Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

Dipl.-Phys. Matthias Günther

born in Sömmerda

Oral examination: October 19, 2016

The role of co-receptor CD8

in ligand discrimination and T cell activation –

insights from

data-driven mathematical modeling

Table of Contents

Acknowledgments

The enterprise of writing the present thesis would have been impossible without people whose support and influence has been of great value. To begin with, I would like to thank my supervisor, Thomas Höfer, who gave me the opportunity to work in his group and provided me with deeper insights in mathematical modeling and advise I could always count on. I also thank the whole Höfer group for plenty discussions and the nice working atmosphere.

A special thank goes to Wolfgang Schamel, not only for being my second supervisor, member of the examination committee, and our collaboration partner, but also for his expertise he granted me on many occasions. Also Sumit Deswal is thanked for conducting the experiments that put the 'data-driven' in the title of this thesis.

I further want to express my gratitude to Ulrich Schwarz for accepting to act as second reviewer of my thesis as well as the chairperson of the examination committee. Beyond that, his inspiring lectures had vital influence on my decision to turn from pure physics towards the mathematical aspects of biology. I would also like to thank Ursula Kummer for kindly agreeing to be a member of the examination committee.

The greatest and surely most pleasant influence came and comes from Katha. Your witty enthusiasm opened my eyes not only for the beauty of biology. Without your deep love, kindness, and endless patience, I would have been lost long time ago. Thank you for being just the way you are.

Lastly, I want to thank my whole family, especially my parents, my brother, and my grandmother, for their unconditioned love, support, and trust.

Summary

The co-receptor CD8 plays an important part in the proper functioning of cytotoxic T lymphocytes. In order to sense stimuli, the T cell surface receptor (TCR) engages peptide-specifically with its ligand pMHC, while CD8 makes peptide-unspecific contact to the MHC subunit. In this work, the entirety of interactions between TCR, CD8, and pMHC are elucidated by confronting of a family of mathematical pMHC-TCR-CD8 interaction models with accurately measured dose response data. The interaction model being in best agreement with the data, termed CBM, consists of a TCR-CD8 complex having the striking property that its CD8 subunit exhibits increased affinity to pMHC compared to CD8 alone.

A T cell triggering model, founded on multivalent binding, is constructed that enables affinity-based ligand discrimination. In combination with CBM, the TCR triggering model is capable to correctly predict key aspects of dose response T cell activation data, and introduces a novel mechanism for the contribution of CD8 in ligand discrimination and T cell activation. The high affinity CD8 binding site of the TCR-CD8 complex prevents low affinity (self) ligands to establish pMHC-TCR contacts and thereby reducing the intracellular signal intensity in response to self peptides. High affinity (foreign) ligands, on the other hand, can counteract by forming pMHC-TCR contacts with TCR-CD8 complexes. Because CD8 also binds the kinase Lck, this leads to enhanced intracellular signal intensity in response to foreign antigen. Thus, CD8 amplifies affinity-based ligand discrimination and the proposed mechanism leads to improved self tolerance as well as sensitivity towards foreign antigen of T cells allocating CD8 a significant contribution to T cell immunity.

Zusammenfassung

Der Korezeptor CD8 trägt zur reibungslosen Funktionsweise von zytotoxischen T-Lymphozyten einen wichtigen Anteil bei. Zur Reizerkennung bindet der T-Zell-Oberflächenrezeptor TCR peptid-spezifisch an den pMHC Liganden, während das CD8 peptid-unspezifisch mit dem MHC Teil interagiert. Um die Gesamtheit der vorliegenden Wechselwirkungen zwischen TCR, CD8 und pMHC zu untersuchen, werden in dieser Arbeit mehrere pMHC-TCR-CD8 Interaktionsmodelle aufgestellt und mit genau bestimmten Dosiswirkungs-Messwerten konfrontiert. Das Interaktionsmodell, welches die größte Übereinstimmung mit den vorliegenden Daten aufweist, wird CBM genannt, und verfügt über einem TCR-CD8-Komplex mit der unerwarteten Eigenschaft einer, im Vergleich zu freiem CD8, erhöhten Affinität der CD8 Untereinheit für pMHC.

Auf multivalenter Bindung basierend, wird ein T-Zell Triggermodell aufgestellt, dass es ermöglicht, Liganden bezüglich ihrer Affinität zu unterscheiden. Durch Kombination der Triggermodells mit CBM können wesentliche Aspekte von Dosiswirkungsdaten der T-Zellaktivierung korrekt wiedergegeben werden. Des Weiteren wird dadurch ein neuartiger Mechanismus des Beitrags von CD8 zur Ligandenunterscheidung sowie zur T-Zellaktivierung eingeführt. Durch die CD8 Untereinheit des TCR-CD8 Komplexes, welche eine hohe Affinität aufweist, wird die Ausbildung von pMHC-TCR Bindungen bei Liganden mit niedriger Affinität (Selbst-Peptide) verhindert und dadurch die intrazelluläre Signalstärke verringert. Andererseits können Liganden mit hoher Affinität (Fremd-Peptide) dem entgegenwirken indem sie pMHC-TCR Bindungen mit dem TCR-CD8 Komplex ausbilden. Da CD8 außerdem die Kinase Lck bindet, ruft fremdes Antigen eine erhöhte intrazelluläre Signalstärke hervor. Somit verstärkt CD8 die affinitätsbezogene Ligandenunterscheidung und der vorgeschlagene Mechanismus führt zu einer Verbesserung sowohl der Selbsttoleranz als auch der

Sensitivität bezüglich Fremdantigenen von T-Zellen. Dadurch wird CD8 ein wesentlicher Beitrag zur T-Zellimmunität zugesprochen.

1 Introduction

1.1 Ligand recognition by T cells

Cytotoxic T cells (CTLs) play a central role in the adaptive immune system, as they constantly scan the surface of body cells for pathogen-derived antigen. Virtually all body cells express major histocompatibility complex of class I molecules (MHC) on their surface to which short peptide fragments (pMHC) of usually 8–10 amino acids are bound. Because the peptides are derived from the intracellular protein pool, such antigen presenting cells (APCs) are mirroring their internal protein state onto their cell surface allowing CTLs to target infected or otherwise damaged cells [52]. The key player on the surface of CTLs participating in the process of antigen recognition is the T cell receptor (TCR), which comprises of a ligand recognition unit and an intracellular signaling module [70]. The ligand recognition unit, which makes contact to pMHC, is a dimer, which in its most common form is composed of a highly variable α and β chain [52]. It possesses a single binding site for pMHC, and typical dissociation constants of TCR binding to agonist ligand were determined to range from $1\,\mu$M to $50\,\mu$M, with the half-life of the interaction on the order of seconds [18]. If a foreign antigen is detected by the TCR, this information is transduced via the signaling module. It comprises of three dimers: CD3$\gamma\epsilon$, CD3$\delta\epsilon$, and CD3$\zeta\zeta$ [70]. All three units carry conserved regions known as immunoreceptor tyrosine-based activation motifs (ITAMs) [64]. Upon appropriate contact between pMHC and TCR, ITAMs get phosphorylated by the Src family kinase Lck [57]. Phosphorylated ITAMs are the target of Syk family kinase ZAP70, which, if bound to an ITAM, gets in turn phosphorylated by Lck [10, 57]. It is only in this state in which the TCR complex possesses enzymatic activity and initiates the intracellular signaling cascade leading to T cell activation [70].

The functionality of the adaptive immune system critically depends on the reliable discrimination of self peptides from foreign antigens. If the presence of antigens, for which the TCR expressed on the surface of the involved CTL possesses a respective specificity, is encountered on the surface of a targeted APC, the information is transmitted to the interior of the T cell via the TCR signaling module. In response, a vast signaling machinery is switched on leading ultimately to the killing of the target APC as well as to the secretion of cytokines into the extracellular environment [52]. Importantly, T cells exhibit a striking self tolerance, i.e. no such actions are carried out if only self peptides are presented. However, even minute amounts of foreign antigens, embedded in a large pool of self peptides, are sufficient to trigger T cell activation [76, 38, 61] indicating that T cells have a remarkable sensitivity towards their cognate antigen. Furthermore, it has been shown that by altering a single amino acid of the peptide, the pMHC ligand can loose its ability to trigger the TCR or can become an antagonist [69]. Nevertheless, T cells are not restricted to a single antigen, as the same TCR can respond to unrelated peptides bound to the same MHC [24]. Thus, T cells are highly self tolerant, show poly-specificity, and a remarkable sensitivity towards their antigens.

The mechanism by which T cells process the information provided by the receptor-ligand interactions is still under debate. In the following, some of these mechanisms are briefly discussed. The kinetic proofreading [49] and discrimination [62] models propose that, upon pMHC-TCR contact, a series of modification steps have to be completed in order to trigger T cell activation. If pMHC dissociates from the TCR before, the molecules will convert to their unmodified form [49, 62]. Hence, the off-rate of the pMHC-TCR interaction plays the key role in these models. However, it has been argued that these models achieve only few false positive events, but pay for this by a considerable loss of sensitivity [11]. Improvements were made by extending kinetic proofreading with additional mechanisms.

Kinetic-segregation proposes that phosphatases exhibiting long extracellular tails, like CD45, are excluded from the sites of pMHC-TCR contact, thereby allowing kinases like Lck to operate more efficiently [19]. The observation that a single antagonist pMHC can engage and trigger up to 200 TCRs led to the development of the serial triggering model [80]. In combination with kinetic proofreading, it hypothesizes that pMHC-TCR interactions must be sufficiently long to induce proximal signaling but also sufficiently short to allow several TCRs to be stimulated, and thereby predicting an 'optimal dwell time' [81, 15]. A further extension includes a negative feedback mediated by the phosphatase SHP-1 [27]. Besides mechanism relying on the off-rate of pMHC-TCR interaction as most critical parameter of T cell antigen recognition, several models exist arguing in favor of affinity. In the conformational change model, it is assumed that ligand engagement with TCR induces a conformational change in the TCR/CD3 complex allowing for initiation of proximal signaling [31, 32]. A refinement of the conformational change model, termed permissive geometry model, has been proposed in that only multimeric pMHC binding induces a conformational change [50]. Furthermore, the aggregation of TCRs into 'microclusters' following ligand engagement has been observed [87, 82] and used to explain T cell activation [67]. This shows that a vast amount of, not necessarily mutually exclusive, models have been considered to explain T cell antigen recognition.

Under physiological conditions, CTLs interact with APCs, but, nevertheless, it was demonstrated that CTLs can also be stimulated by soluble pMHC [4]. At first sight, the conception that the affinity is the most critical parameter poses a serious problem on how discrimination can be achieved. To see why, consider a ligand L binding to receptor R with affinity K. Then, the fraction of ligand-bound receptors is given by $KL/(1 + KL)$. Thus, it is not the affinity alone but the product of affinity and ligand concentration that determines receptor occupancy. Hence, low affinity can be

compensated by high concentration, and ligand discrimination is compromised. For models relying on the off-rate, like kinetic proofreading, this aspect is of no concern, because even at saturated occupancy, the amount of triggered TCRs is low if the dwell time does not match [49]. On the other hand, it has been shown that soluble monomers are not sufficient to activate T cells, but stimulation with dimers or larger oligomers is required [6, 74]. Indeed, for T helper cells, up-regulation of CD69 and CD25 – both being surface marker for T cell activation – as well as down-regulation of CD3 can be accurately described by a multivalent binding model [73]. Multivalent ligand engagement offers the possibility of T cell antigen discrimination, and detailed discussions of multivalent ligands binding to a single receptor type have been published [59, 58]. In its simplest form, the ligand L is a dimer possessing of two identical binding sites interacting with a receptor R having a single binding site, as shown in Fig. 1A. One of the two ligand subunits binds to a receptor with rate k_{on} forming a monovalent complex C_1, which dissociates with rate k_{off}. The unbound ligand subunit of a complex C_1 can further bind a receptor with rate q_{on} forming a bivalent complex C_2. From this complex, one ligand subunit can dissociate with rate q_{off}. The corresponding set of equations is given by

$$\frac{\mathrm{d}C_1}{\mathrm{d}t} = 2\,k_{\mathrm{on}}L\,R - k_{\mathrm{off}}\,C_1 - \left[\,q_{\mathrm{on}}C_1\,R - 2\,q_{\mathrm{off}}\,C_2\,\right], \tag{1a}$$

$$\frac{\mathrm{d}C_2}{\mathrm{d}t} = q_{\mathrm{on}}C_1\,R - 2\,q_{\mathrm{off}}\,C_2\,, \tag{1b}$$

$$R_{\mathrm{tot}} = R + C_1 + 2\,C_2\,, \tag{1c}$$

with R_{tot} the total receptor density. Defining the dimensionless bivalent association constant by $q \equiv R_{\mathrm{tot}}\,q_{\mathrm{on}}/q_{\mathrm{off}}$, the steady state of the fraction of bivalently engaged receptors $y_2 \equiv C_2/R_{\mathrm{tot}}$ is given by

$$y_2(L) = 1 - \frac{(1+2\,KL)^2}{4\,q\,KL} \cdot \left[\sqrt{1 + \frac{8\,q\,KL}{(1+2\,KL)^2}} - 1\right]. \tag{2}$$
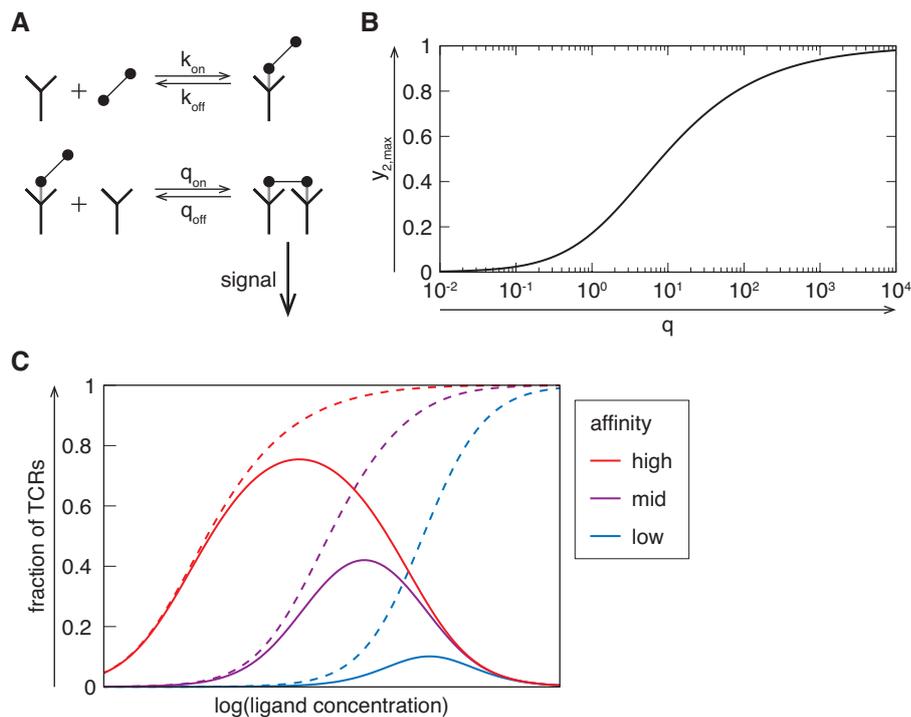
Figure 1: Bivalent ligand engagement

(A) The reaction scheme of bivalent ligand engagement. One of the two identical ligand subunits (black dots) binds to receptor (Y shape) with rate k_{on} to form a monovalent complex. This complex can either dissociate with rate k_{off}, or the unbound ligand subunit binds another receptor with rate q_{on} forming a bivalent complex. From this, one of the two bound ligand subunits can dissociate with rate q_{off}. Only those receptors bound in bivalent complexes are capable of initiating downstream signaling (indicated by the arrow). (B) The maximum fraction of bivalently bound TCRs, $y_{2,textmax}$, in dependence on the bivalent association constant q. (C) The fraction of ligand-bound (dashed curves) and bivalently bound (full curves) receptors in dependence on ligand concentration for high (red), mid (purple), and low (blue) affinity ligands. The bivalent association constant has been set proportional to the binding affinity. The fraction of ligand-bound receptors reaches saturation independent of ligand quality. Bivalent binding discriminates ligands according to affinity since low affinity ligands fail to establish a effectual amount of bivalently bound receptors.

Relation (2) describes a bell-shaped curve with a peak at the concentration value $L^* = 1/(2\,K)$. The bell-shape is a consequence of ligands in solution competing with monovalently bound ligands for unoccupied receptors. At low concentrations, many TCRs are unbound and monovalently bound ligands can engage with a second receptor. However, at high concentra-

tions, most TCRs are bound by a ligand and multivalent engagement is no longer possible. This results in a bell-shaped curve of multivalent binding [59]. The peak value in dependence on the bivalent association constant, $y_{2,\max}(q)$, is given by

$$y_{2,\max}(q) = 1 - \frac{2}{q}\left[\sqrt{1+q} - 1\right],\tag{3}$$

and is a strictly increasing function of q, i.e. $y'_{2,\max}(q) > 0$. Hence, the larger q the larger the peak value, see Fig. 1B. Assuming that only bivalently bound TCRs induce intracellular signaling (arrow in Fig. 1A), ligands with sufficiently large bivalent association constant are capable of activating T cells. This is shown in Fig. 1C for $q \sim K$ [75]. The larger the affinity the more bivalently bound TCRs are established, and low affinity ligands fail to elicit T cell activation. Therefore, ligand discrimination based on affinity can also be achieved by soluble pMHC.

1.2 Impact of co-receptor CD8

Ligand recognition and signal initiation are accompanied by the co-receptor CD8. CD8 is a transmembrane cell surface protein existing in two forms; as $\alpha\alpha$ homodimer and $\alpha\beta$ heterodimer [52], with the $\alpha\beta$ isoform being found on most CTLs [56]. CD8 is thought to contribute to T cell activation by enhancing sensitivity to pMHC [38, 36] and by stabilizing pMHC-TCR interactions [47, 86]. However, CD8 is not always indispensable as some ligands can trigger T cell activation in the absence of CD8 [16, 13, 46]. These effects are attributed to the participation of CD8 in intracellular signal initiation and binding of presented pMHC [45].

The main contribution of CD8 in enhancing sensitivity in T cell antigen discrimination comes from the association of CD8 with the kinase Lck [83, 79], and established the notion of CD8 acting as a T cell co-receptor [40]. This view is also supported by the finding that CD8 binds

MHC at a highly conserved region [66, 65], which is distal to the binding site of pMHC-TCR interactions [30]. This interaction was determined to be of relatively fast kinetics ($k_{\text{off}} \approx 20\,\text{s}^{-1}$) and low affinity ($K_D$ between $100\,\mu\text{M}$ and $220\,\mu\text{M}$ for human, and $15\,\mu\text{M}$ and $135\,\mu\text{M}$ for murine CD8) [14]. Importantly, it was suggested that a TCR and a CD8 co-receptor can bind the same pMHC [30]. Several studies also indicated that there is a physical interaction between the TCR and CD8 on the cell surface [77, 29, 5]. In particular, the α-chain connecting peptide motif of the TCR has been identified as agent recruiting CD8 to the TCR/CD3 complex [55, 48]. Hence, CD8 can deliver Lck to the site of pMHC-TCR contact and therby enhancing sensitivity as well as stabilizing the pMHC-TCR interaction.

1.3   Outline of the thesis

In Sec. 1.2, the contributions of co-receptor CD8 to T cell antigen recognition and activation were outlined. However, the models of antigen recognition described in Sec. 1.1 do not account for CD8 explicitly. It is therefore the aim of this work to identify the mechanistic interactions between pMHC, TCR, and CD8 and to relate them to T cell antigen recognition.

In Chap. 2, the statistical framework used in this work for model fitting, parameter estimation, and model selection is outlined. Based on the remarks made in this chapter, a family of pMHC-TCR-CD8 interaction models is developed in Chap. 3. These models are suited to account for cell surface receptor interactions with multivalent soluble ligands, which is exploited in Chap. 4 to select the binding model that fits accurately measured dose response binding data best. Furthermore, the model parameters are estimated and assigned with confidence intervals. The properties of the chosen model concerning receptor occupancy, multivalent binding and TCR triggering are outlined in Sec. 5. Additionally, in Chap. 6, the model is

placed in the context of antigen presenting cells to investigate whether the found attributes are preserved in more lifelike situations. Finally, the findings are discussed and placed in context in Chap. 7.

2 Statistical methods

A central task of this work is to find a binding model describing the interactions between pMHC, TCR, and CD8 and to estimate the corresponding model parameters. This is achieved by a statistically rigorous confrontation of experimental data with the corresponding observable calculated for each model of a set of candidate models. The determination of this observable is outlined in Chap. 3. In the following, it is described how the models are fitted to the data, parameter confidence intervals are estimated, and the best-fitting model is selected.

2.1 Maximum likelihood estimation

Consider a set of n independent data points, $(x_i)_{i=1,...,n}$, generated by random variables with probability density functions $f_i(x \,|\, \theta)$. For different i, the density functions can assume divers functional froms, but all density functions share a common set of parameters θ. Once all data x_i is gathered and therefore fixed, the likelihood function $\mathcal{L}$, defined as the joint density function of all observations, is regarded as a function of the parameters θ,

$$\mathcal{L}(\theta) := \prod_{i=1}^{n} f_i(x_i \,|\, \theta) \,. \tag{4}$$

The aim is to find an estimate for the true value θ_0 of the parameter vector θ based on the data. The principle of maximum likelihood estimation (MLE) delivers such an estimate by calculating the value θ_{MLE} that maximizes the likelihood function (4) [1, 26]. It is usually more convenient to consider the log-likelihood function l, which is defined as the natural logarithm of the likelihood function,

$$l(\theta) := \ln \mathcal{L}(\theta) = \sum_{i=1}^{n} \ln f_i(x_i \,|\, \theta) \,. \tag{5}$$

Since ln is a monotonically increasing function, maximizing (5) is equivalent to maximizing (4). Therefore, MLE is performed by

$$\theta_{\mathrm{MLE}} \equiv \arg\max_{\theta} l(\theta) \, , \tag{6a}$$

$$l_{\max} \equiv l(\theta_{\mathrm{MLE}}) \, . \tag{6b}$$

The point estimate θ_{MLE} itself does not contain any information about uncertainties of that estimate. Estimation of parameter uncertainties is achieved by using the method of profile likelihood. This procedure, as well as model selection, involves the value $l_{\max}$. But before these two aspects are further outlined, an important special case of MLE is discussed.

2.2 Least squares estimation

Assume all probability density functions f_i are given by normal distributions with unknown mean $y_i(\theta)$ but known standard deviation σ_i,

$$f_i(x \mid \theta) := \frac{1}{\sqrt{2\pi \, \sigma_i^2}} \exp\left[-\frac{1}{2}\left(\frac{y_i(\theta) - x}{\sigma_i}\right)^2\right] \, . \tag{7}$$

This means that, for each i, only the functions y_i depend on the parameters θ. In particular, $y_i(\theta)$ is the observable of condition i calculated from the model under consideration. The corresponding log-likelihood function (5) becomes

$$l(\theta) = -\frac{1}{2} \sum_{i=1}^{n} \left[\left(\frac{y_i(\theta) - x_i}{\sigma_i}\right)^2 + \ln\left(2\pi \, \sigma_i^2\right)\right] \, . \tag{8}$$

Since only parameter-dependent terms impact function maximization, each term of the form $\ln\left(2\pi \, \sigma_i^2\right)$ can be dropped, and maximum likelihood estimation becomes equivalent to least squares estimation, with the least squares function given by

$$\chi^2(\theta) := \sum_{i=1}^{n} \left(\frac{y_i(\theta) - x_i}{\sigma_i}\right)^2 \, . \tag{9}$$

Hence, procedure (6) now reads

$$\theta_{\mathrm{MLE}} \equiv \arg \min_{\theta} \chi^2(\theta) \,, \tag{10a}$$

$$\chi^2_{\min} \equiv \chi^2(\theta_{\mathrm{MLE}}) = -2 \, l_{\max} \,. \tag{10b}$$

In the framework of least squares estimation, the value $\chi^2_{\min}$ can be exploited to access the quality of the fit via the goodness-of-fit (gof) parameter,

$$\mathrm{gof} \equiv \frac{\chi^2_{\min}}{n - |\theta|} \,, \tag{11}$$

with n the total number of data points and $|\theta|$ the number of model parameters. Hence $n - |\theta|$ denotes the degrees of freedom of the fit. A value of gof close to 1 indicates a good agreement between model and data.

2.3 Profile likelihood

To assess the uncertainties of the model parameters in terms of confidence intervals, the method of profile likelihoods is used [53, 84]. This method explores the local neighborhood of the MLE value θ_{MLE}, and considers those points of the parameter space acceptable whose log-likelihood value deviates not too much from $l_{\max}$. The profile likelihood function of the j-th model parameter is defined by [43, 63]

$$\mathrm{PL}_j(p) := \min_{\theta \in \{\theta | \theta_j = p\}} -2 \, l(\theta) \,. \tag{12}$$

Thus, the log-likelihood function is optimized with the j-th parameter θ_j being fixed to the value p. For a sufficiently large number of data points, the confidence interval of the j-th parameter at confidence level α is given by [43, 63]

$$\mathrm{CI}_\alpha(\theta_j) := \{p \mid \mathrm{PL}_j(p) \le -2 \, l_{\max} + \Delta_\alpha\} \,, \tag{13}$$

with Δ_α the α-quantile of the chi-squared distribution with one degree of freedom. For a $95\,\%$ confidence interval it holds $\Delta_a = 3.84$. This procedure is repeated for all model parameters delivering the independent confidence intervals of the model parameters at confidence level α.

