

*Classification:* Biological Sciences/Immunology

## **The Impact of Thymic Antigen Diversity on the Size of the Selected T-cell Repertoire**

*Jose Faro*<sup>1,2</sup>, *Santiago Velasco*<sup>1</sup>, *África González-Fernández*<sup>3</sup>  
*and Antonio Bandeira*<sup>4</sup>

<sup>1</sup>Departamento de Física Aplicada, Universidad de Salamanca, 37008 Salamanca, Spain,

<sup>2</sup>Present address: Instituto Gulbenkian de Ciência, 2781-901 Oeiras, Portugal,

<sup>3</sup>Área de Inmunología, Universidad de Vigo, Lágoas-Marcosende s/n (Vigo), Spain

<sup>4</sup>Unité de Development des Lymphocytes, Institut Pasteur, 25 rue du Dr. Roux,

*F-75724 Paris CEDEX 15, France*

Corresponding author:

Jose FARO  
Estudos Avançados  
Instituto Gulbenkian de Ciência  
Apartado 14  
2781-901 Oeiras  
PORTUGAL  
Tlf: (351) 21 446 46 14  
Fax: (351) 21 440 79 73  
e-mail: jfaro@igc.gulbenkian.pt

*Manuscript Information:* 25 text pages, 4 figures and 3 tables.

*Word and character counts:* 271 words in the abstract; 68804 total characters.

*Abbreviations:* TEC, thymic epithelial cell.

## Abstract

The T-cell receptor (TCR) repertoire of a normal animal is shaped in the thymus by ligand-specific positive and negative selection events. These processes are believed to be determined at the single cell level by the avidity of the TCR/ligand interactions and the duration of TCR signaling. The relationships between the variables involved, in particular the density of TCRs and the number (and density) of different ligands required to select a normal diverse repertoire, are currently unknown, due to the complexity of the interactions involved and the lack of quantitative analysis of those parameters. Here we developed a quantitative model based on the avidity hypothesis of thymic selection, that provides estimates of the fractions of positively and negatively selected thymocytes in the cortex and in the medulla, respectively, as well as general ranges for the number of selecting ligands required for the generation of a normal diverse TCR repertoire. The model predicts counter-intuitive relationships between the key variables that can influence the selection process. Fitting the model to current estimates of positive and negative selected thymocytes leads to specific predictions. The results indicate that: 1) The bulk of thymocyte death takes place in the cortex and it is due to neglect. 2) the probability of a thymocyte to be deleted in the cortex is at least 10–14 fold lower than in the medulla. 3) In the cortex less than 60 ligands are involved in the positive selection process. 4) The moderately high level of negative selection in the medulla is the combined result of a large diversity of selecting ligands on medullary APCs and the increased TCR density on the selected thymocytes.

## Introduction

The thymus plays a central role in the development of T cells (1, 2). There two key processes of TCR repertoire selection take place: first, a broad, random repertoire of TCRs is generated (3); and second, those thymocytes expressing TCRs with weak/medium reactivity binding for self-major histocompatibility complex (MHC) molecules are positively selected (4, 5), while the TCR repertoire is largely purged of cells either not reactive with self-MHC molecules (death by ‘neglect’) or having strong reactivity for self-MHC molecules (negative selection) (6). Those processes take place in two distinct compartments, the cortex and the medulla, as thymocyte migrate (maturing progressively) from the outer cortex (most immature stage) to the cortico-medullary junction to the deep medulla.

The experimental evidence available strongly supports the notion that thymic epithelial cells (TECs) mediate positive selection (7–10), while thymic dendritic cells (DCs) mediate negative selection of thymocytes (11). The cortex is highly enriched for TECs and the cortico-medullary junction and medulla are enriched for DCs (6, 12). Thus, positive selection predominates in the cortex and negative selection is mainly confined to the cortico-medullary junction and the medulla (6, 9–15). The reasons for the poor efficiency of medullary antigen-presenting stromal cells (APCs) in promoting positive selection are not known.

Thymocyte selection relies on specific interactions of the TCR with peptides bound to a selecting MHC protein on the membrane of a thymic APC (16–21). Based on this, it has been proposed that the intrinsic affinity of the TCR to the respective ligand is the key parameter that determines cell fate (22).

An alternative hypothesis is based on the more complex notion that the processes of positive or negative selection are driven by the avidity of the TCR-mediated thymocyte/APC interaction (17–20, 23, 24): thymocytes are rescued from default apoptosis if the avidity of their interaction reaches a given threshold; but they are negatively selected if the avidity exceeds a second threshold. Here the avidity of a thymocyte/APC interaction is assumed to be directly proportional to the density of specific MHC-peptides, the TCR density, and the affinity of binding. For this model it makes then sense to consider the efficiency of selection mediated by single ligands expressed at relatively high copy number versus a peptide mixture adding-up to similar triggering thresholds.

To be built upon these two models is the fact that the TCR specificity is not as stringent as previously thought, that is, a given TCR has a degree of cross-reactivity with distinct MHC-peptides (12, 25). Its estimation, however, has been a matter of debate, because the number of different ligands recognized by a single TCR varies several orders of magnitude depending on the experimental methodology used (26). As a consequence, it is still an open question the minimal number of different MHC-peptides capable to select a normal, diverse mature T cell repertoire (21, 25, 27). At the core of this problem is the consideration that if increased ligand diversity promotes selection of a more diverse TCR repertoire, by the same token negative selection should also increase and therefore there is no obvious advantage over single ligand-mediated selection.

