Supporting Information

Bocharov et al. 10.1073/pnas.1019706108

SI Methods

Section S1. The Core Mathematical Models and Parameters. The rate of change of the number of immature proliferating cells, X_1, is described by

$$\frac{dX_1}{dt} = p_1 \cdot f_p \cdot X_1 - (\alpha_1 + \alpha_{12} \cdot f_{x1x2} \cdot Z_2) \cdot X_1. \qquad [S1]$$

The first term on the right-hand side describes the proliferation of cells with a nominal per-capita rate p_1 that can be shaped by crowding effects at high numbers of X_1 and X_2 cells represented by the function $f_p \equiv f_p(X_1, X_2) = \frac{1}{1+\gamma \cdot (1-e^{-(X_1+X_2)/c})}$. Note that the per-capita rate coefficient in this equation and in the next one represents *net* proliferation rate, the difference between the "actual" rate and an unknown death rate. The second term represents the differentiation of X_1 cells into X_2 with a total flux that includes a constitutive linear component (α_1) and a feedback-regulated differentiation component ($\alpha_{12} \cdot f_{x1x2} \cdot Z_2$). Feedback-imposed differentiation is mediated by Z_2 but assumed to be limited by the total number of differentiated cells Z_1 and, Z_2 as follows: $f_{x1x2} \equiv f_{x1x2}(Z_1, Z_2) = \frac{1}{1+(Z_1+Z_2)/\theta_{x1}}$. Whether Z_1 cells indeed participate in determining the limitation in feedback depends on the nature of the interactions of Z_1 and Z_2 with APC—e.g., on whether the induction of differentiation involves cell-to-cell contact—and on the degree of interference between them, but our results did not change in any significant manner if Z_1 were not included.

The dynamics of the mature proliferating T cells is modeled by

$$\frac{dX_2}{dt} = p_2 \cdot f_p \cdot X_2 + (\alpha_1 + \alpha_{12} \cdot f_{x1x2} \cdot Z_2) \cdot X_1 - (\alpha_2 + \alpha_{22} \cdot f_{x2z1} \cdot Z_2) \cdot X_2. \qquad [S2]$$

Here the first term describes net proliferation of cells of the X_2 subset, the second term is the inflow of cells from the previous compartment, X_1, and the last term represents constitutive and feedback-induced differentiation into nondividing Z_1 cells where that rate is controlled by Z_2 and limited by the sizes of the Z_1 and Z_2 subsets via the function $f_{x2z1} \equiv f_{x2z1}(Z_1, Z_2) = \frac{1}{1+(Z_1+Z_2)/\theta_{x2}}$.

The equations for the nondividing differentiated cells, Z_1 and Z_2,

$$\frac{dZ_1}{dt} = (\alpha_2 + \alpha_{22} \cdot f_{x2z1} \cdot Z_2) \cdot X_2 - (\beta_1 + \beta_{12} \cdot f_{z1z2} \cdot Z_2) \cdot Z_1 \qquad [S3]$$

$$\frac{dZ_2}{dt} = (\beta_1 + \beta_{12} \cdot f_{z1z2} \cdot Z_2) \cdot Z_1 - \delta \cdot Z_2, \qquad [S4]$$

describe the Z_2-induced transition of cells through an intermediate stage Z_1 into Z_2. The induction rate is parameterized using the function $f_{z1z2} \equiv f_{z1z2}(Z_1, Z_2) = \frac{1}{1+(Z_1+Z_2)/\theta_z}$. Z_2 cells are assumed to die or migrate at some constant per-capita rate δ.

The parameters of the core model for activated cell growth and differentiation are listed and described in Table S1. The best-fit parameter values were estimated using the maximum-likelihood approach described in *Section S5*. The confidence intervals are broad, indicating that the information content of the data used for parameter estimation does not allow us to reliably identify all of the parameters of the model. This result implies that for the purpose of quantitative description of the considered data the model is overparameterized and probably can be reduced without loss of description accuracy. This issue of model parsimony is addressed in *Results* in the main text and in *Section S7*. The model parameters of the parsimonious model are listed in Table S2.

Section S2. Mathematical Model for BrdU-Labeling Data. To enable incorporation of the BrdU-labeling data, extension of the model was developed according to Fig. S1*A*. Interestingly, to consistently assimilate the clonal kinetics data along with the BrdU-labeling data, apoptosis of proliferating cells had to be explicitly included in the BrdU-labeling equations. Thus, we distinguish between the net proliferation rates (p_1, p_2) introduced in the basic Eqs. **S1** and **S2** and the genuine proliferation rates ($p_1 + d$, $p_2 + d$), where d is the per-capita death-rate constant. The extended BrdU version describes the population dynamics of labeled and unlabeled cells as follows:

$$\frac{dX_1^U}{dt} = -(p_1 + d) \cdot f_p \cdot X_1^U - (\alpha_1 + \alpha_{12} \cdot f_{x1x2} \cdot (Z_2^U + Z_2^L)) \cdot X_1^U - d \cdot f_p \cdot X_1^U$$