2.4 Prediction profile likelihood

The profile likelihood method allows for estimation of confidence intervals for model parameters. If the quantity of interest, to which a confidence interval has to be assigned, is of the more general form $W(\theta)$, a similar strategy can be applied, which goes under the name prediction profile likelihood [43, 44]. The prediction profile likelihood function of prediction $W(\theta)$ is defined by

$$\mathrm{PPL}(z) := \min_{\theta \in \{\theta | W(\theta) = z\}} - 2\,l(\theta)\,, \tag{14}$$

Thus, optimization is only performed on the subset of parameters obeying $W(\theta) = z$. In analogy to (13), the prediction confidence interval at confidence level α is given by [43, 44]

$$\mathrm{PCI}_\alpha(W) := \{z \mid \mathrm{PPL}(z) \leq -2\,l_{\max} + \Delta_\alpha\}\,. \tag{15}$$

2.5 Model selection

If several different models are considered for describing the same data set, a strategy is required that allows for choosing the model in best agreement with the data. Such a strategy of model selection is provided by the corrected Akaike Information Criterion ($\mathrm{AIC_c}$) [8]. It is defined by

$$\mathrm{AIC_c} \equiv -2\,l_{\max} + 2\,|\theta| + \frac{|\theta|\,(|\theta| + 1)}{n - |\theta| - 1}\,, \tag{16}$$

and balances between the quality of the fit, represented by $-2\,l_{\max}$, and the size of the model in terms of the number of model parameters $|\theta|$. However,

a model's AIC_c value itself has no meaning. Instead, the AIC_c value of the model having the smallest value $\text{AIC}_\text{c,min}$ serves as reference. The difference of a model's AIC_c value to that reference, i.e.

$$\Delta\text{AIC}_\text{c} = \text{AIC}_\text{c} - \text{AIC}_\text{c,min} \,, \tag{17}$$

is the central quantity to perform model selection. Following the recommendation in Burnham and Anderson [8], a model is rejected if $\Delta\text{AIC}_\text{c} > 10$ holds. Thus, model selection via $\Delta\text{AIC}_\text{c}$ is a relative procedure in that always a best model is selected. No statement is made about the actual quality the selected models perform on the data.

3 A family of pMHC-TCR-CD8 interaction models

In Sec. 1.2, several findings concerning the contribution of CD8 to ligand binding were discussed. Based on that, a family of pMHC-TCR-CD8 binding models is developed in this chapter. To assure an unbiased view on the matter, it is started with a simple binding scheme whose complexity is increased stepwise by incorporation of further CD8-dependent reactions. All models are considered at thermodynamic equilibrium, as data acquisition – described in Chap. 4 – was performed under such conditions. At first, the models are introduced for monomeric pMHC clarifying the principal reactions to which the receptors and ligands are subject to. Next, the models are generalized to account for multimeric pMHC and rescaling of model parameters is performed. To obtain a unified treatment, the models are placed in the framework of binding polynomials [21], and conservations of receptor numbers is taken into account. Finally, the amount of bound ligands is determined, which serves as observable for model selection and parameter estimation in Chap. 4.

3.1 Binding models for monomeric pMHC

At first, the binding model family is set up for monomeric pMHC serving as ligands. The reaction schemes of the binding models are depicted in Fig. 2. The first model comprises only a single reaction; TCR binds reversibly to pMHC with affinity K_1 (Model A, Fig. 2A). This model is suited for T cells lacking the expression of CD8.

Incorporation of CD8 into ligand binding is initially done implicitly. The co-receptor does not appear as a binding partner, but instead affects the affinity of pMHC-TCR binding via a scaling factor ε (Model B, Fig. 2B). Hence, ligand binding is described by apparent binding affinities.

15
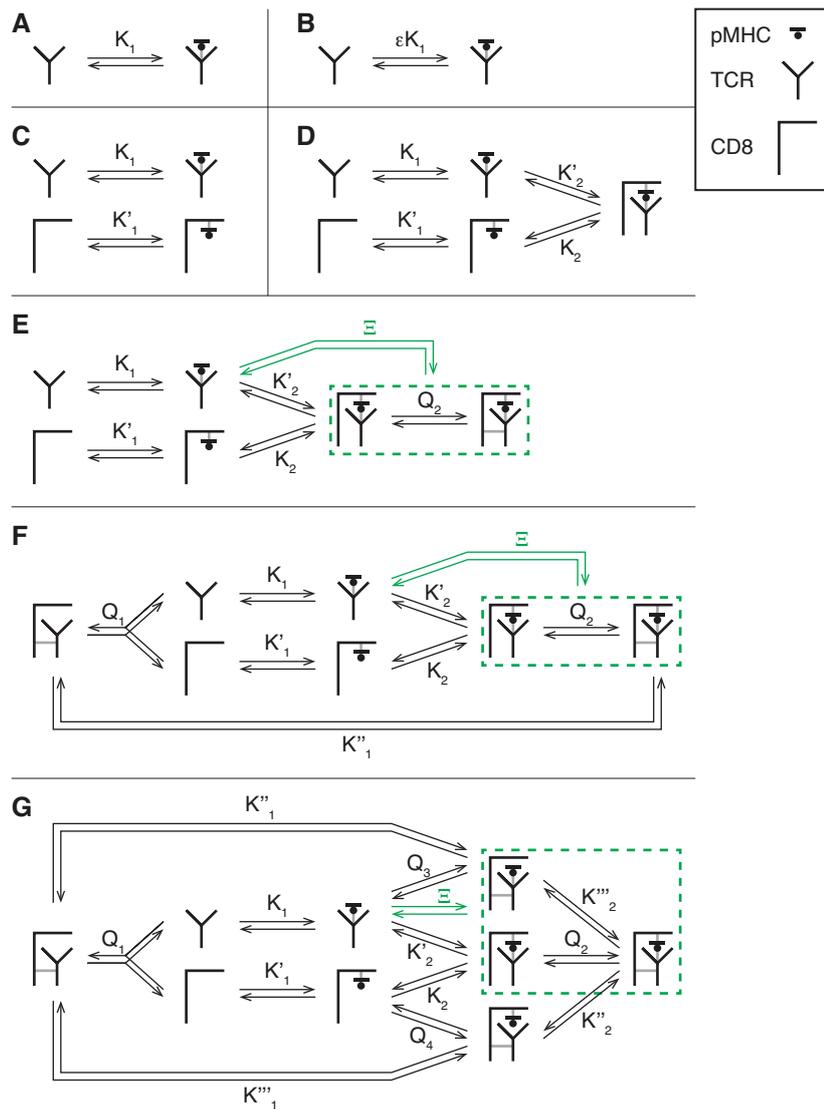
Figure 2: Monomeric binding models

The models of pMHC and TCR interaction (A) as well as pMHC, TCR, and CD8 interactions (B–G) are shown. The details are outlined in the main text. (A) pMHC binds to TCR from solution. (B) pMHC binding to TCR is enhanced by CD8. (C) pMHC can either bind to a TCR or to a CD8 co-receptor. (D) Additionally, TCR and CD8 can bind to the same pMHC. (E) If a TCR and a CD8 are in simultaneous contact with a pMHC, the two receptors can bind to each other. (F) Furthermore, TCR-CD8 complex formation is possible in the absence of pMHC. The TCR-CD8 complex possesses a single binding site. (G) A TCR-CD8 complex with two binding sites, one stemming from the TCR and the other one from CD8, can form. Next to binding to TCR or to CD8, pMHC can bind to the TCR or to the CD8 binding site of a TCR-CD8 complex. TCR and CD8 can further bind pMHC simultaneously, and all three molecules can make complete contact. The green, dashed frames in (E–G) indicate that these states are indistinguishable under the experimental conditions, and merged into a single aggregated state.

The first model accounting for CD8 explicitly comprises of two basic reactions. A pMHC can either bind reversibly to a TCR with affinity K_1 or to a CD8 co-receptor with affinity K_1' (Model C, Fig. 2C). Importantly, K_1' is not affected by altering the peptide of pMHC since CD8 makes only contact to MHC.

Next, the possibility of both receptors binding to the same pMHC is accounted for (Model D, Fig. 2D). In addition to pMHC-TCR and pMHC-CD8 binding, a TCR-bound pMHC can engage with a CD8 with association constant K_2' to form a trimolecular complex. Alternatively, this complex can also be established if a TCR binds to CD8-bound pMHC with association constant K_2. Because K_2' describes a reaction in which CD8 binds pMHC, it is, like K_1', independent of the peptide. Note that detailed balance imposes dependences between the parameters; these are discussed below.

The previous model is extended to include a direct TCR-CD8 interaction (Model E, Fig. 2E). If pMHC is bound simultaneously to TCR and CD8, the two receptors can bind to each other with association constant Q_2. Importantly, the two trimolecular binding states cannot be distinguished by the binding experiments outlined in Sec. 4.1. The reason is that both these states are build up by the same receptors and ligands, and, furthermore, pMHC makes contact to the TCR. These are the principal aspects by which the binding experiments can separate different binding states, because only dose and peptide quality were altered (see Sec. 4.1). All further binding characteristics, like a bond between TCR and CD8, are therefore not resolved. Hence, these states are merged into a single aggregate state, which is indicated by the green, dashed frame. Due to this state aggregation, Model E becomes indistinguishable from Model D.

The next model extension comprises of a ligand-independent interaction between TCR and CD8 (Model F, Fig. 2F). The two receptors can

form a complex with association constant Q_1. Despite both receptors alone can bind pMHC, the TCR-CD8 complex possesses only a single binding site to which pMHC can bind with affinity K_1''. As before, this model contains indistinguishable binding states indicated by the green, dashed frame. The corresponding state aggregation follows the same arguments as above.

The final model also allows for a pMHC-independent formation of a TCR-CD8 complex with association constant Q_1 (Model G, 2G). Unlike the two models before, this complex possesses two binding sites; one originating from the TCR and the other from CD8 making the last one peptide unspecific. Thus, additionally to the reactions described for Model D, pMHC can either bind to the TCR binding site of the TCR-CD8 complex with affinity K_1'' or to its CD8 binding site with affinity K_1'''. Importantly, these binding affinities must not be equal to the corresponding binding affinities of pMHC to TCR, K_1, respectively CD8, K_1'. The binding state with pMHC bound to the TCR binding site of a TCR-CD8 complex can also be formed if CD8 binds to pMHC-bound TCR with association constant Q_3. Similarly, a TCR binding to pMHC-bound CD8 with association constant Q_4 yields the state with pMHC bound to the CD8 binding site of a TCR-CD8 complex. Furthermore, a binding state in which pMHC, TCR, and CD8 are in complete contact with each other can form in three different ways. Either, pMHC makes contact to CD8 with association constant K_2''' if pMHC is already bound to the TCR binding site of a TCR-CD8 complex, or pMHC makes contact to the TCR with association constant K_2'' if pMHC is already bound to the CD8 binding site of a TCR-CD8 complex, or TCR and CD8 bind to each other with association constant Q_2 if both receptors are simultaneously bound to pMHC. Like K_1', K_1''', and K_2', the parameter K_2''' is independent of peptide quality. The model is further considered under the condition $Q_3 \geq Q_1$. The rationale behind this assumption is that, following the arguments given in Sec. 1.2, pMHC binding to TCR should not attenuate CD8 approaching the TCR. As Models E and F, Model G

comprises of a set of states that cannot be distinguished by the binding experiments. As before, these states are marked by a surrounding green, dashed frame and merged into a single aggregate state.

At equilibrium, the number of independent parameters needed for a complete description of the system equals the number of binding states. The remaining parameters are determined by principle of detailed balance. In case of the Model G (Fig. 2G), there are seven binding states (with state aggregation not yet considered) and twelve model parameters. Hence, detailed balance delivers five conditions,

$$K_1 \, K_2' = K_1' \, K_2 \,, \tag{18a}$$

$$K_1 \, Q_3 = K_1'' \, Q_1 \,, \tag{18b}$$

$$K_1' \, Q_4 = K_1''' \, Q_1 \,, \tag{18c}$$

$$K_2' \, Q_2 = K_2''' \, Q_3 \,, \tag{18d}$$

$$K_2 \, Q_2 = K_2'' \, Q_4 \,. \tag{18e}$$

The set of independent parameters describing Model G is chosen to be $\{K_1, K_1', K_2', Q_1, Q_2, Q_3, Q_4\}$ in which only K_1 is peptide specific. The three states in Fig. 2G that are surrounded by the green, dashed frame cannot be distinguished by the experiments outlined in Sec. 4.1, and are thus merged into a single state, which henceforth is termed aggregated state. Therefore, the number of independent parameters reduces by three as well. From the set of independent parameters, it follows that the corresponding effective association constant Ξ (see Fig. 2) describing the aggregated state is given by

$$\Xi \equiv Q_3 + K_2' \, (1 + Q_2) \,, \tag{19}$$

Table 1: Parameter occurrence in monomeric binding models

The independent parameters used to describe the different binding models are shown. The two most left columns comprise the parameter and a brief description thereof. The remaining columns indicate the occurrence of the parameter by a cross. The model names refer to the subfigures of Fig. 2.

Parameter	Comment	Model						
		A	B	C	D	E	F	G
K_1	Binding affinity of pMHC to TCR	×	×	×	×	×	×	×
ε	CD8 related enhancement of pMHC-TCR affinity		×					
K'_1	Binding affinity of pMHC to CD8			×	×	×	×	×
Q_1	TCR-CD8 complex formation						×	×
K'_2 / Ξ	CD8 binding to pMHC-TCR (aggregated state)				×	×	×	×
Q_4	TCR binding to CD8 of pMHC-CD8							×

and the set of independent parameters describing monomeric Model G is given by

$$\Theta_{\mathrm{G,mono}} \equiv \{K_1, K'_1, Q_1, \Xi, Q_4\}. \tag{20}$$

Similar arguments are applied to the other binding models. The independent parameters of Model F are obtained by removing Q_4 from (20) and by setting Q_3 to zero in (19). Models D and E require additional elimination of Q_1 from (20) and, in case of Model D, setting of Q_2 to zero in (19). Further, the parametrization of Model C is attained by additional removal of Ξ from (20). Parametrization of Model B is simply given by K_1 and ε. This is summarized in Tab. 1. The independent parameters occurring in the different models are marked by a cross.

3.2   Incorporation of ligand multivalencies

The next step is to extend the monomeric interaction models to include binding of multimeric ligands. The basic idea of incorporation of ligand multivalencies was already outlined in Sec. 1.1. There, the association constant of multivalent ligand binding was set proportional to the binding affinity. This concept is pushed further in the following. The ligand under consideration is an oligomer having ν identical binding sites or subunits.

The effective ligand valency f is defined to be the maximal number of subunits that can make contact to cell surface receptors simultaneously. At first, incorporation of ligand multivalencies is outlined for a single receptor type as it is the case for Models A and B (Fig. 2A and B). Then, the method is generalized to account for a second receptor type as in Models C–G (Fig. 2C–G). Note that, in the following, no attention is paid to the statistical weights of the binding states. These considerations are postponed to Sec. 3.4.

The general concept of modeling multivalent binding relies on the assumption of a common multivalent engagement parameter K_{MV} [75]. This involves two implications. Consider a ligand X, which binds from solution to a given receptor type with affinity $K_{1,X}$. If already k of the ligand's subunits made contact to k receptors, the first implication states that the association constant of binding to a further receptor is given by $K_{\mathrm{MV}} K_{1,X}$ for each $k = 1, \dots, f - 1$. If, however, f subunits are bound, no further subunit can make contact to another receptor. This means that after a ligand X binds from solution with affinity $K_{1,X}$, each multivalent binding reaction to a further receptor is determined by a single association constant $K_{\mathrm{MV}} K_{1,X}$ until the maximal number of f subunits made contact. The second implication concerns the comparison to another ligand type Y having affinity $K_{1,Y}$. It states that the association constant for ligand Y is given by $K_{\mathrm{MV}} K_{1,Y}$, and beyond that, the first implication holds. Hence, multivalent binding is described for all ligands by the same two parameters: the multivalent engagement parameter K_{MV} and the effective valency f. This is exemplified in Fig. 3A for a tetramer ($\nu = 4$) with effective valency $f = 3$. After binding from solution with affinity K_1, the engagement reactions from monomer to dimer as well as from dimer to trimer are both specified by the same association constant $K_{\mathrm{MV}} K_1$, which in turn is determined by the multivalent engagement parameter. Further binding from trimer to tetramer is prohibited due to the restriction of the effective

ligand valency. In order to describe the same reactions for another ligand, only the affinity needs to be adapted.

Fig. 3B shows all possible bivalent binding states grouped according to the different binding models. The only bivalent binding state of Models A and B, two pMHC subunits bound to TCR, is shown in the first column. Model C comprises of two further bivalent states with one pMHC subunit bound to CD8 (second column); the other subunit is either bound to TCR (first row) or to CD8 (second row). Three additional bivalent states appear in Models D, E, and F with one pMHC subunit bound via aggregated state (third column); the other subunit is either bound to TCR (first row), to CD8 (second row), or via aggregated state (third row). Lastly, Model G contains four more bivalent states with one pMHC subunit bound to the CD8 binding site of a TCR-CD8 complex (fourth column); the other one is either bound to TCR (first row), to CD8 (second row), via aggregated state (third row), or also to the CD8 binding site of a TCR-CD8 complex (fourth row). It is assumed that the same approach used for a single receptor type yields a reasonable outcome if a second receptor type is taken into account. This means that only the receptor, to which a ligand's subunit binds multivalently to, determines the association constant, and not how the ligand was previously bound. This is indicated by the association constants shown below each column and to the right of each row. The allocation is as following. If, for one of the depicted bivalent binding states, the ligand was initially bound via the left subunit, binding to the receptor type to the right is determined by the association constant at the right side of the row in which the binding state is placed. Note that, in the first row, the receptor on the right is always a TCR, in the second row always CD8, in the third row always the aggregated state, and in the fourth row TCR-CD8 complex with pMHC bound to its CD8 binding site. If the ligand was initially bound via the right subunit, the association constant below the column in which the binding state is placed determines multivalent binding with the
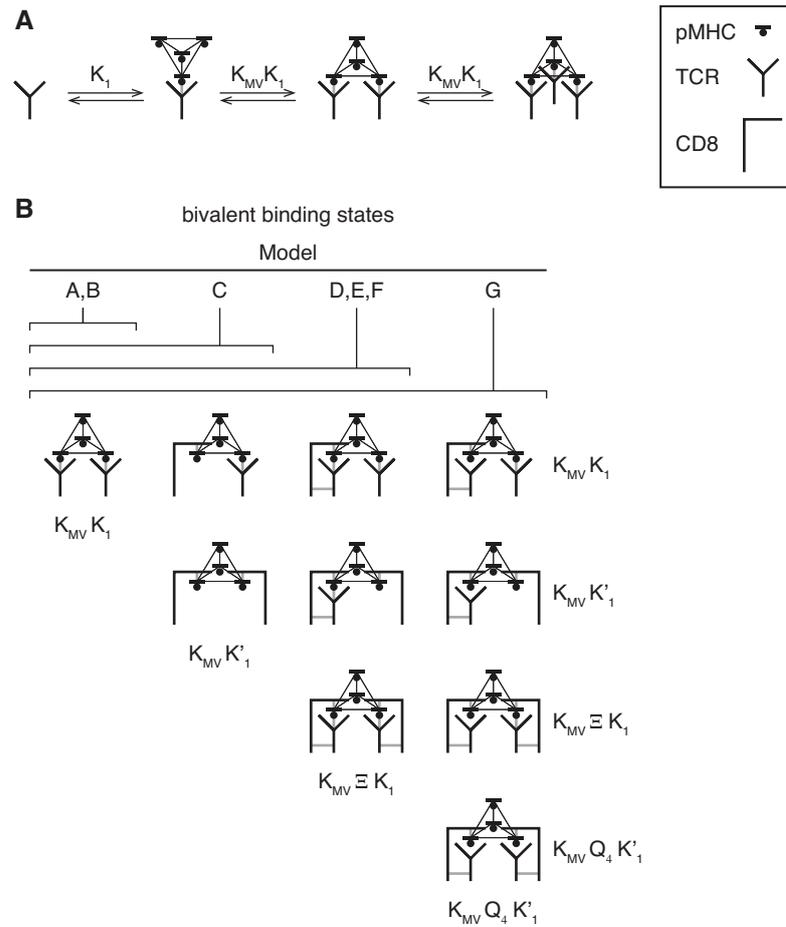
Figure 3: Multivalent binding states

(A) Tetramer with effective valency $f = 3$ binding to TCRs. The tetramer binds from solution with affinity K_1 and subsequent formation of dimers and trimers are determined by a single association constant $K_{MV} K_1$. Larger states than trimers cannot. (B) Bivalent binding states established by tetramers grouped according to the different binding models. First column: Models A and B comprise only of two cross-linked TCRs. Second column: Model C consists on top of states in which CD8 is cross-linked to TCR, or to CD8. Third column: Models D, E, and F also include states in which an aggregated state is cross-linked to TCR, or to CD8, or to an aggregate state. Fourth column: Model G contains further states in which pMHC bound to CD8 binding site of TCR-CD8 complex is cross-linked to TCR, or to CD8, or to an aggregated state, or to pMHC bound to CD8 binding site of TCR-CD8 complex. If the bivalent binding state is formed by binding to the receptor type appearing right in the binding state, the association constant on the right of the row in which the binding state is located is applied. If the bivalent binding state is formed by binding to the receptor type appearing left in the binding state, the association constant below the column in which the binding state is located is applied. In case of Model B, the association constant has further be corrected by the scaling factor ε.

receptor type to the left. The arrangement of the left receptors per column is similar to those for rows. For example, consider the binding state in the first row and second column. It shows a tetramer with the left subunit bound to a CD8 and the right to a TCR. If the tetramer was initially bound to CD8, binding of the right subunit to TCR is determined by the association constant $K_{\mathrm{MV}} K_1$ (shown to the right of the first row), because K_1 denotes the binding affinity of pMHC to TCR. Accordingly, if the tetramer was initially bound to TCR, further binding to CD8 is determined by the association constant $K_{\mathrm{MV}} K_1'$ (shown below the second column). Of course, the occupancy of a given binding state does not depend on the order by which it was formed. Making contact to TCR in the second binding step requires binding to CD8 in the first step with affinity K_1'. Vice versa, binding to CD8 in the second step necessitates contact of the ligand to TCR with affinity K_1. Hence, the occupancy of the binding state with one ligand subunit being in contact with CD8 while another one is in contact with a TCR is determined by $K_{\mathrm{MV}} K_1' K_1$ irrespective of the order of formation. These arguments can be applied to all multivalent binding states including those with larger valency than shown in Fig. 3B, i.e. trivalent or tetravalent binding states. As above, the maximal number of pMHC subunits that can make simultaneous contact to cell surface receptors is restricted by the effective ligand valency f. The outlined methods allows for a treatment of multivalent binding reactions for a ligand with given valency ν by only two parameters, the multivalent engagement parameter K_{MV} and the effective ligand valency f.

The assumption that all multivalent binding reactions can be described by only two parameters is of course very strong, and its validity needs to be shown. The findings described in Sec. 5.3 indicate, however, that the concept of a common multivalent engagement parameter seems to be a reasonable choice. The great advantage of this approach is that a vast number multivalent binding states can be described by a small number of

parameters; this is very desirable in order to avoid overfitting [8].

3.3 Dimensionalities and rescaling of model parameters

The complete set of model parameters comprises the parameters of the monovalent binding models and the two parameters used to describe multivalent binding,

$$\Theta \equiv \{K_1,\, K_1',\, \varepsilon,\, Q_1,\, \Xi,\, Q_4,\, K_{\mathrm{MV}},\, f\}\,. \tag{21}$$

The binding affinities K_1 and K_1' have the unit of an inverse concentration. For any calculations to be performed, the concentration unit is chosen to be µM. Thus the unit of the binding affinities is $\mu\mathrm{M}^{-1}$. The effective ligand valency f and the enhancement factor ε are already dimensionless. The association constants Q_1, Ξ and Q_4 have the unit of an inverse two-dimensional density[1], and the unit of the multivalent binding parameter K_{MV} is given by concentration over two-dimensional density. These four parameters are rescaled with a two-dimensional reference density R_T, which is determined in Sec. 4.1,

$$q_1 \equiv R_T\,Q_1\,, \tag{22a}$$

$$\xi \equiv R_T\,\Xi\,, \tag{22b}$$

$$q_4 \equiv R_T\,Q_4\,, \tag{22c}$$

$$\kappa_{\mathrm{MV}} \equiv R_T\,K_{\mathrm{MV}}\,. \tag{22d}$$

Hence, q_1, ξ, and q_4 are dimensionless, and κ_{MV} has the unit of a concentration, i.e. µM. The full set of model parameters, from which each binding model recruits its parameters (see Tab. 1), is thus given by

$$\theta \equiv \{K_1,\, K_1',\, \varepsilon,\, q_1,\, \xi,\, q_4,\, \kappa_{\mathrm{MV}},\, f\}\,. \tag{23}$$

---

[1]Two-dimensional densities have a unit of number over unit area.

3.4 Formulation via binding polynomials

The binding models are equipped with a proper parameterization and the ability to account for ligand multivalencies. Lacking is a useful framework in which the models are set up for model selection and parameter estimation. Such a framework is provided by the method of binding polynomials. Binding polynomials allow for a unified treatment of equilibrium binding as they represent the corresponding thermodynamic partition function [21]. Having a suitable representation of the model states at hand, the sum over all states is formed yielding the binding polynomial of the model. Based on the binding polynomial, receptor number conservation is taken into account to determine the densities of unbound receptors, which allows for calculating the amount of bound ligands. This quantity is the observable mentioned in Sec. 2.2, and is thus required in Chap. 4 to perform model selection and parameter estimation.

