Design principles of the cytotoxic CD8$^+$ T-cell response

Obinna A. Ukogu,[1] Zachary Montague,[2] Grégoire Altan-Bonnet,[3] and Armita Nourmohammad[1,2,4,5,*]

[1] *Department of Applied Mathematics, University of Washington, 4182 W Stevens Way NE, Seattle, WA 98105, USA*
[2] *Department of Physics, University of Washington, 3910 15th Avenue Northeast, Seattle, WA 98195, USA*
[3] *Immunodynamics Group, Laboratory of Integrative Cancer Immunology, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA*
[4] *Paul G. Allen School of Computer Science and Engineering, University of Washington, 85 E Stevens Way NE, Seattle, WA 98195, USA*
[5] *Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, Seattle, WA 98109, USA*

Cytotoxic T lymphocytes eliminate infected or malignant cells, safeguarding surrounding tissues. Although experimental and systems-immunology studies have cataloged many molecular and cellular actors involved in an immune response, the design principles governing how the speed and magnitude of T-cell responses emerge from cellular decision-making remain elusive. Here, we recast the T-cell response as a feedback-controlled program, wherein the rates of activation, proliferation, differentiation and death are regulated through antigenic, pro- and anti-inflammatory cues. By exploring a broad class of feedback-controller designs as potential immune programs, we demonstrate how the speed and magnitude of T-cell responses emerge from optimizing signal-feedback to protect against diverse infection settings. We recover an inherent trade-off: infection clearance at the cost of immunopathology. We show how this trade-off is encoded into the logic of T-cell responses by hierarchical sensitivity to different immune signals. Notably, we find that designs that balance harm from acute infections and autoimmunity produce immune responses consistent with experimentally observed patterns of T-cell effector expansion in mice. Extending our model to immune-based T-cell therapies for cancer tumors, we identify a trade-off between the affinity for tumor antigens ("quality") and the abundance ("quantity") of infused T-cells necessary for effective treatment. Finally, we show how therapeutic efficacy can be improved by targeted genetic perturbations to T-cells. Our findings offer a unified control-logic for cytotoxic T-cell responses and point to specific regulatory programs that can be engineered for more robust T-cell therapies.

I. INTRODUCTION

Cellular immunity hinges upon the regulated activation, expansion, and contraction of lymphocytes that clear an infection while limiting collateral tissue damage. We focus on CD8$^+$ T-cells—also known as killer or cytotoxic T-cells—which are crucial for immune surveillance and defense against infections and cancer. Pathogen recognition is mediated by T-cell receptors (TCRs) that bind short pathogen-derived peptides (or antigens) presented by major histocompatibility complexes (MHCs) on antigen-presenting cells (APCs) and other target cells. To recognize a multitude of pathogens, the immune system generates a diverse repertoire of TCRs through a stochastic process known as V(D)J recombination, producing hundreds of billions of distinct T-cell clonotypes (i.e., T-cells with unique receptors) [1–3].

Pathogen recognition triggers a response program in naïve T-cells, tuned by antigen signal, pro- and anti-inflammatory signaling molecules (cytokines), and other environmental signals generated by the infection and the host's response (Fig. 1A) [4–8]. These signals determine the extent of burst-like proliferation of short-lived effector T-cells versus differentiation into long-lived, slowly dividing memory subsets. Effector T-cells have the capacity to rapidly divide—up to 19 cell divisions [7, 9–11]—and secrete cytolytic chemical molecules that can degrade infected or unwanted cells [12, 13]. While most activated naïve cells differentiate into terminal effectors that undergo apoptosis once infection is cleared, a small fraction survive as long-lived effector or central memory cells [12, 14–16].

The logic of signal integration for immune decision-making throughout a response remains the central open problem. Simple additive models of antigen and cytokine inputs can reproduce experimental patterns of clonal expansion in some settings [6–8]. However, genetic and single-cell analyses reveal far richer control circuits governing activation, differentiation, and contraction of T-cells that involve intricate genetic programs and probabilistic regulatory rules [11, 15–21].

Pathogens vary in their immunogenicity—the ease with which they are recognized as "non-self"—and in the degree of harm they impose via secreted toxins, inflammation, and disruption of cellular function. The CD8$^+$ T-cell response must have the potential to curb "harm" across diverse infection scenarios [22, 23], yet the evolutionary objectives and biophysical constraints shaping their decision-making program remain unclear. Here, we formalize one plausible objective for T-cell decision-making: to minimize cumulative host healthy-cell loss due to infection and collateral death of healthy cells due

* Correspondence should be addressed to: Armita Nourmohammad: armita@uw.edu.

to the immune response (immunopathology or toxicity).

We model the T-cell decision-making program as a flexible, coarse-grained signal-integration protocol. Specifically, we cast each CD8$^+$ T-cell as a feedback controller whose intra-cellular regulatory program modulates activation, differentiation, proliferation, and death in response to intra-host cues, such as antigen levels and cytokine signals, which themselves are dynamic in the course of an infection [7, 24].

We then ask how alternative immunological objectives give rise to distinct effector programs. First, we parameterize a broad class of feedback-controller designs, with varied signal sensitivities and integration rules. Next, by systematically sampling designs, we optimize for rapid pathogen clearance while minimizing immunopathology. We find that optimal designs lie on a Pareto front: further reductions in infection-induced harm entail increased immunopathology. Accounting for the imperative to separate infection—microbial or malignant—from transient inflammatory events (e.g., tissue damage) [25–29], we show how a trade-off between pathogen-clearance and immunopathology (toxicity) may be encoded into the logic of T-cell responses. Optimizing these competing objectives yields effector programs consistent with experiments with vertebrate organisms. Finally, applying our framework to a minimal tumor-growth model, we demonstrate how modifying the endogenously encoded effector program can improve the effectiveness of T-cell cancer therapies, pointing to signal-processing adjustments—and thus candidate genetic engineering—that could enhance T-cell-based cancer immunotherapy.

The immune system's true performance objective, and how evolution has fine-tuned it, may ultimately be unknowable. Formulating and testing principled heuristics provides a baseline to gauge the performance of real immune systems in responding to infections. This perspective has yielded deep insights in other biological contexts, revealing the architectural logic of genetic and regulatory networks [30–35]. In immunology, high-level mathematical models have demonstrated how a finite lymphocyte repertoire can be distributed to best confront a diverse pathogenic landscape [36–38] and how immune memory should be allocated to counter future threats [39–43]. By extending this perspective to the dynamics of CD8$^+$ T-cell signal integration and response, our work provides a quantitative framework that links observed decision-making to operating principles and trade-offs.

MODEL

To characterize the immune response to an infection, we model the large-population dynamics of susceptible and infected cells in a host as deterministic processes, coupled to a stochastic, time-inhomogeneous birth-death process for the CD8$^+$ T-cell response (Fig. 1A, and SI Appendix for details).

a. Intra-host infection dynamics. We model a population of healthy, susceptible cells S, initially consisting of $S_{\max}$ identical cells. Pathogen exposure seeds an infected subpopulation $I(t)$, which grows via encounters with susceptible cells at rate b_I, and declines by infected-cell death at rate d_I, triggering an immune response (Figs. 1A, S1). The dynamics of the susceptible and infected cells follow,

$$\begin{aligned} \dot{S} &= -b_I I S - d_E \frac{E}{K_S + I\frac{K_S+E}{K_I+E} + E + S} S \\ \dot{I} &= b_I I S - d_I I - d_E \frac{E}{K_I + S\frac{K_I+E}{K_S+E} + E + I} I, \end{aligned} \tag{1}$$

where d_E is the maximal clearance rate of infected/susceptible cells by effector T-cells E, and K_I and K_S are *recognition thresholds*, quantifying how readily effector cells can recognize infected and susceptible cells, respectively, and consequently kill them; a larger K_I or K_S implies more effector cells are needed for recognition [44–46]. The functional form of the last term in these equations can be motivated by considering competitive binding of T-cells to infected and healthy cells (SI Appendix).

We quantify within-host reproductive success of a pathogen by the *intra-host basic reproduction number*,

$$R_0 = \frac{b_I S_{\max}}{d_I}, \tag{2}$$

measuring the expected number of secondary infected cells generated by a single infected cell in an otherwise fully susceptible cell population. This definition mirrors the standard epidemiological reproduction number [47, 48]. We model the T-cell repertoire as initially consisting of a number of identical naïve T-cells, and assume $K_S = S_{\max} \geq K_I$ so that infected cells are typically easier to recognize than susceptible cells. These recognition thresholds determine the *immunogenicity* of an infection, which we define as,

$$\text{immunogenicity:} \quad \kappa = K_I^{-1} K_S. \tag{3}$$

This dimensionless quantity reflects how readily cognate T-cells recognize antigens presented by infected cells as "non-self" relative to those presented by healthy cells (self). We consider infection scenarios with broad ranges of immunogenicity (κ) and basic reproduction numbers (R_0) (SI Appendix).

b. Signal cues to modulate T-cell decision-making. We express damage during an infection in terms of cell deaths,

$$h(t) = h_{\text{inf}}(t) + \underbrace{h_{E,I}(t) + h_{E,S}(t)}_{=h_{\text{res}}(t)}, \tag{4}$$

where $h_{\text{inf}}(t)$ represents the number of cells killed directly by the pathogen (infection), and $h_{E,I}(t)$ and $h_{E,S}(t)$ denote the number of infected and healthy cells, respectively, killed by the effector T-cell response, amounting to the total damage by the response $h_{\text{res}}(t)$.

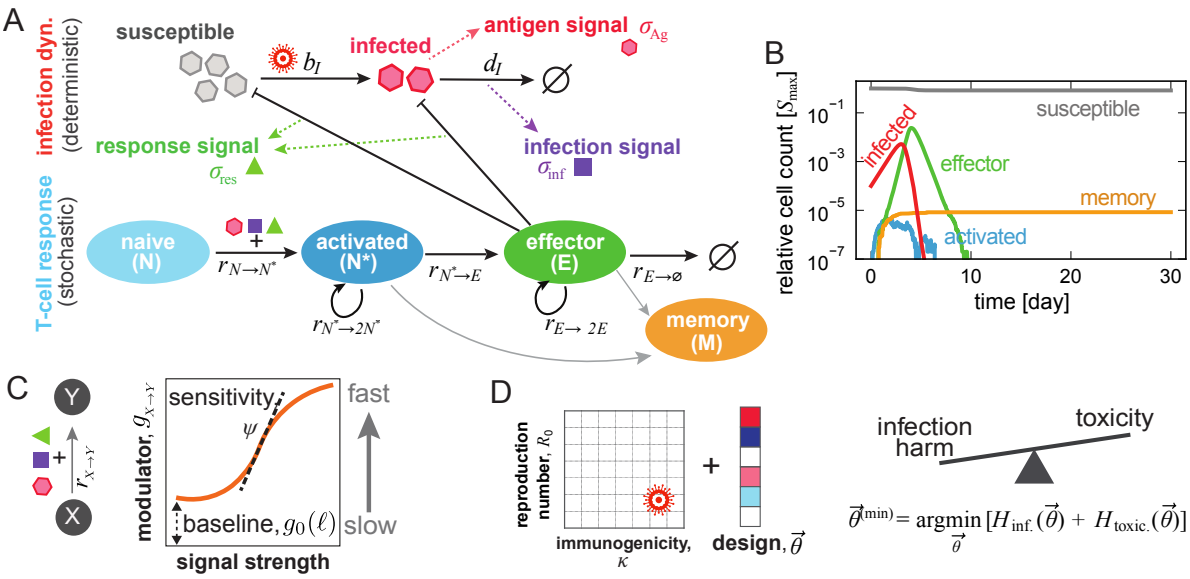

Figure 1. **CD8$^+$ T-cell response to infections. (A)** The schematic shows how upon infection, susceptible cells S become infected I at rate b_I. Infected cells provide an antigen signal $\sigma_{\text{Ag}} \propto I$, and die at rate d_I, which triggers an infection harm signal σ_{inf}. In response to infection, naïve CD8$^+$ T-cells (with initial clone size of $N_{\text{lin.}}$) are activated $N \to N^*$, and can then differentiate into effector or memory fates (Fig. S1). Effector cells can proliferate and kill infected and bystander cells, generating a response-harm signal σ_{res}, and finally become terminal or differentiate into memory. Consequently, T-cell transitions are modulated by the threes signals $(\sigma_{\text{Ag}}, \sigma_{\text{inf}}, \sigma_{\text{res}})$. Infection dynamics (S, I) are modeled by ODEs coupled to a stochastic dynamics of the T-cell response. Memory transitions (gray arrows) are neglected in our model. **(B)** Representative trajectories of susceptible, infected, effector, and memory CD8$^+$ T-cells are shown over time. **(C)** Each T-cell transition has a rate $r_i = r^{\max} g_i(\sigma_{\text{Ag}}, \sigma_{\text{inf}}, \sigma_{\text{res}}; \vec{\psi}, \ell)$, modulated by a sigmoidal function with baseline ℓ and signal sensitivities $\vec{\psi}$ (eqs. 6, 7). The regulatory parameters define the design of the effector program $\vec{\theta}$ (eq. 8). **(D)** We assess design performance over diverse infections sampled from a broad range of reproduction numbers R_0 (eq. 2) and immunogenicities κ (eq. 3). For each infection, we can find an optimal design $\vec{\theta}^{(\min)}$ that minimizes total cell loss (harm). Harm decomposes into infection harm $H_{\text{inf.}}$ (cells killed by the pathogen plus infected cells eliminated by effectors), and toxicity $H_{\text{toxic.}}$ (healthy bystanders killed by effectors), highlighting an intrinsic trade-off (eq. 9). Simulation parameters in (B): $S_{\max} = 10^7, N_{\text{lin.}} = 100, K_I = 10^{-3} S_{\max}, K_S = S_{\max}, K_H = 10^{-3} S_{\max}, b_I = 2.0 \times 10^{-7}\text{day}^{-1}$ per cell, $d_I = 0.5\ \text{day}^{-1}, I_0 = 10^{-4} S_{\max}, \vec{\theta} = (\psi_{Ag} = 2, \psi_I = 2, \psi_E = -0.25, \ell_{N\to N^*} = -0.75, \ell_{N^*\to E} = -0.75, \ell_{E\to\emptyset} = 0.75)$.

Cell deaths can produce different types of signals. Unprogrammed cell death (e.g., pyroptosis triggered by viruses and bacteria) can lead to the the release of pro-inflammatory danger-associated molecular patterns (DAMPs) [25]. On the other hand, cell death initiated by CD8$^+$ T-cells generates a heterogeneous signal that generates pro- and anti-inflammatory cytokines (e.g., IFNγ, TNFα, IL-1ra, IL-4, IL-10, TGF-β) [25, 26, 29].

We simplify this picture by partitioning non-antigen cues into (i) infection-induced and (ii) response-induced harms, whose accumulated effects trigger infection and response harm-induced signals, $\sigma_{\text{inf}}(t)$ and $\sigma_{\text{res}}(t)$, respectively. Antigens presented by cells also serve as a signal; we take the infected-cell burden $I(t)$ as a proxy for the instantaneous antigen signal $\sigma_{\text{Ag}}(t)$. Together, expressed by their recognition thresholds, K_I for antigens and K_H for non-antigens (harm), the three signals available to T-cells follow,

$$\text{antigen signal:}\quad \sigma_{\text{Ag}}(t) = {K_I}^{-1} I(t)$$
$$\text{harm signals:}\quad \sigma_{\text{inf/res}}(t) = K_H^{-1} \int_0^t e^{-(t-s)d_H}\, h_{\text{inf/res}}(s)\, \mathrm{d}s. \tag{5}$$

The time-integrated harm-induced signals are exponentially discounted at rate d_H, reflecting the decay of cytokines in the absence of ongoing stimulation (SI Appendix).

By disentangling these sources of (potentially redundant) signals, we probe how CD8$^+$ T-cells use these cues to make fate decisions. We do not pre-assign any of the signals to a pro- or anti-inflammatory role but allow optimization goals to determine their roles across diverse infection scenarios.

c. Design of T-cell programs for infection response. Naïve T-cells cycle through lymphoid tissues, becoming activated when they durably engage cognate peptide-MHC complexes on activated APCs [49, 50]. Activated

T-cells can differentiate into an effector phenotype, which participates in infection clearance, or into a long-lived memory state (Figs. 1A, S1). Here, we examine the course of $CD8^+$ T-cells response during an infection, dividing the process into four canonical response modules (Fig. 1A):

1. **Activation:** priming of naïve T-cells, $N \to N^*$,
2. **Differentiation:** commitment of the activated T-cell to an effector phenotype, $N^* \to E$,
3. **Expansion:** division of effector cells, $E \to 2E$,
4. **Contraction:** programmed loss of effector cells, $E \to \emptyset$.

These response processes map onto partially distinct genetic programs: the MYC pathway drives activation and proliferation [17]; the Wnt-Tcf7 pathway regulates effector versus memory fate decisions [16, 51]; and the Bcl-2-family of proteins control contraction of the effector cells via apoptosis [17].

For each process i, we represent its genetic regulation by a time-dependent transition rate $r_i(t)$, modulated by the instantaneous signals $\vec{\sigma}(t)$ (eq. 5) received by the T-cell (Fig. 1C),

$$r_i(t) = r^{\max}\, g_i\left(\sigma_{\text{Ag}}, \sigma_{\text{inf}}, \sigma_{\text{res}}; \vec{\psi}, \ell_i\right), \tag{6}$$

where the maximal rate $r^{\max}$ is capped by the cell-cycle time [13, 21, 49], and g_i is a monotonic *response function* of signal input, chosen from the modified Monod-Wyman-Changeux functions (Hill-type family) [52, 53]:

$$g_i(\sigma_{\text{Ag}}, \sigma_{\text{inf}}, \sigma_{\text{res}}; \vec{\psi}_i, \ell_i) = \frac{1}{1 + \exp\left[-L_i(\vec{\sigma}; \vec{\psi}_i, \ell_i)\right]},$$

with, (7)

$$L_i(\vec{\sigma}; \vec{\psi}_i, \ell_i) = \ell_i + \sum_{\alpha \in \{\text{Ag,inf,res}\}} \psi_i^{\alpha} \log\left[1 + \sigma_\alpha\right].$$

The parameter ℓ_i sets the baseline propensity of transition i: in the absence of external signals, the regulatory factor evaluates to $g_{0,i} = (1 + \exp(-\ell_i))^{-1}$ (Fig. 1C). The weights $\{\psi_i^{Ag}, \psi_i^{\text{inf}}, \psi_i^{\text{res}}\}$ encode the sensitivity and polarity (positive / negative = up /down-regulation) of response to the antigen (σ_{Ag}) and harm signals (σ_{inf} and σ_{res}). This formulation abstracts away biochemical details while implementing the fold-change response to stimuli observed in many biological circuits [33, 54–57].

To reflect shared regulatory programs while controlling complexity, we impose two simplifications. First, activation and early expansion are both MYC-driven [58]; we therefore require the modulatory factors, $g_{N\to N^*}$ and $g_{E\to 2E}$, to share parameter values, even though their realized rates will differ because signals vary in time, as cells transition between naive and effector states. Second, logically, since extracellular cues should converge on coherent effector programs, we enforce a common set of signal sensitivities across modules. With the sign convention that positive sensitivities promote effector production, for each signal σ, we set $\psi_\sigma \equiv \psi^{\sigma}_{N\to N^*} = \psi^{\sigma}_{E\to 2E} = \psi^{\sigma}_{N^*\to E} = -\psi^{\sigma}_{E\to\emptyset}$.

The parameters of these constrained response functions define a *response design* $\vec{\theta}$ as,

$$\vec{\theta} = (\underbrace{\psi_{\text{Ag}}, \psi_{\text{inf}}, \psi_{\text{res}}}_{\text{signal feedback}}, \underbrace{\ell_{N\to N^*} = \ell_{E\to 2E}, \ell_{N^*\to E}, \ell_{E\to\emptyset}}_{\text{module-specific baselines}}). \tag{8}$$

We systematically sample $\vec{\theta}$'s to assess performance of designs against different infections (Fig. 1D; SI Appendix). While $\vec{\theta}$ is broadly sampled, we hold the remaining biological rates and design parameters underlying the decision-making program in Fig. 1A within empirically supported ranges reported in SI Appendix, Table S1.

d. Cell economics of T-cell response. Tracking all clinical sequelae of an infection is rarely feasible, but tissue integrity degrades as cells are lost. Cumulative cell death correlates with macroscopic readouts such as weight change, pathogen burden, and temperature dynamics [59, 60]. We therefore adopt a cell-economics objective in which harm is measured by the total number of cells lost over the course of infection.

We evaluate the performance of a given design $\vec{\theta}$ by simulating the response dynamics (Fig. 1D) and computing the *total cell death* (or inflicted harm) as

$$H(\vec{\theta}) = \underbrace{\int_0^T \left[h_{\text{inf}}\left(t, \vec{\theta}\right) + h_{E,I}\left(t, \vec{\theta}\right)\right] \mathrm{d}t}_{H_{\text{inf.}}(\vec{\theta})} + \underbrace{\int_0^T h_{E,S}\left(t, \vec{\theta}\right) \mathrm{d}t}_{H_{\text{toxic.}}(\vec{\theta})}, \tag{9}$$

where $H_{\text{inf.}}(\vec{\theta})$ aggregates infection-related harm (cells killed by the pathogen, and infected cells eliminated by effectors), $H_{\text{toxic.}}(\vec{\theta})$ captures immunopathology (healthy bystander cells killed by effectors), over the time window T (SI Appendix, Table S1). This single, interpretable metric balances clearance benefits against collateral damage, enabling principled comparison across designs and infections (Fig. 1D).

RESULTS

Cell-fate transitions determine effector response and pathogen clearance

How do rates of cell-fate transitions, determined by the regulatory design $\vec{\theta}$ (eq. 8), impact the immune response and pathogen clearance? To isolate the effect of baseline propensities ($\vec{\ell}$) independently of signal feedback ($\vec{\psi}$), we first simulated "hard-wired" designs in which all signal sensitivities were set to zero ($\vec{\psi}_i = 0$). Each transition then happens at a fixed fraction of its maximal rate

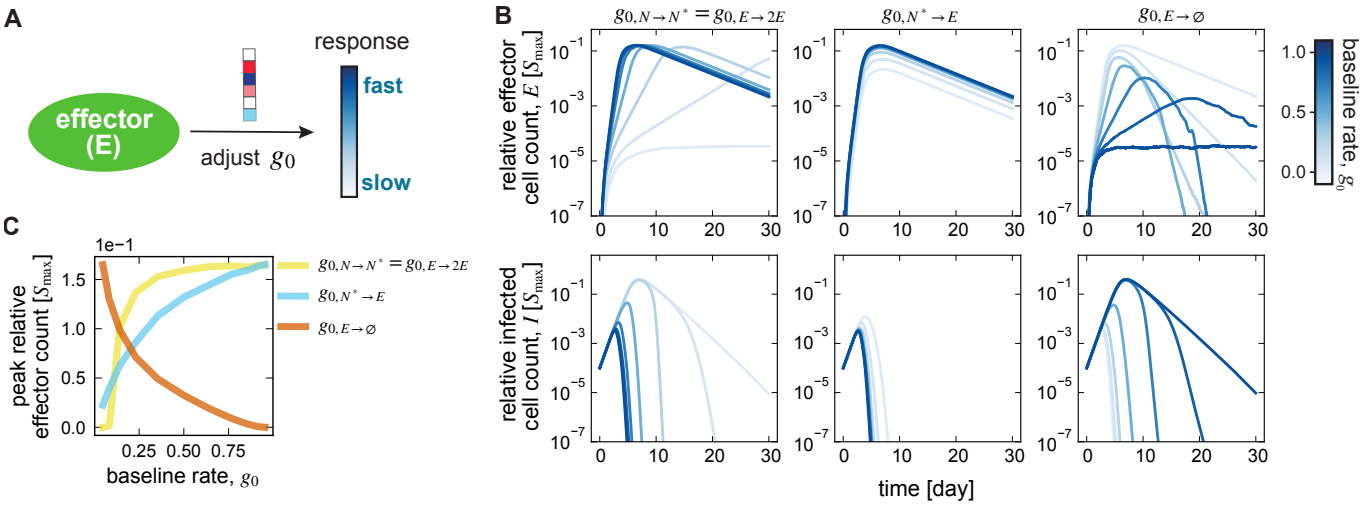

Figure 2. **Design parameters shape T-cell response and infection clearance. (A)** Varying the design parameters influences effector trajectories. **(B)** Trajectories are shown for the population size of effectors (top) and infected cells (bottom), measured in units of maximum number of susceptible cells $S_{\max}$, for designs that vary a single baseline transition rate $g_{0,i}$ (color) indicated atop each panel. All other design parameters are fixed at: fast activation, $g_{0,N\to N^*} = 0.95$; fast differentiation, $g_{0,N^*\to E} = 0.95$; slow contraction, $g_{0,E\to\emptyset} = 0.05$; and zero signal feedback, $\psi_{\text{Ag}} = \psi_{\text{inf}} = \psi_{\text{res}} = 0$. Trajectories reflecting the impact of varying signal sensitivities ψ's are shown in SI Appendix, Fig. S2. **(C)** Peak effector count, measured in units of maximum number of susceptible cells $S_{\max}$, is shown by varying each of the three baseline transition rates (colors), while keeping the rest of the parameters fixed (similar to (B)). Parameters are the same as Fig. 1, except for $K_I = 10^{-2} S_{\max}$.

