), 1093–1101.
- [241] SuJin Hwang *et al.* “TCR ITAM multiplicity is required for the generation of follicular helper T-cells”. *Nature Communications*, vol. 6, no. 1 (May 2015), 6982.
- [242] Manas K. Ray *et al.* “The Cre-loxP System: A Versatile Tool for Targeting Genes in a Cell- and Stage-Specific Manner”. *Cell Transplantation*, vol. 9, no. 6 (Nov. 2000), 805–815.

- [243] Duncan Kirby and Anton Zilman. “Proofreading does not result in more reliable ligand discrimination in receptor signaling due to its inherent stochasticity”. *Proceedings of the National Academy of Sciences* (May 2023).
- [244] Arvind Murugan *et al.* “Discriminatory Proofreading Regimes in Nonequilibrium Systems”. *Physical Review X*, vol. 4, no. 2 (Apr. 2014), 021016.
- [245] Daniel Foreman-Mackey *et al.* “emcee : The MCMC Hammer”. *Publications of the Astronomical Society of the Pacific*, vol. 125, no. 925 (Mar. 2013), 306–312.
- [246] Alan Sokal. “Monte Carlo methods in statistical mechanics: foundations and new algorithms”. *Functional integration*. Springer, 1997, 131–192.
- [247] W. R. (Walter R.) Gilks *et al.* *Markov chain Monte Carlo in practice*. Boca Raton, Fla.: Chapman & Hall, 1998.
- [248] Steve Brooks. *Handbook of Markov Chain Monte Carlo*. Chapman & Hall/CRC handbooks of modern statistical methods. Boca Raton ; CRC Press, 2011.
- [249] Benjamin Ballnus *et al.* “Comprehensive benchmarking of Markov chain Monte Carlo methods for dynamical systems”. *BMC Systems Biology*, vol. 11, no. 1 (June 2017), 63.
- [250] Gloria I. Valderrama-Bahamóndez and Holger Fröhlich. “MCMC Techniques for Parameter Estimation of ODE Based Models in Systems Biology”. *Frontiers in Applied Mathematics and Statistics*, vol. 5 (2019), 55.
- [251] Uri Alon. “Network motifs: theory and experimental approaches”. *Nature Reviews Genetics*, vol. 8 (June 2007), 450.
- [252] Douglas Hanahan and Robert A. Weinberg. “Hallmarks of Cancer: The Next Generation”. *Cell*, vol. 144, no. 5 (Mar. 2011), 646–674.
- [253] F. M. Burnet. “The Concept of Immunological Surveillance”. *Progress in Tumor Research*. Vol. 13. S. Karger AG, Sept. 1970, 1–27.
- [254] Y. Ishida *et al.* “Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death.” *The EMBO Journal*, vol. 11, no. 11 (Nov. 1992), 3887–3895.
- [255] Jackie E. Bader *et al.* “Targeting Metabolism to Improve the Tumor Microenvironment for Cancer Immunotherapy”. *Molecular Cell*, vol. 78, no. 6 (June 2020), 1019–1033.
- [256] Karthik Dhatchinamoorthy *et al.* “Cancer Immune Evasion Through Loss of MHC Class I Antigen Presentation”. *Frontiers in Immunology*, vol. 12 (Mar. 2021), 636568.
- [257] Vijay Shankaran *et al.* “IFN γ and lymphocytes prevent primary tumour development and shape tumour immunogenicity”. *Nature*, vol. 410, no. 6832 (Apr. 2001), 1107–1111.

- [258] Marta Łuksza *et al.* “Neoantigen quality predicts immunoediting in survivors of pancreatic cancer”. *Nature* (May 2022).
- [259] Dana Leach *et al.* “Enhancement of Antitumor Immunity by CTLA-4 Blockade”. *Science*, vol. 271, no. 5256 (Mar. 1996), 1734.
- [260] Luis A. Rojas *et al.* “Personalized RNA neoantigen vaccines stimulate T cells in pancreatic cancer”. *Nature*, vol. 618, no. 7963 (June 2023), 144–150.
- [261] Steven A. Rosenberg *et al.* “Adoptive cell transfer: a clinical path to effective cancer immunotherapy”. *Nature Reviews Cancer*, vol. 8, no. 4 (Apr. 2008), 299–308.
- [262] Frederick L. Locke *et al.* “Axicabtagene Ciloleucel as Second-Line Therapy for Large B-Cell Lymphoma”. *The New England Journal of Medicine*, vol. 386, no. 7 (Feb. 2022), 640–654.
- [263] Daniel W. Lee *et al.* “T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: a phase 1 dose-escalation trial”. *Lancet*, vol. 385, no. 9967 (Feb. 2015), 517–528.
- [264] Sattva S. Neelapu *et al.* “Axicabtagene Ciloleucel CAR T-Cell Therapy in Refractory Large B-Cell Lymphoma”. *The New England Journal of Medicine*, vol. 377, no. 26 (Dec. 2017), 2531–2544.
- [265] Shannon L. Maude *et al.* “Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia”. *The New England Journal of Medicine*, vol. 378, no. 5 (Feb. 2018), 439–448.
- [266] Richard A. Morgan *et al.* “Case report of a serious adverse event following the administration of T cells transduced with a chimeric antigen receptor recognizing ERBB2”. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, vol. 18, no. 4 (Apr. 2010), 843–851.
- [267] Andrew R. Haas *et al.* “Two cases of severe pulmonary toxicity from highly active mesothelin-directed CAR T cells”. *Molecular Therapy* (June 2023).
- [268] Kyle G. Daniels *et al.* “Decoding CAR T cell phenotype using combinatorial signaling motif libraries and machine learning”. *Science*, vol. 378, no. 6625 (Dec. 2022), 1194–1200.
- [269] Victoria M. Giordani *et al.* “Generation of Hinge-Modified Human CD19 Chimeric Antigen Receptor (CAR) T Cells for the Treatment of CD19-Low Leukemia”. *Blood*, vol. 142, no. Supplement 1 (Nov. 2023), 6819–6819.
- [270] Cor H. J. Lamers *et al.* “Treatment of metastatic renal cell carcinoma (mRCC) with CAIX CAR-engineered T-cells-a completed study overview”. *Biochemical Society Transactions*, vol. 44, no. 3 (June 2016), 951–959.

- [271] Kai-chao Feng *et al.* “Cocktail treatment with EGFR-specific and CD133-specific chimeric antigen receptor-modified T cells in a patient with advanced cholangiocarcinoma”. *Journal of Hematology & Oncology*, vol. 10 (Jan. 2017), 4.
- [272] Yelei Guo *et al.* “Phase I Study of Chimeric Antigen Receptor-Modified T Cells in Patients with EGFR-Positive Advanced Biliary Tract Cancers”. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, vol. 24, no. 6 (Mar. 2018), 1277–1286.
- [273] Greg M. Allen *et al.* “Synthetic cytokine circuits that drive T cells into immune-excluded tumors”. *Science*, vol. 378, no. 6625 (Dec. 2022), eaba1624.
- [274] Aidan M. Tousley *et al.* “Co-opting signalling molecules enables logic-gated control of CAR T cells”. *Nature*, vol. 615, no. 7952 (Mar. 2023), 507–516.
- [275] Christopher C. Kloss *et al.* “Combinatorial antigen recognition with balanced signaling promotes selective tumor eradication by engineered T cells”. *Nature Biotechnology*, vol. 31, no. 1 (Jan. 2013), 71–75.
- [276] Evripidis Lanitis *et al.* “Chimeric Antigen Receptor T Cells with Dissociated Signaling Domains Exhibit Focused Antitumor Activity with Reduced Potential for Toxicity *In Vivo*”. *Cancer Immunology Research*, vol. 1, no. 1 (July 2013), 43–53.
- [277] Victor D. Fedorov *et al.* “PD-1– and CTLA-4–Based Inhibitory Chimeric Antigen Receptors (iCARs) Divert Off-Target Immunotherapy Responses”. *Science Translational Medicine*, vol. 5, no. 215 (Dec. 2013).
- [278] Shivani Srivastava *et al.* “Logic-Gated ROR1 Chimeric Antigen Receptor Expression Rescues T Cell-Mediated Toxicity to Normal Tissues and Enables Selective Tumor Targeting”. *Cancer Cell*, vol. 35, no. 3 (Mar. 2019), 489–503.e8.
- [279] Damaris Bausch-Fluck *et al.* “The in silico human surfaceome”. *Proceedings of the National Academy of Sciences*, vol. 115, no. 46 (Nov. 2018).
- [280] Milos Aleksic *et al.* “Different affinity windows for virus and cancer-specific T-cell receptors: Implications for therapeutic strategies”. *European Journal of Immunology*, vol. 42, no. 12 (2012), 3174–3179.
- [281] Dana Stenger *et al.* “Endogenous TCR promotes in vivo persistence of CD19-CAR-T cells compared to a CRISPR/Cas9-mediated TCR knockout CAR”. *Blood*, vol. 136, no. 12 (Sept. 2020), 1407–1418.
- [282] Taylor L. Wilson *et al.* “Common Trajectories of Highly Effective CD19-Specific CAR T Cells Identified by Endogenous T-cell Receptor Lineages”. *Cancer Discovery*, vol. 12, no. 9 (Sept. 2022), 2098–2119.

- [283] Robbie G. Majzner *et al.* “Tuning the Antigen Density Requirement for CAR T-cell Activity”. *Cancer Discovery*, vol. 10, no. 5 (May 2020), 702–723.
- [284] Devraj Basu *et al.* “In vivo antagonism of a T cell response by an endogenously expressed ligand”. *Proceedings of the National Academy of Sciences*, vol. 95, no. 24 (Nov. 1998), 14332–14336.
- [285] R. P. (Richard P.) Brent. *Algorithms for minimization without derivatives*. Prentice-Hall series in automatic computation. Englewood Cliffs, N.J.: Prentice-Hall, 1972.
- [286] Daniel T. Harris *et al.* “Comparison of T Cell Activities Mediated by Human TCRs and CARs That Use the Same Recognition Domains”. *The Journal of Immunology*, vol. 200, no. 3 (Feb. 2018), 1088–1100.
- [287] Nicola Trendel *et al.* “Perfect adaptation of CD8+ T cell responses to constant antigen input over a wide range of affinities is overcome by costimulation”. *Science Signaling*, vol. 14, no. 666 (Jan. 2021), eaay9363.
- [288] Masanori Nakamura *et al.* “Mechanophenotyping of B16 Melanoma Cell Variants for the Assessment of the Efficacy of (-)-Epigallocatechin Gallate Treatment Using a Tapered Microfluidic Device”. *Micromachines*, vol. 10, no. 3 (Mar. 2019), 207.
- [289] Ewa K. Zuba-Surma *et al.* “Morphological characterization of very small embryonic-like stem cells (VSELs) by ImageStream system analysis”. *Journal of Cellular and Molecular Medicine*, vol. 12, no. 1 (Nov. 2007), 292–303.
- [290] John S Davies *et al.* “Non-synergy of PD-1 blockade with T-cell therapy in solid tumors”. *Journal for Immunotherapy of Cancer*, vol. 10, no. 7 (July 2022), e004906.
- [291] David W. Scott. *Statistics: a concise mathematical introduction for students, scientists, and engineers*. Hoboken, NJ: Wiley, 2020.
- [292] Judith Feucht *et al.* “Calibration of CAR activation potential directs alternative T cell fates and therapeutic potency”. *Nature Medicine*, vol. 25, no. 1 (Jan. 2019), 82–88.
- [293] Mohamad Hamieh *et al.* “Programming CAR T Cell Tumor Recognition: Tuned Antigen Sensing and Logic Gating”. *Cancer Discovery*, vol. 13, no. 4 (Apr. 2023), 829–843.
- [294] Christian L. Flugel *et al.* “Overcoming on-target, off-tumour toxicity of CAR T cell therapy for solid tumours”. *Nature Reviews Clinical Oncology*, vol. 20, no. 1 (Jan. 2023), 49–62.
- [295] Yue Liu *et al.* “Chimeric STAR receptors using TCR machinery mediate robust responses against solid tumors”. *Science Translational Medicine*, vol. 13, no. 586 (Mar. 2021), eabb5191.
- [296] Annelisa M. Cornel *et al.* “MHC Class I Downregulation in Cancer: Underlying Mechanisms and Potential Targets for Cancer Immunotherapy”. *Cancers*, vol. 12, no. 7 (July 2020), 1760.