On the other hand, different levels of ligand diversity expressed by distinct sets of thymic stroma cells may differentially affect the process of thymic selection. Thus, while medullary APCs (both epithelial and hemopoietic) can express all or most blood-born, potentially antigenic peptides, the peptide repertoire expressed by cortical TECs appears to be significantly more restricted (8, 28, 29). The question then arises: to what extent the number of different MHC-peptides in the thymus determines whether one selection process predominates over the other. More precisely, is cortical positive selection favored by lower or higher numbers of distinct MHC-peptides? How the different densities of TCRs observed in cortical versus medullary thymocytes affect the process of selection?

Finally, recent data indicate thymocytes may need more than one round of positive signalling by APCs in order to complete their maturation (25, 7, 30, 31), but the minimum number of rounds required and the average number of different APCs with which thymocytes interact in the cortex and medulla remain unknown (7).

Thus, a complex relationship among several different parameters determines the selection of a T-cell repertoire. Because of the complexity and non-linearity of the interactions involved, it becomes very difficult to use intuition to assess how each parameter influences the thymic output (both in terms of repertoire and of cell numbers). As with many other biological processes, a more appropriate way to overcome these intricacies is to develop testable mathematical models (32, 33). Recently, Detours et al. developed a model of thymic selection based on affinity-driven T cell selection (34).

They took into account experimentally determined frequencies of peripheral T cell populations in respect to levels of alloreactivity, self-MHC restriction and frequencies of T cells reactive to conventional antigens. After fitting their model to accommodate those data as well as the experimental demonstration that a majority of positively selected thymocytes are eliminated in the thymic medulla, they concluded that thymic selection is driven by  $10^3 - 10^5$  self peptides.

Here we present a quantitative model of thymocyte selection based on the avidity hypothesis, which discriminates the selection events occurring in the cortex and in the medulla. The model is developed independently of experimental data obtained on antigen reactivities of peripheral T cell populations since these studies addressed activation, proliferation or effector function, rather than peripheral T cell survival which may well be the most equivalent to the process of thymic selection.

### **The model**

The model relies on a series of parameters (described in Table 1) and assumptions schematically depicted in figures 1 and 2. They are grouped into two sets. The first deals with the number of thymocyte/APC interactions and the number of selecting MHC-peptides, and it is illustrated in Fig. 1:

- (a1) A TCR-mediated thymocyte/APC interaction involves a mixture of different MHC-peptides on the APC, each at a concentration able to induce either positive or negative signalling.
- (a2) All APCs belonging to a given compartment express the same number of different MHC-peptides (denoted  $n_c$  for cortical APCs and  $n_m$  for medullary APCs). These quantities represent average numbers per APC.
- (a3) All surviving thymocytes interact, on their way through a given compartment, with equal numbers of different APCs that are assumed to express non-overlapping sets of peptides. In the cortex, the number of APCs encountered by a thymocyte is denoted  $r_c$ , and in the medulla  $r_m$ .
- (a4) To be positively selected a thymocyte must receive signals from a minimum number of APCs (denoted  $z_c$  and  $z_m$  for cortical and medullary thymocytes, respectively). (This is a way to modelize the accumulation —and hence the duration— of positive signalling.)

Contrary to the above assumption, it is conceivable that APCs express the same or a largely overlapping set of peptides. We will also consider this particular case.

The second set of assumptions deals with quantitative aspects of the TCR/ligand interaction (Fig. 2):

- (b1) Randomly generated thymocyte populations have a typical distribution with respect to the affinity of their TCRs for any given MHC-peptide ligand (Fig. 2, right panels) which is considered to be the same for both cortical and medullary thymocytes. Though depicted here as a bell-shaped curve its particular shape is irrelevant for the model presented here.
- (b2) The fate of immature thymocytes (whether cortical or medullary) is determined by a TCR-mediated signal. Signal intensities falling between two excitation thresholds ( $T_1$ ,  $T_2$ ) provide a positive stimulus, while a signal intensity higher than the threshold  $T_2$  leads to negative selection.
- (b3) The intensity of a TCR-mediated signal is proportional to the avidity of the corresponding thymocyte/APC interaction. This avidity is directly proportional to the densities of both TCR and MHC-peptide ligand as well as to the affinity constant of their interaction, ie,  $avidity \propto [\text{TCR}] \times [\text{MHC-peptide}] \times K_a$ .
- (b4) The outcome of distinct TCR-mediated interactions of a thymocyte with different MHC-peptides, on the same or on different APCs, are independent events.

One unscapable consequence follows from the assumptions (b2) and (b3): *for given densities of TCRs and MHC-peptides, excitation thresholds  $T_1$  and  $T_2$  define affinity constants  $K_1$  and  $K_2$ , respectively.* This is illustrated in the right panels of Fig. 2. If activation thresholds are measured in terms of the intrinsic affinity of the TCR to a given Ag, they will change proportionally with TCR and/or ligand density. These affinity thresholds divide the frequency distribution of a given thymocyte population into three meaningful areas: a high ( $K_2 < K$ ), a medium ( $K_1 < K < K_2$ ) and a low ( $K < K_1$ ) affinity area. These areas are the probabilities of a thymocyte being negatively, positively and insufficiently signalled, respectively, upon interaction with that Ag. That is, one can assign a probability to each signalling event both in the cortex ( $p_{c-}$ ,  $p_{c+}$  and  $p_{cN}$ ; see Fig. 2, right, upper panel) and in the medulla ( $p_{m-}$ ,  $p_{m+}$  and  $p_{mN}$ ; see Fig. 2, right, lower panel).

