How a well-adapted immune system is organized

Andreas Mayer[a], Vijay Balasubramanian[b,c], Thierry Mora[d,1], and Aleksandra M. Walczak[a]

[a]Laboratoire de Physique Théorique and [d]Laboratoire de Physique Statistique, CNRS, Université Pierre et Marie Curie and École Normale Supérieure, 75005 Paris, France; [b]Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104; and [c]Initiative for the Theoretical Sciences, The Graduate Center, The City University of New York, New York, NY 10016

Edited by José N. Onuchic, Rice University, Houston, TX, and approved March 3, 2015 (received for review November 14, 2014)

The repertoire of lymphocyte receptors in the adaptive immune system protects organisms from diverse pathogens. A well-adapted repertoire should be tuned to the pathogenic environment to reduce the cost of infections. We develop a general framework for predicting the optimal repertoire that minimizes the cost of infections contracted from a given distribution of pathogens. The theory predicts that the immune system will have more receptors for rare antigens than expected from the frequency of encounters; individuals exposed to the same infections will have sparse repertoires that are largely different, but nevertheless exploit cross-reactivity to provide the same coverage of antigens; and the optimal repertoires can be reached via the dynamics of competitive binding of antigens by receptors and selective amplification of stimulated receptors. Our results follow from a tension between the statistics of pathogen detection, which favor a broader receptor distribution, and the effects of cross-reactivity, which tend to concentrate the optimal repertoire onto a few highly abundant clones. Our predictions can be tested in high-throughput surveys of receptor and pathogen diversity.

immune repertoires | optimal coding | repertoire diversity | competitive exclusion | jamming

The adaptive immune system protects organisms from a great variety of pathogens by maintaining a population of specialized cells, each specific to particular challenges. Together these cells cover the array of potential threats. To recognize pathogens, the immune system relies on receptor proteins expressed on the surface of its main constituents, the B and T lymphocytes. These receptors interact with antigens (small molecular elements making up pathogens), recognize them through specific binding, and initiate the immune response. Each lymphocyte expresses a unique receptor formed from random combinations encoded in the genome. The receptors later undergo selection through the death and division of the lymphocytes that express them, as well as mutations in the case of B lymphocytes. The diversity of the receptor repertoire determines the range of threats that the adaptive immune system can target.

The detailed composition of the immune receptor repertoire, and not just its breadth, is important for conferring effective protection against infections. Broadly speaking, a diverse population of receptors will confer wider immunity, and a larger clonal population of a particular receptor will confer more effective immunity against the pathogens to which it is specific. However, there is a trade-off between diversity and clone sizes because the number of receptors is limited. By selectively proliferating some receptors at the expense of others, the immune system retains a memory of past infections (1), facilitating subsequent immune responses. Furthermore, although infections increase the populations of receptors with the greatest specificity, they can also lead to a reorganization of the immune repertoire as a whole (2).

How should the repertoire be organized to minimize the cost of infections? We develop a framework for answering this question by abstracting key general features of the adaptive immune system: The receptor repertoire is bounded in size, receptors are "cross-reactive" (each antigen binds many receptors; each receptor binds many antigens), and the cost of an infection increases with time. Given these general assumptions, we consider a simplified landscape of pathogens, where infections are drawn from a fixed distribution. By simplifying the setting in this way, and independently of the detailed dynamics of immune responses, we arrive at broad insights about the composition of immune repertoires that are optimal for their pathogenic environments. Our framework is not meant to give a complete account of immunity. To do so we would need to include several other components of the immune system, such as interaction between its innate and adaptive arms and avoidance of autoimmunity. The latter problem—the challenge of discriminating self from nonself—has been the focus of many theoretical studies of the immune system (3, 4). This paper primarily investigates the relation between the adaptive repertoire and the pathogenic environment, but we also discuss how other components and constraints of the immune system can be incorporated into our model.

The theory predicts, counterintuitively, that the number of receptors specific to rare pathogens will be amplified relative to the probability of encounter, at the expense of receptors for common infections. We also find that two organisms responding to a pathogen distribution will display unique populations of immune receptors, even though their coverage of pathogens will be similar. How can the immune system achieve these sorts of optima? Surprisingly, we find that simple competition between receptor clones can drive the population to the optimal composition for minimizing the cost of infections.

New high-throughput methods are making it possible to survey B-cell and T-cell receptor diversity in fish (5, 6), in mice (2, 7), and in humans (8–11). As methods are developed to better characterize pathogenic landscapes and receptor cross-reactivity, predictions for the composition of optimal repertoires derived from our framework can be directly compared with experiments. To arrive at our results we ask how the immune system should be organized to perform its function well, rather than starting with the detailed dynamics of its components. We are proposing that the universal features of the adaptive immune system follow simply from general statistical considerations, whereas the detailed

Significance

The adaptive immune system uses the experience of past infections to prepare its limited repertoire of specialized receptors to protect organisms from future threats. What is the best way of doing this? Building a theoretical framework from first principles, we predict the composition of receptor repertoires that are optimally adapted to minimize the cost of infections from a given pathogenic environment. A naive repertoire can reach these optima through a biologically plausible competitive mechanism. Our findings explain how limited populations of immune receptors can self-organize to provide effective immunity against highly diverse pathogens. Our results also inform the design and interpretation of experiments surveying immune repertoires.

Author contributions: A.M., V.B., T.M., and A.M.W. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. Email: thierry.mora@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1421827112/-/DCSupplemental.

dynamical implementation arises from the historical contingencies of evolution.

Definition of the Problem

To find the optimal repertoire distribution we must consider the nature of antigen–receptor interactions and a penalty that the immune system pays for not recognizing antigens. This penalty must reflect the facts that recognition should happen within a reasonable time, before the pathogen colony can significantly increase its size; the interactions between the immune receptors and antigen are probabilistic; and not all antigens are equally frequent. We assume that, although the immune system cannot predict precisely which antigens it will encounter and when, it incorporates an estimate of the probabilities of their occurrences. We also take these probabilities to be constant in time. This is an idealization grounded in a separation of timescales, which assumes the distribution of antigens remains constant on timescales on which the immune system adapts.

The above ideas are the basis for our cost function, which reflects the penalty of nonrecognition, for a given repertoire and antigenic environment. In Fig. 1 we introduce a quantification of the problem. Given Q_a, the probability that the next infection will be caused by antigen a, we model the immune repertoire by a distribution of receptors P_r, from which lymphocytes with the receptor r are drawn at random.

An antigen a and a receptor r interact with a certain strength set by the binding affinity between the two molecules. This is described by the probability $f_{r,a}$ (which we call the "cross-reactivity function") that an antigen a encountering a receptor r results in a recognition event, leading to the activation of the lymphocyte expressing that receptor. Each encounter of the antigen a with a random receptor has a probability $\tilde{P}_a = \sum_r f_{r,a} P_r$ to lead to recognition and trigger an immune response. Thus, $\tilde{P}_a$ can be viewed as the coverage of antigen a by the repertoire.

Given this coverage, we consider $\overline{F}_a$—the average harm caused by antigen a. We show below that, consistent with intuition, $\overline{F}_a$ is a decreasing function of the coverage $\tilde{P}_a$. The overall expected cost is then just the harm averaged over the antigen distribution:

$$\mathrm{Cost}(\{P_r\}) = \langle F \rangle = \sum_a Q_a \overline{F}_a(\tilde{P}_a). \qquad [1]$$

The need to defend against many antigens at the same time with a limited number of receptors introduces a trade-off. If more receptors recognize an antigen, there are less to protect against other threats.

Finally, we derive an expression for the average harm $\overline{F}_a$ caused by antigen a. During its time in the periphery, an antigen a will encounter and possibly interact with receptors at a rate $\lambda_a(t)$ that increases with time as the pathogen population grows. Each encounter will occur with a different receptor r drawn from P_r. The mean number of encounters between antigens and receptors after a time t, which we call effective time, is defined as $m_a(t) = \int_0^t \mathrm{d}\tau \lambda_a(\tau)$, where $t = 0$ is set by the introduction of the antigen. The time t to the first recognition event, or response time, is random and depends on the coverage $\tilde{P}_a$.

The longer the system fails to detect the antigen, the more likely the infection is to become harmful. We assume that the integrated harm caused by an antigen since the beginning of an infection is an increasing function $F_a(t)$ of the time of first recognition. How exactly F_a grows with time may strongly depend on the type of infection and receptors (12–14). The mean harm inflicted to the organism by the attack of an antigen a is then given by this quantity averaged over the distribution of possible response times, $\overline{F}_a = \tilde{P}_a \int_0^{+\infty} dm\, F_a[t_a(m)]\, e^{-m\tilde{P}_a}$, where $t_a(m)$, the inverse function of $m_a(t)$, is the amount of time it takes for m encounters to occur between the receptors and pathogen a (see *SI Text, Appendix A* for a derivation). The result depends on the cost expressed as a function of the effective time m, $F_a[t_a(m)]$, which we denote $F_a(m)$ to simplify notation. We consider several specific choices of this effective cost function in *Results*.

Our aim here is to propose a general framework for thinking about the repertoire. Thus, we do not explicitly model intracellular communication, cell differentiation, activation of cofactors, coordination of different cell types, the interaction with the innate immune system, and the full complexity of the recognition process. The idea is that $F_a(m)$ implicitly summarizes all of these factors in terms of an effective cost.

In general the cost function $F_a(m)$ depends on the antigen a, reflecting the various virulences of different pathogens. To simplify, we can assume that the cost function takes the factorized form $F_a(m) = \mu_a F(m)$, where μ_a is the pathogen-dependent virulence factor, and $F(m)$ describes how all threats develop with time. The cost will then take the form $\sum_a \mu_a Q_a \tilde{P}_a \int_0^{\infty} dm F(m) e^{-m\tilde{P}_a}$. In this expression, the virulence factor μ_a of a pathogen plays the same role as its likelihood Q_a. Some pathogens are rare but very virulent (like anthrax), whereas others may be common but not very virulent (like the common cold), and an ideal immune system should be able to cope with both. In our model the overall "threat" of a pathogen is expressed as the product of the two, $\mu_a Q_a$. In practice μ_a can be absorbed into the definition of Q_a and is omitted in the rest of this paper.

Given such a model of the recognition process, there exists an optimal adaptive immune system, characterized by the choice of the receptor distribution P_r, that minimizes the expected cost in a given antigenic environment Q_a. The optimal repertoire is found by minimizing the expected cost in Eq. **1** with respect to P_r, subject to constraints of nonnegativity ($P_r \geq 0$) and normalization ($\sum_r P_r = 1$). Simple local extremality conditions are sufficient for optimality because our problem can be shown to be convex (*SI Text, Appendix B*). The condition $\sum_r P_r = 1$ is a normalized version of the constraint that the total number of receptors is limited.

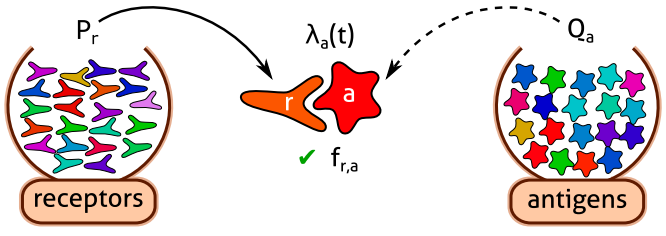

Fig. 1. Schematic of a statistical model of antigen recognition by the adaptive immune system. After infection, antigen a encounters immune receptor r at random with a rate $\lambda_a(t)$. An encounter leads to a successful recognition with a probability $f_{r,a}$ that reflects the matching between a given antigen–receptor pair.