- [297] Sara Bobisse *et al.* “Sensitive and frequent identification of high avidity neo-epitope specific CD8+ T cells in immunotherapy-naive ovarian cancer”. *Nature Communications*, vol. 9 (Mar. 2018), 1092.
- [298] Jason R. Devlin *et al.* “Structural dissimilarity from self drives neoepitope escape from immune tolerance”. *Nature Chemical Biology*, vol. 16, no. 11 (Nov. 2020), 1269–1276.
- [299] Assaf Amitai *et al.* “A Population Dynamics Model for Clonal Diversity in a Germinal Center”. *Frontiers in Microbiology*, vol. 8 (Sept. 2017), 1693.
- [300] Veit R. Buchholz *et al.* “Disparate Individual Fates Compose Robust CD8+ T Cell Immunity”. *Science*, vol. 340, no. 6132 (May 2013), 630.
- [301] Lu Hong *et al.* “Bayesian modeling reveals metabolite-dependent ultrasensitivity in the cyanobacterial circadian clock”. *Molecular Systems Biology*, vol. 16, no. 6 (June 2020), e9355.
- [302] Christopher Rackauckas *et al.* “Universal Differential Equations for Scientific Machine Learning”. *arXiv* (2020).
- [303] Hannah Fronenberg *et al.* “A New Probe of the High-*z* BAO scale: BAO tomography With CMB x LIM-Nulling Convergence”. *arXiv* (Sept. 2023).
- [304] Ville Mustonen and Michael Lässig. “Fitness flux and ubiquity of adaptive evolution”. *Proceedings of the National Academy of Sciences*, vol. 107, no. 9 (Mar. 2010), 4248–4253.
- [305] Abigail Skwara *et al.* “Statistically learning the functional landscape of microbial communities”. *Nature Ecology & Evolution*, vol. 7, no. 11 (Oct. 2023), 1823–1833.
- [306] James Black. “Drugs from Emasculated Hormones: The Principle of Syntopic Antagonism”. *Bioscience Reports*, vol. 24, no. 4-5 (Aug. 2004), 302–322.
- [307] Galileo Galilei. *Dialogue sur les deux grands systèmes du monde*. Trans. by René Fréreau. Seuil. Points science. Paris, 1992.
- [308] Alexandre Koyré. *Études galiléennes*. Hermann. Paris, 1966.
- [309] Albert Einstein. “Physics and reality”. *Journal of the Franklin Institute*, vol. 221, no. 3 (Mar. 1936), 349–382.
- [310] Eugene P. Wigner. “The unreasonable effectiveness of mathematics in the natural sciences.” *Communications on Pure and Applied Mathematics*, vol. 13, no. 1 (Feb. 1960), 1–14.
- [311] Immanuel Kant. *Critique of pure reason*. Trans. by Paul Guyer and Allen W. Wood. The Cambridge edition of the works of Immanuel Kant. Cambridge: Cambridge University Press, 1998.

- [312] Immanuel Kant. *Critique de la raison pure*. Trans. by Alain Renaut. GF-Flammarion. Paris, 2006.
- [313] Howard C. Berg and Douglas A. Brown. “Chemotaxis in *Escherichia coli* analysed by Three-dimensional Tracking”. *Nature*, vol. 239, no. 5374 (Oct. 1972), 500–504.
- [314] Philippe Cluzel *et al.* “An Ultrasensitive Bacterial Motor Revealed by Monitoring Signaling Proteins in Single Cells”. *Science*, vol. 287, no. 5458 (Mar. 2000), 1652.
- [315] George H. Wadhams and Judith P. Armitage. “Making sense of it all: bacterial chemotaxis”. *Nature Reviews Molecular Cell Biology*, vol. 5, no. 12 (Dec. 2004), 1024–1037.
- [316] Álvaro Ortega *et al.* “Sensory Repertoire of Bacterial Chemoreceptors”. *Microbiology and Molecular Biology Reviews*, vol. 81, no. 4 (Dec. 2017), e00033–17.
- [317] Ekaterina A. Korobkova *et al.* “Hidden Stochastic Nature of a Single Bacterial Motor”. *Physical Review Letters*, vol. 96, no. 5 (Feb. 2006), 058105.
- [318] Yuhai Tu. “The nonequilibrium mechanism for ultrasensitivity in a biological switch: sensing by Maxwell’s demons”. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 33 (Aug. 2008), 11737–11741.
- [319] Victor Sourjik and Howard C. Berg. “Receptor sensitivity in bacterial chemotaxis”. *Proceedings of the National Academy of Sciences*, vol. 99, no. 1 (Jan. 2002), 123.
- [320] Mingshan Li and Gerald L. Hazelbauer. “Cellular Stoichiometry of the Components of the Chemotaxis Signaling Complex”. *Journal of Bacteriology*, vol. 186, no. 12 (June 2004), 3687–3694.
- [321] Jean-Daniel Fauny *et al.* “The entire zebrafish blastula-gastrula margin acts as an organizer dependent on the ratio of Nodal to BMP activity”. *Development*, vol. 136, no. 22 (Nov. 2009), 3811–3819.
- [322] Mariela D. Petkova *et al.* “Optimal Decoding of Cellular Identities in a Genetic Network”. *Cell*, vol. 176, no. 4 (Feb. 2019), 844–855.e15.
- [323] Emily J. Mayhew *et al.* “Transport features predict if a molecule is odorous”. *Proceedings of the National Academy of Sciences*, vol. 119, no. 15 (Apr. 2022), e2116576119.
- [324] Leslie B. Vosshall and Reinhard F. Stocker. “Molecular Architecture of Smell and Taste in *Drosophila*”. *Annual Review of Neuroscience*, vol. 30, no. 1 (July 2007), 505–533.
- [325] Elissa A. Hallem and John R. Carlson. “Coding of Odors by a Receptor Repertoire”. *Cell*, vol. 125, no. 1 (Apr. 2006), 143–160.
- [326] Kamesh Krishnamurthy *et al.* “Disorder and the Neural Representation of Complex Odors”. *Frontiers in Computational Neuroscience*, vol. 16 (2022).

- [327] Sanjoy Dasgupta *et al.* “A neural algorithm for a fundamental computing problem”. *Science*, vol. 358, no. 6364 (Nov. 2017), 793–796.
- [328] Joseph D. Zak *et al.* “Antagonistic odor interactions in olfactory sensory neurons are widespread in freely breathing mice”. *Nature Communications*, vol. 11, no. 1 (July 2020), 3350.
- [329] Tobias Ackels *et al.* “Fast odour dynamics are encoded in the olfactory system and guide behaviour”. *Nature*, vol. 593, no. 7860 (May 2021), 558–563.
- [330] Antonio Celani *et al.* “Odor Landscapes in Turbulent Environments”. *Physical Review X*, vol. 4, no. 4 (Oct. 2014), 041015.
- [331] Yang Shen *et al.* “Habituation as a neural algorithm for online odor discrimination”. *Proceedings of the National Academy of Sciences*, vol. 117, no. 22 (June 2020), 12402.
- [332] Nicola Rigolli *et al.* “Alternation emerges as a multi-modal strategy for turbulent odor navigation”. *eLife*, vol. 11 (Aug. 2022), e76989.

Appendix A

Examples of biological chemical sensing

A.1 Bacterial chemotaxis

When seeking food, bacteria need to find regions of high nutrient concentrations (*e.g.*, aspartate, serine), a phenomenon called chemotaxis. They use a characteristic “run and tumble” strategy, stochastically alternating between swimming and stopping to change orientation [313], with the transitions probability set by the chemoattractant concentrations sensed by the bacteria [314]. Each bacterial species expresses a few different types (5 in *E. coli*) of chemoreceptors, called methyl-accepting chemotaxis proteins (MCPs) [315]. While MCPs are typically first discovered for their specific role in detecting some key chemoattractant (*e.g.*, the Tar receptors for aspartate), they are in fact involved in sensing various chemoattractants and repellents [316], giving bacterial cells the potential to perform combinatorial decisions.

Upon binding to their ligand, MCPs trigger a dedicated enzymatic pathway resulting in the phosphorylation of protein CheY. In turn, phosphorylated CheY biases the rotation direction of flagellar motors in an ultra-sensitive switch-like manner [314, 317, 318, 17]. Hence, presence of the chemoattractant leads to more phosphorylated CheY and increases drastically the probability of clockwise motor rotation, corresponding to the “tumbling” or stopping state (counter-clockwise corresponds to running, in the absence of food). MCPs also trigger a biochemical adaptation pathway based on methylation reactions, whereby bacteria only respond to relative changes in concentration and swim in the direction of increasing chemoattractant gradients, *i.e.*, towards food sources [34]. Overall, bacterial chemotactic systems operate close to single-molecule physical limits: they can amplify fractional changes in receptor occupancy as low as 0.2 % [319, 315], corresponding to a change of a few dozen ligand-bound receptors (assuming $\sim 10^4$ receptor copies per bacterium [320]). This is why they have been intensely studied by biophysicists.

A.2 Cell fate decisions in embryos

During embryonic development, individual cells integrate various chemical signals through a few core families of receptors that regulate multiple differentiation pathways. The ligands targeted by

Ligands	Receptors
Delta ligands, Serrate/Jagged ligands	Notch receptors
Wnt proteins	Frizzled family
Fibroblast growth factors (FGF)	FGF receptors
Transforming Growth Factor (TGF) β	TGF- β receptors
Epidermal Growth Factor (EGF)	Receptor tyrosine kinases (RTK)
Nerve Growth Factor (NGF)	RTKs
Bone morphogenetic proteins (BMP)	BMP receptors
Hedgehog (HH)	Patched (PTCH) receptors

Table A.1: Pairs of cell surface receptor and extracellularly secreted ligand families involved in embryogenesis, based on the review in Li and Elowitz [61]. Each of these families typically contain several receptors and ligands and isoforms, with various degrees of cross-reactivity between them. Delta and Serrate are surface ligands while other families are secreted ligands.

these receptors are either presented on the surface of neighbor cells, or secreted in the extracellular medium. Table A.1 lists most of these receptor-ligand systems, which play key roles in embryogenesis. For example, bone morphogenetic protein (BMP) ligand diffusion implements a Turing patterning mechanism for digit specification [72]. Oscillations in three families of receptor-ligand systems— Notch, FGF, and Wnt – generate traveling waves of genetic expression which drive the formation of vertebrae precursors called somites [70]. Gradients of morphogenetic proteins like BMPs and TGF- β establish dorso-ventral and anterior-posterior axes to break symmetry in the embryo [321]. In the early *Drosophila* embryo, cells make fate decisions within minutes based on the concentration gradient of the Bicoid morphogen [35]¹. The Gap genes downstream of these morphogens then provide positional information to individual cells within a 1 % accuracy [53, 322].

Notably, each of the receptor-ligand families in table A.1 typically contain several receptors and ligands or isoforms, with cross-reactivity and promiscuity (*i.e.*, sharing of the same ligands by the same receptors) between them. This enables cells to compute various functions of ligand combinations depending on their receptor expression profile [60], allowing information to be addressed to specific cell types [68]. These functions can decode different features of ligand combinations: their identities, their concentrations, or their dynamical patterns [61]. Ligands can produce synergy or antagonism, that is, elicit responses larger or smaller than the sum of individual ligand

¹In this case, there are no surface receptors involved, since before cell cycle 14, the embryo is in a syncytial state where the 6000 nuclei share the same cytoplasm, but the promoters of downstream genes can be thought of as receptors sensing the morphogen concentration.

responses [63]. Hence, precise chemical sensing of ligand mixtures underlies the proper specification of cell types to carry out developmental plans.

A.3 Olfaction

Odors are a valuable source of information to animals about their surroundings: using smell, they can locate food, avoid predators, recognize a mate, and so on. Odors are caused by small organic compounds sufficiently volatile and hydrophobic to be transported in the air; recent estimates suggest there may be at least 4×10^{10} such possible odorous molecules [323]. Evolution has endowed animals with olfactory receptors (OR) to pick up these chemical cues in the air. These ORs are transmembrane receptors expressed on olfactory receptor neurons (ORN) located in the nose epithelium (humans, mouse) or antenna (fly). Animals have R different OR types ($R \approx 50$ in flies, 1000 in mice) with significant cross-reactivity for many different odorants. In canonical organisms, each ORN expresses a single OR type, and all ORNs with the same receptor project to the same small neural region in the olfactory bulb or antennal lobe, called a glomerulus [324]; hence, an odorant can be encoded crudely as a R -dimensional vector $\vec{x} = (x_1, x_2, \dots, x_R)$ specifying how strongly each ORN type is stimulated by the odor (x_i could be ORN firing rates in response to the odor at a standard concentration, for instance) [325]. This way of encoding odors contrasts with adaptive immunity, where receptor repertoires are much larger, but it implements efficient algorithms – compressed sensing [326] and locality-sensitive hashing [327] – to distinguish sparse odorant mixtures that are signatures of odor sources like preys and predators.