Assuming that MHC-peptides are presented at similar densities in both cortical and medullary APCs and that the intrinsic affinity distributions of TCRs are the same in both cortical and medullary thymocyte populations, the fact that cortical immature double positive (DP,  $CD4^+CD8^+$ ) thymocytes express 90% to 98% less TCRs than medullary thymocytes (35, 36) implies that the TCR/ligand affinity values corresponding to the excitation thresholds  $T_1$  and  $T_2$  (that is,  $K_1$  and  $K_2$ ) are higher in the cortex than in the medulla (ie,  $K_1^m < K_1^c$  and  $K_2^m < K_2^c$ ). This is readily apparent if the respective curves of avidity distributions are compared (left panels of Fig. 2).

The relationship  $K_2^m < K_2^c$  implies that the probability of a thymocyte to be negatively selected by a given Ag is lower in the cortex than in the medulla, considering no intrinsic differences in sensitivity to negative signalling between cortical and medullary thymocytes. Similarly, the relationship  $K_1^m < K_1^c$  implies that the probability of death by neglect is lower in the medulla than in the cortex.<sup>1</sup>

With the above assumptions we build a mathematical model (detailed in Appendix A) which allows to analyse the effects of the key parameters on thymocyte selection. Before performing that analysis, in the next section we will, first of all, apply the compartmental structure of the model to existing experimental data of total positive and negative selected fractions of thymocytes, and derive thereof theoretical estimations of fractions of selected thymocytes in the cortex and the medulla, respectively. Using these theoretical estimations, the model will allow us to establish general relationships between the total number of selecting ligands in the cortex and medulla ( $n_c \times r_c$  and  $n_m \times r_m$ ) and the probabilities of negative selection by a single ligand in the cortex and

---

<sup>1</sup> However, affinity thresholds in the cortex could be lower than in the medulla if the maturation of DP cortical thymocytes to single positive (SP,  $CD4^+CD8^-/CD4^-CD8^+$ ) medullary thymocytes is accompanied by a change in the expression level of adhesion or co-receptor molecules modulating the avidity of TCR/MHC-peptide interactions or the quality of the activation signal (37, 38), or if signalling experience, when moderate, induces in thymocytes adaptation of their signalling thresholds (a concept put forward by Grossman and col. (39) called “tunable activation thresholds”). Nevertheless, the ensuing analysis is not based on any presumed relationship between cortical and medullary threshold affinities of thymocytes but rather suggests specific ways to determine it.

medulla, respectively,  $p_{c-}$  and  $p_{m-}$ . Moreover, by applying the model's equations to recent experimental data obtained from mice expressing virtually a single class II MHC-peptide, we obtain minimal estimates for the probabilities of positive signalling in the cortex and the medulla,  $p_{c+}$  and  $p_{m+}$ , respectively.

## Experimental and theoretical constraints

*Fractions of differentially selected thymocytes in the cortex and in the medulla.* Most estimates of the fractions of neglected, positively selected and negatively selected thymocytes do not distinguish between cortical and medullary thymocytes (see third row in Table 2) (14, 40–42). Only one study has determined the fraction of medullary thymocytes that is negatively selected (14; see also Table 2 and Appendix B). About 95% of daily produced thymocytes die in situ (40–42). How much of this cell loss is due to death by neglect or to negative selection remain a matter of debate (9, 14, 15, 31).

The positively and negatively selected fractions of all thymocytes (40–42; Table 2) are related to the corresponding fractions in the cortex and the medulla as indicated in Eqns. (A5). By using the available experimental estimates mentioned above in these equations, we estimated the remaining fractions of thymocytes in the two major thymic compartments (for details see Appendix C). The theoretical values found are shown in Table 2 together with the experimentally derived data. Our calculations indicate that most thymocytes (75% – 90%) die of neglect in the cortex. Indeed, recent experimental data demonstrate that about 80% of cortical TCR<sup>+</sup> thymocytes do not receive any activation signal (as determined by the expression level of the activation markers CD69 and CD5; ref. 31) and thus are likely to die of neglect. Our results also indicate that in the cortex 10% – 20% of thymocytes are positively selected while less than 5% are negatively selected.

Furthermore, of the fraction of thymocytes that were positively selected and escaped cortical deletion, 50%–70% are negatively selected in the medulla, while 25% – 50% will complete their maturation process. This frequency, which is higher than that one might expect given the general belief that negative selection in the thymus should be stringent, may reflect an optimized degree of TCR cross-reactivity (26). On the other hand, although our estimations cannot be more precise as to indicate that in the medulla less than 25% of thymocytes will die of neglect, experimental observations in K14 mice,

which only express MHC class II molecules in cortical TECs and display normal positive selection of CD4 thymocytes but not negative selection (9), indicate that in the medulla thymocytes do not die of neglect, *i.e.*,  $F_{mN} \sim 0\%$ .

In conclusion, in agreement with observations by Laufer et al. (9) and Surh & Sprent (15) our calculations indicate that the bulk of death of thymocytes is due to neglect and mainly occurs in the cortex (75%–90% neglected cells in the cortex *vs.* 5%–14% negatively selected cells in the medulla).