Results

The Optimal Repertoire Is More Uniform Than the Pathogen Distribution. We can now ask how best to distribute the receptors to minimize the cost (Eq. **1**) for a given antigenic environment. To begin, we neglect cross-reactivity (later we will see that this is equivalent to looking at the structure of the repertoire at scales larger than the cross-reactivity). In this case antigens and receptors can be associated one by one by a cross-reactivity function $f_{r,a} = 1$ if $r = a$ and 0 otherwise. In this case we can analytically determine the optimal distribution (*SI Text, Appendix D, section 2*) $P_r^* = \max[\overline{F}'^{(-1)}(-\lambda/Q_r), 0]$, where $\overline{F}'^{(-1)}$ denotes the inverse function of the derivative of $\overline{F}_a = \overline{F}(\tilde{P}_a)$ expressed as a function of $\tilde{P}_a$, and λ is a positive constant fixed by the normalization $\sum_r P_r^* = 1$. Table 1 presents results for several representative cost functions.

A simple scenario occurs when the pathogen population grows exponentially in time, as do the cost and the encounter rate—reflecting the proliferative nature of pathogens. In this case the cost grows linearly in the number of encounters; i.e., $F(m) = m$ (*SI Text, Appendix C*). Then we find that the optimal fraction of the repertoire taken up by a given receptor is proportional to the

square root of the threat (combination of frequency and virulence) of the corresponding antigen $P_r^* \propto \sqrt{Q_r}$. Intuitively, we expect that the optimal repertoire should focus its resources on receptors recognizing the most common or virulent antigens. However, this enhanced protection against frequent or virulent antigens comes at the cost of a slower response against the uncommon and harmless antigens, and this bias toward more threatening antigens must remain limited. The square-root dependence reflects a particular trade-off between these two opposing constraints, by directing more resources toward more threatening antigens while uniformizing the distribution compared with a linear dependence. Intriguingly, the same square-root dependence has been found as an optimal solution for the size of tRNA pools as a function of codon use (15) and in a model for the screening of suspicious individuals (16).

The extent to which more resources are directed toward more threatening antigens depends on the relative gains and losses of earlier and later recognition events, which are captured in our model by the effective cost function $F(m)$. In general, steeper cost functions imply more flattened distributions of receptors. The cost function $F(m)=m^\alpha$, and its associated optimal distribution $P_r^* \propto Q_r^{1/(1+\alpha)}$, help illustrate this point. Such cost functions can arise when both $m(t)$ and $F(t)$ increase exponentially as a function of time, but with different exponents (*SI Text, Appendix C*). When α is large, the cost of nonrecognition increases very quickly with time, calling for an urgent response. Consequently the optimal immune system tends to cover the space uniformly to get all potential threats, even the unlikely ones, under control. Conversely, when α is small, the harm caused by pathogens does not explode with time, permitting the system to recognize the rarer pathogens late and focus its resources on the common ones.

In some situations, there may be little or even no difference between a late response and no response at all, because the total harm caused by an infection stabilizes. For example, consider the cost $F(m)=1-e^{-\beta m}$ that saturates at large effective times. In this case, the optimal solution (Table 1) relates receptor and antigen through a square root as for linear cost, but with a cutoff at low Q_a. This cutoff occurs because there is little benefit to having receptors recognizing rare or harmless antigens, whose recognition is likely to happen late, when differences in recognition times do not matter anymore. This result is consistent with the observation that the immune system may sometimes ignore infections if the harm they cause is too small, as in, e.g., the case of simian immunodeficiency virus infection of sooty mangabeys (17).

Real harm may occur only when the effective time m crosses a threshold. This situation can be modeled by taking $F(m)=\Theta(m-m_0)=0$ for $m<m_0$ and 1 otherwise. In this case the receptor distribution is organized to maximize the chance of detection before m_0. The optimal repertoire for this cost (Table 1) has no receptors for the least threatening pathogens (cutoff at low probabilities) and a drastically flattened receptor distribution (logarithm of the pathogen distribution).

Is there a cost function for which the receptor distribution is not flattened relative to the pathogen distribution? This occurs in a special case where cost increases very slowly (logarithmically) with effective time. However, in general, cost is minimized by a receptor distribution that is flattened relative to the pathogen distribution.

Table 1. Cost functions and optimal repertoires

$F(m)$	P_r^*	$A(\tilde{N}_a)$
m^α	$C\,Q_r^{1/(1+\alpha)}$	$C'(N_{st}/\tilde{N}_a)^{1+\alpha}$
$\ln m$	$C\,Q_r$	$C'(N_{st}/\tilde{N}_a)$
$1-\exp(-\beta m)$	$\max\{C\sqrt{Q_r}-\beta, 0\}$	$C'/(\beta+\tilde{N}_a/N_{st})^2$
$\Theta(m-m_0)$	$\max\{\ln(Q_r)/m_0-C, 0\}$	$C'\exp(-m_0\tilde{N}_a/N_{st})$

The cost function $F(m)$ measures the harm caused to an organism by the time that immune receptors have had m encounters with a pathogen. The optimal receptor distribution P^* is determined by minimizing this cost, given a pathogen distribution Q and a cross-reactivity function $f_{r,a}$ specifying the probability that receptor r binds to antigen a. The second column gives the form of P^* over scales larger than the cross-reactivity. The optimal P^* can be reached as a steady state resulting from competitive binding between receptors and antigens (last section of *Results*) quantified by an "availability function" A. $\tilde{N}_a=\sum_r N_r f_{r,a}$ represents the coverage of antigen a by the repertoire; $N_{st}=\sum_r N_r$ is the total steady-state population; and C, C', β, and m_0 are positive constants.

Cross-Reactivity Dramatically Reduces Diversity in the Optimal Repertoire. By allowing receptors to bind to a variety of antigens, cross-reactivity should permit the immune system to reduce the number of receptor types required to cover the whole range of possible threats. We show that given sufficient cross-reactivity, the optimal immune repertoire concentrates all its resources on a few receptors, which together tile antigenic space.

Following Perelson and Oster (4), we think of receptors and antigens as points in a common high-dimensional shape space, whose coordinates are associated to unspecified physicochemical properties. For simplicity, assume that cross-reactivity depends only on the relative position of receptor and antigen in shape space $f_{r,a}=f(r-a)$, where f is a decreasing function of the distance between a and r. Short distances in shape space correspond to a good fit between the two molecules, leading to strong recognition, whereas long distances translate into weak interactions and poor recognition.

To build intuition, we first consider an analytically solvable example (Fig. 2). We describe the space of receptors and antigens by a single continuous number and assume a Gaussian antigen distribution with variance σ_Q^2 and Gaussian cross-reactivity of width σ, which sets the typical distance within which a receptor and antigen interact. We derive the optimal receptor distributions for costs of the form $F(m)=m^\alpha$ (*SI Text, Appendix D, section 3b*). For narrow cross-reactivities ($\sigma<\sigma_c=\sigma_Q\sqrt{1+\alpha}$), the optimal receptor distribution is Gaussian with variance $(1+\alpha)\sigma_Q^2-\sigma^2$ and the optimal cost is independent of σ. For wide cross-reactivities ($\sigma>\sigma_c$), the receptors are optimally of a single type with reactivity centered on the pathogen distribution, whereas the optimal normalized cost increases with σ because the receptor is unnecessarily broadly reactive. These results arise from a tension between two opposing tendencies. As in the non-cross-reactive case, the need to cover rare pathogens broadens the optimal receptor distribution relative to the pathogen distribution. However, cross-reactivity has the opposite effect, favoring more concentrated distributions.

Does cross-reactivity generically drive the optimal receptor distribution to cluster into peaks? We investigated this question numerically. For concreteness, we consider a linear cost $F(m)=m$ and random pathogen environments in one or two dimensions constructed by drawing each Q_a from a log-normal distribution characterized by a coefficient of variation κ. The shape space is taken to be bounded and discretized, and we use accelerated gradient projection optimization (*SI Text, Appendix E*). We find that the optimal repertoire P^* is strongly peaked on a discrete forest of receptors (Fig. 3 *A* and *B*). The width of these peaks decreases as numerical precision is increased, suggesting that in a continuous limit the optimum consists of a weighted sum of Dirac delta functions, i.e., distinct, discretely spaced receptors in different amounts (Fig. S1).

By inspection, the peaks tend to repel each other and to organize into local tiling patterns, as further evidenced by the damped oscillations in the radial distribution function (18) (*SI Text, Appendix F* and Fig. S2*A*). This exclusion is a sensible way to distribute resources, as it limits redundant protection against the same pathogens. The spacing between peaks roughly follows the cross-reactivity scale σ, suggesting that P^* is smooth when viewed at scales larger than σ. Confirming this, $\tilde{P}^*$ (i.e., the coverage of the antigenic space by the receptors) smoothly tracks the variations in the antigen distribution Q at a broad scale (Fig. 3*A*). When viewed coarsely in this way, cross-reactivity is irrelevant and P^* tends to the solutions of Table 1. Indeed, at these broad

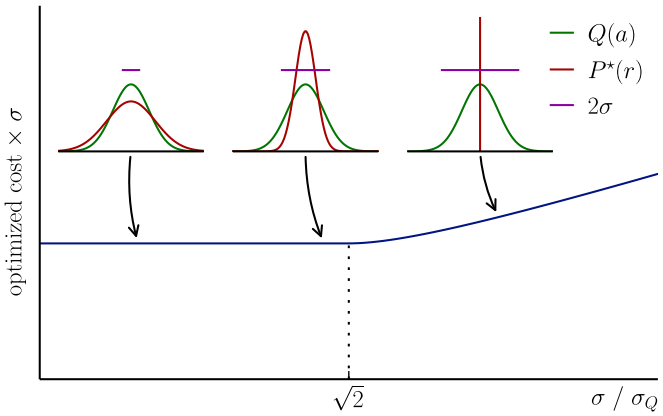

Fig. 2. The optimal cost and receptor distributions for protecting against a one-dimensional Gaussian antigenic landscape $Q(a)$ of variance σ_Q^2, as a function of the cross-reactivity width σ. As σ increases, the optimal distribution $P^\star(r)$ becomes narrower and narrower (*Left* and *Center Insets*), until it concentrates entirely onto a single point, for $\sigma \geq \sqrt{2}\sigma_Q$ (*Right Inset*). The minimal cost (multiplied by σ for a comparison at constant recognition capability) is constant below the transition point, but increases with σ past it. The cross-reactivity function, which quantifies the affinity between receptor r and antigen a as a function of their distance in shape space, has a Gaussian form, $f(r-a) = \exp[-(r-a)^2/2\sigma^2]$, and the cost function is linear in the effective recognition time, $F(m) = m$.

scales, the distribution of peaks is uniform, as demonstrated by the very low power in the spectrum of P_r at small wave vectors (*SI Text*, *Appendix F* and Fig. S2*B*), indicating that the number of receptors contained in any given large area of the shape space is very reproducible. This phenomenon of small-scale randomness with large-scale regularity is called disordered hyperuniformity (19) and arises in jammed packings as evidence of the incompressibility of the material. In biological terms, hyperuniformity means that the distribution of receptor peaks provides a much more uniform coverage of the antigen space than if the peaks were positioned randomly according to a Poisson distribution. For our optimal repertoires small-scale fluctuations get smoothed out by cross-reactivity and can be tolerated, whereas at large scales the fluctuations track variations in the antigenic landscape to provide smooth coverage (Fig. S3).