Various nonlinear and time-dependent effects arise in odor mixtures. A recent biophysical model of ORN activation explained antagonism effects, whereby some odorants reduce the total response compared to the sum of individual responses, and revealed that antagonism normalizes ORN activity to prevent saturation in the olfactory response [66]; these predictions were confirmed experimentally [328]. Moreover, mice can have behavioral responses to fast fluctuations and delays between odorants as short as ~ 40 ms [329], indicating that temporal accuracy plays a role in olfactory sensing performance. Animals can also discriminate turbulent background from sparse relevant odors [330, 33, 331], and turn these signals into tailored behavioral strategies to track broken, intermittent olfactory trails [18, 332]. Theoretical work to understand olfactory circuits can inform us on how fundamental chemical sensing challenges are solved across various biological organisms.

Appendix B

Supplementary results of universal antigen encoding

B.1 Filtering of large background noise in diverse immunological settings

“Given the large number of experimental conditions and repeats produced by our robotic platform, we found that a few time series had very limited signals and needed to be additionally pre-processed and filtered out, as described below.” (*Antigen encoding*, [1], SI)

I developed these processing steps in the later stages of the project, after Thomas Rademaker and I had established the default processing pipeline.

B.1.1 IL-17A in human T cell time series

“Cytokine bead array measurements for the cytokine IL-17A consisted only in background fluorescence noise in *ex vivo* cultures of human T cells with peptide-pulsed tumor cells, with a signal-to-noise ratio (see eq. 2.10 [above]) several fold smaller than for other cytokines. We therefore set this cytokine to zero (in log-transformed scale) in this dataset for the rest of the processing pipeline, for all peptides.” (*Antigen encoding*, [1], SI)

To show why this processing was needed, we plot in figure B.1 the time series of human cytokine concentrations, revealing that IL-17A values are essentially just experimental noise (due to background fluorescence in the flow cytometer): the values oscillate over time in a seemingly random manner, only slightly above the lower limit of detection, similarly for all peptides and activating agents (*e.g.*, aCD3/CD28). Moreover, as shown in figure B.2, the signal-to-noise ratio (SNR) of each cytokine in the human TCR experiment confirms that the IL-17A signal is essentially noise, as opposed to the four other cytokines.

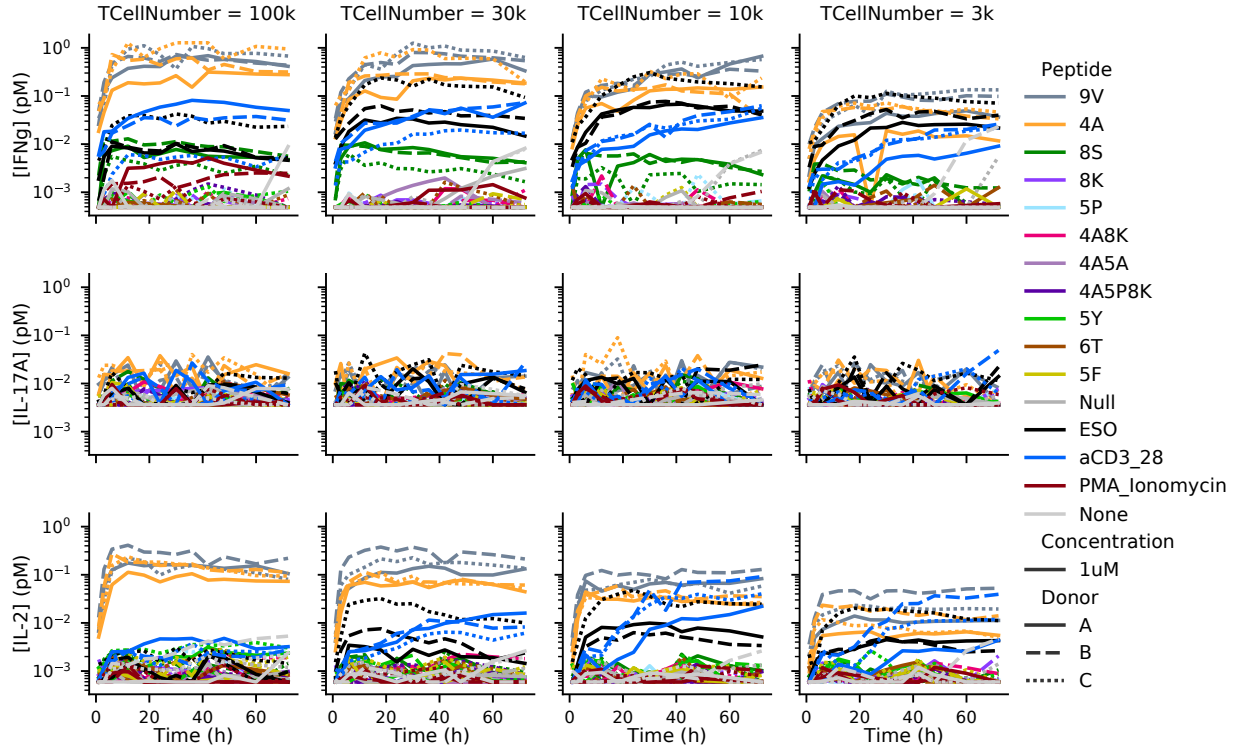


Figure B.1: Calibrated cytokine time series in the human TCR experiment (showing every peptide, T cell number, and cell donor tested). Compared to IFN- γ and IL-2, which exhibit clear dependency on antigen quality, IL-17A trajectories look like background fluctuations, similar for all levels of stimulation. Moreover, comparing IL-2 (first row) and IFN- γ time series (third row) reveals larger background fluctuations in IL-2 than in IFN- γ ; this is why we could base the filtering of IL-2 background values on the level of IFN- γ signal, as described in subsection B.1.2. (*Antigen encoding*, [1], response to reviewers)

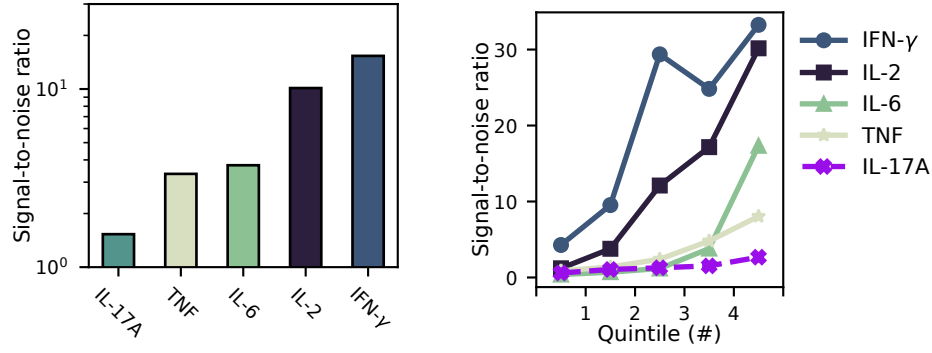


Figure B.2: (Left) Signal-to-noise ratio (SNR), as defined in equation (2.10) computed across all times and all agonist peptides in the human TCR experiment, for each cytokine. “Agonist” peptides include NY-ESO, 4A, 9V, 8S, and 8K, which produced the highest CD69 expression levels at high peptide doses. The graph shows that even for strong stimulation producing clear T cell activation, IL-17A still has a very low SNR for IL-17A, meaning it is essentially background noise only. (Right) SNR computed within each interval between quintiles, for each cytokine. Even when considering only the top 20 % of recorded concentrations for each cytokine (*i.e.*, values above the last quintile, rightmost point on the graph), the SNR of IL-17A is less than 0.5 times that of any other cytokine, again indicating that the IL-17A was simply not detectable in that experiment. (*Antigen encoding*, [1], response to reviewers)

B.1.2 Filtering based on IFN- γ

“The experimental results acquired with human anti-NY-ESO-1 TCR-expressing human T cell blasts, as well as the mouse cell dataset used in section 2.7, exhibited fluorescence background higher than usual. This measurement noise (mostly for IL-2 and IL-17A) could not be entirely removed by our regular spline smoothing process and required specific filtering.

We eliminated this excess background in the IL-2 measurements as follow. We noticed that IFN- γ was the easiest cytokine to detect, even for weak antigens. Hence, we compared each time series to the corresponding “null” condition included in the experiment (un-pulsed tumor cells for the human datasets, antigen presenting cells pulsed with E1 peptide for the mouse datasets). When we found a time series with similar or smaller IFN- γ measurements than this “null” condition, we inferred that the signal in IL-2 was necessarily background noise too, and thus we set IL-2 to zero for that time series. The threshold for similarity was $p > 0.5$ in a two-sample, one-sided Kolmogorov-Smirnov test, *SciPy*’s implementation [187]. When the comparison was inconclusive ($p < 0.5$), we did not filter out that IL-2 time series and proceeded with spline smoothing as described above.” (*Antigen encoding*, [1], SI)

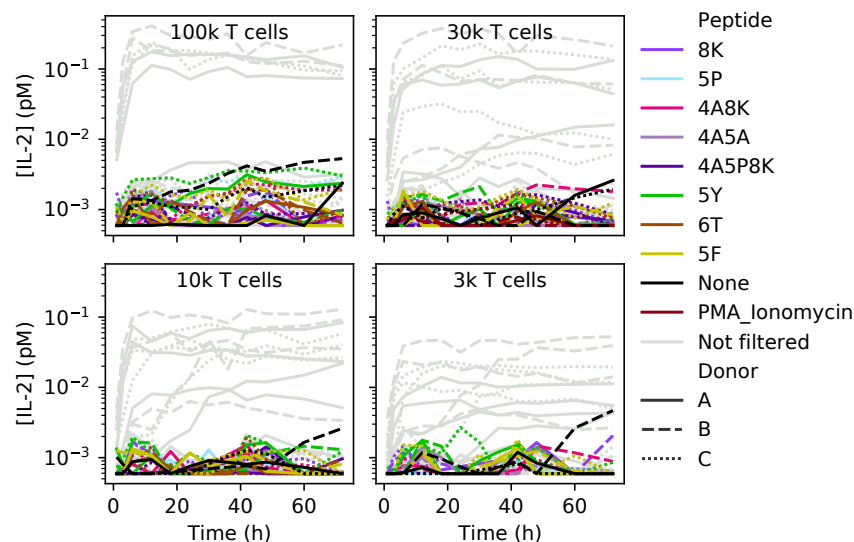


Figure B.3: IL-2 time series filtered out based on a Kolmogorov-Smirnov comparison of IFN- γ with the “null” condition (T cells and un-pulsed tumor cells). The two time series that seem to stand out in the top left panel are those of antigen 5Y (which was the weakest non-agonist in our functional EC_{50} assay of NY-ESO peptides) and of the “None” control condition (T cells, no tumors). No cytokine response is expected in those conditions; hence, all IL-2 trajectories filtered out are certainly background noise, and any trend in the time series must be due to a systematic measurement error, or contamination of the coculture by a stronger peptide. (*Antigen encoding*, [1], response to reviewers)

Figure B.3 highlights the IL-2 trajectories that were filtered out based on this test; none is visibly above background fluorescence fluctuations, confirming that our method reliably identified only time series that essentially consisted of background noise, thanks to the stringent similarity criterion $p > 0.5$.

