

# Quorum sensing allows T cells to discriminate between self and nonself

Thomas Charles Butler<sup>a,b</sup>, Mehran Kardar<sup>a,1</sup>, and Arup K. Chakraborty<sup>c,d,1</sup>

Departments of <sup>a</sup>Physics and <sup>b</sup>Chemical Engineering, Massachusetts Institute of Technology, Cambridge MA 02139; <sup>c</sup>Departments of Chemical Engineering, Physics, Chemistry, and Biological Engineering, Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA 02139; and <sup>d</sup>Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology, and Harvard University, Cambridge, MA

Edited\* by David R. Nelson, Harvard University, Cambridge, MA, and approved June 4, 2013 (received for review December 26, 2012)

**T cells orchestrate pathogen-specific adaptive immune responses by identifying peptides derived from pathogenic proteins that are displayed on the surface of infected cells. Host cells also display peptide fragments from the host's own proteins. Incorrectly identifying peptides derived from the body's own proteome as pathogenic can result in autoimmune disease. To minimize autoreactivity, immature T cells that respond to self-peptides are deleted in the thymus by a process called negative selection. However, negative selection is imperfect, and autoreactive T cells exist in healthy individuals. To understand how autoimmunity is yet avoided, without loss of responsiveness to pathogens, we have developed a model of T-cell training and response. Our model shows that T cells reliably respond to infection and avoid autoimmunity because collective decisions made by the T-cell population, rather than the responses of individual T cells, determine biological outcomes. The theory is qualitatively consistent with experimental data and yields a criterion for thymic selection to be adequate for suppressing autoimmunity.**

statistical mechanics | T-cell-mediated autoimmunity | self tolerance

**T**-cell receptors (TCRs) expressed on the surface of T cells bind (and recognize) peptides displayed on cells. The diversity of the TCR repertoire allows T cells to recognize peptides from a vast range of pathogens (1–4). However, because TCR diversity is generated by stochastic gene rearrangement, newly created TCRs also have the potential to respond to peptides originating from endogenous proteins, possibly leading to autoimmune responses. Immature T cells (thymocytes) are screened for autoreactivity in the thymus before maturation. Within the thymus, thymocytes are exposed to large numbers of ectopically expressed proteins from throughout the body, including peptides derived from all organs. If any of the peptides that a thymocyte encounters binds strongly to its TCR, the thymocyte is deleted, greatly reducing the likelihood that a mature T cell will be activated by self-peptides (5–7).

Encounters between thymocytes and self-peptides during thymic selection are affected by stochastic effects, including how the thymocyte explores the thymus, stochastic ectopic gene expression, and other factors (8–12). These factors imply that thymic selection does not expose thymocytes to every possible self-peptide a mature T cell may encounter in the periphery. This suggests that autoreactive T cells may enter circulation as part of normal T-cell development. In fact, experimental studies of human T-cell repertoire samples confirm that healthy humans contain T cells specific for self-peptides (13, 14).

We present a theory that describes how, despite each individual T cell's potential for autoimmunity, the immune system can avoid T-cell-mediated autoimmunity without the risk of immunodeficiency. Whereas each individual T cell has been only negatively selected against a limited cross-section of self-peptides, other T cells have survived negative selection against different cross-sections of the self repertoire. The key insight is that, if a mechanism exists for activated T cells to check whether neighboring T cells have also been activated before proliferating and

mounting an immune response, then even a repertoire containing many autoreactive individual T cells can avoid autoimmunity. Evidence for such collective decision making exists.

Both experiments and theory indicate that, at least in some circumstances, the decision of T cells to proliferate and mount a response relies on a threshold number of T cells becoming activated by antigen (peptide) stimulation (Fig. 1) (15–17). This is analogous to quorum sensing in bacterial populations (18). One experiment varied the concentration of antigen-stimulated T cells in vitro and measured the resulting levels of pSTAT5, a survival/proliferation marker. Levels of pSTAT5 sufficient to indicate proliferation occurred only when the concentration of antigen-stimulated T cells was sufficiently high (15). The T-cell interactions required for this phenomenon are believed to be mediated by the cytokine IL-2, which is produced by activated T cells (15, 16). Activated T cells can detect IL-2 being secreted by other activated T cells and begin to proliferate and respond only when a threshold number are activated. Experimental evidence suggests that the threshold is controlled by regulatory T cells, which compete with activated T cells for IL-2 (13, 15). Further experiments are needed to identify the conditions under which T cells are required to form a quorum to respond and to determine the precise mechanisms. However, we emphasize that the results in this paper do not depend on a particular mechanism for T-cell communication.

That many T cells together may make a better decision about whether to respond is reminiscent of a classical result from social decision theory, the Condorcet Jury theorem (19). This theorem states that if there are  $N$  voters choosing between two alternatives, and the probability that a voter makes the correct choice is greater than one-half, then for large  $N$ , the probability that the vote returns the correct choice approaches 1. This is true even if individual voters are barely better than coin flips in discriminating between the correct and incorrect choices.

Condorcet's theorem suggests that T cells, through their requirement to form quorums, could amplify their ability to make correct decisions about whether or not to respond. However, T cells respond only to a tiny fraction ( $10^{-4}$ – $10^{-6}$ ) of the peptides they are exposed to (20, 21), so the direct application of Condorcet's theorem with T-cell activation interpreted as a "yes" vote would imply that a T-cell response to infection would essentially never occur, so an alternative procedure is required. Here we develop a theory of T-cell decision making that explicitly accounts for the process of thymic selection and the existence of a quorum rule that requires a sufficient number of T cells to be locally activated for response. Our results highlight the role of intercellular communication between T cells in balancing

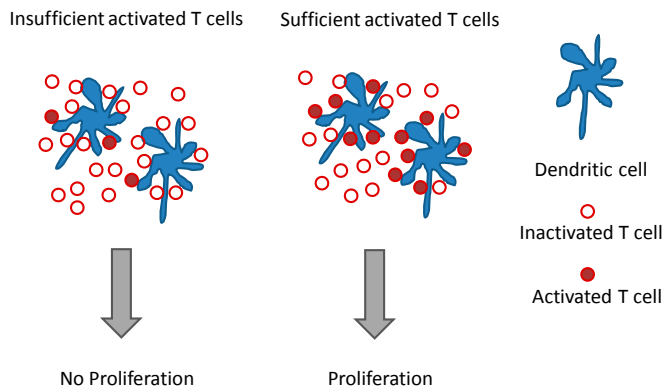
Author contributions: T.C.B., M.K., and A.K.C. designed research, performed research, and wrote the paper.