To test the generality of our findings we tested other choices of cross-reactivity functions (*SI Text*, *Appendix G*). We have so far assumed a unique scale σ for cross-reactivity, consistent with recent reports that cross-reactivity is local in antigenic space (20). However, receptor–antigen recognition can be very specific and sensitive to single mutations (21, 22) or extremely degenerate across very dissimilar antigens (23). To account for these long-range effects, we also tested fat-tailed cross-reactivity functions. We found that, for a variety of non-Gaussian cross-reactivity functions, long tailed or not, the optimal repertoire remains strongly peaked, although the position, number, and strength of the peaks do change (Fig. S4). Next, to relax the assumption that the cross-reactivity width is uniform across receptors, we tested receptor-dependent cross-reactivities σ_r drawn from a log-normal distribution. While the regularity of the local tiling structure is affected by this additional heterogeneity (just as we expect in a packing of spheres of variable size), the large-scale hyperuniformity is nonetheless preserved (Fig. S5). Next we considered distributions of antigens with correlations across shape space (reflecting, e.g., phylogenic correlations between pathogens). Again we find peaked optimal receptor distributions (Fig. S6), similar to those for uncorrelated antigen landscapes. For computational reasons, we restricted our analysis to 2D pathogen landscapes, but the analogy with random packing problems that we discussed above allows us to expect that all of these results will hold generally in higher dimensions. Finally, we incorporated the avoidance of the self by excluding from the optimization all receptors within distance σ of a set of randomly positioned self-antigens (*SI Text*, *Appendix H* and Fig. S7). We find that receptors are likely to be found near the boundary of these exclusion zones, but otherwise keep the same general tiling structure.

In summary, the optimal immune repertoire looks random at scales smaller than the cross-reactivity, but has the structure of a disordered tiling at larger scales so that, after accounting for cross-reactivity, the repertoire smoothly covers the pathogen landscape. These findings have an important consequence for different individuals exposed to the same pathogenic environment. Each individual will experience a slightly different spectrum of antigens because of the statistics of encounters and other sources of variability. These slightly different experiences of the same world lead to optimal repertoires with a striking property—the receptor distributions are largely different, even though their coverage of the pathogen landscape is similar after including cross-reactivity (Fig. 4). This finding can be compared with surveys of "public" repertoires of immune receptors (2, 24).

The Optimal Repertoire Can Be Reached Through Competition for Antigens. The results presented so far have established how repertoires should be structured to provide optimal protection. Given the complex interdependences between receptors arising from local and global trade-offs, one might think that the globally optimal solution could be reached only via some biologically implausible centralized mechanism distributing resources system-wide. In fact, we show that the optimal repertoire can be reached through self-organization, via competitive evolution of receptor populations under antigen stimulation.

We consider a model that is similar to that introduced by de Boer and Perelson (25) and de Boer et al. (26) for competitive dynamics of B and T cells (*SI Text*, *Appendix I*). Its main assumptions are that division of receptor-expressing lymphocytes is driven by antigen stimulation and that receptors compete for the limited supply of antigens. At each time step, a random antigen a is drawn from the distribution Q_a. Each receptor type r responds to it by expanding or shrinking its population N_r according to its specificity, by an amount $N_r \Delta t[A(\sum_{r'} N_{r'} f_{r',a}) f_{r,a} - d]$, where Δt is the time step: Receptors proliferate upon successful recognition of antigens (first term) and die with a constant rate d (second term). In the absence of competition, the proliferation rate should be proportional to $f_{r,a}$, but the antigen a may also bind other receptors, reducing its availability for receptor r. The coverage of antigen a by the repertoire, $\tilde{N}_a = \sum_r N_r f_{r,a}$, quantifies the breadth of the receptor pool competing to bind with a. The availability of antigen a for binding is assumed to be a decreasing function $A(\tilde{N}_a)$ of its coverage. The stimulation of r by a is thus

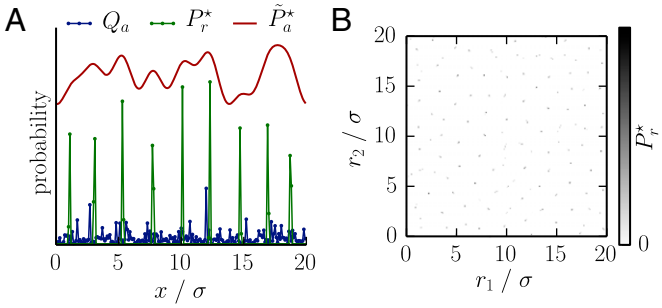

Fig. 3. Cross-reactivity plays an important role in shaping the optimal repertoire, often leading to highly peaked repertoires. (*A* and *B*) The optimal receptor distribution P_r^* for (*A*) 1D and (*B*) 2D random environments. Despite being peaked, the optimal distribution of receptors covers the antigenic space fairly uniformly, as shown by its coverage by the receptors, $\tilde{P}_a^* = \sum_r f_{r,a} P_r^*$, shown in the one-dimensional case (*A*). The cross-reactivity and cost functions are the same as in Fig. 2. The antigenic landscape Q_a is generated randomly from a log-normal distribution with coefficient of variation $\kappa = 1$.

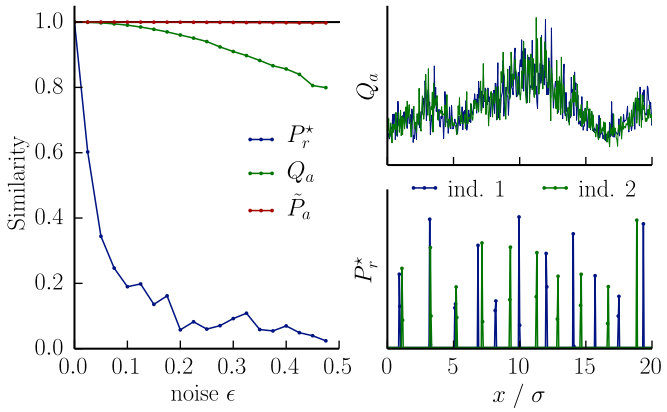

Fig. 4. Two individuals in the same environment Q_a that see it with slightly different noises have similar coverages of the antigenic space, but achieve it with different receptors. This results in largely nonoverlapping repertoires. Shown are the overlaps (normalized to be between 0 and 1) between the experienced pathogen distributions Q_a, the resulting optimal receptor distributions P_r^*, and the corresponding coverages $\tilde{P}_a$, as a function of the noise ε with which individuals perceive the environment. *Right* plots show an example of antigenic environments and optimal receptor distributions for $\epsilon = 0.2$. We calculated the optimal receptor distributions for two individuals 1 and 2 experiencing respective environments Qe^{z_1} and Qe^{z_2}, where Q is a random environment with fluctuations on scales larger than the cross-reactivity σ [power spectrum $\propto 1/(1+(10q\sigma)^2)$] normalized so that its coefficient of variation is 0.5, and z_1, z_2 are Gaussian noises of mean zero and variance ϵ^2. The choice of cost and cross-reactivity functions are the same as in Fig. 2.

modified to $A(\tilde{N}_a)f_{r,a}$ in the equation for the growth rate. In the limit of fast sampling of antigens, or mean-field limit ($\Delta t \to 0$), these stochastic dynamics are well described by the deterministic differential equations

$$\frac{\mathrm{d}N_r}{\mathrm{d}t} = N_r\left[\sum_a Q_a A\left(\sum_{r'} N_{r'} f_{r',a}\right) f_{r,a} - d\right]. \qquad [2]$$

For a given pathogenic environment, the total steady-state receptor population size N will be set by the death rate d, which counterbalances growth at steady state. Although in reality the ability of the system to reorganize itself diminishes with age, for simplicity we take all rates to be constant.

The stable fixed points of the mean-field dynamics (2) realize the optimal repertoires of the previous sections when the availability function A is matched to the cost function $F(m)$ through the relation

$$A(\tilde{N}_a) = -c'\overline{F}'\left(\frac{\tilde{N}_a}{N_{\mathrm{st}}}\right), \qquad [3]$$

where N_{st} is the total number of receptors $\sum_r N_r$ at steady state. Table 1 shows $A(\tilde{N})$ for several cost functions. To understand this result, first note that when binding is not cross-reactive, the dynamical equations for each receptor are independent and read $\mathrm{d}N_r/\mathrm{d}t = N_r(Q_r A(N_r) - d)$. The availability function now depends only on N_r, meaning that receptors compete only with their own kind—they occupy their own antigenic niche. The steady-state size of clone r is thus set by the carrying capacity of that niche, $N_r = A^{(-1)}(d/Q_r)$, or zero if that capacity is negative. With the availability given by Eq. **3**, this reproduces the optimal repertoire. As seen in Table 1, fast-growing cost functions correspond to very load-sensitive availability functions. In these cases, rare infections are almost as threatening as frequent ones; therefore the growth of the receptors that are specific to frequent antigens is actively limited to leave room for other receptors. The correspondence of Eq. **3** holds when receptor binding is cross-reactive (*SI Text*, *Appendix J*).

Cross-reactivity leads to competition among receptor types, effectively enforcing an exclusion between similar receptors. This phenomenon, known in ecology as competitive exclusion, is important for lymphocyte dynamics (25) and provides the mechanism by which our dynamical model reproduces the discrete clustering found in the optimal receptor distribution.

To check that the dynamics do converge to the optimum, we simulated numerically the full stochastic dynamics, as well as their mean-field limit (Eq. **2**), for a random antigenic environment in two dimensions, with $A(\tilde{N}) = 1/(1+\tilde{N}/N_0)^2$. Fig. 5 shows the dynamics of the receptor distribution $P_r(t) = N_r(t)/\sum_{r'} N_{r'}(t)$, as well as its cost relative to the optimal solution, as a function of time. Starting from a random distribution of receptors, the repertoire reorganizes into localized peaks that become increasingly prominent and well separated with time. Three independent runs of the stochastic simulation all converge approximately to the global minimum of the cost, with most of the improvement achieved within a few cell lifetimes. Convergence is exact in the mean-field limit, indicating that the steady-state solution is indeed optimal.

In summary, competitive dynamics can allow the immune repertoire to self-organize into a state that confers high protection against infections. In the special case when the availability A is scale invariant, the expected cost is a Lyapunov function of the dynamics (*SI Text*, *Appendix K*), implying that the optimum is reached regardless of the initial condition. Note, however, that the dynamics of Eq. **2** are expected to slow down with age, as the plasticity of the adaptive system decreases due to the diminishing number of naive cells (27).

