

How persistent infection overcomes peripheral tolerance mechanisms to cause T cell-mediated autoimmune disease

Rose Yin^{1,*}, Samuel Melton^{2,*}, Eric Huseby³, Mehran Kardar^{2,**}, and Arup K. Chakraborty^{1,2,4,5,**}

¹Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

²Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139.

³Department of Pathology, University of Massachusetts Medical School, Worcester, MA, USA.

⁴Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University, Cambridge, MA 02139.

⁵Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139.

*These two authors contributed equally.

**Address correspondence to: arupc@mit.edu or mkardar@mit.edu

November 29, 2023

Abstract

T cells help orchestrate immune responses to pathogens, and their aberrant regulation can trigger autoimmunity. Recent studies highlight that a threshold number of T cells (a quorum) must be activated in a tissue to mount a functional immune response. These collective effects allow the T cell repertoire to respond to pathogens while suppressing autoimmunity due to circulating autoreactive T cells. Our computational studies show that increasing numbers of pathogenic peptides targeted by T cells during persistent or severe viral infections increase the probability of activating T cells that are weakly reactive to self-antigens (molecular mimicry). These T cells are easily re-activated by the self-antigens and contribute to exceeding the quorum threshold required to mount autoimmune responses. Rare peptides that activate many T cells are sampled more readily during severe/persistent infections than in acute infections, which amplifies these effects. Experiments in mice to test predictions from these mechanistic insights are suggested.

Introduction

The key cell types that enable the adaptive immune system to mount pathogen-specific responses to a diverse and evolving world of microbes are T and B lymphocytes (T cells and B cells). Humans have billions of T cells and B cells, each of which expresses a T cell receptor (TCR) or B cell receptor (BCR) on its surface. T and B cell repertoires are characterized by an enormous diversity of TCRs and BCRs generated by VDJ recombination [1, 2, 3]. Upon infection with a pathogen, some receptors from this pool are likely to bind sufficiently strongly to molecular components of a specific pathogen

resulting in T or B cell activation, which can potentially result in an adaptive immune response. For example, T cell receptors (TCRs) bind to peptides (p) derived from a pathogenic protein bound to protein products of the major histocompatibility (MHC) genes. T cell activation is triggered if the TCR-pMHC bond has a sufficiently long half-life [4, 5, 6]. Different TCRs tend to bind to different pMHC molecules with sufficiently long half-lives, thus enabling the T cell repertoire to respond specifically to diverse pathogens. At the same time, the T cell repertoire is largely tolerant to pMHC molecules where the peptide is derived from host proteins. Such self-pMHC molecules are expressed ubiquitously on host cells. Tolerance to self is due to processes that occur during T cell development in the thymus and mechanisms that suppress autoimmune responses in peripheral tissues [7, 8, 9, 10].

Cells in the thymus (especially those in the medulla) express the AIRE gene which enables promiscuous gene expression and results in these cells displaying pMHC molecules with peptides derived from diverse regions of the host proteome [11]. During development, immature T cells (thymocytes) interact with these self-pMHC molecules. To successfully develop into a mature T cell, a thymocyte must bind to at least one self-pMHC molecule it encounters in the thymus with a binding free energy (or half-life) exceeding a threshold in order to receive a survival signal (positive selection) [12]. However, if a thymocyte's TCR binds to any encountered host pMHC molecule with a binding free energy exceeding a higher threshold, it is deleted (negative selection). In this way, positive selection aims to ensure that the mature T cell repertoire expresses TCRs with the ability to bind to pMHC complexes, while negative selection aims to delete self-reactive T cells [13, 14, 15, 16]. Theoretical and experimental studies have shown that negative selection also plays an important role in mediating the peptide-specificity of TCRs; i.e., most point mutations to a TCR's cognate peptide abrogate recognition [17, 18, 19, 20, 21, 22]. Some of these theoretical and experimental studies also showed that the peptide contact residues on TCRs of mature T cells that undergo normal negative selection are statistically enriched in amino acids that are moderately hydrophobic. We note also that thymocytes that successfully mature but express TCRs that bind more strongly to self-pMHC molecules are more likely to differentiate into regulatory T cells (Tregs) that can suppress responses from conventional mature T cells in tissues [22].

On average, a randomly picked conventional T cell from the mature repertoire has a lower probability of being activated by a randomly picked self-pMHC molecule compared to a pathogen-derived pMHC molecule. This is because the activation threshold for mature T cells is slightly higher, but similar, to that for negative selection of thymocytes [23], and every mature T cell interacted with some fraction of the self-pMHC molecules displayed in the thymus and was not deleted by negative selection. However, this difference in probabilities of activation by self and pathogen-derived pMHC molecules is likely to be small in many cases. Furthermore, a given thymocyte does not encounter every self-peptide from its host's proteome during thymic selection. If a mature T cell encounters a self-pMHC molecule that it did not encounter in the thymus, the difference in the probability that it will be activated by this self-pMHC molecule compared to a randomly picked pathogen-derived pMHC molecule is likely to be non-existent. These arguments suggest that imperfect thymic selection and stochastic effects associated with T cell activation should prevent robust discrimination between self and foreign antigens at the level of an individual T cell. Indeed, the mature T cell repertoire is known to include T cells that can be activated by some host-derived peptides [24, 25, 26, 27, 9, 28]. Full blown autoimmune responses to these self-peptides is thought to be suppressed by peripheral tolerance mechanisms, which include the action of Tregs [29, 30] and more recent reports of CD8⁺ suppressor T cells [8]. But, differences in the probabilities of activation of a T cell by self and pathogen-derived pMHC molecules can be small to non-existent. So, how do peripheral tolerance mechanisms suppress self-activated T cells and autoimmunity while not suppressing the ability of the repertoire to respond effectively to pathogens?

In order to mount a functional immune response, T cells must not only be activated but also proliferate and differentiate to acquire effector functions. Proliferation and differentiation require cytokines [31, 32]. For example, IL2 is necessary to promote proliferation [33, 34]. Tregs can compete with activated cells for these resources necessary for growth and differentiation to inhibit activated T cells from mounting a functional immune response [35, 36, 10]. A theoretical analysis [37] posited that if a sufficient number of T cells were activated in the same tissue, they could produce enough of the factors necessary for proliferation and differentiation and share these resources with each other to overcome the competition from suppressive regulatory mechanisms such as those mediated by Tregs. This concept provides a mechanism for how activated autoreactive T cells are inhibited from mounting a functional immune response by peripheral suppressive mechanisms while T cell activation by pathogenic pMHC molecules can overcome this effect. The difference in probability of a T cell being activated by self and pathogen-derived pMHC is amplified by the requirement that a sufficient number of T cells must be activated in the same tissue for an effective immune response to ensue. To illustrate how this collective effect amplifies this difference, as a contrived example let the average probability of a T cell being activated by a pathogen-derived and host-derived pMHC molecule be 0.5 and 0.4, respectively (1.25-fold difference); the difference is due to thymic selection. If the number of T cells that need to be activated for an effective response is 10, the probability of a functional immune response to a pathogenic or self-pMHC molecule is 0.5^{10} and 0.4^{10} , respectively. So, a functional response to a pathogenic pMHC molecule is over 9 (1.25^{10}) times higher than for a host-derived pMHC molecule. Due to stochastic effects, a 1.25-fold difference in the probability of T cell activation by self and pathogenic pMHC molecules is unlikely to robustly differentiate between self and non-self, but a 9-fold difference can be sufficient. Thus, collective effects can provide a robust mechanism enabling the T cell repertoire to mount effective responses against pathogens while suppressing T cells activated by self in spite of imperfect thymic selection and stochastic effects. This mechanism proposed by Butler et. al. can be considered to be akin to quorum sensing by bacteria [37, 38]. The quorum number of activated T cells required for a functional response has to be such that, with high probability, it is not exceeded upon interactions with host pMHC molecules, but is exceeded by pathogen-derived pMHC molecules [Fig. 1].

Note that the amplification of small differences in T cell activation probabilities to result in large differences in the probability of a functional response due to collective effects embodied in the quorum model is a highly non-linear effect. If the quorum number is 10, then the difference in the probability of a functional response if 5 or 6 T cells are activated is negligible, but the difference in outcome if 9 or 10 T cells are activated is very large.

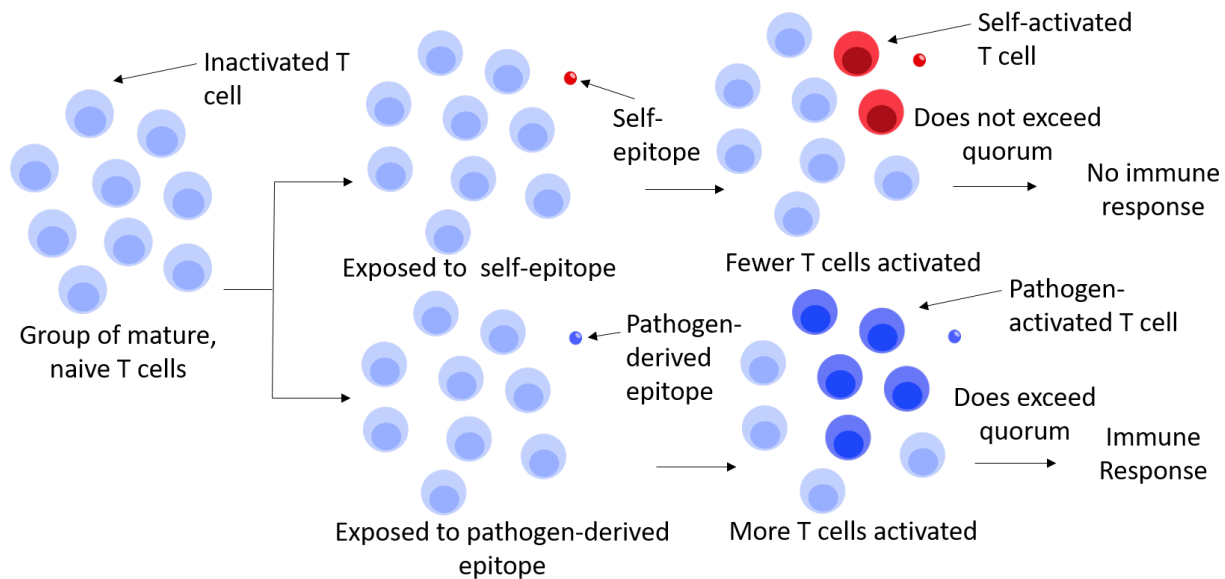


Figure 1: Schematic depiction of how quorum sensing by T cells can result in functional immune responses to pathogen-derived pMHC molecules and not host-derived ones in spite of the existence of autoreactive T cells in the mature repertoire. For the case depicted, the quorum number is taken to be four.

Several experimental observations provide evidence for the importance of collective effects and the quorum model. Following observations of T cells forming clusters by interacting with the same APC [39] and suggestions of potential cooperativity between T cells [40, 41, 42, 43], Gerard et. al. used two-photon microscopy in mice to discover junction formation between CD8 T cells and show that these interactions increased T cell sensitivity to cytokines [44]. This study also showed that interactions between adhesion molecules LFA-1 and ICAM-1, both expressed on T cells, was important for junction formation. Junction formation mediated sharing of the cytokine, $\text{IFN-}\gamma$, and this was important for CD8 T cell differentiation and memory cell development. In a more recent study [45], related results were reported in vitro. CD8 T cell clustering around DCs expressing a stimulatory ligand was mediated by ICAM-1 binding to LFA-1 proteins on other T cells. CD80 expression also increased on activated T cells, enabling binding to CD28 on neighboring T cells. CD28 binding to CD80 results in signaling that mediates the secretion of several cytokines, including IL2 [46, 47, 48, 49]. Experiments showed that clustered T cells expand in an IL2-dependent manner and that the IL2 amount scaled with the density of clustered activated T cells.

Another elegant study combined experiments in mouse models and computation [10]. Tregs require IL2 for mediating their suppressive functions but do not secrete IL2 themselves [50, 36]. The IL2 receptor ($\text{IL2R}\alpha$) on Tregs can bind to IL2 produced by vicinal activated T cells to sequester IL2 and trigger Treg proliferation and effector functions. Wong et. al. showed that this feedback regulation of Treg's suppressive effects could be inhibited by preventing the expression of molecules like $\text{IL2R}\alpha$ in Tregs. Inhibiting Treg activity led to IL2-mediated signaling in activated T cells. The increase in signaling and IL2 production can then outcompete Tregs' ability to bind to IL2. The range of IL2 diffusion is also thus increased, which allows other activated T cells in the vicinity to access IL2. This study also showed that Treg density controls the balance between Treg-mediated suppression and IL2 signaling by activated T cells in a non-linear way. This observation is consistent with the quorum model. Treg density should control the number of activated T cells required to

outcompete the effects of Tregs (the quorum number), and the dependence should be non-linear due to the collective effects inherent in the quorum model (see above). Related observations were made in another study in mice and in vitro on the differentiation of activated CD4 T cells to memory cells [51]. Differentiation into memory T cells in mice was dependent on the precursor frequency of antigen-specific CD4 T cells. In vitro, differentiation into memory cells in microwells was density-dependent. IL2 concentrations were found to decline rapidly with distance from the secreting T cell, thus explaining the density dependence of IL2-mediated signaling that can result in differentiation.

Evidence for quorum sensing is also suggested by comparing the statistical properties of the sequences of the CDR3 regions of mouse T cell repertoires at different stages of thymic development [52]. This study found that post and pre-selection repertoires could be distinguished based on the statistical properties of the ensemble of sequences rather than individual sequences, thus arguing for collective effects in self-nonsel self discrimination.

A T cell-mediated autoimmune condition results when the central and peripheral tolerance mechanisms described above fail. In several contexts, it has been observed that persistent pathogen infections often trigger T cell mediated autoimmunity. Examples include type I diabetes (T1D) [53, 54, 55] and multiple sclerosis (MS) [56, 57]. In the case of MS, longitudinal studies have shown that infection with Epstein-Barr Virus (EBV) is a necessary condition for MS [58] [Fig. 2A]. Recent studies suggest that T1D may be triggered by long-term infections by human enteroviruses (EV) or long COVID [59, 60, 61] [Fig. 2B]. Even though EV has typically been seen as an acute infection, it seems that EV can persistently infect human pancreatic islet cells [62] that are the targets of autoimmune responses in T1D.

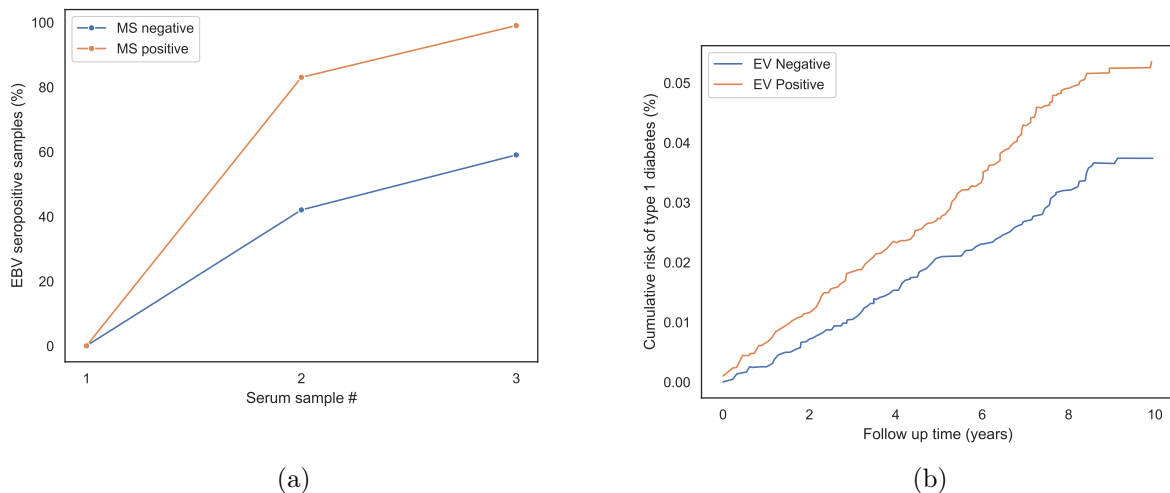


Figure 2: Examples of increasing autoimmunity risk given persistent infection. (a) Figure adapted from Figure 2 in Ref. [58]. This study tracked MS and EBV status of over 10 million individuals in the US military over 20 years. This figure shows the proportion of individuals with and without MS who were EBV-positive at three different times. Serum sample number was determined from the MS cohort. The 1st serum sample corresponds to the first serum sample collected from each MS patient, the 3rd sample corresponds to the one last collected before MS disease onset, and the 2nd sample corresponds to one collected in between. There were two controls (MS negative) samples for each MS positive sample with the same age, sex, race/ethnicity, branch of military service, and blood sample collection dates. As seen, a significantly higher proportion of individuals who later developed MS were EBV-positive in the second and third times of serum collection compared to those who were MS-negative. (b) Figure adapted from Figure 2 in Ref. [63]. This study tracked T1D status of 570,133 children with EV and frequency-matched children without EV over 10 years. Cumulative risk of T1D increased more rapidly in children with EV than those without.

Infection-induced onset of autoimmunity has been attributed to diverse factors including specific MHC genetics, T regulatory cell dysfunction, epitope spreading, and molecular mimicry [64, 65]. The "molecular mimicry" concept suggests that T cells primed by the foreign antigen are crossreactive to similar self-peptides, thus resulting in an autoimmune response. For example, there is evidence that, compared to healthy people, MS patients exhibit higher levels of CD4+ T cells that are crossreactive to EBV peptides and those derived from myelin (a target of autoimmune responses in MS) [66, 67]. Human Leukocyte Antigen (HLA)-B27 is a human MHC allele that is associated with autoimmune diseases like ankylosing spondylitis (AS) and acute anterior uveitis (AAU). Recent studies have identified peptides derived from microbes and self-peptides presented by HLA B27 that activate T cells isolated from AS and AAU patients [67]. The "epitope spreading" mechanism proposes that an antiviral immune response leads to tissue damage and the release of otherwise sequestered self-antigens into the local region, which are then presented by nearby APCs to activate autoreactive T cells [56]. In the case of T1D, it is well known that both CD4+ and CD8+ T cells progressively destroy islet beta cells in the pancreas [68, 69, 70, 71], thus potentially generating higher levels of tissue antigens. Higher levels of inflammation in infected tissues can also result in higher levels of self-antigen presentation on antigen presenting cells [72, 73], including increased cross-priming of MHC class I responses [74].