For a given ligand type with effective valency, all binding models developed above can be viewed as special cases of Model G. Therefore, only the binding polynomial of Model G is derived in the following. The corresponding binding polynomials of the other models are obtained by setting those parameters to zero that are not marked by a cross in Tab. 1. In case of Model B, the additional replacement $K_1 \to \varepsilon K_1$ has to be applied as well. Model G comprises of three states with no bound ligand: TCR (T), CD8 (C) and the TCR-CD8 complex forming with association constant q_1 ($q_1 \, T \, C$). Note that due to the rescaling (22) by the reference density R_T, the receptor densities of TCR, T, and CD8, C, became also dimensionless. The functional form of the density of each ligand-bound state can be written to only depend on the density of unbound TCR T, the density of unbound CD8 C, the ligand concentration L, and a set of parameters describing the way these constituents interact. The parameters are recruited from the set (23). Each ligand-bound state is characterized

by a set of four numbers, $u = \{u_1, u_2, u_3, u_4\}$, describing how many pMHC subunits are bound in which way to the cell surface receptors. If a ligand is bound according to binding state u, it has u_1 pMHC subunits bound only to TCR, u_2 pMHC subunits only to CD8, u_3 pMHC subunits via aggregated state, and u_4 pMHC subunits to the CD8 binding site of a TCR-CD8 complex. This is illustrated in Fig. 4 for the case of a tetramer ($\nu = 4$) with effective ligand valency $f = 3$. The 4 monovalent ligand-bound states were introduced in Sec. 3.1 aside from the fact that the ligand is now a tetramer. The 10 bivalent ligand-bound states were the subject of Sec. 3.2 where ligand multivalencies were incorporated into the binding models. Allowing for trivalent ligand-bound states leads to further 20 states whose occupancies are similarly calculated as those of bivalent ligand-bound states. In the further course, it is convenient to introduce to following scaffold expressions for the four different ways a pMHC subunit can be bound to the cell surface receptors,

$$S_1(T, C \mid \theta) := K_1\, T\,, \tag{24a}$$

$$S_2(T, C \mid \theta) := K_1'\, C\,, \tag{24b}$$

$$S_3(T, C \mid \theta) := \xi\, K_1\, T\, C\,, \tag{24c}$$

$$S_4(T, C \mid \theta) := q_4\, K_1'\, T\, C\,. \tag{24d}$$

These expressions contain the receptor densities and the principal parameter dependencies of the four monovalent binding states of Model G (Fig. 2G), and have the unit of an inverse concentration. To obtain the occupancy of a monovalent ligand-bound state, multiplication of the ligand concentration L with the respective scaffold expression (24) is required[2].

---

[2] An additional term accounting for the statistical weight has to be considered as well. This is postponed to a later part of this section.

monovalent binding states
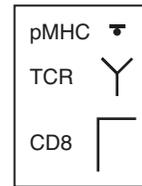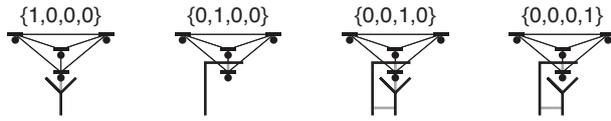
bivalent binding states
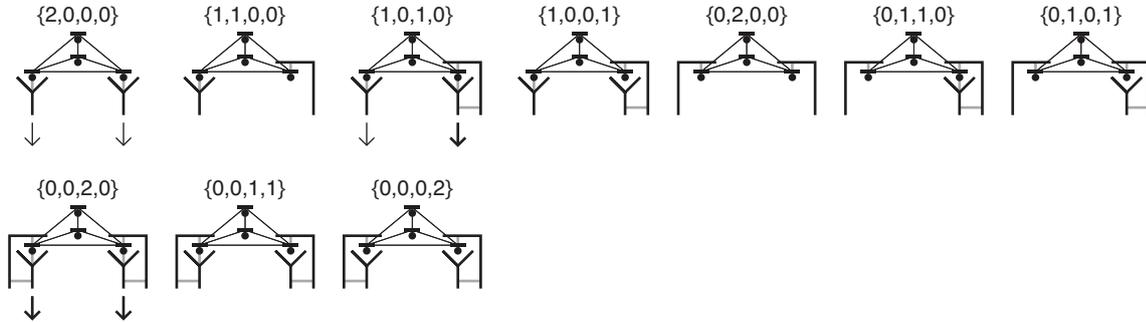
trivalent binding states
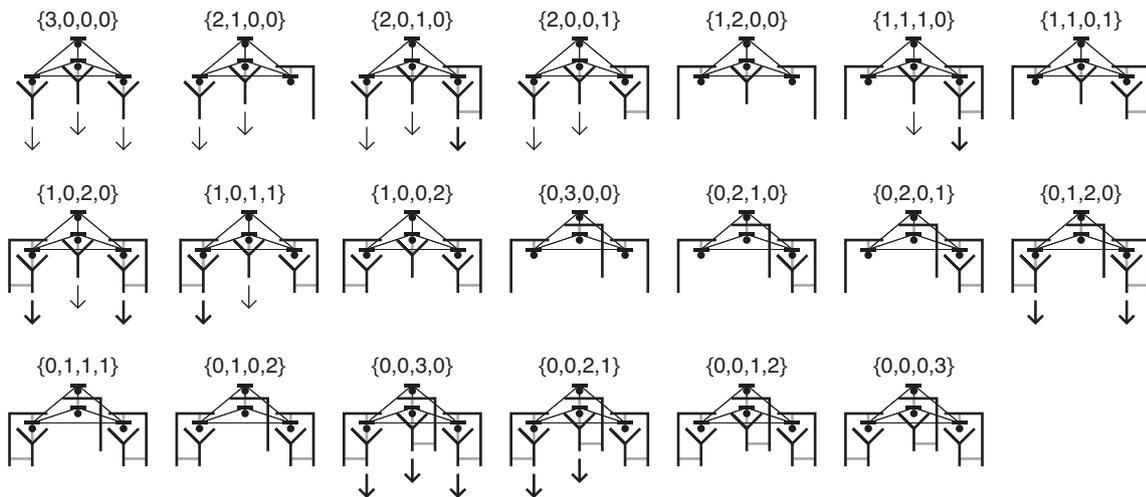
Figure 4: Ligand-bound states of Model G for tetramer with effective valency of 3

Each ligand bound state is characterized by a set $\{u_1, u_2, u_3, u_4\}$ indicating the four different ways a pMHC subunit can be bound to T cell surface receptors. Element u_j represents the number of pMHC subunits bound according to the j-th interaction type ($j = 1, \ldots, 4$). The valency of a state is given by $m = \sum_{j=1}^{4} u_j$. The four monovalent binding states ($m = 1$) show the basic interactions between pMHC, TCR, and CD8. Each pMHC subunit of a ligand can realize the same interactions as in the monovalent case leading to 10 bivalent and 20 trivalent binding states. The corresponding parameterizations are outlined in the main text. The arrows depict below some TCRs indicate the signaling capacity of these TCRs. Thin arrows represent weak signaling TCRs lacking adjacent CD8, while thick arrows represent strong signaling TCRs with increased signaling capacity due to the recruitment of kinase Lck by CD8.

Subsequent multivalent binding is described by multiplying the multivalent binding parameter κ_{MV} with the scaffold expression corresponding to the new contact the pMHC subunit makes[3]. The total number of bound pMHC subunits is $m = \sum_{j=1}^{4} u_j$, which is constrained by the effective ligand valency, $1 \leq m \leq f$. The first pMHC subunit binds from solution at concentration L, and the remaining $m - 1$ pMHC subunits engage multivalently. Thus, the (dimensionless) density of state u, denoted by X_u, is given by

$$X_u(T, C, L \mid \theta) := g_u \cdot L \cdot \kappa_{\mathrm{MV}}^{m-1} \prod_{j=1}^{4} S_j(T, C \mid \theta)^{u_j}, \tag{25}$$

with g_u denoting the statistical weight of state u. It comprises of two terms. The first term counts the number of ways to choose m pMHC subunits for binding with cell surface receptors out of the ν total pMHC subunits, and is therefore given by a binomial coefficient. The second term tallies the number of ways to distribute the four possible bond types among the m pMHC subunits, which is determined by a multinomial coefficient. Thus it holds

$$g_u \equiv \binom{\nu}{m} \cdot \binom{m}{u_1, u_2, u_3, u_4}. \tag{26}$$

The binding polynomial F is obtained by summing over all possible model states,

$$F(T, C, L \mid \theta) := T + C + q_1 \, T \, C + \sum_u X_u(T, C, L \mid \theta). \tag{27}$$

The index set u runs over all non-negative integer values obeying $m = \sum_{j=1}^{4} u_j$ with $1 \leq m \leq f$. Thus one finds

$$F(T, C, L \mid \theta) = T + C + q_1 \, T \, C +$$

$$L \sum_{m=1}^{f} \binom{\nu}{m} \kappa_{\mathrm{MV}}^{m-1} \sum_{\substack{u_1, u_2, u_3, u_4 \\ u_1+u_2+u_3+u_4=m}} \binom{m}{u_1, u_2, u_3, u_4} \prod_{j=1}^{4} S_j(T, C \mid \theta)^{u_j}. \tag{28}$$

---

[3]see footnote 2.

The binding polynomial can be simplified to

$$F(T, C, L \mid \theta) = T + C + q_1\, T\, C +$$

$$L \sum_{m=1}^{f} \binom{\nu}{m} \kappa_{\mathrm{MV}}^{m-1} \left[\sum_{j=1}^{4} S_j(T, C \mid \theta) \right]^{m}, \qquad (29)$$

by exploiting the multinomial theorem.

The experiments described in Sec. 4.1 were performed under conditions that leave the total number of cell surface receptors unchanged during the course of the experiment as well as for different experimental conditions. These conservations of total TCR and total CD8 receptor densities are accounted for by the following relations,

$$\rho_T = T\, \frac{\partial F(T, C, L \mid \theta)}{\partial T}\,, \qquad (30\mathrm{a})$$

$$\rho_C = C\, \frac{\partial F(T, C, L \mid \theta)}{\partial C}\,, \qquad (30\mathrm{b})$$

with ρ_T and ρ_C the dimensionless total TCR and CD8 receptor densities, respectively. The rescaling, resulting in dimensionless receptor densities, is performed by the same reference density used to rescale the model parameters in Sec. 3.3. Next to conservation of total receptor densities, the total number of ligands is conserved as well. However, as the experiments were performed under conditions for which the total number of ligands is in large excess over the total number of available receptors, the number of ligands in solution, and hence the ligand concentration, is not depleted irrespective of how many ligands are bound to T cell surface receptors. The justification of Eqs. (30) is as follows. The binding polynomial tallies the densities of all model states X_u, and multiplying each state with the corresponding number of TCRs respectively CD8 that are within these states results in the total density of TCRs and CD8, respectively. And this is precisely what expressions (30) are doing. The numbers of TCR and CD8

within a model state are given by the exponents of the corresponding receptors in the state density X_u. The partial derivatives with respect to the corresponding receptor brings this exponent in front of the state densities, and multiplication with the receptor density restores the exponent that was lowered by 1 by the derivative. Hence, Eqs. (30) describe the conservations of the total TCR and CD8 densities, respectively.

Solving Eqs. (30) on the intervals $[0, \rho_T]$, respectively $[0, \rho_C]$, for the both receptor densities T and C yield the densities of unbound TCRs, $T^*(L \mid \theta)$, and unbound CD8, $C^*(L \mid \theta)$, in dependence on ligand concentration and model parameters. Evaluating the scaffold expressions (24) for these values leads to

$$S_j^*(L \mid \theta) := S_j(T^*(L \mid \theta), C^*(L \mid \theta) \mid \theta), \quad j = 1, \ldots, 4, \tag{31}$$

and the density of state u in dependence on ligand concentration and model parameters is then given by

$$X_u^*(L \mid \theta) := g_u \cdot L \cdot \kappa_{\mathrm{MV}}^{m-1} \prod_{j=1}^{4} S_j^*(L \mid \theta)^{u_j} . \tag{32}$$

The observable required in the next chapter is the amount of bound ligands in dependence on ligand concentration and model parameters, $y_L(L \mid \theta)$. Because every ligand-bound state possesses exactly one ligand, $y_L(L \mid \theta)$ is determined by tallying all ligand-bound states evaluated for state densities (32),

$$y_L(L \mid \theta) := \sum_u X_u^*(L \mid \theta)$$

$$= L \sum_{m=1}^{f} \binom{\nu}{m} \kappa_{\mathrm{MV}}^{m-1} \left[\sum_{j=1}^{4} S_j^*(L \mid \theta) \right]^m . \tag{33}$$

To obtain the amount of bound ligands for the other binding models, the procedure outlined in this section is carried out the same way but with the parameters not marked by a cross in Tab. 1 set to zero. In case of Model B, the affinity K_1 has additionally to be rescaled with the factor ε.

4 Identification of pMHC, TCR and CD8 interactions

To elucidate the mechanism by which pMHC interacts with the T cell surface receptors, accurate and quantitative binding data is required. Such data was provided by the lab of Prof. Dr. Wolfgang Schamel from the University of Freiburg, Germany. All experiments were conducted by Sumit Deswal. The connection to the modeling framework, outlined in the last chapter, is made by the conditions under which the experiments were performed. First, T cells were put on ice to prevent incorporation of newly synthesized receptors into the cell membrane as well as internalization of cell surface receptors. This assures constant receptor expression throughout the experiments. Second, the number of T cells was chosen small enough to guarantee ligand excess over total receptor numbers even at smallest concentrations. This allows neglecting the depletion of soluble ligands due to ligand binding. Third, equilibrium binding was established by exposing the cells to soluble ligands for 2.5 hrs. These three conditions justify the general framework outlined in Sec. 3.4.

In this chapter, the binding experiments are described and processing of the data is outlined. Then, a statistically-based confrontation of the interaction models and the data is performed. This leads to the selection of a best-fitting binding model whose parameters are estimated and assigned with confidence intervals.

4.1 Binding experiments

The experiments were performed for different experimental conditions given by the triple

$$\alpha = \big\{ \text{cell type, ligand type, tetramer concentration} \big\} . \tag{34}$$

Table 2: Receptor expression levels[†]

The total numbers of TCRs and CD8 co-receptors on T1.CD8- and T1.CD8+ cells were estimated from triplicate measurements. The results are displayed as mean ± standard error.

[†] The experiments were conducted by Sumit Deswal in the group of Prof. Dr. Wolfgang Schamel from the University of Freiburg, Germany.

cell type	receptor	receptors / cell
T1.CD8-	TCR	21 000 ± 400
	CD8	1390 ± 10
T1.CD8+	TCR	17 200 ± 200
	CD8	144 000 ± 2000

Two T cell types, termed T1.CD8- and T1.CD8+, both expressing the same TCR but differing in CD8 expression levels, were used for data acquisition. The expression numbers of TCRs and CD8 co-receptors on both cell types were determined in a separate experiment. The results are shown in Tab. 2. Assuming the same size for both cell types, the expression numbers also reflect the different receptor densities. As reference density that is used for parameter rescaling in Sec. 3.3, the TCR density on T1.CD8+ T cells is employed. The dimensionless receptor densities appearing in Eqs. (30) are then given by $\rho_T = 1$ and $\rho_C = 8.5$ for T1.CD8+ T cells, and $\rho_T = 1.2$ and $\rho_C = 0$ for T1.CD8- T cells. Hence, the small expression level of CD8 on T1.CD8- T cells is neglected.

Soluble tetramers, carrying four identical pMHC subunits, served as ligands. Four different ligand types were used in the experiments. While the amino acid sequence of the peptides were altered at a single position, the MHC molecule remained unchanged for each ligand type. Thus, each kind of pMHC tetramer has distinct affinities to TCR but the same affinity to CD8. The ligands are referred to as 4L, 4P, 4V, and 4S, and were exposed to the cells with concentrations ranging from 0.3 nM to 846 nM.

After equilibrium binding was established and before the amount of bound ligands was detected, it was necessary to remove unspecifically

Table 3: Dose response binding data[†]

The recored median fluorescence intensity values obtained by flow cytometry for the different experimental conditions (34). Data was gather for two cell types (T1.CD8- and T1.CD8+), four ligand types (4L, 4P, 4V, and 4S) and various tetramer concentrations. In total 145 values were measured distributed over 54 experimental conditions. Triplicates were recorded for 39 conditions, duplicates for 13 conditions and single measurements for 2 conditions. The individual measurement outcomes are displayed in Fig. 5.

[†] The experiments were conducted by Sumit Deswal in the group of Prof. Dr. Wolfgang Schamel from the University of Freiburg, Germany.

cell type	ligand type	tetramer concentration [nM]							
		0.3	1	3.3	10	33	100	330	846
T1.CD8-	4L	0.04	0.12	0.22	0.27	0.34	0.41	0.53	---
		0.03	0.09	0.16	0.22	---	0.34	0.56	---
		0.02	0.09	0.17	0.24	0.27	0.34	---	---
	4P	0.04	0.09	0.19	0.24	0.32	0.38	0.52	0.70
		0.03	0.09	0.18	0.21	0.25	0.33	0.48	---
		0.02	0.07	0.14	0.20	0.24	0.32	---	---
	4V	0.00	0.02	0.05	0.10	0.16	0.26	0.40	---
		0.00	0.02	0.05	0.10	0.15	0.24	0.39	---
		---	0.02	0.04	0.09	0.15	0.26	---	---
	4S	---	---	---	---	0.00	0.02	0.06	---
		---	---	---	---	0.01	0.02	0.05	---
		---	---	---	---	---	0.01	---	---
T1.CD8+	4L	0.11	0.23	0.36	0.54	0.80	1.06	1.47	---
		0.08	0.17	0.28	0.36	0.53	0.81	1.33	---
		---	---	0.28	0.44	0.72	1.41	---	---
	4P	0.18	0.24	0.37	0.51	0.83	1.04	1.36	1.31
		0.11	0.18	0.29	0.39	0.57	0.84	---	---
		0.11	0.21	0.31	0.42	0.64	0.97	1.27	---
	4V	0.12	0.22	0.36	0.42	0.64	0.86	1.26	---
		0.11	0.15	0.24	0.33	0.50	0.78	1.24	---
		0.11	0.19	0.28	0.43	0.67	0.95	---	---
	4S	0.07	0.15	0.27	0.31	0.47	0.69	1.06	---
		0.06	0.14	0.23	0.34	0.55	0.83	1.06	---
		0.08	0.15	0.23	0.35	0.59	0.89	---	---
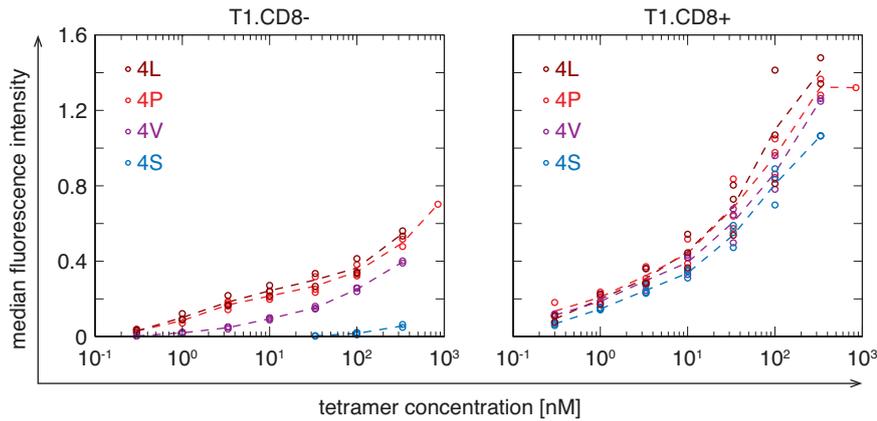
Figure 5: Dose response binding data[†]

The detected median fluorescence intensities (MFI) for T1.CD8- (left) and T1.CD8+ (right) cells in dependence on the concentration of soluble tetramers. For both cell types concentration responses to 4L (brown), 4P (red), 4V (purple), and 4S (blue) ligands were recorded. The circles indicate single measurements whose values are reported in Table 3. The straight dashed lines connect the respective mean values.

[†] The experiments were conducted by Sumit Deswal in the group of Prof. Dr. Wolfgang Schamel from the University of Freiburg, Germany.

bound ligands from the cells. This so called washing procedure also causes a loss of a certain amount of receptor-bound ligands. To keep these losses to a minimum, each pMHC subunit contains a photoreactive 4-azidobenzoic acid that, upon UV-light radiation, possesses a certain probability to co-valently cross-link pMHC to the TCR. No such covalent bond is formed between pMHC and CD8. Only after UV-light exposure, the cells were washed. The remaining amount of cell-bound tetramers was determined as median fluorescence intensity (MFI) values by flow cytometry. The result-ing data is shown in Tab. 3 and Fig. 5.

4.2 Data processing

Reliable outcomes in model selection and parameter estimation require proper uncertainty assignments to the data. Since these uncertainties are

unknown in advance, they have to be estimated from the data. MFI values were measured for 54 different experimental conditions (34). For each condition, n_α repetitive measurements were carried out. The detected MFI value of the k-th experiment repetition of condition α is denoted by $y_{\alpha,k}$, see Tab. 3. It is assumed that the measuring process for each experimental condition α can be modeled by a normal distribution, i.e.

$$Y_{\alpha,k} \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)\,, \tag{35}$$

with $Y_{\alpha,k}$ denoting the random variable describing the measuring process, μ_α the true MFI value, and σ_α the measurement noise. The random variables of the sample mean, $\bar{Y}_\alpha$, and sample variance, $\bar{S}_\alpha^2$, for α are given by

$$\bar{Y}_\alpha = \frac{1}{n_\alpha} \sum_{k=1}^{n_\alpha} Y_{\alpha,k}\,, \tag{36}$$

$$\bar{S}_\alpha^2 = \frac{1}{n_\alpha - 1} \sum_{k=1}^{n_\alpha} (Y_{\alpha,k} - \bar{Y}_\alpha)^2\,. \tag{37}$$

For the sample distribution of the mean and the sample distribution of the variance follows

$$\frac{\bar{Y}_\alpha - \mu_\alpha}{\sigma_\alpha/\sqrt{n_\alpha}} \sim \mathcal{N}(0,1)\,, \tag{38}$$

$$\frac{(n_\alpha - 1)\,\bar{S}_\alpha^2}{\sigma_\alpha^2} \sim \chi_{n_\alpha-1}^2\,, \tag{39}$$

with $\chi_{n_\alpha-1}^2$ the chi-squared distribution with $n_\alpha - 1$ degrees of freedom. Since the number of repetitions is quite small ($n_\alpha \leq 3$), error assignment based on the sample variance is subject to large uncertainties[4]. Therefore, an error model including all data is constructed allowing uncertainty assignments with higher level of precision.

---

[4]The mean and variance of $\chi_{n_\alpha-1}^2$ are given by $n_\alpha - 1$ respectively $2\,(n_\alpha - 1)$.

The error model comprises of a single parameter, the coefficient of variation ϑ, which is assumed to be independent of α,

$$\vartheta \equiv \frac{\sigma_\alpha}{\mu_\alpha} \,. \tag{40}$$

This means that the measurement uncertainty of α is proportional to the detected signal, and that the proportionality factor – the coefficient of variation – is common for each experimental condition. Thus, the data of each α delivers an estimate for ϑ. The corresponding random variable $\bar{T}_{n_\alpha}$ is given by

$$\bar{T}_{n_\alpha} = \frac{\bar{S}_\alpha}{\bar{Y}_\alpha} \,, \tag{41}$$

and depends on the number of repetitions n_α. Starting from distributions (38) and (39), the sample distribution of the coefficient of variation, i.e. the distribution of $\bar{T}_{n_\alpha}$, for n_α repetitions, p_{n_α}, is determined to be,

$$p_n(t \,|\, \vartheta) = C_n \cdot t^{n-2} \cdot \left[\lambda_n(t)\right]^{\frac{n}{2}} \cdot {}_1F_1\left(\tfrac{n}{2}, \tfrac{1}{2}, \tfrac{n^2\,\lambda_n(t)}{2\,\vartheta^2}\right) \cdot \exp\left[-\frac{n}{2\,\vartheta^2}\right], \tag{42}$$

$$\lambda_n(t) = \frac{1}{n + (n-1)\,t^2} \,,$$

$$C_n = \frac{2\sqrt{n}\,(n-1)^{(n-1)/2} \cdot \Gamma\left(\tfrac{n}{2}\right)}{\sqrt{\pi} \cdot \Gamma\left(\tfrac{n-1}{2}\right)} \,,$$

with ${}_1F_1(\cdot,\cdot,\cdot)$ being Kummer's confluent hypergeometric function defined by

$${}_1F_1(a, b, z) := \frac{\Gamma(b)}{\Gamma(a)\,\Gamma(b-a)} \int_0^1 e^{z\,x}\, x^{a-1}\,(1-x)^{b-a-1}\, \mathrm{d}x \,,$$

and $\Gamma(\cdot)$ being the Gamma function,

$$\Gamma(x) := \int_0^\infty t^{x-1}\, e^{-t}\, \mathrm{d}t \,.$$

From the sample distribution of the coefficient of variation (42), the log-likelihood function (5) is computed, serving as starting point for MLE and the profile likelihood-based confidence interval (see Sec. 2.1 respectively
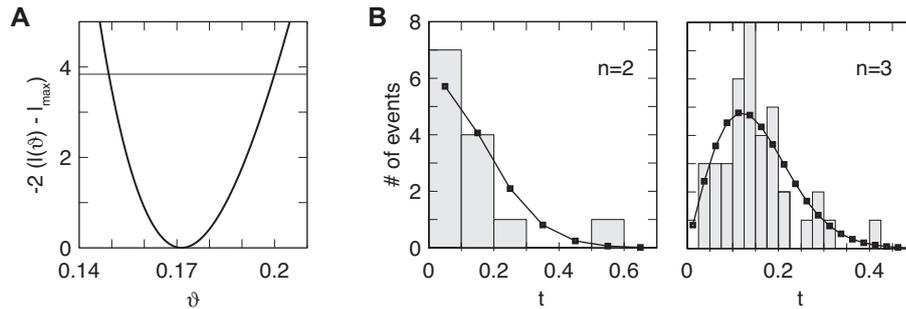
Figure 6: Error model

(A) Profile likelihood of the error model (black curve). The MLE of the coefficient of variation ϑ is found to be 0.17. The intersection of the profile likelihood curve and the gray line, located at 3.84, indicate the lower and upper bounds of the profile likelihood based 95% confidence interval. The value $l_{\max}$ denotes the maximal value of the log-likelihood function $l(\vartheta)$. (B) The binned estimates of the coefficient of variation (gray bars) and the model predictions (black squares) for $n = 2$ (left) and $n = 3$ (right) repetitions.