Discussion

We introduced a general framework for predicting the optimal composition of the immune repertoire to minimize the cost of infections contracted from a given distribution of antigens. This framework can be extended in several ways to be more biologically

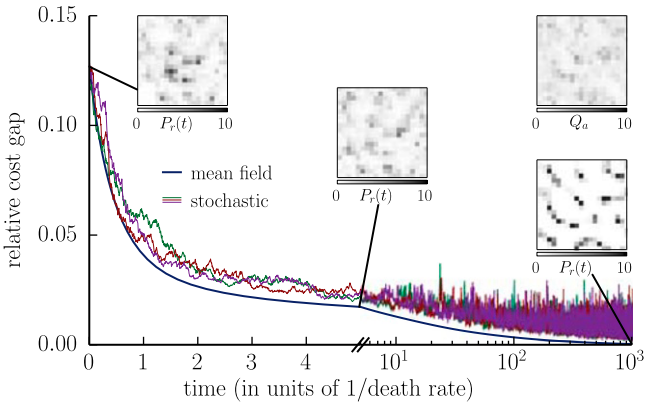

Fig. 5. The immune repertoire can self-organize to a state that minimizes cost and provides protection against infections via competitive evolution of receptor populations stimulated by antigens. Numerical simulations of the population dynamics, as well as its mean-field limit (Eq. **2**), show how competition causes a random initial receptor distribution to fragment into a highly peaked pattern [*Insets* represent $P_r(t) = N_r(t)/\sum_{r'} N_{r'}(t)$]. *Top Right Inset* represents the antigenic environment Q_a driving the dynamics [generated from a lognormal noise of power spectrum $\propto 1/(1+(5q)^2)$ and coefficient of variation 1]. Departure from optimality, as measured by the relative cost gap $[\langle F\rangle(P_r(t)) - \langle F\rangle(P_r^*)]/\langle F\rangle(P_r^*)$, decreases with time and eventually reaches zero in the mean-field limit. The three independent runs of the stochastic dynamics show reproducible results. We use the availability function $A(\tilde{N}) = 1/(1+\tilde{N}/N_0)^2$ with $N_0 = 10^6$, a death rate $d = 0.001$, and a cost function $F(m) = 1 - e^{-\beta m}$ with $\beta = 1/110$. The space size is 10σ. The initial condition was drawn from a lognormal noise of power spectrum $\propto 1/(1+(5q)^2)$, with coefficient of variation 2 and $\sum_r N_r(0) = 1.1\times 10^8$. In the stochastic simulations, the time between antigen presentations is $\Delta t = 0.005 d^{-1}$ (200 infections per cell lifetime).

faithful, e.g., by accounting for antigen-dependent infection dynamics and evolution of the pathogenic landscape. Our predictions can be tested in experiments that study how the environment influences the composition of immune repertoires, either via high-throughput sequencing surveys of receptor populations (2, 28) or by sequencing receptors specific to given antigens (14). The comparison between theory and experiment will provide insight into the functional constraints of antigen recognition by the immune system.

There are many situations where living systems must respond to very diverse and often very high-dimensional spaces of external influences, using strictly limited resources. To sense, internally represent, and then respond to these influences, organisms often use a large diversity of components, such as cell types or genes (29), each sensitive to a small part of the space. For example, the retina supports a diverse population of ganglion cell types, each sensitive to a different visual feature, that collectively represent the behaviorally salient aspects of visual scenes (30). Likewise, the mammalian olfactory system contains some $\sim$ 1,000 distinct receptors that each bind widely to odorants and collectively cover olfactory space (31). In these cases, the limited repertoire of component types provides a key constraint on information processing. Faced with such constraints, living systems must commit resources wisely, adapting to the structure of the environment and balancing breadth of coverage against depth of resolution, in light of priorities, costs, and constraints (32). We have shown that these elements also shape the optimal form of the immune repertoire.

Our finding that cross-reactivity causes the optimal repertoire to fragment is related to the concept of limiting similarity due to competitive exclusion in ecological settings (33). In this context, empty regions of phenotypic space result when competition is important on the scale at which resources vary, and continuous coexistence of species occurs only in exceptional cases (34). In general, niche-space heterogeneity promotes species clustering (35), recalling our finding that any heterogeneous antigen distribution leads to fragmentation of the optimal repertoire. The conceptual connection between the immune repertoire and ecological organization is even clearer in our dynamical model where species compete for an array of resources (the antigens) and grow in relation to their success in securing resources.

Although this study relies on a simple abstraction of the adaptive immune system, we expect that our framework and results will extend to other distributed protection systems where diverse threats are addressed by an array of specific responses. For example, the immune system of bacteria, or the clustered regularly interspaced short palindromic repeats (CRISPR) system (36), for which population dynamics models have already been proposed (37), could be studied within a similar framework to predict the relative abundance of CRISPR spacers and corresponding viruses in a coevolving population of bacteria and viruses.

ACKNOWLEDGMENTS. We thank O. Rivoire and F. Zamponi for helpful discussions. The work was supported by European Research Council Starting Grant 306312. V.B. was supported by the Fondation Pierre-Gilles de Gennes and National Science Foundation (NSF) Grants PHY-1058202 and EF-0928048. Portions of this work were done at the Aspen Center for Physics, supported by NSF Grant PHY-1066293. A.M. was supported by a German Academic Exchange Service (DAAD) Promos stipend.

1. Burnet FM (1976) A modification of Jerne's theory of antibody production using the concept of clonal selection. *CA Cancer J Clin* 26(2):119–121.
2. Thomas N, et al. (2014) Tracking global changes induced in the CD4 T-cell receptor repertoire by immunization with a complex antigen using short stretches of CDR3 protein sequence. *Bioinformatics* 30(22):3181–3188.
3. Forrest S, Perelson A, Allen L, Cherukuri R (1994) Self-nonself discrimination in a computer. *Proceedings of the IEEE Symposium on Research in Security and Privacy* (IEEE, Oakland, CA), pp 202–212.
4. Perelson AS, Oster GF (1979) Theoretical studies of clonal selection: Minimal antibody repertoire size and reliability of self-non-self discrimination. *J Theor Biol* 81(4):645–670.
5. Weinstein JA, Jiang N, White RA, 3rd, Fisher DS, Quake SR (2009) High-throughput sequencing of the zebrafish antibody repertoire. *Science* 324(5928):807–810.
6. Mora T, Walczak AM, Bialek W, Callan CG, Jr (2010) Maximum entropy models for antibody diversity. *Proc Natl Acad Sci USA* 107(12):5405–5410.
7. Ndifon W, et al. (2012) Chromatin conformation governs T-cell receptor $J\beta$ gene segment usage. *Proc Natl Acad Sci USA* 109(39):15865–15870.
8. Robins HS, et al. (2009) Comprehensive assessment of T-cell receptor β-chain diversity in $\alpha\beta$ T cells. *Blood* 114(19):4099–4107.
9. Larimore K, McCormick MW, Robins HS, Greenberg PD (2012) Shaping of human germline IgH repertoires revealed by deep sequencing. *J Immunol* 189(6):3221–3230.
10. Murugan A, Mora T, Walczak AM, Callan CG, Jr (2012) Statistical inference of the generation probability of T-cell receptors from sequence repertoires. *Proc Natl Acad Sci USA* 109(40):16161–16166.
11. Six A, et al. (2013) The past, present, and future of immune repertoire biology - the rise of next-generation repertoire analysis. *Front Immunol* 4:413.
12. Antia R, Ganusov VV, Ahmed R (2005) The role of models in understanding CD8+ T-cell memory. *Nat Rev Immunol* 5(2):101–111.
13. Regoes RR, Barber DL, Ahmed R, Antia R (2007) Estimation of the rate of killing by cytotoxic T lymphocytes in vivo. *Proc Natl Acad Sci USA* 104(5):1599–1603.
14. 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.
15. Xia X (1998) How optimized is the translational machinery in *Escherichia coli*, *Salmonella typhimurium* and *Saccharomyces cerevisiae*? *Genetics* 149(1):37–44.
16. Press WH (2009) From the Cover: Strong profiling is not mathematically optimal for discovering rare malfeasors. *Proc Natl Acad Sci USA* 106(6):1716–1719.
17. Silvestri G, et al. (2003) Nonpathogenic SIV infection of sooty mangabeys is characterized by limited bystander immunopathology despite chronic high-level viremia. *Immunity* 18(3):441–452.
18. Chaikin PM, Lubensky TC (2000) *Principles of Condensed Matter Physics* (Cambridge Univ Press, Cambridge, UK), Vol 1.
19. Torquato S, Stillinger FH (2003) Local density fluctuations, hyperuniformity, and order metrics. *Phys Rev E Stat Nonlin Soft Matter Phys* 68(4 Pt 1):041113.
20. Birnbaum ME, et al. (2014) Deconstructing the peptide-MHC specificity of T cell recognition. *Cell* 157(5):1073–1087.
21. Huseby ES, et al. (2005) How the T cell repertoire becomes peptide and MHC specific. *Cell* 122(2):247–260.
22. Kosmrlj 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.
23. Bhardwaj V, Kumar V, Geysen HM, Sercarz EE (1993) Degenerate recognition of a dissimilar antigenic peptide by myelin basic protein-reactive T cells. Implications for thymic education and autoimmunity. *J Immunol* 151(9):5000–5010.
24. Venturi V, Price DA, Douek DC, Davenport MP (2008) The molecular basis for public T-cell responses? *Nat Rev Immunol* 8(3):231–238.
25. De Boer RJ, Perelson AS (1994) T cell repertoires and competitive exclusion. *J Theor Biol* 169(4):375–390.
26. De Boer RJ, Freitas AA, Perelson AS (2001) Resource competition determines selection of B cell repertoires. *J Theor Biol* 212(3):333–343.
27. Miller RA (1996) The aging immune system: Primer and prospectus. *Science* 273(5271): 70–74.
28. Vollmers C, Sit RV, Weinstein JA, Dekker CL, Quake SR (2013) Genetic measurement of memory B-cell recall using antibody repertoire sequencing. *Proc Natl Acad Sci USA* 110(33):13463–13468.
29. Tkačik G, Walczak AM, Bialek W (2009) Optimizing information flow in small genetic networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 80(3 Pt 1):031920.
30. Gollisch T, Meister M (2010) Eye smarter than scientists believed: Neural computations in circuits of the retina. *Neuron* 65(2):150–164.
31. Buck L, Axel R (1991) A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell* 65(1):175–187.
32. Balasubramanian V, Sterling P (2009) Receptive fields and functional architecture in the retina. *J Physiol* 587(Pt 12):2753–2767.
33. MacArthur R, Levins R (1967) The limiting similarity, convergence, and divergence of coexisting species. *Am Nat* 101:377–385.
34. Pigolotti S, López C, Hernández-García E (2007) Species clustering in competitive Lotka-Volterra models. *Phys Rev Lett* 98(25):258101.
35. Szabó P, Meszéna G (2006) Limiting similarity revisited. *Oikos* 112:612–619.
36. Marraffini LA, Sontheimer EJ (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* 11(3):181–190.
37. He J, Deem MW (2010) Heterogeneous diversity of spacers within CRISPR (clustered regularly interspaced short palindromic repeats). *Phys Rev Lett* 105(12):128102.

Supporting Information

Mayer et al. 10.1073/pnas.1421827112

SI Text

Appendix A: Probability Distribution of the Time of First Recognition

To calculate the cost of not recognizing an antigen a, we need to find the distribution of times when a successful encounter takes place. The probability of having the first recognition of antigen a by receptor r in the time between t and $t+\mathrm{d}t$ reads

$$H_a(t)\mathrm{d}t = \lambda_a(t)\mathrm{d}t \cdot \sum_r P_r f_{r,a} \times \lim_{N\to\infty} \prod_{i=1}^{N} \left(1 - \lambda_a(t_i)\frac{t}{N}\sum_r P_r f_{r,a}\right),$$

where the first term is the probability of having an encounter between t and $t+\mathrm{d}t$, the second term is the probability of this encounter being successful, and the third term is the probability of there not being any prior recognition events. For the calculation of the last term we have decomposed the time leading up to t into N intervals of length t/N. Taking the $N\to\infty$ limit yields

$$H_a(t) = \lambda_a(t)\tilde{P}_a e^{-\int_0^t dt'\lambda_a(t')\tilde{P}_a}, \quad [S1]$$

where we have used the shorthand notation $\tilde{P}_a = \sum_r P_r f_{r,a}$ for the probability that a randomly chosen receptor recognizes antigen a.