B.2 Supplementary figures related to antigen encoding

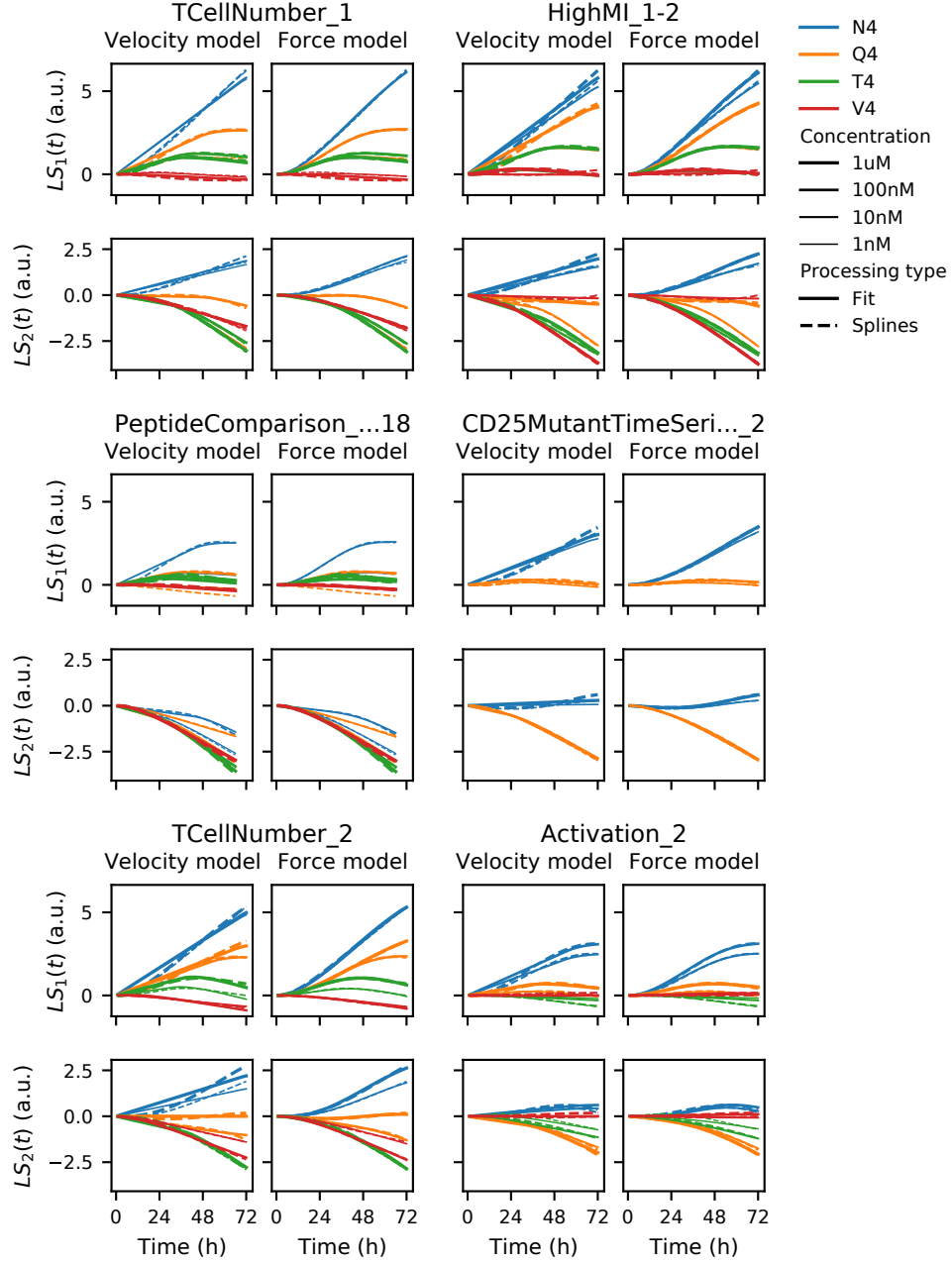


Figure B.4: Comparison between latent space time courses obtained from data (Splines, dashed lines) and fits of our model(s) (Fit, solid) for 6 independent experiments (labelled TCellNumber_1, etc.) chosen randomly out of the 14 datasets used in figure 2.17. For clarity of the plots, two time series for each peptide available in the experiment were selected also randomly. We see fits obtained with both the constant velocity model and the more accurate force model with matching. (*Antigen encoding*, [1], response to reviewers)

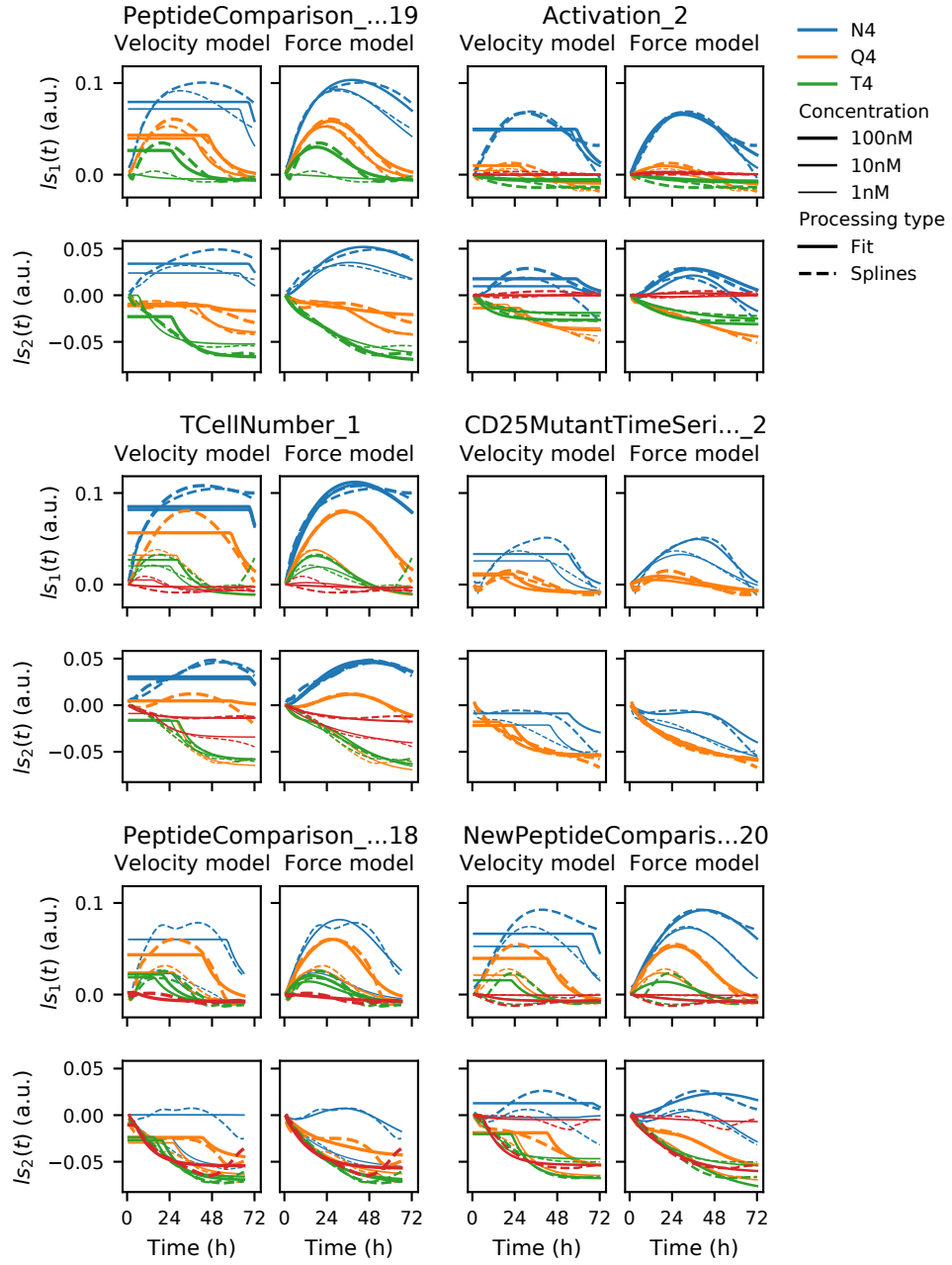


Figure B.5: Similar to figure B.4, but for latent space concentration time series ($ls_1(t)$, $ls_2(t)$), and with a different random selection of datasets. Comparing left and right subplots in each dataset makes particularly clear the greater accuracy of the force model with matching for cytokine concentration dynamics. (*Antigen encoding*, [1], response to reviewers)

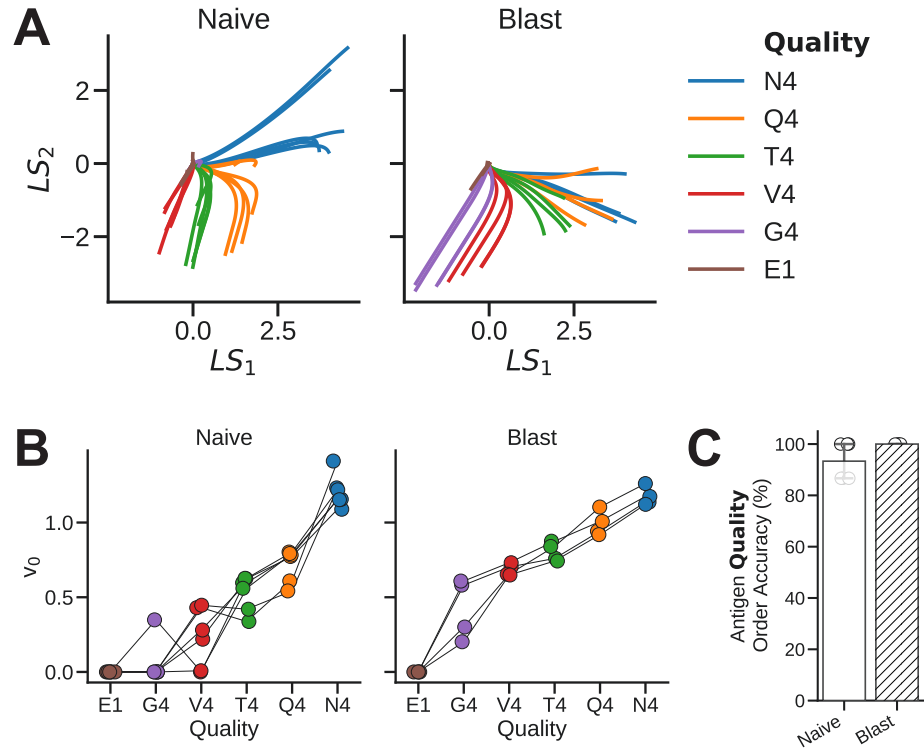


Figure B.6: “Antigen encoding with effector T cells. (A) Latent spaces of cytokine dynamics for OT-1 effector T cells (“Blast”) compared to OT-1 naïve T cells. **(B)** Immune velocity (v_0) as a measure of antigenicity was defined using naïve T cells (left panel) and remained accurate for blast T cells (right panel) Notice that for blast cells the trajectories in the latent space are rotated with respect to trajectories for the naïve cells. **(C)** Despite the rotation in latent space, immune velocity (v_0) is comparable and has high order accuracy in both naïve and blast T cells, showing the relevance of this parameter to compare different immune contexts. The data used in this figure are from three independent experiments, each with two technical replicates.” (*Antigen encoding*, [1], SI)

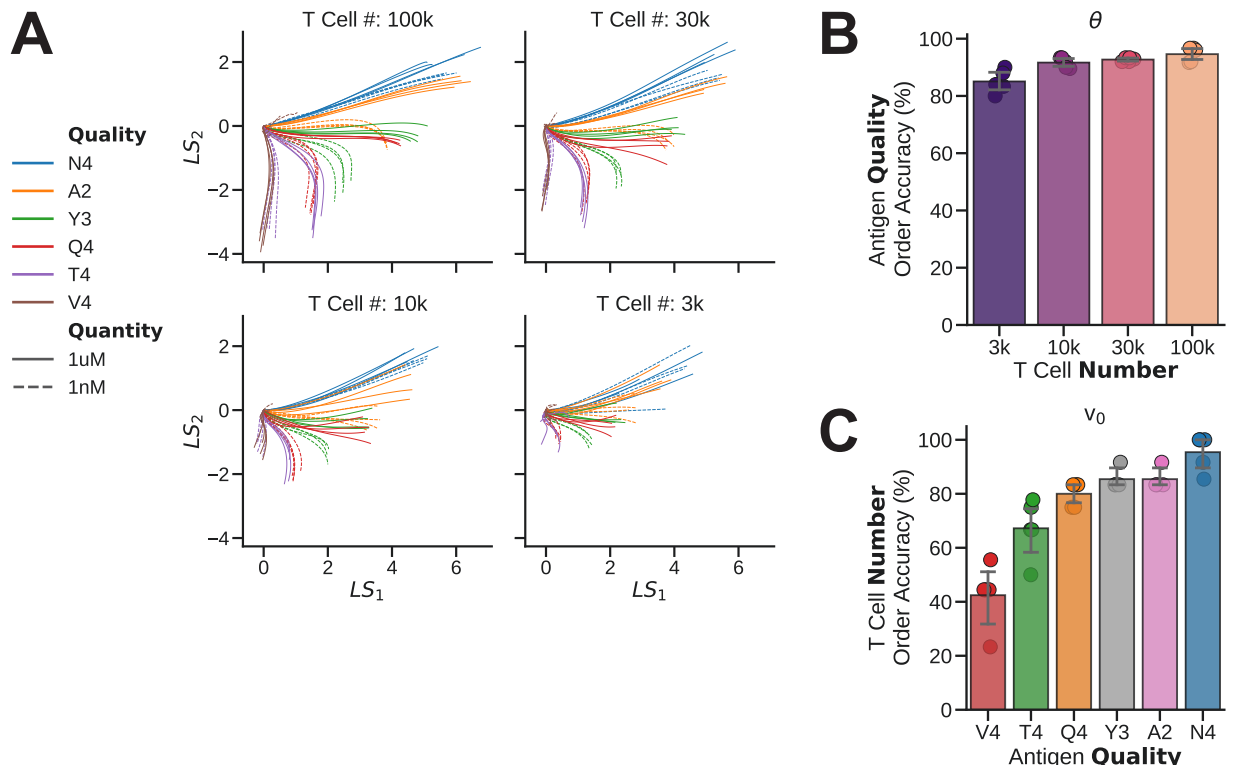


Figure B.7: “Latent space and order accuracy for different number of OT-1 T cells. (A) The latent space projections (trained with 100k T cell series) remain accurate in classifying antigens at lower cell numbers (30k, 10k and 3k). (B) Therefore, for any T cell number, the immune pitch parameter θ still yields high order accuracy to sort trajectories according to antigen quality (see Fig. 2.14 for the definition of order accuracy). (C) The immune velocity parameter v_0 additionally reflects the order of latent space trajectories according to T cell number, for any given antigen, with good accuracy, especially when a stronger antigen is presented. The data used in this figure are representative of three independent experiments. (Antigen encoding, [1], SI)