Sec. 2.3). The results are summarized in Fig. 6. The MLE value $\hat{\vartheta}$ assumes 0.17. The lower and upper bounds of the profile likelihood-based 95 % confidence interval are given by 0.15 and 0.20, respectively (Fig. 6A). Fig. 6B shows the predicted (black squares) and measured (grey bars) numbers of values of the coefficient of variation for $n = 2$ (left) and $n = 3$ (right) repetitions. The good agreement indicates the plausibility of the error model.

Having obtained an estimate for the coefficient of variation, assignments to the mean and standard deviation used in model fitting can be made,

$$\hat{y}_\alpha = \frac{1}{n_\alpha} \sum_{k=1}^{n_\alpha} y_{\alpha,k} \,, \tag{43a}$$

$$\hat{s}_\alpha = \frac{\hat{\vartheta}\, \hat{y}_{\alpha,n}}{\sqrt{n_\alpha}} \,. \tag{43b}$$

In this sense, the measurement uncertainties are taken to be known prior to model fitting.

4.3   Washing model

Before coming to model selection and parameter estimation, the washing procedure has to be taken into account. As described in Sec. 4.1, this is a two step process. At first, the cells were exposed to UV-light causing a certain fraction of pMHC-TCR contacts to cross-link covalently. Only ligands with no such covalent liaison are susceptible to being washed off in the second step. The incorporation of these steps is at first outlined for Model A in the following and subsequently generalized to include the second receptor type CD8.

In order to be detected by flow cytometry, it is sufficient that a single pMHC subunit of a tetramer remains bound until after the washing procedure. Therefore, the fraction of ligands loosing all established bonds is appraised in the following. The probability that a given pMHC-TCR contact forms no covalent bond upon UV-light exposure is denoted by p_{UV}. If no such bond is established, the contact may break during the second washing step, and the chance of this to happen is the higher the weaker the strength of the bond. This is modeled by relating the strength of a bond – taken to be reflected by the pMHC-TCR affinity – with a threshold parameter K_{W} having the unit of a binding affinity. In particular, it is assumed that the probability of a given pMHC bond to break can be described by the following expression,

$$p_{\mathrm{loss},1} \equiv p_{\mathrm{UV}}\, \frac{K_{\mathrm{W}}}{K_{\mathrm{W}} + K_1}\,. \tag{44a}$$

It is composed of the independent probabilities p_{UV} and $K_{\mathrm{W}}/(K_{\mathrm{W}} + K_1)$, with the latter modeling the chance of the pMHC-TCR bond to break during the second washing step. The washing loss is the greater the smaller the affinity is compared to the parameter K_{W}. For a ligand having established m pMHC-TCR contacts, the fraction of such ligands getting lost during the washing procedure is given by $(p_{\mathrm{loss},1})^m$. Hence, it is assumed that each bond breaks independently of the others.

The generalization of the washing model to include CD8 is outlined for Model G in the following. As described in Sec. 3.4, each ligand-bound state of Model G is characterized by a set of four numbers $u = \{u_1, u_2, u_3, u_4\}$ denoting how many pMHC subunits are bound in which way to cell surface receptors (see Fig. 4). Only pMHC subunits in direct contact with TCR, i.e. pMHC bound only to TCR as well as pMHC bound via aggregated state, can benefit from UV-light exposure. To describe the loss in the second washing step for the different ways pMHC can be bound to cell surface receptor, expression (44a) is generalized to include (i) pMHC-CD8 bonds by using K_1' instead of K_1 as affinity, (ii) aggregated states by rescaling K_1 with ξ to account for an impact of adjacent CD8, and (iii) pMHC bound to the CD8 binding site of TCR-CD8 complex by using $q_4 K_1'$ instead of K_1 to account for binding to CD8 and the impact of adjacent TCR. In analogy to expression (44a), the probabilities of being disrupted during the washing procedure for the different pMHC bond types are taken to be

$$p_{\text{loss},2} \equiv \frac{K_{\text{W}}}{K_{\text{W}} + K_1'} \,, \tag{44b}$$

$$p_{\text{loss},3} \equiv p_{\text{UV}} \, \frac{K_{\text{W}}}{K_{\text{W}} + \xi \, K_1} \,, \tag{44c}$$

$$p_{\text{loss},4} \equiv \frac{K_{\text{W}}}{K_{\text{W}} + q_4 \, K_1'} \,. \tag{44d}$$

Also in this case, it is assumed that each bond breaks independently of the others. A ligand in binding state u has to loose all its $u_1 + u_2 + u_3 + u_4$ bonds with respective probabilities (44) in order to be washed off. At concentration L and for parameters θ, state u is occupied with density $X_u^*(L \mid \theta)$, defined in (32). The amount of ligands in state u being washed for these conditions is given by

$$\tilde{X}_{\text{loss},u}(L \mid \theta) := \prod_{j=1}^{4} p_{\text{loss},j}^{u_j} \cdot X_u^*(L \mid \theta) \,. \tag{45}$$

Using the expression for the state density (32) and exploiting the multi-nomial theorem once more, the total amount of ligands lost during the washing procedure is given by

$$y_{\text{loss}}(L \mid \theta) := \sum_{u} \tilde{X}_{\text{loss},u}(L \mid \theta)$$

$$= L \sum_{m=1}^{f} \binom{\nu}{m} \kappa_{\text{MV}}^{m-1} \left[\sum_{j=1}^{4} p_{\text{loss},j}(\theta) \cdot S_j^*(L \mid \theta) \right]^m. \qquad (46)$$

This expression has to be subtracted from the amount of bound ligands (33) to obtain the amount of detected ligands,

$$y_D(L \mid \theta) := y_B(L \mid \theta) - y_{\text{loss}}(L \mid \theta). \qquad (47)$$

The washing losses for the other models are found by setting those parameters to zero that are not marked by a cross in Tab. 1. For Model B the affinity K_1 has further to be replaced by the apparent affinity εK_1.

4.4 Model selection and parameter estimation

Having the observable (47) for each model calculated, it can be confronted with the data (Tab. 3 and Fig. 5) in its processed form (43). The data of T1.CD8- cells is always described by Model A, while the data of T1.CD8+ cells is interpreted in terms of one of the Models B–G. Since four ligand types were used, there are four different pMHC-TCR affinities $K_{1,X}$, with $X = $ 4L, 4P, 4V, or 4S. Hence, each model recruits its parameters from the following set of parameters,

$$\theta \in \big\{ K_{1,4\text{L}}, \, K_{1,4\text{P}}, \, K_{1,4\text{V}}, \, K_{1,4\text{S}}, \, \varepsilon, \, K_1',$$

$$\kappa_{\text{MV}}, \, f, \, q_1, \, \xi, \, q_4, \, p_{\text{UV}}, \, K_{\text{W}}, \, M^* \big\}, \qquad (48)$$

depending on the monomeric binding scheme (Tab. 1), and whether multivalent binding is possible. The parameter set also comprises of the parameters used to describe the washing procedure, p_{UV} and K_{W}, as well as

a parameter M^* denoting a scaling factor relating the predicted amount of cell-bound ligands to the arbitrary MFI scale of the data. The effective ligand valency f was not estimated alongside the other parameters. Instead, for each of the 6 possible binding models for T1.CD8+ cells, it was set to one of the allowed values 1 to 4. Hence, a total of 24 binding model-effective valency combinations had to be tested. Importantly, only the binding affinities $K_{1,X}$, with X = 4L, 4P, 4V or 4S, depend on the presented peptide; all other parameters assume a common value for all ligands. Thus, binding of different ligands is characterized by changing only a single parameter.

As already assumed in Sec. 4.2, the measuring process for individual data acquisition follows a normal distribution. Owing the error model, the estimated means (43a) are also described by a normal distribution with known standard deviation (43b). Hence, maximum likelihood estimation is equivalent to least squares estimation (see Sec. 2.2), and the objective function is given by

$$\chi^2(\theta) = \sum_\alpha \left(\frac{\hat{y}_\alpha - M^* \, y_{D,\alpha}(\theta)}{\hat{s}_\alpha} \right)^2 . \tag{49}$$

The concentration dependence of $y_D(L \mid \theta)$ has been shifted to the index α describing the experimental condition (34) yielding the predicted amount of detected ligand $y_{D,\alpha}(\theta)$ depending only on the model parameters θ. The sum in (49) runs over all experimental conditions meaning that all data is taken into account at once. The minimization of (49) was performed using the built-in function *lsqnonlin* of MATLAB. For each model, the search was started from at least 1000 different initial conditions. Model selection was done by exploiting the corrected Akaike Information Criterion and the quality of the fit was accessed by the goodness-of-fit (see Sec. 2.5 for model selection and Sec. 2.2 for the goodness-of-fit). The fitting results are summarized in Tab. 4. Based on the corrected Akaike Information Criterion, Model G with effective ligand valency 2 and 3 are in best agreement with

Table 4: Fitting Results

The first column denotes the model used to describe the data of T1.CD8+ T cells, and the second column the effective ligand valency. In total, there are 24 model-effective valency combinations. The third column shows the number of parameters used to fit the corresponding model including the washing and scaling parameters. The column with head $\chi^2_{\min}$ gives the value of the objective function (49) obtained at MLE values of the parameters. The goodness-of-fit (gof, see Eq. (11)) is displayed in the fifth column. Finally, the difference of the corrected Akaike Information criterion (ΔAIC_c, see Sec. 2.5) is shown in the last column. The gray frame surrounding Model G with effective ligand valencies of 2 and 3 indicates that these two models are in best agreement with the data, while all other models were ruled out according to $\Delta AIC_c > 10$ [8].

model	effective valency	# of parameters	$\chi^2_{\min}$	gof	ΔAIC_c
	1	8	928	20.2	853
B	2	9	723	16.1	651
	3	9	586	13.0	514
	4	9	534	11.9	462
	1	8	280	6.1	206
C	2	9	219	4.9	147
	3	9	195	4.3	123
	4	9	189	4.2	117
	1	9	270	6.0	198
D	2	10	183	4.2	114
	3	10	147	3.3	78
	4	10	137	3.1	68
	1	9	270	6.0	198
E	2	10	183	4.2	114
	3	10	147	3.3	78
	4	10	137	3.1	68
	1	10	192	4.4	123
F	2	11	123	2.9	57
	3	11	99	2.3	33
	4	11	91	2.1	25
	1	11	104	2.4	38
G	2	12	62	1.5	0
	3	12	67	1.6	5
	4	12	79	1.9	16
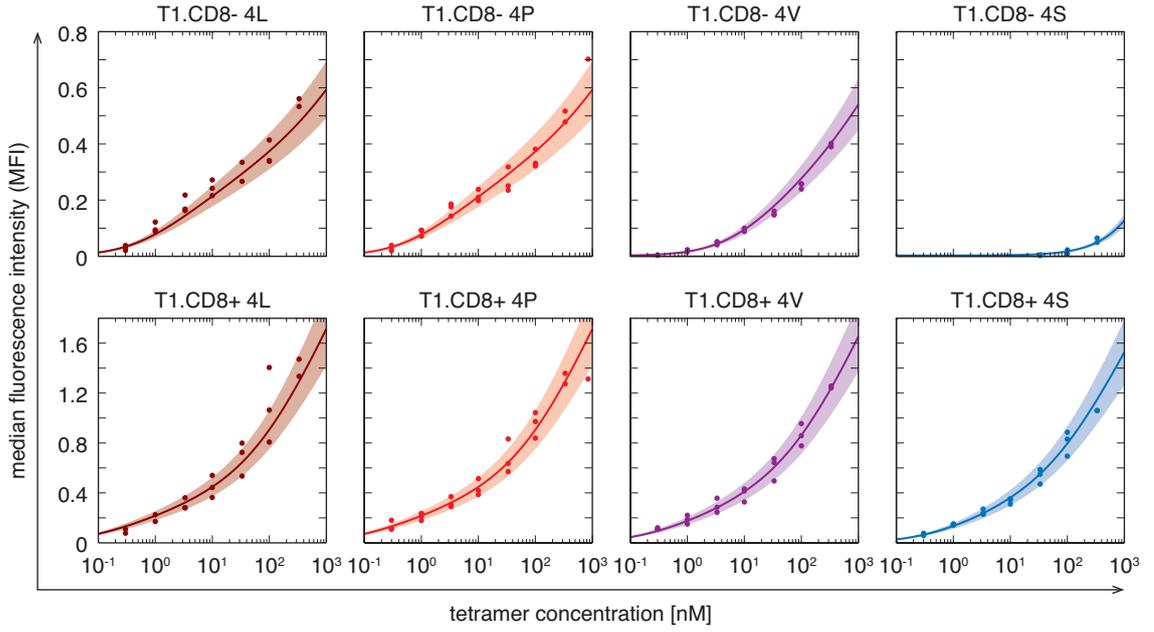
Figure 7: Data fit of Model G3

The measured median fluorescence intensities (dots) and the model fits (full lines) in dependence on tetramer concentration are displayed. The top row shows T1.CD8- and the bottom row T1.CD8+ cells. The different ligand types are indicated by color: 4L in brown, 4P in red, 4V in purple, and 4S in blue. The shaded areas describe the standard deviation of single measurements as estimated from the error model.

the data. All other models were ruled out according to $\Delta\text{AIC}_\text{c} > 10$ [8]. The gof of Models G2 and G3 are 1.5, respectively 1.6, indicating that the data is well described by both models. It is generally believed that tetramers can engage up to three pMHC simultaneously [85], i.e. $f = 3$. Hence, Model G3 is further discussed in the following. The results of Model G2 can be found in Appendix A. The performance of Model G3 on the data is depicted in Fig. 7 showing that the data is well captured by the model.

The uncertainties in the estimated parameter values in terms of 95% confidence intervals were calculated by the profile likelihood method (see Sec. 2.3). The profile likelihood curves are shown in Fig. 8A. All parameters, except for q_1, ξ and p_{UV}, are completely identifiable, i.e. they have an upper and lower bound. While q_1 and ξ have at least an upper
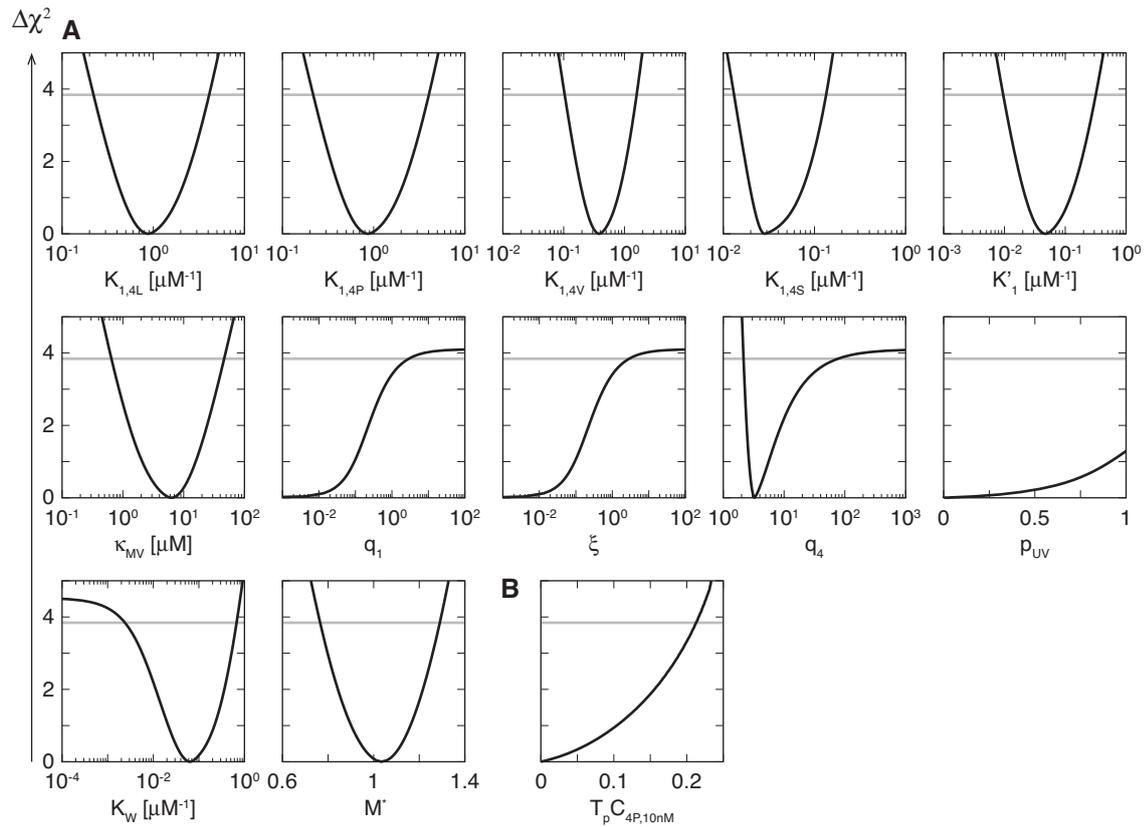
Figure 8: Profile likelihoods of Model G3

(A) The profile likelihood curves (black) of the 12 parameters used for fitting Model G3. The intersections of the profile likelihood curves with the gray lines, located at $\Delta\chi^2 = 3.84$, define the independent 95 % confidence intervals of the model parameters. (B) Prediction profile likelihood of $T_pC_{4P,10\,nM}$, the fraction of TCRs in aggregated state for the 4P ligand at a concentration of $10\,nM$. The gray line plays the same role as in (A).

bound, p_{UV} could not be determined. Thus, the model parameters are well constraint by the data. The best-fit parameters θ_{MLE}, and the lower and upper bounds of the 95 % confidence intervals, θ_{low} respectively θ_{up}, are shown in Tab. 5. The estimated dissociation constants – the inverses of the binding affinities – lie between $0.25\,\mu M$ and $5\,\mu M$ for the 4L and 4P ligand, between $0.5\,\mu M$ and $10\,\mu M$ for the 4V ligand, between $10\,\mu M$ and $100\,\mu M$ for the 4S ligand, and between $3\,\mu M$ and $100\,\mu M$ for CD8 binding to pMHC. These values are in good agreement with previous reports of pMHC-

Table 5: Parameter values of Model G3

The maximum likelihood estimations θ_{MLE}, and the lower and upper bounds of the profile likelihood-based $95\,\%$ confidence intervals, θ_{low} and θ_{up}, of the 12 parameters used for fitting Model G3 are shown. The last column gives the values θ^* obtained by prediction profile likelihood as described in the main text. These values are assigned to the parameters of Model G3.

parameter	θ_{MLE}	θ_{low}	θ_{up}	θ^*
$K_{1,4L}$ [μM^{-1}]	0.9	0.2	4	1.1
$K_{1,4P}$ [μM^{-1}]	0.9	0.2	4	1.1
$K_{1,4V}$ [μM^{-1}]	0.4	0.1	2	0.5
$K_{1,4S}$ [μM^{-1}]	0.03	0.01	0.1	0.03
K'_1 [μM^{-1}]	0.05	0.01	0.3	0.06
κ_{MV} [μM]	6	0.7	45	4.7
q_1	0	0	3	0.18
ξ	0	0	3	0.18
q_4	3.3	2.2	80	6.9
p_{UV}	0	0	1	0
K_W [μM^{-1}]	0.06	0.002	0.7	0.09
M^*	1.0	0.8	1.3	1.0

TCR affinities [18] and murine MHC-CD8 affinity [30]. The multivalent engagement parameter κ_{MV} is estimated lie between $0.7\,\mu M$ and $45\,\mu M$. This value is an indicator for the capability of a ligand to form multivalent binding states. Ligands with dissociation constants larger than κ_{MV} hardly bind multivalently. In case of the cognate 4P ligand, the MLE value of the dimensionless association constant of multivalent binding, $\kappa_{\mathrm{MV}}\,K_{1,4P}$ assumes a value of about 5. This is in good agreement with previously reported values on such dimensionless multivalent binding constants [35, 23, 73]. The accordance of the estimated affinities and multivalent binding parameters with the literature is further evidence on the plausibility of the model and its underlying assumptions.

The association constant for pMHC-independent TCR-CD8 complex formation q_1 is zero for the best fit and has an upper bound of 3. Thus, up to $95\,\%$ of the TCRs on T1.CD8+ cells might be bound to CD8 in the absence of ligands. However, because of this large uncertainty and

due to the fact that this value also depends on the relative expression level of CD8 compared to TCR, a clear estimate of the fraction of TCRs in TCR-CD8 complexes in the absence of ligands cannot be made. A striking finding is that q_4, the association constant of TCR-CD8 complex formation if pMHC is bound to CD8, is significantly larger than q_1. This means that pMHC-bound CD8 has an increased ability to bind TCR compared to CD8 alone, or, put equivalently, the affinity of pMHC binding to CD8 is enhanced if CD8 is bound to a TCR. A similar effect is not observed for pMHC-bound TCRs. As ξ is very similar to q_1, an enhancing impact for pMHC-TCR binding if CD8 is bound to the TCR, or pMHC binding to the TCR binding site of a TCR-CD8 complex, is – if present at all – very small. The probability to not form a covalent bond between a pMHC-TCR contact upon UV-light irradiation, p_{UV}, could not be determined. A reason might be that, in the concentration range used for data acquisition, multivalent binding dominates over monovalent binding – at least for the 4L, 4P, and 4V ligand. Multivalently bound ligands have a better chance to survive cell washing as rebinding of a ligand subunit is possible before all other subunits loose contact [85], and hence UV-light induced cross-linking of pMHC and TCR might not be too relevant. The estimated value of the washing parameter K_{W} indicates that substantial washing loses are expected for the weak binding 4S ligand as well as for pMHC-CD8 bonds. The higher affinity ligands are less impacted by the washing step. Finally, the scaling factor M^* assumes a value of 1. This was expected as the MFI scale in Fig. 5 is of that order of magnitude.

For the further discussion of Model G3, it is convenient to not use θ_{MLE}, since the parameter ξ is zero (Tab. 5). As ξ describes the formation of aggregated states, i.e. TCRs in direct contact with pMHC and adjacent CD8, it is of particular biological interest, and also plays a central role in the subsequent chapters. To find a parameter set, on the one hand in agreement with the data, and on the other hand allowing for the forma-

tion of aggregated states, the method of prediction profile likelihoods was exploited (see Sec. 2.4). As prediction, i.e. as quantity deviated from its MLE value, the fraction of TCRs in aggregated state that establishes for the 4P ligand at $10\,\text{nM}$ concentration ($\text{T}_\text{p}\text{C}_{4\text{P},10\,\text{nM}}$) was chosen. The prediction profile likelihood curve is shown in Fig. 8B. The parameter set θ^*, obtained at $\Delta\chi^2 = 2$, is given in the last column of Tab. 5, and henceforth used to parameterize Model G3.

5 Competition binding model

This section focuses on the discussion of Model G that from here on is termed Competition Binding Model (CBM) for reasons that will become clear later on. The parameterization of CBM is given by the parameter set θ^*, see the last column of Tab. 5. All possible binding states of CBM for soluble tetramer with effective valency $f = 3$ are shown in Fig. 4. In this setup, the impact of CD8 on ligand binding and receptor occupancy is investigated, and multivalent binding is considered as predictor for T cell activation.

5.1 Ligand binding and receptor occupancy

In CBM, each pMHC subunit is bound according to one of four possibilities; three of which involve TCRs. The corresponding fractions of TCRs in dependence on ligand concentration with respective pMHC contacts are given by

$$y_{\mathrm{Tp}}(L) := \sum_u u_1 X_u^*(L\,|\,\theta^*)\,, \tag{50a}$$

$$y_{\mathrm{TpC}}(L) := \sum_u u_3 X_u^*(L\,|\,\theta^*)\,, \tag{50b}$$

$$y_{\mathrm{TCp}}(L) := \sum_u u_4 X_u^*(L\,|\,\theta^*)\,, \tag{50c}$$

where the sums run over all binding states u. Expression (50a) denotes the fraction of TCRs bound only to pMHC, because $X_u^*(L\,|\,\theta^*)$ is the density of binding state u in dependence on ligand concentration (32) evaluated for parameters θ^*, and u_1 is the number of pMHC-TCR bonds in that binding state. Likewise, y_{TpC} is the fraction of TCRs in aggregated states, and y_{TCp} the fraction of TCRs in TCR-CD8 complexes to which pMHC is only bound to the CD8 subunit.

The curve progressions of these fractions as well as their total in dependence on ligand concentration for the 4P, 4V, and 4S ligands are shown in Fig. 9A. The totals are very similar irrespective of the strength of pMHC-TCR binding (full curves). This property is also reflected in the binding data of T1.CD8+ cells (Fig. 5). However, the composition of the totals exhibit great differences between the different ligands. For concentrations less than $1000\,\mathrm{nM}$, the 4S ligand, which has a pMHC-TCR affinity of $0.03\,\mu\mathrm{M}^{-1}$, binds almost exclusively to the CD8 binding site of TCR-CD8 complexes (blue dotted curve) and barely to TCRs directly (blue dashed and blue dashed-dotted curves). Thus, in this concentration range, binding of the weak 4S ligand is dominated by ligand-unspecific interactions. At concentrations larger than $1000\,\mathrm{nM}$, direct contacts between pMHC and TCRs (y_{Tp}) emerge, which is accompanied by the decline of interactions of pMHC with the CD8 binding site of TCR-CD8 complexes. The reason for this behavior is that at such large concentrations, each receptor preferentially binds its own ligand rather than sharing one, as it is the case for the TCR-CD8 complex. This is the same argument explaining the bell-shaped curve of bivalent binding in Fig. 1. Thus, it is only at very large concentrations that low affinity ligands can bind TCRs directly, as otherwise the CD8 binding site of TCR-CD8 complexes offers a strong and ligand-unspecific competitor. This competition property is what gives the name to CBM. Note that at no concentration value appreciable amounts of aggregated states are formed.