Appendix B: Convexity of the Expected Cost

We show that the cost function $\langle F\rangle$ is a convex function of its argument $\{P_r\}$ (the receptor distribution). We start by introducing an alternative expression of $\overline{F}_a$, obtained by integration by parts:

$$\overline{F}_a = \int_0^\infty \mathrm{d}m F_a'(m) e^{-m\tilde{P}_a} + F(0). \quad [S2]$$

We calculate the derivatives of this average cost with respect to $\tilde{P}_a$:

$$\frac{\mathrm{d}\overline{F_a}}{\mathrm{d}\tilde{P}_a} = -\int_0^\infty \mathrm{d}m\, m F_a'(m) e^{-m\tilde{P}_a} \quad [S3]$$

$$\frac{\mathrm{d}^2\overline{F_a}}{\mathrm{d}\tilde{P}_a^2} = \int_0^\infty \mathrm{d}m\, m^2 F_a'(m) e^{-m\tilde{P}_a}. \quad [S4]$$

Because by assumption $F_a'(m)$ is positive, the second derivative of $\overline{F}_a$ with respect to $\tilde{P}_a$ is positive. This establishes the convexity of $\overline{F}_a$ as a function of $\tilde{P}_a$. Because $\langle F\rangle = \sum_a Q_a \overline{F}_a$ (with $Q_a \geq 0$), it is a convex function of $\{\tilde{P}_a\}$. Therefore, it is also a convex function of $\{P_r\}$, as $\{P_r\}$ and $\{\tilde{P}_a\}$ are linearly related.

Appendix C: Biological Motivation of Power-Law Cost Functions

In the main text we have developed a general framework for discussing the antigen–receptor recognition process. To fully specify the model we need to choose an effective cost function $F_a(m) = F_a(t_a(m))$. In the main text we derive optimal receptor distributions for a number of effective cost functions, including power-law functions $F(m) = m^\alpha$. Here we sketch plausible scenarios motivating that choice.

Consider an organism being infected with an antigen a. As long as there is no immune reaction, the antigen divides inside its host and thus increases its population size. If the initial population size is small, it is reasonable to assume exponential growth.

The more antigens there are at the time of the immune reaction, the more damage they can potentially do. Likewise, the more antigens there are, the higher the rate of encounters. These two quantities are also expected to grow exponentially in time:

$$F_a(t) = F_a(0)e^{\nu_a t}, \quad [S5]$$

$$\lambda_a(t) = \lambda_a(0)e^{\nu_a' t}. \quad [S6]$$

The two exponents may be different in general, because the number of pathogenic agents that cause the harm may grow differently than the number of antigens that can be recognized by the immune system. This difference could for example stem from the fact that both the pathogen's antigenic exposure and its virulence are cooperative effects and thus scale as a power of the number of invading individuals. Using $m_a(t) = \lambda_a(0)(e^{\nu_a' t} - 1)/\nu_a'$, and eliminating time $t \approx \ln[m_a/\lambda_a(0)]/\nu_a'$ (for t large compared with $1/\nu_a'$), we rewrite the effective cost function in terms of the number of encounters,

$$F_a(m) = F_a(0)\left(\frac{m}{\lambda_a(0)}\right)^{\nu_a/\nu_a'} \propto m^\alpha, \quad [S7]$$

with $\alpha = \nu_a/\nu_a'$.

Appendix D: Analytical Optimization

1. Optimality Conditions. In the following we give optimality conditions for the optimization problem defined in the main text, which are used for the following analytical determination of optimal receptor distributions. These conditions, called Karush–Kuhn–Tucker conditions (1), are derived from a generalization of the method of Lagrange multipliers to inequality as well as equality constraints.

The Lagrangian for the optimization problem is

$$\mathcal{L}(P,\lambda,\nu) = \langle F\rangle(P) + \lambda\left(\sum_r P_r - 1\right) - \sum_r \nu_r P_r, \quad [S8]$$

with

$$\langle F\rangle = \sum_a Q_a \overline{F}_a. \quad [S9]$$

λ is a Lagrange multiplier enforcing the normalization constraint and ν_r are Lagrange multipliers enforcing the nonnegativity constraint. The optimal P^* is an extremum of this Lagrangian. Therefore, the stationarity conditions

$$\left.\frac{\partial\langle F\rangle}{\partial P_r}\right|_{P*} + \lambda^* - \nu_r^* = 0, \quad [S10]$$

with

$$\frac{\partial\langle F\rangle}{\partial P_r} = \sum_a Q_a \overline{F}'_a(\tilde{P}_a) f_{r,a}, \quad \textbf{[S11]}$$

must hold for some value of λ^* and ν_r^* that enforce the constraints. The inequality constraint $P_r \geq 0$ further requires that

$$\nu_r^* \geq 0 \quad \textbf{[S12]}$$

$$\nu_r^* P_r^* = 0, \quad \textbf{[S13]}$$

where Eq. **S13** is known as the complementary slackness condition. It requires the Lagrange multipliers associated with the nonnegativity to be zero unless the constraint is active, i.e., unless the corresponding receptor probability is zero.

The three conditions may be reformulated as

$$\left.\frac{\partial\langle F\rangle}{\partial P_r}\right|_{P*} + \lambda^* \geq 0 \quad \textbf{[S14]}$$

$$\left(\left.\frac{\partial\langle F\rangle}{\partial P_r}\right|_{P*} + \lambda^*\right) P_r = 0. \quad \textbf{[S15]}$$

For all receptors that are present in the optimal repertoire ($P_r^* > 0$) these conditions imply

$$\left.\frac{\partial\langle F\rangle}{\partial P_r}\right|_{P*} = -\lambda^*. \quad \textbf{[S16]}$$

If a receptor is not present in the optimal repertoire ($P_r^* = 0$), then the less stringent condition holds:

$$\left.\frac{\partial\langle F\rangle}{\partial P_r}\right|_{P*} \geq -\lambda^*. \quad \textbf{[S17]}$$

We note here that $\partial\langle F\rangle/\partial P_r \leq 0$ (because more receptors always yield a lower cost), so that $\lambda^* \geq 0$.

These two conditions can be explained as follows: If a repertoire is optimal, all changes allowed by the constraints will lead to a higher cost; i.e., moving receptors from one type to another will not yield an improvement. All partial derivatives of the cost with respect to the receptor probabilities should thus be equal to the same value (Eq. **S16**). If there are already no receptors of a certain type, i.e., $P_r = 0$, we get a less stringent condition. We can no longer remove receptors away from this type r, but only add some to it, at the expense of other receptor types. The increase in cost due to the depletion of these other types should be higher than the gain of moving them to type r. The partial derivatives of the cost with respect to the receptors that are not present in the repertoire must thus be larger than the partial derivatives of the present receptors, which are given by $-\lambda^*$ (Eq. **S17**).

2. Solution for Uniquely Specific Receptors. We now solve Eqs. **S16** and **S17** for a repertoire of uniquely specific receptors (no cross-reactivity). Eq. **S11** becomes

$$\frac{\partial\langle F\rangle}{\partial P_r} = Q_r \overline{F}'_r(P_r), \quad \textbf{[S18]}$$

where we have used the fact that in the absence of cross-reactivity $\tilde{P}_a = P_a$. If all optimal receptor probabilities are positive, then we can insert this relationship into Eq. **S16** to obtain

$$Q_r \overline{F}'_r(P_r^*) = -\lambda^* \quad \textbf{[S19]}$$

and thus

$$P_r^* = h_r \frac{-\lambda^*}{Q_r}, \quad \textbf{[S20]}$$

where $h_r = \overline{F}'^{(-1)}_r$ denotes the inverse function of $\overline{F}'_r$. Because that function $\overline{F}'_r$ is always negative, h_r must take a negative argument.

For some cost functions, solving this equation may yield some negative receptor probabilities. In these cases some of the nonnegativity constraints need to be active. Setting $P_r = 0$ when Eq. **S20** is negative yields the correct optimal distribution under the nonnegativity constraint. We verify that for these r, Eq. **S17** is satisfied by $P_r = 0$, because

$$Q_r \overline{F}'_r(P_r = 0) \geq Q_r \overline{F}'_r\left[h_r\left(\frac{-\lambda^*}{Q_r}\right)\right] = -\lambda^*, \quad \textbf{[S21]}$$

where we have used the fact that $\overline{F}'_r$ is an increasing function of its argument (due to the positivity of its derivative; compare Eq. **S4**), and $h_r(-\lambda^*/Q_r) \leq 0$.

In summary, the solution to the optimization problem is

$$P_r^* = \max\left\{h_r\left(\frac{-\lambda^*}{Q_r}\right), 0\right\}, \quad \textbf{[S22]}$$

where the value of λ^* is fixed by the normalization condition $\sum_r P_r = 1$.

In Table S1 we give the explicit expressions of $\overline{F}_a$ and h_a, for the particular choices of the cost function $F(m)$ considered in the main text.

3. Solution for Cross-Reactive Receptors. The previous results can be generalized to cross-reactive receptors in a continuous space, using Fourier transforms. This generalization will lead up to the results presented in the *Cross-Reactivity Dramatically Reduces Diversity in the Optimal Repertoire* section of the main text and notably the Gaussian case discussed therein.

a. Deconvoluting the optimality conditions in Fourier space. We consider a continuous receptor–antigen space and we assume a translation invariant cross-reactivity function $f_{r,a} = f(r-a)$. We write the optimality condition Eq. **S16**,

$$\int \mathrm{d}a\, Q(a)\overline{F}'\left[\tilde{P}^*(a)\right] f(r-a) = -\lambda^*, \quad \textbf{[S23]}$$

where in continuous space the coverage is defined as

$$\tilde{P}(a) = \int \mathrm{d}r\, P(r) f(r-a). \quad \textbf{[S24]}$$

We note that both expressions involve integrals, which are convolutions with the cross-reactivity kernel. Because the convolution of a constant is also a constant, a solution of

$$Q(a)\overline{F}'\left(\tilde{P}^*(a)\right) = -\lambda', \quad \text{with } \lambda' > 0, \quad \textbf{[S25]}$$

is also a solution of Eq. **S23**. As in the case of uniquely specific receptors, we can solve this equation for $\tilde{P}^*(a)$,

$$\tilde{P}^*(a) = h\left[\frac{-\lambda'}{Q(a)}\right], \quad \textbf{[S26]}$$

where $h = \overline{F}'^{(-1)}$ as in Eq. **S20**. If there was no cross-reactivity, there would be no difference between P and $\tilde{P}$, and we would be done. Here we need to perform a deconvolution to obtain the optimal receptor distribution P from the optimal coverage $\tilde{P}$. We

do so in Fourier space, where the convolution turns into a product. Deconvolution is therefore much simpler in Fourier space as it corresponds to a division

$$\mathcal{F}[\tilde{P}] = \mathcal{F}[P]\mathcal{F}[f] \quad \Leftrightarrow \quad \mathcal{F}[P] = \frac{\mathcal{F}[\tilde{P}]}{\mathcal{F}[f]}, \qquad \textbf{[S27]}$$

where we have defined the Fourier transform of a function $g(x)$ as $\mathcal{F}[g](k) = \int_{-\infty}^{\infty} dx g(x) e^{ikx}$. To calculate the optimal receptor distribution we insert Eq. **S26** into Eq. **S27** and perform an inverse Fourier transform $\mathcal{F}^{-1}[\tilde{g}](x) = (1/2\pi)\int_{-\infty}^{\infty} dk \tilde{g}(k) e^{-ikx}$ to obtain

$$P^* = \mathcal{F}^{-1}\left[\frac{\mathcal{F}[h(-\lambda'/Q)]}{\mathcal{F}[f]}\right]. \qquad \textbf{[S28]}$$

This result is valid only as long as the above quantity is positive and normalizable, as we shall see below.