Appendix C

Existence and uniqueness of the solution for bound receptors

Recall that in kinetic proofreading models for mixtures of two TCR antigens, the equation to solve for the number of receptors bound to each type of antigen are

$$\begin{aligned} 0 &= \kappa\tau_1(L_1 - R_{b,1})(R_{\text{tot}} - R_{b,1} - R_{b,2}) - R_{b,1} \\ 0 &= \kappa\tau_2(L_2 - R_{b,2})(R_{\text{tot}} - R_{b,1} - R_{b,2}) - R_{b,2} . \end{aligned} \quad (3.8)$$

Then, isolating $R_{b,2}$ in terms of $R_{b,1}$ gives

$$R_{b,2} = R_{\text{tot}} - R_{b,1} - \frac{R_{b,1}}{\kappa\tau_1(L_1 - R_{b,1})} , \quad (3.9)$$

resulting in a cubic equation for $R_{b,1}$,

$$f(R_{b,1}) = p_0(R_{b,1})^3 + p_1(R_{b,1})^2 + p_2R_{b,1} + p_3 = 0 \quad (3.10)$$

where the coefficients are

$$\begin{aligned} p_0 &= \frac{\tau_1}{\tau_2} - 1 \\ p_1 &= -\left(\frac{\tau_1}{\tau_2} - 1\right)\left(R_{\text{tot}} + L_1 + \frac{1}{\kappa\tau_1}\right) - L_2 - \frac{\tau_1}{\tau_2}L_1 \\ p_2 &= \frac{\tau_1}{\tau_2}L_1^2 + \left(2\frac{\tau_1}{\tau_2} - 1\right)R_{\text{tot}}L_1 + \frac{L_1}{\kappa\tau_2} + L_1L_2 \\ p_3 &= -\frac{\tau_1}{\tau_2}R_{\text{tot}}L_1^2 . \end{aligned} \quad (3.11)$$

We now want to prove that for physically meaningful values of parameters – namely, τ_l , L_l , R_{tot} , and κ are all in $(0, \infty)$ – there is a unique real root $(R_{b,1}, R_{b,2})$ which is physically possible,

meaning that

$$R_{b,l} \in [0, L_l] \quad l \in \{1, 2\} \text{ (there can't be more than } L_l \text{ bound ligands of type } l) \quad (\text{C.1})$$

$$R_{b,1} + R_{b,2} \leq R_{\text{tot}} \text{ (there can't be more bound receptors than } R_{\text{tot}}) . \quad (\text{C.2})$$

We only need to consider the case $\tau_1 \geq \tau_2$, else we could relabel ligands $1 \leftrightarrow 2$ to enforce this condition (or equivalently, solve in terms of $R_{b,2}$ first). Hence, it is clear that coefficients p_0 and p_2 are positive while p_1 and p_3 are negative.

Lemma C.0.1. *A root $R_{b,1} \in (0, L_1)$ ensures that (i) $R_{b,1} + R_{b,2} \leq R_{\text{tot}}$ and (ii) $R_{b,2} \in (0, L_2)$.*

Proof. To prove this, notice that isolating $R_{\text{tot}} - R_{b,1} - R_{b,2}$ in both equations 3.8 gives

$$R_{\text{tot}} - R_{b,1} - R_{b,2} = \frac{R_{b,1}}{\kappa\tau_1(L_1 - R_{b,1})} = \frac{R_{b,2}}{\kappa\tau_2(L_2 - R_{b,2})}$$

By hypothesis, $R_{b,1} > 0$ and $L_1 - R_{b,1} > 0$, hence $\frac{R_{b,1}}{\kappa\tau_1(L_1 - R_{b,1})} > 0$ since other parameters are also positive; from the first equality, this means that $R_{\text{tot}} - R_{b,1} - R_{b,2} > 0$, thus proving (i). Moreover, from the second equality, $\frac{R_{b,2}}{\kappa\tau_2(L_2 - R_{b,2})} > 0$ which, since L_2 and other parameters are positive, is only possible if $R_{b,2} > 0$ and $L_2 - R_{b,2} > 0$, thus proving (ii). $\square$

Therefore, the problem reduces to proving (C.1) for $l = 1$, that is, to proving that the cubic equation with coefficients in 3.11 has a unique solution in $[0, L_1]$. We do this in two more steps.

Lemma C.0.2. *The cubic polynomial (3.10) has at least one real root $R_{b,1}$ in $[0, L_1]$.*

Proof. We prove this by evaluating $f(R_{b,1})$ defined in (3.10) at 0 and L_1 .

$$f(0) = p_3 = -\frac{\tau_1}{\tau_2} R_{\text{tot}} L_1^2 < 0$$

while

$$\begin{aligned} f(L_1) &= \left[\frac{\tau_1}{\tau_2} - 1 \right] L_1^3 + \left[- \left(\frac{\tau_1}{\tau_2} - 1 \right) \left(R_{\text{tot}} + L_1 + \frac{1}{\kappa\tau_1} \right) - L_2 - \frac{\tau_1}{\tau_2} L_1 \right] L_1^2 \\ &\quad + \left[\frac{\tau_1}{\tau_2} L_1^2 + \left(2\frac{\tau_1}{\tau_2} - 1 \right) R_{\text{tot}} L_1 + \frac{L_1}{\kappa\tau_2} + L_1 L_2 \right] L_1 - \frac{\tau_1}{\tau_2} R_{\text{tot}} L_1^2 \\ &= \frac{L_1^2}{\kappa\tau_1} > 0 \end{aligned}$$

Since f changes sign between 0 and L_1 , it has at least one real root in between. The root is not at 0 or L_1 exactly, unless $L_1 = 0$, which is not a relevant parameter value here. $\square$

The last step is to show there is never more than one root in $[0, L_1]$.

Lemma C.0.3. *The cubic polynomial (3.10) has only one root in $[0, L_1]$.*

Proof. Depending on the value of the cubic discriminant Δ_3 , the polynomial has either one ($\Delta_3 < 0$), two ($\Delta_3 = 0$) or three ($\Delta_3 > 0$) real roots; the case $\Delta_3 = 0$ means that a root is degenerate. If $\Delta_3 < 0$, the only root is in $[0, L_1]$ by Lemma C.0.2, and we are done. We only need to consider $\Delta_3 \geq 0$.

Just from the exact expression for the roots of a cubic polynomial, it is hard to tell where they are located. To make progress, we examine the derivative,

$$\frac{df}{dR_{b,1}} = 3p_0 R_{b,1}^2 + 2p_1 R_{b,1} + p_2. \quad (\text{C.3})$$

Evaluated at $R_{b,1} = 0$, it gives

$$\left. \frac{df}{dR_{b,1}} \right|_0 = p_2 = \frac{\tau_1}{\tau_2} L_1^2 + \left(2\frac{\tau_1}{\tau_2} - 1 \right) R_{\text{tot}} L_1 + \frac{L_1}{\kappa \tau_2} + L_1 L_2$$

which is always positive. Hence, we must decide between the three cases illustrated in figure C.1, which are the only possibilities for a cubic with $p_0 > 0$, at least one root and a change of sign in $[0, L_1]$, and an upwards branch intersecting at $R_{b,1} = 0$.

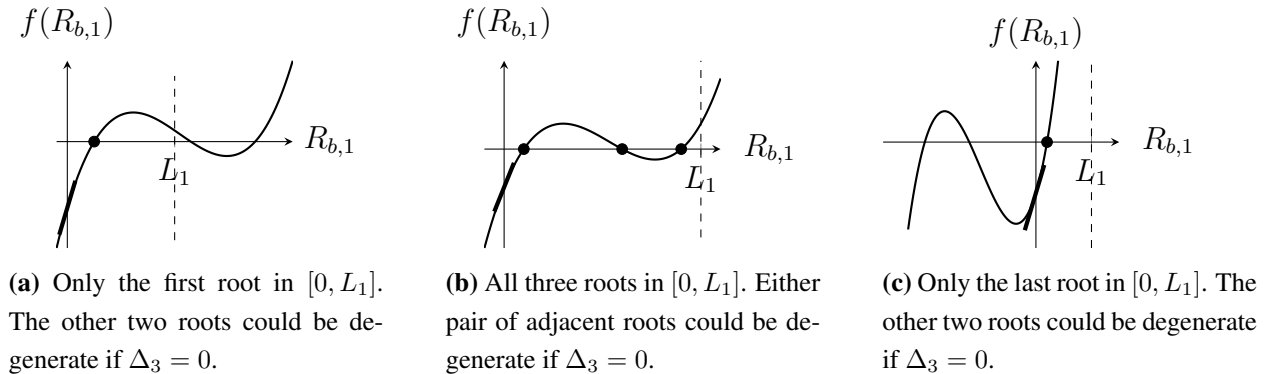


Figure C.1: The three possible ways the real roots of f , in the case $\Delta_3 \geq 0$ relevant here, can be distributed to follow Lemma C.0.2 and to have $\frac{df}{dR_{b,1}} > 0$ at the intercept.

These cases can be distinguished by the positions of the zeros of $f' = \frac{df}{dR_{b,1}}$.

- Case C.1a necessarily has one zero of f' at $R_{b,1} > L_1$, and at most one in $[0, L_1]$.
- Case C.1b has both zeros of f' in $[0, L_1]$.
- Case C.1c has both zeros of f' at $R_{b,1} < 0$.

The zeros of f' are given by the quadratic formula,

$$R_{b,1}^* = \frac{-2p_1}{3p_0} \pm \frac{1}{3p_0} \sqrt{4p_1^2 - 6p_0p_2}$$

Since $p_0, p_2 > 0$ while $p_1 < 0$, both solutions are positive, since $\frac{-2p_1}{3p_0} > 0$ and $\sqrt{4p_1^2 - 6p_0p_2} < |2p_1|$. This rules out case C.1c.

The rightmost zero of f' (positive sign in the quadratic formula) is outside of $[0, L_1]$, since

$$\begin{aligned} R_{b,1,+}^* &= \frac{-2p_1}{3p_0} + \frac{1}{3p_0} \sqrt{4p_1^2 - 6p_0p_2} \\ &> \frac{-2p_1}{3p_0} = \frac{2}{3(\frac{\tau_1}{\tau_2} - 1)} \left(\left(\frac{\tau_1}{\tau_2} - 1 \right) \left(R_{\text{tot}} + L_1 + \frac{1}{\kappa\tau_1} \right) + L_2 + \frac{\tau_1}{\tau_2} L_1 \right) \\ &= \frac{2}{3} \left(L_1 + L_1 \frac{\frac{\tau_1}{\tau_2}}{\frac{\tau_1}{\tau_2} - 1} + R_{\text{tot}} + \frac{1}{\kappa\tau_1} + \frac{L_2}{\frac{\tau_1}{\tau_2} - 1} \right) \\ &> \frac{2}{3} \left(L_1 + L_1 \frac{\frac{\tau_1}{\tau_2}}{\frac{\tau_1}{\tau_2} - 1} \right) \\ &> \frac{2}{3} (L_1 + L_1) = \frac{4}{3} L_1 \\ &> L_1 \end{aligned}$$

This rules out case C.1b; hence, the real roots of the cubic polynomial $f(R_{b,1})$ are distributed as in case C.1a: there is only one in $[0, L_1]$. $\square$

Combining the three lemmas, we can conclude that there is always a unique real root of $f(R_{b,1})$ in $[0, L_1]$, which ensures the system of equations for the number of receptors bound to each type of ligands $(R_{b,1}, R_{b,2})$ always has a unique physically meaningful solution satisfying conditions (C.1) and (C.2).