$$\frac{dX_2^U}{dt} = -(p_2 + d) \cdot f_p \cdot X_2^U + (\alpha_1 + \alpha_{12} \cdot f_{x1x2} \cdot (Z_2^U + Z_2^L)) \cdot X_1^U - (\alpha_2 + \alpha_{22} \cdot f_{x2z1} \cdot (Z_2^U + Z_2^L)) \cdot X_2^U - d \cdot f_p \cdot X_2^U$$

$$\frac{dX_1^L}{dt} = 2 \cdot (p_1 + d) \cdot f_p \cdot X_1^U + (p_1 + d) \cdot f_p \cdot X_1^L - (\alpha_1 + \alpha_{12} \cdot f_{x1x2} \cdot (Z_2^U + Z_2^L)) \cdot X_1^L - d \cdot f_p \cdot X_1^L$$

$$\frac{dX_2^L}{dt} = 2 \cdot (p_2 + d) \cdot f_p \cdot X_2^U + (p_2 + d) \cdot f_p \cdot X_2^L + (\alpha_1 + \alpha_{12} \cdot f_{x1x2} \cdot (Z_2^U + Z_2^L)) \cdot X_1^L - (\alpha_2 + \alpha_{22} \cdot f_{x2z1} \cdot (Z_2^U + Z_2^L)) \cdot X_2^L - d \cdot f_p \cdot X_2^L$$

$$\frac{dZ_1^U}{dt} = (\alpha_2 + \alpha_{22} \cdot f_{x2z1} \cdot (Z_2^U + Z_2^L)) \cdot X_2^U - (\beta_1 + \beta_{12} \cdot f_{z1z2} \cdot (Z_2^U + Z_2^L)) \cdot Z_1^U$$

$$\frac{dZ_2^U}{dt} = (\beta_1 + \beta_{12} \cdot f_{z1z2} \cdot (Z_2^U + Z_2^L)) \cdot Z_1^U - \delta \cdot Z_2^U$$

$$\frac{dZ_1^L}{dt} = (\alpha_2 + \alpha_{22} \cdot f_{x2z1} \cdot (Z_2^U + Z_2^L)) \cdot X_2^L - (\beta_1 + \beta_{12} \cdot f_{z1z2} \cdot (Z_2^U + Z_2^L)) \cdot Z_1^L$$

$$\frac{dZ_2^L}{dt} = (\beta_1 + \beta_{12} \cdot f_{z1z2} \cdot (Z_2^U + Z_2^L)) \cdot Z_1^L - \delta \cdot Z_2^L.$$

Section S3. Mathematical Model for CFSE Dilution Data. To compare the model's prediction with these experimental findings, we replaced the basic set of equations (Eqs. **S1**–**S4** in *Section S1*) with equations describing explicitly the evolving CFSE-labeled cell structure according to Fig. S1*B*. Every subset of cells in the clone is further subdivided into a number of compartments reflecting

cell generations with different numbers of divisions $n = 0, 1, \ldots, N$ they underwent by time t; e.g., $X_1(t) = \sum_{n=0}^{N} X_1^n(t)$. Therefore, the CFSE model equations describe the rates of change in numbers of cells having defined division histories within the cell subsets X_1, X_2, Z_1, Z_2. The kinetics of nondivided labeled cells ($n = 0$) are given by

$$\frac{dX_1^0}{dt} = -p_1 \cdot f_p \cdot X_1^0 - \left(\alpha_1 + \alpha_{12} \cdot f_{x1x2} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot X_1^0$$

$$\frac{dX_2^0}{dt} = -p_2 \cdot f_p \cdot X_2^0 + \left(\alpha_1 + \alpha_{12} \cdot f_{x1x2} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot X_1^0 - \left(\alpha_2 + \alpha_{22} \cdot f_{x2z1} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot X_2^0$$

$$\frac{dZ_1^0}{dt} = \left(\alpha_2 + \alpha_{22} \cdot f_{x2z1} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot X_2^0 - \left(\beta_1 + \beta_{12} \cdot f_{z1z2} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot Z_1^0$$

$$\frac{dZ_2^0}{dt} = \left(\beta_1 + \beta_{12} \cdot f_{z1z2} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot Z_1^0 - \delta \cdot Z_2^0$$

and for cells that divided n times ($n = 1, \ldots, N$) the equations are

$$\frac{dX_1^n}{dt} = 2p_1 \cdot f_p \cdot X_1^{n-1} - p_1 \cdot f_p \cdot X_1^n - \left(\alpha_1 + \alpha_{12} \cdot f_{x1x2} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot X_1^n$$

$$\frac{dX_2^n}{dt} = 2p_2 \cdot f_p \cdot X_2^{n-1} - p_2 \cdot f_p \cdot X_2^n + \left(\alpha_1 + \alpha_{12} \cdot f_{x1x2} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot X_1^n - \left(\alpha_2 + \alpha_{22} \cdot f_{x2z1} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot X_2^n$$

$$\frac{dZ_1^n}{dt} = \left(\alpha_2 + \alpha_{22} \cdot f_{x2z1} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot X_2^n - \left(\beta_1 + \beta_{12} \cdot f_{z1z2} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot Z_1^n$$

$$\frac{dZ_2^n}{dt} = \left(\beta_1 + \beta_{12} \cdot f_{z1z2} \cdot \sum_{j=0}^{N} Z_2^j\right) \cdot Z_1^n - \delta \cdot Z_2^n.$$