Despite these proposed mechanisms, clear explanations for several features of the association of

autoimmune diseases with persistent viral infections are not available. For example, why do the mechanisms that normally suppress autoimmunity discussed above fail for persistent or severe viral infections? Why does EBV infection only increase the probability of developing MS and not always lead to disease? In addition, individuals with the persistent viral infection do not immediately develop the corresponding autoimmune condition. This latency of pathogenesis is not sufficiently explained by the current paradigms.

Based on computational studies, we propose a mechanism for why certain autoimmune conditions are triggered by particular persistent viral infections. This mechanism is compatible with the concepts of molecular mimicry and epitope spreading, but relies critically on understanding why the collective non-linear effects embodied in the quorum model that usually suppress autoimmunity fail upon persistent or severe infections. Consistent with molecular mimicry, our computational studies suggest that specific persistent viral infections can trigger particular autoimmune conditions because T cells activated by certain peptides derived from this pathogen are weakly cross-reactive to specific self-antigens presented in a tissue. Normally, the weak reactivity to these self-antigens does not activate these T cells and so they cannot contribute to exceeding the quorum threshold for mounting functional immune responses to these antigens. Our experimental data show that the activation threshold is lower for previously activated T cells [Fig. 5]. Therefore, once activated by the pathogen-derived peptide, the activation threshold for reacting to the cross-reactive self-antigens is effectively lower. This allows the T cells activated by a pathogenic peptide that are cross-reactive to a self-antigen to be added to the pool of T cells that are usually robustly activated by this self-antigen. This can result in exceeding the quorum threshold, thus resulting in a functional autoimmune response. Pathogen-derived peptides that satisfy these conditions are rare. Persistent or severe infection leads to increasing numbers of pathogen-derived peptides being targeted, which increases the probability of sampling such peptides and overcoming the quorum threshold for responses to certain self-peptides. Local inflammatory conditions and tissue damage during persistent viral infections can increase the presentation of different self-antigens. This increases the probability of finding pairs of pathogen-derived and self-antigens with cross-reactive T cells that can exceed the quorum threshold for the self-antigens, resulting in autoimmunity. Our model provides an explanation for why individuals who are persistently infected with viruses like EBV and EV develop the corresponding autoimmune disease only with some probability, as well as the latency of pathogenesis.

We first develop and employ a computational model of T cell selection in the thymus to show that appropriately set values of positive and negative selection thresholds result in mature repertoires that contain autoreactive T cells while maintaining tolerance to self-antigens with high probability because the quorum threshold is not exceeded by T cells in response to self-antigens alone. We then use our computational model to study the response of such T cell repertoires to infection. Consistent with data [63, 58], we assume that a persistent or severe infection results in the presentation of multiple pathogen-derived peptides either sequentially or at the same time. Our computational results show that the probability of triggering autoimmunity increases monotonically with increasing numbers of pathogen-derived pMHC molecules that are targeted by T cells. Further analyses of our computational results show that rare pathogen-derived peptides that could trigger activation of many T cells may play a significant role in mediating autoimmunity upon persistent viral infection. Importantly, we propose specific experiments in mouse models that could directly test the importance of how collective effects and molecular mimicry are intertwined to break tolerance upon persistent viral infections.

Development of a simple computational model of thymic selection and immune responses of the mature T cell repertoire

In this section, we first describe our computational model for development of the mature T cell repertoire in the thymus. We then utilize our model to characterize the response of the thus generated mature repertoire to infection (i.e., pathogen-derived pMHC molecules), and the impact of infection on the potential emergence of autoimmunity.

Model for development of the T cell repertoire

During development, a thymocyte encounters a number of self-pMHC complexes displayed on cells in the thymus and undergoes positive and negative selection. These processes occur in stages in different compartments of the thymus; for example, negative selection occurs in two different stages. These details of stages of selection are not incorporated in our simple model. Each interaction of a thymocyte’s TCR with a self-pMHC corresponds to a different binding free energy. We use a convention in which lower binding free energy corresponds to higher affinity. To successfully mature and exit the thymus, the binding free energy of a thymocyte’s TCR to at least one of these pMHCs must be lower than E_p (positive selection). The binding free energy must also not be lower than a more negative value, E_n , for interactions with any of the encountered self-pMHC as otherwise the thymocyte is negatively selected. This implies that the minimum value of the binding free energy of a thymocyte with the encountered self-pMHC molecules must lie between E_p and E_n . Therefore, our model for the development of a mature T cell repertoire involves generating a panel of self-pMHC molecules and thymocytes (each expressing a TCR), calculating the binding free energies of every TCR with each encountered self-pMHC molecule, and checking whether the minimum value lies between E_p and E_n . Thymocytes with TCRs that satisfy this condition are added to the mature T cell repertoire [Fig. 3A]. To carry out this procedure, we need to define a model for TCRs expressed on thymocytes and self-pMHC molecules, and their binding free energies. This model will also be used to study the response of the mature repertoire to pathogen-derived pMHC molecules upon infection.

Our model for TCRs and pMHC molecules and their binding free energies follows the framework of Ref. [19], in which TCR and pMHC molecules are represented as strings of amino acids. TCRs interact with the peptide and the MHC. Past work using a similar model has argued that interaction free energies with the MHC are drawn from a relatively narrow distribution, and including these interactions do not change the qualitative results pertinent to the characteristics of the mature repertoire. Therefore, we do not explicitly include MHC, and represent only the peptides and the peptide contact residues of TCRs as strings of amino acids [Fig. 3B]. We assume that each peptide contact residue on a TCR interacts only with a corresponding site on a peptide. Previous work with a similar model that included more complex interaction patterns between these sites showed that most qualitative results are not altered by considering this complexity [75]. For a given TCR-pMHC pair, the interaction free energy is calculated as:

$$E_i = \sum_{i=1}^{\ell} J(l_i, j_i), \tag{1}$$

where ℓ is the length of the variable regions of the TCRs and pMHCs as seen in Figure 3B, and $J(l_i, j_i)$ is the interaction free energy between the amino acid at the i -th peptide contact residue of the TCR (l_i) and the peptide (j_i). For instance, in Figure 3B, the interaction free energy between the first peptide contact residue of the TCR and peptide is that of Proline (P) and Threonine (T), and the total interaction free energy is the summation of the interaction free energies for each pair of amino acids between the TCR (PLNDAENTF) and the peptide (TGACFLRDI). Here, we took ℓ

to be 9, the approximate length of MHC class I binding peptides. However, the model is formulated to be MHC class agnostic and could apply to either class. Moreover, reducing the number of contact residues, as they are typically fewer [Section [Experimentally testable predictions of the model](#)][22] should not change our qualitative results. The interaction free energy between a pair of amino acids in the TCR and peptide is calculated using the Miyazawa-Jernigan (MJ) matrix in units of $k_B T$ [76].

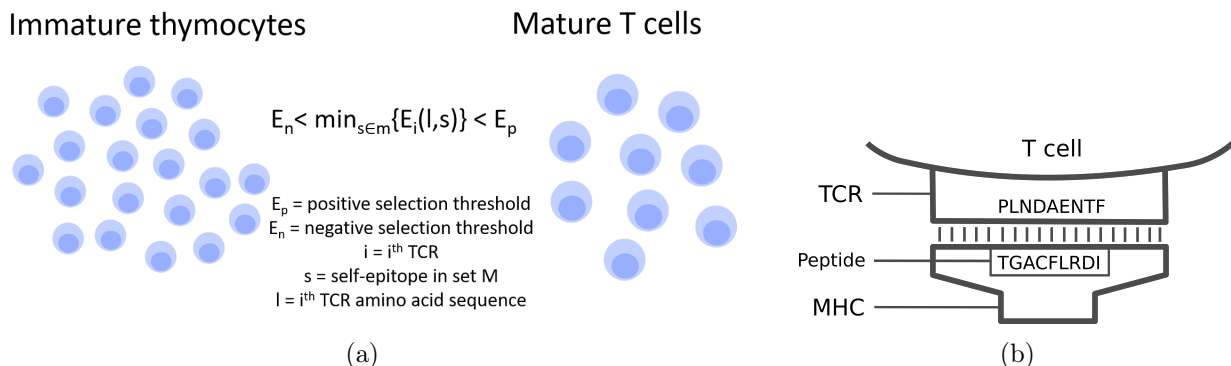


Figure 3: Depiction of thymic selection processes and model for TCR-pMHC binding free energy: (a) Modeling thymic selection: only thymocytes with a minimum binding free energy between E_p and E_n go on to mature. (b) Model for TCR-pMHC binding free energy: The variable residues of the TCR that interact with the peptide in the pMHC complex are represented by strings of amino acids. Addition of pairwise interactions of amino acid along the two strings provides the total binding free energy E_i between a TCR-pMHC pair.

This simple model of TCR-pMHC interactions has been used previously to study thymic development. One prediction that emerged is that positive and negative selection tune the statistical distribution of the biochemical properties of the amino acids in the peptide contact residues of the TCRs on mature T cells. The peptide contact residues of TCRs in the mature repertoire were predicted to be statistically enriched in amino acids that interact moderately with other amino acids [19, 20, 21]. Attesting to the utility of the model, this prediction was tested positively in mouse models [22]. Moderately hydrophobic amino acids were found to be statistically enriched in the key peptide contact residues of TCRs in mature T cells, and this property of the mature T cell repertoire was dependent on normal negative selection. Hydrophobicity is a natural measure of the strength of interactions at an interface, such as for TCR-pMHC binding.

Although statistical models for VDJ recombination are available [3, 77, 78], for simplicity, the sequences of peptide contact residues of the TCRs of thymocytes were generated using just the amino acid frequencies that characterize the human proteome [Supp. 1]; 10^6 unique immature thymocytes were generated. As we will note later, because T cell clonality is not considered in our analyses, our results represent a conservative estimate of the probability of triggering autoimmunity upon persistent viral infection. To generate sequences of self-peptides, we note that not all peptides can be properly presented by all MHCs. To construct a library of self-peptides to present to thymocytes, we screened self-peptides for presentation by human MHC alleles (Human leukocyte Antigen or HLA) using NetMHCpan [79]. This bioinformatics tool is trained on measured peptide-HLA-I interactions and predicts binding with 95% accuracy for each HLA allele. We focused on HLA-A*01:01, HLA-A*02:01, HLA-B*08:01 and HLA-B*07:02, the most commonly found HLAs in the human population. NetMHCpan frameshifts the human proteome from UniProt [80, 81] and determines whether each peptide of length 9 thus generated binds to a specified HLA. We then

took the sequences of peptides predicted to bind to a particular HLA and compared their binding affinities to the distribution of affinities calculated for a set of randomly chosen peptides that have been found to be presented on HLA in human samples; only those predicted binders with HLA binding affinities in the top 2% of the latter affinity distribution were chosen to be among our set of self-pMHC molecules. Thus, we generated 10,000 self-peptides (N) with equal numbers of peptides that bind to HLA-A*01:01, HLA-A*02:01, HLA-B*08:01 and HLA-B*07:02.

To generate the mature T cell repertoire as described above, we expose each TCR generated as described above to 8,000 self-peptides chosen randomly from our set of 10,000 self-peptides. We generate several T cell repertoires to obtain our results (see later), and each time a different random set of 8,000 self-peptides is chosen. We need to pick appropriate values for E_p and E_n to obtain the results of selection. To do so, we simulated thymic selection with different choices of E_n and the gap $\Delta = E_p - E_n$. For each value of E_n , the value for the gap Δ was chosen such that 4% of thymocytes survived selection, as observed in experiments in mouse models [82]. As there are typically two values of E_n that satisfy this requirement [Supp. 2], the value for which negative selection is limiting was chosen [20]. This is consistent with experiments in mouse models [22] that show that defects in negative selection change important statistical characteristics of the mature T cell repertoire, such as the frequency of hydrophobic amino acids that make up TCR peptide contact residues. Our results are not sensitive to the exact value of E_n (as long as negative selection is limiting) [Supp. 2]; we used $E_n = -38k_B T$ for the results described in the main text.

Modeling infection and its impact on autoimmunity

During an acute infection, only a few pathogen-derived peptides are immunodominantly targeted by T cells [83]. During persistent or severe infection, either as time ensues or at the same time, many more peptides can be targeted by T cells. We thus consider infection with different numbers of pathogen-derived peptides (N_f), and ask whether the T cell response to these peptides can also trigger a functional autoimmune response. The chosen values of N_f range from 0 to 10, with higher values of N_f representing a persistent or severe infection. Taking *Listeria monocytogenes* as a model pathogen and its proteome from UniProt [84], we used NetMHCpan to generate 40,000 peptides that can be presented by the same human HLA alleles used in thymic selection of the mature T cell repertoire.

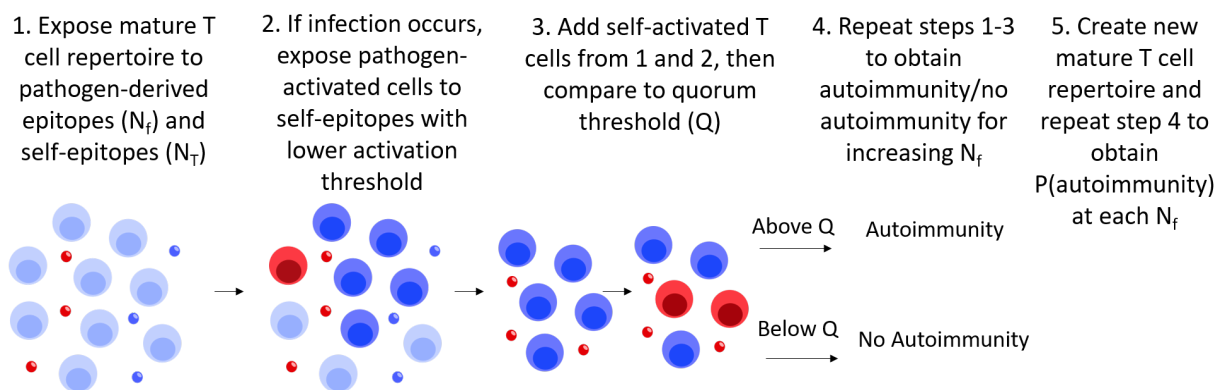


Figure 4: Steps in modeling infection: We follow the steps shown and described in the text for increasing numbers of pathogen-derived peptides (or epitopes), N_f , and the same number, N_T , of self-peptides to calculate the probability of a functional autoimmune response.

Figure 4 outlines the steps in our method for modeling infection. We first generate a particular

realization of a mature naive T cell repertoire using the method in the preceding section. Then, we randomly pick one of the 40,000 *Listeria* peptides, as well as N_T self-peptides from our panel of 10,000, assuming that both sets can be presented by HLAs in the infected tissue. Next, we calculate the binding free energies of the mature T cells in the repertoire with the pathogen-derived peptide. If this free energy is lower (higher affinity) than the activation threshold, taken to be E_n , we count the T cell as being activated. Thus, we determine how many T cells are activated by the chosen pathogen-derived peptide. We chose a sharp threshold of activation because in individual T cells the membrane-proximal signaling network exhibits a digital, all-or-nothing response [85, 86, 87]. Setting the activation threshold to equal E_n ensures that mature T cells that survive selection are not activated by the self-peptides they encountered during thymic selection [23]. If the number of T cells activated by the pathogen-derived peptide is below the quorum number Q , then this event is counted as non-infectious, and also does not trigger autoimmunity. If the number of T cells activated by the pathogen-derived peptide exceeds Q , we study the impact of this functional immune response on triggering autoimmunity. Following indications from experiment [45, 51], we choose the quorum number to be $Q = 10$. Note that this value is likely the average value of Q , a point that we will elaborate upon later.

When there is a functional response to the infection, we take each T cell activated by the pathogen-derived peptide and calculate its binding free energy with the N_T self-peptides presented in the tissue. The goal is to determine if these activated T cells are cross-reactive to the self-peptides in the tissue. Checking for cross-reactivity is consistent with the concept of molecular mimicry. To determine whether or not an activated T cell is crossreactive to each self-peptide in the tissue, we account for the fact that activated T cells have a lower threshold for re-activation. Therefore, for T cells previously activated by a pathogen-derived peptide, the binding free energy threshold for reactivation upon interactions of its TCR with the self-pMHC molecules is taken to be higher than E_n . The value of this lower binding free energy, E_{weak} , is chosen based on our experimental data [Fig. 5]. We measured the response of a population of B3K506 TCR Tg T cells with receptors specific for an antigen labelled 3K. The response of both naive and activated T cells were measured using titrating concentrations of peptide presented by dendritic cells. The fraction of responding T cells was determined by measuring whether they produce the cytokine $\text{TNF}\alpha$ [Fig. 5]. Another example of this type of data for a different peptide (P5R) is shown in Supp. 3.

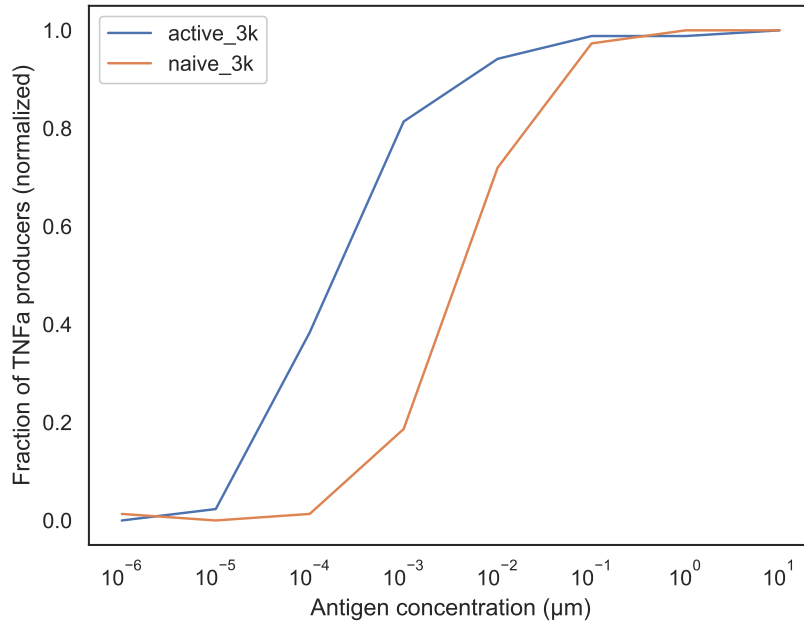


Figure 5: Graph of antigen concentration and % of T cell population that are TNF α producers when exposed to the 3k antigen. The red and blue curves correspond to the responses of naive and previously activated T cells, respectively.