Appendix D

Parameter estimation methods

“For parameter estimations, since the revised TCR-CAR model expands on a simpler TCR model, our general strategy was to first estimate the TCR parameters on TCR/TCR data, before estimating CAR parameters on part of the TCR/CAR dataset, and predicting the remaining part of that dataset. Additionally, some parameters were fixed, either taken from the literature, derived from additional calibration experiments, or chosen to perform simplifying assumptions ensuring a better convergence of optimization.” (*TCR/CAR antagonism*, [3])

D.1 Calibration of receptor and ligand abundances

To faithfully compare our model to experimental data, we wanted to use realistic values of the antigen quantity parameters, L^p , in the model. Our collaborators therefore experimentally calibrated the absolute number of ligands per APC (for a typical APC cell size). To calibrate the abundance of TCR antigens, which are short peptides loaded on major histocompatibility complex (MHC) molecules, our collaborators measured the MHC loading efficiency of peptide pulsing, *i.e.*, the fraction of MHC loaded with a peptide, as a function of the concentration of the peptide solution. They performed the assay with RMA-S antigen presenting cells, which have a deficiency in their pathway to present intracellular-derived peptides, such that they only load peptides present in the extracellular medium. They pulsed them with different doses of OVA-derived (OT-1) peptides and measured the fraction of loaded MHC for each dose, using a fluorescent antibody.

“A Michaelis-Menten model with background was fit on the H-2Kb mean fluorescence intensity (MFI) curves of all RMA-S loading experiment replicates and each ovalbumin-derived (OVA) peptide to determine their loading EC_{50} s (figure 3.13c) in terms of the peptide pulsing concentration $[P]$:

$$MFI = MFI_{background} + \text{Amplitude} \frac{[P]}{[P] + EC_{50}}. \quad (D.1)$$

Cell	Molecule	CI _{2.5%}	Geo. mean	CI _{97.5%}
B16	CD19	7.7×10^5	7.9×10^5	8.0×10^5
B16	MHC	4.7×10^2	4.8×10^2	5.0×10^2
B6 Splenocyte	CD19	7.6×10^3	8.0×10^3	8.4×10^3
B6 Splenocyte	MHC	5.20×10^4	5.27×10^4	5.35×10^4
BEAS2B	Her2	1.68×10^4	1.71×10^4	1.73×10^4
BEAS2B	MHC	2.997×10^6	3.008×10^6	3.016×10^6
E2aPBX 19KO	CD19	3.74×10^2	3.80×10^2	3.87×10^2
E2aPBX 19KO	MHC	4.42×10^4	4.48×10^4	4.54×10^4
E2aPBX WT	CD19	1.05×10^5	1.07×10^5	1.08×10^5
E2aPBX WT	MHC	5.73×10^4	5.80×10^4	5.88×10^4
Nalm6 19KO	CD19	4.57×10^2	4.64×10^2	4.72×10^2
Nalm6 19KO	MHC	6.19×10^5	6.23×10^5	6.27×10^5
Nalm6 19hi	CD19	4.60×10^5	4.62×10^5	4.64×10^5
Nalm6 19hi	MHC	8.09×10^5	8.14×10^5	8.19×10^5
Nalm6 19int	CD19	3.12×10^4	3.15×10^4	3.18×10^4
Nalm6 19int	MHC	5.86×10^5	5.91×10^5	5.95×10^5
Nalm6 19low	CD19	4.31×10^3	4.39×10^3	4.47×10^3
Nalm6 19low	MHC	8.1×10^5	8.3×10^5	8.4×10^5
OT1 Blast	CAR	4.07×10^2	4.16×10^2	4.27×10^2
OT1 Blast	TCR	1.515×10^5	1.519×10^5	1.523×10^5
OT1 CAR	CAR	9.5×10^5	9.8×10^5	10.1×10^5
OT1 CAR	TCR	1.211×10^5	1.218×10^5	1.225×10^5
OT1 Naive	TCR	9.9×10^4	10.0×10^4	10.1×10^4
PC9	Her2	3.28×10^4	3.33×10^4	3.37×10^4
PC9	MHC	7.3×10^5	7.4×10^5	7.5×10^5

Table D.1: “Abundance of surface molecules on various cell types, related to figure 3.13. Molecule numbers are as determined in figure 3.13 and used in the model. Bootstrapped 95 % confidence intervals on the geometric mean computed across cells are also given (lower limit: CI_{2.5%}, upper limit: CI_{97.5%}).” (*TCR/CAR antagonism*, [3], Table S2)

Since all OVA peptides had similar loading EC_{50} s, the geometric average was used to determine a single loading EC_{50} parameter, $K_{D,load}$, to build a universal Michaelis-Menten loading curve from which all peptide-pulsing concentrations $[P]$ were converted to fractions of loaded MHCs,

$$\frac{\text{pMHC}}{\text{Total MHC}} = \frac{[P]}{[P] + K_{D,load}}, \quad (\text{D.2})$$

and ultimately to absolute numbers of TCR ligands, using total MHC numbers quantified for each type of antigen-presenting cell (APC) in figure 3.13 and table D.1. We found $\log_{10}(K_{D,load}/1\text{nM}) = 2.0^{+0.2}_{-0.3}$ (95 % CI of all replicates and peptides, figure 3.13C-D)." (*TCR/CAR antagonism*, [3])

The abundance of other surface molecules, either CAR antigens or TCR and CAR receptors, was also estimated using antibody-based fluorescence assays for each cell type used in this project. The resulting surface antigen numbers are shown in figure 3.13A–D and table D.1.

D.2 Peptide binding time conversion

"In our models, we assumed that all ligands of receptor type ρ have the same binding on-rate (κ^ρ) and are distinguished only by their off-rates, or equivalently by their binding times τ (inverse of the off-rate). We therefore needed a systematic, model-independent way of assigning a binding time τ to TCR antigens based on their experimental EC_{50} s. We assumed EC_{50} s are defined by an idealized KPR power law passing a threshold of activation Θ , that is,

$$EC_{50}\tau^N = \Theta \quad (\text{D.3})$$

where EC_{50} is measured in antigen pulse concentration. The threshold Θ is set by choosing a reference τ_{ref} for one peptide with known $EC_{50,\text{ref}}$. Then, the τ of any other antigen is set by

$$\tau_{\text{pep}} = \tau_{\text{ref}} \left(\frac{EC_{50,\text{pep}}}{EC_{50,\text{ref}}} \right)^{-1/N} \quad (\text{D.4})$$

We used $N^T = 6$, as in our improved AKPR model. For OT-1 antigens, we used the same EC_{50} s as in Achar *et al.* [1] and we set as a reference $\tau = 10$ s for the N4 antigen. We furthermore set $\tau = 0.5$ s for the E1 antigen since it produces

no measurable response in our assays, causing its EC_{50} to be ill-determined. This conversion rule is illustrated in figure 3.13E.

For other antigens (NY-ESO antigens, HHAT/p8F), we carried a dose response experiment based on the percentage of $CD25^+$ cells at 24 h. We used the $CD25$ EC_{50} of N4 (OVA) as the reference point at $\tau = 10$ s for these other antigens. For figure 4.13, we used the $4-1BB^+$ EC_{50} s provided in Łuksza *et al.* [258], and we inferred missing values for CMV-derived antigens with the correction method described below in section 4.4.3." (*TCR/CAR antagonism*, [3])

D.3 Parameter values in the initial AKPR model

"Before diving into our main parameter estimation efforts with the revised model, a few words about parameters in the initial AKPR model. We set most TCR parameters to their original value in François *et al.* [129], except for φ , $C_{m,th}$, I_{tot} , and m , which we estimated by MCMC as described below (section D.5 and table D.4). For the extension of this model to TCR/CAR antagonism, we did not perform an exhaustive parameter search; we only tested representative values (figure 4.5) to conclude the model could evidently not fit all data. We kept CAR parameters in the same relationship to their TCR counterparts as in the revised model (see table D.3), *e.g.*, we set $\kappa^C = 10\kappa^T$." (*TCR/CAR antagonism*, [3])

D.4 Fixed parameter values

"Some parameters of the revised model were determined from the literature, as summarized in the first part of table D.3.

- The TCR ligand on-rate, κ^T , was taken from François *et al.* [129], since it only influences the absolute scale of all C_n^T s via R_b^T , as seen in equations (4.13)–(4.17); the threshold $C_{m,th}^T$ could be rescaled to compensate.
- We set $N^T = 6$ for regular 6Y TCRs and $N^T = 4$ for 6F TCRs, to enforce the idea that TCRs with fewer ITAMs need to undergo fewer proofreading steps before activation (note: this number of steps N is not literally the number of ITAMs).
- The CAR ligand on-rate κ^C was made $10\times$ larger than κ^T , and the CAR antigen unbinding time τ^C was set to $\tau^C = 500$ s, $100\times$ larger than typical TCR τ^T ,

such that overall CAR ligands have an affinity $1000\times$ larger than TCR ligands, as in Harris *et al.* [286].

- Reference Harris *et al.* [286] also found that the CAR proofreading rate, φ^C , should be $100\times$ - $1000\times$ smaller than φ^T , thus we set $\varphi^C = \varphi^T/200$.
- We set diagonal elements $\gamma^\rho_\rho = 1$, as any other value can be compensated by rescaling off-diagonal elements and I^ρ_{th} , as seen in equation (4.9). For the same reason, I_{tot} was normalized to 1.

Receptor (R^T , R^C) and ligand (L^T , L^C) numbers were determined from surface molecule abundance calibration experiments (figure 3.13 and section D.1). table D.1 lists these numbers for the different cell types used in this work. The number of loaded pMHC at a given pulse concentration was computed by multiplying the total MHC number by the fraction of loaded MHCs, given by equation (D.2).

Parameter ψ_0^C was small and ill-determined in preliminary MCMC simulations, but it needed a non-zero value to ensure some minimal antagonism and to help convergence. Its value could be chosen arbitrarily as long as it was small enough but non-zero, but to be more rigorous, we estimated its order of magnitude by adding a minimal condition on inhibition: when I is fully activated, such that $\psi^C \approx \psi_0^C$, the reduction¹ in C_{NC}^C must be less than the minimum $FC_{\text{TCR/CAR}}$ in the data used for parameter fitting (table D.2). Working this condition out from equation (4.15), and making ψ_0^C $3\times$ smaller than this upper bound, we thus set

$$\psi_0^C = \frac{1}{3\tau^C} \frac{\Phi^C FC_{\min}}{1 - \Phi^C FC_{\min}} \quad (\text{D.5})$$

where Φ^C is defined in eq. (4.16) and evaluated at $\tau^C = 500$ s and $\varphi^C = \varphi^T/100$.¹ (TCR/CAR antagonism, [3])

D.5 Overview of the parameter estimation strategy and of the generation of predictions

“Remaining parameters were estimated using the Markov Chain Monte Carlo (MCMC) method. Our log-likelihood and posterior functions for each MCMC simulation, defined in section D.7, compared model outputs to experimental measurements in terms

¹For $f^C = 1$; ψ_0^C could be larger when more than 1 step has rate ψ^C , in which case the absolute limit we set is certainly small enough.

antagonism fold-change FC_{R_1/R_2} , formally defined below in section D.6. Section D.8 details the MCMC algorithm employed (simulation lengths, proposal distribution, convergence, etc.). Table D.2 summarizes which parameters were fitted in each MCMC run and provides information on their convergence, as well as the datasets used for each run. Table D.3 lists the bounds imposed on parameters fixed or fitted, and table D.4 provides their best fit values and confidence intervals.

We used the following **fitting** strategy:

1. We fitted TCR-related parameters of both the François *et al.*, 2013 model (subsection 3.3.3) and the revised AKPR model for TCR/TCR antagonism (section 3.5.3) on cytokine data from T cells with 6Y (10 ITAMs) TCRs (figure 3.12). We noticed how inadequate the fitting of the data was for the François *et al.*, 2013 model
2. Using TCR-related parameters found in step 1, we fitted parameters related to CARs and their interaction with TCRs in the revised AKPR model (section 4.3.3) on TCR/CAR antagonism data for 6Y (10 ITAMs) TCRs, 3-ITAM CARs at 1 μ M and 1 nM TCR antigen densities (figure 4.3D–E).
3. We fitted the revised AKPR model’s TCR-related parameters for 6F (4 ITAMs) TCRs on TCR/TCR antagonism data (figure 3.19) . These were combined with the values of CAR-related parameters found in step 2 for prediction of TCR/CAR antagonism in these altered TCR T cells (figure 4.3F).

We used the estimated parameter distributions to compare the model to data and to make various predictions in T cells with altered receptors and in various tumor cell lines. We used the following **validation/prediction** strategy:

4. We compared models of TCR/TCR antagonism (François *et al.*, 2013 vs revised AKPR) using parameter distributions found in MCMC simulations 1 (figure 3.12).
5. We checked that the model captures accurately CAR/TCR antagonism for 3-ITAM CAR as a function of TCR antigen density with parameters of MCMC simulations 1 and 2 (figure 4.3E, top).
6. We generated predictions for 1-ITAM CAR and 6Y TCR (figure 4.3E, bottom and figure 4.9B–D) using parameters of MCMC simulations 1 and 2.