*General boundaries for the total number of selecting Ags in the cortex and the medulla.* According to the present model the cortical and medullary fractions of negatively selected thymocytes are related to the respective probabilities of negative signaling ( $p_{c-}$  and  $p_{m-}$ ) and the respective numbers of selecting ligands ( $n_c \times r_c$  and  $n_m \times r_m$ ) in the following way (see Appendix A):  $1 - F_{c-} = (1 - p_{c-})^{n_c r_c}$  and  $1 - F_{m-} = (1 - p_{m-})^{n_m r_m}$ . The theoretical estimations shown in Table 2 of the fractions of negatively selected thymocytes in the cortex and medulla are, consequently, the basis to estimate the maximum numbers of naturally selecting ligands in the cortex and the medulla. Taking logarithms at both sides of the above two equations and using now the well known approximation  $\log(1 - p) \approx -p$  for  $|p| \ll 1$ , one is led to the following boundaries for the total number of Ags ‘seen’ in normal mice by a positively selected thymocyte in its journey through the cortex ( $n_c \times r_c$ ) and the medulla ( $n_m \times r_m$ ):

$$n_c \times r_c \leq \frac{0.05}{p_{c-}} \quad (1)$$

and

$$\frac{0.7}{p_{m-}} \leq n_m \times r_m \leq \frac{1.2}{p_{m-}}, \quad i.e., \quad n_m \times r_m \approx \frac{1}{p_{m-}} \quad (2)$$

Note that each of these boundaries is determined by only one basic parameter: the probability of negative signalling ( $p_{c-}$  in the cortex and  $p_{m-}$  in the medulla). Thus, this model provides simple formulae (Eqns. (1) and (2)) to estimate general boundaries for the total number of cortical and medullary Ags screened by positively selected thymocytes. Clearly, the maximum value predicted for the number of Ags in the cortex  $n_c \times r_c$  would be relevant only if the probability  $p_{c-}$  is not too low; for instance, Eqn. (1) predicts that for  $p_{c-} \geq 10^{-3}$  the total number of selecting ligands  $n_c \times r_c$  must be lower

than 50, while if  $p_{c-} \geq 10^{-5}$  then  $n_c \times r_c \leq 5000$ . In contrast, the boundaries for  $n_m \times r_m$  (Eqn. (2)) define a quite narrow range, and hence it must be relevant for virtually any value of the probability of negative selection  $p_{m-}$ : for example, if  $p_{m-} = 10^{-4}$ – $10^{-3}$  it follows that medullary thymocytes should interact with  $10^3$ – $10^4$  different ligands. A relationship very similar to that between  $n_m \times r_m$  and  $p_{m-}$  has been found by D. Mason for optimized reactivity in a theoretical study on T-cell crossreactivity (26). In Mason’s model no attempt was made to distinguish between cortical and medullary negative selection. The present result suggests that Mason’s findings reflect medullary selection rather than total thymic selection.

*Estimating minimum values for the probabilities of positive signalling ( $p_{c+}$  and  $p_{m+}$ ) driven by a single peptide.* Recently, different transgenic mice expressing a single class II MHC-peptide complex have been produced (43–45). Using different semi-quantitative approaches, it has been estimated that in these mice the diversity of the positively selected T-cell repertoire ( $F'_+$ ) is 10%–50% of the corresponding value in normal mice (43, 46–48), that is,  $F'_+ = 0.2\% - 2.5\%$ . Applying these values to the model Eqn. A5b and assuming that the MHC density in the thymus is comparable in both single peptide and wild type mice (an aspect which was only directly addressed in the spleen of those animals) it can be estimated that the probabilities of positive signalling in the cortex and the medulla,  $p_{c+}$  and  $p_{m+}$ , respectively, must be in both cases larger than 0.002 (see Appendix D). Note, however, that for  $p_{c+} = 0.002$  the value of the corresponding medullary probability must be  $p_{m+} = 1$ , and vice versa. Since these extreme values are highly unlikely, one can assume safely a minimum value for each probability of 0.01. In conclusion, the model predicts that at least 1% of cortical and medullary thymocytes are positively signalled after interacting with any given MHC-peptide expressed at high density, assuming that the MHC density in the thymus is comparable in both single peptide and wild type mice, an aspect which was only directly addressed in the spleen of those animals.

The value of these estimates, however, is questionable since some of those genetically engineered mouse strains the single peptide occupies about 95% of the MHC molecules. The roles of the most abundant peptide and the few (but diverse) “contaminating” peptides in thymic selection is not known (49).

## Thymic antigenic diversity and thymocyte selection

To understand thymic selection it is important to determine the number of distinct selecting Ags in the cortex and in the medulla. While some researchers suggested that a diverse T-cell repertoire could be selected by just a few Ags (25, 27) —perhaps the MHC molecules themselves—, others believe that many different Ags are necessary for the selection of a diverse set of TCRs (25, 27, 43, 46, 47). This later view was questioned by the intuitive argument that a large variety of selecting Ags increase not only positive selection but, even more so, negative selection. It is our view, however, that it is not possible to predict even qualitatively how those selection processes may vary with increasing numbers of selecting Ags without additional ad hoc assumptions. On equally intuitive grounds it is possible to conceive that the relationship between selection and peptide diversity is not linear and that there exists a “window” of ligand diversity in which positive selection prevails over deletion.

*Impact of thymic antigenic diversity on thymocyte selection.* As shown in Fig. 3, we studied the dependence of the different fractions of surviving or dying cortical and medullary thymocytes on the number of selecting Ags per APC ( $n_c$  and  $n_m$ , respectively) while varying the number of rounds required for positive signaling ( $z_c, z_m$ ). The results show that: (1) The effect of the number of peptides per APC on the fractions of positively selected thymocytes is, in general, biphasic, rising first with decreasing rate, then reaching a maximum and afterwards decreasing steadily (dashed lines in Fig. 3). This behavior is dramatically modulated by parameters such as the required minimum number of positive interactions  $z_c, z_m$  or the probabilities of positive signaling  $p_{c+}, p_{m+}$ . (2) In contrast, the fraction of negatively selected thymocytes increases steadily, with decreasing rate, as the number of peptides per APC increases (solid lines in Fig. 3). As expected from Eqn. (A4), this increase is only modulated by the probabilities of negative signaling  $p_{c-}, p_{m-}$ . We notice that in most of the cases analyzed the impact of the numbers of encountered APCs,  $r_c$  ( $r_m$ ), on thymocyte selection is very similar to that of  $n_c$  ( $n_m$ ) not only for the fraction of negatively selected thymocytes (as expected from the model’s equations) but also for the neglected and positively selected fractions.