$g_{0,i} = [1+\exp(-\ell_i)]^{-1}$, independently of antigen or harm signals; here, $i \in \{N \to N^*/E \to 2E, N^* \to E, E \to \emptyset\}$ indicates the transition module.

Varying these baseline rates produced strikingly different effector trajectories (Fig. 2). Promoting activation/proliferation ($g_{0,N\to N^*}, g_{0,E\to 2E}$ $\uparrow$), or lowering the contraction baseline rate ($g_{0,E\to\emptyset}$ $\downarrow$) generated a rapid, high-amplitude effector burst that cleared the pathogen quickly, after which effector numbers decay (Fig. 2B, C). In contrast, slowing activation and division ($g_{0,N\to N^*}, g_{0,E\to 2E}$ $\downarrow$), or accelerating contraction ($g_{0,E\to\emptyset}$ $\uparrow$), yielded sluggish effector accumulation, delayed clearance, and in extreme cases pathogen persistence, accompanied by a prolonged plateau of few effectors. Interestingly, varying the baseline rate of differentiation $g_{0,N^*\to E}$ had a weaker impact on the overall effector profile and pathogen clearance (Fig. 2B, C), suggesting relevance more downstream in the response program.

Because effector and pathogen abundances determine the levels of antigen and cumulative harm-associated cytokine signals, these baseline choices not only shape the immediate effector response but also alter the available antigen and cytokine signals, which in turn impact the T-cell response through feedback regulation (SI Appendix, Fig. S2).

Effector designs trade infection harm for toxicity

Previous research has identified a trade-off between maximizing infection clearance and minimizing response toxicity (immunopathology) in the immune response [28, 29, 61]. Consistently, by systematically sampling the space of designs (SI Appendix), we found this trade-off in the form of a Pareto front: plotting infection harm $H_{\text{inf.}}(\theta)$ (i.e., infected cells killed by pathogen or by effector response) versus the toxicity $H_{\text{toxic.}}(\theta)$ (i.e., bystander cells killed by effector response) revealed that decreasing $H_{\text{inf.}}$ comes at the cost of increasing $H_{\text{toxic.}}$ (Fig. 3A). The extent of this trade-off depends on pathogen characteristics. This trade-off is stronger for infections with a higher intra-host reproductive number R_0 or lower immunogenicity κ. Interestingly, for scenarios with lower R_0, many of the hard-wired response programs without signal feedback ($\vec{\psi}_i = 0$) lie along the Pareto front, and perform near-optimally. For aggressive high-R_0 threats—be it acute infections (high κ), or highly metastatic cancers (mid to low κ)—these same hard-wired designs fall far inside the sub-optimal region. Proportionate control in these cases requires designs with appropriate sensitivities to antigen and cytokine/harm signals ($\vec{\psi} \neq 0$) to minimize infection harm while limiting immunopathology.

Next, we explored regulatory design space by evaluating responses of different designs against diverse infection scenarios, sampled on a grid of immunogenicities (κ) and reproduction numbers (R_0) (Fig. 3B). For each infection scenario (κ, R_0), we identified the ensemble of near-optimal designs, $\Theta(\kappa, R_0)$. A design $\vec{\theta}$ belongs to this ensemble if its incurred total harm (eq. 9) is within $10^{-2} \times S_{\max}$ of the minimum realized total harm from all sampled designs for that infection. Finally, we clustered infections based on their ensemble-averaged de-

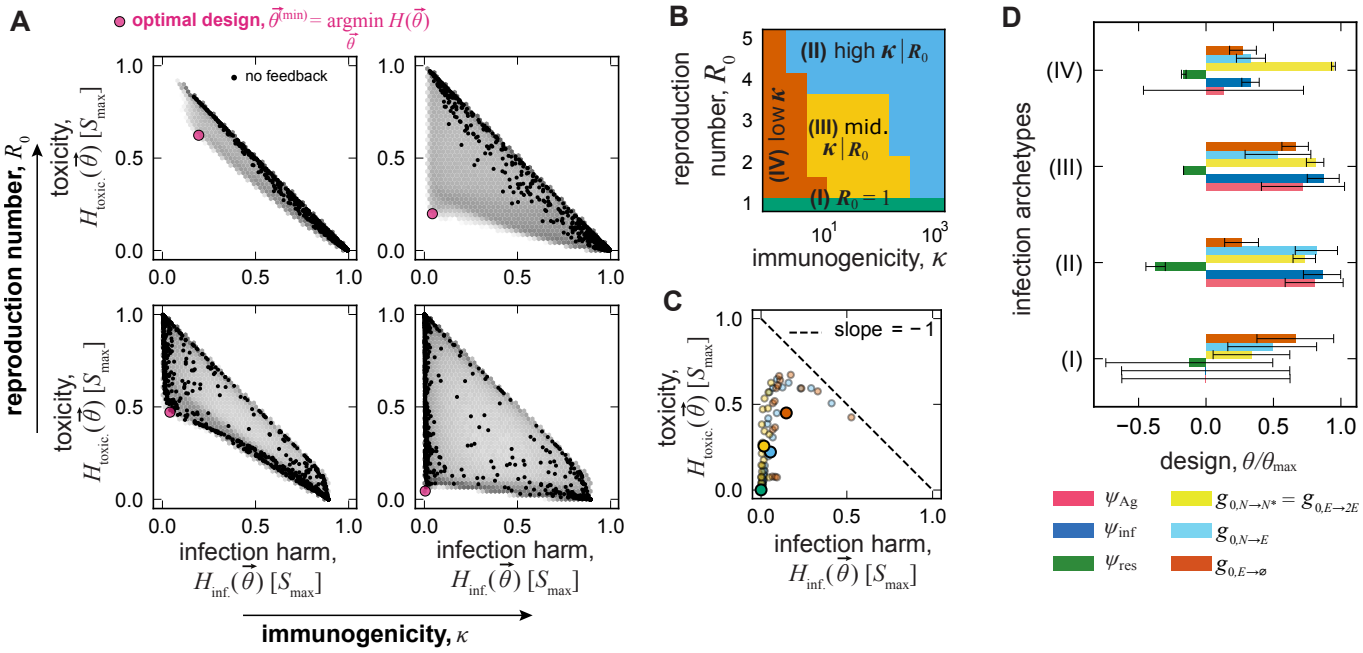

Figure 3. **Toxicity-infection harm trade-off across T-cell regulatory designs.** **(A)** Density plots of toxicity $H_{\text{toxic.}}$ versus infection harm $H_{\text{inf.}}$ (eq. 9)—measured in units of $S_{\max}$—for 4.8×10^6 T-cell regulatory designs, $\vec{\theta}$, uniformly grid-sampled across four infection scenarios (panels; columns: $\kappa = 5.6$ (left), $\kappa = 178$ (right); rows: $R_0 = 2.5$ (bottom), $R_0 = 5.0$ (top)). Designs are sampled uniformly across signal sensitivities ψ's $\in [-3, 3]$ and baseline rates ℓ's $\in [-3, 3]$ (eq. 8). Infection scenarios are sampled by changing K_I and b_I, while holding $S_{\max}$, d_I, and K_S fixed. Black points denote no-feedback designs ($\vec{\psi} = 0$), and the large magenta circle in each panel marks the design that minimizes total harm, $\vec{\theta}^{(\min)}$. **(B)** For many different infection scenarios (κ, R_0), we identify an ensemble of near-optimal designs $\Theta(\kappa, R_0)$, whose total harm is within $10^{-2} \times S_{\max}$ of the minimum sampled total harm (as in the magenta point in A). By computing the ensemble-averaged design parameters $\langle\vec{\theta}\rangle_{\Theta(\kappa,R_0)}$ for each infection scenario, and using these averages to cluster infections, we identify four infection "archetypes" (colors): I–IV, as indicated (SI Appendix). **(C)** Toxicity vs. infection harm associated with optimal designs for each archetype is shown. Large dots represent cluster averages and small transparent dots represent individual infections, sampled in (B). **(D)** Bars indicate the ensemble-mean of design parameters, $\langle\vec{\theta}\rangle_{\Theta^{(i)}}$, computed by identifying the ensemble of designs $\Theta^{(i)}$ which minimize cluster-averaged total harm for archetype i to within $10^{-2} \times S_{\max}$ of the minimum sampled cluster-averaged total harm (similar to (B)). Error bars denote the standard deviation within this ensemble. All design variables are normalized by their respective bounds ($\psi^{\max} = 3$, $\ell^{\max} = 3$). Unless otherwise noted, simulation parameters: $S_{\max} = 10^7, N_{\text{lin.}} = 10^{-5} S_{\max}, K_S = S_{\max}, K_H = 10^{-3} S_{\max}, d_I = 0.5\ \text{day}^{-1}, I_0 = 10^{-4} S_{\max}$.

sign parameters $\langle\vec{\theta}\rangle_{\Theta(\kappa,R_0)}$ (SI Appendix, Fig. S3). Four coherent design classes emerged, loosely corresponding to intuitive pathogen "archetypes" (Fig. 3B): **I**—non-replicating, self-like/auto-immune inducing or allergenic antigens ($R_0 = 1$); **II**—highly immunogenic bacteria and viruses (high κ or high R_0); **III**—intermediate replication and detectability, typifying mixed-strategy intracellular pathogens (mid-high R_0, mid-high κ); **IV**—difficult to detect threats such as growing tumors (low κ, broad range of R_0). The Pareto trade-off between infection harm and toxicity differs across these archetypes: designs optimized for archetypes I-III nearly eliminate infection damage with limited immunopathology, while designs in archetype IV show a strong trade-off (Fig. 3C).

Design parameters of effector programs minimizing harm vary across archetypes (Fig. 3D). For archetype I—non-replicating antigens—the optimal design is largely pre-programmed to suppress the immune response: signal sensitivities $\vec{\psi}$ are close to 0, and the baseline rates are tuned to slow activation, proliferation, and differentiation, and enforce rapid effector deletion. This damped response completely avoids toxicity when the threat does not grow so infection harm is determined by the initial exposure and cannot be decreased. Across all other archetypes a common pattern emerges—the response-derived cytokine signal σ_{res} acts as an anti-inflammatory brake ($\psi_{\text{res}} < 0$), with its magnitude tuned to context. Moreover, antigen sensitivity ψ_{Ag} grows with immunogenicity: the effector program optimized for highly immunogenic pathogens in archetype II shows the largest ψ_{Ag}, whereas the programs for the poorly immunogenic archetype IV show diminished antigen sensitivity. Compared to archetype II, the baseline propensities ℓ's (or equivalently g_0's) for archetype III accelerate effector contraction ($g_{0,E\to\emptyset}$ ↑). For the weakly immunogenic archetype IV, near-optimal designs distinctively tune both the baseline rates g_0's and the infection-signal sensitivity ψ_{inf} to balance clearance against toxicity and min-

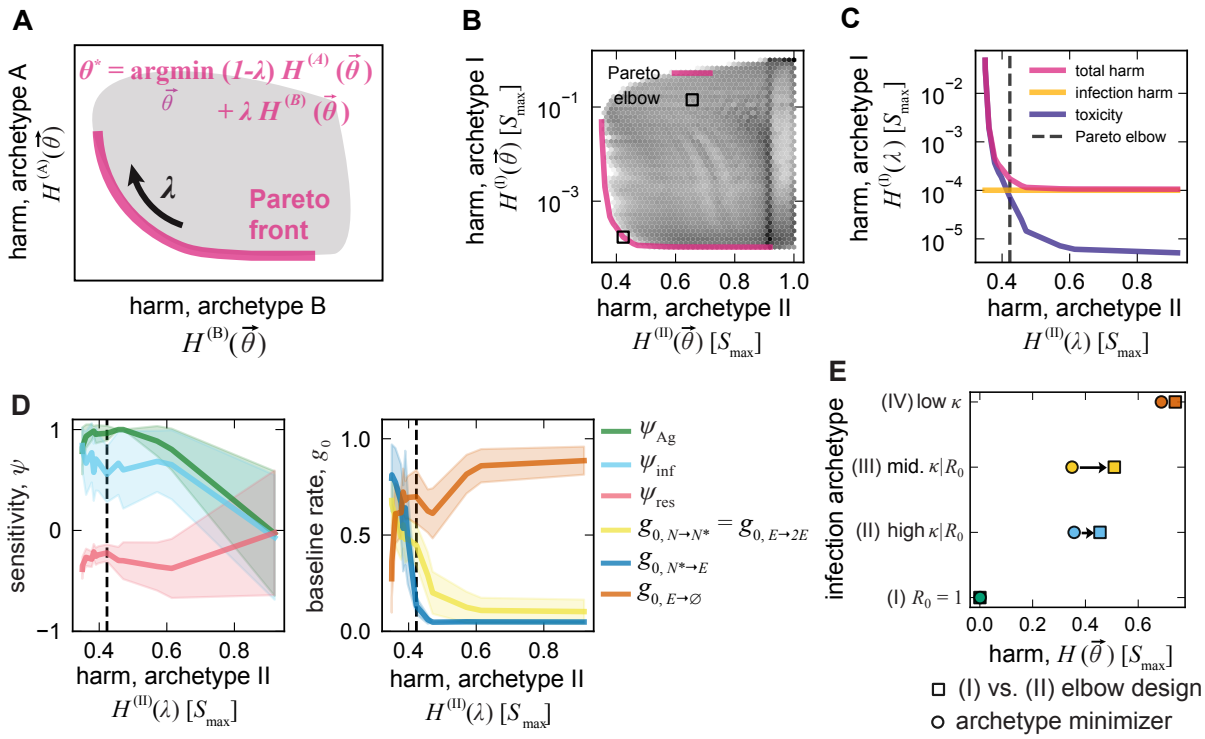

Figure 4. **Pareto-optimal designs balancing harm from acute infections and autoimmunity.** **(A)** Bi-objective optimization to minimize the weighted sum of harms associated with T-cell regulatory designs $\vec{\theta}$ across two infection archetypes (A, B) yields a Pareto front (magenta curve); eq. 10. Points on the curve are Pareto-optimal designs, parameterized by the weight $\lambda \in [0, 1]$, quantifying the emphasis on archetype B; moving left to right corresponds to decreasing λ (less emphasis on B). **(B)** The density plot shows harms incurred by 4.8×10^6 uniformly sampled designs in their response to archetype I (non-replicating, self-like/ auto-immune inducing or allergenic antigens; $R_0 = 1$) versus archetype II (immunogenic infection; high κ), as defined in Fig. 3B. Designs are sampled on a uniform Cartesian grid of sensitivities $\psi_\sigma \in [-3, 3]$ and baseline rates $\ell_i \in [-3, 3]$ (as in Fig. 3A). For many values of λ, we minimize the objective in (A), and trace out the Pareto front (magenta). For each point λ on the front, we identify the ensemble of designs closest to the Pareto front ($\Theta(\lambda)$), where a design belongs to the near-Pareto ensemble at λ if its incurred harm against archetypes I and II simultaneously lie within 5% of the λ-Pareto-optimal harm of the respective archetypes. The black square indicates the elbow of the Pareto curve λ^*: the point where further reductions in archetype-I harm incur larger increases in archetype-II harm (see SI Appendix). The average design parameter values over the near-elbow ensemble are $\langle\vec{\theta}\rangle_{\Theta(\lambda^*)} = (2.89, 1.67, -0.67, -0.22, -1.94, 0.94)$. **(C)** Decomposition of total harm along the Pareto front in (B) into infection-related harm and toxicity (eq. 9). The black dashed line indicates the position of the elbow from (B). **(D)** Solid lines show parameter magnitudes for the signal sensitivities ψ (left) and the baseline rates $g_0(\ell)$ (right), each scaled to variable bounds ($\psi^{\max} = 3$, $\ell^{\max} = 3$), as a function of harm incurred under archetype II along the Pareto curve (parameterized by λ). Shaded envelopes denote the standard deviation of the near-Pareto ensemble of designs $\Theta(\lambda)$, $\lambda \in [0, 1]$ traced out by the Pareto front. **(E)** For each archetype in Fig. 3B, total harm is shown for the archetype-specific minimizer $\vec{\theta}^{(\min)}$ (circle) and for the average over near-elbow ensemble in (B). See SI Appendix, Figs. S6, S7 for Pareto fronts formed by other pairs of archetypes. All simulation parameters other than the varied K_I, b_I, and $\vec{\theta}$ match Fig. 1.

imize harm (SI Appendix, Figs. S4, S5).

Endogenous T-cell programs are shaped by minimizing harm from autoimmunity and acute infections

Although the relative evolutionary pressures exerted by different threats are unknown, we can still pose a principled multi-objective problem for the effector program, i.e., find a design $\vec{\theta}$ that minimizes the weighted sum of harms across pathogen archetypes:

$$\vec{\theta}'(\{\lambda_i\}) = \arg\min_{\vec{\theta}} \sum_{i:\text{archetypes}} \lambda_i H^{(i)}(\vec{\theta}), \qquad (10)$$

where $\lambda_i \in [0, 1]$ with the constraint $\sum_i \lambda_i = 1$. The weights $\{\lambda_i\}$ determine the relative importance of each archetype in determining the optimal design $\vec{\theta}'$. The solution to this family of problems forms a simplex over which $\lambda_i' s$ vary (Fig. 4A; SI Appendix). This approach accommodates uncertainty about the empirical prevalence of each type of infection and the preferences/contingencies of evolution.

Even though optimization can be done simultaneously over all four archetypes, we proceed by examining pairwise trade-offs that generate a one dimensional Pareto front, parameterized by λ (Fig. 4A). A biologically salient case contrasts archetype I (non-replicating antigens) with archetype II (immunogenic pathogens). Non-replicating antigens plausibly arise from unprogrammed cell death events or external trauma inducing inflammation and may be implicated in autoimmune flares. This bi-objective harm minimization aims to find designs balancing effective clearance of immunogenic viral or bacterial infections and suppression of autoimmune responses; see SI Appendix, Figs. S6, S7, for Pareto fronts formed by optimizing other pairs of clusters.

Fig. 4B shows harm from archetype I plotted against that of archetype II for a broad range of sampled designs and highlights the Pareto front (magenta). Designs close to the Pareto front achieve near-best total protection for a given compromise weight λ; moving to the right along the Pareto front shifts the emphasis from clearance of immunogenic pathogens ($\lambda \uparrow$) to tolerance ($\lambda \downarrow$). Decomposing total harm into infection-related harm and toxicity (eq. 9) shows that along this Pareto front it is relatively easy to reduce toxicity to non-replicating antigens without substantially worsening outcomes against immunogenic infections (Fig. 4C).

Next, we evaluated immune designs near the Pareto curve (Fig. 4D; SI Appendix). For each point λ, we identified an ensemble of designs near the Pareto front, $\Theta(\lambda)$. For a design to belong to the near-Pareto ensemble, its incurred harm against archetypes I and II must simultaneously lie within 5% of the λ-Pareto-optimal harm of those respective archetypes. The *elbow* of the Pareto curve represents the point on the curve, λ^*, at which further decreases in harm from archetype I will incur disproportionally large increases in harm from archetype II (see SI Appendix). The ensemble of designs near the elbow $\Theta(\lambda^*)$ encodes a compelling logic: sensitivity from antigen and infection cues promotes effector production ($\psi_{\text{Ag}}, \psi_I > 0$), whereas response-derived cues act like a brake with a predominantly anti-inflammatory effect on the immune response ($\psi_{\text{res}} < 0$). Sensitivity to these signals diminishes as λ decreases (i.e., as avoidance of autoimmunity is prioritized); see Fig. 4D. Baseline rates

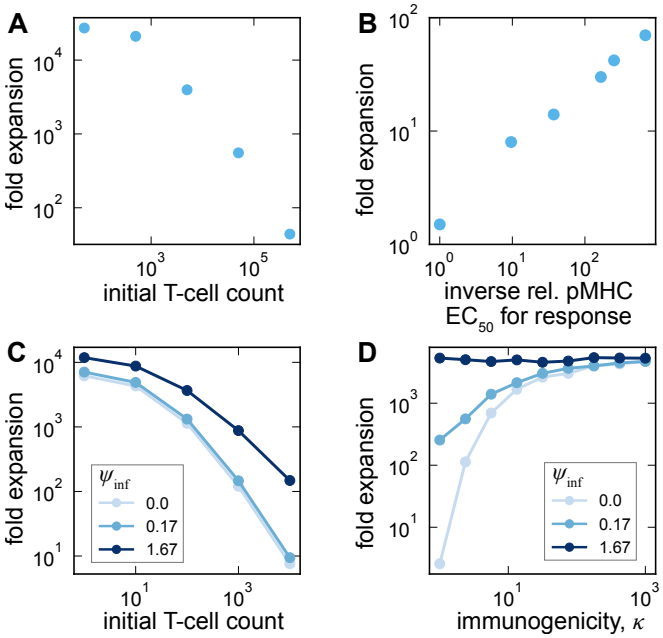

Figure 5. **Elbow-optimal designs recapitulating macro-dynamics of T-cell expansion. (A,B)** Fold expansion (i.e, peak clone size divided by the initial precursor number) of antigen-specific CD8$^+$ T-cells in mice plotted against (A) the initial precursor number (data from [9]), and (B) the relative pMHC concentration required for half-maximal activation EC_{50} (data from [62]); in our model, EC_{50} is proxied by K_I. **(C,D)** Simulated mean fold expansion for designs belonging to the near-elbow ensemble from Fig. 4B ($\Theta(\lambda^*)$) when their sensitivity to the infection-harm signal (ψ_{inf}) is varied (colors), shown as a function of (C) the initial precursor T-cell number, and (D) the immunogenicity of the antigen κ. For each design in the near-elbow ensemble identified in Fig. 4, $\vec{\theta} \in \Theta(\lambda^*)$, we change $\psi_{\text{inf}} \to c \times \psi_{\text{inf}}$ for constants $c \in \{0, 0.1, 1.0\}$ (colors), and simulate responses against the broad sample of infections (Fig. 3B). Trend lines are obtained by averaging over infections and designs in the near-elbow ensemble (SI Appendix, Figs. S8, S9). Elbow designs with reduced sensitivity to infection-derived harm qualitatively match the experimental macro-dynamics of T-cell expansion. Unless otherwise indicated, simulation parameters: $S_{\max} = 10^7, K_S = S_{\max}, K_H = 10^{-3} S_{\max}, d_I = 0.5\ \text{day}^{-1}, I_0 = 10^{-4} S_{\max}$.

Table I. Endogenous design of CD8$^+$ T-cell response.

Design Parameter	Effect	Source
baseline rates:		
1. activation & prolif., $g_{N\to N^*}, g_{E\to 2E}$	low	[4, 17]
2. differentiation, $g_{N^*\to E}$	low	[16]
3. contraction, $g_{E\to\emptyset}$	moderate	[17, 63]
sensitivity to signal:		
4. antigen, ψ_{Ag}	promote	[4, 6, 62]
5. infection, ψ_{inf}	promote	[5, 29, 64]
6. response, ψ_{res}	suppress	[28, 29]

(g_0's) of the designs at the elbow favor moderate to slow activation and proliferation of effectors, slow differentiation to effector state, and moderate contraction (Fig. 4D). As λ increases (leftward; clearance weighed more), all baseline rates shift toward promoting more effector cells. As λ decreases below the elbow (rightward), the already slow differentiation baseline rate falls toward zero, while activation/proliferation and contraction are tuned to suppress effector production, reflecting designs that minimize collateral damage even at the expense of incomplete clearance. Notably, despite being optimized only for archetypes I and II, designs in the near-elbow ensemble $\Theta(\lambda^*)$ perform competitively for other archetypes,

incurring harm close to the archetype-specific optimal designs (Fig. 4E; SI Appendix, Figs. S6, S7).

From these analyses alone, it is not possible to ascertain the compromises encoded in the endogenous T-cell response program. However, experiments on T-cell fate decisions in endogenous immunological programs (Table I) are qualitatively consistent with designs that jointly minimize harm from immunogenic pathogens and auto-immune responses to transient injuries (elbow in Fig. 4B).