7. We generated predictions of CAR/TCR antagonism for T cells with 6F TCR (4 ITAMs[2]) (figure 4.3E, bottom and figure 4.9A–D), using parameters of MCMC simulations 2 and 3.
8. We generated predictions of CAR/TCR antagonism against various tumor cell lines (figures 4.11, 4.13, 4.15; figure 4.14) with appropriate model parameters, and with tumor antigen numbers detailed in section D.1 and table D.1 .

In addition to the main MCMC simulations, predictions in points 6 to 8 required adjusting amplitudes and thresholds of Z^T or Z^C , which was done on independent measurements of T cell responses to single CAR or TCR antigens (*i.e.*, not on antagonism data in response to mixtures of TCR and CAR antigens). Sections D.9 and 4.4.1 provide details about these model predictions.” (*TCR/CAR antagonism*, [3])

D.6 Calculation and statistics of antagonism fold-change

“To compare model predictions of antagonism with T cell activation measurements (cytokines IFN γ , IL-2, TNF; % CD25⁺ cells, etc.), we needed a dimensionless quantity that could be computed from both model and data. We used the fold-change ratio introduced in the main text, $FC_{\text{TCR/TCR}}$ for TCR/TCR antagonism, or $FC_{\text{TCR/CAR}}$, for TCR/CAR antagonism. For antagonism of receptor kind X by the TCR,

$$FC_{\text{TCR/X}} = \frac{\text{Output}(X \text{ agonist} + \text{TCR antagonist})}{\text{Output}(X \text{ agonist})} \quad (\text{D.6})$$

where Output is a measure of T cell activation. A ratio $FC < 1$ indicates antagonism: the response is lowered by adding a TCR antigen. A ratio $FC > 1$ indicates enhancement of the response by the TCR signal.

For receptor type $X = \text{TCR}$ (figure 3.12), we directly used C_{NT}^T as the model output. The denominator is $C_{NT,1}^T (L_2^T = 0)$, the number of active receptors bound to agonist ligands in absence of antagonist ligands; the numerator is $C_{NT,1}^T + C_{NT,2}^T$, the number of active receptors bound to either ligand type in the mixture.

For $X = \text{CAR}$ (figure 4.3), in the model, Output is Z^{tot} . In the denominator, $Z^{\text{tot}} = Z^C$ for the CAR antigen (agonist) alone ($L^T = 0$). In the numerator, $Z^{\text{tot}} = Z^C + Z^T$ in response to the mixture of the CAR antigen and TCR antigen.” (*TCR/CAR antagonism*, [3])

Run	Agonist lig- and	TCR antago- nists	Nb. experiments	Replicates per exp.	Timepoints	Data per cond. (n_p)
6Y TCR/TCR (both models)	N4, 10 pM to 1 nM	None to Q4, 1 nM & 1 μ M	3	1	6	18
6F TCR/TCR	N4 10 pM to 1 μ M	None to V4, 1 nM to 1 μ M	2	3 biological, 1-2 technical	1-12	6-42
TCR/CAR	CD19	None to N4, 1 nM and 1 μ M	1 - 2*	3	12	36-72
Run	Fitted parameters		Steps	$\max_{\theta} \tau_{\text{int},\theta}$ (steps)	Burn-in ($20\tau_{\text{int},\max}$) (fraction of simulation)	
6Y TCR/TCR, initial AKPR	$\varphi^T, C_{m,\text{th}}^T, I_{\text{tot}}$		40,000	234.9	0.117	
6Y TCR/TCR, revised AKPR	$\varphi^T, C_{m,\text{th}}^T, I_{\text{th}}^T, \psi_0^T$		40,000	79.2	0.040	
6F TCR/TCR, revised AKPR	$\varphi^T, C_{m,\text{th}}^T, I_{\text{th}}^T, \psi_0^T$		40,000	123.1	0.062	
TCR/CAR	$C_{m,\text{th}}^C, I_{\text{th}}^C, \gamma^T_C,$ $\gamma^C_T, \tau_c^T, \tau_c^C$		80,000	64.1	0.016	

Table D.2: “Information on MCMC simulations, related to figures 3.12 and 4.3. **Top section:** details of datasets used for each MCMC parameter estimation run. *The 1 nM TCR antigen condition was tested in one experiment; 1 μ M was present in both. **Bottom section:** information on parameter space, length, and convergence of MCMC simulations. 6Y: 10-ITAM TCR, 6F: 4-ITAM TCR.” (*TCR/CAR antagonism*, [3], Table S4)

	Parameter	Constraints	Value	Units
Fixed	κ^T	[129]	10^{-4}	molec. $^{-1}$ s $^{-1}$
	κ^C	$10 \times \kappa^T$ [286]	10^{-3}	molec. $^{-1}$ s $^{-1}$
	N^T	Choice	6Y TCR: 6	-
			6F TCR: 4	-
	N^C	Choice	3-ITAM CAR: 3	-
			1-ITAM CAR: 1	-
	I_{tot}	Normalized	1	-
	ψ_0^C	(section D.4)	8.0×10^{-5}	s $^{-1}$
	γ_{ρ}^{ρ}	Normalized	1	-
	τ_{CD19}^C	Intermediate agonist [1] with $\sim$ nM K_D [286]	500	s
Fitted by MCMC on TCR/TCR antagonism	φ^T	6Y: [0.1, 5.0] 6F: [0.05, 5.0]	Supp. Table D.4	s $^{-1}$
	$C_{m,th}^T$	[1, 100 R_{tot}^T]		molec
	I_{th}^T	[10^{-5} , 10^3 I_{tot}]		-
	ψ_0^T	[10^{-8} , 10^{-2}]		s $^{-1}$
	m^T	6Y: Grid {2...5}	4	-
		6F: Grid {1...4}	3	-
	f^T	Grid {1, 2}	6Y & 6F: 1	-
	k_I^T	6Y: Grid {1, 2}	1	-
		6F: choice	1	-
From 6Y TCR fit	φ^C	Less than 0.01 φ^T [286]	$\varphi^T/200$	s $^{-1}$
Fitted by MCMC on TCR/CAR antagonism	$C_{m,th}^C$	[1, 1000 R_{tot}^C]	Supp. Table D.4	molec
	I_{th}^C	[10^{-5} , 10^3 I_{tot}]		-
	$\gamma_{C,T}^T$	[0.1, 1.0]		-
	$\gamma_{C,T}^C$	[0.01, 10^4]		-
	τ_c^T	[1, 30]		s
	τ_c^C	[50, 5000]		s
	m^C	3-ITAM: Grid {1, 2, 3}	2	-
		1-ITAM: choice	1	-
	f^C	Grid {1, 2, 3}	1	-
	k_I^C	Grid {1, 2}	1	-
Fitted on single antigen response	$\alpha_A^{T,6F}$	6F TCR	0.51	-
	$\alpha_{\tau}^{T,6F}$	6F TCR	0.36	-
	$\alpha_A^{C,1CAR}$	1-ITAM CAR	1.16	-
	$\alpha_{\tau}^{C,1CAR}$	1-ITAM CAR	2.60	-

Table D.3: “Constraints on mathematical model parameters in MCMC simulations, related to figures 3.12 and 4.3. Fixed and estimated parameters in the revised AKPR model, with the parameter fit boundaries or the constraints from which other parameters were fixed. The last columns provide the value of fixed parameters, or the value of integer parameters estimated by grid search (“Grid”), by running MCMC simulations for each possible integer value. Real-valued parameters estimated by maximum a posteriori probability are provided in table D.4 with confidence intervals.” (*TCR/CAR antagonism*, [3], Table S5)

Model	Parameter	P _{5%}	Best	P _{95%}
6Y TCR/TCR, initial AKPR	$\log_{10} \varphi^T$	-0.20	0.70	0.68
	$\log_{10} C_{m,th}^T$	0.03	0.01	1.84
	$\log_{10} I_{tot}^T$	7.220	8.167	8.165
6Y TCR/TCR, revised AKPR	$\log_{10} \varphi^T$	-0.99	-0.93	-0.69
	$\log_{10} C_{m,th}^T$	2.5	3.3	4.2
	$\log_{10} I_{th}^T$	-4.9	-4.2	-3.6
	$\log_{10} \psi_0^T$	-7.8	-5.5	-4.8
6F TCR/TCR, revised AKPR	$\log_{10} \varphi^T$	-1.295	-1.301	-1.020
	$\log_{10} C_{m,th}^T$	1.7	2.0	4.3
	$\log_{10} I_{th}^T$	-4.8	-2.8	-2.5
	$\log_{10} \psi_0^T$	-7.7	-4.6	-3.8
TCR/CAR	$\log_{10} C_{m,th}^C$	3.6	4.5	5.2
	$\log_{10} I_{th}^C$	-4.5	-3.7	-2.8
	$\log_{10} \gamma_C^T$	-0.937	-0.058	-0.050
	$\log_{10} \gamma_T^C$	1.7	1.9	3.8
	$\log_{10} \tau_c^T$	0.55	0.61	0.75
	$\log_{10} \tau_c^C$	2.43	2.50	2.56

Table D.4: “Value of mathematical model parameters estimated by MCMC, related to figures 3.12 and 4.3. Parameter values estimated in each MCMC run, for the best grid search value of integer parameters (k_I, m, f) . The best value is the maximum a posteriori probability point, while $P_{5\%}$ and $P_{95\%}$ are limits of the 90 % confidence interval on each parameter, *i.e.*, the 5th and 95th percentiles of the marginal posterior distribution. Shaded rows are those where the best estimate lies outside the 5th-95th percentile range, because the marginal distributions are highly skewed. This occurs in particular for the initial AKPR model, whose fit to data is poor.” (*TCR/CAR antagonism*, [3], Table S6)

Antagonism ($FC < 1$) occurs when a weak TCR signal lowers Z^C via the inhibitory module I . Enhancement occurs when the TCR signal is strong enough for Z^T to compensate the reduction in Z^C .

“To compute experimental values of $FC_{\text{TCR}/X}$, we collected cytokine time series of T cells (naive OT-1 T cells for $X = \text{TCR}$, OT-1/CAR T cells for $X = \text{CAR}$) responding to X agonists alone or in mixture with TCR antagonists. For each replicate of a given antigen mixture, we computed the FC ratio at each acquired time point, with the corresponding “no TCR antagonist” replicate in the denominator. We then computed the geometric average of FC across time points and replicates of each condition p (i.e. a choice of TCR antigen, TCR antigen density, CAR and TCR constructs).

As an error bar σ_p on these experimental values, we computed the 95% confidence interval on the mean of $\log_2 FC$ across replicates and time points, supposing normally distributed residuals. In equation,

$$\sigma_p = t_{(0.975, n_p-1)} \times \sqrt{\frac{\widehat{\text{Var}}[\log_2 FC_p]}{n_p}} \quad (\text{D.7})$$

where n_p is the number of time points and replicates of a given condition p , $\widehat{\text{Var}}$ is the sample variance estimator, and $t_{(0.975, n_p-1)}$ is the 97.5th percentile of the Student's t-distribution with $n_p - 1$ degrees of freedom [291]. The numbers of points in each dataset we used for MCMC are provided in table D.2.

For MCMC parameter estimation, we computed FC from IL-2 concentration time series: since all cytokines gave identical antagonism patterns (Figure 4.2d, bottom), we focused on one of them, IL-2, to have consistent FC amplitudes to fit.” (*TCR/CAR antagonism*, [3])

D.7 Likelihood function and prior distribution

“We used a least-squares log-likelihood function, considering residuals of model FC compared to data FC in $\log_2$ scale, meaning that

$$\log P(\text{data} \mid \theta) \propto - \sum_{\text{data } p} w_p \left(\frac{\log_2 FC_{\text{model}}(\tau_p^T, L_p^T) - \log_2 FC_{\text{data},p}}{\sigma_p} \right)^2 \quad (\text{D.8})$$

where $FC = FC_{\text{TCR/X}}$, defined in eq. (D.6), where, w_p is the weight assigned to point p , and where σ_p is the confidence interval in $\log_2$ scale of the FC of data point p , defined in eq. (D.7). For TCR/TCR antagonism, used to estimate TCR-related parameters, we gave more weight in eq. (D.8) to data points coming from the smallest agonist concentration (10 pM), since this is the condition giving antagonism patterns most similar to TCR/CAR antagonism, and thus the most important to capture. We found that applying a weight factor of $w_p = 3$ to residuals of these data points, and $w_p = 1$ for other agonist concentrations, provided fits that best transferred to TCR/CAR antagonism.