We used the response curves in Figure 5 to estimate $E_n - E_{\text{weak}}$. While individual T cell activation is digital, activation at the population level has a sigmoidal shape [85]. However, a typical T cell activation threshold can still be inferred from the population response curve [23] by noting that the experimental curves resemble standard binding fractions as a function of chemical potential μ and binding free energy E_B , given by

$$f = \frac{1}{1 + e^{\beta(\mu - E_B)}}. \quad (2)$$

The chemical potential at a response fraction of 1/2 corresponds to the binding free energy. Since in a dilute solution concentration (c) is proportional to $e^{\beta\mu}$, the difference in concentrations at a response fraction of 1/2 for naive and activated T cells in Figure 5 can be related to differences in binding free energies by

$$e^{\beta(E_B - E'_B)} = \frac{c_{1/2}}{c'_{1/2}}. \quad (3)$$

By comparing the concentrations at a response fraction of 1/2 in Figure 5, we estimate $E_{\text{weak}} - E_n \approx 3k_B T$, and using $E_n = -38k_B T$ we obtain $E_{\text{weak}} = -35k_B T$.

Using this lower activation threshold, we tested whether any of the pathogen-activated T cells are also activated by the panel of N_T self-peptides presented in the tissue. Let us denote the total number of crossreactive T cells activated by all N_T self-peptides by N'_{self} . If the sum of T cells activated by these N_T self-peptides before infection and N'_{self} exceeds the quorum threshold, this event is counted as one that triggers autoimmunity upon infection. This procedure is repeated if additional pathogenic peptides are targeted by T cells during infection. If we used peptide A for the one pathogen-derived peptide ($N_f = 1$), for $N_f = 2$ we add a second peptide, B. If B also activates a larger number of T cells than Q , these T cells are also tested for crossreactivity with the N_T

self-peptides in the tissue. The crossreactive ones are added to the pool of T cells that could exceed the quorum threshold against the self-antigens, and if the quorum threshold is exceeded this counts as an autoimmune event if the two pathogen-derived peptides, A and B, are both targeted. This procedure is continued sequentially until a desired value of N_f is reached. To obtain statistically meaningful results for the probability of triggering autoimmunity for each value of N_f , we repeat the above calculations 10,000 times, with each trial carried out with new choices of the N_f pathogen-derived peptides and the N_T self-peptides displayed in the tissue, but with the same mature T cell repertoire. The probabilities thus obtained represent the average chance that autoimmunity will be triggered in a particular individual with a specific T cell repertoire for each value of N_f . Finally, all the calculations described above are carried out for 30 different realizations of the mature T cell repertoire (or individuals), and the average value for triggering autoimmunity is reported for each value of N_f . Table 1 summarizes the parameter values used in our simulations.

Parameter	Value
Number of immature thymocyte	1,000,000
Fraction of surviving thymocytes	0.04
Number of surviving thymocytes	40,000
E_n	$-38k_B T$
Number of self-peptides presented during selection (M)	8,000
Number of self-peptides in total (N)	10,000
Number of pathogenic peptides in total	40,000
Number of self-peptides presented in tissue (N_T)	10
E_{weak}	$-35 k_B T$
Quorum threshold	10
Number of trials	10,000

Table 1: Parameters used in our simulations.

Results

As noted earlier, consistent with data [63, 58], we assume that a persistent or severe infection results in the presentation of multiple pathogen-derived peptides either sequentially or at the same time. Thus, the variation of the probability of triggering autoimmunity as a function of the number of pathogen-derived peptides targeted in a tissue (N_f) should shed light on differences in the chance of triggering autoimmunity upon persistent infection or severe infections and more usual infections, with severe or persistent infections corresponding to larger values of N_f . In the subsections to follow, we will first describe the results of our simulations and then elaborate the mechanistic reasons that underlie our results.

The probability of triggering autoimmunity grows with the number of pathogen-derived peptides targeted by T cells during infection

As depicted in Figure 6A, simulations show that increasing the number of pathogen-derived peptides targeted by T cells also increases the probability of triggering autoimmunity. The increased chance of autoimmunity with N_f provides an explanation for why persistent or severe infections can increase the chance of developing autoimmune conditions. As the probability of developing

autoimmunity is not large, not everyone with a persistent infection develops autoimmunity. Our results also explain why there is a lag time between the establishment of persistent infection and the onset of autoimmunity. As time ensues, increasing numbers of pathogen-derived peptides are presented and targeted by T cells, and the chance of triggering autoimmunity grows as per the results reported in Figure 6.

Our results also explain why only certain viral infections trigger particular autoimmune conditions. A key component of our model is that T cells activated by pathogen-derived peptides are weakly cross-reactive to self-peptides presented in the same tissue. Such crossreactivity or molecular mimicry is only possible for certain self-antigens and pathogen-derived peptides, such as the observed crossreactivity between EBV and myelin-derived peptides in the case of MS [66, 88, 67] and microbial and self-peptides in the case of AS [67]. However, such crossreactivity exists regardless of the collective effects inherent in the quorum model. Figures 6A and 6B show the influence of collective effects as we compare what happens when $Q = 1$ (no collective effects) and $Q = 10$. Increasing the quorum number, Q , not only has the desired effect of reducing the intrinsic probability of autoimmunity due to circulating autoreactive T cells in the absence of infection ($N_f = 0$), but also decreases the relative chance of triggering autoimmunity upon viral infections [Fig. 6A]. This is because two conditions that depend on collective effects must be satisfied to trigger autoimmunity. First, the number of T cells activated by a pathogen-derived peptide must exceed the quorum threshold and then the ones among these T cells that are crossreactive to self-antigens in the same tissue must contribute to exceeding the quorum threshold with the self-antigens. In some instances, it is possible that a pathogen-derived peptide can trigger a sufficiently large number of T cells that the ones among these that are proliferating and cross-reactive to a self-antigen are the major contributors to exceeding the quorum threshold with the self-antigen and triggering autoimmunity. We will discuss this point more fully in the next section. As noted earlier, we do not consider multiple clones of the same T cell. Including clonal T cell populations while still using exactly our model would make the probability of triggering autoimmunity higher compared to that reported in Figure 6.

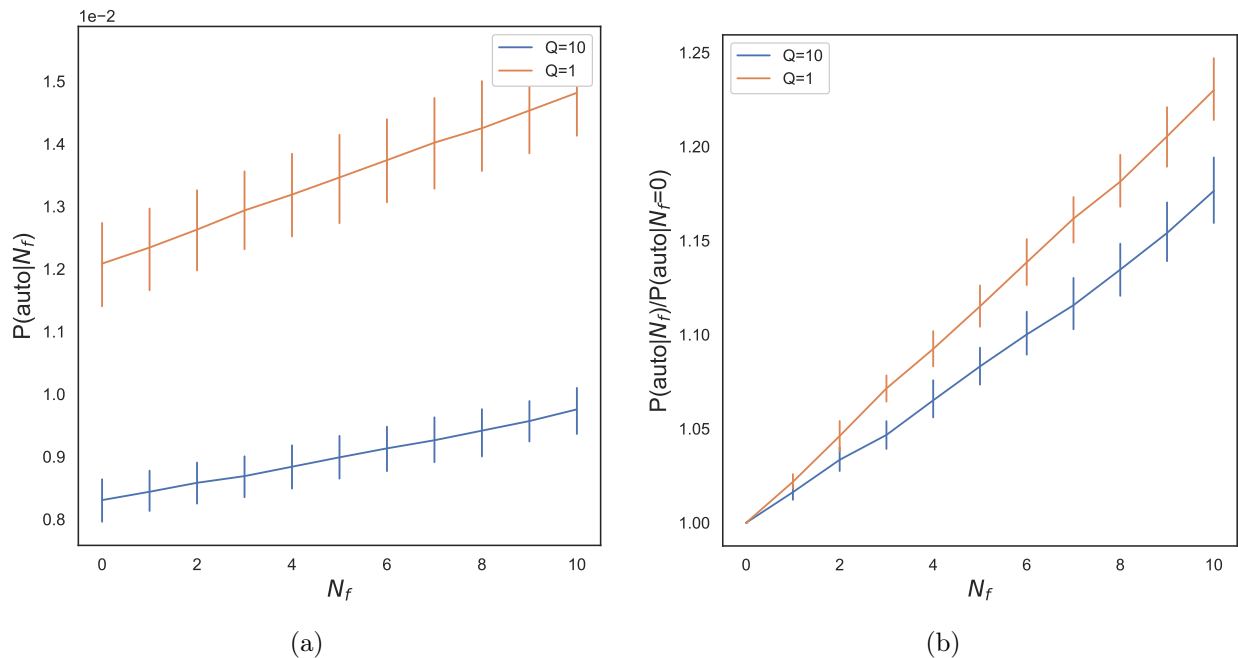


Figure 6: (a) Probability of triggering autoimmunity, $P(\text{auto}|N_f)$, as a function of the number of pathogen-derived peptides targeted for two different values of Q , $Q = 1$ (no collective effects) and $Q = 10$ (collective effects as per quorum model). (b) Probability of triggering autoimmunity, $P(\text{auto}|N_f)$ normalized by the probability of triggering autoimmunity in the case of no infection, $P(\text{auto}|N_f=0)$. Parameters used are as shown in Table 1. The error bars were calculated based on the number of simulations noted in Figure 4.

The number of self-peptides that T cells encounter in the thymus can vary from person to person and is different for different T cells. Also, experimental results indicate that a HLA allele associated with diabetes binds self-peptides in a less stable way [89], suggesting that T cells restricted by this HLA allele are likely to encounter fewer self-peptides during development. We investigated whether the collective effects embodied in the quorum model can make the immune system more robust by inhibiting autoimmune responses upon persistent or severe infections even in the face of such variations. To address this question, we carried out simulations with different values for the average number of self-peptides T cells encounter in the thymus (M) and calculated the change in probability of triggering autoimmunity upon persistent viral infection ($N_f = 10$) for different values of Q . Figure 7A shows the results of simulations for $M = 8,000$ (as in Fig. 6) and a lower value, $M = 6,500$. The increase in the probability of autoimmunity being triggered if $M = 6,500$ is shown for various values of the quorum threshold, Q . As the value of the quorum threshold increases, the increase in the probability of autoimmunity decreases. Thus, the collective effects embodied in the quorum model makes the immune system more robust to inter-person variations in thymic development by lowering the change in the probability of triggering autoimmunity upon persistent infections due to such variations. We also find similar results if N_T , the number of self-peptides presented in a tissue, increases from 10 to 20 [Fig. 7B]. Our results in Figure 7 indicate that the majority of benefits of the robustness due to collective effects is realized at $Q \sim 10$. While our calculations are not meant to be quantitatively accurate, it is interesting to note that this value is consistent with experimental observations that the number of proximal activated T cells required for proliferation and differentiation is of the same order. [45][51].

The effector functions of activated T cells that target infected cells can lead to damage to healthy

tissue and the release of otherwise sequestered self-antigens into the local tissue, thus increasing the presentation of self-antigens [90, 91]. We wondered whether the collective effects inherent to the quorum model might also inhibit the increased chance of autoimmunity being triggered due to enhanced presentation of self-antigens. Figure 7 shows results of our simulations for the increase in probability of triggering autoimmunity as a function of Q when the number of self-antigens presented in the tissue (N_T) increases from $N_T = 10$ to $N_T = 20$. We find that increasing values of Q leads to a smaller increase in the risk of autoimmunity given an increase in self-antigen presentation. These results suggest that autoimmunity being triggered by epitope spreading and bystander activation [56] is reduced due to collective effects [Fig. 7]. Again, the influence of collective effects in suppressing autoimmunity upon strong perturbations diminishes once the quorum threshold is sufficiently high [Fig. 7]. It is important to note, however, that increasing N_T leads to a monotonically increasing risk of autoimmunity in spite of the collective effects embodied in the quorum model [Supp. 4]. This result shows that, in spite of the robustness conferred by the quorum model, substantial changes in the number of self-peptides presented in a tissue enhances the probability of triggering autoimmunity. So, if increased inflammation and tissue damage result in presentation of new self-antigens, autoimmunity would be more likely to be triggered upon infection.

The results described in this section suggest that collective effects embodied in the quorum model not only confer robustness to the immune system from the standpoint of discriminating between self and pathogen-derived antigens, but also serve to inhibit autoimmunity in spite of inter-person variations in thymic development and when strongly challenged such as upon severe or persistent infection.

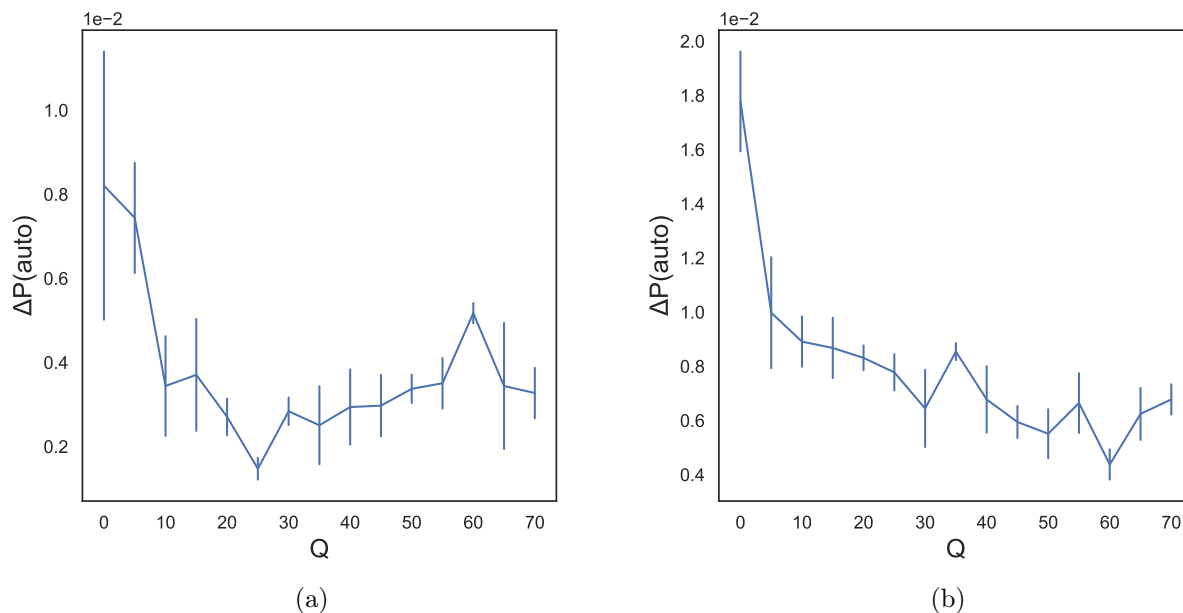


Figure 7: Collective effects in the quorum model confer robustness against autoimmunity. All simulations unless otherwise denoted were performed with the same parameters as in Table 1. (a) Simulations were performed with $M=8,000$ and $M=6,500$ at increasing values of Q at $N_f = 10$. The calculated change in the probability of triggering autoimmunity, $P(\text{auto})$, is graphed as a function of Q . (b) Simulations were performed with $N_T = 10$ and $N_T=20$ for increasing values of Q at $N_f = 10$. The calculated change in $P(\text{auto})$ is graphed as a function of Q .

The importance of rare, highly reactive pathogen-derived peptides for triggering autoimmunity

In order to obtain additional mechanistic insights, we first attempted to analyze the simulation results reported above in terms of a simple probabilistic model. The model is framed in terms of the following probabilities:

- q , the probability that a pathogen-derived peptide activates a T cell. Thymic selection against M peptides leads to $q \sim 1/M$ [37].
- q_s , the probability that a self-peptide activates a T cell. If M out of N possible self-peptides are seen in the thymus, $q_s = q(1 - M/N)$.
- $q_w = \alpha q_s$, the probability that a pathogen-activated T cell can be reactivated by a weakly crossreactive self-peptide. Cross-reactivity is thus encoded in the parameter $\alpha > 1$ that can be deduced from Figure 5.

An important consideration is how many of the T cells patrolling a tissue are on average activated when a self or pathogen-derived peptide is presented. These numbers are estimated using the parameters of Table 1, as $\bar{n} = Tq \approx 10^6 \times 0.04 \times (8,000)^{-1} \approx 5$ for pathogen-derived peptides, and $\bar{n}_s \approx \bar{n}(1 - 8,000/10,000) \approx 1$ for self-peptides. *Assuming that T cell activation events are independent* leads to a Poisson distribution for the probability of activating m T cells, and a probability of exceeding the quorum number Q of

$$P_Q(\bar{n}) = \sum_{m \geq Q} e^{-\bar{n}} \frac{\bar{n}^m}{m!} \approx e^{-\bar{n}} \frac{\bar{n}^Q}{Q!}. \quad (4)$$

The last approximation is correct as long as quorum number $Q > \bar{n}$ which holds for the choice of $Q = 10$. This leads to an estimate of a roughly 2% probability for the quorum threshold being exceeded by a pathogen-derived peptide (on average), and a probability of 10^{-7} for a self-peptide (obtained as $P_Q(\bar{n}_s)$).

However, the base value of autoimmunity in the absence of infection ($N_f = 0$) in Figure 6A is orders of magnitude higher than the 10^{-7} predicted by Eq. 4 for $P_Q(\bar{n}_s)$. This discrepancy can be traced back to the Poisson distribution not capturing the characteristics of the T cell repertoire. Figure 8A shows results of our simulations for the distribution of the number of T cells in the mature repertoire that are activated upon being challenged by our set of $N_f = 40,000$ pathogen-derived peptides and $N_s = 10,000$ self-peptides. These simulations were carried out with many mature T cell repertoires and choices of the self and pathogen-derived peptides. For both self and pathogen-derived antigens, the distribution of the number of activated T cells is very different from a Poisson distribution and is characterized by a long tail. That is, rare peptides can activate a large number of T cells. For example, a peptide composed of highly hydrophobic amino acids would be able to activate many T cells with TCRs containing similar moderately or poorly hydrophobic peptide contact residues. Thus, the activation of these T cells would be correlated and not independent, and the assumption of a Poisson distribution in Eq. 4 would be incorrect. This characteristic of rare peptides being able to activate many T cells can be illustrated with a much simpler model of TCR-peptide interactions than the string model.