The choice of design impacts the dynamics of the immune response, which we can compare to experimental data. Prior experiments in mice [9, 62] indicate that peak fold-expansion of T-cell populations in response to a pathogen scales inversely with the initial number of cognate T-cells, and with the pMHC concentration for half-maximal response (proxied by K_I in our model); see Fig. 5A, B. This pattern was recapitulated by explicitly modeling competition for antigens in polyclonal T-cell responses [65]. Near-elbow designs $\vec{\Theta}(\lambda^*)$ (Fig. 4) reproduce this inverse dependence on initial clone size but underestimate the steepness of the trend with respect to immunogenicity (κ) (Fig. 5C, D). Reducing the sensitivity to infection harm ψ_{inf} of these designs by an order of magnitude restores the observed dynamic range (Fig. 5D) while having only a minor effect on total harm, and hence, our optimization objective (Fig. S9). Interestingly, adjusting the other sensitivities ($\psi_{\text{Ag}}, \psi_{\text{res}}$), or the baseline rates g_0's does not reproduce the dynamic range observed in the experiments (Figs. S8, S9).

To align responses with the observed macro-dynamic trends, subsequent analyses consider near-elbow designs $\vec{\theta} \in \Theta(\lambda^*)$ with their infection-signal sensitivity reduced to one-tenth of their original value: for $\psi_{\text{inf}} \mapsto 0.1\,\psi_{\text{inf}}$. This ensemble of *tuned* designs is indicated by $\tilde{\Theta}(\lambda^*)$. See SI Appendix, Figs. S11 and S12 for extensive model parameter sensitivity analyses with this ensemble.

Designing an effector response to a growing tumor

CD8$^+$ T-cells are central to cancer surveillance and elimination of mutated pre-cancerous cells. We therefore asked how the tuned elbow-optimal designs $\vec{\theta} \in \tilde{\Theta}(\lambda^*)$ identified for infectious threats (Fig. 4B, Fig. 5) respond to cancer, which we modeled as a growing, low-immunogenic tumor that kills healthy cells as it expands (Fig. 6A; SI Appendix).

Similar to the case of infection control, a clearance-toxicity trade-off arises, but it is sharpened in cancer because many tumor antigens are only marginally distinct from self. Even high-affinity recognition risks cross-reactivity and bystander damage. Tuned near-elbow designs incur high total harm for low-immunogenic tumors (Fig. 6B,C). As tumor immunogenicity increases, harm declines due to decreases in cross-reactive toxicity and increases in clearance (Fig. 6C).

Cancer immunotherapy leverages and improves upon the body's own immune system to target tumors [66]. Patient-specific tumor-infiltrating lymphocytes (TILs) and TCR-engineered T-cells (TCR-T's) purified and expanded ex-vivo, and infused in patients to recognize tumor peptide-MHC's have shown clinical benefits. Chimeric antigen receptor (CAR) T-cells, on the other hand, target surface antigens without MHC restriction and can be deployed more broadly when targets are shared across patients. Contemporary CARs incorporate co-stimulatory domains (e.g., CD28 or 4-1BB) that enhance activation, proliferation, survival, and effector function [67]. In practice, three levers are pulled to engineer a T-cell therapy: (i) the antigen-recognition module (native TCR, TCR-T, or CAR) to specifically target the tumor; (ii) the intracellular signaling architecture (co-stimulation and engineered circuits) to sustains a robust and durable response; and (iii) the infused T-cell dose/schedule, tuned to maximize tumor control while limiting toxicity.

Our model predicts that the optimal dose should be tuned to tumor immunogenicity (Fig. 6B), consistent with prior findings on antigen-dependent T-cell dosing [68]. Examining the tuned near-elbow designs, we find: at low immunogenicity, clearance is inefficient and toxicity dominates, so smaller infusion doses minimize harm; at high immunogenicity, recognition and in-situ expansion are efficient, so even modest doses suffice; at intermediate immunogenicity, larger initial doses minimize total harm by accelerating tumor elimination while keeping toxicity tolerable (Fig. 6C).

TCR engineering raises an additional nuance: increasing TCR-tumor affinity often increases affinity to self [69], so that increasing tumor immunogenicity is difficult. To capture this, we evaluated treatment scenarios where the affinities/reactivities of tumor and self to the engineered TCR/CAR are proportional, leaving tumor immunogenicity unchanged (SI Appendix, Fig. S10). We quantify affinity or reactivity of a cell by the ratio of the maximum healthy cell counts $S_{\max}$ to the number of target cells required for half-maximal T-cell response (EC_{50}), i.e., $K_I^{-1} S_{\max}$ for tumor cells, and $K_S^{-1} S_{\max}$ for self targets. Provided that tumor immunogenicity ($\kappa = K_I^{-1} K_S$) is not too small and self-reactivity remains moderate, the dose-response as a function of tumor antigenicity (Fig. S10A) mirrors its dependence on immunogenicity in Fig. 6B. Nevertheless, these cross-reactive regimes (Fig. S10B) yield higher total harm than scenarios where self-interaction fixed so that increasing tumor antigenicity increases immunogenicity (Fig. 6C). This highlights the cost of affinity-driven cross-reactivity in TCR engineering for cancer therapy [70–72].

Finally, we asked which aspects of the endogenous CD8$^+$ T-cell program could be modulated to improve tumor control. Using high-dose infusions for a broad range of immunogenicities, we computed the sensitivity of negative harm (i.e., protection) to small changes in design parameter values, i.e., $-\nabla H(\vec{\theta})$ for $\vec{\theta} \in \tilde{\Theta}(\lambda^*)$ (Fig. 6D). For very high-affinity T-cells (those able to respond at

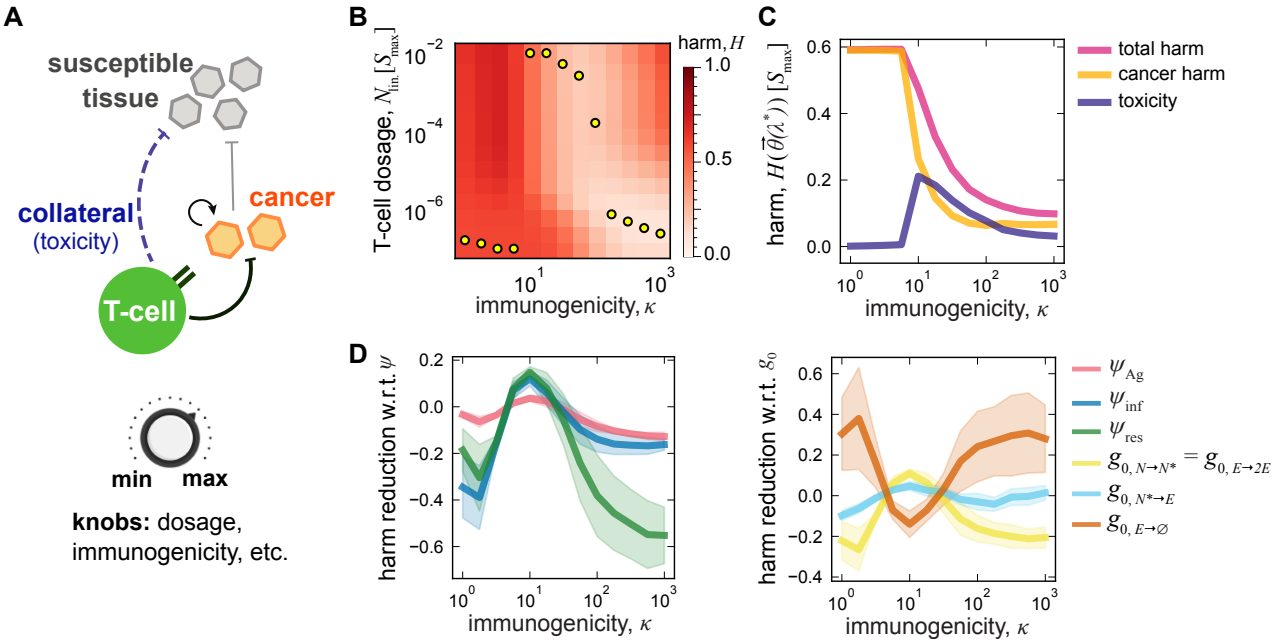

Figure 6. **Engineering immune response for cancer immunotherapy. (A)** The schematic shows the potential clearance of replicating cancer tumor cells by T-cells (e.g., infused during immunotherapy), and the potential collateral damage of therapy on susceptible tissue cells. The accessible levers in an immunotherapy include the dosage and antigenicity of the infused cells. Antigenicity is a property of the interaction between a specific TCR-antigen pair, and it can be increased by either changing the antigen or the TCR. Given a specific cognate T-cell, immunogenicity is the ratio of the antigenicity of a cancer cell to that of a healthy cell. **(B)** Averaged over the tuned near-elbow ensemble $\tilde{\Theta}(\lambda^*)$, the heatmap displays harm as a function of cancer antigenicity and dosage of administered T-cells. Yellow points correspond to the optimal dosage of T-cells for a given immunogenicity/antigenicity. **(C)** Total harm, harm from cancer and toxicity as functions of cancer immunogenicity (κ) are shown at the optimal T-cell dosage (yellow circles in B). **(D)** Clearance response—defined as the derivative of negative harm (i.e., protection) with respect to each design parameter θ_i, $-\partial H(\vec{\theta})/\partial\theta_i$ for $\vec{\theta} \in \tilde{\Theta}(\lambda^*)$—averaged over the tuned near-elbow designs $\tilde{\Theta}(\lambda^*)$, is shown for different levels of tumor immunogenicity κ; Left: $\vec{\psi}$ parameters; right: $\vec{g}_0$ parameters (colors). Shaded envelopes denote the standard deviation of outcomes over $\theta \in \tilde{\Theta}(\lambda^*)$. For this analysis, we assume that T-cells are administered at high dosage ($N_{\text{lin.}} = 10^{-2} S_{\max}$). Simulation parameters: $S_{\max} = 10^7, K_S = S_{\max}, K_H = 10^{-3} S_{\max}, d_I = 0.0\ \text{day}^{-1}, I_0 = 10^{-1} S_{\max}, b_I = 0.1\ \text{day}^{-1}$.

$\leq 1/100$ of the number of healthy cells), gains from further engineering are limited. For intermediate affinities—arguably more relevant clinically—several levers are impactful: increasing activation/proliferation baseline rates ($g_{0,N\to N^*}, g_{0,E\to 2E} \uparrow$) [73] and decreasing effector death ($g_{0,E\to\emptyset} \downarrow$) provide strong improvements but may entail perturbing oncogene-linked pathways (e.g., MYC, BIM) and warrant caution [74]. Boosting antigen and infection sensitivities ($\psi_{\text{Ag}}, \psi_{\text{inf}} \uparrow$), and tempering the anti-inflammatory brake ($|\psi_{\text{res}}| \downarrow$) also enhance clearance, albeit with a risk of runaway inflammation [75]. Increasing the baseline rate of differentiation ($g_{0,N^*\to E} \uparrow$) yields smaller gains. This suggests that reducing effector differentiation—e.g., via Tcf7 modulation [16, 73]—may provide a tractable knob to promote memory production and reduce relapse risk. Reduced effector differentiation may then be compensated for by tuning effector expansion and longevity.

Interestingly, recent CRISPR screens show that among other candidates, knocking out FAS and PTPN2 genes in CAR T-cells reduces apoptosis ($g_{0,E\to\emptyset} \downarrow$) and redirects differentiation toward effector programs ($g_{0,E\to 2E} \uparrow$), thereby enhancing CAR T-cell efficacy [76–79]. Our modeling framework can help prioritize targets for such screens to rationally identify candidate genes that improve T-cell immunotherapies. It should be noted that the achievable benefit of any perturbation for T-cell engineering will depend on patient- and tumor-specific constraints. These include the extent to which affinity-driven TCR cross-reactivity can be mitigated (Fig. S10) and may require patient-specific measurements of T-cell interactions and signaling to calibrate designs for robust, durable therapy.

DISCUSSION

In this work, we modeled the $CD8^+$ T-cell response as a feedback-controlled program that integrates three cues—instantaneous antigen, cumulative infection-derived harm, and cumulative response-derived harm—to regulate activation, expansion, differentiation, and contraction of T-cells. Modulation of response proportional to the antigen signal and the time-integrated harm signals mirrors proportional-integral-derivative (PID) control of engineered systems, a strategy known for its stable and robust steering of dynamical processes toward desired outcomes [80].

Exploring an ensemble of biologically motivated controllers (designs) and optimizing for pathogen clearance while minimizing immunopathology exposed a Pareto trade-off: a reduction in infection harm generally leads to an increased collateral damage to healthy tissues. This pattern persisted across diverse pathogenic challenges defined by their immunogenicity and within-host basic reproduction number.

Three design principles emerged: (i) Pre-programmed (baseline-tuned) designs help restrain autoimmunity and slowly proliferating threats but are inefficient against fast-proliferating infections or metastatic tumors; effective control there requires signal integration and feedback. (ii) Among signal sources, antigen and response-harm dominate; sensitivity to infection-harm contributes comparatively little once $CD8^+$ effector T-cell expansion is underway. (iii) Feedback from response-harm acts as an anti-inflammatory brake, enabling strong antigen-driven responses while limiting collateral damage when the response overshoots—a characteristic consistent with endogenous effector programs.

We posited that effector programs should attempt to solve a multi-objective problem to counter diverse pathogenic challenges. Interestingly, we found that a Pareto-optimal design that suppresses responses to frequent, self-like injuries while clearing recognizable, replicating (e.g., acute) infections mirrors the endogenous T-cell programs in vertebrates: strongly antigen-driven with an anti-inflammatory brake, and conservative cell-fate transition speeds that prevent runaway effector responses.

We assumed that, although memory formation is essential, its immediate impact on primary effector magnitude is modest: $5-10\%$ of peak effector numbers [16, 81]. Memory "quality" (longevity, proliferative capacity) depends on differentiation timing and the avidity of the selected clones [15, 21, 82, 83]. To isolate how signal integration shapes acute control, we did not explicitly model memory differentiation or recall. An important next step is to extend the framework to the allocation trade-off in which diverting cells to memory lowers primary effector counts but seeds future protection, thereby linking our high-level control design to memory formation and the host's evolutionary survival pressures that could have shaped the effector program.

Using a minimal tumor model, we showed how the design objective informs rational immunotherapy protocols. First, it suggests that infused T-cell dosage should depend on tumor immunogenicity—balancing clearance against collateral damage, as previously emphasized in the literature [66, 68, 71]. Second, treating the design parameters as engineering targets yields testable strategies to improve efficacy and safety, which is highly desirable [84–86]. For moderately immunogenic tumors, promoting effector output by raising baseline rates of activation/proliferation or increasing antigen and infection sensitivities, while tempering the anti-inflammatory brake, can produce robust responses.

To operationalize these ideas, abstract signal sensitivities and baseline rates must be mapped to gene programs as targets of cell engineering (Table I). Combining principled control-theoretic design with patient-specific measurements of antigenicity and damage signaling may enable tailored effector programs that respect the fundamental trade-offs of response from this study, while bending them toward better outcomes [87].

Key questions follow from our study. Our coarse-grained model elides many details of tissue biology and micro-environment: how do tissue contexts impact the state of resident cells and shift clearance-toxicity trade-offs [88, 89]? Endogenous immune responses are polyclonal [90]: how do the proposed designs shape information flow and selection of dominant clones in polyclonal responses? Do the optimization principles posited here predispose the immune response to certain vulnerabilities, and what mechanisms might have evolved to mask them? Addressing these issues through the lens of cell economics and immunological trade-offs may clarify the roles of chance, contingency, and necessity in the evolution of effector programs.

II. DATA AND CODE AVAILABILITY

Data and code for all the analyses can be found in the Github repository: https://github.com/StatPhysBio/designimmune.

ACKNOWLEDGEMENT

This work has been supported by the CAREER award from the National Science Foundation grant 2045054 (A.N., O.A.U., Z.M.), the National Institutes of Health MIRA award R35 GM142795 (A.N.), and the Intramural Research Program of the National Cancer Institute, Project # ZIA BC 012007 (G.A.-B.). This work is also supported, in part, through the Departments of Physics and Applied Mathematics and the College of Arts and Sciences at the University of Washington. The numerical analyses in this work were completed on Hyak, the UW's high performance computing cluster, which is funded by the UW student technology fee. This work benefited from

discussions during the 2023 summer workshop "Statistical Physics and Adaptive Immunity" at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-2210452, and the 2024 program "Interactions and Co-evolution between Viruses and Immune Systems" at the Kavli Institute for theoretical physics (KITP), which is supported by National Science Foundation grant PHY-2309135, and the Gordon and Betty Moore Foundation Grant No. 2919.02.

[1] Q. Qi, Y. Liu, Y. Cheng, J. Glanville, D. Zhang, J.-Y. Lee, R. A. Olshen, C. M. Weyand, S. D. Boyd, and J. J. Goronzy, Proc. Natl. Acad. Sci. U. S. A. **111**, 13139 (2014).
[2] G. Lythe, R. E. Callard, R. L. Hoare, and C. Molina-París, J. Theor. Biol. **389**, 214 (2016).
[3] T. Mora and A. M. Walczak, Curr. Opin. Syst. Biol. **18**, 104 (2019).
[4] M. J. B. van Stipdonk, G. Hardenberg, M. S. Bijker, E. E. Lemmens, N. M. Droin, D. R. Green, and S. P. Schoenberger, Nat. Immunol. **4**, 361 (2003).
[5] N. S. Joshi, W. Cui, A. Chandele, H. K. Lee, D. R. Urso, J. Hagman, L. Gapin, and S. M. Kaech, Immunity **27**, 281 (2007).
[6] J. M. Marchingo, A. Kan, R. M. Sutherland, K. R. Duffy, C. J. Wellard, G. T. Belz, A. M. Lew, M. R. Dowling, S. Heinzel, and P. D. Hodgkin, Science **346**, 1123 (2014).
[7] J. M. Marchingo, G. Prevedello, A. Kan, S. Heinzel, P. D. Hodgkin, and K. R. Duffy, Nat. Commun. **7**, 13540 (2016).
[8] I. Eizenberg-Magar, J. Rimer, I. Zaretsky, D. Lara-Astiaso, S. Reich-Zeliger, and N. Friedman, Proc. Natl. Acad. Sci. U. S. A. **114**, E6447 (2017).
[9] V. P. Badovinac, J. S. Haring, and J. T. Harty, Immunity **26**, 827 (2007).
[10] N. Zhang and M. J. Bevan, Immunity **35**, 161 (2011).
[11] V. R. Buchholz, M. Flossdorf, I. Hensel, L. Kretschmer, B. Weissbrich, P. Gräf, A. Verschoor, M. Schiemann, T. Höfer, and D. H. Busch, Science **340**, 630 (2013).
[12] S. Sarkar, V. Kalia, W. N. Haining, B. T. Konieczny, S. Subramaniam, and R. Ahmed, J. Exp. Med. **205**, 625 (2008).
[13] M. R. Jenkins, J. Mintern, N. L. La Gruta, K. Kedzierska, P. C. Doherty, and S. J. Turner, J. Immunol. **181**, 3818 (2008).
[14] R. A. Seder, P. A. Darrah, and M. Roederer, Nat. Rev. Immunol. **8**, 247 (2008).
[15] B. Youngblood, J. S. Hale, H. T. Kissick, E. Ahn, X. Xu, A. Wieland, K. Araki, E. E. West, H. E. Ghoneim, Y. Fan, P. Dogra, C. W. Davis, B. T. Konieczny, R. Antia, X. Cheng, and R. Ahmed, Nature **552**, 404 (2017).
[16] K. Abadie, E. C. Clark, R. M. Valanparambil, O. Ukogu, W. Yang, R. M. Daza, K. K. H. Ng, J. Fathima, A. L. Wang, J. Lee, T. H. Nasti, A. Bhandoola, A. Nourmohammad, R. Ahmed, J. Shendure, J. Cao, and H. Y. Kueh, Immunity **57**, 271 (2024).
[17] S. Heinzel, T. Binh Giang, A. Kan, J. M. Marchingo, B. K. Lye, L. M. Corcoran, and P. D. Hodgkin, Nat. Immunol. **18**, 96 (2017).
[18] Y.-L. Cho, M. Flossdorf, L. Kretschmer, T. Höfer, D. H. Busch, and V. R. Buchholz, Cell Rep. **20**, 806 (2017).
[19] M. Plambeck, A. Kazeroonian, D. Loeffler, L. Kretschmer, C. Salinno, T. Schroeder, D. H. Busch, M. Flossdorf, and V. R. Buchholz, Proceedings of the National Academy of Sciences **119**, e2116260119 (2022).
[20] A. Xin, F. Masson, Y. Liao, S. Preston, T. Guan, R. Gloury, M. Olshansky, J.-X. Lin, P. Li, T. P. Speed, G. K. Smyth, M. Ernst, W. J. Leonard, M. Pellegrini, S. M. Kaech, S. L. Nutt, W. Shi, G. T. Belz, and A. Kallies, Nat. Immunol. **17**, 422 (2016).
[21] L. Kretschmer, M. Flossdorf, J. Mir, Y.-L. Cho, M. Plambeck, I. Treise, A. Toska, S. Heinzel, M. Schiemann, D. H. Busch, and V. R. Buchholz, Nat. Commun. **11**, 113 (2020).
[22] A. E. Downie, A. Mayer, C. J. E. Metcalf, and A. L. Graham, PLoS Comput. Biol. **17**, e1009714 (2021).
[23] A. Mayer, T. Mora, O. Rivoire, and A. M. Walczak, Proc. Natl. Acad. Sci. U. S. A. **113**, 8630 (2016).
[24] H. A. van den Berg and Y. N. Kiselëv, Bull. Math. Biol. **66**, 1345 (2004).
[25] D. R. Green, T. Ferguson, L. Zitvogel, and G. Kroemer, Nat. Rev. Immunol. **9**, 353 (2009).
[26] K. Labbé and M. Saleh, Cell Death Differ. **15**, 1339 (2008).
[27] T. C. Butler, M. Kardar, and A. K. Chakraborty, Proc. Natl. Acad. Sci. U. S. A. **110**, 11833 (2013).
[28] J. M. Cicchese, S. Evans, C. Hult, L. R. Joslyn, T. Wessler, J. A. Millar, S. Marino, N. A. Cilfone, J. T. Mattila, J. J. Linderman, and D. E. Kirschner, Immunol. Rev. **285**, 147 (2018).
[29] R. Medzhitov, Science **374**, 1070 (2021).
[30] P. François and V. Hakim, Proc. Natl. Acad. Sci. U. S. A. **101**, 580 (2004).
[31] G. Shinar, R. Milo, M. R. Martínez, and U. Alon, Proc. Natl. Acad. Sci. U. S. A. **104**, 19931 (2007).
[32] J.-B. Lalanne and P. François, Phys. Rev. Lett. **110**, 218102 (2013).
[33] Y. Hart, S. Reich-Zeliger, Y. E. Antebi, I. Zaretsky, A. E. Mayo, U. Alon, and N. Friedman, Cell **158**, 1022 (2014).
[34] U. Alon, *An introduction to systems biology: Design principles of biological circuits* (Chapman and Hall/CRC, Second edition. — Boca Raton, Fla. : CRC Press, [2019], 2019).
[35] T. R. Sokolowski, T. Gregor, W. Bialek, and G. Tkačik, Proceedings of the National Academy of Sciences **122**, e2402925121 (2025), https://www.pnas.org/doi/pdf/10.1073/pnas.2402925121.
[36] A. S. Perelson and G. F. Oster, J. Theor. Biol. **81**, 645 (1979).
[37] A. Mayer, V. Balasubramanian, T. Mora, and A. M. Walczak, Proc. Natl. Acad. Sci. U. S. A. **112**, 5950 (2015).
[38] S. Bradde, A. Nourmohammad, S. Goyal, and V. Balasubramanian, Proc. Natl. Acad. Sci. U. S. A. **117**, 5144 (2020).
[39] A. Mayer, V. Balasubramanian, A. M. Walczak, and T. Mora, Proc. Natl. Acad. Sci. U. S. A. **116**, 8815 (2019).
[40] O. H. Schnaack and A. Nourmohammad, Elife **10** (2021).

[41] V. Chardès, M. Vergassola, A. M. Walczak, and T. Mora, arXiv [q-bio.PE] , 2021.07.26.453765 (2021).

[42] O. H. Schnaack, L. Peliti, and A. Nourmohammad, Phys. Rev. X **12**, 021063 (2022).

[43] Y. Qin, E. M. Mace, and J. P. Barton, Proc. Natl. Acad. Sci. U. S. A. **120**, e2305859120 (2023).

[44] D. L. Chao, M. P. Davenport, S. Forrest, and A. S. Perelson, J. Theor. Biol. **228**, 227 (2004).