The ligands 4P and 4V, having pMHC-TCR affinities of $1.1\,\mu\mathrm{M}^{-1}$, respectively $0.5\,\mu\mathrm{M}^{-1}$, manage to form direct contacts between pMHC and TCRs. This includes TCRs binding pMHC alone (dashed curves), as well as aggregated states (dash-dotted curves). Importantly, the higher the pMHC-TCR affinity the more such bonds form. This coincides with less pMHC binding to the CD8 binding site of TCR-CD8 complexes (dotted curves). Thus, high affinity ligands are able to counteract the strong and
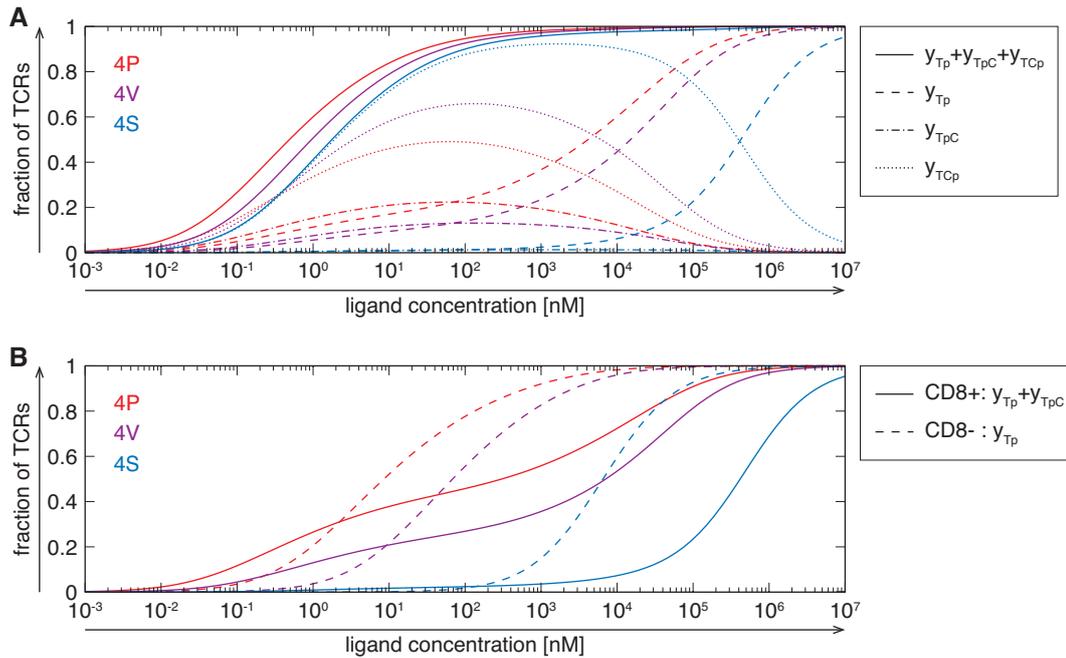
Figure 9: Receptor occupancy for CBM

(A) Fraction of TCRs bound according to the different interaction possibilities of CBM for the 4P (red), 4V (purple) and 4S (blue) ligands in dependence on ligand concentration. The dashed lines show TCRs bound to pMHC alone, the dash-dotted lines TCRs in aggregated state, the dotted lines TCRs in TCR-CD8 complexes to which pMHC is bound to the CD8 binding site, and the full lines the total of the three. (B) Impact of CD8 on the fraction of TCRs in direct contact with pMHC in dependence on ligand concentration using the same ligands and color code as in (A). The full lines show binding for T1.CD8+ cells and the dashed lines for T1.CD8- cells.

ligand-unspecific interactions with the CD8 binding site of TCR-CD8 complexes showing that despite similar total binding, ligand-specific binding is manifested in CBM. The decline of pMHC bonds with TCR-CD8 complexes, i.e. y_{TpC} and y_{TCp}, at high concentration has the same reason as above; each ligand prefers binding its own receptor.

The finding that total binding is very similar for all ligands, but the individual contributions are distinct can be understood qualitatively in the framework of a toy model of CBM. Assume a receptor possessing two binding sites; one for productive and one for unproductive binding. Importantly, only one of the two binding sites can be bound simultaneously.

The productive binding site is supposed to reflect binding in which pMHC is in direct contact with a TCR, i.e. y_{Tp} and y_{TpC}, while unproductive binding mirrors binding to the CD8 binding site of the TCR-CD8 complex y_{TCp}. The affinity of a (monovalent) ligand for productive binding is denominated with K_P and depends on ligand quality, while the affinity of unproductive binding, termed K_U, is independent thereof. Unproductive binding is of high affinity; i.e. it holds $K_P \leq K_U$. In particular, for ligands with low productive affinity one has $K_P \ll K_U$, while ligands with high productive affinity fulfill $K_P \approx K_U$. The fraction of occupied receptors (total binding), is given by

$$y_{\mathrm{total}} \equiv \frac{(K_P + K_U)\,L}{1 + (K_P + K_U)\,L} \approx \begin{cases} \dfrac{2\,K_U\,L}{1 + 2\,K_U\,L} & \text{if } K_P \approx K_U, \\[2ex] \dfrac{K_U\,L}{1 + K_U\,L} & \text{if } K_P \ll K_U. \end{cases} \tag{51}$$

Thus, total binding is very similar for all ligands independent of their ability to bind productively, as the overall affinity $K_P + K_U$ differs at most by a factor of 2 between different affinity ligands. Nevertheless, the fraction of productively bound ligands, given by

$$y_{\mathrm{prod}} \equiv \frac{K_P\,L}{1 + (K_P + K_U)\,L} \leq \frac{K_P}{K_P + K_U}\,, \tag{52}$$

shows a distinct dependence on affinity K_P. For ligands with low productive affinity, a proportion of at most $K_P/K_U \ll 1$ receptors are bound productively. On the other hand, high affinity ligands can achieve a non-zero fraction of productive interactions. This shows how the existence of a TCR-CD8 complex with a ligand-unspecific high affinity binding site explains the great similarity in ligand binding observed, while the ability to distinguish between different pMHC-TCR affinity ligands is maintained.

The impact of CD8 on direct pMHC-TCR binding is shown in Fig. 9B. In general, on T1.CD8+ cells, more bonds between pMHC and TCR are formed at lower concentrations than on T1.CD8- cells, and fewer at

higher concentrations. The points of intersection, however, depend on lig-
and affinity. The lower the affinity the higher the concentration value, and
the smaller the fraction of pMHC-bound TCRs at the intercept. Due to
this property, the fraction of TCRs in direct contact with pMHC for the
4S ligand remains at a relatively low level much longer than in the absence
of CD8. Additionally, high affinity ligands achieve a substantially higher
degree of direct pMHC-TCR contacts at low concentrations in the presence
of CD8. Because foreign (high affinity) antigen is expected to be presented
at relatively low doses while self (low affinity) peptides are presented at
rather high doses, this finding suggests that CBM possesses the capability
to detect foreign antigen with greater sensitivity, and to improve self tol-
erance, i.e. the ability not not respond to self peptides, if CD8 is present.
However, because all ligands eventually reach full occupancy independent
of ligand affinity and CD8 expression, low affinity can be compensated by
high dose offering the possibility to compromise ligand discrimination.

5.2   Multivalent binding and TCR activity

As pointed out in Sec. 1.1, multivalent engagement, in contrast to receptor
occupancy, allows for affinity-based ligand discrimination and T cell acti-
vation if T cells are probed with soluble ligands. Therefore, multivalent
binding of CBM is investigated in the following. In the absence of CD8,
CBM reduces to Model A, which comprises only of the states $\{m, 0, 0, 0\}$,
with m assuming the values 1 to 3 (see Fig. 4). All states obeying $m \geq 2$,
i.e. all states in which a ligand binds at least two TCRs simultaneously,
are referred to as signaling competent states, and each TCR in such states
is regarded as a signaling TCR. As has been shown by Perelson [59], mul-
tivalent binding curves in dependence on ligand concentration, established
by oligomers with valency ν and effective valency $f \geq 2$, are bell-shaped
(see Fig. 1), and have a single maximum located at concentration value

L_{pot} given by

$$
L_{\text{pot}} \equiv
\begin{cases}
\dfrac{1}{\nu \, K_1} & \text{if} \ \ \kappa_{\text{MV}} \, K_1 \ll 1, \\[3mm]
\dfrac{1}{(f-1) \, \nu \, K_1} & \text{if} \ \ \kappa_{\text{MV}} \, K_1 \gg 1.
\end{cases}
\tag{53}
$$

The (dimensionless) parameter $\kappa_{\text{MV}} \, K_1$ assumes the values 5.2, 2.4, and 0.1 for the 4P, 4V, and 4S ligand, respectively, indicating that L_{pot} lies between $110\,\text{nM}$ to $230\,\text{nM}$ for 4P ligand, $250\,\text{nM}$ to $500\,\text{nM}$ for the 4V ligand, and $4200\,\text{nM}$ to $8300\,\text{nM}$ for the 4S ligand, with the latter expected to be at the upper bound due to the smallness of $\kappa_{\text{MV}} \, K_1$. The fractions of signaling TCRs for the 4P, 4V, and 4S ligand in dependence on ligand concentration are shown in Fig. 10A confirming the afore mentioned. Owing the assumption of a uniform multivalent binding parameter κ_{MV}, high affinity ligands induce signaling at lower concentrations than low affinity ligands, and the maximal amount of signaling TCRs increases with ligand affinity [75]. Considering a model of T cell activation that requires a minimal number of signaling TCRs in order to trigger a T cell, ligands with too low affinity are incapable of activating T cells as they fail to establish this minimal number of signaling TCRs. Hence, multivalent engagement allows for affinity-based ligand discrimination.

In the following, these arguments are generalized to CBM, and it is investigated whether in this case affinity-based ligand discrimination is still possible, and how the influence of CD8 is manifested. CBM comprises of 3 bivalent and 10 trivalent ligand-bound states that contain at least two TCRs in direct contact with pMHC. Each TCR in such a state that is in direct contact with pMHC is regarded as a signaling TCR (indicated by subjacent arrows in Fig. 4). These TCRs are divided into two groups: weak signaling TCRs (thin arrows) and strong signaling TCRs (thick arrows). Strong signaling TCRs have an adjacent CD8 co-receptor (aggregated state) while weak signaling TCRs lacking it. The reason for this nomenclature is that CD8 recruits the kinase Lck to the TCR [83, 79].
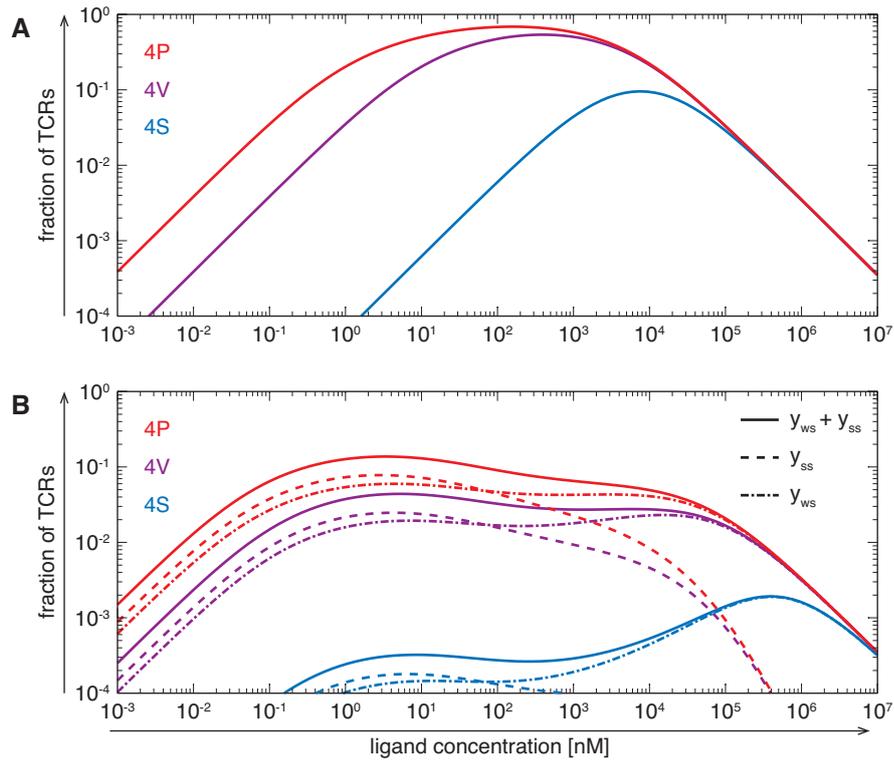
Figure 10: Signaling TCRs for CBM

Fraction of signaling TCRs in dependence on ligand concentration for the 4P, 4V, and 4S ligand as indicated by color. (A) In the absence of CD8. (B) With CD8. Dashed lines show strong signaling TCRs, y_{ss}, dashed-dotted lines weak signaling TCRs, y_{ws}, and the full lines all signaling TCRs, $y_{ws}+y_{ss}$.

Using the notation introduced in Sec. 3.4, the proportions of weak signaling TCRs, y_{ws}, respectively strong signaling TCRs, y_{ss}, are given by

$$y_{ws}(L) := \sum_{\substack{u \\ u_1+u_3 \geq 2}} u_1\, X_u^*(L\,|\,\theta^*)\,, \tag{54}$$

$$y_{ss}(L) := \sum_{\substack{u \\ u_1+u_3 \geq 2}} u_3\, X_u^*(L\,|\,\theta^*)\,. \tag{55}$$

The sum runs over all binding states u that contain at least to TCRs in direct contact with pMHC, i.e. pMHC-TCR (u_1) or aggregated state (u_3). The total fraction of signaling TCRs in dependence on ligand concentration is given by the full lines in Fig. 10B for the 4P, 4V, and 4S ligand. As for T1.CD8- cells, the higher the affinity the earlier signaling TCRs emerge

and the larger the maximal number of signaling TCRs becomes. Thus, affinity-based ligand discrimination is also possible in the presence of CD8. However, there are notable differences between the presence and absence of CD8. To begin with, in the presence of CD8, fewer TCRs are bound appropriate for signaling than without CD8 (compare full lines in Figs. 10A and B). The reason is that CBM comprises of several multivalent TCR binding states that are not signaling competent (see Fig. 4). Especially states in which a pMHC subunit is bound to the CD8 binding site of a TCR-CD8 complex (y_{TCp} in Fig. 9A) are prominently occupied leading to a depleted availability of TCRs able of forming a signaling competent state. A further difference is that, for T1.CD8+ cells, the total fractions of signaling TCRs are bimodal functions of ligand concentration. The total fraction of signaling TCRs is composed of weak and strong signaling TCRs (dashed and dash-dotted lines in Fig. 10B). As outlined in Sec. 5.1, simultaneous binding of pMHC to TCR and CD8, as in binding states with u_3 or $u_4 > 0$, disappears at high concentrations. Thus, TCRs from states with $u_4 > 0$, which are not signaling, are released allowing the formation of states containing weak signaling TCRs (dash-dotted lines in Fig. 10B). This is especially pronounced for low affinity ligands, because for those, pMHC binding is dominated by binding to the CD8 binding site of TCR-CD8 complex. Hence, a second peak is formed at high concentrations, which is the more evident the lower the affinity.

Despite the finding that fewer signaling TCRs are established for T1.CD8+ cells, it does not compulsorily result in diminished T cell activation. Among the signaling TCRs of CBM, strong signaling TCRs (dashed lines in Fig. 10B) are expected to be more potent than weak signaling TCRs due to the ability of CD8 to recruit the kinase Lck to the TCR. This leads to the concept of TCR activity, which denotes the overall intracellular signal intensity established in response to ligand encounter. The following

simple model for TCR activity is considered

$$y_{\text{ws}} \xleftrightarrow{\eta} y_{\text{ws}}^*, \tag{56a}$$

$$y_{\text{ss}} \xleftrightarrow{\psi\,\eta} y_{\text{ss}}^*. \tag{56b}$$

The asterisk denotes the active state of weak and strong signaling TCRs. The parameter η describes the signal intensity a weak signaling TCR establishes. The impact of CD8 on the signaling capacity of a TCR is simply modeled by an enhancement factor $\psi > 1$. The TCR activity in dependence on ligand concentration is now defined as

$$
\begin{aligned}
y_{\text{act}}(L, \psi) &:= y_{\text{ws}}^*(L) + y_{\text{ss}}^*(L) \\[2mm]
&= \eta \left[y_{\text{ws}}(L) + \psi\, y_{\text{ss}}(L) \right],
\end{aligned}
\tag{57}
$$

and describes the overall signal emanating from signaling TCRs. In principle, the parameter η can depend on ligand properties. For example, kinetic proofreading predicts a dependence on the rate of pMHC-TCR dissociation [49]. However, no such dependencies are considered in the remainder of this work. Therefore, η reduces to a simple scaling factor and is thus set to one, $\eta \equiv 1$. Note that this implies that the TCR activity in case of T1.CD8- cells is identical to the proportion of signaling TCRs.

From Fig. 10, it becomes apparent that strong signaling TCRs for T1.CD8+ cells have a peak at lower concentrations than signaling TCRs for T1.CD8- cells. Thus, for sufficiently large values of ψ, a given ligand can induce a larger TCR activity at smaller concentrations for T1.CD8+ cells than for T1.CD8- cells leading to an increased sensitivity for that ligand in the presence of CD8. Such an increase in TCR activity, as desired it might be for high affinity, i.e. foreign, ligands, could also lower self tolerance, because low affinity, i.e. self, ligands experience the same CD8-related enhancement effect. However, the peak value of the strong

signaling TCRs of the 4S ligand for T1.CD8+ cells is almost 3 orders of magnitude smaller than the peak value for T1.CD8- cells, while for the 4P ligand, it is only 1 order of magnitude. Hence, there is a window of values for ψ in which TCR activity is increased for high affinity ligands in the presence of CD8 while TCR activity remains suppressed for low affinity ligands. These findings are striking as they suggest that CBM exhibits the capability of sensing small amounts of foreign, i.e. high affinity, antigen more sensitively than CD8 negative T cells and additionally tolerating self, i.e. low affinity, peptides with higher specificity. This means that the usual trade-off between sensitivity and specificity, i.e. gaining one reduces the other, is not encountered.

To quantify this observation, two measures for a ligand's ability to activate T cells are introduced – termed potency and efficacy. In contrast to their usual definitions, efficacy y_{eff} is defined to be the maximal TCR activity a ligand can establish, and potency L_{pot} is defined as the concentration at which maximal TCR activity is reached,

$$y_{\text{eff}}(\rho_C, \psi) := \max_L y_{\text{act}}(L, \psi) \,. \tag{58a}$$

$$L_{\text{pot}}(\rho_C, \psi) := \arg\max_L y_{\text{act}}(L, \psi) \,. \tag{58b}$$

The rationale behind these definitions is the assumption that a minimal TCR activity is needed in order to trigger T cell activation. Expression (58a) determines whether this minimal TCR activity can be established by a given ligand that causes a TCR activity y_{act}, and expression (58b) gives the dose at which the ligand induces optimal TCR activity. The unit of the efficacy is such that it equals 1 if all TCRs on a T cell are weak signaling and ψ if all TCRs are strong signaling. The definition of the potency (58b) is such that it yields expression (53) if T cells lack CD8. Next to the dependence on the enhancement factor ψ, efficacy and potency are chosen to be further depend on the CD8 expression level ρ_C. This dependence
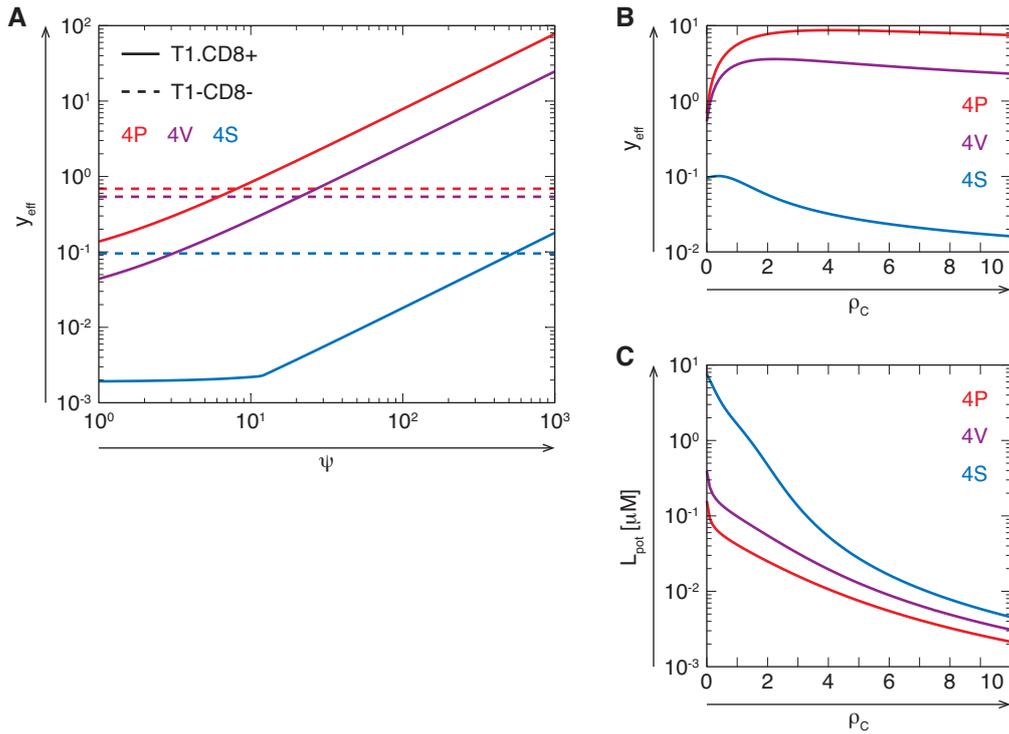
Figure 11: Efficacy and potency of CBM for soluble ligands

(A) Ligand efficacy in dependence on enhancement factor ψ for 4P, 4V, and 4S ligand as indicated by color. Dashed lines represent T1.CD8- cells and full curves T1.CD8+ cells having $\rho_C = 8.5$. (B) Ligand efficacy in dependence on the relative CD8 expression level ρ_C for 4P, 4V, and 4S ligand as indicated by color and $\psi = 100$. (C) Same as (B), but potency L_{pot} versus ρ_C is shown.

is inherited from the TCR activity as it depends on ρ_C via the weak and strong signaling TCRs.

The efficacy in dependence on enhancement factor ψ for 4P, 4V, and 4S ligand is shown in Fig. 11A. In the absence of CD8 (dashed lines), there is of course no dependence of the efficacy on ψ, and the efficacy assumes the peak value of signaling TCRs given in Fig. 10A of the corresponding ligand. For T1.CD8+ cells, having an CD8 expression level of $\rho_C = 8.5$, the efficacy increases almost linearly with ψ for 4P and 4V ligand. For the 4S ligand, a linear increase sets in at $\psi > 10$ while the efficacy stays almost constant below. The reason is that, for 4S ligand, the peak value of strong signaling TCRs is 1 order of magnitude smaller than the peak value of weak

signaling TCRs (see Fig. 10B). Thus, efficacy is almost exclusively given by weak signaling TCRs if $\psi < 10$, whereas for $\psi > 10$, strong signaling TCRs dominate efficacy leading to the linear increase. The value of ψ at which the efficacies of T1.CD8- and T1.CD8+ cells become equal depends on ligand affinity. It is given by 8, 21, and 530 for 4P, 4V, and 4S ligand, respectively. Hence, there is indeed a window of values for ψ in which high affinity ligands benefit from the CD8-related enhancement of TCR activity, while low affinity ligands fail to become more efficacious due to interactions of pMHC with the CD8 binding site of TCR-CD8 complexes.

Next, the impact of CD8 expression level is investigated. The enhancement factor is set to $\psi = 100$ yielding an increased efficacy for 4P and 4V ligand as well as lowered efficacy for 4S ligand at CD8 expression level $\rho_C = 8.5$ compared to $\rho_C = 0$. As shown in Fig. 11B, the beneficial effect of CD8 on 4P and 4V ligand sets in immediately and is most pronounced at $\rho_C \approx 4$ and $\rho_C \approx 2$ for 4P and 4V ligand, respectively. For larger values a slight decrease is found. For the 4S ligand, the efficacy remains almost constant until $\rho_C \approx 1$ and starts to decrease for larger values of ρ_C. Hence, for $\psi = 100$, CD8 exhibits the respective beneficial effects concerning efficacy on both, high and low affinity ligands, for CD8 expression levels $\rho_C > 1$.

Not only the amount of TCR activity is important but also which ligand dose elicits it. Therefore, the potency in dependence on CD8 expression level for $\psi = 100$ is considered. As shown in Fig. 11C, the potency is a decreasing function of ρ_C for all ligands. At $\rho_C = 0$, the potency is given by (53), and hence, is as distinct as the ligand affinities. This means that high affinity ligands are optimally sensed at low concentrations, whereas the potency of low affinity ligands is found at high concentrations. With increasing ρ_C, the potencies become more and more similar indicating that the maximal TCR activity is established at rather small concentration values for all ligands. As a consequence, smaller doses of high affinity ligands

4P and 4V can be sensed in the presence of CD8. This is of great relevance as foreign antigen is expected to be presented with minute amounts. The decrease of potency for the 4S ligand has also beneficial effect, because the rather large amounts with which low affinity self peptides are presented are suboptimal, which further improves self tolerance.

In summary, CBM possesses a TCR-CD8 complex that binds pMHC with high affinity only with its CD8 binding site independent of peptide quality, which is in competition with signal-inducing pMHC-TCR engagement. A further contribution of CD8 arises from the recruitment of kinase Lck leading to an increased TCR activity for pMHC-TCR-CD8 complexes in which pMHC is in direct contact with the TCR. Hence, CD8 plays a dual role in CBM. On the one hand, it prevents the formation of signaling competent TCRs, and on the other hand enables enhanced signaling once such signaling TCRs have formed. Whether a ligand benefits from signal enhancement or suffers from diminished pMHC-TCR binding depends on the ligand affinity. High affinity ligands are able to counteract the peptide-independent high affinity interaction leading to enhanced signaling in the presence of CD8. Furthermore, the optimal dose of T cell stimulation is also reduced with CD8 leading to increased sensitivity towards high affinity antigens. Low affinity ligands, on the other hand, fail to establish signaling competent TCRs as they cannot counteract the peptide independent high affinity interaction. Moreover, CD8 related enhancement effects are expected to deploy at small concentrations, and self peptides, usually of low affinity and presented at high doses, would hardly benefit from them. Thus, as self peptides are not supposed to trigger T cell activation, the contribution of CD8 leads to increased self tolerance. Taken together, the findings indicate that the presence of CD8 enhances sensitivity towards foreign antigens as well as self tolerance in response to self peptides assigning CD8 an important role in ligand discrimination and T cell activation.