Section S4. Experimental Characterization of Clonal CD4 T-Cell Dynamics. As outlined in the main text, a series of experiments using the murine system were performed to analyze the clonal dynamics of transgenic 5C.C7 CD4 T cells, adoptively transferred into B10.A mice and then immunized with 100 μg of PCC/25 μg of LPS (1). The core time series characterizes the expansion and contraction following transfer of either 5,000 or 500,000 cells. Six percent of the cells homed to lymph nodes, establishing initial ("precursor") populations of 3×10^2 and 3×10^4 responding cells, respectively. The experimental conditions ensured that antigen availability was not a limiting factor. BrdU- and CFSE-labeling procedures were used to characterize the dependence of cell cycling and division age structure on precursor number and on the time since response initiation. The effects of adjuvants, cytokines (IL-1, IL-2, IL-7, and IL-15), and Fas ligands on expansion were examined to explore the robustness of the observed regularities. The most interesting regularity was an approximate square-root law for the dependence of FE on PN. For further data details we refer the reader to the original source (1).

Section S5. Data Assimilation and Parameter Identification. The process of combining diverse information into a unified and consistent description of a physical system is known as data assimilation (2). Here, to consistently integrate the heterogeneous sampled data into the core mathematical model of the system's dynamics, we optimize a "misfit function," which expresses the distance between the whole range of observations and the corresponding model estimates. The vector of 15 model parameters (see Table S1 for definitions), $\mathbf{p}$, $\mathbf{p} \in R^L$, $L = 15$, is $\mathbf{p} = [p_1, p_2, d, \gamma, c, \alpha_1, \alpha_{12}, \alpha_2, \alpha_{22}, \beta_1, \beta_{12}, \theta_{x1}, \theta_{x2}, \theta_z, \delta]$. For global adjustment, we search for a vector of best-fit parameters $\mathbf{p}^*$ following a maximum-likelihood approach as described previously (3).

Four blocks of data characterize different aspects of clonal dynamics. These blocks, along with the constitutive relationships between the observables and the model state variables, are formally specified below:

The time course of clonal expansion and contraction (Table S3): We denote the corresponding observation pairs for two different initial numbers of adoptively transferred cells as $\{t_i^j, C_i^j\}_{i=1}^5, j = 1, 2$, $\{N_0^j\}_{j=1}^2$, where the total size of the antigen-specific clone at times t_i^j is the sum

$$C(t_i^j, N_0^j; \mathbf{p}) = X_1(t_i^j, N_0^j; \mathbf{p}) + X_2(t_i^j, N_0^j; \mathbf{p}) + Z_1(t_i^j, N_0^j; \mathbf{p}) + Z_2(t_i^j, N_0^j; \mathbf{p}).$$

The size of the clone at day 7 (the expansion magnitude) for different initial numbers of transferred cells (the "dose-effect" type data) $\{C^j\}_{j=1}^2$ (Table S6, first row): The observed characteristics are related to the model variables at day 7 as follows:

$$C(t, N_0^j; \mathbf{p})|_{t=7} = X_1(t, N_0^j; \mathbf{p}) + X_2(t, N_0^j; \mathbf{p}) + Z_1(t, N_0^j; \mathbf{p}) + Z_2(t, N_0^j; \mathbf{p})|_{t=7}.$$

The cell-cycling fraction (% BrdU+) at times $\{t_i^j\}_{i=1,j=1}^{5,2}$ (Table S4): Denote the data pairs $\{t_i^j, \boldsymbol{\varphi}_i^j\}_{i=1}^5, j = 1, 2$, with the proportion of cells that get labeled after 6 h BrdU pulse defined as follows:

$$\boldsymbol{\varphi}(t_i^j, N_0^j; \mathbf{p}) = \frac{X_1^L(t_i^j, N_0^j; \mathbf{p}) + X_2^L(t_i^j, N_0^j; \mathbf{p}) + Z_1^L(t_i^j, N_0^j; \mathbf{p}) + Z_2^L(t_i^j, N_0^j; \mathbf{p})}{C(t_i^j, N_0^j; \mathbf{p})}.$$

CFSE dilution (Table S5) and factor of expansion (Table S6, last row) are used for validation of the model rather than for parameter estimation.

The observation data blocks enter the misfit function as separate additive terms: $\Phi_{\text{time course}}(\mathbf{p}) = \sum_{j=1}^{2} \sum_{i=1}^{5} (\log C_i^j - \log C(t_i^j, N_0^j; \mathbf{p}))^2$, $\Phi_{\text{BrdU}}(\mathbf{p}) \equiv \sum_{j=1}^{2} \sum_{i=1}^{5} (\log \boldsymbol{\varphi}_i^j - \log \boldsymbol{\varphi}(t_i^j, N_0^j; \mathbf{p}))^2$, and $\Phi_{\text{magnitude}}(\mathbf{p}) = \sum_{j=1}^{2} (\log C_{d7}^j - \log C(7, N_0^j; \mathbf{p}))^2$, respectively, so that the total misfit between the available observations and the model predictions is expressed by

$$\Phi(\mathbf{p}) = \Phi_{\text{time course}}(\mathbf{p}) + \Phi_{\text{magnitude}}(\mathbf{p}) + w \cdot \Phi_{\text{BrdU}}(\mathbf{p}).$$

Here w is some scaling constant introduced to give appropriate weight to BrdU data, which are in percentages rather than cell numbers. The value $w = 10^2$ was found appropriate to achieve uniform consistency of the model with both the clonal kinetics and the labeling data.