[45] V. V. Ganusov, D. L. Barber, and R. J. De Boer, PLoS One **6**, e15959 (2011).

[46] S. Halle, K. A. Keyser, F. R. Stahl, A. Busche, A. Marquardt, X. Zheng, M. Galla, V. Heissmeyer, K. Heller, J. Boelter, K. Wagner, Y. Bischoff, R. Martens, A. Braun, K. Werth, A. Uvarovskii, H. Kempf, M. Meyer-Hermann, R. Arens, M. Kremer, G. Sutter, M. Messerle, and R. Förster, Immunity **44**, 233 (2016).

[47] G. N. Milligan and A. D. T. Barrett, *Vaccinology: An Essential Guide*, edited by G. N. Milligan and A. D. T. Barrett (Wiley-Blackwell, Hoboken, NJ, 2015).

[48] P. Abuin, A. Anderson, A. Ferramosca, E. A. Hernandez-Vargas, and A. H. Gonzalez, Annu. Rev. Control **50**, 457 (2020).

[49] M. J. van Stipdonk, E. E. Lemmens, and S. P. Schoenberger, Nat. Immunol. **2**, 423 (2001).

[50] P. François, G. Voisinne, E. D. Siggia, G. Altan-Bonnet, and M. Vergassola, Proc. Natl. Acad. Sci. U. S. A. **110**, E888 (2013).

[51] L. Gattinoni, E. Lugli, Y. Ji, Z. Pos, C. M. Paulos, M. F. Quigley, J. R. Almeida, E. Gostick, Z. Yu, C. Carpenito, E. Wang, D. C. Douek, D. A. Price, C. H. June, F. M. Marincola, M. Roederer, and N. P. Restifo, Nat. Med. **17**, 1290 (2011).

[52] W. de Ronde, P. Rein ten Wolde, and A. Mugler, Biophys. J. **103**, 1097 (2012).

[53] A. M. Walczak, G. Tkacik, and W. Bialek, Phys. Rev. E Stat. Nonlin. Soft Matter Phys. **81**, 041905 (2010).

[54] A. Oyler-Yaniv, J. Oyler-Yaniv, B. M. Whitlock, Z. Liu, R. N. Germain, M. Huse, G. Altan-Bonnet, and O. Krichevsky, Immunity **46**, 609 (2017).

[55] S. R. Achar, F. X. P. Bourassa, T. J. Rademaker, A. Lee, T. Kondo, E. Salazar-Cavazos, J. S. Davies, N. Taylor, P. François, and G. Altan-Bonnet, Science **376**, 880 (2022).

[56] G. Voisinne, B. G. Nixon, A. Melbinger, G. Gasteiger, M. Vergassola, and G. Altan-Bonnet, Cell Rep. **11**, 1208 (2015).

[57] M. Adler, A. Mayo, and U. Alon, PLoS Comput. Biol. **10**, e1003781 (2014).

[58] S. Heinzel, J. M. Marchingo, M. B. Horton, and P. D. Hodgkin, Curr. Opin. Immunol. **51**, 32 (2018).

[59] A. S. Gupta, V. Chevée, A. S. Kirosingh, N. M. Davis, and D. S. Schneider, bioRxiv , 2021.09.29.462483 (2021).

[60] Y. Lebel, A. S. Gupta, V. Chevée, U. Alon, and D. S. Schneider, bioRxiv , 2025.03.04.641508 (2025).

[61] P. L. F. Johnson, B. F. Kochin, M. S. McAfee, I. M. Stromnes, R. R. Regoes, R. Ahmed, J. N. Blattman, and R. Antia, J. Virol. **85**, 5565 (2011).

[62] D. Zehn, S. Y. Lee, and M. J. Bevan, Nature **458**, 211 (2009).

[63] R. J. De Boer, M. Oprea, R. Antia, K. Murali-Krishna, R. Ahmed, and A. S. Perelson, J. Virol. **75**, 10663 (2001).

[64] J. N. Blattman, J. M. Grayson, E. J. Wherry, S. M. Kaech, K. A. Smith, and R. Ahmed, Nat. Med. **9**, 540 (2003).

[65] A. Mayer, Y. Zhang, A. S. Perelson, and N. S. Wingreen, Proc. Natl. Acad. Sci. U. S. A. **116**, 5914 (2019).

[66] E. Baulu, C. Gardet, N. Chuvin, and S. Depil, Sci. Adv. **9**, eadf3700 (2023).

[67] I. Zugasti, L. Espinosa-Aroca, K. Fidyt, V. Mulens-Arias, M. Diaz-Beya, M. Juan, Ã. Urbano-Ispizua, J. Esteve, T. Velasco-Hernandez, and P. Menéndez, Signal Transduct. Target. Ther. **10**, 210 (2025).

[68] A. Rotte, M. J. Frigault, A. Ansari, B. Gliner, C. Heery, and B. Shah, J. Immunother. Cancer **10**, e005678 (2022).

[69] M. M. Hoffmann and J. E. Slansky, Mol. Carcinog. **59**, 862 (2020).

[70] B. Liu, N. F. Greenwood, J. E. Bonzanini, A. Motmaen, J. Meyerberg, T. Dao, X. Xiang, R. Ault, J. Sharp, C. Wang, G. M. Visani, D. K. Vafeados, N. Roullier, A. Nourmohammad, D. A. Scheinberg, K. C. Garcia, and D. Baker, Science **389**, 386 (2025).

[71] T. Kondo, F. X. P. Bourassa, S. Achar, J. DuSold, P. F. Céspedes, M. Ando, A. Dwivedi, J. Moraly, C. Chien, S. Majdoul, A. L. Kenet, M. Wahlsten, A. Kvalvaag, E. Jenkins, S. P. Kim, C. M. Ade, Z. Yu, G. Gaud, M. Davila, P. Love, J. C. Yang, M. L. Dustin, G. Altan-Bonnet, P. François, and N. Taylor, Cell **0** (2025).

[72] G. M. Visani, M. N. Pun, A. A. Minervina, P. Bradley, P. Thomas, and A. Nourmohammad, bioRxivorg , 2025.02.28.640903 (2025).

[73] G. Escobar, K. Tooley, J. P. Oliveras, L. Huang, H. Cheng, M. L. Bookstaver, C. Edwards, E. Froimchuk, C. Xue, D. Mangani, R. K. Krishnan, N. Hazel, C. Rutigliani, C. M. Jewell, L. Biasco, and A. C. Anderson, Cancer Cell **41**, 1662 (2023).

[74] N. Muthalagu, M. R. Junttila, K. E. Wiese, E. Wolf, J. Morton, B. Bauer, G. I. Evan, M. Eilers, and D. J. Murphy, Cell Rep. **8**, 1347 (2014).

[75] J. R. Tisoncik, M. J. Korth, C. P. Simmons, J. Farrar, T. R. Martin, and M. G. Katze, Microbiol. Mol. Biol. Rev. **76**, 16 (2012).

[76] T. Yoshikawa, Z. Wu, S. Inoue, H. Kasuya, H. Matsushita, Y. Takahashi, H. Kuroda, W. Hosoda, S. Suzuki, and Y. Kagoya, Blood **139**, 2156 (2022).

[77] F. Yi, T. Cohen, N. Zimmerman, F. Dündar, P. Zumbo, R. Eltilib, E. J. Brophy, H. Arkin, J. Feucht, M. V. Gormally, C. S. Hackett, K. N. Kropp, I. Etxeberria, S. S. Chandran, Z. Zhao, W. Cai, A. F. Daniyan, J. H. Park, C. A. Lareau, K. C. Hsu, M. Sadelain, D. Betel, and C. A. Klebanoff, Nat. Cancer , 1 (2025).

[78] N. H. Knudsen, G. Escobar, F. Korell, T. Kienka, C. Nobrega, S. Anderson, A. Y. Cheng, M. Zschummel, A. Armstrong, A. Bouffard, M. C. Kann, S. Goncalves, H. W. Pope, M. Pezeshki, A. Rojas, J. S. M. T. Suermondt, M. Phillips, T. R. Berger, S. Park, D. Salas-Benito, E. P. Darnell, F. Birocchi, M. B. Leick, R. C. Larson, J. G. Doench, D. Sen, K. B. Yates, R. T. Manguso, and M. V. Maus, Nature , 1 (2025).

[79] P. Datlinger, E. V. Pankevich, C. D. Arnold, N. Pranckevicius, J. Lin, D. Romanovskaia, M. Schaefer, F. Piras, A.-C. Orts, A. Nemc, P. N. Biesaga, M. Chan, T. Neuwirth, A. V. Artemov, W. Li, S. Ladstätter, T. Krausgruber, and C. Bock, Nature , 1 (2025).

[80] K. H. Ang, G. Chong, and Y. Li, IEEE Trans. Control Syst. Technol. **13**, 559 (2005).

[81] S. M. Kaech and W. Cui, Nat. Rev. Immunol. **12**, 749 (2012).

[82] K. Bresser, L. Kok, A. C. Swain, L. A. King, L. Jacobs, T. S. Weber, L. Perié, K. R. Duffy, R. J. de Boer, F. A. Scheeren, and T. N. Schumacher, Nat. Immunol. **23**, 791 (2022).

[83] S. Grassmann, L. Mihatsch, J. Mir, A. Kazeroonian, R. Rahimi, S. Flommersfeld, K. Schober, I. Hensel, J. Leube, L. O. Pachmayr, L. Kretschmer, Q. Zhang, A. Jolly, M. Z. Chaudhry, M. Schiemann, L. Cicin-Sain, T. Höfer, D. H. Busch, M. Flossdorf, and V. R. Buchholz, Nat. Immunol. **21**, 1563 (2020).

[84] S. Xia, A. C. Lu, V. Tobin, K. Luo, L. Moeller, D. J. Shon, R. Du, J. M. Linton, M. Sui, F. Horns, and M. B. Elowitz, Cell **187**, 2785 (2024).

[85] H.-S. Li, D. V. Israni, K. A. Gagnon, K. A. Gan, M. H. Raymond, J. D. Sander, K. T. Roybal, J. K. Joung, W. W. Wong, and A. S. Khalil, Science **378**, 1227 (2022).

[86] G. M. Allen, N. W. Frankel, N. R. Reddy, H. K. Bhargava, M. A. Yoshida, S. R. Stark, M. Purl, J. Lee, J. L. Yee, W. Yu, A. W. Li, K. C. Garcia, H. El-Samad, K. T. Roybal, M. H. Spitzer, and W. A. Lim, Science **378**, eaba1624 (2022).

[87] M. Lässig, V. Mustonen, and A. Nourmohammad, Nat. Rev. Genet. **24**, 851 (2023).

[88] C. H. van Dorp, J. I. Gray, D. H. Paik, D. L. Farber, and A. J. Yates, PLoS Comput. Biol. **21**, e1013242 (2025).

[89] E. R. Jerison, N. Romeo, and S. R. Quake, bioRxiv , 2025.01.28.635318 (2025).

[90] A. Straub, S. Grassmann, S. Jarosch, L. Richter, P. Hilgendorf, M. Hammel, K. I. Wagner, V. R. Buchholz, K. Schober, and D. H. Busch, Immunity (2023).

Supplementary Information

Design principles of the cytotoxic CD8$^+$ T-cell response

Obinna A. Ukogu, Zachary Montague, Grégoire Altan-Bonnet & Armita Nourmohammad

CONTENTS

I. Model of T-cell response to infections 2
- Competitive binding of T-cells to healthy and infected cells 2
- Dynamics of growing (acute) infections 3
- Dynamics of slow-growing tumors 3
- Sampling pathogen space 4
- Harm during T-cell response to pathogens 4
- Signal cues for T-cell response 5
- Lineage-based model of CD8$^+$ T-cell response 5

II. Numerical implementation of stochastic T-cell response 8

III. Optimization of CD8$^+$ T-cell response to pathogens 12
- Design architecture of T-cell programs for infection response 12
- Optimizing the performance of response designs 12
- Clustering of infection archetypes 13

References 27

I. MODEL OF T-CELL RESPONSE TO INFECTIONS

Competitive binding of T-cells to healthy and infected cells

T-cells are cross-reactive and can bind to a broad range of antigens. A cognate T-cell for a pathogen-derived antigen can be cross-reactive to antigens derived from self-proteins. Consequently, when analyzing T-cell recognition, it is essential to account for the potential competition between binding to cells presenting cognate non-self antigens and to self-antigens.

To model this effect in a simple way, we consider the following coupled reactions:

$$\begin{aligned} E^{(u)} + I^{(u)} &\underset{k_I^d}{\overset{k_I^a}{\rightleftharpoons}} C_I \\ E^{(u)} + S^{(u)} &\underset{k_S^d}{\overset{k_S^a}{\rightleftharpoons}} C_S \end{aligned} \tag{1}$$

where $E^{(u)}$ denotes the number of unbound effector T-cells, $I^{(u)}$ and $S^{(u)}$ represent the numbers of unbound infected and susceptible cells, respectively, and C_I and C_S correspond to the bound complexes of effector T-cells with those cells. Note that we take the number of infected and susceptible cells as a proxy for the amount non-self and self antigens presented, respectively. The parameters $k^a_{(\cdot)}$ and $k^d_{(\cdot)}$ are the rate constants for the association and dissociation of complexes $C_{(\cdot)}$, where the placeholder $(\cdot)$ can be substituted with either I (infected) or S (healthy). The dynamics of these complexes over time follows,

$$\begin{aligned} \frac{\mathrm{d}C_I}{\mathrm{d}t} &= k_I^a E^{(u)} I^{(u)} - k_I^d C_I \\ \frac{\mathrm{d}C_S}{\mathrm{d}t} &= k_S^a E^{(u)} S^{(u)} - k_S^d C_S. \end{aligned}$$

In this picture, we assume a separation of timescales: binding and unbinding events occur much more rapidly than T-cell proliferation and effector killing of bound cells. The bound complexes can be treated as being in quasi-equilibrium at all times, i.e., $\mathrm{d}C_{(\cdot)}/\mathrm{d}t = 0$, which yields,

$$E^{(u)} I^{(u)} = K_I C_I, \qquad E^{(u)} S^{(u)} = K_S C_S, \tag{2}$$

where $K_I = k_I^d / k_I^a$ and $K_S = k_S^d / k_S^a$ are the dissociation constants of the effector cells from infected and healthy cells, respectively.

Expressing eq. 2 in terms of the total concentrations of effectors $E = E^{(u)} + C_I + C_S$, infected cells $I = I^{(u)} + C_I$, and healthy cells $S = S^{(u)} + C_S$, we arrive at,

$$\begin{aligned} 0 &= (E - C_I - C_S)(I - C_I) - K_I C_I \\ 0 &= (E - C_I - C_S)(S - C_S) - K_S C_S \end{aligned}$$

Assuming that the number of bound complexes is small relative to that of unbound cells or T-cells, we can neglect the second order terms $\mathcal{O}(C_I^2), \mathcal{O}(C_S^2), \mathcal{O}(C_I C_S)$, and solve for these abundances in terms of the total number of effectors, and the infected and healthy cells,

$$C_I = \frac{EI}{E + I + K_I + \frac{E+K_I}{E+K_S} S} \quad , \quad C_S = \frac{ES}{E + S + K_S + \frac{E+K_S}{E+K_I} I}. \tag{3}$$

The correction terms in the denominators ($\frac{E+K_I}{E+K_S} S$, and $\frac{E+K_S}{E+K_I} I$) capture the competition between binding to self- and non-self-derived antigens. In the limit $K_S \gg K_I$ (i.e., when T-cells bind more strongly to infected cells than healthy ones), the antagonistic effect vanishes and C_I approaches the form for response in the presence of a single-antigen only, consistent with ref. [1, 2]. However, as $K_S \to K_I$ (self and non-self antigens become comparable in affinity), the fraction of effector cells bound to infection-derived antigens decreases, reflecting the impact of self/non-self cross-reactivity and antagonism on recognition. Prior work has explored the impact of antagonistic effects on T-cell response in much more detail [3, 4]. Here, we only adopt a high-level perspective to capture how such effects shape responses to low-immunogenicity infections and cancer.

The relative magnitude of infection- to self-derived antigenicities (i.e, the ratio of the respective recognition thresholds) determine the *immunogenicity* of an infection, which we define as,

$$\text{immunogenicity:} \quad \kappa = K_I^{-1} K_S. \tag{4}$$

This dimensionless measure reflects how readily cognate T-cells recognize antigens presented by infected cells as "non-self" relative to those presented by uninfected cells (self).

Dynamics of growing (acute) infections

In the absence of infection, the population of susceptible (tissue) cells (S) are homeostatically maintained at a steady state, $S(t=0) = S_{\max}$. Upon infection, there are initially a small number of infected cells—$I(t=0) = I_0$—but susceptible cells gradually become infected as they come into contact with other infected cells at rate b_I. The infected state of cells and their death, at rate d_I, produce chemical signals—$\sigma_{\text{Ag}}(t)$ and $\sigma_{\text{inf}}(t)$ respectively—which are detected by the innate immune system. The innate response then triggers naïve cells in adaptive immune response (N) to differentiate into effector cells (E). Differentiated effector cells kill infected and susceptible cells, producing chemical signals ($\sigma_{\text{res}}(t)$) that feedback on the response. The dynamics of the susceptible and infected cells follow,

$$\begin{aligned} \dot{S} &= -b_I I S - d_E C_S, \\ \dot{I} &= b_I S I - d_I I - d_E C_I, \end{aligned} \tag{5}$$

where d_E is the maximal clearance rate of infected/susceptible cells by the effector T-cells, and $C_S = \frac{ES}{E+S+K_S+\frac{E+K_S}{E+K_I}I}$ and $C_I = \frac{EI}{E+I+K_I+\frac{E+K_I}{E+K_S}S}$ represent the number of susceptible or infected cells to be bound to an effector cell at time t, respectively; see section I for derivation of these quantities. This framework closely reflects viral or intracellular bacterial infections, but also captures features of extracellular pathogens that damage or invade cells indirectly (e.g., via toxins); we do not distinguish these mechanisms here.

The extent of pathogen reproduction during the lifetime of an infected cell is determined by the product of the infection rate and the characteristic survival time of the cell. Specifically, over the infected-cell lifetime $\tau_{\text{death}} = d_I^{-1}$, an infected cell produces on average $b_I S \cdot \tau_{\text{death}} = b_I S / d_I$ new infections. In epidemiology, this term is often referred to as the *effective reproduction number*, whose maximum value—attained when $S = S_{\max}$—is the *basic reproduction number*, $b_I S_{\max}/d_I$ [5, 6]. This motivates defining the *intra-host basic reproduction number*,

$$R_0 = \frac{b_I S_{max}}{d_I}, \tag{6}$$

measuring the expected number of secondary infected cells generated by a single infected cell in an otherwise fully susceptible cell population. Note that R_0 scales with $S_{\max}$, the total number of cells in the modeled tissue. For related approaches see refs. [1, 7, 8].

Dynamics of slow-growing tumors

In Fig. 6 we discuss the efficacy of different immune networks for tumor clearance. In this context, we a model a tumor as a slowly growing pathogen—maximum rate b_I—which grows by out-competing healthy tissue cells for limited space/resources, resulting in a logistic growth of tumor cells [9]. We assume that at the time of treatment, the tumor is already at detectable size, $I_0 = 10\%\, S_{\max}$, and that tumor cells die at negligible rates during our simulation window, $d_I = 0$. The resulting dynamics are

$$\begin{aligned} \dot{S} &= -b_I \left(1 - \frac{I}{S_{\max}}\right) I - d_E C_S \\ \dot{I} &= b_I \left(1 - \frac{I}{S_{\max}}\right) I - d_E C_I. \end{aligned} \tag{7}$$

Sampling pathogen space

To model diverse infections, we sample infections across a broad range of immunogenicities (κ), and basic reproduction numbers (R_0). We take $K_S = S_{max} \geq K_I$ so that effector cells typically bind more strongly to infected cells than healthy cells. We note that K_I and K_S are effective parameters reflecting the complex dynamics of antigen detection in the complex tissue environment. The killer T-cell response is delayed with respect to pathogen exposure—typically around a few days [10, 11]—so there are limits on its effectiveness against fast-replicating pathogens. We sample a range of reproduction numbers (R_0) to include non-replicating pathogens $R_0 = 1$ and fast replicating infections which deplete the reservoir of susceptible cells. Quantitatively, we can consider when the time infected cell population peaks in the absence of an immune response $T_{I_{\max}}$ by setting $\mathrm{d}I/\mathrm{d}t = 0$ and $E = 0$ in eq. 5 to find:

$$S(T_{I_{\max}}) = \frac{S_{\max}}{R_0}.$$

In our simulations, we choose the maximum R_0 such that $S(T_{I_{\max}}) \ll S_{\max}$, representing a pathogen that quickly infects most susceptible cells. Here, we chose $(K_I, R_0) \in [10^{-3}S_{\max}, S_{\max}] \times [1, 5]$ to reflect the aforementioned considerations; see Table S1 for details on simulation parameter ranges.

Harm during T-cell response to pathogens

We express harm (damage) to a host in terms of cell deaths. This depends on the details of the programmed immune response, which we parametrize by $\vec{\theta}$. The rate of harm at a given time point can be expressed as

$$h(t) = h_{\inf}(t, \vec{\theta}) + \underbrace{h_{E,I}(t, \vec{\theta}) + h_{E,S}(t, \vec{\theta})}_{=h_{\mathrm{res}}(t)}, \tag{8}$$

where $h_{E,I}(t) = d_E C_I(t)$ and $h_{E,S}(t) = d_E C_s(t)$ denote the loss of infected and healthy cells, respectively, caused by the effector T-cell response, amounting to the total damage by the response $h_{\mathrm{res}}(t)$. Here, $C_I(t)$ and $C_S(t)$ are the expected number of infected and susceptible cells bound to effectors at time t (eq. 3).

- **Harm associated with growing (acute) infections.** The infection harm $h_{\inf}(t)$ captures pathogen-induced harm, i.e., cells directly killed by the infection itself. Its functional form depends on the biological context: For an acute, proliferating infection (eq. 5), it follows:

$$h_{\inf}(t, \vec{\theta}) = d_I I(t, \vec{\theta}) \tag{9}$$

reflecting cell loss due to the death of infected cells. The *total cell death* (or inflicted harm) then follows

$$H(\vec{\theta}) = \underbrace{\int_0^T \left[h_{\inf}\left(t, \vec{\theta}\right) + h_{E,I}\left(t, \vec{\theta}\right)\right] \mathrm{d}t}_{H_{\inf.}(\vec{\theta})} + \underbrace{\int_0^T h_{E,S}\left(t, \vec{\theta}\right) \mathrm{d}t}_{H_{\mathrm{toxic.}}(\vec{\theta})}, \tag{10}$$

where $H_{\inf.}(\vec{\theta})$ aggregates infection-related harm (cells killed by the pathogen and infected cells eliminated by effectors), $H_{\mathrm{toxic.}}(\vec{\theta})$ captures immunopathology (healthy bystander cells killed by effectors), and $T > 0$ is the simulation window. This single, interpretable metric balances clearance benefits against collateral damage, enabling principled comparison across designs and infections.

- **Harm associated with slow-growing tumors.** For logisticaly growing cancer tumors (eq. 7), the harm from cancer takes the form

$$h_{\mathrm{cancer}}(t) = b_I \left(1 - \frac{I(t)}{S_{\max}}\right) I(t), \tag{11}$$

reflecting net healthy cell loss due to tumor growth (eq. 7). In this case, we define the total harm as

$$H(\vec{\theta}) = \underbrace{\int_0^T h_{cancer}\left(t, \vec{\theta}\right) \mathrm{d}t}_{H_{\mathrm{cancer}}(\vec{\theta})} + \underbrace{\int_0^T h_{E,S}\left(t, \vec{\theta}\right) \mathrm{d}t}_{H_{\mathrm{toxic.}}(\vec{\theta})}. \tag{12}$$

Here, $H_{\text{cancer}}(\vec{\theta})$ quantifies the loss of healthy cells as the tumor expands and replaces normal tissue, while $H_{\text{toxic.}}(\vec{\theta})$ captures collateral healthy-cell death caused by the ensuing immune response. Note that in the case of acute infections (eq. 10), loss of healthy cells is represented by $H_{\text{inf.}}$, the sum of cells killed directly by the pathogen and infected cells eliminated by effectors. In cancer, by contrast, a slowly growing tumor replaces healthy cells with malignant cells. Accordingly, harm is tied to the "birth" of cancer cells captured by the time-integral of the birth term $h_{\text{cancer}}(t)$ in eq. 11.