An important consequence of the distinct impact of antigenic diversity on the fractions of positively and negatively selected thymocytes is that there is an optimum value

for the number of peptides per APC in respect to the ratios of positively to negatively selected thymocytes in the cortex ( $F_{c+}/F_{c-}$ ) and in the medulla ( $F_{m+}/F_{m-}$ ). This is shown in Fig. 4 as well as the fact that the particular value of the optimum depends on the model parameters.

In conclusion, this analysis provides a rationale for the view that there is a range of peptide diversity in which positive selection increases faster than negative selection with increasing peptide diversity. An optimum for ligand diversity has also been predicted by other authors based on different theoretical grounds (26, 50, 51).

*Estimating normal thymic antigenic diversity in the cortex and the medulla.* Next, the impact of each parameter on the model's behavior was analyzed under the special case of imposing on the model's behavior the requirement to fit all the experimental and the theoretical constraints here determined (Table 2 and Appendix D). In this way we obtained ranges of admissible values for each parameter as a function of the other parameters. This procedure is illustrated in Fig. 3, where the admissible values for  $n_c$  are those for which the curves representing the fractions  $F_{cN}$ ,  $F_{c+}$  and  $F_{c-}$  lay within the horizontal bands labelled, respectively, 'd.a.', '+' and '-' (similarly indicated in the left panels for  $n_m$ ). Note that those bands correspond to the estimated ranges in Table 2. Remarkably, the number of selecting ligands obtained for the cortex (indicated by small squares in Fig. 4, upper panel) is always lower than the corresponding optimum number. In contrast, for the medulla the reverse is true, namely, the number of selecting ligands is always larger than the optimum (grey thick lines in Fig. 4, lower panel). If the optimum for the cortex and the medulla are similar, the above finding suggests that the number of selecting ligands per APC might be lower in the cortex than in the medulla.

The results of a systematic computer analysis using this procedure are summarized in Table 3. The results indicate that, in the cortex, there is a direct correlation between the minimum number of positive-signalling cell interactions required for positive selection ( $z_c$ ) and the number of different selecting peptides per cortical APC ( $n_c$ ). However, there is an inverse correlation between this number,  $n_c$ , and both the average number of interacted cortical APCs ( $r_c$ ) and the probability of positive signalling by a single ligand  $p_{c+}$ . In contrast, these parameters are virtually independent of  $p_{c-}$  for values of this probability equal or lower than  $10^{-4}$ . Conversely, in the medulla, the values of  $n_m$  that

are consistent with experimental observations are quite independent of the parameters  $p_{m+}$  and  $z_m$ , but are highly dependent on the probability of negative signalling  $p_{m-}$  and the number of different APCs,  $r_m$ .

Thus, the present analysis clearly delineates those scenarios that entail large, medium and small numbers of selecting Ags ( $n \times r$ ) in terms of a few key parameters:  $p_{c+}$  and  $z_c$  in the cortex, and  $p_{m-}$  in the medulla.

In engineered mice expressing class II molecules exclusively on cortical TECs (K14 mice) thymic negative selection is severely abrogated. In the cortical TECs of those mice the density of class II molecules is about 10% the value observed in normal mice (T. Laufer, personal communication), and the fraction of SP CD4<sup>+</sup> thymocytes is reduced 2 to 3-fold compared to normal mice (52, and A.B., unpublished observations). In the present model a 10-fold reduction in ligand density in the cortex implies a 10-fold increase in the corresponding threshold affinities and a 10–15 fold decrease in the values of the probabilities  $p_{c+}$  and  $p_{c-}$ . On the other hand, a 2–3 fold reduction of the total fraction of positively selected cells implies that in those mice  $F'_+ = 0.3 \times F_+$  to  $0.5 \times F_+ = 0.007 - 0.025$ . Since in K14 mice there is no negative selection in the medulla, the fraction of total positively selected cells must equal the fraction of positively selected cells in the cortex, that is  $F'_{c+} \approx F'_+ = 0.007 - 0.025$ , which amounts to a 8 to 15-fold reduction of the estimated values for  $F_{c+}$  in normal mice (see Table 2). This prompted us to ask under what parameter regime of the model it verifies the condition that a 10-fold reduction of  $p_{c+}$  and  $p_{c-}$  leads to a 8–15 fold reduction of  $F_{c+}$ . To answer that question, we performed a systematic analysis for the cortex similar to that summarized in Table 3 using the same parameter ranges but this time setting  $p_{c+}$  to 10-fold lower values,  $p_{c-} = 0$  and  $F_{c+} = 0.007 - 0.025$ . The striking results indicate that that condition is only fulfilled for  $p_{c+} < 0.1$ ,  $r_c < 30$ ,  $z_c = 1$  and  $n_c = 1 - 2$ . In summary, the model predicts that in the cortex less than 60 different MHC/peptide ligands are involved in the selection process, while it is compatible with several thousands of selecting ligands in the medulla. In a theoretical work mentioned above, Detours et al. estimated that thymocyte selection is driven by several thousands of peptides (34). As argued above in respect to Mason's model, this result is in line with ours if interpreted as reflecting medullary, but not cortical selection processes, a distinction not taken into consideration by the authors.