The authors declare no conflict of interest.

\*This Direct Submission article had a prearranged editor.

<sup>1</sup>To whom correspondence may be addressed. E-mail: arupc@mit.edu or kardar@mit.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1222467110/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1222467110/-DCSupplemental).



**Fig. 1.** Schematic representation of the role of a quorum-sensing-like mechanism for T-cell proliferation. Peptides bound to major histocompatibility complex (MHC) molecules on the surface of dendritic cells in the lymph nodes are presented to T cells. If an insufficient number of T cells are activated, no response is generated. If a sufficient number of T cells are activated, proliferation and response result.

the conflicting demands of self-tolerance and responsiveness to diverse pathogens.

### T-Cell Activation Model

Each T cell usually expresses a single T-cell receptor clonotype. The receptor encounters and potentially binds to large numbers of peptides, both self- and pathogen-derived, mounted on major histocompatibility complex (MHC) molecules on the surface of antigen-presenting cells (APCs). TCR:peptide MHC (TCR:pMHC) complexes bind with energies taken from a TCR-dependent probability distribution  $q_i(E)$ , which is not known. Related models, as in our previous work, assume that the distribution of binding energies is Gaussian (22–24). This is justified by the Central Limit Theorem and the fact that the binding energy of a TCR:pMHC complex is the result of many factors (22–25), including the sequence of the TCR and the sequence of the peptide. However, the results presented here do not require any specific functional form for the distribution of binding energies to be specified.

If the binding energy required to activate a T cell is indicated by  $E_a$ , then the probability that the T-cell  $i$  is activated by a random peptide is given by

$$P_i(\text{Activation}) = \int_{-\infty}^{E_a} q_i(z) dz, \quad [1]$$

taking the sign convention that more negative binding energies are stronger.

Recent studies have shown that the cross-reactivity [represented by the magnitude of  $P(\text{Activation})$  in Eq. 1] of naive T cells is decreased by thymic selection (22, 26–28). Following ref. 22 by using extreme value theory to model thymic selection, we can explicitly model the impact of thymic selection on the properties of the T-cell repertoire. In the process of negative selection, if the magnitude of the binding energy of a thymocyte with a self-peptide in the thymus exceeds the magnitude of the negative selection energy  $E_n$ , then the thymocyte does not enter the repertoire. On the other hand, if a thymocyte does not bind with any self-peptide at least as strongly as  $E_p$  (positive selection), it is also deleted.  $E_p$  is known to correspond to slightly weaker TCR:pMHC interactions compared with  $E_n$  (29). Thus, as  $E_p$  approaches  $E_n$ , the strongest interaction,  $E^*$ , that a thymocyte can have with a set of  $M$  self-peptides in the thymus and still survive, must be approximately equal to  $E_n$  (22). Combined

with the fact that the activation threshold of T cells in the periphery is approximately equal to the negative selection threshold ( $E_a = E_n$ ) (30), the strongest interaction,  $E_i^*$ , in the thymus must be nearly equal to the activation threshold,  $E_a$ , so that the activation probability for T cells that survive thymic selection is given by

$$P_i(\text{Activation}) = \int_{-\infty}^{E_i^*} q_i(z) dz. \quad [2]$$

Using the theory of extreme values, the integral in Eq. 2 can be carried out (*SI Text I*), even though the underlying distribution  $q_i(E)$  in Eq. 1 is unknown and potentially varies between T cells. Carrying out the integral, we have the very simple result of

$$P(\text{Activation}) \equiv p = \frac{1}{M}. \quad [3]$$

The decreasing value of  $p$  as a function of increasing exposure to self-peptides in the thymus is consistent with experiments and models showing that T cells gain specificity from thymic selection (26, 28). However, if  $E_p$  is not asymptotically close to  $E_n$ , the T cells that enter the periphery may be more difficult to activate and Eq. 3 represents an upper bound of activation probability. Rather than deal with the resulting distribution of activation probabilities, we rely on experiments to estimate the correction to Eq. 3 for a typical T cell. For fixed  $E_n$ , increasing the difference between  $E_n$  and  $E_p$  increases the likelihood that less reactive T cells that easily survive negative selection will be positively selected and enter the periphery. If the probability of negative selection is  $1/K$ , the result is (*SI Text I*)

$$p = \frac{1}{KM}. \quad [4]$$

The probability of negative selection for thymocytes is estimated experimentally to be of order  $10^{-1}$  (31, 32), leading to the estimate  $p \approx 1/10M$ . We also investigated the dependence of activation probability on thymic selection, using simulations. To carry out the simulations, we adopted the string model of TCR:pMHC binding and thymic selection described in ref. 26. In this model, strings represent amino acid sequences for the complementarity-determining region 3 (CDR3) loop of the TCR, as well as the peptide antigen. The binding energy of the TCR to the peptide is then computed using an empirical potential (33). To simulate thymic selection, we generated random TCR sequences and used sequences from the mouse proteome from UNIPROT (34) to generate random ensembles of self-peptides that the TCRs were negatively selected against. Testing binding probabilities for the resulting TCRs yielded results consistent with the estimate in Eq. 4 (*SI Text II*).

When a T cell encounters a self-peptide, the probability that the self-peptide is one of those that the T cell encountered in the thymus is  $M/N$ , where  $N$  is the total number of self-peptides. Thus, the probability that a T cell is activated by a random self-peptide is given by

$$P(\text{Activation}|\text{Self}) = p \left(1 - \frac{M}{N}\right) \equiv px \quad [5]$$

with  $p$  as defined in Eq. 4.

Eqs. 4 and 5 assume that the binding energy distribution of peptides to a given T cell does not depend on whether the peptide originates from self or pathogen. Simulations of thymic selection similar to those described in refs. 26 and 35 were used to create a set of T-cell receptors *in silico*. These TCRs were then used to obtain simulated binding energy distributions from human

peptides that were not present in the thymus and *Salmonella typhi* peptides. The results confirm that the binding energy distributions for human and pathogenic peptides binding to TCRs are approximately equal (SI Text IV).