Signal cues for T-cell response

We assume that T-cells can sense three types of signals during infection:

- **Antigen signal** $\sigma_{\text{Ag}}(t)$, which is proportional to the number of infected cells, and expressed in units of the specificity constant K_I

$$\text{antigen signal:}\ \ \sigma_{\text{Ag}}(t) = {K_I}^{-1} I(t) \leq K_I^{-1} S_{\max} \tag{13}$$

- **Infection-induced harm signal** $\sigma_{\text{inf}}(t)$, reflecting cumulative damage caused directly by the pathogen, and expressed in units of the specificity constant K_H

$$\text{infection harm signal:}\ \ \sigma_{\text{inf}}(t) = K_H^{-1} \int_0^t e^{-(t-s)d_H}\, h_{\text{inf}}(s)\, \mathrm{d}s \leq K_H^{-1} S_{\max}. \tag{14}$$

- **Response-induced harm signal** $\sigma_{\text{res}}(t)$, reflecting cumulative damage caused by the effector response, and expressed in units of the signal recognition threshold K_H

$$\text{response harm signal:}\ \ \sigma_{\text{res}}(t) = K_H^{-1} \int_0^t e^{-(t-s)d_H}\, h_{\text{res}}(s)\, \mathrm{d}s \leq K_H^{-1} S_{\max}. \tag{15}$$

The harm-derived signals are time-integrated, reflecting the accumulation of cytokines during an immune response, and are exponentially discounted at rate d_H, reflecting the natural decay of cytokines in the absence of ongoing stimulation. Note that K_I plays a dual role: it represents both the recognition threshold of effectors binding to and killing infected cells (see section I) and the recognition threshold for binding to and stimulating T-cells.

Lineage-based model of CD8$^+$ T-cell response

To characterize the immune response to an infection, we model the large-population dynamics of susceptible and infected cells in a host as deterministic processes, coupled to a stochastic, time-inhomogeneous birth-death process for the CD8$^+$ T-cell response.

The T-cell repertoire typically contains many precursors, and responses to a given pathogen are often polyclonal. An interesting problem would be to assess how interactions between different T-cell clones impact the response in an individual, which we defer to future work. In our simulations, we model a response mediated by a single clone of naïve T-cells (i.e., cells sharing the same TCR). Upon activation, each naïve cell can seed a lineage whose progeny differentiate into effector or memory cells that act against infection or tumor.

In what follows, we consider an initial population of $N_{\text{lin.}}$ monoclonal naïve T-cells, and track the lineages of their progeny in response to a pathogenic threat. Each cell's decisions—to engage an antigen-presenting cell (APC), become activated, divide, differentiate, or die—are governed by time-dependent antigen and infection harm signals (eqs. 19). We assume these signals are lineage-independent: cells in the same state but belonging to different lineages share identical transition rates. Inter-lineage variability therefore arises solely from stochastic differences in transition timing, consistent with experimental observations [8, 12, 13].

a. Recruitment of naïve T-cells to interact with APCs ($N \to [APC{\cdot}N]$). Naïve T-cells recirculate through blood and lymph and may also reside in tissues [14]. Upon infection, APCs become activated and acquire antigen as infected cells die [15]. This process is coupled to the production of an infection harm signal σ_{inf}, which serves as a proxy for the "danger" associated with tissue damage caused by the pathogen. As a result, naïve T-cells can be recruited to

the sites of infection and interact with APCs to initiate an immune response. We model the recruitment rate of naïve T-cells to bind to the APCs as

$$r_{N\to \mathrm{APC}\cdot N}(t) \equiv r_{\mathrm{bind}}(t) = \frac{r_{\mathrm{bind}}^{\max}}{1+\exp\left[-\log(1+\sigma_{\mathrm{inf}}(t))\right]}, \tag{16}$$

where $r_{\mathrm{bind}}^{\max} \approx 1\ [\mathrm{day}]^{-1}$ is the upper bound on the recruitment rate of T-cells to lymph nodes for sustained APC binding [16–18] (see Table S1 for the biological range). The denominator modulates this rate according to the strength of the infection harm signal activating the APCs. We assume that recruitment of all the naïve cells to APCs is mediated by the same infection harm signal, and that these cells do not compete for APCs or infection harm signal.

b. Activation of naïve T-cells ($[APC \cdot N] \to N^*$). When a naïve T-cell engages an APC to form a complex $[\mathrm{APC}\cdot N]$, the two cells may remain bound for several hours, leading to activation of the bound T-cell ("priming phase"). During this time, the immunological synapse formed between the T-cell and APC initiates a cascade of intracellular signaling events in the naïve T-cell, enabling it to surpass a sharp activation threshold [19]. The strength of antigenic stimulation, together with co-stimulatory signals, modulates the activation rate of APC-bound naïve T-cells $r_{N\to N^*}(t)$.

Experiments have shown that naïve T-cells can remain stably bound for 12-20 hours before the first division [13] to become activated $[\mathrm{APC}\cdot N] \to N^*$. Following activation, cells commit to a rapid proliferative program [20, 21], resulting in $k = 2-3$ burst-like divisions at a rate $r_{N^*\to 2N^*}$ [13] (Fig. S1). Once initiated, these early doublings proceed largely independently of continued antigen or cytokine input [13, 22], and set an upper bound $N^*_{\max} = 2^k$ for the number of activated cells generated prior to differentiation in a lineage; we use $N^*_{\max} = 2^2$ in our simulations (Table S1).

c. Differentiation of activated T-cells to effector phenotype ($N^* \to E$). During an infection, $\mathrm{CD8}^+$ T-cells follow a pathway from long-lived naïve/memory cells to short-lived effector cells. This fate is partially "reversible": effectors can reacquire memory-like traits (e.g., longevity), and become so-called effector memory [23–25]. We consider differentiation from naïve to effector state as the primary transition for activated cells, with un-transitioned cells retaining the relative longevity and slow proliferation of naïve cells. Consistent with prior work [13, 23, 25, 26], we take the burst-like early proliferation phase as the window during which effector commitment occurs (Fig. S1).

Formally, for a given activated cell N^*, we model effector commitment as the first arrival of a time-inhomogeneous Poisson process with instantaneous rate, $r_{N^*\to E}(t)$. Its probability of becoming an effector cell is computed as the probability that the first arrival time is less than the stochastic time interval τ_{burst} to complete the initial burst-like divisions,

$$p_{N^*\to E} = 1 - \exp\left[-\int_0^{\tau_{\mathrm{burst}}} r_{N^*\to E}(s)\mathrm{d}s\right]. \tag{17}$$

Given N^* activated cells available at the end of the burst, the number that differentiate into effectors is drawn independently as $e_{\mathrm{differentiate}} \sim \mathrm{Binom}(N^*, p_{N^*\to E})$, with the remaining $N^* - e_{\mathrm{differentiate}}$ cells adopting a memory fate. This fate decision is executed independently across lineages (Algorithms S1-S3; Fig. S1).

d. Division and death of effector T-cells ($E \to 2E$, and $E \to \emptyset$). Experiments have shown that T-cell lineages do not divide indefinitely [27]. By tracking proliferation markers in cell lineages, it has been estimated that a naïve precursor undergoes at most $K_{\max} \approx 10-19$ division rounds before terminal arrest [12, 26, 28, 29], with differentiation into effector possible along the way [12, 26, 27]. This sets an upper bound on the number of effector cells in a lineage (if all cells turn into effector), which is $E_{\max} = 2^{K_{\max}}$, which we set to 2^{15} in our simulations (Table S1). A comparable bound for memory cells is less certain, as establishing it would require years of longitudinal tracking.

Because effector commitment is stochastic, the maximum additional effector output beyond the initial committed pool varies by lineage. At the end of the early burst, the lineage has completed $\log_2 N^*_{\max}$ semi-synchronous divisions. Each newly committed effector cell can therefore undergo at most $K_{\max} - \log_2 N^*_{\max} = \log_2 E_{\max} - \log_2 N^*_{\max}$ additional rounds. Thus, starting from $e_{\mathrm{differentiate}}$ effectors at the end of the burst phase, proliferation halts once the effector count reaches, $e_{\text{stop prolif.}} = e_{\mathrm{differentiate}} \times \left(\frac{E_{\max}}{N^*_{\max}}\right)$; see Algorithms S1-S3. Each differentiated effector cell undergoes division and death events at rates $r_{E\to E}(t)$ and $r_{E\to\emptyset}(t)$, respectively, which are modulated by the time-varying antigen and harm signals (see below). In lieu of tracking the division history/tree of cells in the lineage, as effector cells divide, we keep a running count of the total number of effector division events in that lineage, $B_{\mathrm{proliferation}}$, and set the division rate to zero once $B_{\mathrm{proliferation}} \geq e_{\text{stop prolif.}}$ (Algorithms S1-S3).

In summary, the dynamics for a single responding T-cell from exposure to an antigen, activation by engaging APC's, differentiation, expansion, and contraction are summarized as follows,

$$
\begin{aligned}
\text{exposure} &\begin{cases} S + I & \xrightarrow{b_I} I + I \to \sigma_{\text{Ag}} \\ I & \xrightarrow{d_I} \emptyset \to \sigma_{\text{inf}} \end{cases} \\
\text{clearance} &\begin{cases} I/S + E & \xrightarrow{d_E} E + \emptyset \to E + \sigma_{\text{res}} \end{cases} \\
\text{activation} &\begin{cases} \text{APC} + N & \xrightarrow{r_{\text{bind}}(\sigma_{\text{Ag}},\sigma_{\text{inf}},\sigma_{\text{res}})} \text{APC} \cdot N \\ \text{APC} \cdot N & \xrightarrow{r_{N\to N^*}(\sigma_{\text{Ag}},\sigma_{\text{inf}},\sigma_{\text{res}})} N^* + \text{APC} \\ N^* & \xrightarrow{r_{N^*\to 2N^*}} 2N^* \xrightarrow{r_{N^*\to 2N^*}} 4N^* \end{cases} \\
\text{differentiation} &\begin{cases} N^* & \xrightarrow{r_{N^*\to E}(\sigma_{\text{Ag}},\sigma_{\text{inf}},\sigma_{\text{res}})} E \end{cases} \\
\text{expansion} &\begin{cases} E & \xrightarrow{r_{E\to 2E}(\sigma_{\text{Ag}},\sigma_{\text{inf}},\sigma_{\text{res}})} E + E \end{cases} \\
\text{contraction} &\begin{cases} E & \xrightarrow{r_{E\to \emptyset}(\sigma_{\text{Ag}},\sigma_{\text{inf}},\sigma_{\text{res}})} \emptyset \end{cases}
\end{aligned}
\tag{18}
$$

Note that the transition rates are dependent on the three time-dependent signals $(\sigma_{\text{Ag}}, \sigma_{\text{inf}}, \sigma_{\text{res}})$; see below for details. The programmatic details of these reactions, when keeping track of different T-cell lineages, are provided in Algorithms S1-S3.

e. Regulation of T-cell transition rates. Activated T-cells undergo differentiation, proliferation, and eventual contraction in response to antigen and and other signaling molecules like cytokines [30–32]. These processes are shaped by both extrinsic noise (e.g., variability in cytokine levels and spatial heterogeneity of APCs) and intrinsic noise (e.g., stochastic gene expression and chromatin remodeling) [33]. We model the effector CD8$^+$ T-cell response as a sequence of four canonical modules:

1. **Activation:** priming of naïve T-cells, $N \to N^*$,
2. **Differentiation:** commitment of the activated T-cells to effector phenotype, $N^* \to E$,
3. **Expansion:** division of effector cells, $E \to 2E$,
4. **Contraction:** programmed loss of effector cells, $E \to \emptyset$.

We model the complex genetic regulation of each transition i as an inhomogeneous Poisson process with a transition rate $r_i(t)$, modulated by the signals that T-cells receive at a given time,

$$
r_i(t) = r^{\max}\, g_i\left(\sigma_{\text{Ag}}, \sigma_{\text{inf}}, \sigma_{\text{res}}; \vec{\psi}_i, \ell_i\right), \tag{19}
$$

where the maximal rate $r^{\max}$ is bounded by the cell-cycle time ($\sim 5-6$ hrs [22, 34, 35]), and g is a monotonic *response function* chosen from the modified Monod-Wyman-Changeux (Hill-type) family [36, 37]:

$$
\begin{aligned}
g_i(\sigma_{\text{Ag}}, \sigma_{\text{inf}}, \sigma_{\text{res}}; \vec{\psi}_i, \ell_i) &= \frac{1}{1 + \exp\left[-L_i(\sigma_{\text{Ag}}, \sigma_{\text{inf}}, \sigma_{\text{res}}; \vec{\psi}_i, \ell_i)\right]}, \\
&\text{with} \\
L_i(\sigma_{\text{Ag}}, \sigma_{\text{inf}}, \sigma_{\text{res}}; \vec{\psi}_i, \ell_i) &= \ell_i + \sum_{\alpha \in \{\text{Ag,inf,res}\}} \psi_i^{\alpha} \log\left[1 + \sigma_\alpha\right].
\end{aligned}
\tag{20}
$$

The parameter ℓ_i sets the baseline propensity of transition i: in the absence of external signals ($\sigma = 0$), the regulatory factor evaluates to $g_{0,i} = (1 + \exp(-\ell_i))^{-1}$. As ℓ_i increases, this baseline transition rate approaches its ceiling $r^{\max}$. The weights $\{\psi_i^{Ag}, \psi_i^{\text{inf}}, \psi_i^{\text{res}}\}$ encode the sensitivity and polarity (positive / negative = up /down-regulation) of response to the antigen (σ_{Ag}) and harm-induced signals (σ_{inf}, and σ_{res}). Each logarithmic term $\log[1 + \sigma_\alpha]$ becomes significant only once its corresponding cue $\sigma_\alpha \sim 1$. This formulation abstracts away many layers of biochemical detail, while implementing the fold-change response to stimuli observed in many biological circuits [38–42].

The transition rates governing each of the four modules are parameterized by a response function (eq. 20), whose parameters we collect into a *design vector* $\vec{\theta}$ (up to 4×4 degrees of freedom). These functions are expressive enough to include many three-signal logic circuits [37]. Introducing explicit non-linear interactions between signals would broaden the expressivity of the model but at the cost of greater complexity and limited interpretability, an extension that we leave for future work.

II. NUMERICAL IMPLEMENTATION OF STOCHASTIC T-CELL RESPONSE

Parameters governing our model and their biological ranges are given in Table S1, and the pseudo-code describing our numerical approach is provided in Algorithms S1-S3. Simulating our lineage-based model of T-cell response with a fully agent-based Gillespie algorithm with time-dependent rates is computationally challenging given the large number of agents ($\sim 10^7$) that must be tracked and updated [43]. We employ two approximations. First, we model the dynamics of the susceptible and infected cells, and harm signals using a piecewise, deterministic-stochastic system of ODEs, implemented numerically with first-order difference equations, and specified in Algorithms S1-S3. Second, we model the cellular immune response as a time-inhomogeneous Markov process implemented with a tau-leaping scheme [44].

As long as the simulation time-step Δt is chosen carefully, this approach provides a good compromise between computational speed and accuracy. In our case, it is sufficient to choose Δt such that transition rates do not change too much during this small time interval, i.e, $r(t+\Delta t) \approx r(t)$. Thus, we seek to bound

$$\frac{\mathrm{d}r}{\mathrm{d}t} = r^{\max} \sum_{\sigma \in \{\mathrm{Ag,inf,res}\}} \frac{\mathrm{d}g}{\mathrm{d}\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}t}, \tag{21}$$

where

$$\frac{\mathrm{d}g}{\mathrm{d}\sigma} = \frac{\psi_\sigma}{1+\sigma}\frac{1}{1+e^{-L(\sigma)}}\frac{e^{-L(\sigma)}}{1+e^{-L(\sigma)}} \quad \Longrightarrow \quad \left|\frac{\mathrm{d}g}{\mathrm{d}\sigma}\right| \leq \frac{|\psi_\sigma|}{1+\sigma}\frac{1}{2^2}. \tag{22}$$

The upper bound is observed by noting that the product of the last two terms is equivalent to $x(1-x)$ for $x \in (0,1)$. Next we compute

$$\begin{aligned}
\frac{\mathrm{d}\sigma_{\mathrm{Ag}}}{\mathrm{d}t} &= K_I^{-1}\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma_{\mathrm{Ag}} \times \left(b_I S - d_I - d_E C_I/I\right), \\
\frac{\mathrm{d}\sigma_{\mathrm{inf}}}{\mathrm{d}t} &= K_H^{-1} h_{\mathrm{inf}} - d_H \sigma_{\mathrm{inf}} = K_H^{-1} d_I I - d_H \sigma_{\mathrm{inf}}, \\
\frac{\mathrm{d}\sigma_{\mathrm{res}}}{\mathrm{d}t} &= K_H^{-1} h_{\mathrm{res}} - d_H \sigma_{\mathrm{res}} = K_H^{-1} d_E C_S - d_H \sigma_{\mathrm{res}},
\end{aligned} \tag{23}$$

from which we conclude

$$\begin{aligned}
\left|\frac{\mathrm{d}\sigma_{\mathrm{Ag}}}{\mathrm{d}t}\right| &\leq \sigma_{\mathrm{Ag}} \times (b_I S + d_I + d_E C_I/I), \\
\left|\frac{\mathrm{d}\sigma_{\mathrm{inf}}}{\mathrm{d}t}\right| &\leq K_H^{-1} d_I I(t) + d_H \sigma_{\mathrm{inf}} \sim 2 d_H \sigma_{\mathrm{inf}}, \\
\left|\frac{\mathrm{d}\sigma_{\mathrm{res}}}{\mathrm{d}t}\right| &\leq K_H^{-1} d_E C_S(t) + d_H \sigma_{\mathrm{res}} \sim 2 d_H \sigma_{\mathrm{res}}
\end{aligned} \tag{24}$$

where we rely on the homeostatic forms of the dynamical equations for σ_{inf} and σ_{res} to argue that typically, $K_H^{-1} d_I I(t) \sim d_H \sigma_{\mathrm{inf}}(t)$ and $K_H^{-1} d_E C_S(t) \sim d_H \sigma_{\mathrm{res}}(t)$. Thus, we see that typically,

$$\begin{aligned}
\left|\frac{\mathrm{d}r}{\mathrm{d}t}\right| &\leq r^{\max}\frac{1}{4}\left(\frac{\sigma_{\mathrm{Ag}}}{1+\sigma_{\mathrm{Ag}}}|\psi_{\mathrm{Ag}}|(b_I S + d_I + d_E C_I/I) + \frac{\sigma_{\mathrm{inf}}}{1+\sigma_{\mathrm{inf}}} 2 d_H |\psi_{\mathrm{inf}}| + \frac{\sigma_{\mathrm{res}}}{1+\sigma_{\mathrm{res}}} 2 d_H |\psi_{\mathrm{res}}|\right) \\
&\leq r^{\max}\frac{1}{4}\left(|\psi_{\mathrm{Ag}}|(b_I S + d_I + d_E C_I/I) + 2 d_H |\psi_{\mathrm{inf}}| + 2 d_H |\psi_{\mathrm{res}}|\right).
\end{aligned} \tag{25}$$

With parameters chosen according to Table S1, each term in eq. 25 is at most on the order of $\mathcal{O}(10)$. Thus, taking $\Delta t = \mathcal{O}(10^{-2})$ guarantees that transition rates typically change only minimally over each time interval,

$$|r(t+\Delta t) - r(t)| \approx \left|\frac{\mathrm{d}r}{\mathrm{d}t}\right| \Delta t < \mathcal{O}(10^{-1}). \tag{26}$$

AlgorithmS1 Piecewise, deterministic-stochastic simulation of CD8$^+$ T-cell response to an infection.
Simulating the trajectories of susceptible and infected cells, and T-cells.

Input: $\vec{\psi}, \vec{\ell}, b_I, K_I, K_S, K_H, r^{\max}, r^{\max}_{\text{bind}}, r_{N^*\to 2N^*}, d_I, d_E, N_{\text{lin.}}, N^*_{\max}, E_{\max}, \Delta t, T$
Initialize: $S \leftarrow S_{\max} - I_0,\ I \leftarrow I_0,\ \vec{N} \leftarrow \vec{1},\ \vec{\sigma} \leftarrow \vec{0},\ \vec{E} \leftarrow \vec{0},\ \vec{e}_{\text{stop prolif.}} \leftarrow 0; \vec{B}_{\text{differentiate}} \leftarrow 0,\ \vec{B}_{\text{proliferate}} \leftarrow 0$

for $t = 0, \ldots, T-1$ **step** Δt **do**

▷ *(i) deterministic update of infection dynamics, using* INFECTION (Algorithm S2)

$(\vec{r}, h_{\text{inf}}, h_{E,I}, h_{E,S}) \leftarrow \textsc{Infection}(S, I, \vec{E}, \vec{\sigma};\ I_0, N_{\text{lin.}}, \vec{\psi}, \vec{\ell}, b_I, K_I, K_S, K_H, r^{\max}, r^{\max}_{\text{bind}}, r_{N^*\to 2N^*}, d_I, d_E, \Delta t)$
▷ —- shorthand notation: $\vec{r} = \{r_{\text{bind}}, r_{N^*\to 2N^*}, r_{N\to N^*}, r_{N^*\to E}, r_{E\to 2E}, r_{E\to\emptyset}\}$—-

▷ *(ii) stochastic T-cell response for all lineages $l = 1, \ldots, N_{lin.}$, using* RESPONSE (Algorithm S3)

for $l = 1, \ldots, N_{\text{lin.}}$ **do**

$(N^{(l)}, [\text{APC}\cdot N]^{(l)}, N^{*,(l)}, E^{(l)}, M^{(l)}, e^{(l)}_{\text{stop prolif.}}, B^{(l)}_{\text{differentiate}}, B^{(l)}_{\text{proliferate}})$
$\leftarrow \textsc{Response}(N, N^*, [APC\cdot N], E, M, \vec{r}, N^*_{\max}, E_{\max}, e_{\text{stop prolif.}}, B_{\text{differentiate}}, B_{\text{proliferate}};\ \Delta t)$

end for

▷ *(iii) record trajectories of observables at time t*

Record $S, I, \vec{N}, \vec{N}^*, \vec{E}, \vec{M}, h_{\text{inf}}, h_{E,I}, h_{E,S}$ into $[\,\cdot\,]_t$

end for

return $([S]_t, [I]_t, [\vec{N}]_t, [\vec{N}^*]_t, [\vec{E}]_t, [\vec{M}]_t, [h_{\text{inf}}]_t, [h_{E,I}]_t, [h_{E,S}]_t)$

AlgorithmS2 Deterministic update of infection dynamics. Infection variables (S, I), harms $(h_{\text{inf}},\ h_{E,I},\ h_{E,S})$, signals $(\sigma_{\text{Ag}}, \sigma_{\text{inf}}, \sigma_{\text{res}})$, and transition rates $(r_{\text{bind}}, r_{N^*\to 2N^*},\ r_{N\to N^*},\ r_{N^*\to E},\ r_{E\to 2E},\ r_{E\to\emptyset})$ are computed.

function INFECTION$(S, I, \vec{E}, \vec{\sigma};\ I_0, N_{\text{lin.}}, \vec{\psi}, \vec{\ell},\ b_I, K_I, K_S, K_H,\ r^{\max}, r^{\max}_{\text{bind}}, r_{N^*\to 2N^*},\ d_I, d_E, d_H,\ \Delta t)$

▷ *(i) compute the number of T-cell bound infected and susceptible cells C_I and C_S*

$$C_I \leftarrow \frac{I \sum_{l=1}^{N_{\text{lin.}}} E^{(l)}}{\sum_{l=1}^{N_{\text{lin.}}} E^{(l)} + I + K_I + \frac{\sum_{l=1}^{N_{\text{lin.}}} E^{(l)} + K_I}{\sum_{l=1}^{N_{\text{lin.}}} E^{(l)} + K_S} S} \quad , \quad C_S \leftarrow \frac{S \sum_{l=1}^{N_{\text{lin.}}} E^{(l)}}{\sum_{l=1}^{N_{\text{lin.}}} E^{(l)} + S + K_S + \frac{\sum_{l=1}^{N_{\text{lin.}}} E^{(l)} + K_S}{\sum_{l=1}^{N_{\text{lin.}}} E^{(l)} + K_I} I}.$$

▷ *(ii) compute the infection harm h_{inf}, and response harms $h_{E,I}$, $h_{E,S}$*

$$h_{\text{inf}} \leftarrow d_I I\, \Delta t, \qquad h_{E,I} \leftarrow d_E C_I\, \Delta t, \qquad h_{E,S} \leftarrow d_E C_S\, \Delta t$$

▷ *(iii) update the number of susceptible and infected cells*

$$S \leftarrow S + (-\mathbf{1}_{I\geq I(0)} b_I I S - d_E C_S)\Delta t$$
$$I \leftarrow I + (\mathbf{1}_{I\geq I(0)} b_I I S - d_I I - d_E C_I)\Delta t$$