5.3   Comparison to experiment

The analysis of the preceding section gives rise to four testable predictions. First, the dose response curves of T cell activation are bell-shaped. This is a consequence of the notion that TCR activity is triggered by multivalent binding rather than receptor occupancy. Second, the higher the ligand affinity the more signaling competent TCRs are formed, and the more TCR activity is established. This prediction results from the assumption of a uniform multivalent engagement parameter (see Sec. 3.2). Third, CD8 increases TCR activity only for higher affinity ligands. This reflects the impact of enhancement factor ψ on TCR activity. And fourth, for a given ligand, the potency is the lower the larger the CD8 expression level. Importantly, none of the stated predictions are specific for CBM. The first two predictions make general statements about multivalent binding, and how it is incorporated in the general modeling concept outlined in Sec. 3.2. The third prediction is also made by any model allowing for simultaneous binding of TCR and CD8 to the same pMHC, i.e. Models D, E, and F, because a similar model of TCR activity can be applied. And lastly, the dependence of the potency on ρ_C is also expected for these other models. Take Model D (see Chap. 3) as example. Next to binding of pMHC to TCR with concentration-rescaled affinity $K_1 L$, TCR can also bind CD8-bound pMHC with effective association constant $K_2' C \cdot K_1 L$. Thus, binding of pMHC to TCR is described by an apparent affinity $(1 + K_2' C) K_1 > K_1$. Following Eq. (53), a shift in potency towards smaller values in dependence on CD8 expression level is expected as well. This shows that the predictions make general statements about the underlying modeling concepts rather than a specific remark on CBM that no other model makes. One such CBM-specific prediction, for which, however, no data is at hand, is to prevent Lck to bind to CD8 and thereby extinguish the enhancement effect of CD8 on TCR activity. CBM predicts in this situation that even

high affinity ligands exhibit the lesser TCR activity, and hence the lesser ability to trigger T cell activation, the larger the CD8 expression level. As in no other model CD8 inhibits pMHC-TCR binding, this prediction is specific for CBM.

To test for the four stated predictions above, T cell activation induced by soluble tetramers was detected as changes in expression level of the T cell surface marker CD69 [78, 17]. All experiments, including the ones described further below, were performed by Sumit Deswal in the group of Prof. Dr. Wolfgang Schamel from the University of Freiburg, Germany. T1.CD8- and T1.CD8+ cells were exposed to different concentrations of 4P, 4V, and 4S ligand, and the CD69 expression level was measured after 12 hours. Assuming that CD69 up-regulation is a strictly increasing function of TCR activity, the concentration dependence of CD69 up-regulation follows the qualitative progression of TCR activity. Thus, observed peaks in CD69 up-regulation coincide with a ligand's potency. However, the experiments were performed at $37\,°C$; by contrast, the binding measurements were performed at $4\,°C$. This means that temperature dependencies of the model parameters might interfere with the predictions. In particular, pMHC-TCR binding affinities have been observed to decline with increasing temperature [54, 72]. As a consequence, T cell activation curves are expected to be shifted towards larger concentration values by an amount that depends on the CD8 expressing level (see Fig. 11C).

The results of the measurements of CD69 up-regulation as means of T cell activation are shown in Fig. 12A and B. The amount of detected CD69 follows a bell-shaped curve strongly pointing towards multivalent binding as origin of T cell activation. Further, a clear hierarchy with pMHC-TCR affinities is observed. Moreover, the 4P and 4V ligand are more efficacious on T1.CD8+ cells than on T1.CD8- cells. Also the 4V ligand induces more CD69 expression on T1.CD8+ cells than the 4P
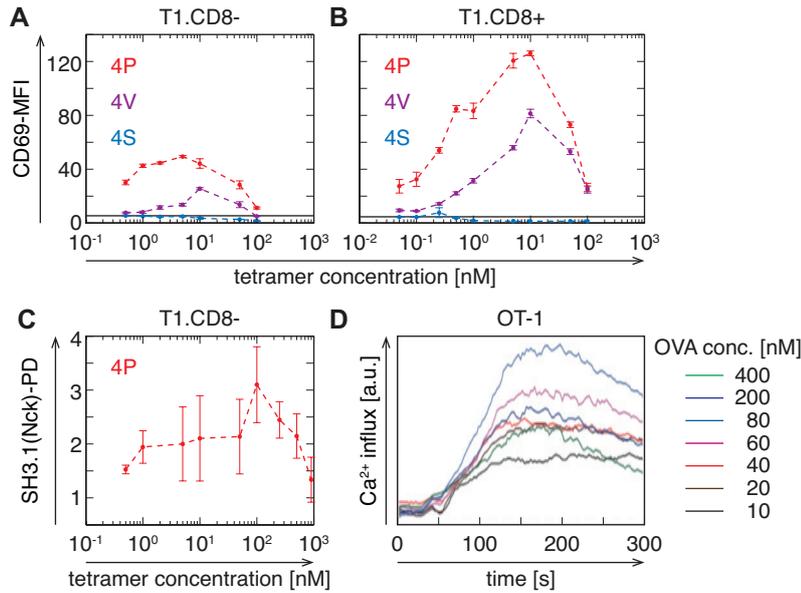
Figure 12: T cell activation data[†]

(A) CD69 up-regulation in dependence on tetramer concentration for 4P, 4V, and 4S ligands as indicated by color for T1.CD8- cells measured at 37 °C. The dots denote mean values, and the error bars standard errors of 3 experiment repetitions. The dashed lines are straight connectors of the mean values. The black line shows the background MFI in the absence of ligands. (B) Same as (A) but for T1.CD8+ cells having a CD8 expression level of $\rho_C = 8.5$. (C) Conformational change of the TCR determined by SH3.1 (Nck) pull down assay in response to 4P ligand for T1.CD8- cells measured at 4 °C. Dots represent means, and error bars standard errors of 3 experiment repetitions. The dashed line connects the mean values. (D) Calcium influx over time for OT-1 cells for various concentrations of OVA ligand as indicated measured at 37 °C.

[†] All experiments were performed by Sumit Deswal in the group of Prof. Dr. Wolfgang Schamel of the University of Freiburg, Germany.

ligand on T1.CD8- cells. The 4S ligand fails to trigger CD69 up-regulation on both cell types. This indicates that the value of $\psi = 100$ chosen in Fig. 11B and C is reasonable. Taken together, these observations confirm the first three predictions made above. However, a CD8 dependence of the potency is not found. The predicted potencies at 4 °C of 4P and 4V ligand for T1.CD8+ cells ($\rho_C = 8.5$) are 3 nM and 4 nM, respectively (see Fig. 11C). As the potency depends only weakly on affinity for high CD8 expression levels, the predicted potencies should also be valid at 37 °C. Indeed, the

peaks of CD69 up-regulation for T1.CD8+ cells are found at about 10 nM in good agreement with the predictions. The potencies at 4 °C of 4P and 4V ligand for T1.CD8- cells are predicted to be 160 nM and 390 nM, respectively. Without CD8, the potency depends on the affinity K_1 according to $\sim 1/K_1$. Hence, the predictions give a lower bound on the concentration value of the peak of CD69 up-regulation if affinities decline with temperature. However, the measured peak values are 5 nM and 10 nM for 4P and 4V ligand, respectively, exhibiting a poor agreement with the predictions. Indeed, no dependence of the potency on CD8 expression level was found in the experiment in contradiction with the fourth prediction.

To have a closer look at T cell activation of T1.CD8- cells, the conformational change of the TCR by a SH3.1 (Nck) pull down assay [31, 51] was measured. Importantly, these measurements were performed under the exact same conditions as the dose response binding data, and hence, no temperature effects are expected. The detected conformational change induced by 4P ligand is shown in Fig. 12C. The measured peak value, located at about 100 nM, is in good agreement with the prediction of 160 nM. However, similar data for the other ligands and T1.CD8+ cells are not available.

The finding of the measurement of TCR conformational change indicates that the potency might exhibit a dependence on CD8 expression level. To further consider this possibility, T cell activation by means of calcium influx for OT-1 cells in response to OVA ligand – the cognate ligand of OT-1 cells – has been recorded at 37 °C. The measured time courses for various concentrations of OVA ligand is shown in Fig. 12D. One finds that calcium influx is maximal at concentrations of about 80 nM. Importantly, primary OT-1 T cells have a lower CD8 expression level than T1.CD8+ cells ranging from 1 to 3 [72]. Using the 4P ligand affinity as surrogate for the OVA ligand affinity, the potency is expected to range from 15 nM to 40 nM. However, there is a certain affinity dependence of the potency

for the stated range of CD8 expression level 11C. Using the 4V affinity instead, to account for the temperature dependence of the affinity, yields a predicted range of the potency from $30\,\mathrm{nM}$ to $100\,\mathrm{nM}$ in agreement with the measurements. Thus, the data indicates that the potency of cognate ligand is largest for T1.CD8- cells (CD8 low, Fig. 12C), becomes smaller for OT-1 cells (CD8 mid, Fig. 12D), and assumes the lowest value for T1.CD8+ cells (CD8 high, Fig. 12B). Taken together, the findings provide evidence that the potency does indeed depend on the CD8 expression level in agreement with the fourth prediction. However, the issue that the peaks of CD69 up-regulation of T1.CD8- cells (Fig. 12A) are at such low concentrations remains unresolved.

6 Competition binding model and antigen presenting cells

CBM indicates that the interactions between pMHC, TCR, and CD8 enhance sensitivity for foreign antigen as well as self tolerance in response to self peptides. Thus, CD8 does not only facilitate T cell activation in response to an infection, but also contributes to self tolerance via the high affinity interaction between pMHC and the CD8 binding site of the TCR-CD8 complex. However, the findings so far were obtained for soluble ligands. Under natural conditions, T cells interact with APCs. For this reason, CBM is put in context of APCs via a model of pMHC presentation and infection of APCs that allows for bivalent engagement between pMHC and T cell surface receptors TCR and CD8.

6.1 Modeling of pMHC presentation

In Sec. 5.2, TCR activity was determined by the amount of multivalently bound TCRs in direct contact with pMHC. To keep this notion, the model of pMHC presentation assumes that pMHC can physically engage each other leading to the formation of pMHC clusters on the surface of APCs [25]. For the sake of simplicity, only single pMHC and bivalent clusters are considered. Furthermore, the presented pMHC can either arise from low affinity self peptides or from high affinity foreign antigen. Hence, five distinct ligand states can form on the surface of APCs (Fig. 13A): single self pMHC S, single foreign pMHC A, homo-dimerized self pMHC SS, homo-dimerized foreign pMHC AA, and hetero-dimerized self and foreign pMHC AS. The densities of these five ligand states are determined by the total density of pMHC on an APC surface P_{tot}, the degree of infection ω, and the association constant for cluster formation K_C. The degree of infection is defined as the fraction of pMHC on the APC surface stemming

69
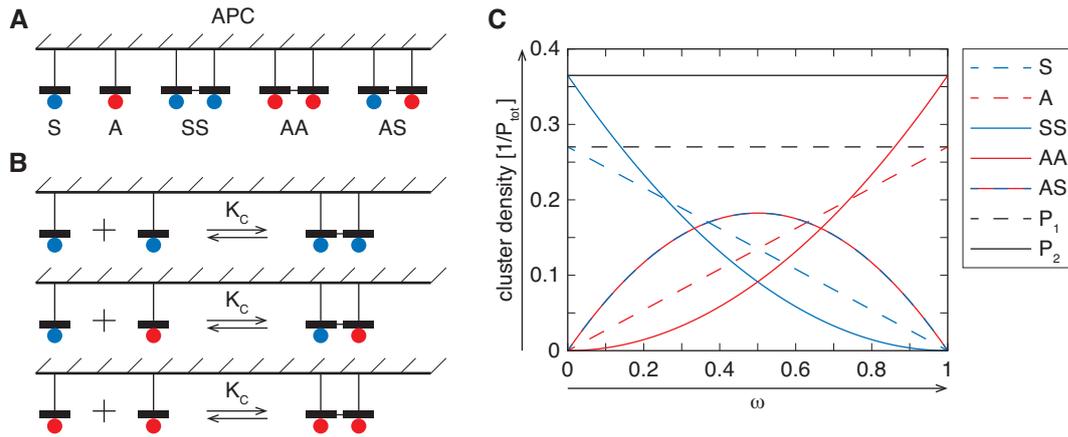
Figure 13: Model of pMHC presentation by APCs

(A) The five possible ligand states on infected APCs: single self pMHC (blue), single foreign pMHC (red), homo-dimerized self pMHC, homo-dimerized foreign pMHC, and hetero-dimerized self and foreign pMHC. (B) Dimerization of pMHC is determined by a single association constant K_C that cannot discriminate between self peptide and foreign antigen. (C) Cluster density in units of total pMHC density P_{tot} in dependence on degree of infection ω for the five ligand states as indicated. The dashed and full black lines represent the total densities of single and dimerized pMHC, which are independent of ω. For the association constant holds $K_C P_{\text{tot}} = 10$.

from foreign antigens, and the association constant does not distinguish between self peptides and foreign antigen (Fig. 13B). At steady state, the set of equations describing the ligand states on the APC is given by

$$\omega P_{\text{tot}} = A + 2 \cdot AA + AS \,, \tag{59a}$$

$$(1 - \omega) P_{\text{tot}} = S + 2 \cdot SS + AS \,, \tag{59b}$$

$$AA = \frac{1}{2} K_C A^2 \,, \tag{59c}$$

$$SS = \frac{1}{2} K_C S^2 \,, \tag{59d}$$

$$AS = K_C A \cdot S \,. \tag{59e}$$

The first two equations describe the conservation of foreign antigen respectively self peptides. The factor of 2 in front of the term AA in Eq. (59a),

respectively SS in Eq. (59b), arises from the number of foreign, respectively self, pMHC in the homo-dimerized complex. The next three equations are the relations between the densities of dimerized and single pMHC. The prefactor of $1/2$ in Eqs. (59c) and (59d) is due to the stoichiometry of homo-dimerization. Note that the sum of Eqs. (59a) and (59b) expresses the conservation of total pMHC. Furthermore, defining the total density of single pMHC by $P_1 \equiv A + S$ and the total density of dimerized pMHC by $P_2 \equiv AA + SS + AS$, one finds

$$P_{\text{tot}} = P_1 + 2\,P_2\,, \tag{60a}$$

$$P_2 = \frac{1}{2}\,K_C\,P_1^2\,. \tag{60b}$$

This means that the total densities of single pMHC and bivalent clusters are independent of the degree of infection. This is a reflection of K_C not depending on the nature of pMHC, and implies that APCs are not themselves able to distinguish self peptides from foreign antigen. Hence, in terms of cluster size distribution, APCs present an unbiased pool of pMHC, and ligand discrimination falls entirely in the realm of T cells.

The densities of the ligand states in dependence on the degree of infection are shown in Fig. 13C for an association constant $K_C\,P_{\text{tot}} = 10$. Varying the degree of infection from $\omega = 0$ to $\omega = 1$, the ligand states containing only self pMHC (S and SS) decline while the ligands states build up from foreign antigen (A and AA) increase continuously in density. The density of bivalent cluster containing one self and one foreign pMHC (AS) exhibits a peak at $\omega = 1/2$, because at this condition the individual constituents are present at equal amounts. As pointed out above, the total densities of single pMHC (P_1) and dimerized pMHC (P_2) are independent of the degree of infection.

6.2 Model parameters and rescaling

The model of pMHC presentation comprises of three parameters: the total pMHC density P_{tot}, the degree of infection ω, and the association constant for cluster formation K_C. Including CBM adds further eight parameters: the self pMHC-TCR affinity K_S, the foreign pMHC-TCR affinity K_A, the pMHC-CD8 affinity K', the association constant of TCR-CD8 complex formation Q_1, the effective association constant for the formation of aggregated state Ξ, the association constant of TCR binding to pMHC-bound CD8 Q_4, the total TCR density R_T, and the total CD8 density R_C. Importantly, the affinities are distinct from those used for soluble ligands as now the interaction between two cell surfaces is considered. Hence, the affinities have the unit of an inverse two-dimensional density, just as Q_1, Ξ, Q_4, and K_C. The TCR density is used as reference density leading to the following definitions of dimensionless model parameters,

$$\kappa_S \equiv R_T\,K_S\,, \tag{61a}$$

$$\kappa_A \equiv R_T\,K_A\,, \tag{61b}$$

$$\kappa' \equiv R_T\,K'\,, \tag{61c}$$

$$q_1 \equiv R_T\,Q_1\,, \tag{61d}$$

$$\xi \equiv R_T\,\Xi\,, \tag{61e}$$

$$q_4 \equiv R_T\,Q_4\,, \tag{61f}$$

$$k_C \equiv R_T\,K_C\,, \tag{61g}$$

$$\rho_C \equiv R_C/R_T\,, \tag{61h}$$

$$\rho_P \equiv P_{\text{tot}}/R_T\,. \tag{61i}$$

Note that the definitions of q_1, ξ, and q_4 match the ones in Sec. 3.3.

6.3   Binding polynomial

Having established a model for pMHC presentation by APCs, the interactions with the T cell surface receptors via CBM need to be determined. As in Sec. 3.4, this is carried out in the framework of binding polynomials. CBM allows for four different ways of how pMHC can be bound to T cell surface receptors (see Chap. 3). These are reflected in the scaffold expressions (24). Tallying these expressions, on the one hand for self peptide with low affinity κ_S and on the other hand for foreign antigen with high affinity κ_A, leads to auxiliary binding polynomials for self, respectively foreign, pMHC,

$$F_S(T, C) := \kappa_S \, T \left(1 + \xi \, C\right) + \kappa' \, C \left(1 + q_4 \, T\right), \tag{62a}$$

$$F_A(T, C) := \kappa_A \, T \left(1 + \xi \, C\right) + \kappa' \, C \left(1 + q_4 \, T\right). \tag{62b}$$

Using expressions (62), the complete binding polynomial of CBM in the context of APCs is given by

$$F(T, C, S, A) := T + C + q_1 \, T \, C + S \cdot \left[1 + F_S(T, C)\right] + A \cdot \left[1 + F_A(T, C)\right] +$$

$$\frac{1}{2} \, k_C \, S^2 \cdot \left[1 + F_S(T, C)\right]^2 + \frac{1}{2} \, k_C \, A^2 \cdot \left[1 + F_A(T, C)\right]^2 +$$

$$k_C \, A \cdot S \cdot \left[1 + F_S(T, C) + F_A(T, C) + F_S(T, C) \cdot F_A(T, C)\right]. \tag{63}$$

The first three terms describe unbound TCR T, unbound CD8 C, and the TCR-CD8 complex forming with association constant q_1. The subsequent terms denote the different ligand states and their interactions with T cell surface receptors (Fig. 14). First, single self pMHC S and its binding to T cell surface receptors $S \cdot F_S(T, C)$ is specified. This is followed by a similar term for single foreign pMHC. Next, homo-dimerized self pMHC is described. It can be unbound, SS (see Eq. (59d)), one of the two subunits
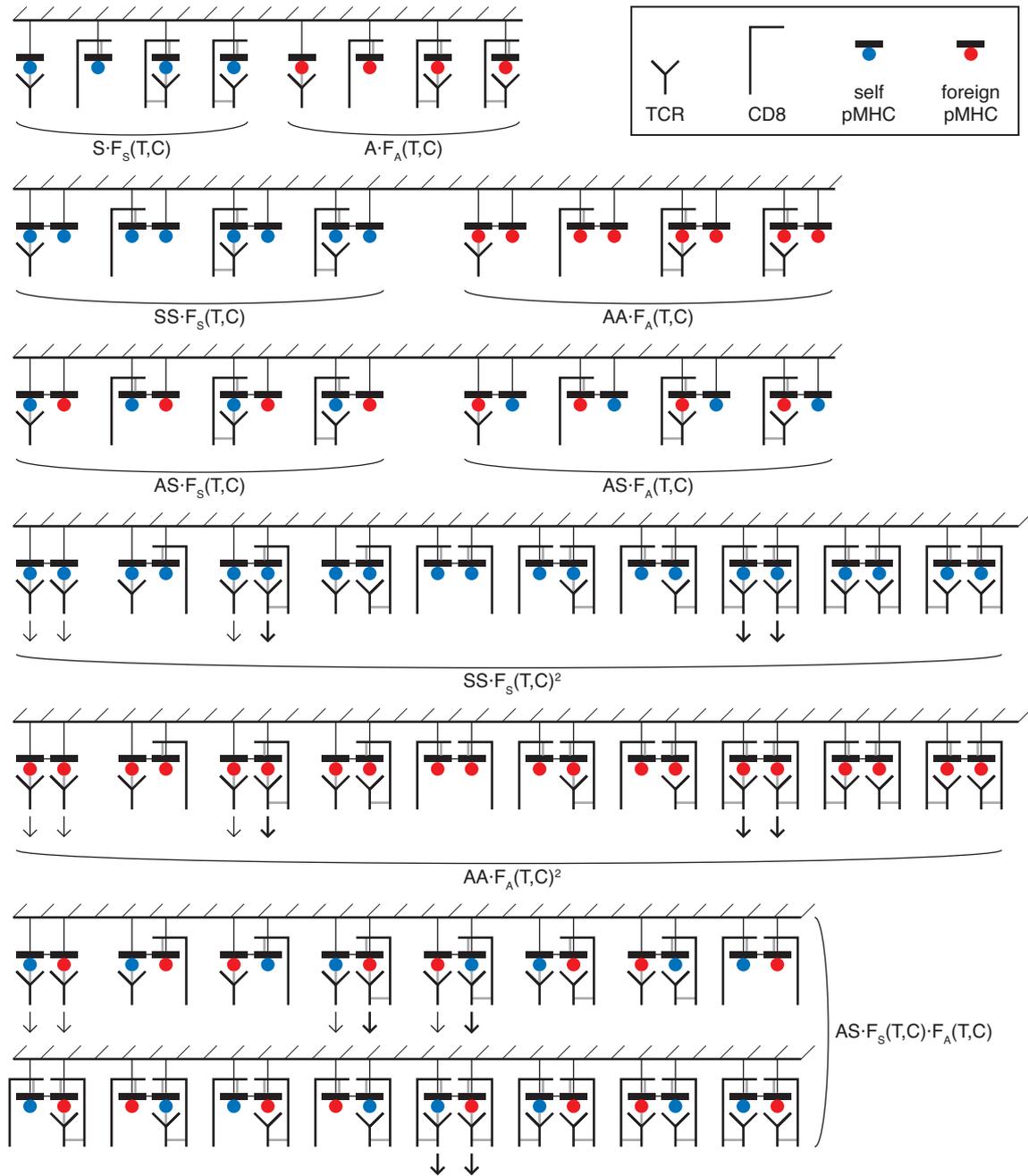
Figure 14: Binding states of CBM in context of APCs

In the first three rows, the four basic ligand-receptor binding states of CBM for single self and single foreign pMHC (first row), homo-dimerized self and foreign pMHC (second row), and hetero-dimerized pMHC (third row) are show. The remaining rows give the bivalent states of CBM with homo-dimerized self pMHC (fourth row), homo-dimerized foreign pMHC (fifth row), and hetero-dimerized pMHC (last two rows). The respective state densities are expressed in terms of the auxiliary binding polynomials (62) and the corresponding pMHC densities (see, Sec. 6.1) indicated by the parentheses.

can be bound, $2\,SS \cdot F_S(T,C)$, or both subunits are bound, $SS \cdot F_S(T,C)^2$. The same holds for homo-dimerized foreign pMHC, which is given thereafter. Finally, hetero-dimerized self and foreign pMHC is characterized. It can be unbound, AS (see Eq. (59e)), only the self pMHC subunit can be bound, $AS \cdot F_S(T,C)$, only the foreign pMHC subunit can be bound, $AS \cdot F_A(T,C)$, or both subunits made contact with T cell surface receptors, $AS \cdot F_S(T,C) \cdot F_A(T,C)$. Thus, expression (63) tallies all possible model states and is therefor the binding polynomial of CBM in the context of APCs.

For fixed parameters, the unknown variables are the densities of unbound TCR T, unbound CD8 C, unbound single self pMHC S, and unbound single foreign pMHC A. As outlined in Sec. 3.4, these variables are determined by solving a set of equations describing the conservation of the respective total densities. These equations are given by

$$1 = T\,\frac{\partial F(T,C,S,A)}{\partial T}\,, \tag{64a}$$

$$\rho_C = C\,\frac{\partial F(T,C,S,A)}{\partial C}\,, \tag{64b}$$

$$(1-\omega)\,\rho_P = S\,\frac{\partial F(T,C,S,A)}{\partial S}\,, \tag{64c}$$

$$\omega\,\rho_P = A\,\frac{\partial F(T,C,S,A)}{\partial A}\,. \tag{64d}$$

From top to bottom, these relations describe the conservation of total TCR, CD8, self pMHC, and antigenic pMHC. The solutions of Eqs. (64) are denoted with $\bar{T}$, $\bar{C}$, $\bar{S}$, and $\bar{A}$. Based on these, the density of each binding state can be calculated and consequently of each observable of interest.