b. The Gaussian case. In this section we apply the general results of the previous section to a concrete example. To find the optimal receptor distribution analytically we use Eq. **S28**, we assume the antigen distribution and cross-reactivity function are Gaussian

$$Q(a) = \frac{1}{\sqrt{2\pi\sigma_Q^2}} \exp\left(\frac{-a^2}{2\sigma_Q^2}\right), \qquad \textbf{[S29]}$$

$$f(r-a) = \exp\left[\frac{-(r-a)^2}{2\sigma^2}\right], \qquad \textbf{[S30]}$$

and we take

$$F(m) = m^{\alpha}. \qquad \textbf{[S31]}$$

Inserting h from Table S1 into Eq. **S28** allows us to write

$$P^* \propto \mathcal{F}^{-1}\left[\frac{\mathcal{F}[Q^{1/(1+\alpha)}]}{\mathcal{F}[f]}\right] \qquad \textbf{[S32]}$$

as an equivalent equation determining the optimal repertoire. We can calculate the modified antigen distribution as

$$Q(a)^{1/(1+\alpha)} \propto \exp\left(-\frac{a^2}{2(1+\alpha)\sigma_Q^2}\right). \qquad \textbf{[S33]}$$

The Fourier transform of a Gaussian function of variance σ^2 is a Gaussian function of variance $1/\sigma^2$ (2). Therefore, we have

$$\mathcal{F}\left[Q^{1/(1+\alpha)}\right](q) \propto \exp\left[\frac{-(1+\alpha)\sigma_Q^2 q^2}{2}\right], \qquad \textbf{[S34]}$$

$$\mathcal{F}[f](q) \propto \exp\left[\frac{-\sigma^2 a^2}{2}\right], \qquad \textbf{[S35]}$$

from which

$$\frac{\mathcal{F}[Q^{1/(1+\alpha)}]}{\mathcal{F}[f]} \propto \exp\left\{\frac{-\left[(1+\alpha)\sigma_Q^2 - \sigma^2\right]q^2}{2}\right\} \qquad \textbf{[S36]}$$

follows. Taking the inverse Fourier transform and normalizing, we obtain

$$P^*(r) = \frac{1}{\sqrt{2\pi\left[(1+\alpha)\sigma_Q^2 - \sigma^2\right]}} \exp\left(-\frac{r^2}{2\left[(1+\alpha)\sigma_Q^2 - \sigma^2\right]}\right). \qquad \textbf{[S37]}$$

Normalization is possible only for $\sigma < \sigma_Q\sqrt{1+\alpha} \equiv \sigma_c$. In the limit $\sigma \to \sigma_c$ the Gaussian converges to a Dirac delta function. Intuition suggests that a Dirac delta function centered on the peak position should remain optimal for further increases in σ. To prove this assertion we note that a Dirac delta function is zero everywhere, except in one point. Because all but one receptor probabilities are at the boundary defined by the nonnegativity constraints, we need only to check Eq. **S17**. We compute the left-hand side of Eq. **S23** as a function of r,

$$\int dp\, Q(a)\overline{F}'[\tilde{P}^*(a)]f(r-a) \propto -\exp\left\{\frac{-r^2\left[\sigma^2 - (1+\alpha)\sigma_Q^2\right]}{2\sigma^2\left(\sigma^2 - \alpha\sigma_Q^2\right)}\right\}, \qquad \textbf{[S38]}$$

and note that it has a minimum for $r = 0$. This shows that the partial derivatives of the expected cost at $r \neq 0$ are greater than at $r = 0$, implying that Eq. **S17** holds.

The cost of the optimal repertoires as a function of the cross-reactivity width σ is given by

$$\langle F\rangle(P^*) = \left(\frac{\sigma_Q}{\sigma}\right)^{\alpha} \begin{cases} (1+\alpha)^{(1+\alpha)/2} & \text{if } \sigma < \sigma_c, \\ \dfrac{(\sigma/\sigma_Q)^{\alpha}}{\sqrt{1-\alpha(\sigma_Q/\sigma)^2}} & \text{otherwise.} \end{cases} \qquad \textbf{[S39]}$$

Both expressions give the same cost at the transition $\sigma = \sigma_c$. After multiplying by $(\sigma/\sigma_Q)^{\alpha}$ to compare at constant recognition capability $\int f = \sqrt{2\pi}\sigma$, this expression is constant for $\sigma < \sigma_c$ and grows for $\sigma > \sigma_c$.

c. General argument for peakedness. A simple argument can help us understand why cross-reactivity generically leads to peaked optimal solutions. The convolution with a kernel is a smoothening operation, represented by a low-pass filter in the Fourier domain. The optimal solution in the absence of the nonnegativity constraints requires that $\tilde{P}_a = h(Q_a)$. As $\tilde{P}_a$ is the low-pass filtered version of P_r, the high-frequency components of $h(Q_a)$ will be magnified by the deconvolution. These high-frequency wiggles can lead to negative values of $\mathcal{F}^{-1}[h(Q_a)]$, which are not allowed, leading us to set many values of $P(r)$ to zero. This effect results in a peaked solution. Because the size of the cross-reactivity kernel is inversely proportional to the cutoff frequency in the Fourier domain, we expect the spacing of the peaks to be related to the size of the cross-reactivity kernel.

Appendix E: Numerical Optimization

We numerically minimize the cost function subject to the normalization and nonnegativity constraints by using a fast projected-gradient algorithm. In the following we provide details on this numerical algorithm. To facilitate notations let us define the function to minimize as $g(x)$, where x is a vector in a Euclidean space, and the convex set C is defined by the constraints. In these notations the problem we want to solve can be stated as

$$\min_{x \in C} g(x). \qquad \textbf{[S40]}$$

Given an arbitrary starting point $x_0 \in C$ the algorithm performs the iterative procedure

$$y^{k+1} = x^k + \omega^k\left(x^k - x^{k-1}\right), \quad \text{[S41]}$$

$$x^{k+1} = \mathcal{P}\left(y^{k+1} - s^k \nabla g\left(y^{k+1}\right)\right), \quad \text{[S42]}$$

where ∇ denotes the gradient. Here $\mathcal{P}$ denotes a projection onto C, ω^k is an extrapolation step size, and s^k is the step size taken in the direction of the gradient. The extrapolation step size has to be chosen carefully to ensure the faster convergence of this method with respect to an ordinary gradient method. Following ref. 3 we use

$$\omega_k = \frac{k}{k+3}. \quad \text{[S43]}$$

The step size s is determined by backtracking (4): We iteratively decrease s by multiplication by $\beta < 1$ until $g(z) \le g(y^k) + (z - y^k)\cdot\nabla g(y^k) + (1/2s)(z-y^k)^2$, where $x\cdot y$ denotes the inner dot product between x and y, and $z = \mathcal{P}(y^k - s\nabla g(y^k))$. In practice we determine s in this way at the first step of the optimization and then keep it fixed based on this initial estimate.

The projection of a point y onto a convex set C is defined by the following quadratic programming problem:

$$\mathcal{P}(y) = \underset{x\in C}{\operatorname{argmin}} \frac{1}{2}(x-y)^2. \quad \text{[S44]}$$

If the convex set is a simplex as is the case for our problem, efficient algorithms fortunately exist for solving this problem. We use the algorithm described in ref. 5.

To stop the iteration one needs to define suitable stopping criteria. As the problem is convex we can establish a lower bound for the cost by solving a linear programming problem as follows:

$$g_{lb} = g\left(x^k\right) + \min_{x\in C}\left[\left(x - x^k\right)\cdot\nabla g\left(x^k\right)\right] \le g(x^*). \quad \text{[S45]}$$

The linear programming problem $\bar{x}^k = \operatorname{argmin}_{x\in C}\nabla g(x^k)^T(x - x^k)$ is solved explicitly (6) by

$$\bar{x}^k = e_{i^*},\ i^* = \underset{i}{\operatorname{argmin}}\left(\nabla g\left(x^k\right)\right)_i, \quad \text{[S46]}$$

where e_i denotes the ith unit vector. We can use this lower bound to define a stopping criterion for the numerical optimization

$$\frac{g\left(x^k\right) - g_{lb}}{g_{lb}} < \epsilon. \quad \text{[S47]}$$

For all reported numerical results we have chosen $\epsilon = 10^{-8}$.

To minimize finite size effects in the simulations we have used periodic boundary conditions in the receptor/antigen space. The discretization steps used in the figures are listed below:

Step	Figure
0.5σ	Fig. 5
0.1σ	Fig. 3 and Figs. S2–S4
0.05σ	Fig. 4

All source code associated with this paper is available online at dx.doi.org/10.5281/zenodo.16796.

Appendix F: Tiling Properties: Radial Distribution Function and Power Spectral Density of the Receptor Distribution

We analyze the tiling structure of the peaks in the optimal distribution P_r^* found in Fig. 3 of the main text. A useful technique, borrowed from condensed matter physics, is to measure the radial distribution function (7), $g(R) = \langle P(r)P(r')\rangle_{|r-r'|=R}$, where $|r-r'|$ is the distance between points r and r'. Fig. S2*A* presents $g(R)$ for P^* in two dimensions. The initial drop at small r indicates that peaks in P^* are rarely close—i.e., peaks in the optimal repertoire tend to repel each other. This exclusion, which operates over the range of strong cross-reactivity, is a sensible way to distribute resources, as it limits redundant protection against the same pathogens. The damped oscillation of the peaks of $g(R)$ confirms that the receptors in P^* are organized into a disordered tiling pattern. A similar radial distribution function is seen in high-density random packings of hard spheres where the spheres must cover as much space as possible but exclude each other. In both cases, the tiling ensures uniform coverage of space at large scales.

To quantify the regularity of the tiling, we calculate the normalized power spectral density of the 2D pattern, $S(q) = \sum_{r,r'} P_r P_{r'} e^{iq(r-r')}/\sum_r P_r^2$, where q is a wave vector. Large (small) $|q|$ correspond to short (long) distances in antigen shape space. When P_r is made of Dirac delta peaks of uniform heights, $S(q)$ coincides with the structure factor familiar in physics and satisfies $S(q\to\infty) = 1$. Fig. S2*B* shows $S(q)$ averaged over many realizations of the antigen landscape and over all directions of q so that it depends only on its modulus $|q|$. $S(q)$ approaches 1 for large q, showing that the precise local positions of the peaks are random. (The small departure from 1 is attributable to numerical discretization.) $S(q)$ is very low for small q, indicating that the number of receptors contained in any given large area of the shape space is very reproducible, providing uniform coverage—a property called hyperuniformity in the context of jammed materials (8–10). For our optimal repertoires small-scale fluctuations (large q) get smoothed out by cross-reactivity and can be tolerated, whereas at large scales the fluctuations track variations in the antigenic landscape to provide smooth coverage (Fig. S3).

Appendix G: Non-Gaussian, Long-Tailed, and Nonuniform Cross-Reactivity Functions

To assess the impact of different assumptions about the nature of cross-reactivity on the results we performed a number of simulations with different kernel functions.

First, we investigated the family of kernel functions defined by $f(r-a) = \exp[-(|r-a|/\eta)^\gamma]$ (Fig. S4 *A–D*, *Left*). By changing the parameter γ we can go from an exponential ($\gamma = 1$) via a Gaussian $\gamma = 2$ to a top-hat kernel ($\gamma\to\infty$). Up to $\gamma = 2$ all such kernels have positive Fourier transforms, whereas for $\gamma > 2$ the Fourier transforms also take negative values. The positive definiteness has been shown to be an important property in a related problem in ecology (11). Second, we also investigated how long-tailed kernel functions change the optimal repertoire by considering the functional form $f(r-a) = 1/(1 + (|r-a|\eta)^2)$ (Fig. S4 *A–D*, *Right*).

Dropping one further assumption, we investigated the influence of varying the width of the cross-reactivity function between receptors (Fig. S5). The width of the cross-reactivity was drawn randomly from a log-normal distribution with different coefficients of variation (corresponding to different amounts of scatter in the width). Biologically, the overall stimulatory capacity of receptors is constrained, and we rescaled the cross-reactivity so that all receptors had the same overall stimulatory potency.