▷ *(iv) update the three signals $\vec{\sigma}$*

$$\sigma_{\text{Ag}} \leftarrow K_I^{-1} I$$
$$\sigma_{\text{inf}} \leftarrow \sigma_{\text{inf}} + (K_H^{-1} h_{\text{inf}} - d_H \sigma_{\text{inf}})\Delta t$$
$$\sigma_{\text{res}} \leftarrow \sigma_{\text{res}} + (K_H^{-1} h_{E,I} + K_H^{-1} h_{E,S} - d_H \sigma_{\text{res}})\Delta t$$

▷ *(v) update the recruitment rate of naïve cells to bind to APC's r_{bind} and the transition rates r_x for all the processes $x \in \{N \to N^*, N^* \to E, E \to 2E, E \to \emptyset\}$}*

$$r_{\text{bind}} \leftarrow \frac{r^{\max}_{\text{bind}}}{1 + \exp[-\log(1 + \sigma_{\text{inf}})]}$$

for $x \in \{N \to N^*, N^* \to E, E \to 2E, E \to \emptyset\}$ **do**

$$L_x \leftarrow \ell_x + \sum_{\alpha:\{\text{Ag,inf,res}\}} \psi_\alpha \log(1 + \sigma_\alpha), \quad g_x \leftarrow (1 + \exp[-L_x])^{-1}, \quad r_x \leftarrow r^{\max} g_x$$

end for

return $\vec{r} = \{r_{\text{bind}}, r_{N^*\to 2N^*},\ r_{N\to N^*},\ r_{N^*\to E},\ r_{E\to 2E},\ r_{E\to\emptyset}\},\ h_{\text{inf}},\ h_{E,I},\ h_{E,S},\ S,\ I$

end function

AlgorithmS3 Stochastic response of a T-cell lineage. Stochastic recruitment, activation, proliferation, differentiation, and death of CD8$^+$ T-cells are simulated for each lineage (lineage superscripts dropped). The number of cells from different states $(N, [\text{APC} \cdot N], N^*, E, M)$, and the variables determining the lineage's fate $(e_{\text{stop prolif.}}, B_{\text{differentiate}}, B_{\text{proliferate}})$ are evaluated.

function RESPONSE$(N, N^*, [APC \cdot N], E, M, \vec{r}, N^*_{\max}, E_{\max}, e_{\text{stop prolif.}}, B_{\text{differentiate}}, B_{\text{proliferate}};\ \Delta t)$
▷ —- shorthand notation: $\vec{r} = \{r_{\text{bind}}, r_{N\to N^*}, r_{N^*\to 2N^*}, r_{N^*\to E}, r_{E\to 2E}, r_{E\to\emptyset}\}$—-

Initialize: $(n_{\text{bound}}, n_{\text{activated}}, n^*_{\text{proliferate}},\ n^*_{\text{differentiate}},\ e_{\text{differentiate}},\ m_{\text{differentiate}},\ e_{\text{contract}}) \leftarrow 0$

▷ *(i) recruitment of naïve cells to interact with APC for activation*

if $N > 0$ **then**
$n_{\text{bound}} \leftarrow \text{Binom}(N,\ r_{\text{bind}}\Delta t)$ ▷ —- binomial sampling of naïve cells to bind to APC, rate r_{bind}—-
$n_{\text{activated}} \leftarrow \text{Binom}([\text{APC} \cdot N],\ r_{N\to N^*}\Delta t)$ ▷ —- binomial sampling to activate APC-bound naïves, rate $r_{N\to N^*}$—-
end if

▷ *(ii) burst-like proliferation of activated naïves N^*, differentiation to effectors E, or memory M*

if $0 < N^* < N^*_{\max}$ **then**
$n^*_{\text{proliferate}} \leftarrow \text{Binom}(N^*,\ r_{N^*\to 2N^*}\Delta t)$ ▷ —- binomial sampling to replicate N^*, rate $r_{N^*\to 2N^*}$ (initial burst)—-
$B_{\text{differentiate}} \leftarrow B_{\text{differentiate}} + r_{N^*\to E}\Delta t$ ▷ —- updating likelihood of N^* differentiating to E by time t —-
else if $N^* \geq N^*_{\max}$ **then**
$n^*_{\text{differentiate}} \leftarrow N^*$ ▷ —- the number of activated naïves that can differentiate—-
$e_{\text{differentiate}} \leftarrow \text{Binom}\left(N^*,\ 1 - \mathrm{e}^{-B_{\text{differentiate}}}\right)$ ▷ —- binomial sampling to differentiate N^* to E during burst—-
$m_{\text{differentiate}} \leftarrow N^* - e_{\text{differentiate}}$ ▷ —- number of activated cells differentiating to memory—-
$e_{\text{stop prolif.}} \leftarrow \min(e_{\text{differentiate}}, N^*_{\max}) \times \frac{E_{\max}}{N^*_{\max}}$ ▷ —- maximum number of allowed effectors in the lineage—-
end if

▷ *(iii) contraction and proliferation of effectors with rates $r_{E\to\emptyset}$, and $r_{E\to 2E}$, respectively*

$e_{\text{contract}} \leftarrow \text{Binom}(E,\ r_{E\to\emptyset}\Delta t)$ ▷ —- binomial sampling to remove E—-
if $B_{\text{proliferate}} < e_{\text{stop prolif.}}$ **then**
$e_{\text{proliferate}} \leftarrow \text{Binom}(E,\ r_{E\to 2E}\Delta t)$ ▷ —- binomial sampling to replicate E—-
$B_{\text{proliferate}} \leftarrow B_{\text{proliferate}} + e_{\text{proliferate}}$ ▷ —- running count of the number of effector cell divisions—-
end if

▷ *(iv) update the populations of naïve N, naïve bound to APC $[APC \cdot N]$, activated N^*, effector E, and memory M cells for a given T-cell lineage responding to infection:*

$N \leftarrow N - n_{\text{bound}}$
$[\text{APC} \cdot N] \leftarrow [\text{APC} \cdot N] + n_{\text{bound}} - n_{\text{activated}}$
$N^* \leftarrow N^* + n_{\text{activated}} + n^*_{\text{proliferate}} - n^*_{\text{differentiate}}$
$E \leftarrow E + e_{\text{proliferate}} - e_{\text{contract}} + e_{\text{differentiate}}$
$M \leftarrow M + m_{\text{differentiate}}$

return $N, [\text{APC} \cdot N], N^*, E, M, e_{\text{stop prolif.}}, B_{\text{differentiate}}, B_{\text{proliferate}}$
end function

III. OPTIMIZATION OF CD8$^+$ T-CELL RESPONSE TO PATHOGENS

Design architecture of T-cell programs for infection response

Infection clearance results from coordination of the effector trajectory $E(t)$ in response to an infection $I(t)$. This corresponds to controlling/regulating the rates, $\vec{r}(t)$ of cell state transitions, divisions and deaths, potentially varying $4 \times 4 = 16$ module-specific sensitivity and baseline parameters,

$$\begin{aligned} \text{activation:}\ & \psi^{\sigma_{\rm Ag}}_{N\to N^*},\ \psi^{\sigma_{\rm inf}}_{N\to N^*},\ \psi^{\sigma_{\rm res}}_{N\to N^*},\ \ell_{N\to N^*} \in \mathbb{R}, \\ \text{differentiation:}\ & \psi^{\sigma_{\rm Ag}}_{N^*\to E},\ \psi^{\sigma_{\rm inf}}_{N^*\to E},\ \psi^{\sigma_{\rm res}}_{N^*\to E},\ \ell_{N^*\to E} \in \mathbb{R}, \\ \text{proliferation:}\ & \psi^{\sigma_{\rm Ag}}_{E\to 2E},\ \psi^{\sigma_{\rm inf}}_{E\to 2E},\ \psi^{\sigma_{\rm res}}_{E\to 2E},\ \ell_{E\to 2E} \in \mathbb{R}, \\ \text{contraction:}\ & \psi^{\sigma_{\rm Ag}}_{E\to \emptyset},\ \psi^{\sigma_{\rm inf}}_{E\to \emptyset},\ \psi^{\sigma_{\rm res}}_{E\to \emptyset},\ \ell_{E\to \emptyset} \in \mathbb{R}, \end{aligned} \tag{27}$$

to generate effector response trajectories.

Accounting for shared regulatory programs and controlling the combinatorial complexity of our model parameter space, we impose two simplifications. First, activation and proliferation are both MYC-driven [45]; we therefore constrain the modulatory factors, $g_{N\to N^*}$ and $g_{E\to 2E}$, to share parameter values. Note, however, that their realized rates will differ because signals vary in time, as cells transition between naive and effector states. Second, we make the assumption that extracellular cues converge on coherent effector programs, and thus enforce a common set of signal sensitivities across modules. With the sign convention that positive sensitivities promote effector production, for each signal σ, we set

$$\psi_\sigma \equiv \psi^{\sigma}_{N\to N^*} = \psi^{\sigma}_{E\to 2E} = \psi^{\sigma}_{N^*\to E} = -\psi^{\sigma}_{E\to \emptyset}.$$

Each module i nonetheless retains its own baseline parameter ℓ_i, or equivalently the baseline rate $g_{0;i} = (1 + \exp(-\ell))^{-1}$. With this simplification we reduce the number of degrees of freedom in our model from 16 to a total of six sensitivity and baseline rates,

$$\begin{aligned} \text{activation:}\ & \psi^{\sigma_{\rm Ag}}_{N\to N^*} = \psi_{\sigma_{\rm Ag}},\ \psi^{\sigma_{\rm inf}}_{N\to N^*} = \psi_{\sigma_{\rm inf}},\ \psi^{\sigma_{\rm res}}_{N\to N^*} = \psi_{\sigma_{\rm res}},\ \ell_{N\to N^*} \in \mathbb{R}, \\ \text{differentiation:}\ & \psi^{\sigma_{\rm Ag}}_{N^*\to E} = \psi_{\sigma_{\rm Ag}},\ \psi^{\sigma_{\rm inf}}_{N^*\to E} = \psi_{\sigma_{\rm inf}},\ \psi^{\sigma_{\rm res}}_{N^*\to E} = \psi_{\sigma_{\rm res}},\ \ell_{N^*\to E} \in \mathbb{R}, \\ \text{proliferation:}\ & \psi^{\sigma_{\rm Ag}}_{E\to 2E} = \psi_{\sigma_{\rm Ag}},\ \psi^{\sigma_{\rm inf}}_{E\to 2E} = \psi_{\sigma_{\rm inf}},\ \psi^{\sigma_{\rm res}}_{E\to 2E} = \psi_{\sigma_{\rm res}},\ \ell_{E\to 2E} = \ell_{N\to N^*} \in \mathbb{R}, \\ \text{contraction:}\ & \psi^{\sigma_{\rm Ag}}_{E\to \emptyset} = -\psi_{\sigma_{\rm Ag}},\ \psi^{\sigma_{\rm inf}}_{E\to \emptyset} = -\psi_{\sigma_{\rm inf}},\ \psi^{\sigma_{\rm res}}_{E\to \emptyset} = -\psi_{\sigma_{\rm res}},\ \ell_{E\to \emptyset} \in \mathbb{R}. \end{aligned} \tag{28}$$

Note the relationship, $\psi^{\sigma}_{E\to\emptyset} = -\psi_\sigma$ reflects coordinated signal-regulation of effector-promoting and effector-suppressing transitions.

The parameters of these constrained response functions define a *response design* $\vec{\theta}$ as

$$\vec{\theta} = (\underbrace{\psi_{\rm Ag}, \psi_{\rm inf}, \psi_{\rm res}}_{\text{signal feedback}}, \underbrace{\ell_{N\to N^*} = \ell_{E\to 2E}, \ell_{N^*\to E}, \ell_{E\to\emptyset}}_{\text{module-specific baselines}}). \tag{29}$$

For each transition i, we choose the baseline response parameter, ℓ_i, such that $|\ell_i| \le \ell^{\max}$ with

$$r_i(\ell^{\max}) \sim r^{\max},\ r_i(-\ell^{\max}) \sim T^{-1}_{\rm sim},$$

This range is set to be broad enough that at the lower baseline rate ($-\ell^{\max}$) transitions are seldom observed in our simulation time window, whereas at the upper baseline rate ($\ell^{\max}$) they occur in nearly every cell cycle. Here, we take $\ell^{\max} = 3.0$, so $r_i(\ell^{\max} = 3) \approx 0.95 \times r^{\max}$. For the signal-sensitivity parameter ψ_σ, we require that under weak stimulation ($\sigma = 1$) the signal term in the response function (eq. 20) is below the threshold of response, set by the maximal baseline term, i.e., $\psi^{\max}_\sigma \cdot \log(1+1) < \ell^{\max}$, and thus take $\psi^{\max}_\sigma = 3.0$.

Optimizing the performance of response designs

a. Optimal design for a given infection scenario. We assume that the performance of a design $\vec{\theta}$ is negatively related to the pathogen-associated harm, when that design is used in response. For an infection characterized by immunogenicity κ and basic reproduction number R_0, we define the optimal design as,

$$\vec{\theta}^{\min}(\kappa, R_0) = \arg\min_{\vec{\theta}} H(\vec{\theta}). \tag{30}$$

where the harm $H(\vec{\theta})$ is computed by eq. 10 for growing (acute) infections, and by eq. 12 for slow-growing cancer tumors.

To do this optimization for a fixed pathogenic threat (κ, R_0), we perform a uniform grid search over $\vec{\theta} \in [-\psi^{\max}, \psi^{\max}]^3 \times [-\ell^{\max}, \ell^{\max}]^3$. We use a Cartesian grid with $m = 13$ equally spaced points per coordinate, yielding a total of $13^6 \approx 4.8 \times 10^6$ candidate designs. For each $\vec{\theta}$, we simulate the infection dynamics to compute the associated harm and identify the design that minimizes harm $\vec{\theta}^*(\kappa, R_0)$.

Multi-objective design optimization across infection archetypes. To find designs that can perform well across infection archetypes, we pose a principled multi-objective problem for the immune program, i.e., choose a design $\vec{\theta}$ that minimizes the weighted sum of harms across pathogen archetypes:

$$\vec{\theta}'(\{\lambda_i\}) = \arg\min_{\vec{\theta}} \sum_{i:\text{archetypes}} \lambda_i H^{(i)}(\vec{\theta}), \tag{31}$$

where $\lambda_i \in [0, 1]$ with the constraint $\sum_i \lambda_i = 1$. The weights $\{\lambda_i\}$ determine the relative importance of each archetype in determining the optimal design $\vec{\theta}'$. The solution to this family of problems forms a simplex over which $\lambda_i's$ vary, called the *Pareto front*.

Similar to the case of optimization for a single infection, we estimate the Pareto front by sampling designs on a grid and identifying the design that minimizes the weighted harm $\sum_{i:\text{archetypes}} \lambda_i H^{(i)}(\vec{\theta})$ for a given instance of $\{\lambda_i\}$.

b. Finding the elbow of the two-dimensional Pareto curve. Even though optimization can be done simultaneously over all five archetypes, we examine trade-offs between pairs of archetypes A and B, generating one-dimensional Pareto fronts, with single independent weights $\lambda = \lambda_A = 1 - \lambda_B$. The curvature of the Pareto front quantifies the lowest achievable cost of decreasing harm from one archetype by tuning design parameters, in terms of a corresponding increase in harm from the other archetype. Given a convex Pareto front for archetypes A and B, in some regions of the curve (values of λ) the benefit of decreasing harm from one archetype may be larger than the cost. A common heuristic for multi-objective optimization in engineering problems is to identify the point along the curve, $(H[\vec{\theta}(\lambda^*)]^{(A)}, H[\vec{\theta}(\lambda^*)]^{(B)})$, called the *elbow* or *knee*, at which no additional benefit can be obtained by tuning parameters [46]. To estimate the elbow of the Pareto front, we first interpolate the discretely sampled Pareto points to estimate the curve. This procedure was sensitive to the density of sampled points near the Pareto front and requires fine-tuning. For plots involving archetype I (Fig. 4, S6), we found that subsetting to points where $H(\vec{\theta})^{(A)} < 10^{-2} S_{\max}$, and linearly interpolating this subset of performed best. Finally, we estimated the elbow point (λ^*) along the interpolated curve using the Kneedle algorithm [46, 47].

c. Identifying the ensemble of near-optimal designs. Since we are sampling over a finite grid and our simulations are stochastic, we must contend with the likelihood that discretely sampled points will not coincide with the true minimizer in the continuous space of designs and that stochastic effects will add noise to our measurements of harm. To capture this uncertainty, we identify a *near-optimal ensemble* of designs defined as the set of designs θ such that $H(\vec{\theta}) \leq H(\vec{\theta}') + \epsilon$, where $\vec{\theta}'$ is the sampled minimizer and $0 < \epsilon$ is a threshold. We typically choose $\epsilon = 10^{-2} S_{\max}$, and compute the mean and variance of designs in this ensemble, as shown in Fig. 3D.

Having estimated the Pareto front for a pair of archetypes, $\{A, B\}$, for each point on the Pareto front, $\vec{\theta}'(\lambda)$ (equation 31), we identify an ensemble of near-Pareto designs

$$\Theta(\lambda) = \{\vec{\theta} | H(\vec{\theta})^{(A)} \leq (1+\alpha) H[\vec{\theta}'(\lambda)]^{(A)}, H(\vec{\theta})^{(B)} \leq (1+\alpha) H[\vec{\theta}'(\lambda)]^{(B)}\}, \tag{32}$$

with $\alpha = 5 \times 10^{-2}$.

Clustering of infection archetypes

For each infection scenario (κ, R_0), we identify the ensemble of sampled designs $\vec{\theta}$ such that $H(\vec{\theta}) \leq H(\vec{\theta}^{(\min)}) + \epsilon$, where $\vec{\theta}^{(\min)}$ is the harm-minimizing design for that scenario and $\epsilon = 10^{-2} S_{\max}$. We call this ensemble $\Theta(\kappa, R_0)$, and compute its mean, $\langle \vec{\theta} \rangle_{\Theta(\kappa, R_0)}$ (Fig. S3). Next, we pool the mean vectors of all infection scenarios and normalize their distribution so each design parameter has zero mean and unit variance. Finally, we use a spectral clustering algorithm implemented in Python [48] to cluster scaled mean design parameters to identify distinct infection archetypes. We selected the cluster number $n_{\text{cluster}} = 4$ that maximized the mean silhouette coefficient (implemented in Python [48]),

$$\sum_{(\kappa, R_0)} [b(\kappa, R_0) - a(\kappa, R_0)] / \max[a(\kappa, R_0), b(\kappa, R_0)],$$

where $a(\kappa, R_0)$ is the mean distance between $\langle\vec{\theta}\rangle_{\Theta(\kappa,R_0)}$ and every other point in its assigned cluster, and $b(\kappa, R_0)$ is the mean distance between $\langle\vec{\theta}\rangle_{\Theta(\kappa,R_0)}$ and every other point in its nearest, neighboring cluster. The mean silhouette coefficient measures cohesiveness of clusters and how distinct they are from each other.

---

[1] Chao DL, et al. A stochastic model of cytotoxic T cell responses. J Theor Biol. 2004;228(2):227–240.
[2] Mayer A, et al. Regulation of T cell expansion by antigen presentation dynamics. Proc Natl Acad Sci U S A. 2019;116(13):5914–5919.
[3] Altan-Bonnet G, et al. Modeling T cell antigen discrimination based on feedback control of digital ERK responses. PLoS Biol. 2005;3(11):e356.
[4] François P, et al. Phenotypic model for early T-cell activation displaying sensitivity, specificity, and antagonism. Proc Natl Acad Sci U S A. 2013;110(10):E888–97.
[5] Milligan GN, et al. Vaccinology: An Essential Guide. Milligan GN, et al., editors. Hoboken, NJ: Wiley-Blackwell; 2015.
[6] Abuin P, et al. Characterization of SARS-CoV-2 dynamics in the host. Annu Rev Control. 2020;50:457–468.
[7] Deenick EK, et al. Stochastic model of T cell proliferation: a calculus revealing IL-2 regulation of precursor frequencies, cell cycle time, and survival. J Immunol. 2003;170(10):4963–4972.
[8] Cheon H, et al. Cyton2: A Model of Immune Cell Population Dynamics That Includes Familial Instructional Inheritance. Front Bioinform. 2021;1:723337.
[9] Yin A, et al. A review of mathematical models for tumor dynamics and treatment resistance evolution of solid tumors. CPT Pharmacometrics Syst Pharmacol. 2019;8(10):720–737.
[10] De Boer RJ, et al. Recruitment times, proliferation, and apoptosis rates during the CD8(+) T-cell response to lymphocytic choriomeningitis virus. J Virol. 2001;75(22):10663–10669.
[11] Lessler J, et al. Incubation periods of acute respiratory viral infections: a systematic review. Lancet Infect Dis. 2009;9(5):291–300.
[12] Marchingo JM, et al. T-cell stimuli independently sum to regulate an inherited clonal division fate. Nat Commun. 2016;7:13540.
[13] Plambeck M, et al. Heritable changes in division speed accompany the diversification of single T cell fate. Proceedings of the National Academy of Sciences. 2022;119(9):e2116260119.
[14] Henrickson SE, et al. T cell sensing of antigen dose governs interactive behavior with dendritic cells and sets a threshold for T cell activation. Nat Immunol. 2008;9(3):282–291.
[15] Uhl M, et al. Autophagy within the antigen donor cell facilitates efficient antigen cross-priming of virus-specific CD8+ T cells. Cell Death Differ. 2009;16(7):991–1005.
[16] Bousso P, et al. Dynamics of CD8+ T cell priming by dendritic cells in intact lymph nodes. Nat Immunol. 2003;4(6):579–585.
[17] Belz GT, et al. Killer T cells regulate antigen presentation for early expansion of memory, but not naive, $CD8^+$ T cell. Proc Natl Acad Sci U S A. 2007;104(15):6341–6346.
[18] Ozga AJ, et al. pMHC affinity controls duration of CD8+ T cell-DC interactions and imprints timing of effector differentiation versus expansion. J Exp Med. 2016;213(12):2811–2829.
[19] Au-Yeung BB, et al. A sharp T-cell antigen receptor signaling threshold for T-cell proliferation. Proc Natl Acad Sci U S A. 2014;111(35):E3679–88.
[20] van Heijst JWJ, et al. Recruitment of antigen-specific CD8+ T cells in response to infection is markedly efficient. Science. 2009;325(5945):1265–1269.
[21] van Stipdonk MJB, et al. Dynamic programming of CD8+ T lymphocyte responses. Nat Immunol. 2003;4(4):361–365.
[22] van Stipdonk MJ, et al. Naïve CTLs require a single brief period of antigenic stimulation for clonal expansion and differentiation. Nat Immunol. 2001;2(5):423–429.
[23] Kaech SM, et al. Memory CD8+ T cell differentiation: initial antigen encounter triggers a developmental program in naïve cells. Nat Immunol. 2001;2(5):415–422.
[24] Youngblood B, et al. Effector CD8 T cells dedifferentiate into long-lived memory cells. Nature. 2017;552(7685):404–409.
[25] Abadie K, et al. Flexible and scalable control of T cell memory by a reversible epigenetic switch; 2023.
[26] Buchholz VR, et al. Disparate individual fates compose robust CD8+ T cell immunity. Science. 2013;340(6132):630–635.
[27] Heinzel S, et al. A Myc-dependent division timer complements a cell-death timer to regulate T cell and B cell responses. Nat Immunol. 2017;18(1):96–103.
[28] Badovinac VP, et al. Initial T cell receptor transgenic cell precursor frequency dictates critical aspects of the CD8(+) T cell response to infection. Immunity. 2007;26(6):827–841.
[29] Zhang N, et al. CD8(+) T cells: foot soldiers of the immune system. Immunity. 2011;35(2):161–168.
[30] Hiam-Galvez KJ, et al. Transient dendritic cell activation diversifies the T cell response to acute infection; 2021.
[31] Reiner SL, et al. Lymphocyte fate specification as a deterministic but highly plastic process. Nat Rev Immunol. 2014;14(10):699–704.
[32] Liao W, et al. Interleukin-2 at the crossroads of effector responses, tolerance, and immunotherapy. Immunity. 2013;38(1):13–25.
[33] Guillemin A, et al. Noise and the molecular processes underlying cell fate decision-making. Phys Biol. 2021;18(1):011002.