Let us assume that the T-cell repertoire in the periphery is composed of  $T$  distinct cells. Experiments indicate that  $T = 5 \times 10^6$  is a conservative lower bound on T-cell diversity (4), although larger numbers are likely. Indeed, given the small activation probability of an individual T cell in Eq. 4, a large  $T$  is essential to avoid immunodeficiency; i.e.,  $T > 1/p$  for mean activation number  $Tp > 1$ . However, a large  $T$  also increases the chance of activation by self-peptides. The probability that no member of the repertoire is activated by the entirety of  $N$  self-peptides is

$$P(\text{Autoimmunity fully avoided}) = \left[ (1 - px)^T \right]^N \simeq \exp(-xNTp). \quad [6]$$

Even in a hypothetical case where thymic selection is nearly perfect, with each thymocyte exposed to all but one self-peptide ( $M = N - 1$  and  $x = 1/N$ ), according to Eq. 6 the probability that every T cell is not activated by self-peptides  $e^{-Tp}$  is still quite small (because  $Tp > 1$ ), equal to  $10^{-4}$  for  $M = 9,999$ ,  $N = 10,000$ , and  $T = 5 \times 10^6$ . This argument shows that a T-cell repertoire diverse enough to avoid immunodeficiency (requiring  $Tp > 1$ ) cannot be fully self-tolerant unless thymic selection is perfect, which is inconsistent with experiments showing that healthy humans contain autoreactive T cells (13, 14).

However, as conjectured above, satisfactory determination of self and nonself relies on a collective mechanism of T-cell decision making through the quorum rule. This requires accounting for not only the diversity,  $T$ , of the T-cell population, but also interactions between T cells. Incorporating the latter can be accomplished by introducing an additional parameter,  $t$ , denoting the threshold number of T cells that must be activated before a response is initiated. The resulting theory described below shows how the challenge of tolerating a diverse proteome ( $N$ ) while maintaining responsiveness to infection requires a sufficient amount of thymic selection ( $M$ ) and T-cell diversity ( $T$ ) to make effective decisions.

### Results

To analyze the effects of the quorum rule and thymic selection on the ability to discriminate between self and nonself, we must compute the probability that the T-cell population makes a mistake. We explicitly include the assumption that a quorum of activated T cells of threshold size is required for a response. The theory of T-cell activation probability and thymic selection developed above provides the basis for the calculation.

Consider the probabilities of mutually exclusive mistakes of T-cell responses under different assumptions about the threshold number of activated T cells ( $t$ ) required to initiate a response. The first mistake ( $M1$ ) occurs when there is no T-cell response even though there is infection, and pathogenic peptides are presented by APCs in the lymph nodes. The lower the threshold number of T cells required to be activated before a response is initiated, the smaller the probability of  $M1$ . This is because the requirement for proliferation in response to any stimuli is less stringent, so response to pathogen is more likely. In the opposite limit, where infinitely many activated T cells are required for a response,  $M1$  occurs upon every exposure to pathogen because no response can occur.

The second type of mistake ( $M2$ ) occurs when inflammation or injury leads to tissue-specific self-peptides being presented more abundantly in the lymph nodes, and activated T cells proliferate, respond, and attack the body's own tissues. In this case, the larger the number of T cells required to be activated before a response occurs (the larger  $t$  is), the lower the probability of

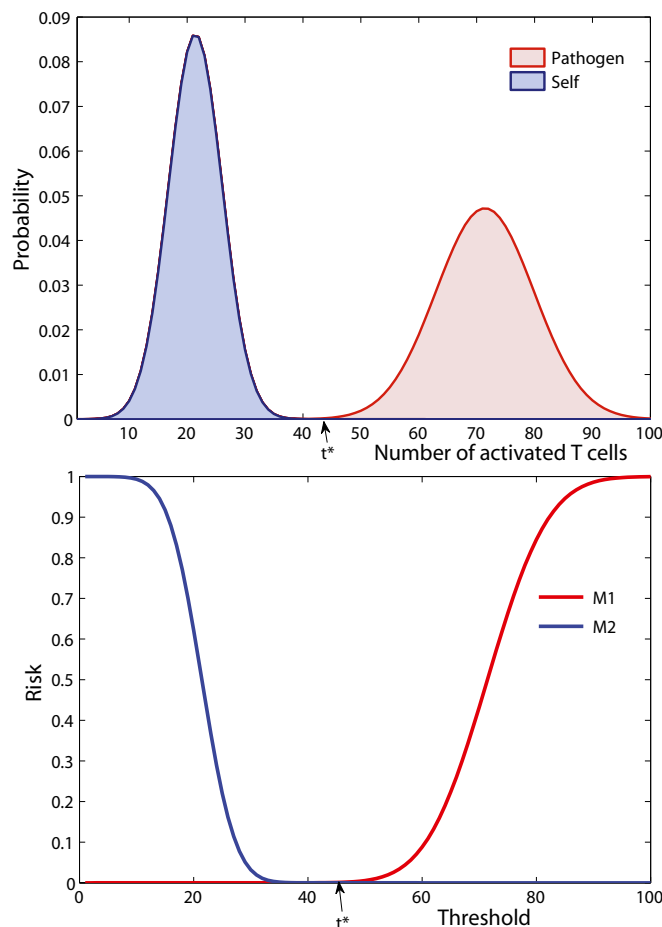
$M2$ . In the opposite limit, where only one T cell is required for a response,  $M2$  is very likely to occur because if any of the T cells in the body recognize a self-peptide as presented in the lymph node, then response will occur (Eq. 6).

Without thymic selection, no possible balance of  $M1$  and  $M2$  exists. For any threshold, one or the other mistake will happen with high probability. However, thymic selection ensures that, on average, peptides derived from self have fewer precursor T cells than those derived from pathogens as implied by Eq. 5. Thus, the mean number of T cells activated by a display of self will be lower than the mean number activated by a pathogen. If the quorum threshold is chosen to lie between the two means, both classes of mistakes can be suppressed. This is illustrated in Fig. 2, *Upper*.

