## Text S2: The Two Phase Model

We propose that an interaction with agonist peptide present at frequency $t$ triggers $T_{reg}$ development during Phase A but deletion in Phase B, when the same peptide is capable of inducing a stronger downstream TCR signal. The adjusted probabilities of a TCR-pMHC encounter within the four selecting regions are $((1-t)\alpha_1, (1-t)\alpha_2, (1-t)\alpha_3+t, (1-t)\alpha_4)$ in phase A, and $((1-t)\beta_1, (1-t)\beta_2, (1-t)\beta_3, (1-t)\beta_4+t)$ in phase B. Each cell has up to $n_A$ and $n_B$ encounters in the respective phases.

The probabilities of selection into each lineage as a function of agonist frequency $t$ are then

$$\begin{aligned}
P[T_{conv}(t)] &= (1-t)^{n_A+n_B}(\alpha_1+\alpha_2)^{n_A}(\beta_1+\beta_2)^{n_B} - (1-t)^{n_A+n_B}\alpha_1^{n_A}\beta_1^{n_B} \\
P[T_{reg}(t)] &= [(1-t)(1-\alpha_4)+t]^{n_A}\,(1-t)^{n_B}(1-\beta_4)^{n_B} \\
&\quad -(1-t)^{n_A+n_B}(\alpha_1+\alpha_2)^{n_A}(\beta_1+\beta_2)^{n_B} \\
&= ((1-t)^{n_A}(1-\alpha_4)^{n_A} + [(1-t)(1-\alpha_4)+t]^{n_A} - (1-t)^{n_A}(1-\alpha_4)^{n_A}) \\
&\quad \times(1-t)^{n_B}(1-\beta_4)^{n_B} - (1-t)^{n_A+n_B}(\alpha_1+\alpha_2)^{n_A}(\beta_1+\beta_2)^{n_B}, \\
&= (1-t)^{n_A+n_B}(1-\alpha_4)^{n_A}(1-\beta_4)^{n_B} - (1-t)^{n_A+n_B}(\alpha_1+\alpha_2)^{n_A}(\beta_1+\beta_2)^{n_B} \\
&\quad +(1-t)^{n_B}(1-\alpha_4)^{n_A}(1-\beta_4)^{n_B}\left[\left(1-t+\frac{t}{1-\alpha_4}\right)^{n_A} - (1-t)^{n_A}\right].
\end{aligned}$$

By definition,

$$P[T_{conv}(WT)] = (\alpha_1+\alpha_2)^{n_A}(\beta_1+\beta_2)^{n_B} - \alpha_1^{n_A}\beta_1^{n_B}, \quad \text{(S2-1)}$$

$$P[T_{reg}(WT)] = (1-\alpha_4)^{n_A}(1-\beta_4)^{n_B} - (\alpha_1+\alpha_2)^{n_A}(\beta_1+\beta_2)^{n_B} \quad \text{(S2-2)}$$

and if we assume that failure to positively select is negligible for $AND^+$ T cells (*i.e.* $\alpha_1^{n_A}\beta_1^{n_B} \ll 1$), then $(1-\alpha_4)^{n_A}(1-\beta_4)^{n_B} \approx P[T_{conv}(WT)] + P[T_{reg}(WT)]$. Thus six unknown parameters $\beta_{1,2,3}$ and $\alpha_{1,2,3}$ are removed from the calculation, and the model becomes:

$$\begin{aligned}
P[T_{conv}(t)] &= (1-t)^{n_A+n_B}P[T_{conv}(WT)] \\
P[T_{reg}(t)] &= (1-t)^{n_A+n_B}P[T_{reg}(WT)] + \\
&\quad (1-t)^{n_B}(P[T_{conv}(WT)] + P[T_{reg}(WT)])\left[\left(1-t+\frac{t}{1-\alpha_4}\right)^{n_A} - (1-t)^{n_A}\right].
\end{aligned}$$

So we obtain a prediction for conventional and regulatory T cell numbers as a function of agonist frequency, number of interactions in each phase, the probability of encountering a negatively selecting ligand in Phase A, and wild-type conventional and regulatory T cell numbers. These parameters are insensitive to the scaling constant derived from the proportion of AND TCR that are negatively selected in controls (see Text S1). The model's predictions do not depend on the temporal order of phase A and phase B, and so allow for TCR sensitivity to either increase or decrease with time during development.