We performed MCMC in a log-scaled parameter space. We used a uniform prior distribution on the log of parameters, since we lacked good prior knowledge of the parameter values in our phenomenological model. The boundaries are detailed in the “Constraints” column of table D.3. From Bayes’ theorem, the posterior distribution sampled by the MCMC algorithm was therefore equal to the likelihood within these parameter boundaries, and zero outside.” (*TCR/CAR antagonism*, [3])

D.8 MCMC simulation details

“At each main parameter estimation step outlined in section D.5, we performed a grid search over integer-valued parameters m^ρ , f^ρ , k_I^ρ : for each possible combination of these integers, we ran a separate MCMC simulation in the space of the remaining real-valued parameters. For TCR/TCR antagonism results (steps 1, 3), we selected the (k_I^T, m^T, f^T) combination yielding the best fit (highest posterior probability). For TCR/CAR antagonism (step 2), all combinations gave equally good fits (within 5 % of each other), hence we selected $m^C = 2$, $f^C = k_I^C = 1$, as it also provided the best predictions for 4-ITAM CARs and 10-ITAM TCRs.

MCMC simulations were performed with the *emcee* package [245] in Python, which relies on an ensemble of walkers and affine-invariant proposal distributions. We used $n_w = 32$ walkers and a proposal distribution mixing so-called “walk” (affine-invariant), differential evolution, and “snooker” moves with proportions (0.2, 0.6, 0.2) (we found this combination to sample our parameter space more efficiently than pure walk moves).”

We set the mean stretch factor (γ_0) of the differential evolution move to $2\sqrt{2D}$, where D is the number of parameters, and the γ parameter of the snooker move to 1.5. These were slightly lower than the default values, to favor smaller steps and avoid missing narrow maxima.

“The n_w walkers were initialized at uniformly random positions in the log-scaled parameter space within the prior distribution boundaries, to reduce the risk of getting all walkers stuck in a local optimum.

We ran MCMC simulations for durations (in numbers of steps) listed in table D.2. To ascertain convergence to the stationary distribution, we computed normalized 1D autocorrelation functions, $\rho_{\theta\theta}(t)$, and the integrated correlation time, $\tau_{\text{int},\theta}$ for each parameter and each walker. We used the automatic windowing estimator described in Sokal [246]:

$$\hat{\tau}_{\text{int},\theta}(M) = \frac{1}{2} + \sum_{t=0}^M \hat{\rho}_{\theta\theta}(t) \quad (\text{D.9})$$

where M is the smallest window size such that $M \geq 5\hat{\tau}_{\text{int},\theta}(M)$. We averaged the autocorrelation estimators $\hat{\rho}_{\theta\theta}(t)$ and $\hat{\tau}_{\text{int},\theta}$ across walkers to reduce their variance, following Foreman-Mackey *et al.* [245]. We checked convergence of the estimator $\hat{\tau}_{\text{int},\theta}$ itself by also computing it over the first half of the simulation only; we considered $\hat{\tau}_{\text{int},\theta}$ an accurate estimate of the true $\tau_{\text{int},\theta}$ if the two values agreed within 10 %. We also verified the algorithm’s proper behaviour by checking that the acceptance fraction of proposed transitions was between 0.2 and 0.5. Finally, to establish parameter distributions from MCMC samples, we dropped a burn-in phase of $20 \max_{\theta} (\hat{\tau}_{\text{int},\theta})$. With the duration of our simulations, this corresponded to only 2–20 % of the total length. Hence, our simulations were run longer than the recommended minimum of $50\hat{\tau}_{\text{int}}$. These MCMC convergence metrics are shown in figures 3.14, 3.17, 4.8, 3.19B.

The posterior distributions sampled by MCMC are shown in figures 3.12E, H, 4.3D, and 3.19A. The non-normal character of these distributions comes from possible compensations between combinations of parameters. For instance, when $C_{m,th}^T \gg$ typical values of C_m^T , only the combination $I_{th}^T C_{m,th}^T / I_{tot}$ determines the strength of the inhibitory module in $\psi^T(I)$ (combining equations 3.12 and 3.21). This reflects the “sloppiness” inherent to most biological models [158].

To estimate the best parameter fits, we took the MCMC sample with the best posterior probability. We generated 90 % confidence intervals on model fits and predictions, for instance in figure 4.3E, by randomly pulling 1000 parameter vector samples from the MCMC results, computing model outputs for each, and computing the 5th–95th percentiles of model FC values at each antigen τ^T ." (*TCR/CAR antagonism*, [3])

By construction of the MCMC algorithm, this parameter vector is in the most densely sampled region of parameter space, around the global maximum of the posterior distribution. Visual inspection of MCMC corner plots in Figures 2, 3 confirms it is the case, keeping in mind that the mode of any skewed multivariate distribution is different from the marginal modes.

D.9 Predictions in other T cell types

“MCMC parameter sampling was primarily performed for T cells with regular TCRs (10 ITAMs) and 3-ITAM CARs. For other cell types, a few parameters had to be adjusted on separate data to generate TCR/CAR antagonism predictions in figure 4.3. The rest of parameters related to TCR/CAR interactions – $C_{m,th}^C$, I_{th}^C , γ_C^T , γ_T^C – estimated by MCMC on 3-ITAM CAR, 10-ITAM TCR data (step 2) were kept the same throughout. Parameter φ^C was also kept unchanged, equal to $\varphi^{T,6Y}/200$.” (*TCR/CAR antagonism*, [3])

D.9.1 6F (4 ITAM) TCRs

“As stated in section D.5, step 3, a separate MCMC simulation established the value of TCR-related parameters – φ^T , $C_{m,th}^T$, I_{th}^T , ψ_0^T , k_I^T , m^T , and f^T – for 6F TCRs (figure 3.19A), and we set $N^T = 4$ for 6F TCRs. Moreover, 6F TCRs displayed more sensitivity to low-affinity TCRs but a lesser maximal response amplitude. To match this effect quantitatively, we defined correction factors $\alpha_A^{T,6F}$ and $\alpha_\tau^{T,6F}$ to apply to the amplitude A^T and threshold binding time τ_c^T of the Hill activation function in

Z^T , that is,

$$Z^{T,6F} = \alpha_A^{T,6F} \frac{(C_{NT}^{T,6F})^2}{(C_{NT}^{T,6F})^2 + (\Theta^{T,6F})^2}$$

with $\Theta^{T,6F} = C_{NT}^{T,6F} (\alpha_\tau^{T,6F} \tau_c^T, L^T \rightarrow \infty, L^C = 0)$ (D.10)

We estimated these correction factors as illustrated in figure 4.9A. We used the IL-2 response (geometric average across time points and CAR constructs over 72 hours) of 6Y or 6F CAR T cells responding to single TCR antigens in the absence of a CAR antigen, as a function of TCR antigen τ . The factor $\alpha_A^{T,6F}$ was set based on the ratio of responses of 6F and 6Y T cells to strong TCR agonists N4 and A2; the factor $\alpha_\tau^{T,6F}$ was set by fitting Hill functions on the response curves and taking the ratio of Hill thresholds of 6F and 6Y. These α factors were then applied in Z^T for model predictions of $FC_{\text{TCR}/\text{CAR}}$ in figure 4.3." (*TCR/CAR antagonism*, [3])

D.9.2 1-ITAM CARs

"For 1-ITAM CARs, we used the same real-valued rate parameters (φ^C , $C_{m,\text{th}}^C$, etc.) as for 3-ITAM CARs, but we set $N^{C,1} = m^{C,1} = f^{C,1} = 1$ and computed the threshold Θ^C with these integer parameters. We adjusted the amplitude of Z^C since 1-ITAM CARs consistently produced weaker responses than 3-ITAM CARs (observable in vivo and in vitro, across cytokines). We defined a correction factor $\alpha_A^{C,1\text{CAR}}$, analogous to equation $\alpha_A^{T,6F}$ in eq. (D.10), and fitted it to ensure that the ratio of 1-ITAM versus 3-ITAM CAR T cell IL-2 responses to CD19 only (geometric average of IL-2 concentration across time points, replicates, and TCR constructs) was equal to the ratio $Z^{C,1}/Z^{C,3}$ of CAR outputs in response to CD19 alone. This fit is illustrated in figure 4.9B.

Moreover, the mere presence of 3-ITAM CARs on T cells, even in absence of a CAR antigen, produced a somewhat lower TCR-mediated response than 1-ITAM CAR T cells. We accounted for this passive influence by defining a correction factor $\alpha_A^{T,1\text{CAR}}$ to apply to Z^T in 1-ITAM CAR T cells, as in eq. (D.10). We fitted this factor on the response of 1-ITAM versus 3-ITAM CAR T cells to TCR agonists N4 and A2 in absence of CD19 (geometric average of IL-2 concentration across time points, replicates, and TCR constructs). This fit is illustrated in figure 4.9C. These α_A factors were then applied to Z^T and Z^C for model predictions of $FC_{\text{TCR}/\text{CAR}}$ in figure 4.3." (*TCR/CAR antagonism*, [3])

D.9.3 Combination of both altered receptors

“For T cells with 6F TCRs and 1-ITAM CARs, we applied all the parameter adjustments described above. The Z^T amplitude corrections were multiplied with one another in Z^T : $A^T = \alpha_A^{T,6F} \alpha_A^{T,1CAR}$.” (*TCR/CAR antagonism*, [3])

D.10 Antagonism predictions from TCR antigen libraries

“We used the publicly available dataset of peptide EC_{50} s from Łuksza *et al.* [258]. It contains EC_{50} s of all single amino acid substitution variants of three different TCR antigens, for 1–3 different TCR clones each (CMV: TCRs C1–C3, gp100: TCRs G1–G3, neoantigen: TCR N1).” (*TCR/CAR antagonism*, [3], SI)

D.10.1 Correcting missing CMV-derived peptide EC_{50} s

“In the dataset, all CMV-derived peptides with 4-1BB⁺ response below $\sim 48\%$ at maximal (100 $\mu\text{g}/\text{ml}$) dose were not measured at lower doses, as indicated by the cumulative distribution functions shown in Inline Figure D.1, left panels (see also, in the original reference, Extended Data Fig. 5). Therefore, no direct EC_{50} s were available for these peptides. Since these weaker peptides are precisely in the expected range of antagonists, not including them would have severely underestimated the abundance of null and antagonist CMV-derived peptides in figure 4.13H–K. Since the EC_{50} for all other peptides (derived from gp100 and the neoantigen) were completely measured in the dataset, we built a linear regression of $\log EC_{50}$ as a function of 4-1BB % response at maximal dose (figure D.1, right panel). We then used this linear relationship to infer the EC_{50} s of the CMV peptides missing full dose response measurements, based on their maximum 4-1BB⁺ response (available even for these peptides).” (*TCR/CAR antagonism*, [3], SI)

D.10.2 Model predictions

“We converted all $\mu\text{g}/\text{ml}$ EC_{50} s to molar units, then converted to binding times τ with equation (D.4). For each TCR-peptide combination, we generated predictions of $FC_{\text{TCR}/\text{CAR}}$, modelling the peptide being pulsed at 1 μM on PC9 tumors presenting CD19 (assuming that these cells are endowed with the right HLA haplotype to present these peptides).” (*TCR/CAR antagonism*, [3], SI)

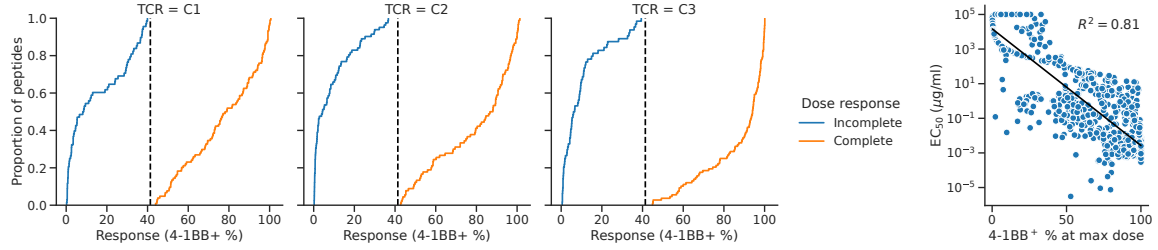


Figure D.1: “(Left) Cumulative distribution function of 4-1BB⁺ % response of each pair of TCR and CMV-derived peptide at maximal dose (100 μg/ml), split depending on the availability of a complete dose response curve. Other peptides (gp100, neoantigen) would only have a ‘Complete dose response’ line. (Right) Correlation between 4-1BB⁺ response at maximal dose and log EC₅₀, for all peptides with an EC₅₀ in the dataset ($R^2 = 0.81$). We performed a linear regression on this distribution to infer the missing EC₅₀s of CMV-derived peptides.” (TCR/CAR antagonism, [3], SI)