The probabilities of each of these mistakes can be computed explicitly:

$$P_t(M1|S) = P(\sim R|\sim S) = \sum_{k=0}^{t-1} \binom{T}{k} p^k (1-p)^{T-k}, \quad [7]$$

$$P_t(M2|S) = P(R|S) = \sum_{k=t}^T \binom{T}{k} (px)^k (1-px)^{T-k}.$$



**Fig. 2.** (*Upper*) Probability distributions for the number of T cells activated by a collection of self-peptides vs. a collection of pathogenic peptides. (*Lower*) Risk functions for all three kinds of error computed as conditional risks in Eq. 7. An optimal threshold requires more T cells to be activated than are likely to be activated by self but a sufficiently low number that at least the threshold number will essentially always be activated by pathogen. The parameters are  $T = 5 \times 10^6$ ,  $M = 7 \times 10^3$ , and  $N = 10^4$ .

The subscript  $t$  on the probabilities indicates the quorum threshold, the state  $R$  indicates a response,  $S$  represents self-peptide display,  $T$  represents the number of unique T-cell-antigen interactions,  $p$  and  $x$  are as defined above in Eq. 5, and  $\sim$  indicates negation.

Assuming that the cost to the organism of any mistake is equal, and without loss of generality setting that cost to be 1, each of the conditional mistake probabilities is equal to the risk of loss associated with the corresponding mistake. Combining the risks, there exists an optimal threshold  $t^*$  that minimizes the risks. To solve for  $t^*$ , a decision rule must be chosen. The simplest decision rule minimizes the expected loss [Bayes expected loss principle (36)]. However, computing the expected loss correctly requires specifying the probabilities  $P(S)$  and  $P(\sim S)$ . In the absence of sufficient information to assign the probabilities  $P(S)$  and  $P(\sim S)$ , we chose to adopt the minimax decision principle introduced originally by John Von Neumann (36, 37). The robustness of the results with respect to the decision principle and costs is discussed below and in *SI Text III*.

The minimax principle, as applied to this problem, corresponds to a threshold chosen to minimize the largest risk at the value of the threshold. Mathematically, the minimax decision rule gives the optimum threshold

$$t^* = \operatorname{argmin}_t \sup_{\theta} \sum_{m \in M} P_t(m|\theta), \quad [8]$$

with  $M = \{M1, M2\}$  and  $\theta \in \Theta$ ,  $\Theta = \{S, \sim S\}$ . To evaluate Eq. 8, we consider the Gaussian approximation of the probabilities (or risks) in Eq. 7,

$$P_t(M1|\sim S) = \sum_{k=0}^{t-1} \binom{T}{k} p^k (1-p)^{T-k} \approx \frac{1}{2} \operatorname{erfc}(-z_1), \quad [9]$$

$$P_t(M2|S) = \sum_{k=t}^T \binom{T}{k} px^k (1-px)^{T-k} \approx \frac{1}{2} \operatorname{erfc}(z_2), \quad [10]$$

where  $z_1 = (t - Tp)/(Tp(1-p))$  and  $z_2 = (t - Tpx)/(Tpx(1-px))$ . To satisfy the minimax principle, we seek the point  $-z_1 = z_2$ . This gives the solution

$$t^* = Tp - Tp(1-x) \frac{1}{\sqrt{\frac{x(1-px)}{1-p} + 1}} \quad [11]$$

(Fig. 2). To gain insight into how the threshold depends on the original parameters, we note that  $p$  is small and take  $x$  close to 1 (corresponding to sparse coverage in thymic selection). Expanding Eq. 11, we obtain

$$t^* \approx Tp - \frac{Tp(1-x)}{2} = \frac{T}{KM} - \frac{T}{2KN}. \quad [12]$$

Eq. 12 shows that a greater degree of thymic selection (larger  $M$ ) allows a lower threshold to be chosen, whereas a greater degree of diversity in the proteome (larger  $N$ ) requires a larger threshold for self-tolerance.

Even with an optimal choice of threshold, it may be that costs are not significantly reduced. For the T-cell quorum rule, the decision rule performs well if the risks of each of the mistakes are

both substantially depressed at the threshold  $t^*$ . This can be formalized by requiring that  $|z_i(t^*)| \gg 1$  for both  $M1$  and  $M2$ , resulting in the criterion

$$\frac{TM}{K} \gg 4N^2. \quad [13]$$

Eq. 13 defines the conditions required for the T-cell population to be both self-tolerant and immune-competent even with imperfect thymic selection. This result has the surprising implication that T-cell diversity (proportional to  $T$ ) plays a key role in self-tolerance by allowing a larger threshold (Eq. 12). It further shows the interplay between the degree of thymic selection (parameterized by  $M$  as defined above), the diversity of self-peptides ( $N$ ), and the diversity of T cells. For a highly diverse proteome, the amount of T-cell diversity multiplied by the amount of self seen in the thymus must scale as approximately the square of the proteome diversity.

As noted above, it is not clear that the minimax rule is the best choice for the problem. However, if the criterion in Eq. 13 is satisfied, then the results are qualitatively robust to variations in both the decision rule and losses assigned to each mistake (*SI Text III*).

Accurate estimates of the independent parameters of this model ( $T, M, N$ ) for human and mouse are difficult to determine. For  $T$ , as a first approximation, we use the estimate  $T = 5 \times 10^6$  discussed above. We use the total number of distinct T-cell receptors because circulation of T cells through the body is surprisingly efficient, allowing a large fraction of T cells to participate in the decision, even if the infection is local (38). For the number of self-epitopes that can be displayed,  $N$ , the proteome of a human or mouse is of order  $10^4 - 10^5$  distinct peptides. Each protein can be broken up into of order 100 distinct epitopes. However, only a small fraction, roughly 1% (39), will be displayed, suggesting the estimate in the range  $N = 10^4 - 10^5$ . For numerical study, we adopt  $N = 10^4$ . Estimating  $M$  is more difficult. Thymocytes scan around 500 distinct APCs during thymic selection (10, 40). However, it is very difficult to assess how many unique proteins they are exposed to in each contact. To avoid this difficulty, we estimate  $M$  indirectly using Eq. 3, which relates  $M$  to the more experimentally accessible quantity  $p$ , the probability that a random peptide activates a given T cell. Experimental measurements of  $p$  are typically in the range  $10^{-4} - 10^{-6}$  (20, 21). Seeking values of  $M$  consistent with these data, using Eq. 3, suggests a rough estimate for  $M$  centered near  $M = 0.7 \times 10^4$ . These parameter estimates satisfy the criterion of Eq. 13 and underscore the importance of high T-cell diversity in maintaining self-tolerance (because  $M < N$ ). With these parameters, the quorum threshold is estimated to be  $t^* \approx 46$ . The risks as a function of threshold and the calculated optimal threshold using the estimated parameters are shown in Fig. 2. Further experiments are needed to better estimate the values of the parameters. However, we emphasize that our results depend only upon the experimentally justified assumption of a quorum rule  $t > 1$  and the condition in Eq. 13 being satisfied.