Appendix H: Excluding Strongly Self-Binding Receptors

The presence of self-antigens that should not be recognized puts constraints on which receptors the repertoire might contain. As a first step to understand how such a requirement interacts with the trade-off considered in this paper we analyzed a simple model: A number of self-antigens are picked at random positions. The repertoire is not allowed to have receptors that are too highly reactive to any of the self-antigens. In practice this is ensured by adding a constraint to the optimization that none of the receptors in the repertoire can have a distance smaller than σ to any self-antigen. Introducing this constraint changes the optimal repertoire,

but key features such as the fragmentation of the repertoire and the tiling are conserved (Fig. S7).

Appendix I: Model for Receptor Dynamics

Here we describe our model for competitive receptor dynamics. We then show how, in a mean-field limit where antigen encounters are very frequent, this model reduces to a system of differential equations for the population dynamics.

At every step we update the number of receptors according to

$$\Delta N_r = \Delta t \cdot N_r \left[A\left(\sum_{r'} N_{r'} f_{r',a} \right) f_{r,a} - d \right], \qquad \text{[S48]}$$

where the antigen a is drawn randomly with probability Q_a and Δt is a parameter determining how much the repertoire changes per step.

In the limit where Δt is small the dynamics cycle through different encountered pathogens so fast that they effectively become the following dynamics:

$$\frac{\mathrm{d}N_r}{\mathrm{d}t} = N_r \left[\sum_a Q_a A\left(\sum_{r'} N_{r'} f_{r',a} \right) f_{r,a} - d \right]. \qquad \text{[S49]}$$

These dynamics are of mean-field type; i.e., they neglect the effect of the stochasticity in the encounter of pathogens.

Appendix J: The Stable Fixed Point of the Mean-Field Population Dynamics Minimizes the Cost Function

In this section we show that the stable fixed point $\{N_r^*\}$ of the population dynamics, Eq. **S49**, gives a probability distribution $P_r = N_r/N_{\mathrm{tot}}$ (with $N_{\mathrm{tot}} = \sum_r N_r$) that minimizes the cost $\langle F \rangle$. For this correspondence to be exact, the availability function of the dynamics and the effective cost function of the optimization must be related by

$$A(\tilde{N}_a) = -c' \overline{F}' \frac{\tilde{N}_a}{N_{\mathrm{st}}}, \qquad \text{[S50]}$$

where $\tilde{N}_a = \sum N_r f_{r,a}$, and N_{st} is the total number of receptors N_{tot} at the fixed point.

A fixed point is characterized by $\mathrm{d}N_r/\mathrm{d}t = 0$. If $N_r > 0$, this translates into

$$\sum_a Q_a A\left(\sum_{r'} N_{r'} f_{r',a} \right) f_{r,a} - d = 0. \qquad \text{[S51]}$$

Using the correspondence between availability and cost function given by Eq. **S50**, we rewrite this condition as

$$\sum_a Q_a \overline{F}'(\tilde{P}_a) f_{r,a} = -c' d, \qquad \text{[S52]}$$

which is equivalent to the optimality condition Eq. **S16**, with the identification $\lambda^* = c' d$.

For $N_r = 0$ we need to work a bit harder to show that the optimality condition at the boundary Eq. **S17** is satisfied. Here the key assumption establishing the minimization of the cost function is the stability of the fixed point. A fixed point is stable if the real parts of the Jacobian's eigenvalues are all negative. The Jacobian reads

$$\begin{aligned} J_{r,r'} = {} & \delta_{r,r'} \left(\sum_a Q_a A\left(\sum_{r''} N_{r''} f_{r'',a} \right) f_{r,a} - d \right) \\ & + N_r \sum_a Q_a A'\left(\sum_{r''} N_{r''} f_{r'',a} \right) f_{r,a} f_{r',a}. \end{aligned} \qquad \text{[S53]}$$

We remark that for $N_r = 0$ the rth row of the Jacobian is nonzero only on the diagonal. That value on the diagonal is an eigenvalue of the Jacobian and must be negative:

$$\sum_a Q_a A\left(\sum_{r'} N_{r'} f_{r',a} \right) f_{r,a} - d < 0. \qquad \text{[S54]}$$

Again we replace $A(\sum_r N_r f_{r,a})$ by $-\overline{F}'_a(\tilde{P}_a)$ according to Eq. **S50** to obtain

$$\sum_a Q_a \overline{F}'(\tilde{P}_a) f_{r,a} > -c' d, \qquad \text{[S55]}$$

which is equivalent to the optimality condition at the boundary Eq. **S17**, provided that $\lambda^* = c' d$.

Appendix K: Cost Function as a Lyapunov Function of the Mean-Field Dynamics

Here we show rigorously that, when the availability function is scale invariant, as in the case for the simple cost function $F(m) = m^\alpha$, the dynamics must converge toward a fixed point. This fixed point is unique and corresponds to the optimal of the cost $\langle F \rangle$, as we have shown in the previous section.

$A(x)$ is scale invariant if a function v exists such that $A(\gamma x) = v(\gamma) A(x)$. In this case we will see that the changes of relative frequencies P_r in the repertoire over time depend only on the total number of receptors through a prefactor. Below we derive the equations governing these dynamics and then prove that these dynamics are ensured to converge to a stable fixed point. We do so by showing that the dynamics admit the expected cost $\langle F \rangle$ as a Lyapunov function, i.e., a function that continually decreases under the dynamics.

For ease of notation we rewrite Eq. **S49** as

$$\frac{\mathrm{d}N_r}{\mathrm{d}t} = N_r [\pi_r(N) - d], \qquad \text{[S56]}$$

where N is shorthand for $\{N_r\}$, and $\pi_r = \sum_a Q_a A(\sum_{r'} N_{r'} f_{r',a}) f_{r,a}$ is the growth rate of receptor type r. The relative frequencies $P_r = N_r/N_{\mathrm{tot}}$ evolve according to

$$\frac{\mathrm{d}P_r}{\mathrm{d}t} = \frac{1}{N_{\mathrm{tot}}} \frac{\mathrm{d}N_r}{\mathrm{d}t} - \frac{N_r}{N_{\mathrm{tot}}^2} \frac{\mathrm{d}N_{\mathrm{tot}}}{\mathrm{d}t} \qquad \text{[S57]}$$

$$= P_r \left[\pi_r(N) - \sum_{r'} P_{r'} \pi_{r'}(N) \right]. \qquad \text{[S58]}$$

If A is scale invariant, so is π_r and $\pi_r(N) = \pi_r(N_{\mathrm{tot}} P) = v(N_{\mathrm{tot}}) \pi_r(P)$. Then the equations further simplify to

$$\frac{\mathrm{d}P_r}{\mathrm{d}t} = v(N_{\mathrm{tot}}) P_r \left[\pi_r(P) - \sum_{r'} P_{r'} \pi_{r'}(P) \right], \qquad \text{[S59]}$$

$$= v(N_{\mathrm{tot}}) P_r (\pi_r - \overline{\pi}), \qquad \text{[S60]}$$

where $\overline{\pi} = \sum_r P_r \pi_r$.

We can now write how the expected cost $\langle F \rangle$ evolves in time:

$$\frac{\mathrm{d}\langle F \rangle}{\mathrm{d}t} = \sum_r \frac{\partial \langle F \rangle}{\partial P_r} \frac{\mathrm{d}P_r}{\mathrm{d}t} \qquad \text{[S61]}$$

$$= v(N_{\mathrm{tot}}) \sum_r P_r \left[\sum_a Q_a \overline{F}'_a(\tilde{P}_a) f_{r,a} \right] (\pi_r - \overline{\pi}) \qquad \text{[S62]}$$

$$= -\frac{v(N_{\mathrm{tot}})}{c'} \sum_r P_r \left[\sum_a Q_a A\left(N_{\mathrm{st}} \tilde{P}_a\right) f_{r,a} \right] (\pi_r - \overline{\pi}) \qquad \textbf{[S63]}$$

$$= -\frac{v(N_{\mathrm{tot}}) v(N_{\mathrm{st}})}{c'} \sum_r P_r \pi_r (\pi_r - \overline{\pi}) \qquad \textbf{[S64]}$$

$$= -\frac{v(N_{\mathrm{tot}}) v(N_{\mathrm{st}})}{c'} \sum_r P_r (\pi_r - \overline{\pi})^2 \leq 0. \qquad \textbf{[S65]}$$

This proves that the cost always decreases with time, i.e., is a Lyapunov function of the dynamics. Therefore, the dynamics will reach a stable fixed point at steady state, which is guaranteed to be the global minimum of the expected cost $\langle F \rangle$.

1. Boyd SP, Vandenberghe L (2004) *Convex Optimization* (Cambridge Univ Press, Cambridge, UK).
2. Bronshtein IN, Semendyayev KA, Musiol G, Muehlig H (2007) *Handbook of Mathematics* (Springer, Berlin Heidelberg), Vol 3.
3. Parikh N, Boyd S (2013) Proximal algorithms. *Found Trends Optim* 1(3):123–231.
4. Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J Imaging Sci* 2(1):183–202.
5. Duchi J, Shalev-Shwartz S, Singer Y, Chandra T (2008) Efficient projections onto the l_1 ball for learning in high dimensions. *Proceedings of the International Conference on Machine Learning* (ICML, Helsinki), pp 272–279.
6. Bertsekas DP (1999) *Nonlinear Programming* (Athena Scientific, Belmont, MA).
7. Chaikin PM, Lubensky TC (2000) *Principles of Condensed Matter Physics* (Cambridge Univ Press, Cambridge, UK), Vol 1.
8. Torquato S, Stillinger FH (2003) Local density fluctuations, hyperuniformity, and order metrics. *Phys Rev E Stat Nonlin Soft Matter Phys* 68(4 Pt 1):041113.
9. Donev A, Stillinger FH, Torquato S (2005) Unexpected density fluctuations in jammed disordered sphere packings. *Phys Rev Lett* 95(9):090604.
10. Berthier L, Chaudhuri P, Coulais C, Dauchot O, Sollich P (2011) Suppressed compressibility at large scale in jammed packings of size-disperse spheres. *Phys Rev Lett* 106(12):120601.
11. Pigolotti S, López C, Hernández-García E (2007) Species clustering in competitive Lotka-Volterra models. *Phys Rev Lett* 98(25):258101.

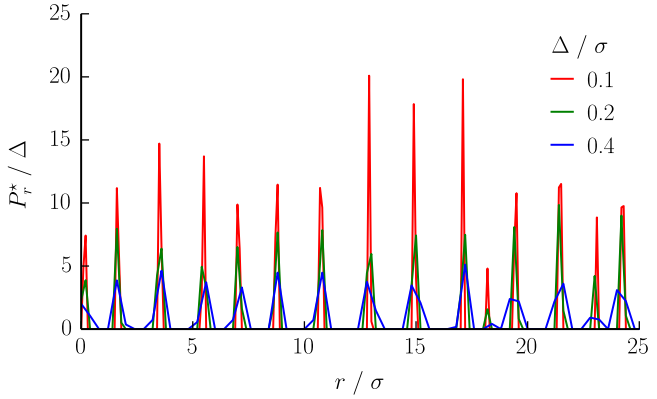

Fig. S1. Solving the optimization problem with a finer and finer discretization step suggests that the peaks found in the optimal receptor distributions converge to true Dirac delta functions. Starting from a problem with a discretization step of $\Delta = 0.1\sigma$, we construct coarse-grained versions of it by downsampling the antigen distribution two- and fourfold, yielding $\Delta = 0.2\sigma$ and 0.4σ, respectively. The resulting coarse-grained optimization problems are then solved, and the optimal distributions P_r^*/Δ are represented (after appropriate normalization by the step size). The random antigen distribution is log-normal with coefficient of variation $\kappa = 0.25$.