[34] Jenkins MR, et al. Cell cycle-related acquisition of cytotoxic mediators defines the progressive differentiation to effector status for virus-specific CD8+ T cells. J Immunol. 2008;181(6):3818–3822.
[35] Kretschmer L, et al. Differential expansion of T central memory precursor and effector subsets is regulated by division speed. Nat Commun. 2020;11(1):113.
[36] Walczak AM, et al. Optimizing information flow in small genetic networks. II. Feed-forward interactions. Phys Rev E Stat Nonlin Soft Matter Phys. 2010;81(4 Pt 1):041905.
[37] de Ronde W, et al. Protein logic: a statistical mechanical study of signal integration at the single-molecule level. Biophys J. 2012;103(5):1097–1107.
[38] Hart Y, et al. Paradoxical signaling by a secreted molecule leads to homeostasis of cell levels. Cell. 2014;158(5):1022–1032.
[39] Oyler-Yaniv A, et al. A Tunable Diffusion-Consumption Mechanism of Cytokine Propagation Enables Plasticity in Cell-to-Cell Communication in the Immune System. Immunity. 2017;46(4):609–620.
[40] Achar SR, et al. Universal antigen encoding of T cell activation from high-dimensional cytokine dynamics. Science. 2022;376(6595):880–884.
[41] Voisinne G, et al. T Cells Integrate Local and Global Cues to Discriminate between Structurally Similar Antigens. Cell Rep. 2015;11(8):1208–1219.
[42] Adler M, et al. Logarithmic and power law input-output relations in sensory systems with fold-change detection. PLoS Comput Biol. 2014;10(8):e1003781.
[43] Gillespie DT. Stochastic simulation of chemical kinetics. Annu Rev Phys Chem. 2007;58:35–55.
[44] Cao Y, et al. Efficient step size selection for the tau-leaping simulation method. J Chem Phys. 2006;124(4):044109.
[45] Heinzel S, et al. The regulation of lymphocyte activation and proliferation. Curr Opin Immunol. 2018;51:32–38.
[46] Satopaa V, et al. Finding a "kneedle" in a haystack: Detecting knee points in system behavior. In: 2011 31st International Conference on Distributed Computing Systems Workshops. IEEE; 2011.
[47] Arvai K. Kneed: Knee-point detection in Python; 2023. Available from: https://doi.org/10.5281/zenodo.8127224.
[48] Pedregosa F, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12:2825–2830.
[49] Ikeda H, et al. Improving the estimation of the death rate of infected cells from time course data during the acute phase of virus infections: application to acute HIV-1 infection in a humanized mouse model. Theor Biol Med Model. 2014;11(1):22.
[50] Ke R, et al. On the death rate of abortively infected cells: Estimation from simian-human immunodeficiency virus infection. J Virol. 2017;91(18).
[51] Barchet W, et al. Direct quantitation of rapid elimination of viral antigen-positive lymphocytes by antiviral CD8(+) T cells in vivo. Eur J Immunol. 2000;30(5):1356–1363.
[52] Yiu HH, et al. Dynamics of a cytokine storm. PLoS One. 2012;7(10):e45027.
[53] Straub A, et al. Recruitment of epitope-specific T cell clones with a low-avidity threshold supports efficacy against mutational escape upon re-infection. Immunity. 2023;.
[54] van Stipdonk MJ, et al. Naïve CTLs require a single brief period of antigenic stimulation for clonal expansion and differentiation. Nat Immunol. 2001;2(5):423–429.
[55] Zehn D, et al. Complete but curtailed T-cell response to very low-affinity antigen. Nature. 2009;458(7235):211–214.

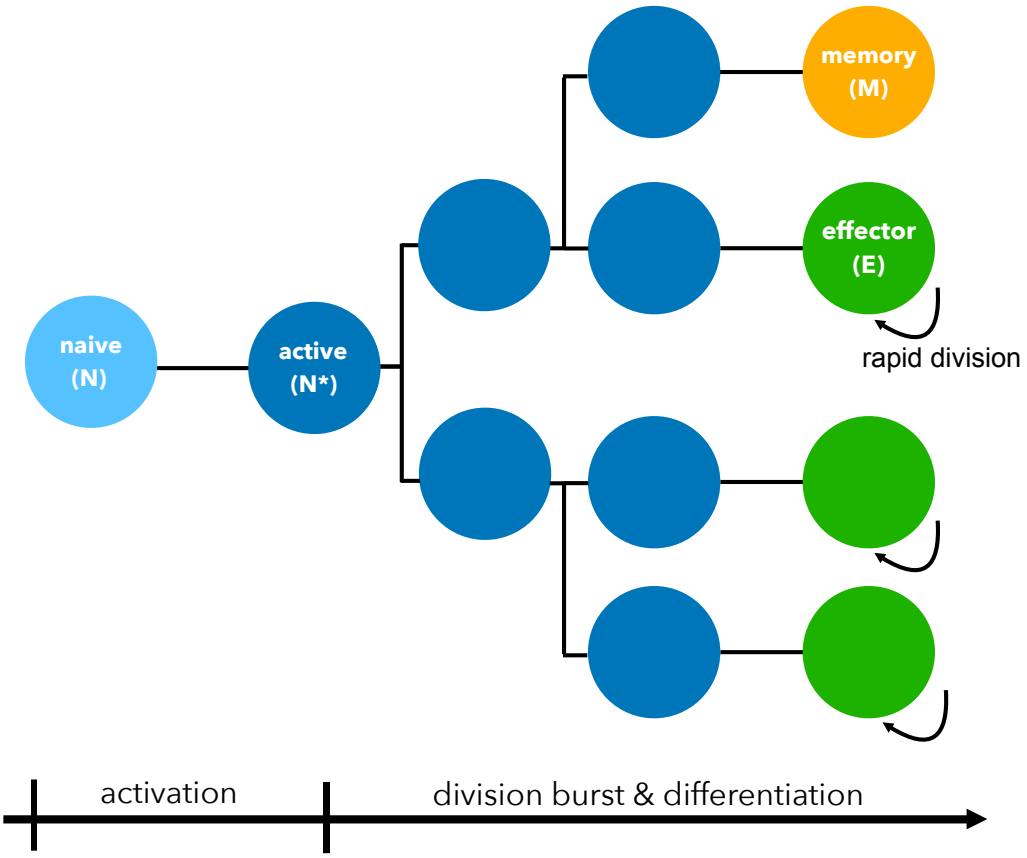

Figure S1. **Schematic of modeled early T-cell division bursts and differentiation.**

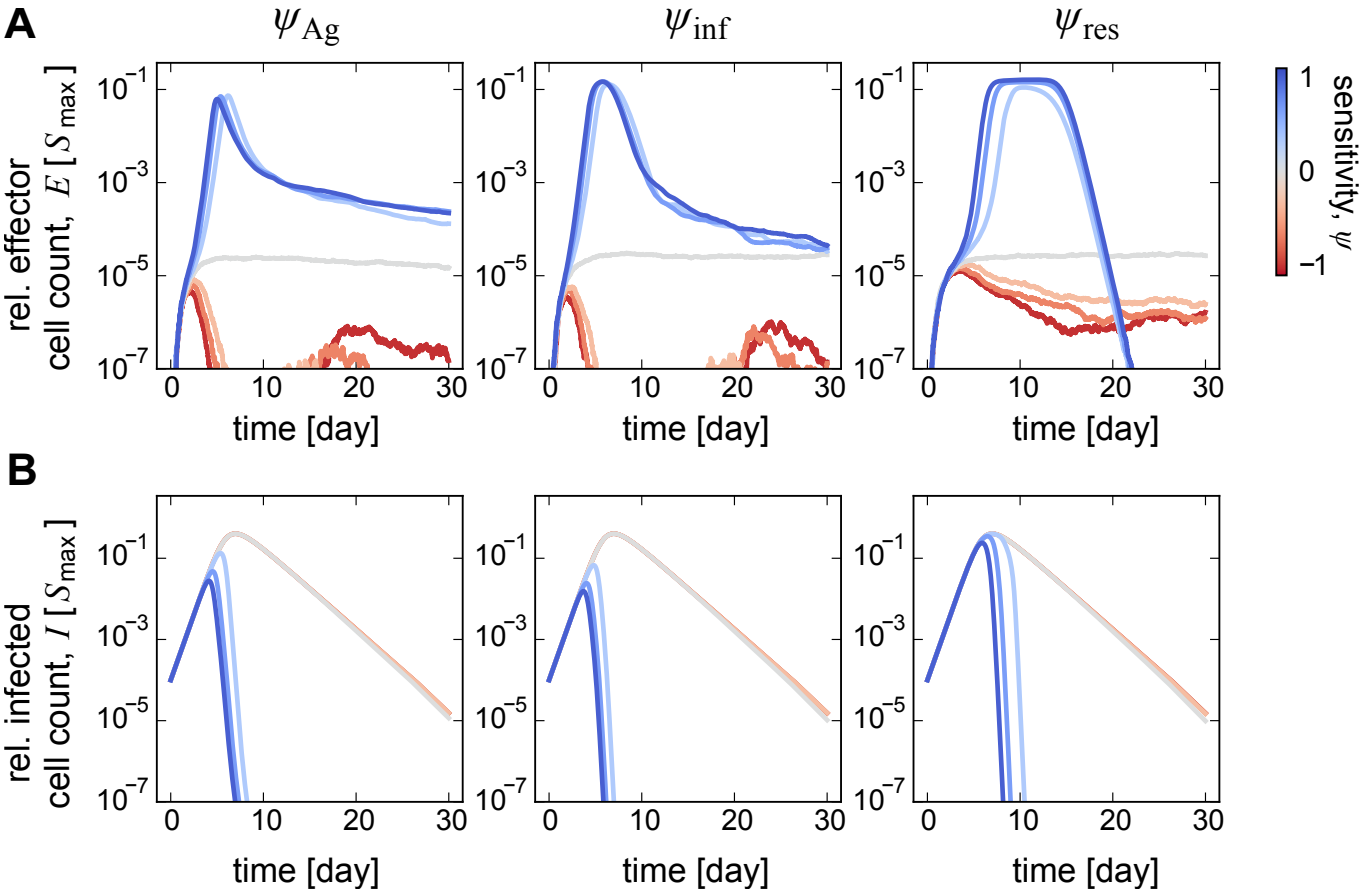

Figure S2. **Design parameters shape T-cell response and infection clearance. (A, B)** Time course of the relative population size of (A) effector cells E and (B) infected cells is shown for designs when varying (color) a single sensitivity parameter ψ_σ (indicated at the top of each column), while keeping the rest fixed at: moderate activation / proliferation $g_{0,N\to N^*} = g_{0,E\to 2E} = 0.5$; moderate differentiation $g_{0,N^*\to E} = 0.5$; moderate contraction $g_{0,E\to\emptyset} = 0.5$; and zero sensitivity to the other signals $\psi_{\mathrm{Ag}} = \psi_{\mathrm{inf}} = \psi_{\mathrm{res}} = 0$. All population sizes are measured in units of maximum number of susceptible cells S_{max} and sensitivities in units of ψ_{max}. Other simulation parameters as in Fig. 2.

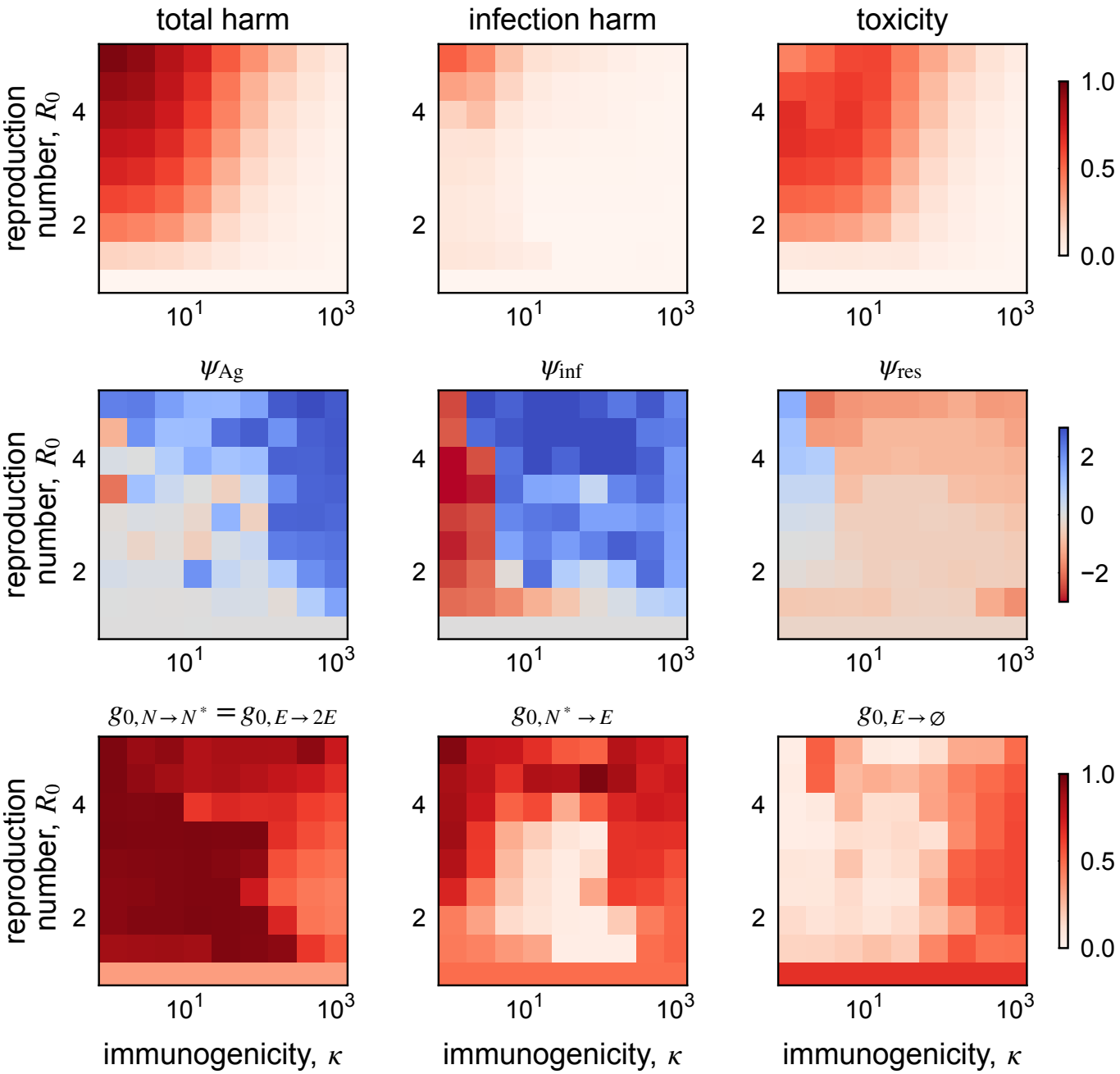

Figure S3. **Near-optimal performance and designs for each infection scenario.** Heat maps showing the near-optimal outcomes and design values that achieve them for each infection scenario (κ, R_0). The first row displays the average total harm, infection harm and toxicity incurred by the near-optimal designs for each infection scenario. The last two rows display the average near-optimal design values for each scenario, $\langle\vec{\theta}\rangle_{\Theta(\kappa,R_0)}$. The data in the last two rows is used for clustering infection scenarios in Fig. 3B.

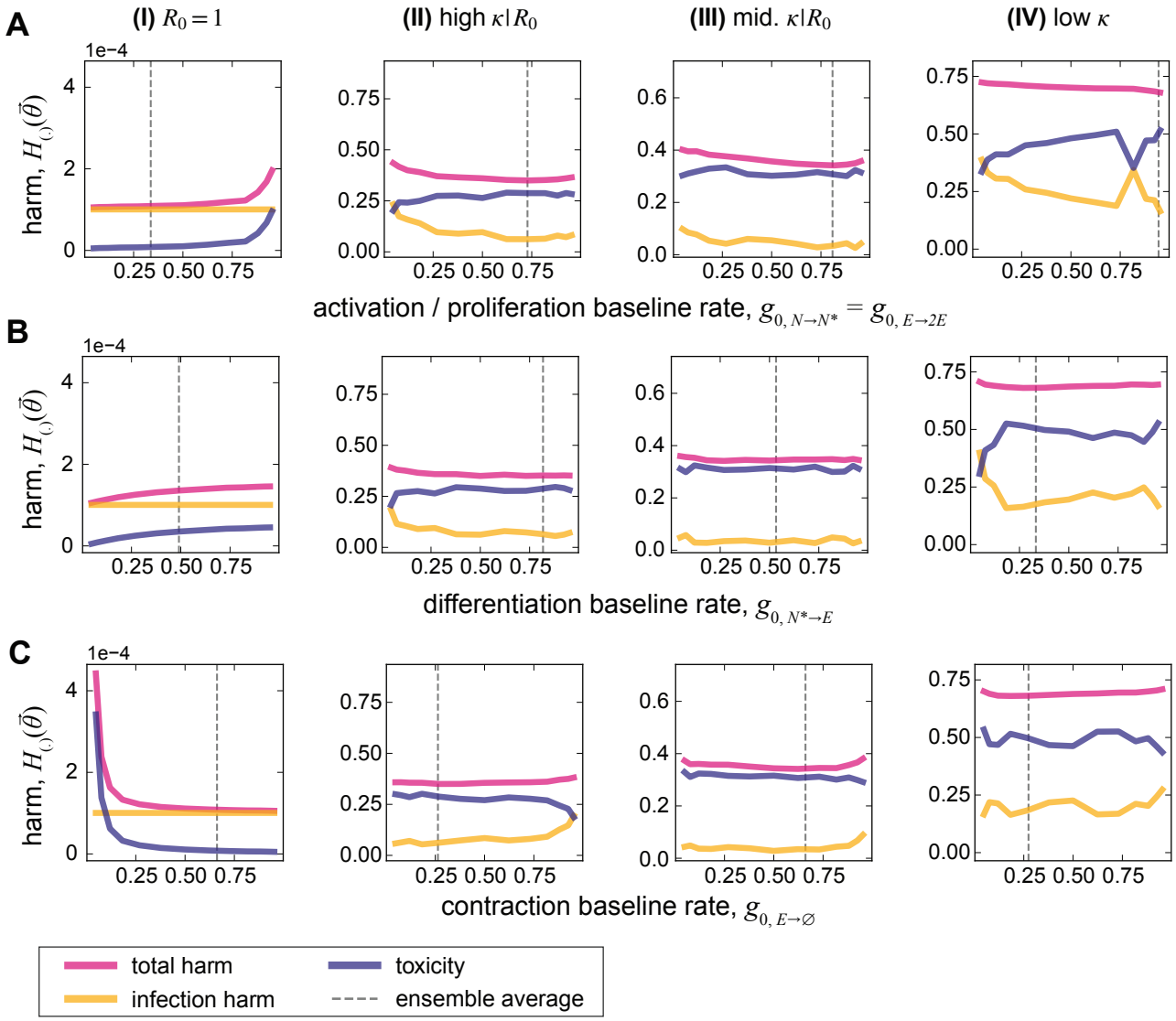

Figure S4. **Dependence of harm on baseline rates in different archetypes.** Total harm, infection harm, and toxicity (colors) are shown as a function of baseline rates for **(A)** activation / proliferation $g_{0,N\to N^*} = g_{0,E\to 2E}$, **(B)** differentiation $g_{0,N^*\to E}$ and **(C)** contraction $g_{0,E\to\emptyset}$, in different archetypes as indicated above each column. In each panel, the value of the design parameter on the horizontal axis is fixed and total harm is minimized by varying the other design parameters. The dashed line in each panel indicates the near-optimal ensemble average values shown in bar charts in Fig. 3D. Simulation parameters as in Fig. 3.

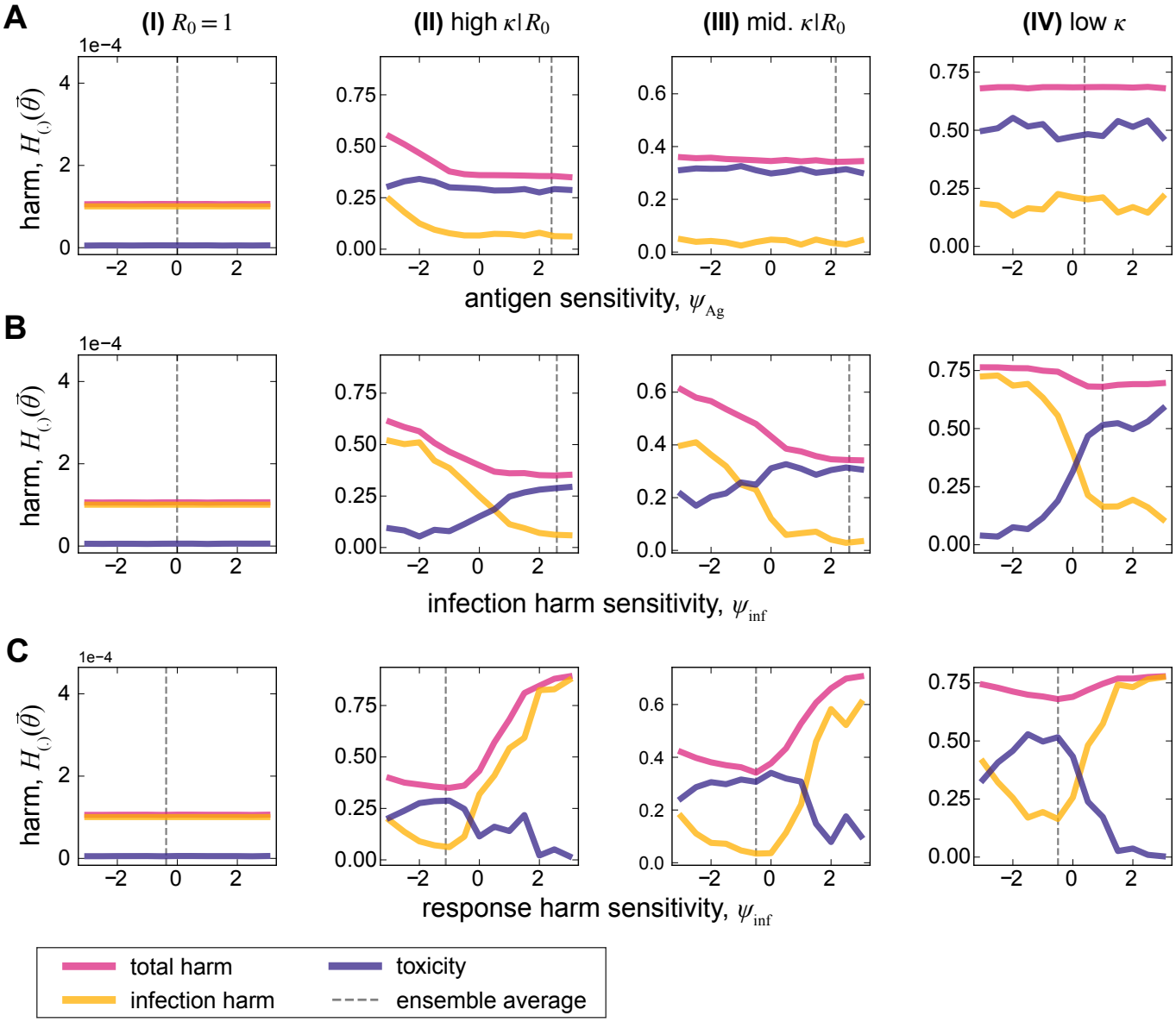

Figure S5. **Dependence of harm on signal sensitivities in different archetypes.** Similar to Fig. S4, but when varying the sensitivities to **(A)** the antigen signal ψ_{Ag}, **(B)** the infection harm signal ψ_{inf}, and **(C)** the response harm signal ψ_{res}.