## Discussion

T-cell activation functions at the single-cell level and plays an essential role in self-nonself discrimination. Negative selection in the thymus ensures that the chance that a T cell is activated by a self-peptide is less than the chance that it is activated by a pathogenic peptide. We have proposed that this difference is amplified by T cells making collective decisions according to a quorum rule to realize robust self-tolerance in the face of autoreactive T cells. If such collective decision-making mechanisms functioning at the many-cell level play this critical role in self-tolerance, how can this be affirmed in vivo? Consider an idealized experiment. Genetically modified mice with altered

T-cell signaling machinery such that no quorum is required for response should develop autoimmune disease spontaneously if collective decision making is critical for self-tolerance. However, it is difficult to see how to carry out such an experiment.

In the absence of such an idealized experiment, it is possible that IL-2 dosage experiments similar to those reported in ref. 41 on nonobese diabetic (NOD) mice (42) may provide some insight into how an experimental test might be developed. These experiments showed that low dosages of IL-2 supported the regulatory T-cell population and delayed the onset of diabetes compared with that in untreated NOD mice. In contrast, high dosages of IL-2 promoted the proliferation of effector T cells and led to the rapid onset of diabetes concomitant with an excess of effector cells in the pancreas. If, as suggested in ref. 15 and discussed in the Introduction, IL-2 mediates the communication between T cells that underlies the quorum rule, the results of these experiments can be interpreted in the light of collective T-cell decision making. Regulatory T cells are known to be more sensitive to IL-2 dosage than other CD4 and CD8 T cells (41). In the low-dosage experiments, the regulatory T cells become more active whereas the effector cells are unaffected, thus enhancing the quorum threshold. In the high-dosage experiments, there was sufficient IL-2 to allow the effector T cells to proliferate on their own without need for IL-2 secretion by other cells, leading directly to diabetes by circumventing the quorum rule. Cytokines might dynamically and locally regulate the number of T cells required to form a quorum. However, in the absence of more direct evidence of a quorum threshold *in vivo*, the primary value of these experiments is in suggesting a possible experimental path.

An experimental test of the quorum threshold rule suggested by the NOD IL-2 dosage experiments might inject normal mice with self- and pathogenic peptides for which there are known amounts of T-cell precursors (20, 21). Along with the peptides, the mice would be injected with varying doses of IL-2. The resulting T-cell response could then be measured as a function of the number of precursors and the dosage of IL-2. Varying the dosage of IL-2 will, albeit in the nonmonotonic way described in the NOD experiments described above, adjust the quorum threshold. Varying the peptide will vary the number of activated T cells. We anticipate that such an experiment would lead to the discovery of all or nothing immune responses as a function of both IL-2 dosage and peptide identity in the event that the quorum rule functions *in vivo*.

By injecting a mouse with a heterogeneous mixture of self-peptides and varying dosages of IL-2, autoimmunity could be induced above a certain IL-2 dosage. In principle, the curve corresponding to *M2* in Fig. 2 could be completely reconstructed. Injecting a mouse with a heterogeneous mix of pathogenic peptides and different IL-2 dosages should result in an immune response above an IL-2 dosage that is offset from that required for autoimmunity due to self-peptide injection, allowing in principle the reproduction of the curve corresponding to *M1* in Fig. 2.

Experiments closely related to the ones we are proposing have been performed with a fixed amount of adjuvant coinjected with peptides in place of a varying amount of IL-2. These studies examined the proliferation of T cells in response to injection of antigen for which it is known that there are few precursors (20, 21). Although there are few precursors, there is still proliferation in response to the peptides. The number of T-cell precursors in the above studies is a minimum of about 15 cells total, and perhaps all of them participate in the decision, with a quorum threshold of less than 15. This would require that the 15 cells end up in the same lymph node and that they are able to signal each other. Some evidence that this is possible comes from experiments that give mice a highly localized infection and track the resulting T-cell response. Within 4 d nearly all of the T-cell precursors specific to the pathogen are recruited from circulation. However, in this experiment, there were much larger numbers of precursor cells (38). A more likely possibility is that the quorum rule was bypassed because of high dosages of adjuvant that were used in the experiments (20, 21). Under milder inflammatory conditions during natural infection, immune responses can occur, even if precursor frequencies are low, because many pathogenic peptides are presented by APCs. Thus, the quorum rule can be satisfied by T cells specific for different pathogenic peptides in the same lymph node.

We anticipate that the results we have reported will lead to further experimental work on how inter-T-cell communication and quorum sensing are connected with self-tolerance, leading to a deeper understanding of how the immune system is able to be both responsive to pathogen and tolerant of self.

**ACKNOWLEDGMENTS.** This work was partially supported by a National Institutes of Health Pioneer Award (to A.K.C.) and National Science Foundation Grant 12-06323 (to M.K.).