The maximum-likelihood approach to parameter estimation requires knowledge of the statistical distribution of the observation errors. The available data vary substantially in terms of sample sizes, which is common for multiparameter studies of lymphocyte dynamics. The measurements of clonal sizes at day 7 postimmunization (Tables S3 and S6) are numerous enough to apply the normality test. The log-transformed values of cell numbers measured at day 7 postimmunization for different frequencies of adoptively transferred cells (3, 30, 300, and 30,000 cells homing to LNs) all passed the Kolmogorov–Smirnov normality test using GraphPad Prism

version 4 software (http://www.graphpad.com). Assuming further that the errors in successive observations are independent and the variance of observation errors is constant for all observation times and for different numbers of transferred cells, the maximum-likelihood approach can be used for parameter estimation (3). We sought a vector of best-fit parameters $\mathbf{p}^*$ by maximizing the likelihood function $L(\mathbf{p})$ specifying the probability of obtaining the observed data given the model. Under the above assumptions this is equivalent to minimizing the value of the composite function $\Phi(\mathbf{p}) = \Phi_{\text{time course}}(\mathbf{p}) + \Phi_{\text{magnitude}}(\mathbf{p}) + w \cdot \Phi_{\text{BrdU}}(\mathbf{p})$. Whereas the first two terms specify the misfit between model and data in terms of cell numbers, the third one is in terms of the weighted fraction of labeled cells. The 95% confidence intervals (CIs) were calculated via the profile-likelihood method (4). Parameter estimation and CI analysis were carried out using MATLAB 7.0 routines (http://www.mathworks.com).

Section S6. Sensitivity Analysis. The analysis was performed using a sampling-based sensitivity index, the partial rank correlation coefficient (PRCC) (5) (the code is available at http://malthus.micro.med.umich.edu/lab/usadata/) combined with the Latin hypercube sampling (LHS) method. The PRCC determines the statistical relationships between each model parameter (input variable) and the specified outcome variable, with the sign indicating the qualitative relation and the magnitude quantifying the effect of uncertainty in parameter value on the imprecision in predicting the model output value (6). We considered the factor of expansion as integrative model output. The LHS was implemented using triangle probability density functions for the model parameters with the peak value being the best-fit estimate. The ranges were taken to be the intervals specified by ±50% of the best-fit parameter values. The sample sizes were 10^4. The correlation coefficients computed for the low- and high-precursor number cases are shown in Table S7.

The most critical parameters appeared to be the proliferation rate of nondifferentiated cells p_1 and the terminal-differentiation shaping parameters, α_{22}, β_{12}, and θ_z. A few parameters, such as γ, c, and d, display a weak linkage to the factor of expansion. Interestingly, the sensitivity of FE to the proliferation rate of more differentiated cells p_2 changes its sign with increasing PN.

The uncertainty in the estimates of some parameters is rather large (Table S1), also indicating that the information content of the data imposes a restriction on the complexity of the model if variance of the estimated parameters is an issue. The straightforward candidates for elimination were the nonlinear feedback regulation functions. The corresponding choice is justified by careful analysis of the parameter values entering the above functions. The optimized value of the misfit function, $\Phi(\mathbf{p})$, for the reduced model at the computed minimum increased only by 10% compared with the complete version of the model (*Section S5*).

Section S7. Information-Theoretic and Statistical Comparison of the Models. In general, biological model architectures are not comprehensively justified on the basis of proven mechanisms, but are merely parsimonious characterizations of the system under study (7). The two versions of the mathematical model derived in this article are not the only plausible models for studying clonal dynamics in vivo. Other formulations may be preferable given a different context or different goals. It is important not only to evaluate plausible models with respect to their consistency with the data as measured by the values of the "misfit function" (described above in *Section S5*), but also to assess their distance, called "information loss," from an unknown "true model" underlying the specific data set. The modern information-theoretic framework provides a basis for such assessment of information loss and for the ranking of different mathematical models (8).

One of the most powerful criteria for comparing the parsimony of models given the maximum-likelihood estimation of their misfit function is the Akaike information-loss criterion (AIC) (8),

$$\mu_{\text{AIC}} = n_d \ln\left(\Phi\left(\mathbf{p}^*\right)\right) + 2(L+1) + \frac{2(L+1)(L+2)}{n_d - L - 2},$$

where n_d is the total number of scalar measurements used for parameter estimation and L is the number of parameters in the model. Models with a larger value of the AIC are less consistent with the data. For the basic model developed above the relative information loss (distance to the true model of the data) is equal to $\mu_{\text{AIC}} \sim 237$. This value provides a scale for comparing the parsimony of other mathematical models of the given data sets should they be formulated, for example, models with fewer parameters.