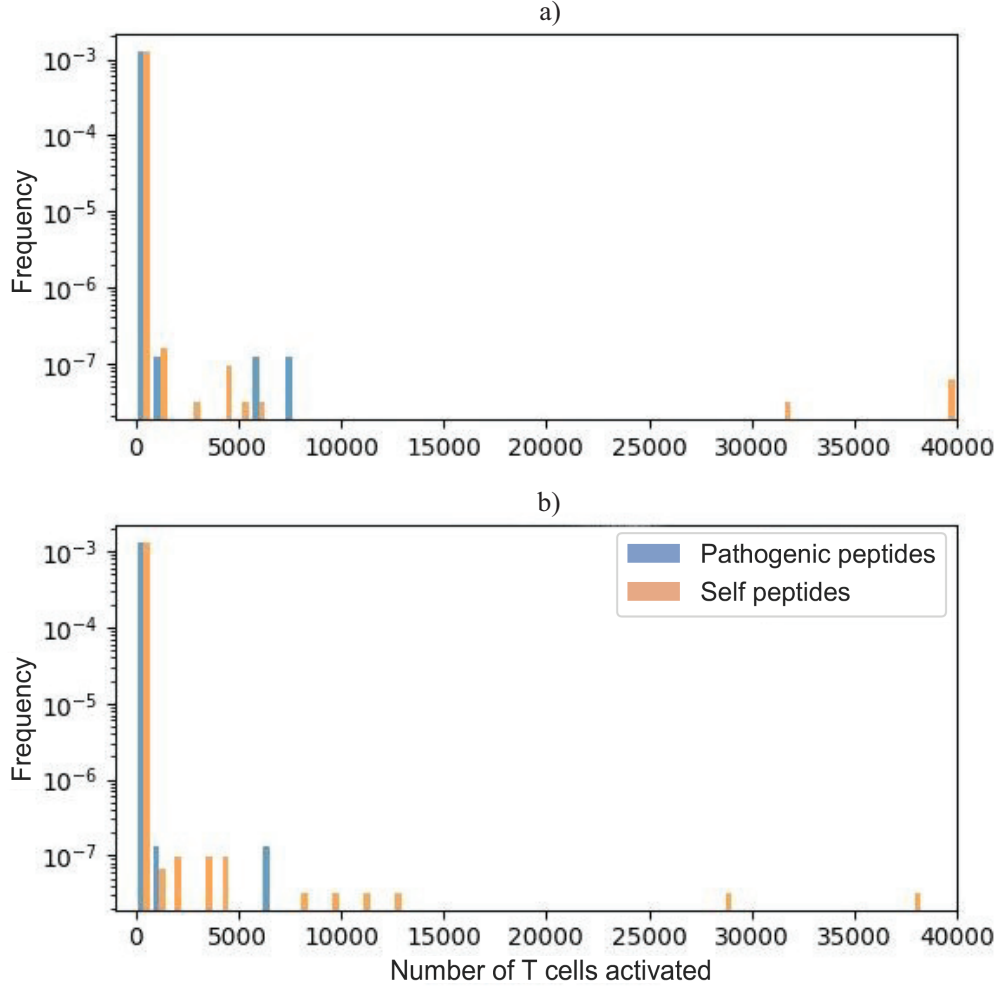


Figure 8: (a) Distribution of number of T cells activated by pathogen-derived and self-peptides as obtained from our simulations. Each simulation was conducted with $M = 8,000$ and $40,000$ pathogen-derived peptides, and $10,000$ self-peptides. $9,987$ self-peptides activated 0 T cells, 5 self-peptides activated between 1 and 10 , and 8 self-peptides activated 10 or more. $39,901$ pathogen-derived peptides activated 0 T cells, 43 activated between 1 and 10 , and 56 activated 10 or more. (b) Distribution of number of T cells activated by pathogen-derived and self-peptides as obtained from the simpler coarse-grained model described in the text. $9,995$ self-peptides activated 0 T cells, 1 self-peptide activated between 1 and 10 , and 4 self-peptides activated 10 or more. $39,983$ pathogen-derived peptides activated 0 T cells, 4 activated between 1 and 10 , and 13 activated 10 or more.

For a given peptide sequence j , in the string model we can construct a reactivity parameter β_j as the mean of the strength of its interactions with all possible TCR sequences. The ensemble of possible peptides can be characterized by the probability distribution $p(\beta)$ of reactivity. We can construct a similar distribution $p(\alpha)$ describing the mean strength of interactions of pre-selection thymocyte sequences (each sequence, i corresponds to α_i). To illustrate the role of reactivity, we consider a simplified model in which thymocyte sequences α_i and β_j are taken from the same Gaussian distribution, and their binding free energy is computed simply as

$$E_{ij} = \alpha_i + \beta_j. \quad (5)$$

Positive and negative selection prunes the initial TCR ensemble to create a post-selection set of TCRs characterized by a probability distribution, $p(\bar{\alpha}_i)$, which is concentrated in the range $\alpha_{\min} < \bar{\alpha}_i < \alpha_{\max} < 0$ [Supp. 5]. Possible values of β_j , however, are not similarly restricted, and it is possible (within the model) to encounter rare self-peptides β_j such that $\beta_j + \alpha_{\min} < E_n$; such peptides would then activate many T cells. We carried out calculations with this "coarse-grained" model with parameters close to that used in simulations with the "string" model [Supp. 5]. The resulting distributions for number of activated T cells are presented in Figure 8B. The distribution of T cell reactivity to peptides is similar to that obtained using the string model, including a tail of rare highly reactive peptides that activate many T cells that can overcome the quorum threshold. The chance of sampling such rare peptides from this tail of the distribution grows with the number of pathogen-derived peptides that are targeted as is the case during persistent or severe infections, but not usual acute infections. Such a tail of highly immunogenic antigens has also been observed in the context of SARS-CoV-2 infection. Out of a panel of HLA-I SARS-CoV-2 epitopes, 122 were considered immunogenic [92]. This is approximately 0.4% of total possible HLA-I restricted epitopes, which is comparable to the percentage of pathogen-derived peptides above the quorum number in our simulations (roughly 0.14%).

We emphasize that our goal is not quantitative recapitulation of known experimental facts. Our results are not quantitatively accurate as they are derived from simple models that aim to uncover new mechanistic insights that lead to experimentally testable predictions (see next section). For example, Figure 8 suggests that rare self-peptides would trigger autoimmunity by overcoming the quorum threshold, and so over long times, they should cause autoimmunity in everyone. But, this result is due to a limitation of our simple model. In our model the quorum threshold, Q , is fixed. In reality, T cells that bind to self-peptides in the thymus with long half-lives are likely to differentiate into Tregs [93]. Therefore, many more conventional effector T cells would have to be activated to beat the suppressive effects of more numerous Tregs that are activated by the same rare self-peptides; i.e., the quorum threshold is not a fixed number, but depends upon the T cell reactivity of the self-peptides. We also note that the Mirazawa-Jerningan model we have used for interactions between amino acids [76] is known to overemphasize the strength of hydrophobic interactions, which would result in an overestimate of the strength of TCR-pMHC interactions.

The main insight from the results presented in this section is that rare crossreactive self and pathogen-derived peptides may work together to trigger autoimmunity by exceeding the pertinent quorum thresholds. Persistent or severe infection makes the chance of sampling these peptides more likely.

Experimentally testable predictions of the model

It has been proposed that pathogen-derived peptides with homologs in the host proteome should be associated with higher likelihoods of triggering autoimmunity [94]. But, this notion does not account for the importance of collective effects embodied in the quorum model and the results we have described above. We propose immunizing mice with increasing numbers of pathogen-derived peptides that are homologous to peptides derived from the mouse proteome. The prediction of our model is that, as the number of such pathogen-derived peptide immunogens introduced into the animals increases, the chance of triggering autoimmunity should also increase.

To experimentally test this prediction, we have designed immunogens using the following approach. We propose immunizing C57BL/6 (B6) mice that express the I-Ab MHC molecule on their APCs. We used NetMHCIPan to find 15-mer I-Ab binding *mycobacterium tuberculosis* peptides and peptides from the mouse proteome, respectively [79, 95, 96]. Then, following past work [94], we

identified mouse homologs for each of the tuberculosis peptides by identifying the mouse peptides with 9-mer cores that shared the same amino acids in the 2,3,5,7, and 8 positions with the tuberculosis peptide’s 9-mer core. The number of mouse homologs per tuberculosis peptide ranged from 0 to 3,000 [Supp. 6]. These peptides that have similarity to the I-Ab binding mouse peptides are promising candidates for our proposed experiment, as we expect them to be more likely to activate cross-reactive T cells than peptides with low similarity to self-peptides. However, the pathogen-derived peptides should also be chosen to be different enough from self-peptides such that thymic selection would not have negatively selected against T cells that could recognize them with high probability. Based on these considerations, we chose tuberculosis peptides that are in neither of the tails of the similarity distribution (i.e., not too similar and not too distinct from self-peptides). We thus selected tuberculosis peptides that have 17 mouse homologs, which lies approximately at the mean of the similarity distribution shown in Supp. 6. These 122 peptides are listed in the Table provided in Supp. 6. Because we expect the rare peptides that activate many T cells to have more hydrophobic amino acids [22], we have also indicated the peptides in our list of homologs that have the most hydrophobic amino acids. Hydrophobicity was determined using multiple methods (Kyte-Doolittle [97], Hopp-Woods [98], Eisenburg [99], Rose [100], Engelman [101], and Wimley-White [102]) and essentially the same peptides were identified as being comprised of the most hydrophobic amino acids.

We suggest a protocol wherein mice are immunized with randomly picked tuberculosis peptides from our list. We predict that immunization with an increasing number of these peptides is more likely to trigger autoimmunity. We also anticipate that autoimmunity is more likely to be triggered upon immunizing with peptides from our list that we have identified as having more hydrophobic amino acids.

Discussion

In recent years, based on theoretical modeling, experiments in mice and in vitro, and analyses of sequences of T cell repertoires, evidence for the importance of collective effects in mediating T cell proliferation, differentiation and a functional response has accumulated [37, 44, 45, 10, 52]. These data suggest that a threshold number (or density) of T cells must be activated in a tissue in order for activated T cells to proliferate and mount an immune response. This threshold number, or quorum, is necessary to beat out the suppressive effects of peripheral tolerance mechanisms mediated by Tregs and perhaps the newly reported suppressor CD8+ T cells [37, 10]. The quorum threshold was postulated to be necessary for preventing autoimmunity due to circulating autoreactive T cells without preventing effective T cell responses to pathogens [37]. There is significant empirical evidence that some T cell mediated autoimmune diseases such as MS and T1D are triggered by persistent viral infections. The focus of this paper was to study how persistent or severe infections might cause the breakdown of the mechanism by which collective effects embodied in the quorum model suppresses autoimmunity.

We studied a computational model for T cell development in the thymus, and the response of the resulting T cell repertoire to infection. Upon infection, pathogen-derived peptides are targeted by T cells, and an effective functional response results because the quorum threshold is exceeded to some of these pathogen-derived peptides because they were not encountered by the T cell repertoire during development in the thymus. Our results show that some of the activated T cells that normally exhibit weak cross-reactivity to host-derived pMHC molecules can now be efficiently reactivated by these host antigens in the same tissue because previously activated T cells have a lower activation threshold. Evidence exists for such cross-reactivity between host pMHC molecules and

pathogen-derived ones in the case of MS. CD4⁺ T cells have been isolated that are cross-reactive to both EBV-derived peptides and self-antigens like myelin basic protein (MBP), anoctamin 2, alpha-crystallin B, and glial cell adhesion [66, 103, 88, 104, 105, 106]. For MBP, a self-antigen present in myelin sheaths, there is evidence that MS patients have autoreactive T cells that can be activated by APCs presenting EBV peptides [103]. As further evidence, not only is there an increase in the presence of MBP-specific CD8⁺ T cells in MS patients [107], but EBV-specific CD8⁺ T cells isolated from these patients were cytotoxic against cells that were pulsed with MBP and cells transfected to endogenously express MBP [108]. But, our results show that such "molecular mimicry" leads to autoimmunity with low probability for typical infections when just a few pathogen-derived peptides are targeted by the T cell repertoire. For autoimmunity to develop, the T cells activated by the pathogen-derived peptides that are crossreactive to a self-antigen must be sufficiently large in number such that, when added to other T cells that are activated by the same self-antigen, the quorum threshold is exceeded. When just a few pathogen-derived pMHC molecules are targeted, the probability of this happening is small. In other words, peripheral tolerance mechanisms that determine the quorum threshold are able to suppress autoimmunity upon typical infections. This concept is related to the observation that CD8⁺ suppressor T cells expand upon infection to kill activated autoreactive T cells [8] and shows that, if the balance is tilted toward Tregs, they expand and effector cells do not [10].

Our results also describe how this mechanism goes awry upon persistent or severe infection. In this case, larger numbers of pathogen-derived pMHC molecules are targeted by the T cell repertoire [64, 65]. The T cells that target each of these pMHC molecules has a chance of being weakly cross-reactive to a host pMHC molecule expressed in the same tissue. So, as the number of targeted pathogen-derived pMHC molecules increases, the total probability of activating T cells cross-reactive to the pathogen-derived pMHC and a self-antigen grows. Our results show that this, in turn, results in increasing the chance that the quorum threshold is exceeded by autoreactive T cells and T cells activated by pathogen-derived pMHC molecules that are weakly cross-reactive to a self-antigen. Thus, we find that the probability of autoimmunity being triggered grows with the number of pathogen-derived pMHC molecules targeted during persistent infection [Fig. 6].

Our model brings together the concepts of molecular mimicry and collective effects to provide a consistent explanation for why viral infections such as EBV do not always result in autoimmunity, why certain viral infections trigger certain autoimmune diseases, why there is a lag time between establishment of a persistent viral infection and onset of the corresponding autoimmune condition and why this lag time exhibits large variations. Only certain viruses will result in APCs displaying immunogenic pMHC molecules that activate T cells that are crossreactive to particular self-antigens. This is why certain viral infections are more likely to induce certain autoimmune conditions (e.g., EBV and MS). Our results show that these crossreactive T cells will contribute to exceeding the quorum threshold for a self-antigen to result in autoimmunity only with a finite probability; therefore, a persistent viral infection will only lead to a corresponding autoimmune disease only with some probability [Fig. 6]. One of our findings is the importance of rare peptides that can activate many T cells. If a pathogen-derived peptide triggers many T cells, the chance that T cells cross-reactive to a self-antigen are activated and contribute to exceeding the quorum threshold with respect to self-antigens also increases. Such events are rare [Figs. 6 and Supp. 5] and so autoimmunity may not occur for a very long time, if ever. However, the chance that such an event occurs increases with time during a persistent infection as the chance of T cells targeting more pathogen-derived peptides increases, thus providing an explanation for the lag time. As the lag time is dependent on a stochastic event occurring, it exhibits large variations across individuals.

Our results also show that collective effects embodied in the requirement that the quorum threshold must be exceeded for a functional response lowers the probability of triggering autoimmunity

because of variations in thymic selection across individuals, and when infection leads to strong perturbations [Fig. 7]. For example, we found that collective effects embodied in the quorum model make the immune system more robust to autoimmunity when the number of self-antigen T cells are exposed to in the thymus is lower. This situation can be realized either stochastically or due to certain HLA alleles, such as one associated with T1D binds self-peptides less stably. The increase in the probability of triggering autoimmunity upon T cells encountering fewer host-derived peptides in the thymus is reduced by the existence of collective effects and the quorum threshold [Fig. 7A]. However, the influence of collective effects in suppressing autoimmunity does not matter much beyond a threshold value of the quorum threshold, Q [Fig. 7]. The quorum number, Q , is set by the number (or density) of Tregs and other suppressor cells. Thus, this result from our model suggests that beyond a point having more Tregs will not add to robustness of the immune system against triggering autoimmunity.

Inflammation and T cell-mediated cell killing leads to tissue damage, which promotes presentation of self-epitopes that would otherwise not be presented. It has been postulated in the case of T1D that, through this mechanism, also known as epitope spreading, self-reactive T cells are recruited and can initiate an autoimmune response [109]. For example, persistent EV infection leads to inflammation and destruction of islet cells by T cells [110] and an increase in inflammatory cytokines (IFN) [111]. Increase in IFN leads to HLA class 1 over-expression in islet cells and cell death. These factors contribute to increased vulnerability to $CD8^+$ T cell bystander activation and targeting of novel self-antigens [112]. More generally, inflammation leads to various effects, such as enhanced crosspriming, etc, that can result in the presentation of self-antigens that are typically “hidden”. This concept of epitope spreading contributing to autoimmunity has also been noted in the context of MS. EBV has been suggested to induce an antiviral immune response against infected cells in the CNS which leads to the release of sequestered self-antigens [113]. In addition, it has also been suggested that EBV-infected autoreactive B cells mediate an autoimmune attack by increasing self-antigen presentation and activating autoreactive $CD4^+$ T helper cells that can activate these B cells [114]. In the context of our model, the increased presentation of self-peptides is reflected as an increase in N_T . Our results show that increasing N_T increases the number of T cells activated by self-antigens [Supp. 4], thus contributing to exceeding the quorum threshold. However, consistent with the role of collective effects making the immune system more robust against onset autoimmunity, our results show that the effect of epitope spreading on triggering autoimmunity is mitigated by the requirement that the quorum threshold be exceeded for a functional immune or autoimmune response [Fig. 7B]. Note also that inflammation and more cell death has a higher chance of resulting in expression of rare self-pMHC molecules that can activate many T cells [Fig. 8], thus exceeding the quorum threshold.

The results we have described provide a unifying framework to understand various phenomena, but the model requires further testing. We make a specific prediction to test the veracity of our model. We carried out calculations to identify peptides derived from the proteome of *mycobacterium tuberculosis* that are homologs of peptides from the proteome of C57BL/6 (B6) mice; both sets of peptides bind to the I-Ab MHC molecule expressed by these mice. We predict that immunizing these mice with increasing numbers of the identified *mycobacterium tuberculosis* peptides should result in a higher chance of inducing autoimmune disease. We also expect that the homologous *mycobacterium tuberculosis* peptides that have more hydrophobic amino acids at the sites that contact the TCR should serve as proxies for rare peptides that trigger more T cells, and so we also identify these peptides. Immunizing with a larger number of homolog *mycobacterium tuberculosis* peptides that include these more hydrophobic ones is predicted to increase the probability of triggering autoimmunity. We hope that these experiments will soon be done as they will likely shed new light on the aberrant regulation of peripheral tolerance mechanisms.