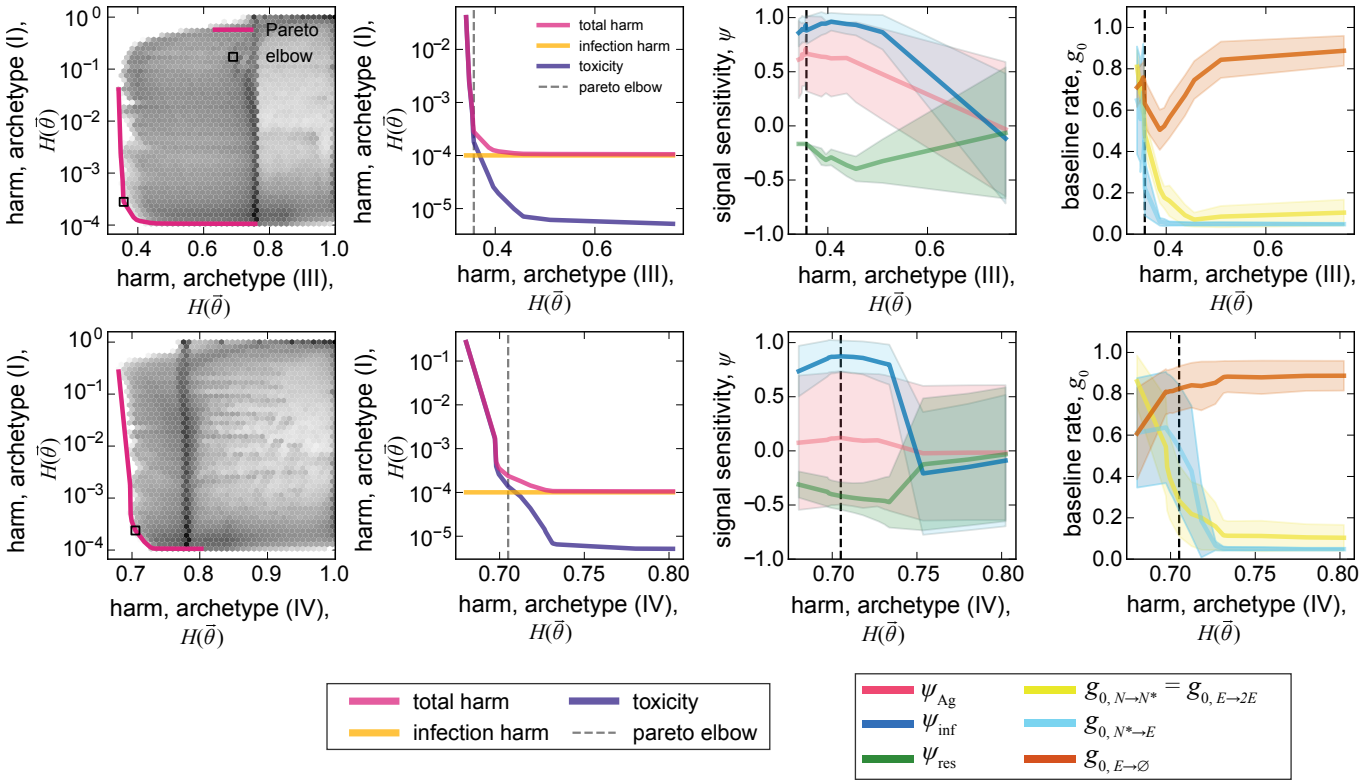

Figure S6. **Pareto-optimal design balancing harm from autoimmunity and other infection archetypes.** Similar to Fig. 4, but showing the Pareto optimal designs resulting from joint optimization of infection archetypes I and III, and infection archetypes I and IV (rows); joint optimization for archetypes I and II is shown in Fig. 4. Simulation parameters as in Fig. 4.

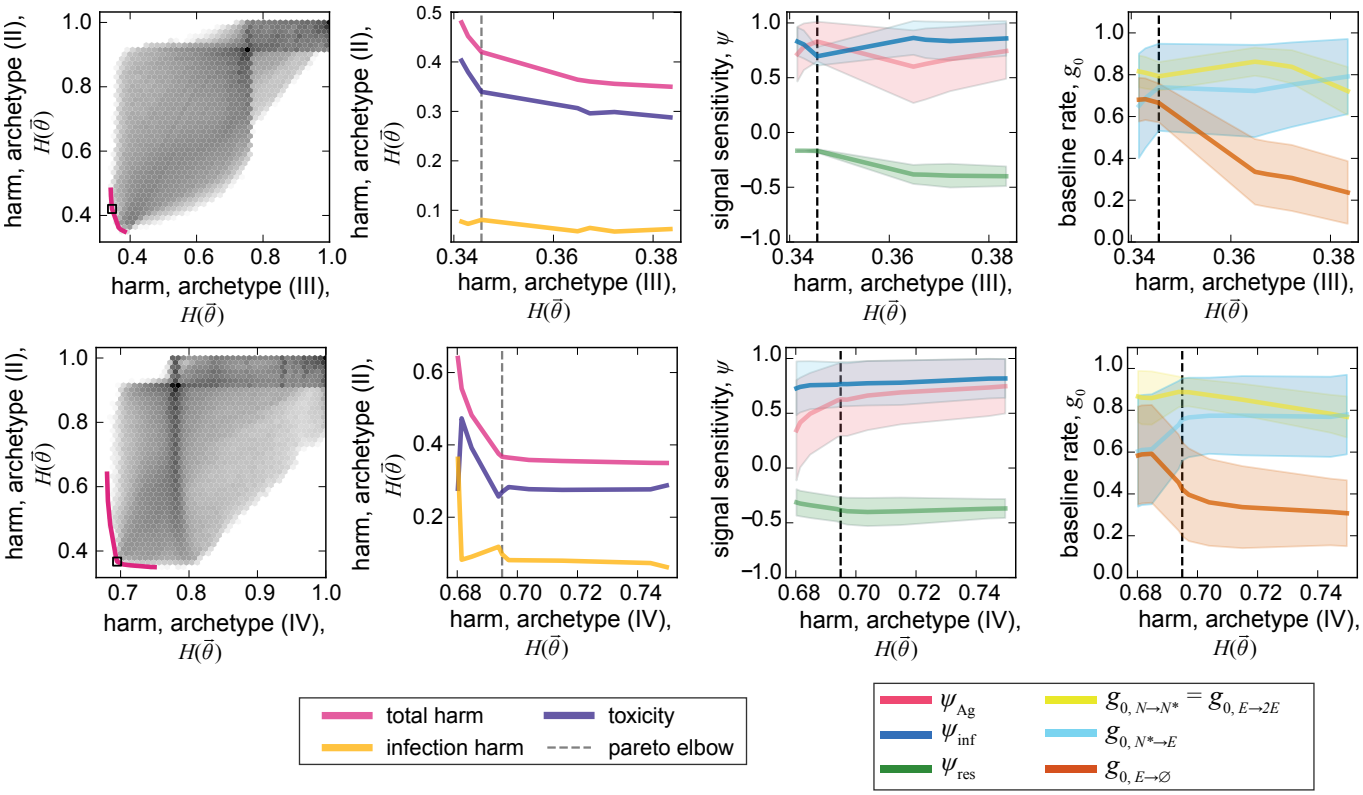

Figure S7. **Pareto-optimal design balancing harm from acute infection and other infection archetypes.** Similar to Fig. 4, but showing the Pareto optimal designs resulting from joint optimization of infection archetype II and III, and archetypes II and IV (rows); joint optimization for archetypes I and II is shown in Fig. 4. Simulation parameters as in Fig. 4.

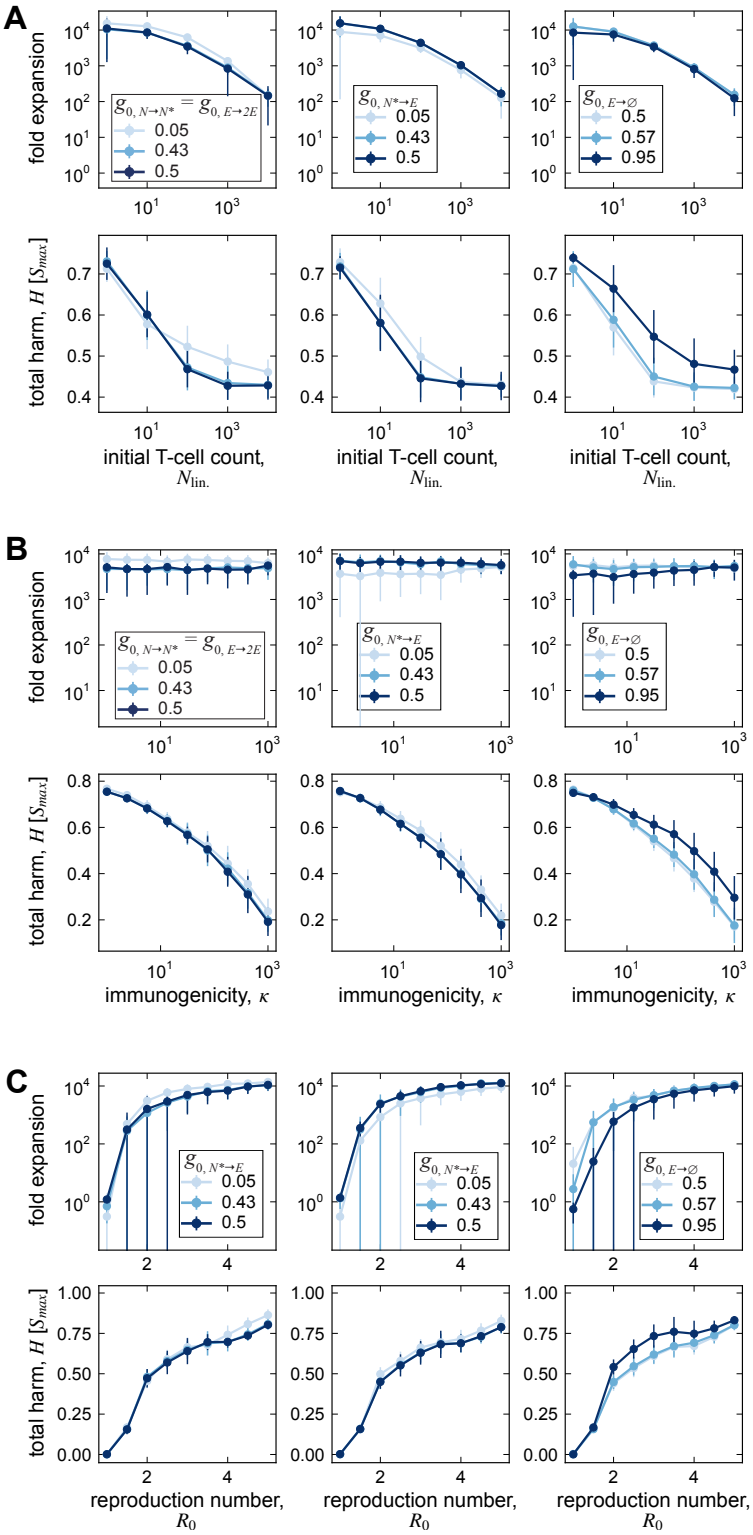

Figure S8. **Dependence of T-cell macro-dynamics on baseline rates.** T-cell fold expansion (top) and total harm (bottom) are shown as functions of **(A)** initial T-cell number $N_{\text{lin.}}$, **(B)** immunogenicity κ, **(C)** and the intra-host basic reproduction of the infection R_0. In each panel, outcomes are shown for the "elbow" Pareto-optimal program from Fig. 4B while varying a single baseline rate (colors)—activation/proliferation baseline rate $g_{0,N\to N^*} = g_{0,E\to 2E}$ (left), differentiation baseline rate $g_{0,N^*\to E}$ (center), or contraction baseline rate $g_{0,E\to\emptyset}$ (right)—with all other parameters held fixed. Trend lines and standard deviation error bars are obtained by averaging over infection scenarios (in Fig. 3B) and designs in the near-elbow ensemble. Simulation parameters as in Fig. 5.

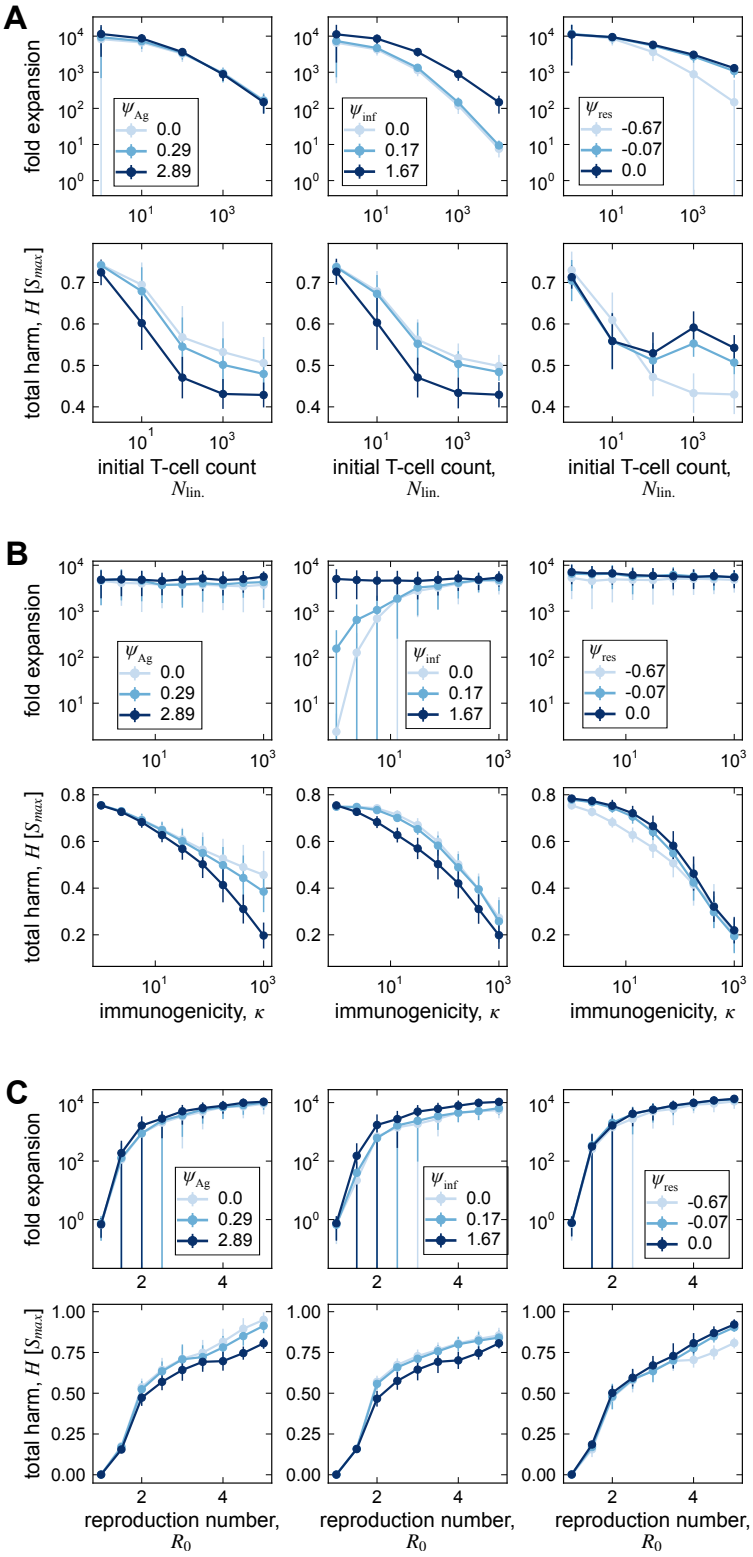

Figure S9. **Dependence of T-cell macro-dynamics on signal sensitivities.** T-cell fold expansion (top) and total harm (bottom) are shown as functions of **(A)** initial T-cell number $N_{\text{lin.}}$, **(B)** immunogenicity κ, **(C)** and the intra-host basic reproduction of the infection R_0. In each panel, outcomes are shown for the "elbow" Pareto-optimal program from Fig. 4B while varying a single sensitivity parameter (colors)—antigen sensitivity ψ_{Ag} (left), infection-harm sensitivity ψ_{inf} (center), or response-harm sensitivity ψ_{res} (right)—with all other parameters held fixed. Trend lines and standard deviation error bars are obtained by averaging over infection scenarios (in Fig. 3B) and designs in the near-elbow ensemble. Simulation parameters as in Fig. 5.

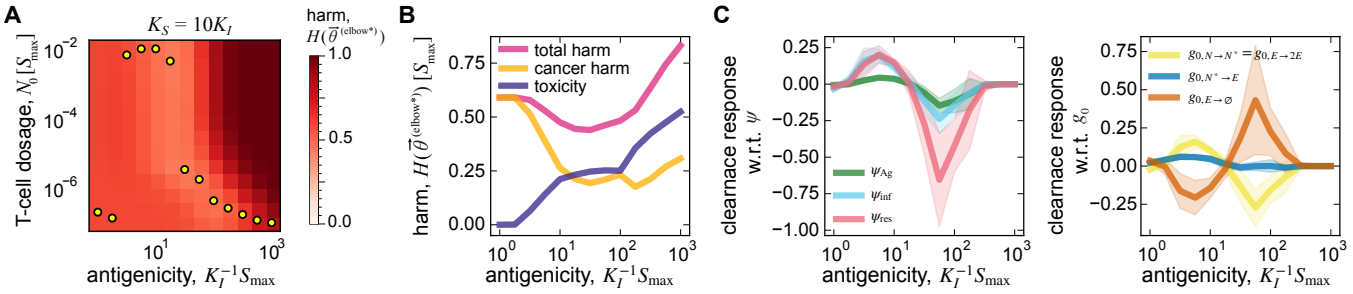

Figure S10. **Impact of self-antigenicity on the engineering of immune response for cancer immunotherapy.** As in Fig. 6, but incorporating cross-reactivity of engineered TCRs with self and tumor antigens. We model cross-reactivity by assuming tumor and self reactivities are proportional, fixing the immunogenicity $\kappa = K_I^{-1}K_s = 10$ while varying the antigenicity of the tumor defined as $K_I^{-1}S_{\max}$ (proxy for inverse EC_{50}, where larger values indicate higher antigenicity). **(A)** With responses averaged over the tuned near-elbow ensemble $\tilde{\Theta}(\lambda^*)$, the heatmap shows harm as a function of T-cell dosage and tumor antigenicity. **(B)** Total harm, cancer harm, and toxicity are shown as functions of tumor antigenicity at the optimal T-cell dosage (yellow circles in A). **(C)** Clearance response, defined as sensitivity of protection (negative harm) to each design parameter—evaluated and averaged over $\theta \in \tilde{\Theta}(\lambda^*)$—is shown for different levels of tumor antigenicity for signal sensitivity parameters ψ's (left), and for baseline parameters g_0's (right). For this analysis, we assume that T-cells are administered at high dosage ($N_{\text{lin.}} = 10^{-2}S_{\max}$). Simulation parameters as in Fig. 6.

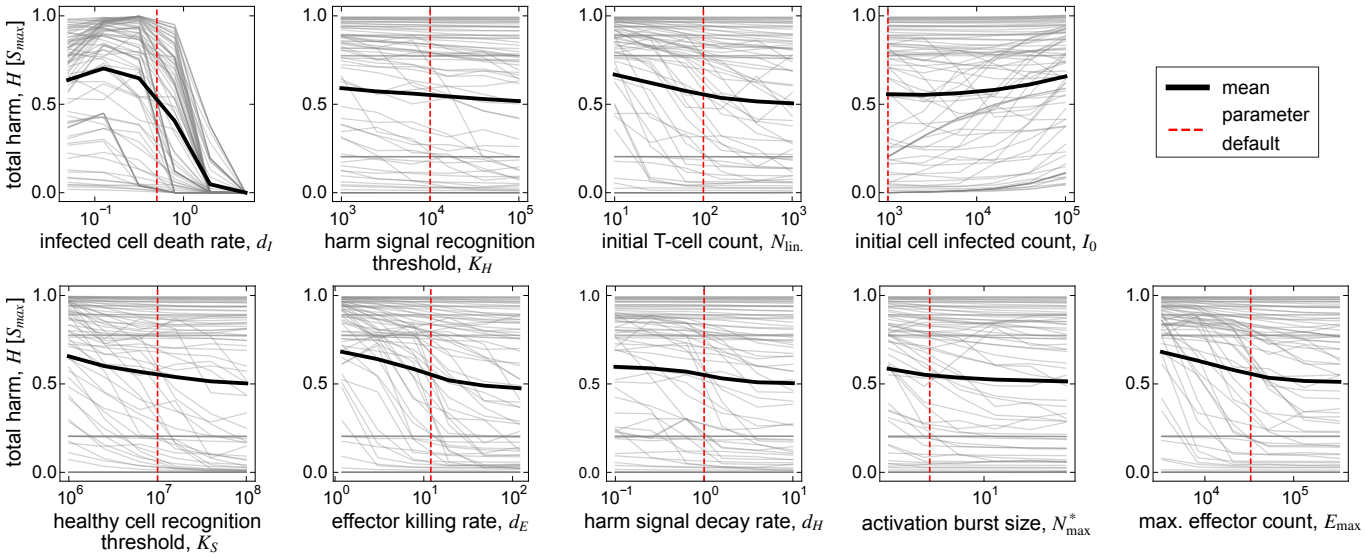

Figure S11. **Parameter sensitivity analysis for total harm.** Total harm, evaluated and averaged over the near-elbow ensemble $\theta \in \tilde{\Theta}(\lambda^*)$, is shown for different levels of the indicated parameter value. Gray lines represent trend lines for each infection condition (κ, R_0) in Fig. 3B, and the thick black line is the average trend over all infections. The red dotted line indicates the default parameter value used in simulations (SI Table S1). Except when explicitly varied, simulation parameters are as in Fig. 3.

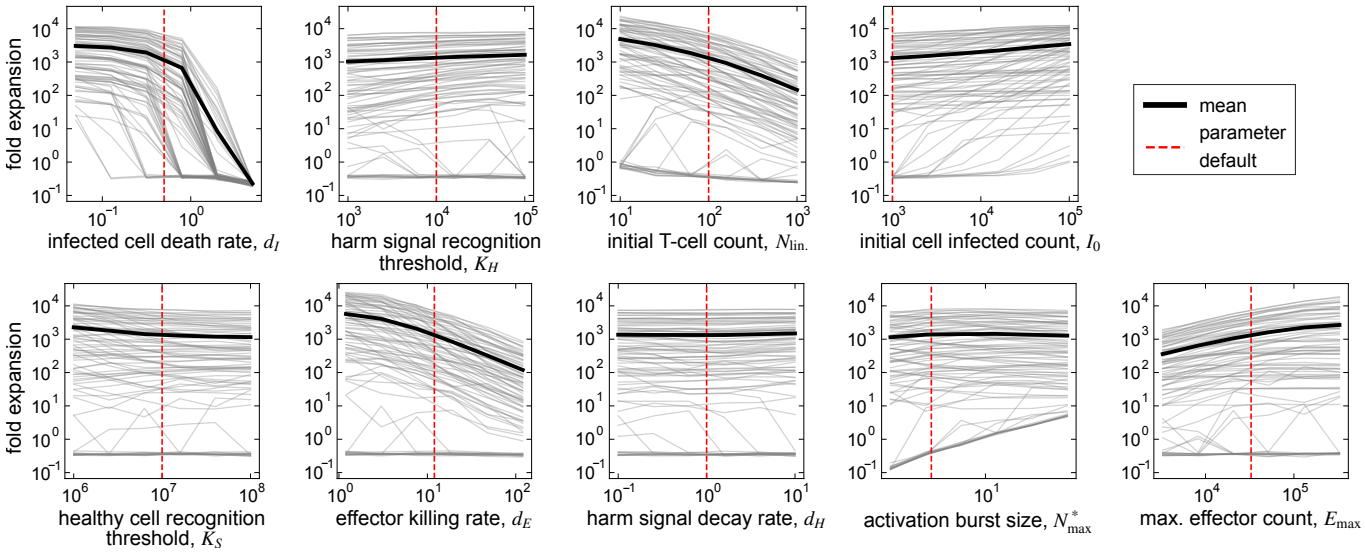

Figure S12. **Parameter sensitivity analysis for effector fold expansion.** Similar to Fig. S11 but showing the trend for effector fold expansion.

Table S1. Model parameters for simulations. The * symbol indicates parameters that were systematically varied in simulations.

Parameter	Value	Reference(s)
initial susceptible cell count (S_0)	10^7	
initial infected cell count (I_0)	10^3	
infected cell death rate (d_I)	0.5 day^{-1}	[49, 50]
susceptible cell infection rate (b_I)*	$1 - 5 \times 10^{-7}$ infected cell^{-1}day^{-1}	[49]
maximum cancer growth rate (b_C)	0.1 day^{-1}	
effector cell killing rate (d_E)	12 day^{-1}	[51]
infected cell recognition threshold (K_I)*	$10^4 - 10^7$ cells	[1]
healthy cell recognition threshold (K_S)*	10^7 cells	
harm signal recognition threshold (K_H)	10^4 cells	
max. binding rate of APCs to Naive cells ($r_{\text{bind}}^{\max}$)	1 day^{-1}	[2, 10, 17]
decay rate of harm signals (d_H)	1.0 day^{-1}	[52]
initial naive cognate T-cell count ($N_{\text{lin.}}$)*	10^2	[53]
activation burst size ($N_{\max}^*$)	2^2	[13]
max. effector count per lineage ($E_{\max}$)	2^{15}	[12, 26, 28, 29]
max. rate of T-cell state transition ($r^{\max}$)	4 day^{-1}	[13, 34, 35, 54]
rate of activated naive cell division ($r_{N^* \to 2N^*}$)	$\frac{24}{8.8}$ day^{-1}	[13]
signal sensitivity (ψ_σ)*	$\in [-3, 3]$	
Baseline parameter (ℓ_i)*	$\in [-3, 3]$	
simulation duration (T)	30 days	[11, 28, 55]
time step (Δt)	10^{-2} days	