The Akaike index for the simplified model is reduced substantially, $\mu_{\text{AIC}} \sim 123$, and this reduction rewards the model for parsimony. The computed 95% confidence intervals are presented in Table S2. The variance of the parameter estimates is much smaller than that of the parameters of the more complete model (Table S1), which is consistent with the smaller value of the Akaike index. Obviously, the more complete model pays a price for having extra free parameters in view of the limited information content of the considered datasets.

The reduced mathematical model (which we refer to as R) is nested within the more general model (denoted G). They have a different number of parameters, $L_{\text{R}} = 10$ and $L_{\text{G}} = 15$, respectively. The model with more parameters fits the data better than the simpler model, the relative difference in the value of the objective function $\Phi(\mathbf{p}^*)$ being ~10%. The important question in comparing the models is whether the difference in the fit in terms of the sum of squared differences between the data and the model solution is statistically significant. Model discrimination for nested models is based upon standard hypothesis tests, such as the F-test (9). Consider the null hypothesis $\text{H}_0 : \gamma = \theta_z = 0, \theta_{x_2} = \infty$. Note that under these conditions $f_p = 1$, $f_{x_2 z_1} = 1$, $f_{z_1 z_2} = 0$ and the model does not depend on the parameters c and β_{12}. The test statistic to check the validity of this hypothesis is defined as $F = \frac{\Phi_R - \Phi_G}{\Phi_G} \cdot \frac{n_d - L_G}{L_G - L_R}$. Under the null hypothesis that the reduced model does not provide a significantly better fit than the general model, we calculate from the data the estimate of $F = 0.143$ and compare it with the critical value of the F-distribution at the level of significance 0.05 with (5, 6) degrees of freedom, which is $F_{0.05,5,6} = 4.39$. As we have $F << F_{0.05,5,6}$, the null hypothesis is rejected. This result suggests that the set of processes parameterized in the reduced model R, being a subhypothesis of model G, is sufficient to describe the processes underlying the considered data sets.

Accordingly, given equal biological plausibility and utility to inform future experiments, the more parsimonious model ranks higher. Importantly, such ranking is far from providing a proof of validity, which can be approached only when a critical body of biological data becomes available.

1. Quiel J, et al. (2011) Antigen -stimulated CD4 T-cell expansion is inversely and log-linearly related to precursor number. *Proc Natl Acad Sci USA*, 10.1073/pnas.1018525.
2. Daley R (1997) Atmospheric data assimilation. *J Meteorol Soc Jpn* 75:319–329.
3. Bard Y (1974) *Nonlinear Parameter Estimation* (Academic, New York).
4. Venzon DJ, Moolgavkar SH (1988) A method for computing profile-likelihood-based confidence intervals. *Appl Stat* 37:87–94.
5. Marino S, Hogue IB, Ray CJ, Kirschner DE (2008) A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J Theor Biol* 254:178–196.
6. Blower SM, Dowlatabadi H (1994) Sensitivity and uncertainty analysis of complex models of disease transmission: An HIV model, as an example. *Int Stat Rev* 62:229–243.
7. Wood SN, Thomas MB (1999) Super-sensitivity to structure in biological models. *Proc R Soc Lond B Biol Sci* 266:565–570.
8. Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference—A Practical Information-Theoretic Approach* (Springer, New York), 2nd Ed.
9. Horn R (1987) Statistical methods for model discrimination. Applications to gating kinetics and permeation of the acetylcholine receptor channel. *Biophys J* 51:255–263.