We emphasize again that the purpose of our model is not quantitative recapitulation of a specific experimental result. Rather, our goal was to develop a model that could explore the significant consequences of collective effects that are important for a functional T cell response on mitigating the danger of autoimmunity, and how these tolerance mechanisms can be mis-regulated upon persistent or severe infections. Our model has several limitations. In this manuscript, we have focused on the initial stages of T cell activation and proliferation. It is recognized that the affinity of self-reactive T cell responses also determines whether activated clones upregulate tissue homing integrins and induce MHC-I cross-priming of self-antigens [115, 22]. As details of these processes are further quantified, second-generation computational models will incorporate these attributes to better understand when overt autoimmunity ensues, or a given proliferative burst simply expands self-reactive clones. We do not explicitly treat the effects of innate immunity and inflammation, and implicitly treat their effects on increasing self-antigen presentation by increasing the parameter in our model (N_T) that represents the number of self-antigens presented in the infected tissue. We do not treat Tregs and suppressor CD8+ T cells explicitly either. Their effects are embodied in the quorum threshold number, Q . As we have noted earlier, the value of Q is not going to be a constant as we have assumed, but rather it will be larger for self-peptides that bind more avidly to T cells as interactions with these self-peptides in the thymus are more likely to result in the development of Tregs [93]. Accounting for Tregs and suppressor CD8+ T cells and concomitant differences in the values of Q for different peptides is an important next step for developing a better understanding of the dynamic instability that results in breaking the quorum threshold and initiating autoimmunity. This will require the development of a dynamical version of our model of infection, which is an important next step.

Github

Link to github with source code and tuberculosis peptide files: https://www.github.com/pinkzephyr/autoimmune_model

Acknowledgements

This research was supported by the Ragon Institute of MGH, MIT, and Harvard. RY was also supported by a National Science Foundation Graduate Research Fellowship. MK acknowledges support from NSF grant DMR-2218849. SM was supported by a Physics of Living Systems Fellowship at MIT. We are very grateful for fruitful discussions with Marc Jenkins, Shiv Pillai, and Ruslan Medhizitov.

Competing Interests

There are no competing interests. However, for completeness, it is noted that A.K.C. is a consultant (titled Academic Partner) for Flagship Pioneering, serves as consultant and on the Strategic Oversight Board of its affiliated company, Apriori Bio, and is a consultant and SAB member of another affiliated company, Metaphore Bio.

References

- [1] N Hozumi and S Tonegawa. “Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions.” In: *Proceedings of the National Academy of Sciences* 73.10 (Oct. 1976). Publisher: Proceedings of the National Academy of Sciences, pp. 3628–3632. DOI: [10.1073/pnas.73.10.3628](https://doi.org/10.1073/pnas.73.10.3628). URL: <https://www.pnas.org/doi/10.1073/pnas.73.10.3628> (visited on 06/17/2023).
- [2] David G. Schatz, Marjorie A. Oettinger, and David Baltimore. “The V(D)J recombination activating gene, RAG-1”. In: *Cell* 59.6 (Dec. 22, 1989). Publisher: Elsevier, pp. 1035–1048. ISSN: 0092-8674, 1097-4172. DOI: [10.1016/0092-8674\(89\)90760-5](https://doi.org/10.1016/0092-8674(89)90760-5). URL: [https://www.cell.com/cell/abstract/0092-8674\(89\)90760-5](https://www.cell.com/cell/abstract/0092-8674(89)90760-5) (visited on 06/17/2023).
- [3] Yuval Elhanati et al. “Quantifying selection in immune receptor repertoires”. In: *Proceedings of the National Academy of Sciences of the United States of America* 111.27 (July 8, 2014), pp. 9875–9880. ISSN: 0027-8424. DOI: [10.1073/pnas.1409572111](https://doi.org/10.1073/pnas.1409572111). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4103359/> (visited on 06/17/2023).
- [4] Arup K. Chakraborty and Arthur Weiss. “Insights into the initiation of TCR signaling”. In: *Nature Immunology* 15.9 (Sept. 2014). Number: 9 Publisher: Nature Publishing Group, pp. 798–807. ISSN: 1529-2916. DOI: [10.1038/ni.2940](https://doi.org/10.1038/ni.2940). URL: <https://www.nature.com/articles/ni.2940> (visited on 04/10/2023).
- [5] Christopher Szeto et al. “Covalent TCR-peptide-MHC interactions induce T cell activation and redirect T cell fate in the thymus”. In: *Nature Communications* 13.1 (Aug. 23, 2022). Number: 1 Publisher: Nature Publishing Group, p. 4951. ISSN: 2041-1723. DOI: [10.1038/s41467-022-32692-4](https://doi.org/10.1038/s41467-022-32692-4). URL: <https://www.nature.com/articles/s41467-022-32692-4> (visited on 01/18/2023).
- [6] Pablo A. González et al. “T cell receptor binding kinetics required for T cell activation depend on the density of cognate ligand on the antigen-presenting cell”. In: *Proceedings of the National Academy of Sciences* 102.13 (Mar. 29, 2005). Publisher: Proceedings of the National Academy of Sciences, pp. 4824–4829. DOI: [10.1073/pnas.0500922102](https://doi.org/10.1073/pnas.0500922102). URL: <https://www.pnas.org/doi/10.1073/pnas.0500922102> (visited on 01/18/2023).
- [7] Steven Z. Josefowicz, Li-Fan Lu, and Alexander Y. Rudensky. “Regulatory T Cells: Mechanisms of Differentiation and Function”. In: *Annual review of immunology* 30 (2012), pp. 531–564. ISSN: 0732-0582. DOI: [10.1146/annurev.immunol.25.022106.141623](https://doi.org/10.1146/annurev.immunol.25.022106.141623). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6066374/> (visited on 04/24/2023).
- [8] Jing Li et al. “KIR+CD8+ T cells suppress pathogenic T cells and are active in autoimmune diseases and COVID-19”. In: *Science* 376.6590 (Mar. 8, 2022). Publisher: American Association for the Advancement of Science, eabi9591. DOI: [10.1126/science.abi9591](https://doi.org/10.1126/science.abi9591). URL: <https://www.science.org/doi/10.1126/science.abi9591> (visited on 05/11/2022).
- [9] Ludger Klein et al. “Positive and negative selection of the T cell repertoire: what thymocytes see (and don’t see)”. In: *Nature Reviews Immunology* 14.6 (June 2014). Number: 6 Publisher: Nature Publishing Group, pp. 377–391. ISSN: 1474-1741. DOI: [10.1038/nri3667](https://doi.org/10.1038/nri3667). URL: <https://www.nature.com/articles/nri3667> (visited on 01/18/2023).
- [10] Harikesh S. Wong et al. “A local regulatory T cell feedback circuit maintains immune homeostasis by pruning self-activated T cells”. In: *Cell* 184.15 (July 22, 2021), 3981–3997.e22. ISSN: 1097-4172. DOI: [10.1016/j.cell.2021.05.028](https://doi.org/10.1016/j.cell.2021.05.028).

- [11] Diane Mathis and Christophe Benoist. “Aire”. In: *Annual Review of Immunology* 27 (2009), pp. 287–312. ISSN: 0732-0582. DOI: [10.1146/annurev.immunol.25.022106.141532](https://doi.org/10.1146/annurev.immunol.25.022106.141532).
- [12] Gerald P. Morris and Paul M. Allen. “How the TCR balances sensitivity and specificity for the recognition of self and pathogens”. In: *Nature Immunology* 13.2 (Jan. 19, 2012), pp. 121–128. ISSN: 1529-2916. DOI: [10.1038/ni.2190](https://doi.org/10.1038/ni.2190).
- [13] Harald von Boehmer and Pawel Kisielow. “Self-Nonself Discrimination by T Cells”. In: *Science* 248.4961 (June 15, 1990). Publisher: American Association for the Advancement of Science, pp. 1369–1373. DOI: [10.1126/science.1972594](https://doi.org/10.1126/science.1972594). URL: <https://www.science.org/doi/abs/10.1126/science.1972594> (visited on 06/17/2023).
- [14] Philip G. Ashton-Rickardt et al. “Peptide contributes to the specificity of positive selection of CD8+ T cells in the thymus”. In: *Cell* 73.5 (June 4, 1993), pp. 1041–1049. ISSN: 0092-8674. DOI: [10.1016/0092-8674\(93\)90281-T](https://doi.org/10.1016/0092-8674(93)90281-T). URL: <https://www.sciencedirect.com/science/article/pii/S009286749390281T> (visited on 06/17/2023).
- [15] K. A. Hogquist et al. “T cell receptor antagonist peptides induce positive selection”. In: *Cell* 76.1 (Jan. 14, 1994), pp. 17–27. ISSN: 0092-8674. DOI: [10.1016/0092-8674\(94\)90169-4](https://doi.org/10.1016/0092-8674(94)90169-4).
- [16] Eric S. Huseby, John W. Kappler, and Philippa Marrack. “Thymic selection stifles TCR reactivity with the main chain structure of MHC and forces interactions with the peptide side chains”. In: *Molecular Immunology* 45.3 (Feb. 2008), pp. 599–606. ISSN: 0161-5890. DOI: [10.1016/j.molimm.2006.03.025](https://doi.org/10.1016/j.molimm.2006.03.025).
- [17] Eric S. Huseby et al. “How the T Cell Repertoire Becomes Peptide and MHC Specific”. In: *Cell* 122.2 (July 29, 2005), pp. 247–260. ISSN: 0092-8674. DOI: [10.1016/j.cell.2005.05.013](https://doi.org/10.1016/j.cell.2005.05.013). URL: <http://www.sciencedirect.com/science/article/pii/S0092867405004599> (visited on 11/05/2020).
- [18] Eric S. Huseby et al. “Interface-disrupting amino acids establish specificity between T cell receptors and complexes of major histocompatibility complex and peptide”. In: *Nature Immunology* 7.11 (Nov. 2006). Number: 11 Publisher: Nature Publishing Group, pp. 1191–1199. ISSN: 1529-2916. DOI: [10.1038/ni1401](https://doi.org/10.1038/ni1401). URL: <https://www.nature.com/articles/ni1401> (visited on 05/24/2023).
- [19] Andrej Košmrlj et al. “How the thymus designs antigen-specific and self-tolerant T cell receptor sequences”. In: *Proceedings of the National Academy of Sciences* 105.43 (Oct. 28, 2008), pp. 16671–16676. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.0808081105](https://doi.org/10.1073/pnas.0808081105). URL: <https://www.pnas.org/content/105/43/16671> (visited on 01/15/2020).
- [20] Andrej Kosmrlj et al. “Thymic selection of T-cell receptors as an extreme value problem”. In: *Physical Review Letters* 103.6 (Aug. 7, 2009), p. 068103. ISSN: 0031-9007. DOI: [10.1103/PhysRevLett.103.068103](https://doi.org/10.1103/PhysRevLett.103.068103).
- [21] Andrej Košmrlj et al. “Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection”. In: *Nature* 465.7296 (May 2010), pp. 350–354. ISSN: 1476-4687. DOI: [10.1038/nature08997](https://doi.org/10.1038/nature08997). URL: <https://www.nature.com/articles/nature08997> (visited on 01/15/2020).
- [22] Brian D. Stadinski et al. “Hydrophobic CDR3 residues promote the development of self-reactive T cells”. In: *Nature immunology* 17.8 (Aug. 2016), pp. 946–955. ISSN: 1529-2908. DOI: [10.1038/ni.3491](https://doi.org/10.1038/ni.3491). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4955740/> (visited on 11/05/2020).

- [23] Hugo A. van den Berg and David A. Rand. “Dynamics of T cell activation threshold tuning”. In: *Journal of Theoretical Biology* 228.3 (June 7, 2004), pp. 397–416. ISSN: 0022-5193. DOI: [10.1016/j.jtbi.2004.02.002](https://doi.org/10.1016/j.jtbi.2004.02.002). URL: <https://www.sciencedirect.com/science/article/pii/S0022519304000517> (visited on 01/19/2023).
- [24] Kohei Ota et al. “T-cell recognition of an immuno-dominant myelin basic protein epitope in multiple sclerosis”. In: *Nature* 346.6280 (July 1990). Number: 6280 Publisher: Nature Publishing Group, pp. 183–187. ISSN: 1476-4687. DOI: [10.1038/346183a0](https://doi.org/10.1038/346183a0). URL: <https://www.nature.com/articles/346183a0> (visited on 04/10/2023).
- [25] Yuka Maeda et al. “Detection of self-reactive CD8+ T cells with an anergic phenotype in healthy individuals”. In: *Science* 346.6216 (Dec. 19, 2014). Publisher: American Association for the Advancement of Science, pp. 1536–1540. DOI: [10.1126/science.aaa1292](https://doi.org/10.1126/science.aaa1292). URL: <https://www.science.org/doi/10.1126/science.aaa1292> (visited on 04/10/2023).
- [26] Laura F. Su et al. “Virus-Specific CD4+ Memory-Phenotype T Cells Are Abundant in Unexposed Adults”. In: *Immunity* 38.2 (Feb. 2013), pp. 373–383. ISSN: 10747613. DOI: [10.1016/j.immuni.2012.10.021](https://doi.org/10.1016/j.immuni.2012.10.021). URL: <https://linkinghub.elsevier.com/retrieve/pii/S1074761313000526> (visited on 04/10/2023).
- [27] Wong Yu et al. “Clonal Deletion Prunes but Does Not Eliminate Self-Specific CD8+ T Lymphocytes”. In: *Immunity* 42.5 (May 19, 2015), pp. 929–941. ISSN: 1074-7613. DOI: [10.1016/j.immuni.2015.05.001](https://doi.org/10.1016/j.immuni.2015.05.001). URL: <https://www.sciencedirect.com/science/article/pii/S1074761315001818> (visited on 04/10/2023).
- [28] Deepali Malhotra et al. “Tolerance is established in polyclonal CD4+ T cells by distinct mechanisms, according to self-peptide expression patterns”. In: *Nature Immunology* 17.2 (Feb. 2016). Number: 2 Publisher: Nature Publishing Group, pp. 187–195. ISSN: 1529-2916. DOI: [10.1038/ni.3327](https://doi.org/10.1038/ni.3327). URL: <https://www.nature.com/articles/ni.3327> (visited on 04/24/2023).
- [29] Shohei Hori, Takeshi Takahashi, and Shimon Sakaguchi. “Control of autoimmunity by naturally arising regulatory CD4+ T cells”. In: *Advances in immunology* 81 (Jan. 1, 2003), pp. 331–371. ISSN: 1557-8445. DOI: [10.1016/s0065-2776\(03\)81008-8](https://doi.org/10.1016/s0065-2776(03)81008-8). URL: [https://doi.org/10.1016/s0065-2776\(03\)81008-8](https://doi.org/10.1016/s0065-2776(03)81008-8) (visited on 08/10/2023).
- [30] Kajsa Wing and Shimon Sakaguchi. “Regulatory T cells exert checks and balances on self tolerance and autoimmunity”. In: *Nature Immunology* 11.1 (Jan. 2010). Number: 1 Publisher: Nature Publishing Group, pp. 7–13. ISSN: 1529-2916. DOI: [10.1038/ni.1818](https://doi.org/10.1038/ni.1818). URL: <https://www.nature.com/articles/ni.1818> (visited on 04/10/2023).
- [31] Jens Geginat, Federica Sallusto, and Antonio Lanzavecchia. “Cytokine-driven Proliferation and Differentiation of Human Naive, Central Memory, and Effector Memory CD4+ T Cells”. In: *The Journal of Experimental Medicine* 194.12 (Dec. 17, 2001), pp. 1711–1720. ISSN: 0022-1007. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2193568/> (visited on 04/10/2023).
- [32] Ananda W. Goldrath et al. “Cytokine requirements for acute and Basal homeostatic proliferation of naive and memory CD8+ T cells”. In: *The Journal of Experimental Medicine* 195.12 (June 17, 2002), pp. 1515–1522. ISSN: 0022-1007. DOI: [10.1084/jem.20020033](https://doi.org/10.1084/jem.20020033).
- [33] K. A. Smith. “Interleukin-2: inception, impact, and implications”. In: *Science (New York, N.Y.)* 240.4856 (May 27, 1988), pp. 1169–1176. ISSN: 0036-8075. DOI: [10.1126/science.3131876](https://doi.org/10.1126/science.3131876).