*Signalling duration in the cortex and the medulla.* It has been suggested recently that positive selection of cortical thymocytes requires repeated signalling (7, 30, 53). The present model predicts that selection of cortical thymocytes requires a single encounter with APCs. This does not preclude, however, the possibility that this single cell encounter be of long duration such that prolonged signalling takes place. In its present stage this model cannot be more precise on this issue. In the medulla, however, there is no restriction as to the minimum number of positive selection hits.

## **Experimental tests and predictions**

*Heterogeneity of APC and diversity of the peptide repertoire.* To our knowledge it is still not solved the issues on the magnitude of peptide diversity of cortical and medullary APCs, the degree of overlap between the respective peptide repertoires and the heterogeneity among APCs belonging to each compartment. Nevertheless, studies by Rudensky and colleagues on the differential tissue expression of a number of proteases led to the demonstration that the intracellular proteolytic environment of cortical TECs and of hemopoietic-derived APCs, is quite different (54). This indicates that different sets of peptides and/or different concentrations of peptides are likely to be expressed by each of these cell types. This view is further supported by the drastic differences observed in I-A<sup>b</sup> mice between cortical TECs and hemopoietic-derived APCs, which show inverse levels of expression of major peptides like CLIP (high level in TECs) and Eα52 – 58 (high level in BM-derived APCs) (29).

If all APCs within a given compartment express the same set of peptides and with similar densities, this implies that thymocytes interact (though perhaps repeatedly) with only one type of cortical APC and only one type of medullary APC, *ie*,  $r_c, r_m = 1$ . In this case, equations A4 can be approximated by:  $\log(1 - F_{c-}) \approx -p_{c-} \times n_c$  and  $\log(1 - F_{m-}) \approx -p_{m-} \times n_m$ . That is, in this scenario the model predicts a linear correlation between  $\log(1 - F_{c-})$  and  $n_c$  and between  $\log(1 - F_{m-})$  and  $n_m$ , with  $p_{c-}$  and  $p_{m-}$  being the respective proportionality constants. Following this approach,  $p_{c-}$  and  $p_{m-}$  could be estimated from the slopes of the above linear relationships. There are already mouse lines expressing a single class II MHC-peptide on the membrane of all APCs (19, 43). The generation of mouse lines expressing 2 to 10 different peptides on the membrane of all APCs, will allow to test this prediction by measuring the frequencies of

negatively selected thymocytes in the cortex ( $F_{c-}$ ) and the medulla ( $F_{m-}$ ). It will be of particular interest to construct thymic and bone marrow chimeras in which cortical and medullary APCs express sets of peptides differing by one order of magnitude in their respective diversity. This type of experiments, however, may be difficult to interpret if they do not take into account two fundamental aspects. First, the membrane density of total MHC molecules and distinct MHC-peptides must be determined for the various cell types involved in selection. In one study, no correlation was found between MHC expression in thymic epithelial cells (assessed in histological sections) and splenic cells (analyzed by flow cytometry) (29). In addition, in the same study MHC expression was quite heterogenous in the thymus and in some cases only scattered cells in the cortex scored positive. Second, it is possible that selection is influenced by different classes of MHC molecules. For example, if few CD4 T cells are selected under conditions of restricted peptide density or diversity, a role of class I molecules on the selection process must be excluded.

*Testing the avidity hypothesis of thymocyte selection.*

As shown above, for fixed ligand densities one can derive threshold affinities from the avidity thresholds. This is convenient since densities of membrane molecules and affinities are measurable parameters. In a previous section it was shown that an affinity threshold can be mapped to a basic probability, which implies that, for given densities of TCRs and MHC-peptides, the avidity thresholds determine the critical probabilities  $p_{cN}$ ,  $p_{c+}$ ,  $p_{c-}$  in the cortex and  $p_{mN}$ ,  $p_{m+}$ ,  $p_{m-}$  in the medulla. Interestingly, when all APCs in the cortex and the medulla express a single MHC-peptide, the thymocyte fractions  $F_{c+}$ ,  $F_{c-}$  and  $F_{m+}$ ,  $F_{m-}$  reduce to the probabilities  $p_{c+}$ ,  $p_{c-}$  and  $p_{m+}$ ,  $p_{m-}$ , respectively (see Eqns. A3 and A4). Consequently, these basic probabilities may be estimated experimentally as follows: first, the fractions  $F_N$ ,  $F_+$ ,  $F_-$  and  $F_{m-}$  could be measured, following a kinetic method like bromodeoxyuridine pulse-chase type of experiments (see refs. 40–42), in mice expressing a single class II MHC-peptide (19, 29, 43, 46) in either a normal background or in a background that targets the expression of class II MHC molecules only to cortical TECs (14); moreover, using the method in Appendix C the thymocyte fractions  $F_{c+}$ ,  $F_{c-}$  and  $F_{m+}$  could be derived from the previous ones. Since the avidity hypothesis predicts similar values for any particular

peptide, this study conducted with different peptides would be a critical test.

In its present form the avidity hypothesis does not consider developmental changes both in the intracellular signal transducer device coupled to the TCR, and in the up-regulation of co-receptors or co-inhibitors (38, 55). However, although the importance of those changes has been clearly established when comparing thymocytes with peripheral T cells (37, 38), it remains to be shown when cortical and medullary thymocytes are compared.