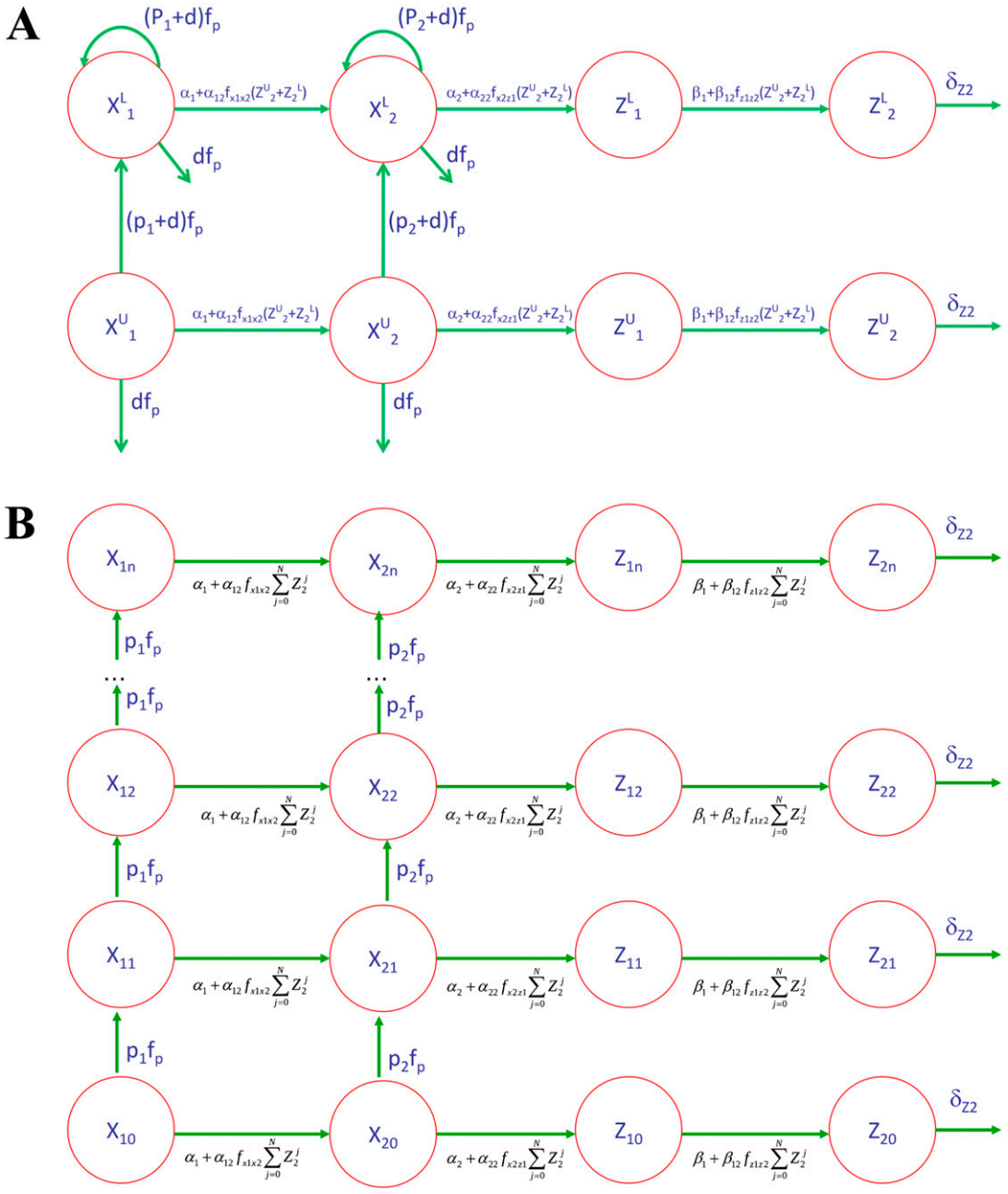

Fig. S1. Biological schemes of the extended versions of the core model for BrdU and CFSE data assimilation. (*A*) Adaptation of the model for BrdU-labeling data analysis. (*B*) Adaptation of the model for CFSE dilution.