- [34] Sarah H. Ross and Doreen A. Cantrell. “Signaling and Function of Interleukin-2 in T Lymphocytes”. In: *Annual review of immunology* 36 (Apr. 26, 2018), pp. 411–433. ISSN: 0732-0582. DOI: [10.1146/annurev-immunol-042617-053352](https://doi.org/10.1146/annurev-immunol-042617-053352). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6472684/> (visited on 04/10/2023).
- [35] Ethan M. Shevach. “Mechanisms of Foxp3+ T Regulatory Cell-Mediated Suppression”. In: *Immunity* 30.5 (May 22, 2009). Publisher: Elsevier, pp. 636–645. ISSN: 1074-7613. DOI: [10.1016/j.immuni.2009.04.010](https://doi.org/10.1016/j.immuni.2009.04.010). URL: [https://www.cell.com/immunity/abstract/S1074-7613\(09\)00197-6](https://www.cell.com/immunity/abstract/S1074-7613(09)00197-6) (visited on 04/24/2023).
- [36] Zhiduo Liu et al. “Immune homeostasis enforced by co-localized effector and regulatory T cells”. In: *Nature* 528.7581 (Dec. 2015). Number: 7581 Publisher: Nature Publishing Group, pp. 225–230. ISSN: 1476-4687. DOI: [10.1038/nature16169](https://doi.org/10.1038/nature16169). URL: <https://www.nature.com/articles/nature16169> (visited on 04/24/2023).
- [37] Thomas Charles Butler, Mehran Kardar, and Arup K. Chakraborty. “Quorum sensing allows T cells to discriminate between self and nonself”. In: *Proceedings of the National Academy of Sciences* 110.29 (July 16, 2013), p. 11833. DOI: [10.1073/pnas.1222467110](https://doi.org/10.1073/pnas.1222467110). URL: <http://www.pnas.org/content/110/29/11833.abstract>.
- [38] Melissa B. Miller and Bonnie L. Bassler. “Quorum Sensing in Bacteria”. In: *Annual Review of Microbiology* 55.1 (2001). _eprint: <https://doi.org/10.1146/annurev.micro.55.1.165>, pp. 165–199. DOI: [10.1146/annurev.micro.55.1.165](https://doi.org/10.1146/annurev.micro.55.1.165). URL: <https://doi.org/10.1146/annurev.micro.55.1.165> (visited on 04/10/2023).
- [39] Philippe Bousso and Ellen Robey. “Dynamics of CD8+ T cell priming by dendritic cells in intact lymph nodes”. In: *Nature Immunology* 4.6 (June 2003). Number: 6 Publisher: Nature Publishing Group, pp. 579–585. ISSN: 1529-2916. DOI: [10.1038/ni928](https://doi.org/10.1038/ni928). URL: <https://www.nature.com/articles/ni928> (visited on 06/17/2023).
- [40] Flora Castellino et al. “Chemokines enhance immunity by guiding naive CD8+ T cells to sites of CD4+ T cell–dendritic cell interaction”. In: *Nature* 440.7086 (Apr. 2006). Number: 7086 Publisher: Nature Publishing Group, pp. 890–895. ISSN: 1476-4687. DOI: [10.1038/nature04651](https://doi.org/10.1038/nature04651). URL: <https://www.nature.com/articles/nature04651> (visited on 06/17/2023).
- [41] Stéphanie Hugues et al. “Dynamic imaging of chemokine-dependent CD8+ T cell help for CD8+ T cell responses”. In: *Nature Immunology* 8.9 (Sept. 2007). Number: 9 Publisher: Nature Publishing Group, pp. 921–930. ISSN: 1529-2916. DOI: [10.1038/ni1495](https://doi.org/10.1038/ni1495). URL: <https://www.nature.com/articles/ni1495> (visited on 06/17/2023).
- [42] Catherine A. Sabatos et al. “A Synaptic Basis for Paracrine Interleukin-2 Signaling during Homotypic T Cell Interaction”. In: *Immunity* 29.2 (Aug. 15, 2008), pp. 238–248. ISSN: 1074-7613. DOI: [10.1016/j.immuni.2008.05.017](https://doi.org/10.1016/j.immuni.2008.05.017). URL: <https://www.sciencedirect.com/science/article/pii/S1074761308003166> (visited on 06/17/2023).
- [43] Kathryn E. Foulds and Hao Shen. “Clonal Competition Inhibits the Proliferation and Differentiation of Adoptively Transferred TCR Transgenic CD4 T Cells in Response to Infection1”. In: *The Journal of Immunology* 176.5 (Mar. 1, 2006), pp. 3037–3043. ISSN: 0022-1767. DOI: [10.4049/jimmunol.176.5.3037](https://doi.org/10.4049/jimmunol.176.5.3037). URL: <https://doi.org/10.4049/jimmunol.176.5.3037> (visited on 06/17/2023).

- [44] Audrey Gérard et al. “Secondary T cell–T cell synaptic interactions drive the differentiation of protective CD8+ T cells”. In: *Nature Immunology* 14.4 (Apr. 2013). Number: 4 Publisher: Nature Publishing Group, pp. 356–363. ISSN: 1529-2916. DOI: [10.1038/ni.2547](https://doi.org/10.1038/ni.2547). URL: <https://www.nature.com/articles/ni.2547> (visited on 06/07/2023).
- [45] Simon Zenke et al. “Quorum Regulation via Nested Antagonistic Feedback Circuits Mediated by the Receptors CD28 and CTLA-4 Confers Robustness to T Cell Population Dynamics”. In: *Immunity* 52.2 (2020). tex.ids= zenke2020a, 313–327.e7. ISSN: 1097-4180. DOI: [10.1016/j.immuni.2020.01.018](https://doi.org/10.1016/j.immuni.2020.01.018).
- [46] James D. Fraser et al. “Regulation of Interleukin-2 Gene Enhancer Activity by the T Cell Accessory Molecule CD28”. In: *Science* 251.4991 (Jan. 18, 1991). Publisher: American Association for the Advancement of Science, pp. 313–316. DOI: [10.1126/science.1846244](https://doi.org/10.1126/science.1846244). URL: <https://www.science.org/doi/10.1126/science.1846244> (visited on 07/04/2023).
- [47] Carl H. June et al. “T-Cell Proliferation Involving the CD28 Pathway Is Associated with Cyclosporine-Resistant Interleukin 2 Gene Expression”. In: *Molecular and Cellular Biology* 7.12 (Dec. 1, 1987). Publisher: Taylor & Francis _eprint: <https://doi.org/10.1128/mcb.7.12.4472-4481.1987>, pp. 4472–4481. ISSN: null. DOI: [10.1128/mcb.7.12.4472-4481.1987](https://doi.org/10.1128/mcb.7.12.4472-4481.1987). URL: <https://doi.org/10.1128/mcb.7.12.4472-4481.1987> (visited on 07/04/2023).
- [48] C B Thompson et al. “CD28 activation pathway regulates the production of multiple T-cell-derived lymphokines/cytokines.” In: *Proceedings of the National Academy of Sciences* 86.4 (Feb. 1989). Publisher: Proceedings of the National Academy of Sciences, pp. 1333–1337. DOI: [10.1073/pnas.86.4.1333](https://doi.org/10.1073/pnas.86.4.1333). URL: <https://www.pnas.org/doi/abs/10.1073/pnas.86.4.1333> (visited on 07/04/2023).
- [49] M K Jenkins et al. “CD28 delivers a costimulatory signal involved in antigen-specific IL-2 production by human T cells.” In: *The Journal of Immunology* 147.8 (Oct. 15, 1991), pp. 2461–2466. ISSN: 0022-1767. DOI: [10.4049/jimmunol.147.8.2461](https://doi.org/10.4049/jimmunol.147.8.2461). URL: <https://doi.org/10.4049/jimmunol.147.8.2461> (visited on 07/04/2023).
- [50] Yongqing Wu et al. “FOXP3 controls regulatory T cell function through cooperation with NFAT”. In: *Cell* 126.2 (July 28, 2006), pp. 375–387. ISSN: 0092-8674. DOI: [10.1016/j.cell.2006.05.042](https://doi.org/10.1016/j.cell.2006.05.042).
- [51] Michal Polonsky et al. “Induction of CD4 T cell memory by local cellular collectivity”. In: *Science* 360.6394 (June 15, 2018). Publisher: American Association for the Advancement of Science, eaaj1853. DOI: [10.1126/science.aaj1853](https://doi.org/10.1126/science.aaj1853). URL: <https://www.science.org/doi/full/10.1126/science.aaj1853> (visited on 05/19/2022).
- [52] Francesco Camaglia et al. *Population based selection shapes the T cell receptor repertoire during thymic development*. Pages: 2022.02.14.480309 Section: New Results. Feb. 15, 2022. DOI: [10.1101/2022.02.14.480309](https://doi.org/10.1101/2022.02.14.480309). URL: <https://www.biorxiv.org/content/10.1101/2022.02.14.480309v1> (visited on 11/29/2022).
- [53] Sarah J Richardson and Noel G Morgan. “Enteroviral infections in the pathogenesis of type 1 diabetes: new insights for therapeutic intervention”. In: *Current Opinion in Pharmacology. • Endocrine & Metabolic Diseases • Gastrointestinal* 43 (Dec. 1, 2018), pp. 11–19. ISSN: 1471-4892. DOI: [10.1016/j.coph.2018.07.006](https://doi.org/10.1016/j.coph.2018.07.006). URL: <https://www.sciencedirect.com/science/article/pii/S1471489217302205> (visited on 01/18/2023).

- [54] Juliana Casqueiro, Janine Casqueiro, and Cresio Alves. “Infections in patients with diabetes mellitus: A review of pathogenesis”. In: *Indian Journal of Endocrinology and Metabolism* 16 (Suppl1 Mar. 2012), S27–S36. ISSN: 2230-8210. DOI: [10.4103/2230-8210.94253](https://doi.org/10.4103/2230-8210.94253). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3354930/> (visited on 01/18/2023).
- [55] Magloire Pandoua Nekoua, Enagnon Kazali Alidjinou, and Didier Hober. “Persistent coxsackievirus B infection and pathogenesis of type 1 diabetes mellitus”. In: *Nature Reviews Endocrinology* 18.8 (Aug. 2022). Number: 8 Publisher: Nature Publishing Group, pp. 503–516. ISSN: 1759-5037. DOI: [10.1038/s41574-022-00688-1](https://doi.org/10.1038/s41574-022-00688-1). URL: <https://www.nature.com/articles/s41574-022-00688-1> (visited on 01/18/2023).
- [56] Robert S. Fujinami et al. “Molecular Mimicry, Bystander Activation, or Viral Persistence: Infections and Autoimmune Disease”. In: *Clinical Microbiology Reviews* 19.1 (Jan. 2006), pp. 80–94. ISSN: 0893-8512. DOI: [10.1128/CMR.19.1.80-94.2006](https://doi.org/10.1128/CMR.19.1.80-94.2006). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1360274/> (visited on 01/18/2023).
- [57] Gunnar Houen and Nicole Hartwig Trier. “Epstein-Barr Virus and Systemic Autoimmune Diseases”. In: *Frontiers in Immunology* 11 (2021). ISSN: 1664-3224. URL: <https://www.frontiersin.org/articles/10.3389/fimmu.2020.587380> (visited on 01/18/2023).
- [58] Kjetil Bjornevik et al. “Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis”. In: *Science* 375.6578 (Jan. 21, 2022). Publisher: American Association for the Advancement of Science, pp. 296–301. DOI: [10.1126/science.abj8222](https://doi.org/10.1126/science.abj8222). URL: <https://www.science.org/doi/10.1126/science.abj8222> (visited on 03/20/2023).
- [59] Marc S. Horwitz et al. “Diabetes induced by Coxsackie virus: Initiation by bystander damage and not molecular mimicry”. In: *Nature Medicine* 4.7 (July 1998). Number: 7 Publisher: Nature Publishing Group, pp. 781–785. ISSN: 1546-170X. DOI: [10.1038/nm0798-781](https://doi.org/10.1038/nm0798-781). URL: <https://www.nature.com/articles/nm0798-781> (visited on 03/17/2023).
- [60] Balqis Okta Putry et al. “Can SARS-CoV-2 trigger new onset of autoimmune disease in adults? A case-based review”. In: *Heliyon* 8.11 (Nov. 1, 2022). Publisher: Elsevier. ISSN: 2405-8440. DOI: [10.1016/j.heliyon.2022.e11328](https://doi.org/10.1016/j.heliyon.2022.e11328). URL: [https://www.cell.com/heliyon/abstract/S2405-8440\(22\)02616-0](https://www.cell.com/heliyon/abstract/S2405-8440(22)02616-0) (visited on 11/29/2022).
- [61] Elena Ortona and Walter Malorni. “Long COVID: to investigate immunological mechanisms and sex/gender related aspects as fundamental steps for tailored therapy”. In: *The European Respiratory Journal* 59.2 (Feb. 3, 2022), p. 2102245. ISSN: 0903-1936. DOI: [10.1183/13993003.02245-2021](https://doi.org/10.1183/13993003.02245-2021). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8462012/> (visited on 01/18/2023).
- [62] Wassim Chehadeh et al. “Persistent Infection of Human Pancreatic Islets by Coxsackievirus B Is Associated with Alpha Interferon Synthesis in Cells”. In: *Journal of Virology* 74.21 (Nov. 2000). Publisher: American Society for Microbiology, pp. 10153–10164. DOI: [10.1128/JVI.74.21.10153-10164.2000](https://doi.org/10.1128/JVI.74.21.10153-10164.2000). URL: <https://journals.asm.org/doi/10.1128/jvi.74.21.10153-10164.2000> (visited on 03/21/2023).
- [63] Hsiao-Chuan Lin et al. “Enterovirus infection is associated with an increased risk of childhood type 1 diabetes in Taiwan: a nationwide population-based cohort study”. In: *Diabetologia* 58.1 (Jan. 1, 2015), pp. 79–86. ISSN: 1432-0428. DOI: [10.1007/s00125-014-3400-z](https://doi.org/10.1007/s00125-014-3400-z). URL: <https://doi.org/10.1007/s00125-014-3400-z> (visited on 03/20/2023).

- [64] Yasmine Belkaid. “Regulatory T cells and infection: a dangerous necessity”. In: *Nature Reviews Immunology* 7.11 (Nov. 2007). Number: 11 Publisher: Nature Publishing Group, pp. 875–888. ISSN: 1474-1741. DOI: [10.1038/nri2189](https://doi.org/10.1038/nri2189). URL: <https://www.nature.com/articles/nri2189> (visited on 01/18/2023).
- [65] Shaye Kivity et al. “Infections and autoimmunity – friends or foes?” In: *Trends in Immunology* 30.8 (Aug. 1, 2009), pp. 409–414. ISSN: 1471-4906. DOI: [10.1016/j.it.2009.05.005](https://doi.org/10.1016/j.it.2009.05.005). URL: <https://www.sciencedirect.com/science/article/pii/S1471490609001252> (visited on 01/18/2023).
- [66] Jan D. Lünemann et al. “EBNA1-specific T cells from patients with multiple sclerosis cross react with myelin antigens and co-produce IFN-gamma and IL-2”. In: *The Journal of Experimental Medicine* 205.8 (Aug. 4, 2008), pp. 1763–1773. ISSN: 1540-9538. DOI: [10.1084/jem.20072397](https://doi.org/10.1084/jem.20072397).
- [67] Xinbo Yang et al. “Autoimmunity-associated T cell receptors recognize HLA-B*27-bound peptides”. In: *Nature* 612.7941 (Dec. 2022). Number: 7941 Publisher: Nature Publishing Group, pp. 771–777. ISSN: 1476-4687. DOI: [10.1038/s41586-022-05501-7](https://doi.org/10.1038/s41586-022-05501-7). URL: <https://www.nature.com/articles/s41586-022-05501-7> (visited on 06/22/2023).
- [68] Andréa Toma et al. “Recognition of a subregion of human proinsulin by class I-restricted T cells in type 1 diabetic patients”. In: *Proceedings of the National Academy of Sciences* 102.30 (July 26, 2005). Publisher: Proceedings of the National Academy of Sciences, pp. 10581–10586. DOI: [10.1073/pnas.0504230102](https://doi.org/10.1073/pnas.0504230102). URL: <https://www.pnas.org/doi/full/10.1073/pnas.0504230102> (visited on 03/21/2023).
- [69] Ania Skowera et al. “CTLs are targeted to kill beta cells in patients with type 1 diabetes through recognition of a glucose-regulated preproinsulin epitope”. In: *The Journal of Clinical Investigation* 118.10 (Oct. 2008), pp. 3390–3402. ISSN: 0021-9738. DOI: [10.1172/JCI35449](https://doi.org/10.1172/JCI35449).
- [70] Gabrielle G. M. Pinkse et al. “Autoreactive CD8 T cells associated with cell destruction in type 1 diabetes”. In: *Proceedings of the National Academy of Sciences* 102.51 (Dec. 20, 2005). Publisher: Proceedings of the National Academy of Sciences, pp. 18425–18430. DOI: [10.1073/pnas.0508621102](https://doi.org/10.1073/pnas.0508621102). URL: <https://www.pnas.org/doi/full/10.1073/pnas.0508621102> (visited on 03/21/2023).
- [71] Deborah Kronenberg et al. “Circulating Preproinsulin Signal Peptide-Specific CD8 T Cells Restricted by the Susceptibility Molecule HLA-A24 Are Expanded at Onset of Type 1 Diabetes and Kill β -Cells”. In: *Diabetes* 61.7 (June 15, 2012), pp. 1752–1759. ISSN: 0012-1797. DOI: [10.2337/db11-1520](https://doi.org/10.2337/db11-1520). URL: <https://doi.org/10.2337/db11-1520> (visited on 03/21/2023).
- [72] Charles T. Spencer et al. “Viral infection causes a shift in the self peptide repertoire presented by human MHC class I molecules”. In: *Proteomics. Clinical applications* 9.11 (Dec. 2015), pp. 1035–1052. ISSN: 1862-8346. DOI: [10.1002/prca.201500106](https://doi.org/10.1002/prca.201500106). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4920078/> (visited on 04/10/2023).
- [73] Mollie M. Jurewicz and Lawrence J. Stern. “Class II MHC Antigen Processing in Immune Tolerance and Inflammation”. In: *Immunogenetics* 71.3 (Mar. 2019), pp. 171–187. ISSN: 0093-7711. DOI: [10.1007/s00251-018-1095-x](https://doi.org/10.1007/s00251-018-1095-x). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6377339/> (visited on 04/10/2023).
- [74] Brian D. Stadinski et al. “I-Ag7 56/57 polymorphisms regulate non-cognate negative selection to CD4⁺ T cell orchestrators of type 1 diabetes”. In: *Nature Immunology* 24.4 (Apr. 2023), pp. 652–663. ISSN: 1529-2916. DOI: [10.1038/s41590-023-01441-0](https://doi.org/10.1038/s41590-023-01441-0).