### *Potential implications for thymic selection during development*

During ontogeny there is a dramatic increase of the thymus, both in terms of stromal cellularity and stromal cell diversity (12). This increase, which very likely correspond in the present model to an increase of  $r_c$  and  $r_m$ , suggests some interesting possibilities that are amenable for experimental testing. For instance, it is still unknown whether the fractions of positively and negatively selected thymocytes reach values comparable to adult ones early during thymus ontogeny. If this were the case then, according to the model, the numbers of Ags per APC in the cortex,  $n_c$ , and the medulla,  $n_m$ , must decrease during thymic ontogeny concomitantly with the increase of  $r_c$  and  $r_m$ , so that their combined potential impact on thymocyte selection is continuously counter-balanced. Alternatively, it might be the case that during the ontogeny of thymic medulla the asymptotic increase of  $r_m$  is unparalleled by the other parameters which remain virtually unchanged. In this scenario the model predicts that, except for large values of  $z_m$ , there would be an initial increase of the fraction of positively selected thymocytes until it reaches a large maximum followed by a monotonous decrease as the size of the medulla ( $r_m$ ) increases further, while the fraction of negatively selected cells will increase steadily until reaching adult values. These fractions will follow a behavior similar to that of the corresponding thymocyte fractions in the right, top panel in Fig. 3, but with  $r_m$  in the horizontal axis, instead of  $n_m$ . We consider this last possibility attractive, since it has been shown that, contrary to the cortex, the medulla forms and expands as a consequence of interactions between TCR<sup>+</sup> thymocytes and stromal cells, a process that continues in mice at least during the first week after birth (12). Moreover, this scenario would explain why thymocyte deletion is defective in neonatal mice (56).

## Concluding remarks

Although the present model is based on a simplified view of thymocyte development, it provides mechanistic boundaries for the selective process in the thymus. Surely, as mentioned above, there are additional biological constraints, some of which developmentally programmed (for instance, CD5 dampens TCR activation signals in thymocytes but enhances them in mature T cells (57)) and some adaptively acquired (38, 39, 58); but they operate very likely within the more general, mechanistic boundaries provided by the present model.

Remarkably, when the model was applied to experimental data it made possible to estimate the different fractions of selected thymocytes in both the cortex and the cortico-medullary junction + the medulla (Table 2). These results indicate that, while 70%–90% of thymocytes die of neglect in the cortex, the probability for a thymocyte to be negatively selected in the cortico-medullary junction + the medulla is at least 10–14 fold higher than in the cortex. Moreover, the present analysis has provided very simple and general formulae for the maximum values of the total number of ligands in the cortex and the medulla (Eqns. (1) and (2)). Finally, a major prediction of the model (obtained after fitting it to experimental data) is that the number of selecting ligands in the cortex is less than 60.

A possibly important shortcoming of the present simulations is that we did not take into account the likely diversity of ligand densities in a normal thymus. The model equations imply a strongly non-linear relationship between ligand densities and the number of selecting ligands required to produce a given frequency of positively selected thymocytes. Because of this, if high density ligands do not have distinguishing features for binding to TCRs compared to other ligands, it is to be expected that thymocyte selection is not influenced in an important manner by very low density ligands (*ie*, being at a density that is lower than 1% of that of high density ligands). In other words, it is expected that thymocyte selection be driven essentially by high and medium density ligands.

In summary, the present model, although based on a rather simplified view of the thymic selection processes, predicts in a testable way particular patterns of complex and often counter-intuitive relationships among measurable parameters that influence the selection processes (Table 3), and suggests a potential scenario for thymocyte selection

during development that can explain the observed defective deletion in neonatal mice. The complexity of that fabric of relationships emphasizes the increasing need of quantitative conceptual approaches amenable to mathematical analysis, complementary to experimental studies of the basic immunological phenomena.

**Acknowledgments.** We thank A. Coutinho J. Carneiro, M.A.R. Marcos for discussions and useful suggestions and W. Haas and Z. Grossman for critical reading and comments on the manuscript. This work was partially supported by the University of Vigo, Xunta de Galicia and the Ministry of Education and Science of Spain.

## Appendix A

The computation of the probabilities that a thymocyte interacting with a cortical APC be negatively signalled,  $f_{c-}$ , and insufficiently signalled,  $f_{cN}$ , follows readily from the independence-of-events assumption, while the probability of being positively signalled ( $f_{c+}$ ) follows from the general relationship  $f_{c+} + f_{c-} + f_{cN} = 1$ . Thus, we have

$$f_{c-} = 1 - (1 - p_{c-})^{n_c}, \quad f_{cN} = p_{cN}^{n_c}, \quad f_{c+} = (1 - p_{c-})^{n_c} - p_{cN}^{n_c}. \quad (A1)$$

where  $p_{cN}$  is the probability that a randomly picked cortical thymocyte interacting with a given Ag be insufficiently signalled, and  $p_{c-}$  the probability that the interaction results in negative selection.

Consider a cortical thymocyte interacting with  $r_c$  APCs. Since these are also independent events, it follows that the probability that such a thymocyte not be negatively selected in the cortex is

$$(f_{cN} + f_{c+})^{r_c} = \sum_{i=0}^{z_c-1} \binom{r_c}{i} f_{cN}^{(r_c-i)} f_{c+}^i + \sum_{i=z_c}^{r_c} \binom{r_c}{i} f_{cN}^{(r_c-i)} f_{c+}^i. \quad (A2)$$

The second term on the right hand side of Eqn. (A2) is the probability that a thymocyte interacting with  $r_c$  APCs experiences at least  $z_c$  cell encounters resulting in positive signalling. This can be identified, therefore, with the fraction of thymocytes positively selected in the cortex,  $F_{c+}$ . In terms of  $p_{cN}$  and  $p_{c-}$  this fraction reads:

$$F_{c+} = \sum_{i=z_c}^{r_c} \binom{r_c}{i} p_{cN}^{n_c(r_c-i)} \left[ (1 - p_{c-})^{n_c} - p_{cN}^{n_c} \right]^i. \quad (A3)$$

Note also that the first term in (A2) can be interpreted as the fraction of thymocytes that die of ‘neglect’ in the cortex (denoted  $F_{cN}$ ). A similar calculation gives for the medullary compartment the corresponding probabilities  $f_{m+}$ ,  $f_{m-}$  and  $f_{mN}$  as well as  $F_{m+}$  and  $F_{mN}$ .