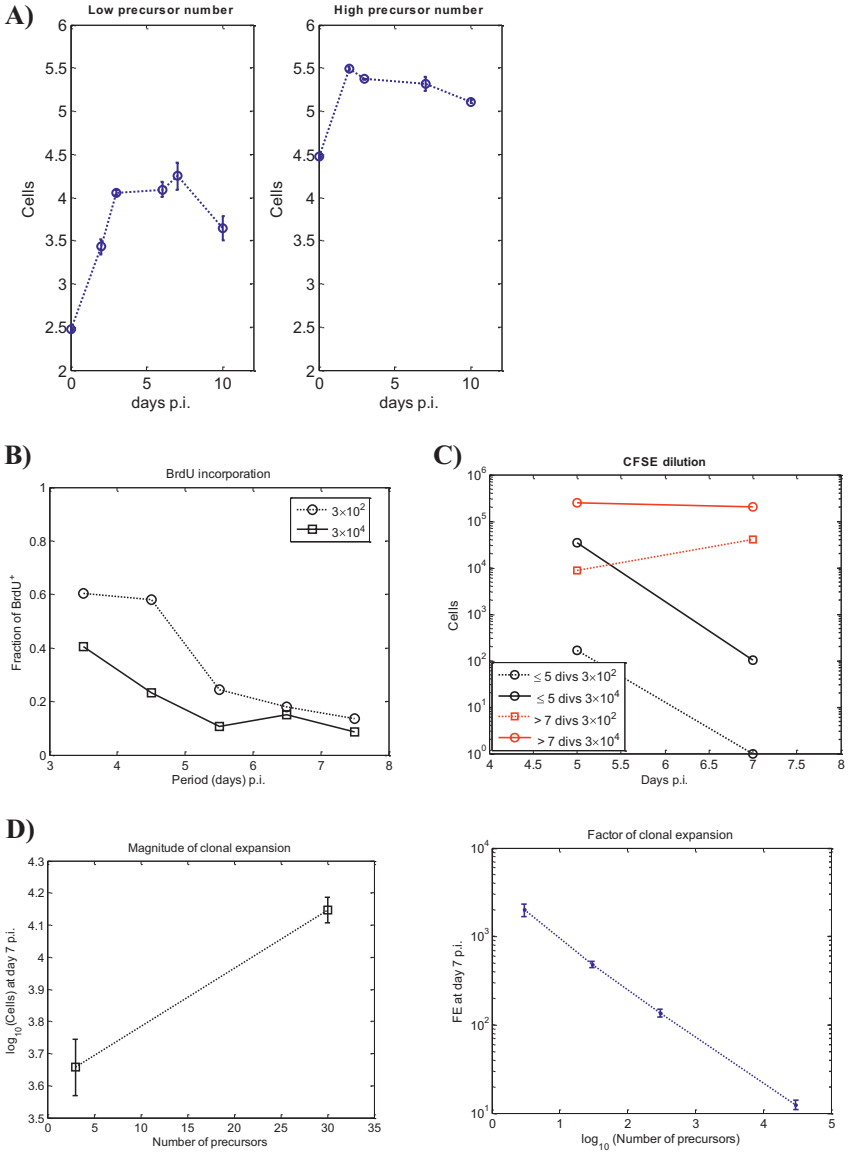

Fig. S2. Observation data: Blocks of experimental datasets used for model identification and validation. (*A*) The overall kinetics of T-cell expansion in lymph nodes of mice for 300 and 30,000 of antigen-specific CD4 T cells homing to the LNs. (*B*) The proportion of BrdU incorporation for 300 and 30,000 CD4 T precursors. (*C*) Proliferation profile via CFSE-label dilution. (*D*) The magnitude and factor of expansion at day 7.

Table S1. Model parameters, their best-fit estimates, and 95% confidence intervals

Parameter	Notation (units)	Best-fit estimate	95% confidence interval
Net proliferation rate of CD4 T cells at initial stage X_1	p_1 (d^{-1})	0.81	[0.80*, 1.12]
Net proliferation rate of CD4 T cells at transit stage X_2	p_2 (d^{-1})	2.0*	[1.0, 2.0*]
Death rate of proliferating CD4 T cells	d (d^{-1})	0.95	[0.45, 1.0]
Maximal effect parameter of cell crowding on proliferation rate	γ	5.0×10^{-4}	[0, ∞)
Threshold cell number for the onset of the effect of cell crowding on proliferation rate	c (cell)	2.8×10^{5}	(0, ∞)
Constitutive maturation rate of proliferating cells	α_1 (d^{-1})	0.0038	[0, 0.23]
Rate constant of the feedback-regulated maturation of proliferating cells	α_{12} (d^{-1}·$cell^{-1}$)	0.06	[10^{-4}, >10^{10}]
Constitutive differentiation rate of proliferating cells into nonproliferating stage	α_2 (d^{-1})	0.001*	[10^{-3}, >1.0]*
Rate constant of the feedback-regulated differentiation of proliferating cells into nonproliferating cells	α_{22} (d^{-1}·$cell^{-1}$)	0.023	[3×10^{-4}, >10^{10}]
Constitutive terminal differentiation rate of nonproliferating cells	β_1 (d^{-1})	0.5	[10^{-3}, >1.0]*
Rate constant of the feedback-regulated terminal differentiation of nonproliferating cells	β_{12} (d^{-1}·$cell^{-1}$)	9.0×10^{1}	[0, ∞)
Threshold nonproliferating cell number for the onset of the saturation in the maturation rate of proliferating cells	θ_{x1} (cell)	2.4×10^{1}	[<10^{-10}, ∞)
Threshold nonproliferating cell number for the onset of the saturation in the differentiation rate of proliferating cells	θ_{x2} (cell)	3.5×10^{2}	[<10^{-10}, ∞)
Threshold nonproliferating cell number for the onset of the saturation in the maturation rate of nonproliferating cells	θ_z (cell)	8.9×10^{-2}	(0, ∞)
Death/migration rate of terminally differentiated cells	δ (d^{-1})	0.59	[0.21, 1.25]

*Lower or upper bound constrained estimation.

Table S2. Parameters of the parsimonious model with their best-fit estimates and 95% confidence intervals

Parameter	Notation (units)	Best-fit estimate	95% confidence interval
Net proliferation rate of CD4 T cells at initial stage X_1	p_1 (d^{-1})	0.80*	[0.80*, 1.02]
Net proliferation rate of CD4 T cells at transit stage X_2	p_2 (d^{-1})	2.0*	[1.17, 2.0*]
Death rate of proliferating CD4 T cells	d (d^{-1})	0.99	[0.41, 1.0*]
Constitutive maturation rate of proliferating cells	α_1 (d^{-1})	0.0059	[0.0018, 0.104]
Rate constant of the feedback-regulated maturation of proliferating cells	α_{12} (d^{-1}·$cell^{-1}$)	0.012	[9×10^{-4}, 90]
Constitutive differentiation rate of proliferating cells into nonproliferating stage	α_2 (d^{-1})	0.001*	[10^{-3}*, 0.35]
Rate constant of the feedback-regulated differentiation of proliferating cells into nonproliferating cells	α_{22} (d^{-1}·$cell^{-1}$)	0.045	[4×10^{-3}, 0.25]
Constitutive terminal differentiation rate of nonproliferating cells	β_1 (d^{-1})	0.52	[0.15, 1.0*]
Threshold nonproliferating cell number for the onset of the saturation in the maturation rate of proliferating cells	θ_{x1} (cell)	842	[0.05, 5.2×10^{3}]
Death/immigration rate of terminally differentiated cells	δ (d^{-1})	3.45	[0.51, 17.2]

*Lower or upper bound constrained estimation.