- [75] Hanrong Chen, Arup K. Chakraborty, and Mehran Kardar. “How nonuniform contact profiles of T cell receptors modulate thymic selection outcomes”. In: *Physical Review. E* 97.3 (Mar. 2018), p. 032413. ISSN: 2470-0053. DOI: [10.1103/PhysRevE.97.032413](https://doi.org/10.1103/PhysRevE.97.032413).
- [76] Sanzo Miyazawa and Robert L. Jernigan. “Residue – Residue Potentials with a Favorable Contact Pair Term and an Unfavorable High Packing Density Term, for Simulation and Threading”. In: *Journal of Molecular Biology* 256.3 (Mar. 1, 1996), pp. 623–644. ISSN: 0022-2836. DOI: [10.1006/jmbi.1996.0114](https://doi.org/10.1006/jmbi.1996.0114). URL: <http://www.sciencedirect.com/science/article/pii/S002228369690114X> (visited on 03/20/2020).
- [77] Quentin Marcou, Thierry Mora, and Aleksandra M. Walczak. “High-throughput immune repertoire analysis with IGoR”. In: *Nature Communications* 9.1 (Feb. 8, 2018). Number: 1 Publisher: Nature Publishing Group, p. 561. ISSN: 2041-1723. DOI: [10.1038/s41467-018-02832-w](https://doi.org/10.1038/s41467-018-02832-w). URL: <https://www.nature.com/articles/s41467-018-02832-w> (visited on 07/04/2023).
- [78] Anand Murugan et al. “Statistical inference of the generation probability of T-cell receptors from sequence repertoires”. In: *Proceedings of the National Academy of Sciences* 109.40 (Oct. 2, 2012). Publisher: Proceedings of the National Academy of Sciences, pp. 16161–16166. DOI: [10.1073/pnas.1212755109](https://doi.org/10.1073/pnas.1212755109). URL: <https://www.pnas.org/doi/10.1073/pnas.1212755109> (visited on 07/04/2023).
- [79] Birkir Reynisson et al. “NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data”. In: *Nucleic Acids Research* 48 (W1 July 2, 2020), W449–W454. ISSN: 0305-1048. DOI: [10.1093/nar/gkaa379](https://doi.org/10.1093/nar/gkaa379). URL: <https://doi.org/10.1093/nar/gkaa379> (visited on 11/29/2022).
- [80] *Homo sapiens (Human) | Taxonomy | UniProt*. URL: <https://www.uniprot.org/taxonomy/9606> (visited on 01/19/2023).
- [81] *UniProt: the universal protein knowledgebase in 2021 | Nucleic Acids Research | Oxford Academic*. URL: <https://academic.oup.com/nar/article/49/D1/D480/6006196> (visited on 01/19/2023).
- [82] M. Merckenschlager et al. “How many thymocytes audition for selection?” In: *The Journal of Experimental Medicine* 186.7 (Oct. 6, 1997), pp. 1149–1158. ISSN: 0022-1007. DOI: [10.1084/jem.186.7.1149](https://doi.org/10.1084/jem.186.7.1149).
- [83] Ali Akram and Robert D. Inman. “Immunodominance: A pivotal principle in host response to viral infections”. In: *Clinical Immunology* 143.2 (May 1, 2012), pp. 99–115. ISSN: 1521-6616. DOI: [10.1016/j.clim.2012.01.015](https://doi.org/10.1016/j.clim.2012.01.015). URL: <https://www.sciencedirect.com/science/article/pii/S152166161200037X> (visited on 04/24/2023).
- [84] *Listeria monocytogenes serovar 1/2a (strain ATCC BAA-679 / EGD-e) | Taxonomy | UniProt*. URL: <https://www.uniprot.org/taxonomy/169963> (visited on 01/19/2023).
- [85] Julie Zikherman and Byron Au-Yeung. “The role of T cell receptor signaling thresholds in guiding T cell fate decisions”. In: *Current Opinion in Immunology* 33 (Apr. 2015), pp. 43–48. ISSN: 1879-0372. DOI: [10.1016/j.coi.2015.01.012](https://doi.org/10.1016/j.coi.2015.01.012).
- [86] Jayajit Das et al. “Digital signaling and hysteresis characterize Ras activation in lymphoid cells”. In: *Cell* 136.2 (Jan. 23, 2009), pp. 337–351. ISSN: 0092-8674. DOI: [10.1016/j.cell.2008.11.051](https://doi.org/10.1016/j.cell.2008.11.051). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2662698/> (visited on 04/24/2023).

- [87] Morgan Huse et al. “Spatial and Temporal Dynamics of T Cell Receptor Signaling with a Photoactivatable Agonist”. In: *Immunity* 27.1 (July 27, 2007), pp. 76–88. ISSN: 1074-7613. DOI: [10.1016/j.immuni.2007.05.017](https://doi.org/10.1016/j.immuni.2007.05.017). URL: <https://www.sciencedirect.com/science/article/pii/S1074761307003299> (visited on 04/24/2023).
- [88] Katarina Tengvall et al. “Molecular mimicry between Anoctamin 2 and Epstein-Barr virus nuclear antigen 1 associates with multiple sclerosis risk”. In: *Proceedings of the National Academy of Sciences* 116.34 (Aug. 20, 2019). Publisher: Proceedings of the National Academy of Sciences, pp. 16955–16960. DOI: [10.1073/pnas.1902623116](https://doi.org/10.1073/pnas.1902623116). URL: <https://www.pnas.org/doi/full/10.1073/pnas.1902623116> (visited on 03/21/2023).
- [89] O. Kanagawa et al. “Autoreactivity of T cells from nonobese diabetic mice: an I-Ag7-dependent reaction”. In: *Proceedings of the National Academy of Sciences of the United States of America* 95.4 (Feb. 17, 1998), pp. 1721–1724. ISSN: 0027-8424. DOI: [10.1073/pnas.95.4.1721](https://doi.org/10.1073/pnas.95.4.1721).
- [90] Paul V. Lehmann et al. “Determinant spreading and the dynamics of the autoimmune T-cell repertoire”. In: *Immunology Today* 14.5 (May 1, 1993), pp. 203–208. ISSN: 0167-5699. DOI: [10.1016/0167-5699\(93\)90163-F](https://doi.org/10.1016/0167-5699(93)90163-F). URL: <https://www.sciencedirect.com/science/article/pii/016756999390163F> (visited on 07/05/2023).
- [91] Paul V. Lehmann et al. “Spreading of T-cell autoimmunity to cryptic determinants of an autoantigen”. In: *Nature* 358.6382 (July 1992). Number: 6382 Publisher: Nature Publishing Group, pp. 155–157. ISSN: 1476-4687. DOI: [10.1038/358155a0](https://doi.org/10.1038/358155a0). URL: <https://www.nature.com/articles/358155a0> (visited on 07/05/2023).
- [92] Sunil Kumar Saini et al. “SARS-CoV-2 genome-wide T cell epitope mapping reveals immunodominance and substantial CD8+ T cell activation in COVID-19 patients”. In: *Science Immunology* 6.58 (Apr. 14, 2021). Publisher: American Association for the Advancement of Science, eabf7550. DOI: [10.1126/sciimmunol.abf7550](https://doi.org/10.1126/sciimmunol.abf7550). URL: <https://www.science.org/doi/full/10.1126/sciimmunol.abf7550> (visited on 03/12/2023).
- [93] Shawn A. Mahmud et al. “Tumor necrosis factor receptor superfamily costimulation couples T cell receptor signal strength to thymic regulatory T cell differentiation”. In: *Nature immunology* 15.5 (May 2014), pp. 473–481. ISSN: 1529-2908. DOI: [10.1038/ni.2849](https://doi.org/10.1038/ni.2849). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4000541/> (visited on 07/17/2023).
- [94] Ryan W. Nelson et al. “T cell receptor cross-reactivity between similar foreign and self peptides influences naive cell population size and autoimmunity”. In: *Immunity* 42.1 (Jan. 20, 2015), pp. 95–107. ISSN: 1097-4180. DOI: [10.1016/j.immuni.2014.12.022](https://doi.org/10.1016/j.immuni.2014.12.022).
- [95] *Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) | Taxonomy | UniProt*. URL: <https://www.uniprot.org/taxonomy/83332> (visited on 01/19/2023).
- [96] *Mus musculus (Mouse) | Taxonomy | UniProt*. URL: <https://www.uniprot.org/taxonomy/10090> (visited on 01/19/2023).
- [97] J. Kyte and R. F. Doolittle. “A simple method for displaying the hydropathic character of a protein”. In: *Journal of Molecular Biology* 157.1 (May 5, 1982), pp. 105–132. ISSN: 0022-2836. DOI: [10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0).
- [98] Thomas P. Hopp and Kenneth R. Woods. “A computer program for predicting protein antigenic determinants”. In: *Molecular Immunology* 20.4 (Apr. 1, 1983), pp. 483–489. ISSN: 0161-5890. DOI: [10.1016/0161-5890\(83\)90029-9](https://doi.org/10.1016/0161-5890(83)90029-9). URL: <https://www.sciencedirect.com/science/article/pii/0161589083900299> (visited on 06/27/2023).

- [99] David Eisenberg et al. “Hydrophobic moments and protein structure”. In: *Faraday Symposia of the Chemical Society* 17.0 (Jan. 1, 1982). Publisher: The Royal Society of Chemistry, pp. 109–120. ISSN: 0301-5696. DOI: [10.1039/FS9821700109](https://doi.org/10.1039/FS9821700109). URL: <https://pubs.rsc.org/en/content/articlelanding/1982/fs/fs9821700109> (visited on 06/27/2023).
- [100] George D. Rose et al. “Hydrophobicity of Amino Acid Residues in Globular Proteins”. In: *Science* 229.4716 (Aug. 30, 1985). Publisher: American Association for the Advancement of Science, pp. 834–838. DOI: [10.1126/science.4023714](https://doi.org/10.1126/science.4023714). URL: <https://www.science.org/doi/10.1126/science.4023714> (visited on 06/27/2023).
- [101] D. M. Engelman, T. A. Steitz, and A. Goldman. “Identifying Nonpolar Transbilayer Helices in Amino Acid Sequences of Membrane Proteins”. In: *Annual Review of Biophysics and Biophysical Chemistry* 15.1 (1986). _eprint: <https://doi.org/10.1146/annurev.bb.15.060186.001541>, pp. 321–353. DOI: [10.1146/annurev.bb.15.060186.001541](https://doi.org/10.1146/annurev.bb.15.060186.001541). URL: <https://doi.org/10.1146/annurev.bb.15.060186.001541> (visited on 06/27/2023).
- [102] William C. Wimley and Stephen H. White. “Experimentally determined hydrophobicity scale for proteins at membrane interfaces”. In: *Nature Structural Biology* 3.10 (Oct. 1996). Number: 10 Publisher: Nature Publishing Group, pp. 842–848. ISSN: 1545-9985. DOI: [10.1038/nsb1096-842](https://doi.org/10.1038/nsb1096-842). URL: <https://www.nature.com/articles/nsb1096-842> (visited on 06/27/2023).
- [103] Kai W Wucherpfennig and Jack L Strominger. “Molecular mimicry in T cell-mediated autoimmunity: Viral peptides activate human T cell clones specific for myelin basic protein”. In: *Cell* 80.5 (Mar. 1995), pp. 695–705. ISSN: 00928674. DOI: [10.1016/0092-8674\(95\)90348-8](https://doi.org/10.1016/0092-8674(95)90348-8). URL: <https://linkinghub.elsevier.com/retrieve/pii/0092867495903488> (visited on 03/20/2023).
- [104] Gijsbert P van Nierop et al. “Intrathecal CD8 T-cells of multiple sclerosis patients recognize lytic Epstein-Barr virus proteins”. In: *Multiple Sclerosis Journal* 22.3 (Mar. 1, 2016). Publisher: SAGE Publications Ltd STM, pp. 279–291. ISSN: 1352-4585. DOI: [10.1177/1352458515588581](https://doi.org/10.1177/1352458515588581). URL: <https://doi.org/10.1177/1352458515588581> (visited on 03/21/2023).
- [105] Tobias V. Lanz et al. “Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GialCAM”. In: *Nature* 603.7900 (Mar. 2022). Number: 7900 Publisher: Nature Publishing Group, pp. 321–327. ISSN: 1476-4687. DOI: [10.1038/s41586-022-04432-7](https://doi.org/10.1038/s41586-022-04432-7). URL: <https://www.nature.com/articles/s41586-022-04432-7> (visited on 03/21/2023).
- [106] Arianne C. van Sechel et al. “EBV-Induced Expression and HLA-DR-Restricted Presentation by Human B Cells of B-Crystallin, a Candidate Autoantigen in Multiple Sclerosis1”. In: *The Journal of Immunology* 162.1 (Jan. 1, 1999), pp. 129–135. ISSN: 0022-1767. DOI: [10.4049/jimmunol.162.1.129](https://doi.org/10.4049/jimmunol.162.1.129). URL: <https://doi.org/10.4049/jimmunol.162.1.129> (visited on 03/21/2023).
- [107] Michael P. Crawford et al. “High prevalence of autoreactive, neuroantigen-specific CD8+ T cells in multiple sclerosis revealed by novel flow cytometric assay”. In: *Blood* 103.11 (June 1, 2004), pp. 4222–4231. ISSN: 0006-4971. DOI: [10.1182/blood-2003-11-4025](https://doi.org/10.1182/blood-2003-11-4025).
- [108] Ying C. Q. Zang et al. “Increased CD8+ Cytotoxic T Cell Responses to Myelin Basic Protein in Multiple Sclerosis1”. In: *The Journal of Immunology* 172.8 (Apr. 15, 2004), pp. 5120–5127. ISSN: 0022-1767. DOI: [10.4049/jimmunol.172.8.5120](https://doi.org/10.4049/jimmunol.172.8.5120). URL: <https://doi.org/10.4049/jimmunol.172.8.5120> (visited on 03/21/2023).

- [109] Carol L. Vanderlugt and Stephen D. Miller. “Epitope spreading in immune-mediated diseases: implications for immunotherapy”. In: *Nature Reviews Immunology* 2.2 (Feb. 2002). Number: 2 Publisher: Nature Publishing Group, pp. 85–95. ISSN: 1474-1741. DOI: [10.1038/nri724](https://doi.org/10.1038/nri724). URL: <https://www.nature.com/articles/nri724> (visited on 12/08/2022).
- [110] Francesco Dotta et al. “Coxsackie B4 virus infection of β cells and natural killer cell insulinitis in recent-onset type 1 diabetic patients”. In: *Proceedings of the National Academy of Sciences* 104.12 (Mar. 20, 2007). Publisher: Proceedings of the National Academy of Sciences, pp. 5115–5120. DOI: [10.1073/pnas.0700442104](https://doi.org/10.1073/pnas.0700442104). URL: <https://www.pnas.org/doi/full/10.1073/pnas.0700442104> (visited on 03/21/2023).
- [111] Laura Marroqui et al. “Interferon- γ mediates human beta cell HLA class I overexpression, endoplasmic reticulum stress and apoptosis, three hallmarks of early human type 1 diabetes”. In: *Diabetologia* 60.4 (Apr. 1, 2017), pp. 656–667. ISSN: 1432-0428. DOI: [10.1007/s00125-016-4201-3](https://doi.org/10.1007/s00125-016-4201-3). URL: <https://doi.org/10.1007/s00125-016-4201-3> (visited on 03/21/2023).
- [112] Richard E. Lloyd, Manasi Tamhankar, and Åke Lernmark. “Enteroviruses and Type 1 Diabetes: Multiple Mechanisms and Factors?” In: *Annual Review of Medicine* 73.1 (2022). eprint: <https://doi.org/10.1146/annurev-med-042320-015952>, pp. 483–499. DOI: [10.1146/annurev-med-042320-015952](https://doi.org/10.1146/annurev-med-042320-015952). URL: <https://doi.org/10.1146/annurev-med-042320-015952> (visited on 03/21/2023).
- [113] Barbara Serafini et al. “Dysregulated Epstein-Barr virus infection in the multiple sclerosis brain”. In: *Journal of Experimental Medicine* 204.12 (Nov. 5, 2007), pp. 2899–2912. ISSN: 0022-1007. DOI: [10.1084/jem.20071030](https://doi.org/10.1084/jem.20071030). URL: <https://doi.org/10.1084/jem.20071030> (visited on 03/21/2023).
- [114] Ivan Jelcic et al. “Memory B Cells Activate Brain-Homing, Autoreactive CD4⁺ T Cells in Multiple Sclerosis”. In: *Cell* 175.1 (Sept. 20, 2018). Publisher: Elsevier, 85–100.e23. ISSN: 0092-8674, 1097-4172. DOI: [10.1016/j.cell.2018.08.011](https://doi.org/10.1016/j.cell.2018.08.011). URL: [https://www.cell.com/cell/abstract/S0092-8674\(18\)31027-4](https://www.cell.com/cell/abstract/S0092-8674(18)31027-4) (visited on 03/21/2023).
- [115] Carolyn G. King et al. “T Cell Affinity Regulates Asymmetric Division, Effector Cell Differentiation, and Tissue Pathology”. In: *Immunity* 37.4 (Oct. 19, 2012), pp. 709–720. ISSN: 1074-7613. DOI: [10.1016/j.immuni.2012.06.021](https://doi.org/10.1016/j.immuni.2012.06.021). URL: <https://www.sciencedirect.com/science/article/pii/S1074761312004293> (visited on 08/10/2023).

Supplementary Materials

Supp. 1

Amino acid frequencies used to generate TCRs:

Amino Acid	Homo Sapiens
C	0.0225
M	0.0215
F	0.0359
I	0.0434
L	0.0985
V	0.0598
W	0.0123
Y	0.0263
A	0.0692
G	0.0658
T	0.0536
S	0.0836
N	0.0360
Q	0.0481
D	0.0476
E	0.0718
H	0.0261
R	0.0568
K	0.0576
P	0.0636

Supp. 2

E_n gap

Dependence of the E_n - E_p gap on E_n for the 4% probability of emerging from thymic selection:

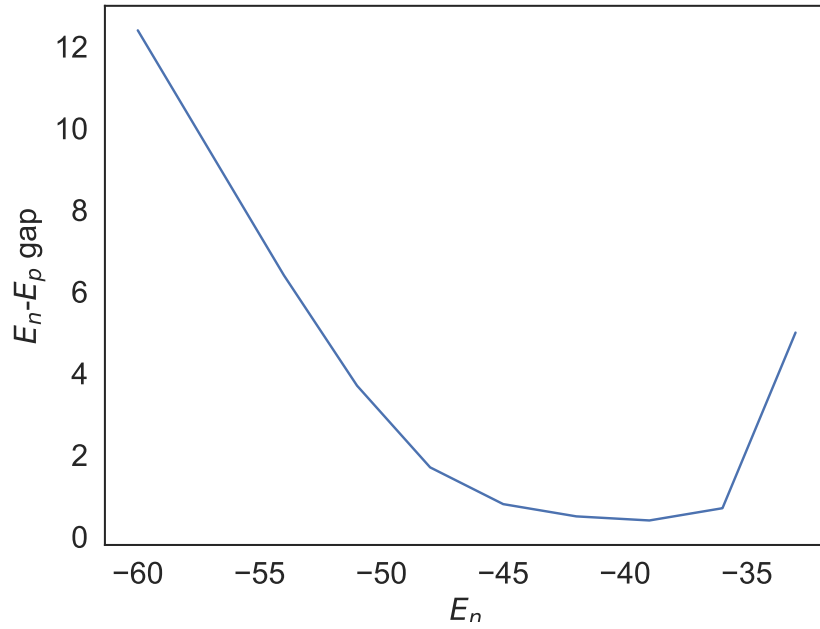
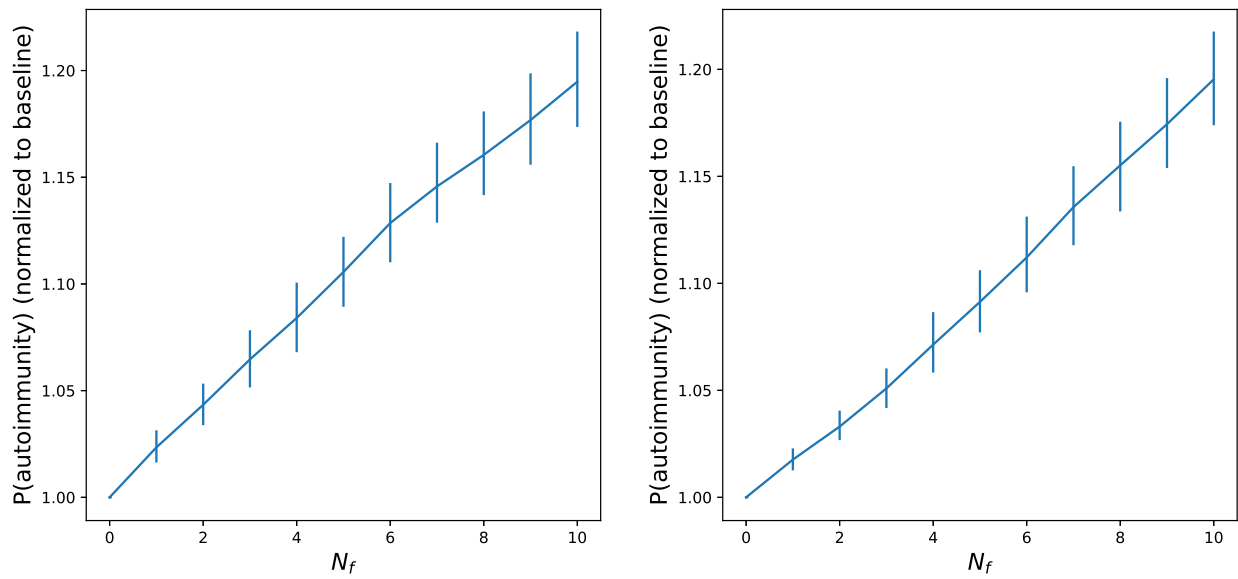


Figure 9: Graph showing the relationship between the value of E_n and the corresponding gap between the positive and negative selection thresholds such that there is 4% survival of immature thymocytes during selection.