- Arstila TP, et al. (1999) A direct estimate of the human  $\alpha\beta$  T cell receptor diversity. *Science* 286(5441):958–961.
- Nikolich-Zugich J, Slifka MK, Messaoudi I (2004) The many important facets of T-cell repertoire diversity. *Nat Rev Immunol* 4(2):123–132.
- Robins HS, et al. (2009) Comprehensive assessment of T-cell receptor  $\beta$ -chain diversity in alpha-beta T cells. *Blood* 114(19):4099–4107.
- Warren RL, et al. (2011) Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome Res* 21(5):790–797.
- Hogquist KA, Baldwin TA, Jameson SC (2005) Central tolerance: Learning self-control in the thymus. *Nat Rev Immunol* 5(10):772–782.
- Klein L, Hinterberger M, Wirnsberger G, Kyewski B (2009) Antigen presentation in the thymus for positive selection and central tolerance induction. *Nat Rev Immunol* 9(12):833–844.
- Kyewski B, Derbinski J (2004) Self-representation in the thymus: An extended view. *Nat Rev Immunol* 4(9):688–698.
- Derbinski J, Pinto S, Rösch S, Hexel K, Kyewski B (2008) Promiscuous gene expression patterns in single medullary thymic epithelial cells argue for a stochastic mechanism. *Proc Natl Acad Sci USA* 105(2):657–662.
- Venanzi ES, Melamed R, Mathis D, Benoist C (2008) The variable immunological self: Genetic variation and nongenetic noise in Aire-regulated transcription. *Proc Natl Acad Sci USA* 105(41):15860–15865.
- Le Borgne M, et al. (2009) The impact of negative selection on thymocyte migration in the medulla. *Nat Immunol* 10(8):823–830.
- Li J, Park J, Foss D, Goldschneider I (2009) Thymus-homing peripheral dendritic cells constitute two of the three major subsets of dendritic cells in the steady-state thymus. *J Exp Med* 206(3):607–622.
- Proietto AI, et al. (2008) Dendritic cells in the thymus contribute to T-regulatory cell induction. *Proc Natl Acad Sci USA* 105(50):19869–19874.
- Danke NA, Koelle DM, Yee C, Beheray S, Kwok WW (2004) Autoreactive T cells in healthy individuals. *J Immunol* 172(10):5967–5972.
- Su LF, Kidd BA, Han A, Kotzin JJ, Davis MM (2013) Virus-specific CD4 memory-phenotype T cells are abundant in unexposed adults. *Immunity* 38(2):373–383.
- Feinerman O, et al. (2010) Single-cell quantification of IL-2 response by effector and regulatory T cells reveals critical plasticity in immune response. *Mol Syst Biol* 6(1):437.
- Busse D, et al. (2010) Competing feedback loops shape IL-2 signaling between helper and regulatory T lymphocytes in cellular microenvironments. *Proc Natl Acad Sci USA* 107(7):3058–3063.
- Burroughs NJ, Miguel Paz Mendes de Oliveira B, Adrego Pinto A (2006) Regulatory T cell adjustment of quorum growth thresholds and the control of local immune responses. *J Theor Biol* 241(1):134–141.
- Miller MB, Bassler BL (2001) Quorum sensing in bacteria. *Annu Rev Microbiol* 55(1):165–199.
- de Condorcet M (1785) *Essay on the Application of Analysis to the Probability of Majority Decisions* (De L'imprimerie Royale, France).
- Moon JJ, et al. (2007) Naive CD4(+) T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. *Immunity* 27(2):203–213.
- Jenkins MK, Moon JJ (2012) The role of naive T cell precursor frequency and recruitment in dictating immune response magnitude. *J Immunol* 188(9):4135–4140.
- Košmrlj A, Chakraborty AK, Kardar M, Shakhnovich EI (2009) Thymic selection of T-cell receptors as an extreme value problem. *Phys Rev Lett* 103(6):068103.
- Derrida B (1980) Random-energy model: Limit of a family of disordered models. *Phys Rev Lett* 45(2):79–82.
- Derrida B (1981) Random-energy model: An exactly solvable model of disordered systems. *Phys Rev B* 24(5):2613.
- Wolfson MY, Nam K, Chakraborty AK (2011) The effect of mutations on the all-or-none T cell receptor/peptide-MHC interface structure: A molecular dynamics study. *J Phys Chem B* 115(25):8317–8327.

26. Košmrlj A, Jha AK, Huseby ES, Kardar M, Chakraborty AK (2008) How the thymus designs antigen-specific and self-tolerant T cell receptor sequences. *Proc Natl Acad Sci USA* 105(43):16671–16676.
27. Košmrlj A, et al. (2010) Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. *Nature* 465(7296):350–354.
28. Huseby ES, et al. (2005) How the T cell repertoire becomes peptide and MHC specific. *Cell* 122(2):247–260.
29. Daniels MA, et al. (2006) Thymic selection threshold defined by compartmentalization of Ras/MAPK signalling. *Nature* 444(7120):724–729.
30. Naeher D, et al. (2007) A constant affinity threshold for T cell tolerance. *J Exp Med* 204(11):2553–2559.
31. van Meerwijk JPM, et al. (1997) Quantitative impact of thymic clonal deletion on the T cell repertoire. *J Exp Med* 185(3):377–383.
32. Palmer E (2003) Negative selection—clearing out the bad apples from the T-cell repertoire. *Nat Rev Immunol* 3(5):383–391.
33. Miyazawa S, Jernigan R (1985) Estimation of effective interresidue contact energies from protein crystal structures: Quasi-chemical approximation. *Macromolecules* 18:534–552.
34. UniProt Consortium (2012) Reorganizing the protein space at the universal protein resource (uniprot). *Nucleic Acids Res* 40(Database issue):D71–D75.
35. Detours V, Mehr R, Perelson AS (1999) A quantitative theory of affinity-driven T cell repertoire selection. *J Theor Biol* 200(4):389–403.
36. Berger J (1985) *Statistical Decision Theory and Bayesian Analysis* (Springer, New York).
37. Von Neumann J (1928) Zur theorie der gesellschaftspiele [The theory of games]. *Math Ann* 100(1):295–320.
38. Stock AT, Jones CM, Heath WR, Carbone FR (2011) Rapid recruitment and activation of CD8+ T cells after herpes simplex virus type 1 skin infection. *Immunol Cell Biol* 89(1):143–148.
39. Rao X, Costa AI, van Baarle D, Keşmir C (2009) A comparative study of HLA binding affinity and ligand diversity: Implications for generating immunodominant CD8+ T cell responses. *J Immunol* 182(3):1526–1532.
40. Klein L (2009) Dead man walking: How thymocytes scan the medulla. *Nat Immunol* 10(8):809–811.
41. Tang Q, et al. (2008) Central role of defective interleukin-2 production in the triggering of islet autoimmune destruction. *Immunity* 28(5):687–697.
42. Makino S, et al. (1980) Breeding of a non-obese, diabetic strain of mice. *Jikken Dobutsu* 29(1):1–13.