Table S3. Clonal expansion and contraction of transgenic 5C.C7 CD4 T cells adoptively transferred into B10.A mice (either 5,000 or 500,000 cells) and then immunized with 100 μg PCC/25 μg LPS

Time after immunization, d	No. of antigen-specific $CD4^+$ T cells in LNs	
0	3×10^{2}	3×10^{4}
2	2.70×10^{3} [1.12, 6.51] $\times 10^{3}$	3.066×10^{5} [2.44, 3.86] $\times 10^{5}$
(SS = 3)	3.43 ± 0.089 [3.05, 3.81]	5.49 ± 0.023 [5.39, 5.59]
3	1.14×10^{4} [0.53, 2.43] $\times 10^{4}$	2.35×10^{5} [1.93, 2.86] $\times 10^{5}$
(SS = 2)	4.06 ± 0.026 [3.73, 4.39]	5.37 ± 0.007 [5.29, 5.46]
6	1.24×10^{4} [0.11, 14.22] $\times 10^{4}$	—
(SS = 2)	4.09 ± 0.083 [3.03, 5.15]	
7	1.77×10^{4} [0.81, 3.85] $\times 10^{4}$	2.08×10^{5} [1.39, 3.12] $\times 10^{5}$
(SS = 16)	4.247 ± 0.159 [3.91, 4.59]	5.318 ± 0.083 [5.14, 5.49]
(SS = 17)		
10	4.40×10^{3} [0.07, 262.0] $\times 10^{3}$	1.28×10^{5} [0.78, 2.11] $\times 10^{5}$
(SS = 2)	3.64 ± 0.14 [1.87, 5.41]	5.11 ± 0.017 [4.89, 5.33]

The data are given both as geometric mean (GM) (with 95% confidence intervals) and arithmetic mean (AM) ± SEM of the log of the individual measurements with sample sizes (SS) ranging from 2 to 17 depending on the time.

Table S4. Proportion of BrdU incorporation during clonal expansion and contraction of transgenic 5C.C7 CD4 T cells adoptively transferred into B10.A mice (either 5,000 or 500,000 cells) and then immunized with 100 μg PCC/25 μg LPS

	No. of antigen-specific $CD4^+$ T cells in LNs	
Time after immunization, d	3×10^2, % BrdU+	3×10^4, % BrdU+
3.5	60.5	40.4
4.5	58.1	23.2
5.5	24.5	10.7
6.5	17.9	15.1
7.5	13.5	8.5

The data represent the arithmetic mean (±SEM) of the individual measurements from five mice.

Table S5. CFSE dilution during clonal expansion and contraction of transgenic 5C.C7 CD4 T cells adoptively transferred into B10.A mice (either 5,000 or 500,000 cells) and then immunized with 100 μg PCC/25 μg LPS

	No. of antigen-specific $CD4^+$ T cells in LNs			
	3×10^2		3×10^4	
Time after immunization, d	CFSE 0–5 divisions	CFSE >7 divisions (%)	CFSE 0–5 divisions	CFSE >7 divisions (%)
5	1.63×10^2	8.821×10^3 (98)	3.439×10^4	2.44×10^5 (87)
7	0	4.043×10^4 (100)	1.01×10^2	2.00×10^5 (99)

Table S6. Clonal expansion size evaluated at day 7 postimmunization with 100 μg PCC/25 μg LPS of transgenic 5C.C7 CD4 T cells adoptively transferred into B10.A mice

	No. of antigen-specific $CD4^+$ T cells in LNs			
Clonal expansion	3 (SS = 20)	3×10 (SS = 10)	3×10^2 (SS = 44)	3×10^4 (SS = 18)
Magnitude (cells)	4.541×10^3 [2.968, 6.946] $\times 10^3$ 3.657 ± 0.088 (GM)	1.402×10^4 [1.14, 1.72] $\times 10^4$ 4.147 ± 0.04 (GM)	—	—
Factor of expansion	1.514×10^3 [0.99, 2.315] $\times 10^3$ 3.657 ± 0.203 (GM)	4.67×10^2 [3.81, 5.74] $\times 10^2$ 6.147 ± 0.091 (GM)	(AM) $1.357 \times 10^2 \pm 0.139 \times 10^2$	(AM) $1.25 \times 10^1 \pm 0.16 \times 10^1$

AM, arithmetic mean; GM, geometric mean; SS, sample size.

Table S7. Latin hypercube sampling-based analysis of the partial rank correlation between the model parameters and the FE for low and high PN

Model parameter	Correlation coefficient, 3×10^2 cells	Correlation coefficient, 3×10^4 cells
p_1	0.91	0.91
p_2	−0.27	0.12
d	−0.007	−0.0009
γ	−0.006	0.002
c	−0.006	−0.009
α_1	−0.25	−0.19
α_{12}	−0.39	−0.76
α_2	−0.14	−0.11
α_{22}	−0.56	−0.47
β_1	−0.24	−0.28
β_{12}	−0.66	−0.66
θ_{x1}	−0.37	−0.74
θ_{x2}	−0.39	−0.39
θ_z	−0.63	−0.65
δ	−0.17	−0.34