6.4   T cell activation

The notion of T cell receptor activity is directly adapted from Sec. 5.2. There, only states with at least two TCRs in direct contact with pMHC

were assumed to be signaling states, and in these states only TCRs in direct contact with pMHC were taken to be signaling. Further, it was distinguished between strong and weak signaling TCRs depending on whether a CD8 molecule was bound to the signaling TCR or not. The same is assumed for CBM in context of APCs. In Fig. 14, weak and strong signaling TCRs are indicated by subjacent thin, respectively thick, arrows. The corresponding densities, y_{ws} and y_{ss}, are given by

$$y_{\mathrm{ws}}(\omega, \kappa_A, \rho_C) := k_C \left(1 + \xi \, \bar{C}\right) \left[\left(\kappa_S \, \bar{S} + \kappa_A \, \bar{A}\right) T\right]^2, \tag{65a}$$

$$y_{\mathrm{ss}}(\omega, \kappa_A, \rho_C) := \xi \, \bar{C} \cdot y_{\mathrm{ws}}(\omega, \kappa_A, \rho_C), \tag{65b}$$

and taken to be functions of the degree of infection ω, the affinity of foreign antigen κ_A, and the CD8 expression level ρ_C. The remaining parameters are fixed in the following . The parameters q_1, ξ, and q_4 are exactly defined as in the case of soluble ligands, and their values are thus assumed to be the same as in Chap. 5 (see the most right column of Tab. 5). The total pMHC density [20] and total TCR density [15] were estimated to be roughly the same. Hence, the total dimensionless total pMHC density is set to $\rho_P = 1$. No values of two-dimensional absolute affinities for pMHC-TCR or pMHC-CD8 interactions have been reported so far. Thus, these quantities have to be guessed. In order to be positively selected, TCRs must weakly interact with self pMHC [52]. Therefore, a small pMHC-TCR affinity is imposed for self peptides, $\kappa_S = 10^{-2}$. The pMHC-TCR affinities of cognate ligands are assumed to be 10^2, and antigen affinities are expected to lie in the range between self peptides and cognate ligands. The pMHC-CD8 affinity is assumed to have a moderate value. Hence, $\kappa' = 1$ is chosen. This yields an affinity of ≈ 40 for the CD8 binding site of a TCR-CD8 complex. Finally, the association constant for pMHC cluster formation on APCs is set to the same value as in Fig. 13C, i.e. $k_C = 10$.

The model of TCR activity is exactly the same as in Sec. 5.2, and thus given by the weighted sum of weak and strong signaling TCRs (see

Eq. (57) with $\eta \equiv 1$),

$$y_{\mathrm{act}}(\omega, \kappa_A, \rho_C, \psi) := y_{\mathrm{ws}}(\omega, \kappa_A, \rho_C) + \psi\, y_{\mathrm{ss}}(\omega, \kappa_A, \rho_C)\,. \qquad (66)$$

and is taken to be a function of the degree of infection ω, the affinity of foreign antigen κ_A, the CD8 expression level ρ_C, and the enhancement factor ψ describing the impact of CD8 on TCR signaling. In the following, it is investigated how different values of ρ_C and ψ influence self tolerance and recognition of foreign antigen of different affinities. As measure for self tolerance, the TCR activity that establishes if T cells interact with healthy cells is used. It is defined by

$$y_{\mathrm{act,self}}(\rho_C, \psi) := y_{\mathrm{act}}(0, 0, \rho_C, \psi)\,. \qquad (67a)$$

If $y_{\mathrm{act,self}}(\rho_C, \psi) < y_{\mathrm{act,self}}(\rho_C', \psi)$, it is said that the CD8 expression level ρ_C improves self tolerance compared to ρ_C'. Next, a measure for sensitivity towards foreign antigen is required. The basic idea is that a minimal TCR activity y_{act}^* needs to be established in order to trigger T cell activation. The smallest degree of infection ω^* for which y_{act}^* is reached shall serve as measure for sensitivity towards foreign antigen with affinity κ_A. It is given by

$$y_{\mathrm{act}}(\omega, \kappa_A, \rho_C, \psi) \overset{!}{=} y_{\mathrm{act}}^* \quad \Rightarrow \quad \omega = \omega^*(\kappa_A, \rho_C, \psi, y_{\mathrm{act}}^*)\,. \qquad (67b)$$

A T cell with CD8 expression level ρ_C is said to be the more sensitive towards an antigen of affinity κ_A the smaller $\omega^*(\kappa_A, \rho_C, \psi, y_{\mathrm{act}}^*)$.

Using the parameter values specified above, a healthy APC establishes a TCR activity of 7×10^{-5} in the absence of CD8. To assure that this activity is well below threshold, a minimal TCR activity of $y_{\mathrm{act}}^* = 10^{-3}$ is used in the following. The TCR activity of healthy cells in dependence on CD8 expression level ρ_C and for various values of enhancement factor ψ, ranging from 50 to 1000, is shown in Fig. 15A. The dashed lines indicate the minimal TCR activity y_{act}^*. For all values of ρ_C and ψ, the amount
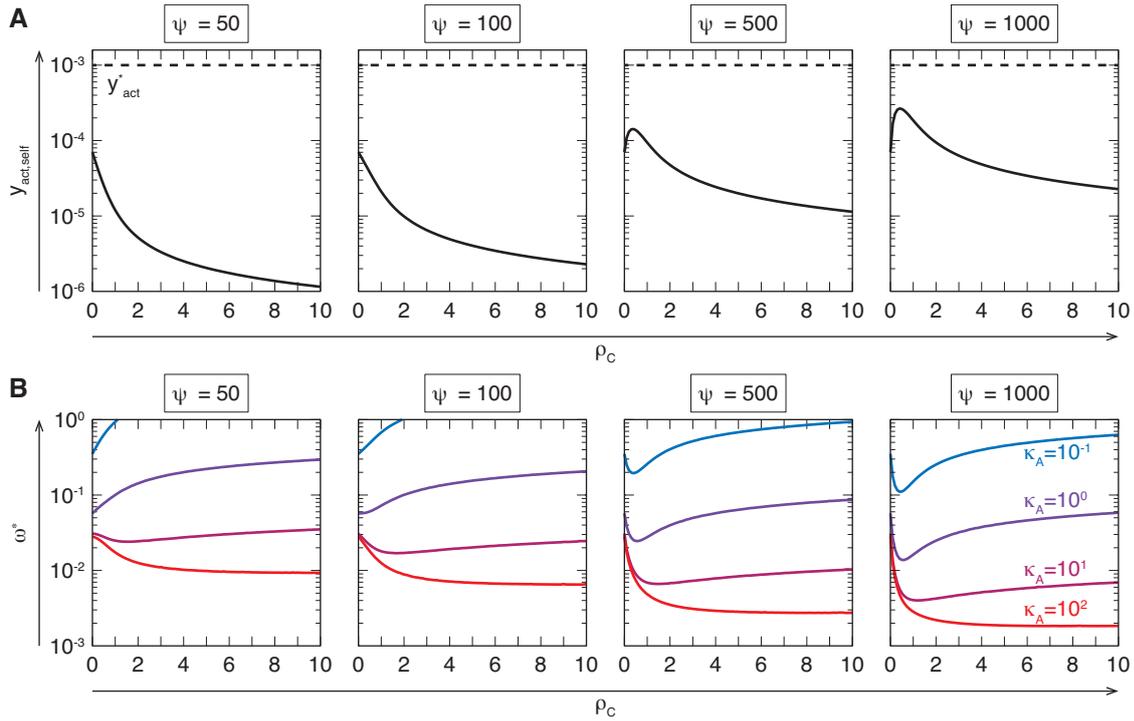
Figure 15: Self tolerance and sensitivity for CBM in context of APCs

(A) Full curves show TCR activity of healthy APCs ($\omega = 0$) in dependence on the CD8 expression level ρ_C for various values of ψ. The dashed lines indicate the minimal TCR activity needed to trigger T cell activation, $y_{\text{act}}^* = 10^{-3}$. The remaining parameters are specified in the main text on p. 76. (B) The smallest detectable degree of infection ω^* in dependence on the CD8 expression level ρ_C for various values of ψ. Different antigens are given by different affinities κ_A and are indicated by color. The remaining parameters are as in (A).

of TCR activity established is insufficient to trigger T cell activation. For $\psi = 50$ and $\psi = 100$, self tolerance continuously improves with increasing ρ_C approaching about 10^{-6} ($\psi = 50$), respectively 2×10^{-6} ($\psi = 100$), as $\rho_C \to 10$. However, for $\psi = 500$ and $\psi = 1000$, the TCR activity initially increases before a decline sets in. The TCR activity found for $\rho_C = 0$ is assumed again at CD8 expression level $\rho_C = 1.4$ ($\psi = 500$), respectively $\rho_C = 2.7$ ($\psi = 1000$). As $\rho_C \to 10$, the TCR activity approaches 10^{-5} ($\psi = 500$), respectively 2×10^{-5} ($\psi = 1000$). Thus, self tolerance is in general improved by CD8, showing that this aspect of the findings in Chap. 5 is recovered in the context of APCs.

The smallest detectable degree of infection ω^* in dependence on CD8 expression level ρ_C and for various values of enhancement factor ψ, ranging from 50 to 1000, and different antigen affinities between 10^{-1} and 10^2 is shown in Fig. 15B. As specified above, the affinity of self peptides is $\kappa_S = 10^{-2}$. In the absence of CD8, the different affinity antigens yield values for ω^* of 0.35 ($\kappa_A = 10^{-1}$), 0.057 ($\kappa_A = 10^0$), 0.031 ($\kappa_A = 10^1$), and 0.028 ($\kappa_A = 10^2$). The lowest affinity antigen ($\kappa_A = 10^{-1}$) cannot be detected for $\rho_C > 1.1$ if $\psi = 50$, respectively $\rho_C > 2$ if $\psi = 100$, because the T cell fails to establish the minimal required TCR activity due to the competitive effect of CD8. For $\psi = 500$ and $\psi = 1000$, ω^* initially improves to 0.19 at $\rho_C = 0.4$ ($\psi = 500$), respectively 0.11 at $\rho_C = 0.4$ ($\psi = 1000$), after which an increase sets in, intersecting the values of $\rho_C = 0$ at $\rho_C = 1.6$ ($\psi = 500$) and $\rho_C = 3.2$ ($\psi = 1000$), to finally approach 0.94 ($\psi = 500$), respectively 0.63 ($\psi = 1000$), as $\rho_C \to 10$. Thus, for $\kappa_A = 10^{-1}$, improved sensitivity is encountered only for $\psi = 500$ and $\psi = 1000$ if the CD8 expression level is not too high.

For antigens with affinity $\kappa = 10^0$, the smallest detectable degree of infection continuously increases with ρ_C for $\psi = 50$, and reaches a value of 0.3 as $\rho_C \to 10$. For the larger values of ψ, ω^* initially declines before an increase sets in. While only a little bump is found for $\psi = 100$, the minimal value of ω^* is given by 0.025 at $\rho_C = 0.5$ for $\psi = 500$, respectively 0.014 at $\rho_C = 0.6$ for $\psi = 1000$. The value at $\rho_C = 0$ is assumed again at $\rho_C = 3.9$ ($\psi = 500$) and at $\rho_C = 9.5$ ($\psi = 1000$), respectively. For $\rho_C \to 10$, ω^* is given by 0.20 ($\psi = 100$), 0.087 ($\psi = 500$), and 0.058 ($\psi = 1000$). As for $\kappa_A = 10^{-1}$, only for $\psi = 500$ and $\psi = 1000$ improved sensitivity is found, though on a broader range of CD8 expression levels.

For antigen affinity of $\kappa_A = 10^1$, an initial decline of ω^*, followed by an increase, is found for all values of ψ. The minimal values are 0.024 at $\rho_C = 1.5$ for $\psi = 50$, 0.017 at $\rho_C = 1.7$ for $\psi = 100$, 0.007 at $\rho_C = 1.5$

for $\psi = 500$, and 0.004 at $\rho_C = 1.2$ for $\psi = 1000$. As $\rho_C \to 10$, the smallest detectable degrees of infection approach 0.035 for $\psi = 50$, 0.024 for $\psi = 100$, 0.010 for $\psi = 500$, and 0.007 for $\psi = 1000$. Thus, only for $\psi = 50$, ω^* rises above the value at $\rho_C = 0$, which happens at $\rho_C = 6.4$, and sensitivity is only improved if ρ_C becomes not too large. For the other values of ψ, increased sensitivity due to CD8 is found for all CD8 expression levels considered, with minimal values are found for ρ_C between 1 and 2.

Finally, antigens with affinity $\kappa_A = 10^2$ cause a continuous decline in ω^* with increasing ρ_C for all values of ψ. The values ω^* is approaching as $\rho_C \to 10$ are given by 0.009 ($\psi = 50$), 0.007 ($\psi = 100$), 0.003 ($\psi = 500$), and 0.002 ($\psi = 1000$), respectively. Though, most of the decline is found for CD8 expression levels $\rho_C < 3$ after which ω^* remains almost constant.

The preceding results indicate that, in the presence of CD8, the highest affinity ligands can be sensed if their proportion in the pMHC pool is only a few per mille. Experimentally, it was found that a 1 to 10 cognate antigens are sufficient to trigger a T cell [38]. To translate the computed proportions into the number of cognate pMHC that can be detected, the following considerations shall serve as estimate. The part of the T cell surface in contact with an APC was specified to be about 5 % [33]. Further, the total number of TCRs on the T cell surface is approximately 20 000 (Tab. 2). Thus, the number of TCRs in the contact zone is roughly 1000. Assuming the contact area of the APC to be of the same size as the contact area of the T cell, the total number of pMHC with which the TCRs interact is also roughly 1000, due to the similarity of the TCR and pMHC densities [15, 20]. Hence, detection of foreign pMHC with a proportion of only a few per mille in an otherwise vast pool of self peptides corresponds to sensing a number of order 1 foreign antigens.

The obtained levels of CD8 expression for which improved self tolerance and improved sensitivity are found are summarized in Fig. 16. The
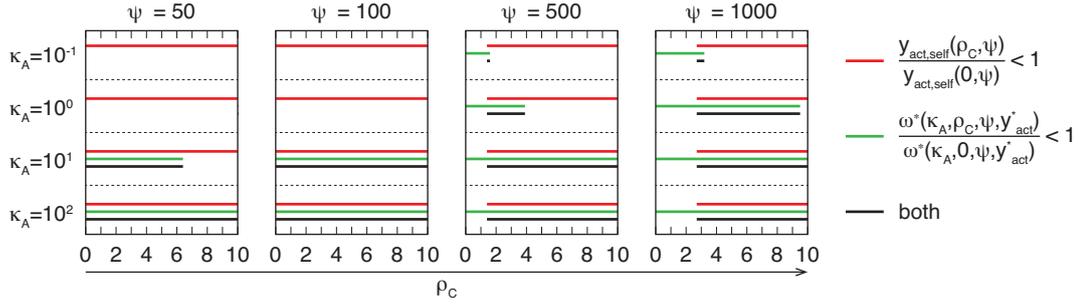
Figure 16: Impact of CD8 on self tolerance and sensitivity

CD8 expression levels ρ_C for which self tolerance and sensitivity are improved compared to the absence of CD8 for various antigen affinities κ_A and enhancement factors ψ. Red lines show regions of ρ_C with improved self tolerance, defined by $y_{\text{act,self}}(\rho_C, \psi) < y_{\text{act,self}}(0, \psi)$, green lines regions with improved sensitivity, defined by $\omega^*(\kappa_A, \rho_C, \psi, y^*_{\text{act}}) < \omega^*(\kappa_A, 0, \psi, y^*_{\text{act}})$, and black lines regions where both quantities are improved. The TCR activity threshold is set to $y^*_{\text{act}} = 10^{-3}$. The remaining parameters are specified in the main text on p. 76.

red lines indicate CD8 expression levels with improved self tolerance compared to the absence of CD8, i.e. $y_{\text{act,self}}(\rho_C, \psi) < y_{\text{act,self}}(0, \psi)$. The green lines show the same but for improved sensitivity, i.e. $\omega^*(\kappa_A, \rho_C, \psi, y^*_{\text{act}}) < \omega^*(\kappa_A, 0, \psi, y^*_{\text{act}})$. As expected, the two effects are complementary. But there are expression levels of CD8 for which both beneficial effects are established simultaneously (black lines). This includes the value $\psi = 100$ used in Chap. 5. However, the sensitivity of the highest affinity ligand drops hardly below the percent limit (Fig. 15B), which is not in good agreement with requirement to sense an order 1 number of foreign antigens. For $\psi = 500$, on the other hand, the sensitivity requirements are fulfilled (Fig. 15B). Further, the finding in Sec. 5.3 that the 4S ligand does not trigger CD69 up-regulation in the presence of CD8 is also in agreement with this value. Thus, simultaneous improvements of self tolerance and sensitivity are recovered in the context of APCs.

7 Discussion and Conclusions

The focus of this work is on the role of the co-receptor CD8 in ligand discrimination and T cell activation. In particular, the interplay of TCR and CD8 with pMHC was scrutinized by developing a family of mathematical binding models, and the contribution of CD8 on intracellular signaling was investigated. Based on findings reported in the literature, a family of pMHC-TCR-CD8 binding models was developed that systematically incorporates CD8-related interactions. Each probed pMHC-TCR-CD8 interaction model was confronted with data from accurately measured dose response assays, which were conducted by Sumit Deswal at the University of Freiburg, Germany, in the group of Prof. Dr. Wolfgang Schamel. Although T cells react with APCs under physiological conditions, soluble pMHC tetramers were used as ligands. This allowed high control over the dose to which the T cells were exposed. Further aspects of data acquisition, like prevention of receptor internalization, were readily taken into account by the modeling approach. The data was gathered for a wide range of experimental conditions, i.e. different cell and ligand types as well as various ligand doses, allowing for the identification of a best-fitting model as well as the corresponding model parameters.

The best-fitting model, termed Competition Binding Model (CBM, see Fig. 2G), comprises peptide-specific binding between pMHC and TCR, peptide-unspecific binding between pMHC and CD8, simultaneous binding of TCR and CD8 to the same pMHC, as well as an interaction between the two types of T cell surface receptors yielding a TCR-CD8 complex possessing two binding sites – one from the TCR and the other one from CD8, making the latter peptide-unspecific. The most notable finding revealed in the course of parameter estimation was that the CD8 binding site of the TCR-CD8 complex exhibits a much larger affinity for MHC than CD8 alone (about 40-fold for parameters θ^* in Tab. 5). As pMHC, if it is only bound

83

to the CD8 binding site of a TCR-CD8 complex, is not in direct contact with the TCR, no proximal TCR signaling (e.g. ITAM phosphorylation) is expected. However, because this interaction involves a TCR and is of high affinity, it is in direct competition with productive, i.e. potentially signal-inducing, binding of pMHC to TCR.

The notion that pMHC can be simultaneously bound by a TCR and a CD8 has been proposed before [30, 86]. Multiple evidence also suggests that TCR and CD8 can engage in some way [77, 29, 5]. One possible source of such an interaction is the α-chain connecting peptide motif found on the TCR [55, 48]. A further study proposed that engagement of TCR and CD8 is mediated by the kinase Lck [41]. Further, co-localization of TCR and CD8 in the absence of ligands has also been reported [28]. Hence, the way TCR-CD8 interaction has been included into CBM is in agreement with previous findings. However, a TCR-related increase of pMHC-CD8 affinity has not yet been claimed. In this respect, CBM is a hypothesis requiring further independent proof by experiments.

Based on a simple model of TCR triggering, CBM is capable to qualitatively predict the up-regulation of key markers of T cell activation. The data was provided by our collaboration partners Sumit Deswal and Prof. Dr. Wolfgang Schamel from the University of Freiburg, Germany. The TCR triggering model builds up on the observation that soluble monomeric pMHC is insufficient to induce T cell responses [6, 74] considering only those TCRs as signaling competent that are multivalently bound and in direct contact with pMHC (see Fig. 4). It further connects the association constant of multivalent binding to the pMHC-TCR affinity enabling affinity-based ligand discrimination. The TCR triggering model is related to the permissive geometry model [50], which states that only multivalently bound TCRs undergo a conformational change making them susceptible to further biochemical modifications (e.g. phosphorylations) and hence initiators of

intracellular signaling [70]. The dynamics of these modifications were not outlined in any detail in this work. Instead, it was simply assumed that the signal intensity emanated by a conformationally changed TCR does not depend on any ligand property, and was thus set to 1. Hence, it is only the pMHC-TCR affinity determining whether a ligand is capable of inducing a sufficient intracellular signal intensity to launch a T cell response. However, extensions to this simplified point of view are possible. For example, kinetic proofreading predicts that the signal intensity also depends on the ligand's rate of dissociation from the TCR and the number of modification steps necessary to activate the TCR [49]. Such further dependencies of the intracellular signal can not be excluded, but there is no clear evidence in the activation data used in this work pointing towards additional mechanisms of signal assembly.

The impact of co-receptor CD8 on intracellular signaling was incorporated in the simplest conceivable way. If CD8 is bound to a conformationally changed TCR, the emanated signal intensity is increased by a factor $\psi > 1$. This assumption is justified by the observation that CD8 binds the kinase Lck [83, 79], and thereby recruits the most critical kinase of ITAM phosphorylation to the TCR [7]. This enhancement of signal generation, together with the property of CBM to provide a high affinity competitor for productive pMHC-TCR contacts, introduces a novel mechanism for the role of CD8 in ligand discrimination and T cell activation that improves self tolerance as well as sensitivity towards foreign antigen of T cells. Low affinity (self) ligands fail to overcome the obstacle provided by the high affinity CD8 binding site of the TCR-CD8 complex leading to a diminished amount of conformationally changed TCRs. Thus, a reduced intracellular signal intensity in the presence of CD8 is found in response to self peptides. High affinity (foreign) ligands counteract this interaction by forming pMHC-TCR bonds with TCR-CD8 complexes and thereby benefiting from the CD8-related enhancement of signal generation. This induces increased

intracellular signal intensity in the presence of CD8, and thus facilitates T cell activation in response to foreign antigen. Importantly, an enhancing impact of CD8 on T cell activation has been observed before [36, 46] and could also be confirmed experimentally in this work.

The notion that multivalent binding is necessary to induce a cellular response is not limited to T cells. In fact, this requirement is found for many transmembrane receptors, for example epidermal growth factor receptors [68], platelet derived growth factor receptors [34], the high affinity receptor for IgE (FcεRI) [39, 37], B cell receptors [9, 60], and ephrin receptors [71], showing that ligand discrimination based on multivalent binding is a very common mechanism in biology. Thus, the findings in this work might also be of relevance for other systems than T cells.

A further important aspect underlying this work is the consistency of the experimental approach. In many studies, affinity measurements were conducted in one system, e.g. by surface plasmon resonance or pMHC tetramer binding, and were then correlated to T cell activation measured in another one, like T cell stimulation by APCs [42, 36, 3, 46, 12, 2, 22]. By contrast, the relation between ligand affinity and T cell activation was established for the same system in this work enabling the formulation of a quantitative model on the interactions of pMHC, TCR, and CD8 for the first time.

The successful application of CBM to soluble ligands motivated the investigation in the context of APCs. Based on a model of pMHC presentation that includes ligand clustering on the surface of APCs [25], and using the same TCR triggering model as for soluble ligands, simultaneous improvements of self tolerance and antigen sensitivity were recovered in this system. Furthermore, the degree of infection of cognate ligands that can be sensed by a T cell was found to be as low as a few per mille. This small proportion translates into a number of order 1. Strikingly, T cell responses

to such small numbers of cognate ligands embedded in vast number of self peptides have been observed experimentally [38]. Thus, CBM is capable of explaining this finding, and importantly, no sacrifices of self tolerance have to be made.

In conclusion, the statistically rigorous confrontation of a family of pMHC-TCR-CD8 binding models with accurate dose response data resulted in the identification of an interaction model that, in combination with a multivalent TCR triggering model, suggests a novel mechanism to amplify affinity-based ligand discrimination, in which a dual role is attached to CD8. On the one hand, CD8 hampers the formation of productive pMHC-TCR contacts, and on the other hand enhances intracellular signaling once such productive contacts have been made. The crux of the proposed mechanism is that the decision which of the two opposing effects comes into effect is solely determined by the pMHC-TCR affinity. This leads to improved self tolerance in response to self peptides (low affinity) and to increased sensitivity towards foreign antigen (high affinity). Hence, the mechanism by which CD8 contributes to ligand discrimination and T cell activation was refined in this work pointing to an important role of CD8 in T cell immunity.

References

[1] J. Aldrich. R. A. Fisher and the making of maximum likelihood 1912-1922. *Statistical Science*, 12(3):162–176, 1997.

[2] M. Aleksic, O. Dushek, H. Zhang, E. Shenderov, J. L. Chen, V. Cerundolo, D. Coombs, and P. A. van der Merwe. Dependence of T cell antigen recognition on T cell receptor-peptide MHC confinement time. *Immunity*, 32(2):163–174, 2010.

[3] G. Altan-Bonnet and R. N. Germain. Modeling T cell antigen discrimination based on feedback control of digital ERK responses. *PLoS Biology*, 3(11):1925–1938, 2005.

[4] J. D. Altman, P. A. Moss, P. J. Goulder, D. H. Barouch, M. G. McHeyzer-Williams, J. I. Bell, A. J. McMichael, and M. M. Davis. Phenotypic analysis of antigen-specific T lymphocytes. *Science*, 274(5284):94–96, 1996.

[5] A. D. Beyers, L. L. Spruyt, and A. F. Williams. Molecular associations between the T-lymphocyte antigen receptor complex and the surface antigens CD2, CD4, or CD8 and CD5. *Proceedings of the National Academy of Sciences of the United States of America*, 89(7):2945–2949, 1992.

[6] J. J. Boniface, J.D. Rabinowitz, C. Wülfing, J. Hampl, Z. Reich, J. D. Altman, C. Kantor, C. Beeson, H. M. McConnell, and M. M. Davis. Initiation of signal transduction through the T cell receptor requires the multivalent engagement of peptide/MHC ligands. *Immunity*, 9(4):459–466, 1998.

[7] J. M. Bradshaw. The Src, Syk and Tec family kinases: distinct types of molecular switches. *Cellular Signalling*, 22(8):1175–1184, 2010.

[8] K. P. Burnham and D. R. Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* Springer Verlag, New York, 2002.

[9] J. C. Cambier and J. T. Ransom. Molecular mechanisms of transmembrane signaling in B lymphocytes. *Annual Review of Immunology*, 5:175–199, 1987.