# Supporting Information

Butler et al. 10.1073/pnas.1222467110

## SI Text I. Probability of Response to a Random Peptide

In this section we derive the probability that a T-cell receptor (TCR) will be activated by a random peptide. The result is derived by integrating the probability density  $q(E)$  for the binding energy of a random peptide with a TCR given that the TCR survived thymic selection. The distribution,  $q(E)$ , does not need to be specified. If  $M$  random variables  $X_i \sim q(E)$  are generated, the minimum of this set,  $x^*$ , is the negative extreme value. As argued in the main text, the set of random variables can be taken to be the interaction energies of a thymocyte with self-peptides in the thymus. For T cells that survive selection in the limit that  $E_p \rightarrow E_n$  and large  $M$ , the activation energy,  $E_a$ , is the negative extreme value  $x^*$ . Under these assumptions, the probability that a T cell will be activated by a random peptide is given by

$$p_a(x^*) = \int_{-\infty}^{x^*} q(z) dz. \quad [\text{S1}]$$

To evaluate this integral, we note that in the main text it is shown that as a consequence of thymic selection, the activation energy is approximately equal to the extreme value for  $M$  random interactions, with  $M$  the number of peptides the T cell was exposed to in the thymus. The probability distribution for the extreme value itself is given by

$$p_M(x) = Mp(x)(1 - p_a(x))^{M-1}. \quad [\text{S2}]$$

An implicit equation for the most probable value for the extreme value ( $x^*$ ) is obtained by differentiation,

$$\frac{dp_M(x)}{dx} = Mq'(x)(1 - p_a(x))^{M-1} - M(M-1)q(x)^2(1 - p_a(x))^{M-2}. \quad [\text{S3}]$$

Setting the derivative equal to zero, and simplifying, results in

$$q'(x^*)(1 - p_a(x^*)) - (M-1)q(x^*)^2 = 0. \quad [\text{S4}]$$

We now evaluate the integral in Eq. S1:

$$\begin{aligned} p_a(x^*) &= \int_{-\infty}^{x^*} \exp(\ln q(z)) dz \\ &\simeq \int_{-\infty}^{x^*} \exp[\ln q(x^*) + (z - x^*)q'(x^*)/q(x^*)] dz \\ &= \frac{q(x^*)^2}{q'(x^*)}. \end{aligned} \quad [\text{S5}]$$

Substituting into Eq. S4, we obtain the result

$$p_a(E_a) = \frac{1}{M} \quad [\text{S6}]$$

for large  $M$ .

The calculation above assumes that  $E_a = E_n = E_p$ . In reality,  $E_p > E_n$ . Unless, in contradiction to experimental evidence,

negative selection is the only source of selection, then the extreme value experienced by T cells during thymic selection will typically be less than the activation energy. To estimate the true activation probability, we adopt another approach to estimating the activation probability.

Consider an individual thymocyte that has a probability  $p_{ns}$  of being negatively selected during thymic selection with  $M$  self-peptides. Negative selection occurs when the thymocyte encounters a self-peptide and binds with a binding energy exceeding the activation energy. Assuming, consistent with experiment, that the probability of activation ( $p_a$ ) is small,  $p_{ns}$  and  $p_a$  are related by

$$1 - p_{ns} = (1 - p_a)^M \simeq \exp(-Mp_a), \quad [\text{S7}]$$

or, assuming  $p_{ns}$  is small,

$$\begin{aligned} p_a &\simeq \frac{\log(1 - p_{ns})}{M} \\ &\simeq \frac{p_{ns}}{M}. \end{aligned} \quad [\text{S8}]$$

With the assignment  $p_{ns} = 1/K$ , the result of the main text is obtained. The relationship between the result in Eq. S8 and that in Eq. S6 is that as  $E_p \rightarrow E_n$ ,  $p_{ns} \rightarrow 1$ .

## SI Text II. Simulation Details

Following previous work (1–3), we modeled the TCR:peptide major histocompatibility complex (TCR:pMHC) interaction by representing the complementarity-determining region 3 (CDR3) region of the TCR as a string of amino acids represented by the vector  $\mathbf{t} = \{t_1, t_2, \dots, t_L\}$ . The peptide complexed with the MHC is represented by the vector  $\mathbf{s} = \{s_1, s_2, \dots, s_L\}$ . The interaction energy is then given by

$$E_{\text{int}}(\mathbf{s}, \mathbf{t}) = \sum_{i=1}^L J(s_i, t_i) + \frac{1}{2} \sum_{|i-j|=1} J(s_i, t_j) + E_c. \quad [\text{S9}]$$

Here  $E_c$  represents the interaction between the T cell and the MHC independent of the variable region and the peptide. The interaction potential between amino acid residues  $J(\cdot, \cdot)$  is given by the Miyazawa–Jernigan statistical potential (4). Interactions are between corresponding positions on the TCR and the peptide as well as nearest-neighbor positions. The sum is over amino acids on the CDR3 loop of the TCR and the peptide mounted to the MHC. For both CD8 and CD4 T cells, this number is roughly eight, although some sites have more contact than others (5). If the interaction energy exceeds an activation threshold  $E_a$ , the peptide  $\mathbf{s}$  activates the T cell with TCR  $\mathbf{t}$ .

We tested the analytical results in Eqs. S6 and S8 above by carrying out simulations with  $(E_p - E_n)/E_n = 0.008$  as well as with  $(E_p - E_n)/E_n = 0.125$ . The simulations confirm the  $1/M$  scaling for both cases and that increasing  $(E_p - E_n)/E_n$  decreases the activation probability. See Fig. S1.

## SI Text III. Robustness of Results to Changes in Decision Rule

In the main text, it was assumed that the costs of the errors were equal and that using a minimax decision rule was valid. It is possible that the costs are not equal and/or a Bayes expected risk

principle is preferable to the minimax principle. Here we show that if the error condition from the main text

$$TM \gg 4KN^2 \quad [\text{S10}]$$

is satisfied, then the results are robust to the changes. To account for the probabilities of each of the mistakes and different costs, we write the Bayesian expected loss (6) with activation at threshold  $t$  for each state

$$C_1 = K_1 P(\sim S) P_t(M1 | \sim S), \quad [\text{S11}]$$

$$C_2 = K_2 P(S) P_t(M2 | S). \quad [\text{S12}]$$

The costs and probabilities can be merged into a single constant for each mistake ( $K_{1,2} = K'_{1,2} P$ )

$$C_1 = K_1 P_t(M1 | \sim S), \quad [\text{S13}]$$

$$C_2 = K_2 P_t(M2 | S). \quad [\text{S14}]$$

The threshold  $t^*$  adequately suppresses error if both  $C_i$  are small, even with the modifications. The error condition Eq. S10 implies that the complementary error functions in the costs (derived in the main text)

$$P_t(M1 | \sim S) \simeq \frac{1}{2} \operatorname{erfc}(-z_1), \quad [\text{S15}]$$

$$P_t(M2 | S) \simeq \frac{1}{2} \operatorname{erfc}(z_2), \quad [\text{S16}]$$

have arguments larger than 1. Explicitly, we have

$$|z_i| \simeq \frac{1}{2N} \sqrt{TM} \gg 1. \quad [\text{S17}]$$

Proceeding on one case (generalizing is trivial), we have

$$C_2 = K_2 P(\sim S) P_t(M2 | S) \simeq \frac{K_2}{2} \operatorname{erfc}(z_2) \quad [\text{S18}]$$

$$\simeq \frac{K_2}{2z_2} \exp[-z_2^2] = \frac{1}{2z_2} \exp[-z_2^2 + \log K_2].$$