E_n dependency



(a) Simulation with same parameters as Table 1 except with $E_n = -37k_B T$

(b) Simulation with same parameters as Table 1 except with $E_n = -39k_B T$

Figure 10: Results of simulations for two additional values of E_n , $-37k_B T$ and $-39k_B T$. The results are qualitatively the same as in Figure 6.

Supp. 3

Graph of activation by P5R peptide:

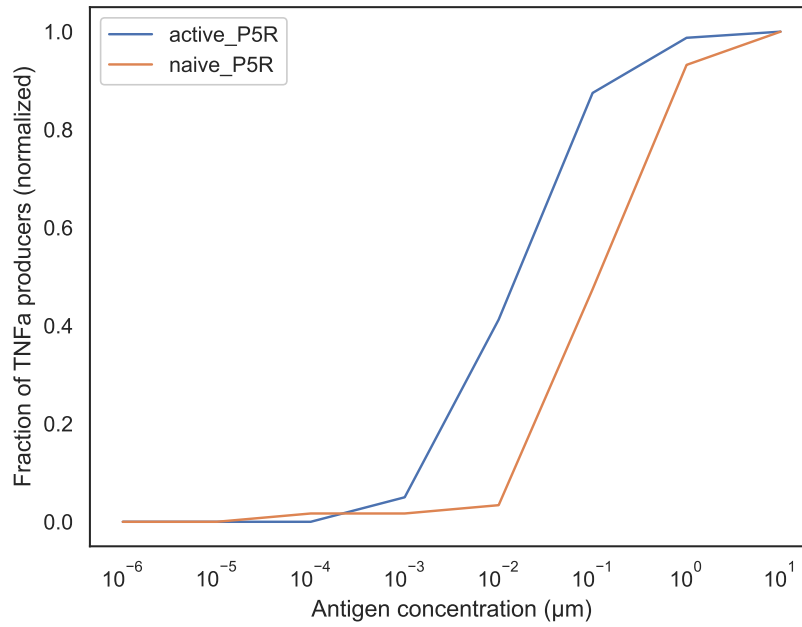


Figure 11: Graph of antigen concentration versus % of T cell population that are TNF α producers when exposed to another antigen, P5R. The blue and red curves are the responses of activated and naive T cells.

Supp. 4

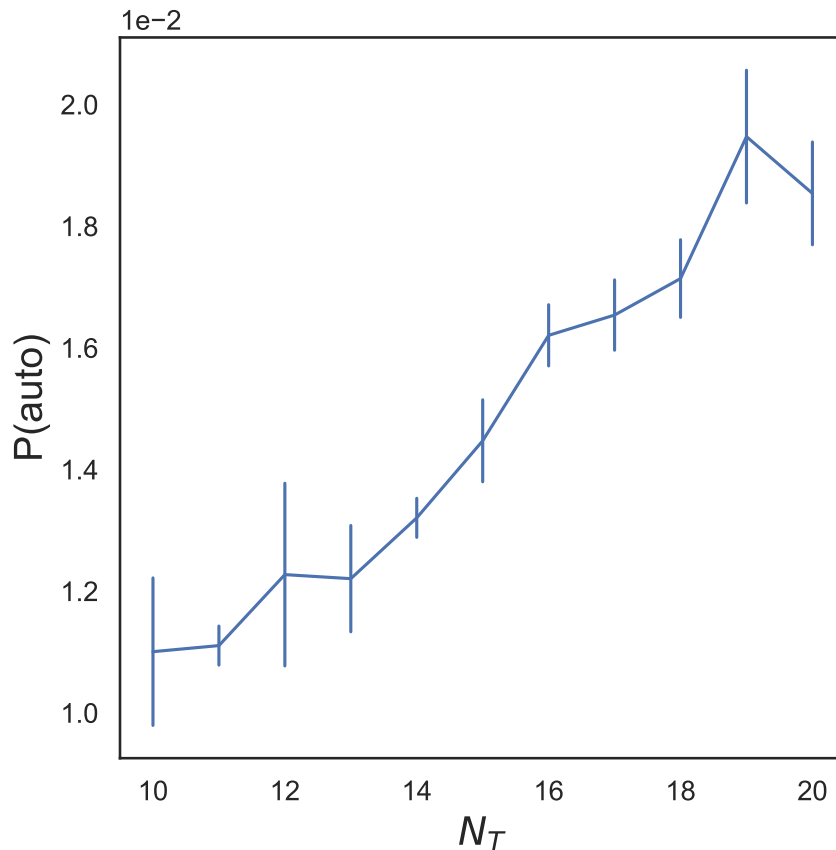


Figure 12: Graph of probability of autoimmunity for increasing N_T . Simulation parameters are the same as in Table 1, with $P(\text{autoimmunity})$ drawn from $N_f = 10$. Error bars were calculated based on the number of simulations noted in Figure 4.

Supp. 5

Hydrophobicity and Rare Peptides

Let each thymocyte i be characterized by a reactivity parameter, $\alpha_i \sim \mathcal{N}(0, 1)$, and each peptide j be identified by a reactivity parameter β_j drawn from the same distribution. This reactivity parameter may be thought of as a proxy for hydrophobicity, such that larger values indicate that the thymocyte or peptide is more likely to bind strongly to its counterpart. We can then write the free energy of an interaction between thymocyte i and peptide j as

$$E_{ij} = -(\alpha_i + \beta_j). \quad (6)$$

Using this framework, we can simulate selection by generating 1,000,000 naive thymocytes, each associated with a reactivity α_i , and 10,000 self-peptides associated with a reactivity β_j . Each thymocyte i encounters a random set, S_j , of M self-peptides ($M = 8,000$) during selection, and the

mature repertoire consists of thymocytes i where

$$E_n < \min_{j \in S_j} E_{ij} < E_p .$$

We can then calculate the number of T cells that bind to a sample of 10,000 self-peptides, and pathogenic peptides, and observe a tail analogous to the one seen in the string model [Fig. 8]. In addition, we can calculate the number of T cells that are activated by pathogenic and self-peptides in the same fashion as the string model, which produces a very similar graph [Fig. 13] as seen by the string model [Fig. 6].

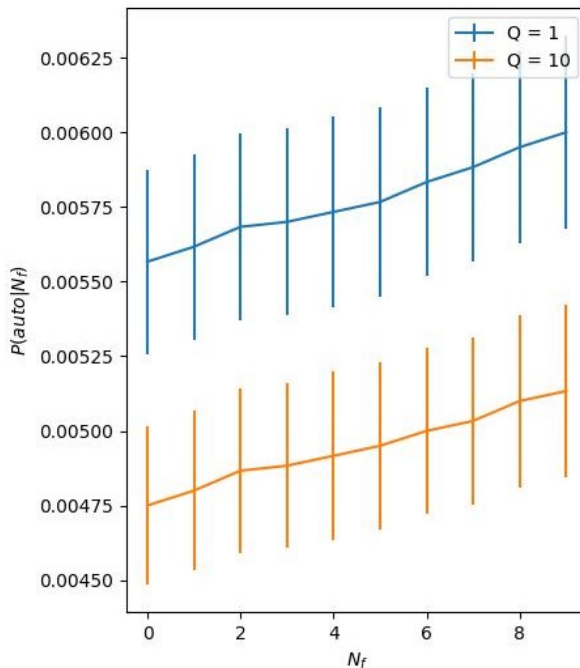


Figure 13: Probability of autoimmunity due to cross-reactivity of pathogen-derived peptides as function of the number of pathogen-derived peptides, simulated based on the coarse grained model described in the text and above. The results are similar to those in Figure 6 in the main text. Parameters used: $\frac{M}{N} = 0.8$, and $N_T = 10$, number of naive T cells $T = 10^6$, $E_n = 2.90$, $E_p = 2.59$, $E_{\text{weak}} = 2.25$.

Supp. 6

Distribution of number of mouse homologs for each tuberculosis peptide:

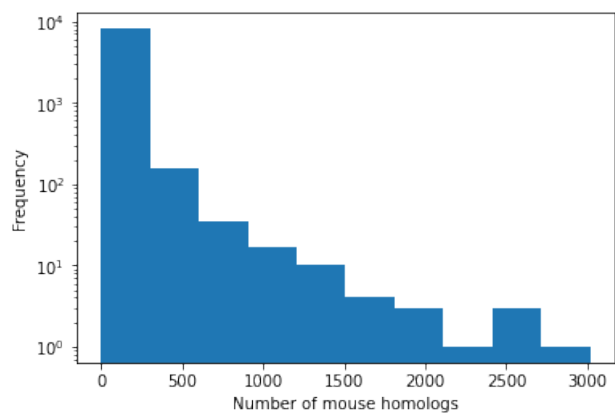


Figure 14: Histogram of mouse homologs per tuberculosis peptide.

Table S1: List of tuberculosis epitopes. The first 31 peptides above the line are the most hydrophobic peptides in this list.

Epitope	9-mer Core		
ARLVAPLAVAAAVDD	LVAPLAVAA	RVAVLFAGPGMNLA	FAGPGMNLA
RRIVAPTUVVIHGRAD	IVAPTUVVIH	DKPFEATVVSVDPSV	FEATVVSVD
AAITAPVAIANFTGS	ITAPVAIAN	KEGFEATVVTDGPAA	FEATVVTDG
VETFTAIALAVGAI	FTAIALAV	DDFGAAEAANMAKYA	FGAAEAANM
ETQMTAAIPLATREV	MTAAIPLAT	VRFHQAVTAAGVDHN	FHQAVTAAG
PGFIGAPLALALGQ	IGAPLALAA	REFLAANAASIPTKS	FLAANAASI
TPEWAAALSGLAAGD	WAAALSGLA	KDKFLAATSSDTVRS	FLAATSSDT
IVRWATPVTAFAQRTA	WATPVTAFAQ	STAFLATTVLESQRH	FLATTVLES
EQTVVSSAATIAAPG	VVSSAATIA	DESFLTGESLPVDKQ	FLTGESLPV
PELDQAPVAAIEEHL	DQAPVAAIE	DQFFNAPANGPAGLF	FNAPANGPA
EPGGYTVAVGASAVA	YTVAVGASA	PFNGATPANSPYPTH	FNGATPANS
SGGFLSGAAGVSGAD	FLSGAAGVS	RKFSANEANAEEQMQ	FSANEANAA
ASYISGAVIPVDGGM	YISGAVIPV	VEFSGPEAVGDYLG	FSGPEAVGD
EWSTPAALVAKIKEI	WSTPAALVA	VPFSYPASLPTPFGP	FSYPASLPT
DPVYAAAFNTAADDA	YAAAFNTAA	ARDFVAGVAQSCEEL	FVAGVAQSC
RIFAGPLAGALKMPY	FAGPLAGAL	AVHRAPKPAQGLTYD	HRAPKPAQG
KADLAGPLSGAGPIV	LAGPLSGAG	ATMIATARALASRAE	IATARALAS
HVTAAPNVLAASAGE	TAAPNVLAA	ADKIRGGKVAAAGAI	IRGGKVAAA
PPVAAPNVLAGVLEH	VAAPNVLAG	EAIRGPLAEALSGIE	IRGPLAEAL
RALRMSTPLALTEEE	MSTPLALTE	PQSITGATATATHGT	ITGATATAT
AYHLFGPAPATDNDL	LFGPAPATD	AGIITGPTVNEYAVY	ITGPTVNEY
VLPRPMFGAAAATAA	MFGAAAATA	RGDGTITGSAINESA	ITGSAINES
VRIAGPVVTGVALAA	IAGPVVTGV	AGGIVPALVNYAPTQ	IVPALVNYA
RPYIGVSANIVAAAQ	YIGVSANIV	AAGLDAPSVAGHERR	LDAPSVAGH
QSGNAGPAVNVLAGA	NAGPAVNVL	PHDLGAAEPAFAPGP	LGAAEPAFA
DRGIQAPAALGLICP	IQAPAALGL	TVDLGSPAAQAIWVV	LGSPAAQAI
AESLLGPGTQLADLK	LLGPGTQLA	RRLHQPVIAAVNGPA	LHQPVIAAV
HYLGPGAQLDGPEIM	YLGPGAQLD	EPLNPPAAAEMAARF	LNPPAAAEM
GLVTGPLTNLALALR	VTGPLTNLA	AAQLREAAAQAQDTQ	LREAAAQAQ
NRWSYPSALATDHHH	WSYPSALAT	AALRSARPNSSPGAA	LRSARPNSS
FNLLGPLTNPARPRA	LLGPLTNPA	WGPLSAPKLEKAIDS	LSAPKLEKA
AAAPAAAPHSTPGGVH	AAAPHSTPG	TELVAPHTVNTMPEK	LVAPHTVNT
AGLAASPSANVLEE	AASPSANVL	MPGMHGTVAAVAALQ	MHGTVAAVA
AGAATPAAGAAPSAG	AATPAAGAA	TLMISPGAARVINPD	MISPGAARV
AGPAGGPAPVESTDN	AGGPAPVES	LPRMSAAIAEPTSHD	MSAAIAEPT
RDDAKGPSASVLSLK	AKGPSASVL	IPAMTAAQAVASRVG	MTAAQAVAS
GVAALGPDAAQVIYS	ALGPDAAQV	DNPASPLTNRAISET	PASPLTNRA
VKANAPTAASDEPNQ	ANAPTAASD	RKRPGSPAAQASKRA	PGSPAAQAS
DWRANGPIASTGLAR	ANGPIASTG	SGAQLAPGTEPSPAP	QLAPGTEPS
APAQAPKVAVAVLVE	AQAPKVAVA	AIDSAAPPAELSMGD	SAAPPAELS
SLNCAQGPTANPCGV	AQGPTANPC	TNRSATPVASEGPAV	SATPVASEG
FHDAAPKVADAVGGN	DAAPKVADA	PGLDSLPPAASEAEA	SLPPAASEA
DCDHGPVAVAGRFB	DHGPVAVAV	IRSTAAAYPNPALRS	TAAAYPNPA
DADDRGPASVSGGA	DRGPASVGS	GGTTAAPGINGSRAR	TAAPGINGS
TEIEGTPVTAEVFDE	EGTPVTAEV	QSTRSPQTVSAVVND	TRSPQTVSA
		PANTTSPTAGQSLDC	TTSPTAGQS

RIVAAALPNSSSSQP	VAAALPNSS
VGQVAASHPEVTFVG	VAASHPEVT
ILVEGPGALVTADTH	VEGPGALVT
GTVQGPQAALVSI GL	VQGPQAALV
EPAVRAALAE LTRAA	VRAALAE L T
MPVRAPAAVRGAGLI	VRAPAAVRG
PARVSGPDPVPGCLS	VSGPDPVPG
PTVTSPQVAVNDIGS	VTSPQVAVN
GNGVVAAALAAKTAVP	VVAALAAKT
TAVVAPAADPVSLQS	VVAPAADPV
VSVVYSSVAENAFVR	VYSSVAENA
VAGWAPAVVSGAAIA	WAPAVVSGA
SSQWISSEAA RLDLH	WISSEAA R L
QFGWQGASAMALNAK	WQGASAMAL
HAWRGSKSAAVKGIM	WRGSKSAAV
TVSWRPSAVAHTKAS	WRPSAVAHT
TGEWSAAVSRLAAAE	WSAAVSRLA
ADSWVEAAPAQGWPA	WVEAAPAQG
NLYAAAESHTVEQMV	YAAAESHTV
EVPYAEIVASVASVS	YAEIVASVA
IPYAGPPIADLFQPQ	YAGPPIADL
EDYEGAEISDGAGEL	YEGAEISDG
IRYEGPKGGPGMREM	YEGPKGGPG
SLYEGTTPVEVNRTP	YEGTTPVEV
AQYIAPEQALGHDAS	YIAPEQALG
APRYQAAAPTGP IEE	YQAAAPTGP
QIYQATAANS DLKRL	YQATAANS D
GTLYQQVSAEAAAVY	YQQVSAEAA
RYRAPAGVRS PAAVP	YRAPAGVRS
QAYRGPLGQRLVVRG	YRGPLGQRL
GDYTGIEPEK TTPAH	YTGIEPEKT