[10] A. C. Chan, B. A. Irving, J. D. Fraser, and A. Weiss. The ζ chain is associated with a tyrosine kinase and upon T-cell antigen receptor stimulation associates with ZAP-70, a 70-kDa tyrosine phosphoprotein. *Proceedings of the National Academy of Sciences of the United States of America*, 88(20):9166–9170, 1991.

[11] C. Chan, A. J. T. George, and J. Stark. T cell sensitivity and specificity – kinetic proofreading revisited. *Discrete and Continuous Dynamical Systems – Series B*, 3(3):343–360, 2003.

[12] A. S. Chervin, J. D. Stone, P. D. Holler, A. Bai, J. Chen, N. H. Eisen, and D. M. Kranz. The impact of TCR-binding properties and antigen presentation format on T cell responsiveness. *Journal of Immunology*, 183(2):1166–1178, 2009.

[13] B. K. Cho, K. C. Lian, P. Lee, A. Brunmark, C. McKinley, J. Chen, D. M. Kranz, and H. N. Eisen. Differences in antigen recognition and cytolytic activity of $CD8^+$ and $CD8^-$ T cells that express the same antigen-specific receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 98(4):1723–1727, 2001.

[14] D. K. Cole, B. Laugel, M. Clement, D. K. Price, L. Wooldridge, and A. K. Sewell. The molecular determinants of CD8 co-receptor function. *Immunology*, 137(2):139–148, 2012.

[15] D. Coombs, A. M. Kalergis, S. G. Nathenson, C. Wofsy, and B. Goldstein. Activated TCRs remain marked for internalization after dissociation from pMHC. *Nature Immunology*, 3(10):926–931, 2002.

[16] M. A. Daniels and S. C. Jameson. Critical role for CD8 in T cell receptor binding and activation by peptide/major histocompatibility complex multimers. *The Journal of Experimental Medicine*, 191(2):335–346, 2000.

[17] M. A. Daniels, E. Teixeiro, B. Hausmann, D. Roubaty, K. Holmbeg, G. Weren, G. A. Holländer, N. R. Gascoigne, and E. Palmer. Thymic selection threshold defined by compartmentalization of Ras/MAPK signalling. *Nature*, 444(7120):724–729, 2006.

[18] M. M. Davis, J. J. Boniface, Z. Reich, D. Lyons, J. Hampl, B. Arden, and Y. Chien. Ligand recognition by $\alpha\beta$ T cell receptors. *Annual Review of Immunology*, 16:523–544, 1998.

[19] S. J. Davis and P. A. van der Merwe. The kinetic-segregation model: TCR triggering and beyond. *Nature Immunology*, 7(8):803–809, 2006.

[20] J. Deeg, M. Axmann, J. Matic, A. Liapis, D. Depoil, J. Afrose, D. Curado, M. L. Dustin, and J. P. Spatz. T cell activation is determined by the number of presented antigens. *Nano Letters*, 13(11):5619–5626, 2013.

[21] K. Dill and S. Bromberg. *Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience*. Garland Science, New York, 2010.

[22] O. Dushek, M. Aleksic, R. J. Wheeler, H. Zhang, S. P. Cordoba, Y. C. Peng, J. L. Chen, V. Cerundolo, T. Dong, D. Coombs, and P. A. van der Merwe. Antigen potency and maximal efficacy reveal a mechanism of efficient T cell activation. *Science Signaling*, 4(176):ra39, 2011.

[23] T. M. Fahmy, J. G. Bieler, M. Edidin, and J. P. Schneck. Increased TCR avidity after T cell activation: a mechanism for sensing low-density antigen. *Immunity*, 14(2):135–143, 2001.

[24] N. J. Felix, Donermeyer D. L., S. Horvath, J. J. Walters, M. L. Gross, A. Suri, and P. M. Allen. Alloreactive T cells respond specifically to multiple distinct peptide-MHC complexes. *Nature Immunology*, 8(4):388–397, 2007.

[25] M. Ferez, M. Castro, B. Alarcon, and H. M. van Santen. Cognate peptide-MHC complexes are expressed as tightly apposed nanoclusters in virus-infected cells to allow TCR crosslinking. *Journal of Immunology*, 192(1):52–58, 2014.

[26] R. A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A*, 222:309–368, 1922.

[27] P. François, G. Voisinne, E. D. Siggia, G. Altan-Bonnet, and M. Vergassola. Phenotypic model for early T-cell activation displaying sensitivity, specificity, and antagonism. *Proceedings of the National Academy of Sciences of the United States of America*, 110(10):E888–E897, 2013.

[28] D. M. Gakamsky, I. F. Luesher, A. Pramanik, R. B. Kopito, F. Lemonnier, H. Vogel, R. Rigler, and I. Pecht. CD8 kinetically promotes ligand binding to the T-cell antigen receptor. *Biophysical Journal*, 89(3):2121–2133, 2005.

[29] P. F. Gallagher, B. Fazekas de St Groth, and J. F. Miller. CD4 and CD8 molecules can physically associate with the same T-cell receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 86(24):10044–10048, 1989.

[30] G. F. Gao, Z. Rao, and J. I. Bell. Molecular coordination of $\alpha\beta$ T-cell receptors and coreceptors CD8 and CD4 in their recognition of peptide-MHC ligands. *Trends in Immunology*, 23(8):408–413, 2002.

[31] D. Gil, W. W. Schamel, M. Montoya, F. Sanchez-Madrid, and B. Alarcon. Recruitment of Nck by CD3ϵ reveals a ligand-induced conformational change essential for T cell receptor signaling and synapse formation. *Cell*, 109(7):901–912, 2002.

[32] D. Gil, A. G. Schrum, B. Alarcon, and E. Palmer. T cell receptor engagement by peptide-MHC ligands induces a conformational change in the CD3 complex of thymocytes. *The Journal of Experimental Medicine*, 201(4):517–522, 2005.

[33] A. Grakoui, S. K. Bromley, C. Sumen, M. M. Davis, A. S. Shaw, P. M. Allen, and M. L. Dustin. The immunological synapse: a molecular machine controlling T cell activation. *Science*, 285(5425):221–227, 1999.

[34] C. H. Heldin, A. Ernlund, Rorsman C., and L. Rönnstrand. Dimerization of B-type platelet-derived growth factor receptors occurs after ligand binding and is closely associated with receptor kinase activation. *The Journal of Biological Chemistry*, 264(15):8905–8912, 1989.

[35] W. S. Hlavacek, A. S. Perelson, B. Sulzer, J. Bold, J. Paar, W. Gorman, and R. G. Posner. Quantifying aggregation of IgE-FcεRI by multivalent antigen. *Biophysical Journal*, 76(5):2421–2431, 1999.

[36] P. D. Holler and D. M. Kranz. Quantitative analysis of the contribution of TCR/pepMHC affinity and CD8 to T cell activation. *Immunity*, 18(2):255–264, 2003.

[37] D. Holowka and B. Baird. Antigen-mediated IGE receptor aggregation and signaling: a window on cell surface structure and dynamics. *Annual Review of Biophysics and Biomolecular Structure*, 25:79–112, 1996.

[38] D. J. Irvine, M. A. Purbhoo, M. Kroosgaard, and M. M. Davis. Direct observation of ligand recognition by T cells. *Nature*, 419(6909):845–849, 2002.

[39] Erickson J., P. Kane, B. Goldstein, D. Holowka, and B. Baird. Cross-linking of IgE-receptor complexes at the cell surface: a fluorescence method for studying the binding of monovalent and bivalent haptens to IgE. *Molecular Immunology*, 23(7):769–781, 1986.

[40] C. A. Janeway Jr. The T cell receptor as a multicomponent signalling machinery: CD4/CD8 coreceptors and CD45 in T cell activation. *Annual Review of Immunology*, 10:645–674, 1992.

[41] N. Jiang, J. Huang, L. J. Edwards, B. Liu, Y. Zhang, C. D. Beal, B. D. Evavold, and C. Zhu. Two-stage cooperative T cell receptor-peptide major histocompatibility complex-CD8 trimolecular interactions amplify antigen discrimination. *Immunity*, 34(1):13–23, 2011.

[42] A. M. Kalergis, N. Boucheron, M. A. Doucey, E. Palmieri, E. C. Goyarts, Z. Vegh, I. F. Luescher, and S. G. Nathenson. Efficient T cell activation requires an optimal dwell time of interaction between the TCR and the pMHC complex. *Nature Immunology*, 2(3):229–234, 2001.

[43] C. Kreutz, A. Raue, D. Kashek, and J. Timmer. Profile likelihood in systems biology. *The FEBS Journal*, 280(11):2564–2571, 2013.

[44] C. Kreutz, A. Raue, and J. Timmer. Likelihood based observability analysis and confidence intervals for predictions of dynamic models. *BMC Systems Biology*, 6(1):120–128, 2012.

[45] B. Laugel, D. K. Cole, M. Clement, L. Wooldridge, D. A. Price, and A. K. Sewell. The multiple roles of the CD8 coreceptor in T cell Biology: opportunities for the selective modulation of self-reactive cytotoxic T cells. *Journal of Leukocyte Biology*, 90(6):1089–1099, 2011.

[46] B. Laugel, H. A. van den Berg, E. Gostick, D. K. Cole, L. Wooldridge, J. Boulter, A. Milicic, D. A. Price, and A. K. Sewell. Different T cell receptor affinity thresholds and CD8 coreceptor dependence govern cytotoxic T lymphocyte activation and tetramer binding properties. *The Journal of Biological Chemistry*, 282(33):23799–23810, 2007.

[47] I. F. Luescher, E. Vivier, A. Layer, J. Mahiou, F. Godeau, B. Malissen, and P. Romero. CD8 modulation of T-cell antigen receptor-ligand interactions on living cytotoxic T lymphocytes. *Nature*, 373(6512):353–356, 1995.

[48] M. Mallaun, D. Naeher, M. A. Daniels, P. P. Yachi, B. Hausmann, I. F. Luescher, N. R. Gascoigne, and E. Palmer. The T cell receptor's α-chain connecting peptide motif promotes close approximation of the CD8 coreceptor allowing efficient signal initiation. *Journal of Immunology*, 180(12):8211–8221, 2008.

[49] T. W. McKeithan. Kinetic proofreading in T-cell receptor signal transduction. *Proceedings of the National Academy of Sciences of the United States of America*, 92(11):5042–5046, 1995.

[50] S. Minguet and W. W. Schamel. Permissive geometry model. *Advances in Experimental Medicine and Biology*, 640:113–120, 2008.

[51] S. Minguet, M. Swamy, B. Alarcon, I. F. Luescher, and W. W. Schamel. Full activation of T cell receptor requires both clustering and conformational changes at CD3. *Immunity*, 26(1):43–54, 2007.

[52] K. Murphy. *Janeway's Immunobiology*. Garland Science, New York, 2012.

[53] S. A. Murphy and A. W. Van Der Vaart. On profile likelihood. *Journal of the American Statistical Association*, 95(450):449–465, 2000.

[54] D. Naeher, M. A. Daniels, B. Hausmann, P Guillaume, I. Luescher, and E. Palmer. A constant affinity threshold for T cell tolerance. *The Journal of Experimental Medicine*, 204(11):2553–2559, 2007.

[55] D. Naeher, I. F. Luescher, and E. Palmer. A role for the α-chain connecting peptide motif in mediating TCR-CD8 cooperation. *Journal of Immunology*, 169(6):2964–2970, 2002.

[56] A. M. Norment and D. R. Littman. A second subunit of CD8 is expressed in human T cells. *The EMBO Journal*, 7(11):3433–3439, 1988.

[57] E. H. Palacios and A. Weiss. Function of the Src-family kinases, Lck and Fyn, in T-cell development and activation. *Oncogene*, 23(48):7990–8000, 2004.

[58] A. S. Perelson. Receptor clustering on a cell surface. III. Theory of receptor cross-linking by multivalent ligands: description by ligand states. *Mathematical Biosciences*, 53:1–39, 1981.

[59] A. S. Perelson and C. DeLisi. Receptor clustering on a cell surface. I. Theory of receptor cross-linking by ligands bearing two chemically identical functional groups. *Mathematical Biosciences*, 48:71–110, 1980.

[60] E. B. Puffer, J. K. Pontrello, J. J. Hollenbeck, J. A. Kink, and L. L. Kiessling. Activating B cell signaling with defined multivalent ligands. *ACS Chemical Biology*, 2(4):252–262, 2007.

[61] M. C. Purbhoo, D. J. Irvine, M. Krogsgaard, and M. M. Davis. T cell killing does not require formation of a stable mature immunological synapse. *Nature Immunology*, 5(5):524–530, 2004.

[62] J. D. Rabinowitz, C. Beeson, D. S. Lyons, M. M. Davis, and H. M. McConnell. Kinetic discrimination in T-cell activation. *Proceedings*

of the National Academy of Sciences of the United States of America, 93(4):1401–1405, 1996.

[63] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. Structural and practical identifiability analysis of partially observed dynamic models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, 2009.

[64] M. Reth. Antigen receptor tail clue. *Nature*, 338(6214):383–384, 1989.

[65] R. D. Salter, R. J. Benjamin, P. K. Wesley, S. E. Buxton, T. P. Garrett, C. Clayberger, A. M. Krensky, A. M. Norment, D. R. Littman, and P. Parham. A binding site for the T-cell co-receptor CD8 on the α 3 domain of HLA-A2. *Nature*, 345(6270):41–46, 1990.

[66] R. D. Salter, A. M. Norment, B. P. Chen, C. Clayberger, A. M. Krensky, D. R. Littman, and P. Parham. Polymorphism in the α 3 domain of HLA-A molecules affects binding to CD8. *Nature*, 338(6213):345–347, 1989.

[67] W. W. Schamel, I. Arechaga, R. M. Risueño, H. M. van Santen, P. Cabezas, C. Risco, J. M. Valpuesta, and Alarcón. Coexistence of multivalent and monovalent TCRs explains high sensitivity and wide range of response. *The Journal of Experimental Medicine*, 202(4):493–503, 2005.

[68] A. B. Schreiber, T. A. Libermann, I. Lax, Y. Yarden, and J. Schlessinger. Biological role of epidermal growth factor-receptor clustering: investigation with monoclonal anti-receptor antibodies. *The Journal of Biological Chemistry*, 258(2):846–853, 1983.

[69] J. Sloan-Lancaster, B. D. Evavold, and P. M. Allen. Induction of T-cell anergy by altered T-cell-receptor ligand on live antigen-presenting cells. *Nature*, 363(6425):156–159, 1993.

[70] J. E. Smith-Garvin, G. A. Koretzky, and M. S. Jordan. T cell activation. *Annual Review of Immunology*, 27:591–619, 2009.

[71] E. Stein, A. A. Lane, D. P. Cerretti, H. O. Schoecklmann, A. D. Schroff, R. L. Van Etten, and T. O. Daniel. Eph receptors discriminate specific ligand oligomers to determine alternative signaling complexes, attachment, and assembly responses. *Genes & Development*, 12(5):667–678, 1998.

[72] J. D. Stone, M. N. Artyomov, A. S. Chervin, A. K. Chakraborty, H. N. Eisen, and D. M. Kranz. Interaction of streptavidin-based peptide-MHC oligomers (tetramers) with cell-surface TCRs. *Journal of Immunology*, 187(12):6281–6290, 2011.

[73] J. D. Stone, J. R. Cochran, and L. J. Stern. T-cell activation by soluble MHC oligomers can be described by a two-parameter binding model. *Biophysical Journal*, 81:2547–2557, 2001.

[74] J. D. Stone and L. J. Stern. CD8 T cells, like CD4 T cells, are triggered by multivalent engagement of TCRs by MHC-peptide ligands but not by monovalent engagement. *Journal of Immunology*, 176(3):1498–1505, 2006.

[75] B. Sulzer, R. J. De Boer, and A. S. Perelson. Cross-linking reconsidered: binding and cross-linking fields and the cellular response. *Biophysical Journal*, 70(3):1154–1168, 1996.

[76] Y. Sykulev, M. Joo, I. Vturina, T. J. Tsomides, and N. H. Eisen. Evidence that a single peptide-MHC complex on a target cell can elicit a cytolytic T cell response. *Immunity*, 4(6):565–571, 1996.

[77] S. Takada and E. G. Engleman. Evidence for an association between CD8 molecules and the T cell receptor complex on cytotoxic T cells. *Journal of Immunology*, 139(10):3231–3235, 1987.

[78] R. Testi, D. D'Ambrosio, R. De Maria, and A. Santoni. The CD69 receptor: a multipurpose cell-surface trigger for hematopoietic cells. *Immunology Today*, 15(10):479–483, 1994.

[79] J. M. Turner, M. H. Brodsky, B. A. Irving, S. D. Levin, R. M. Perlmutter, and D. R. Littman. Interactions of the unique N-terminal region of tyrosine kinase p56lck with cytoplasmic domains of CD4 and CD8 is mediated by cysteine motifs. *Cell*, 60(5):755–765, 1990.

[80] S. Valitutti, S. Müller, M. Cella, E. Padovan, and A. Lanzavecchia. Serial triggering of many T-cell receptors by a few pMHC complexes. *Nature*, 375(6527):148–151, 1995.

[81] S. Valitutti, S. Müller, M. Salio, and A. Lanzavecchia. Degradation of T cell receptor (TCR)-CD3-ζ complexes after antigenic stimulation. *The Journal of Experimental Medicine*, 185(10):1859–1864, 1997.

[82] R. Varma, G. Campi, T. Yokosuka, T. Saito, and M. L. Dustin. T cell receptor-proximal signals are sustained in peripheral microclusters and terminated in the central supramolecular activation cluster. *Immunity*, 25(1):117–127, 2006.

[83] A. Veillette, M. A. Bookman, E. M. Horak, and J. B. Bolen. The CD4 and CD8 T cell surface antigens are associated with the internal membrane tyrosine-protein kinase p56lck. *Cell*, 55(2):301–308, 1988.

[84] D. J. Venzon and S. H. Moolgavkar. A method for computing profile-likelihood-based confidence intervals. *Applied Statistics*, 37(1):87–94, 1988.

[85] L. Wooldridge, A. Lissina, D. K. Cole, H. A. van den Berg, D. A. Price, and A. K. Sewell. Tricks with tetramers: how to get the most from multimeric peptide-MHC. *Immunology*, 126(2):147–164, 2009.

[86] L. Wooldridge, H. A. van den Berg, M. Glick, E. Gostick, B. Laugel, S. L. Hutchinson, A. Milicic, J. M. Brenchley, D. C. Douek, D. A. Price, and A. K. Sewell. Interaction between the CD8 coreceptor and major histocompatibility complex class I stabilizes T cell receptor-antigen complexes at the cell surface. *The Journal of Biological Chemistry*, 280(30):27491–27501, 2005.

[87] T. Yokosuka, K. Sakata-Sogawa, W. Kobayashi, M. Hiroshima, A. Hashimoto-Tane, M. Tokunaga, M. L. Dustin, and T. Saito. Newly generated T cell receptor microclusters initiate and sustain T cell activation by recruitment of ZAP70 and SLP-76. *Nature Immunology*, 6(12):1253–1262, 2005.

A Results of Model G2

In the following, the results obtained for Model G2, the binding model G (see Fig. 2G) with effective ligand valency $f = 2$, are discussed. In Fig. 17A, the best fit (solid lines) of Model G2 together with the standard deviation of individual measurements (shaded areas), as obtained from the error model (see Sec. 4.2), and the binding data (dots) are shown. The top row shows T1.CD8- cells and the bottom row T1.CD8+ cells. The ligand types are indicated by color. As for Model G3 (see Fig. 7), the agreement between model and data is very good.

The 95 % confidence intervals were estimated using the profile likelihood method (see Sec. 2.3), and are shown in Fig. 17B. All parameters are completely identifiable, except of q_1 and ξ, which have only an upper bound. The values of the dissociation constants, i.e. the inverses of the affinities, of the 4L and 4P ligand lie between $0.1\,\mu M$ and $1\,\mu M$, between $0.3\,\mu M$ and $3\,\mu M$ for the 4V ligand, between $6\,\mu M$ and $40\,\mu M$ for the 4S ligand, and between $1.5\,\mu M$ and $20\,\mu M$ for pMHC-CD8. These intervals contain the reported values of cognate ligands and murine CD8 [18, 30]. The multivalent engagement parameter κ_{MV} lies between $5\,\mu M$ and $30\,\mu M$. The MLE value of the dimensionless association constant of multivalent binding for the 4P ligand, $\kappa_{MV} K_{1,4P}$ assumes a value of about 10, which is similar to the value found for Model G3, and is in good agreement with previously reported values on such dimensionless multivalent binding constants [35, 23, 73].

The association constant for pMHC-independent TCR-CD8 complex formation q_1 is zero for the best fit and has an upper bound of 0.1. Thus, up to 55 % of the TCRs on T1.CD8+ cells might be bound to CD8 in the absence of ligands. This fraction does also depends on the CD8 expression level in a monotonically decreasing way. As for Model G3 , the

103
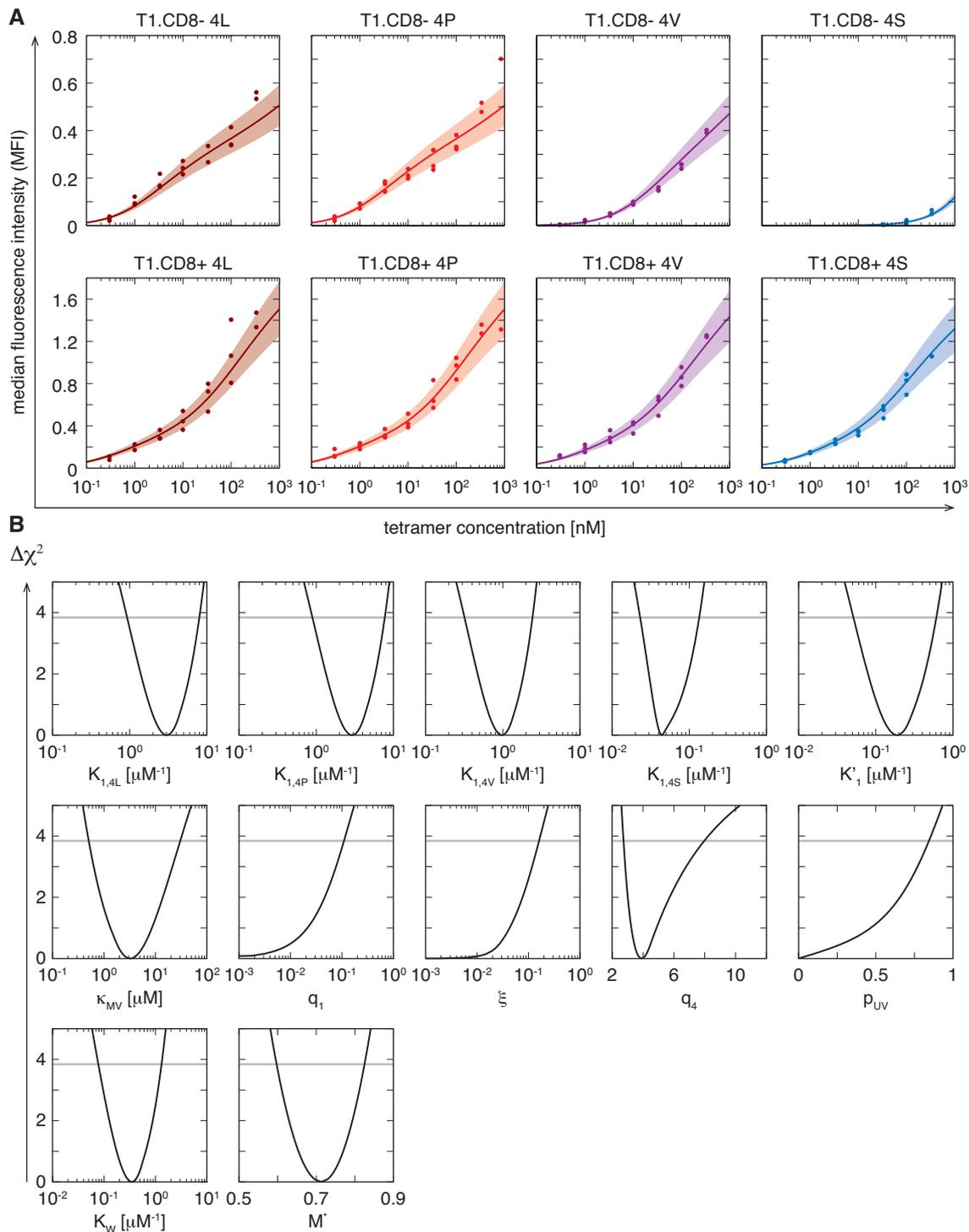
Figure 17: Fitting results of Model G2

(A) Best fit: the same as Fig. 7 but for Model G2. (B) Profile likelihoods: the same as Fig. 8 but for Model G2.

association constant of TCR-CD8 complex formation if pMHC is bound to CD8, q_4, is significantly larger than q_1, while ξ and q_1 are quite similar. This leads to the same implications as for Model G3. Hence, Model G2 and G3 behave qualitatively identical.

The probability to not form a covalent bond between a pMHC-TCR contact upon UV-light irradiation, p_{UV}, has an upper bound of 0.8. This means that at least $20\,\%$ of the pMHC-TCR contacts develop a covalent connection during UV-light exposure. The value of K_W is located in the range from $0.1\,\mu M^{-1}$ to $1\,\mu M^{-1}$. And lastly, the scaling factor M^* lies between 0.6 and 0.8.

List of Figures

107

List of Tables

List of Abbreviations

AIC_c	corrected Akaike Information Criterion
APC	antigen presenting cell
CBM	Competition Binding Model
CD	cluster of differentiation
CTL	cytotoxic T cell
gof	goodness-of-fit
ITAM	immunoreceptor tyrosine-based activation motif
MHC	major histocompatibility complex
MFI	median fluorescence intensity
MLE	maximum likelihood estimation
Lck	lymphocyte-specific tyrosine kinase
pMHC	peptide MHC
SHP-1	Src homology region 2 domain-containing phosphatase-1
TCR	T cell receptor
ZAP70	ζ-chain associated protein of 70 kDa