This implies a new error suppression criterion

$$\frac{1}{2N} \sqrt{\frac{TM}{K}} \gg 1 + \log K_{\max}, \quad [\text{S19}]$$

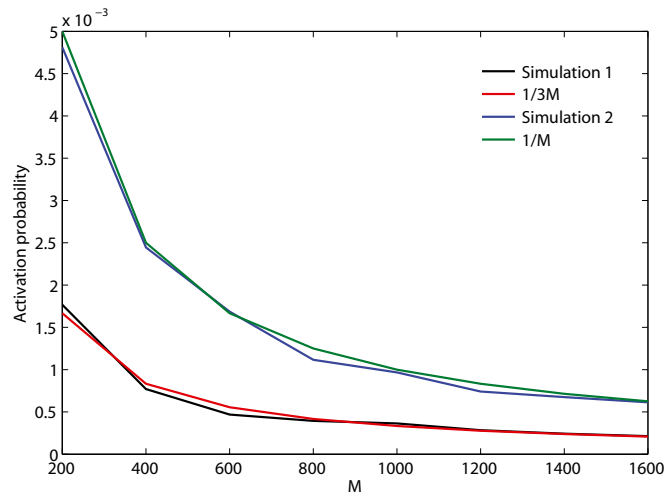
where  $K_{\max}$  is the maximum of the two  $K_i$ . Due to the slow growth of the logarithm, unless  $K_{\max}$  is extremely large, error is adequately suppressed by the value of  $t^*$  solved for in the main text, even if the decision rule or costs are changed a great deal, as shown in the criterion Eq. S19. Thus, the results of the main text are robust under substantial changes of decision criteria.

#### SI Text IV. Binding Energy Distributions for Self-Peptides and Typhoid Peptides

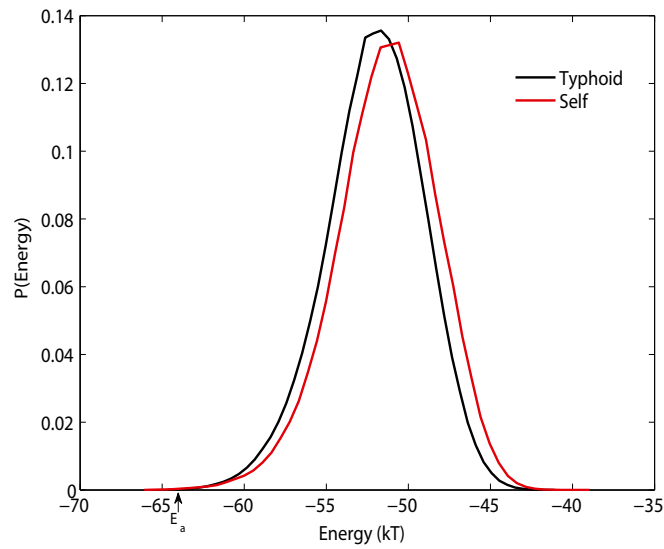
An assumption of the calculations in the main text is that the binding energy distributions of random self- and nonself-peptides to a given TCR are similar. In this section, we examine the validity of this assumption *in silico*. Fig. S2 shows simulated binding energy distributions. The distributions are measured using simulations for 5,555 peptides of length 8 from *Salmonella typhi* (Typhoid) and for 5,555 peptides of length 8 from the human proteome, each tested against the same set of 1,000 TCRs, which were selected against 500 self-peptides distinct from those used to measure the binding energy distribution. Near the activation energy ( $E_a = -64$  kT) the binding energy distributions are essentially identical. Given that T cells are sensitive only to energies near the activation or negative selection energies, the assumption that the distributions are approximately the same is consistent with the results of our simulations.

1. Košmrlj A, Jha AK, Huseby ES, Kardar M, Chakraborty AK (2008) How the thymus designs antigen-specific and self-tolerant T cell receptor sequences. *Proc Natl Acad Sci USA* 105(43):16671–16676.
2. Košmrlj A, Chakraborty AK, Kardar M, Shakhnovich EI (2009) Thymic selection of T-cell receptors as an extreme value problem. *Phys Rev Lett* 103(6):068103.
3. Košmrlj A, et al. (2010) Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. *Nature* 465(7296):350–354.
4. Miyazawa S, Jernigan RL (1996) Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *J Mol Biol* 256(3):623–644.
5. Huseby ES, et al. (2005) How the T cell repertoire becomes peptide and MHC specific. *Cell* 122(2):247–260.
6. Berger J (1985) *Statistical Decision Theory and Bayesian Analysis* (Springer, New York).





**Fig. S1.** Average probability of activation by a random peptide as a function of the number ( $M$ ) of self-peptides T cells are exposed to in the thymus. For simulation 1,  $L=8$ ,  $E_n = -64$ ,  $E_p = -59$ ,  $E_c = -32$  in units of  $kT$ , so there is significant separation between the negative and positive selection energies. For simulation 2, energies for  $L=8$ ,  $E_n = -64$ ,  $E_p = -63.6$ ,  $E_c = -32$  in units of  $kT$ , so that the positive and negative selection thresholds are very close to each other. As claimed, the gap between the positive and negative selection energies reduces the activation probability. All data points are averages over 500 T cells trained on uncorrelated sets of self-peptides.



**Fig. S2.** Binding energy histogram for self-peptides and typhoid peptides. Each data point is the average of 1,000 T cells, with  $L=8$ ,  $E_n = -64$ ,  $E_p = -59$ ,  $E_c = -24$ .