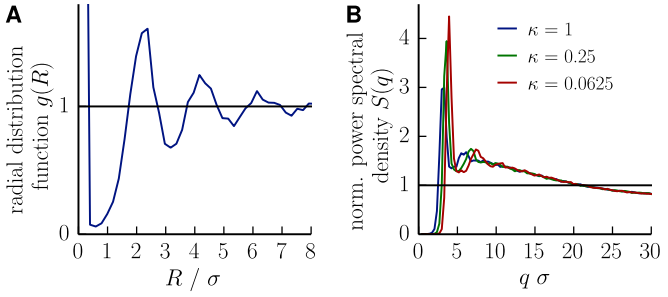

Fig. S2. Radial distribution function and normalized power spectral density of the optimal receptor distribution P_r^* for random environments in two dimensions. (*A*) The radial distribution function of P_r^* shows an exclusion zone around each peak, followed by oscillations characteristic of a local tiling pattern. (*B*) Normalized power spectral density $S(q)$ of P_r^* for different values of the parameter κ quantifying the heterogeneity of the antigenic landscape. The high suppression of fluctuations at large scales (small q) indicates that the pattern has very little fluctuation in the number of receptors used to cover large surface areas.

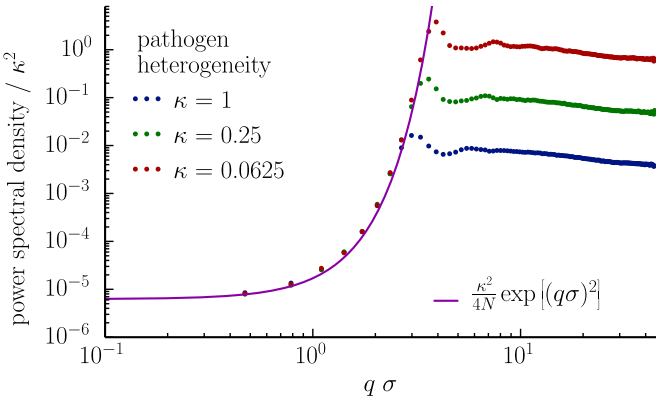

Fig. S3. Power spectral density normalized by the squared antigenic environment heterogeneity index κ: $|\sum_r P_r e^{iqr}|^2/\kappa^2$. The data collapse for different κ shows that the fluctuations at large scale are entirely attributable to those of the antigenic environment and scale with them. At these large scales, the power spectrum of the receptor distribution is approximately given by $\exp[(q\sigma)^2]\kappa^2/4N$. The exponential term stems from the inverse of the Fourier transform of f (Eq. **S28**). In other words, the coverage of the antigenic space exactly tracks the distribution of antigens, with no additional fluctuations due to the random positioning of peaks (which would be present if this positioning was Poisson distributed). This property is called disordered hyperuniformity in the physics of jammed materials (8–10). Parameters are the same as in Fig. 3.

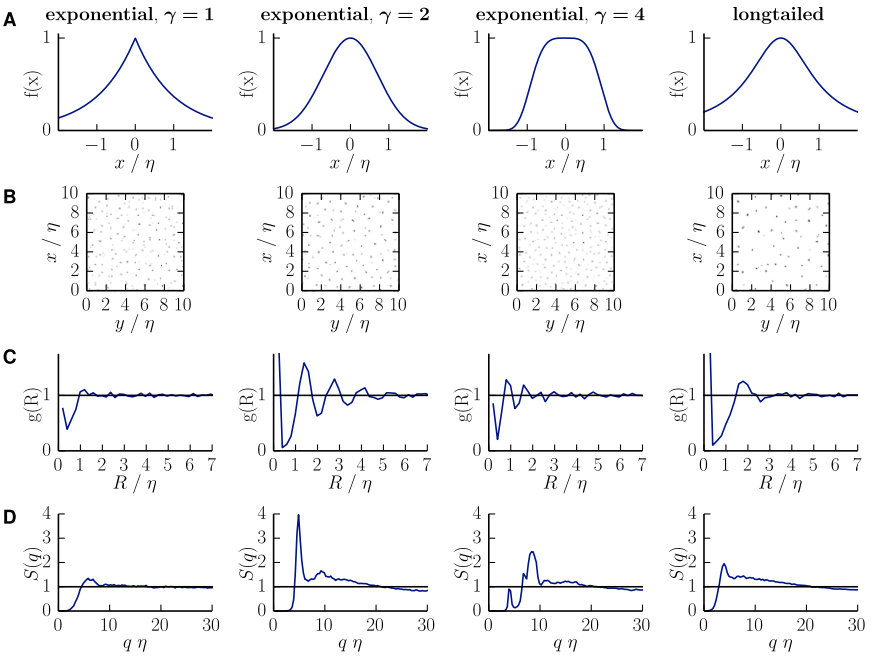

Fig. S4. Influence of the choice of the cross-reactivity kernel $f(a-r)$ on the optimization problem. Regardless of the kernel choice the optimal repertoire is peaked for nonuniform antigen distributions. The details of distribution depend on the cross-reactivity kernel. (*A*) Kernel functions used to describe cross-reactivity. *A*, *Left*, *Center Left*, and *Center Right* show exponential kernels of the form $f(r-a)=\exp[-(|r-a|/\eta)^\gamma]$ with different values of the parameter γ. *A*, *Right* shows a long-tailed kernel of the form $f(r-a)=1/(1+(|r-a|/\eta)^2)$. (*B*) Examples of optimal receptor distributions in two dimensions, for antigenic environments generated as in Fig. 3*B* (with coefficient of variation $\kappa=0.25$). (*C*) Radial distribution function of the optimal distribution. (*D*) Structure factor of the optimal distribution. The results in *C* and *D* are averaged over 10 independent runs. A linear effective cost function $F(m)=m$ is assumed throughout.

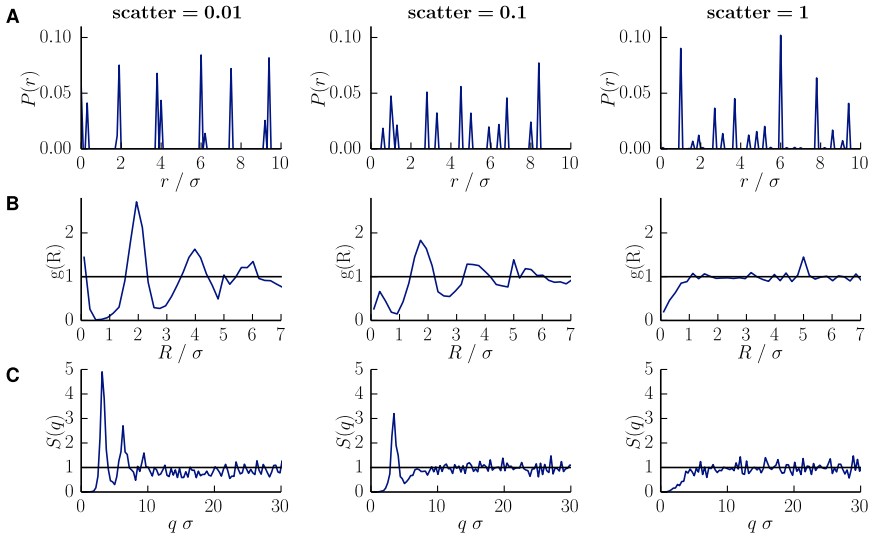

Fig. S5. Influence of varying amounts of nonuniformity in the widths of the cross-reactivity kernel on the optimization problem. (*A*) Examples of optimal receptor distributions in one dimension, for antigenic environments generated as in Fig. 3*B* (with coefficient of variation $\kappa = 0.25$). A linear effective cost function $F(m) = m$ was assumed. (*B*) Radial distribution function of the optimal distribution. (*C*) Structure factor of the optimal distribution. The results in *B* and *C* are averaged over 30 independent runs.

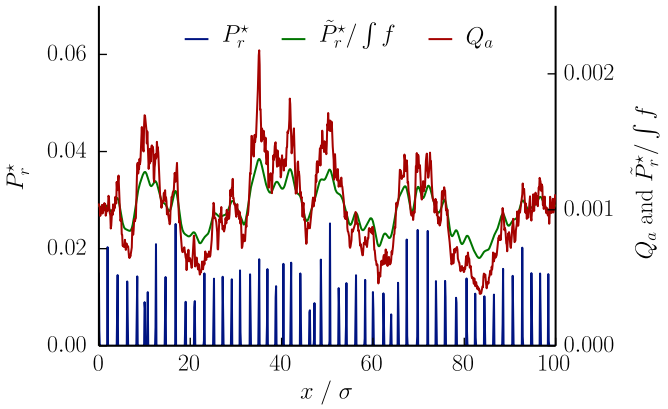

Fig. S6. Adding correlations to the antigen distribution does not change the peakedness of optimal receptor distributions. The result of the optimization is shown for a random antigen landscape with correlations. The antigen distribution is generated by Fourier filtering. First we generate an uncorrelated, normally distributed random series. This series is then filtered to obtain a power spectrum $\propto 1/(1+(10q\sigma)^2)$. Finally, the filtered series is exponentiated to ensure the nonnegativity of the generated values.

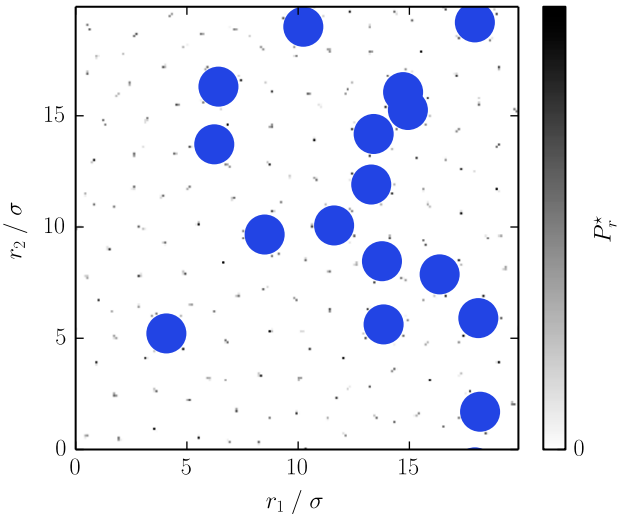

Fig. S7. Effect of exclusion zones around self-antigens on the optimization problem. No receptors are allowed in exclusion regions around self-antigens (shaded in blue).

Table S1. Intermediate results in the derivation of the optimal solution

$F(m)$	$\bar{F}(\tilde{P}_a)$	$h(x)$
m^α	$\Gamma(1+\alpha)/\tilde{P}_a^\alpha$	$(-x/(\alpha\Gamma(1+\alpha)))^{\frac{1}{1+\alpha}}$
$\ln m$	$\gamma - \ln\tilde{P}_a$	$-1/x$
$1-\exp(-\beta m)$	$\beta/(\beta+\tilde{P}_a)$	$\sqrt{-\beta/x}-\beta$
$\Theta(m-m_0)$	$\exp(-m_0\tilde{P}_a)$	$-\ln(-x/m_0)/m_0$

The first column shows several choices of the effective cost function, $F(m)$. For these cost functions the second column shows the average cost of a pathogenic attack, $\bar{F}(\tilde{P}_a)$, and the third column shows the inverse of its derivative, $h=(\bar{F}')^{-1}$. Γ is the Gamma function, γ is Euler's constant, and β and m_0 are positive constants.