Let’s now denote the fractions of thymocytes that are negatively selected in the cortex and medulla, respectively,  $F_{c-}$  and  $F_{m-}$ . Then, using the general relations  $F_{c+} + F_{cN} + F_{c-} = F_{m+} + F_{mN} + F_{m-} = 1$ , one has:

$$F_{c-} = 1 - (1 - p_{c-})^{n_c r_c} \quad \text{and} \quad F_{m-} = 1 - (1 - p_{m-})^{n_m r_m}. \quad (A4)$$

Finally, from the assumption that positively selected thymocytes move unidirectionally from the cortex to the medulla to periphery, a straightforward computation shows that the total expected fractions (cortex plus medulla) of thymocytes following default apoptosis ( $F_N$ ), positive selection ( $F_+$ ) and negative selection ( $F_-$ ), are related to the above ones through the equations:

$$F_N = F_{cN} + F_{c+} \times F_{mN}, \quad (A5a)$$

$$F_+ = F_{c+} \times F_{m+}, \quad (A5b)$$

$$F_- = F_{c-} + F_{c+} \times F_{m-}. \quad (A5c)$$

## Appendix B

In radiation bone-marrow chimeras reconstituted with cells from genetically engineered mice not expressing class II MHC molecules on bone-marrow derived cells (14), it has been estimated that the total fraction of positively selected thymocytes is  $F'_+ \approx 0.10 - 0.15$  (that is, 2 to 3 times its value in normal mice ( $F_+$ )). Assuming that in those chimeric mice  $F'_{m-} = 0$  and  $F'_{m+} \approx F_{m+}/(F_{m+} + F_{mN})$  (where  $F_{m+}$  and  $F_{mN}$  are frequencies in normal mice) and that the corresponding thymocyte frequencies in the cortex are unaffected (i.e.,  $F'_{c+} = F_{c+}$  and  $F'_{cN} = F_{cN}$ ), and using Eq. A5 one has  $F'_+ = F'_{c+} F'_{m+} = F_{c+} F_{m+}/(F_{m+} + F_{mN}) = F_+/(1 - F_{m-})$ . From this relationship and the experimental estimates of  $F'_+$  and  $F_+$  it follows directly that

$$0.5 \leq F_{m-} < 0.7. \quad (B1)$$

## Appendix C

The derivation of estimates for the different fractions in the cortex and medulla follows from a straightforward algebraic manipulation of inequalities. Thus, from Eqns. (A5c) and (B1) one has:  $0.1 > F_{c+} \times F_{m-} > F_{c+} \times 0.5$  which implies:

$$F_{c+} < 0.2. \quad (C1)$$

From Eqn. (B1) one has  $F_{mN}, F_{m+} < 0.5$ , which together with Eqn. (A5b) yields:  $0.05 \approx F_+ = F_{c+} \times F_{m+} < F_{c+} \times 0.5$ , and therefore

$$0.1 < F_{c+}. \quad (C2)$$

From Eqns. (A5b) and (C3) one has:  $0.05 \approx F_+ = F_{c_+} \times F_{m_+} < 0.2 \times F_{m_+}$ , and therefore

$$0.25 < F_{m_+}. \quad (C3)$$

Eqns. (B1) and (C6) yield:  $0.5 > F_{m_N} + F_{m_+} > F_{m_N} + 0.25$ , which implies:

$$F_{m_N} < 0.25. \quad (C4)$$

Using Eqns. (B1), (C1) and (C2) into Eqn. (A5c) one has,  $0.1 \geq F_- = F_{c_-} + F_{c_+} \times F_{m_-} > F_{c_-} + 0.1 \times 0.5$  and therefore

$$F_{c_-} \leq 0.05 \quad (C5)$$

This, together with Eqns. (C1) and (C2) and the fact that  $F_{c_N} + F_{c_+} + F_{c_-} = 1$ , imply:

$$0.75 < F_{c_N} < 0.9. \quad (C6)$$

### Appendix D

In normal mice  $F_+ = F_{c_+} \times F_{m_+} = 0.02 - 0.05$  (40–42). Since in transgenic mice expressing a single class II MHC-peptide it verifies  $n_c, n_m = 1$  and  $r_c, r_m = 1$ , it follows that  $F'_{c_+} = p_{c_+}$  and  $F'_{m_+} = p_{m_+}$ . On the other hand, if  $F'_+ (= F'_{c_+} \times F'_{m_+}) \approx 10\% - 50\%$  of  $F_+$  in normal mice, that is, if  $F'_+ = 0.002 - 0.025$ , then  $p_{c_+} \times p_{m_+} \geq 0.002$ , and hence  $p_{c_+}, p_{m_+} \geq 0.002$ .

## REFERENCES

1. Burnet, M. (1969) *Cellular Immunology* (Cambridge University Press, Cambridge).
2. Jerne, N.K. (1971) *Eur. J. Immun.* **1**, 1–9.
3. Fehling, H.J. & von Boehmer, H. (1997) *Curr. Opin. Immun.* **9**, 263–275.
4. Bevan, M.J. (1977) *Nature* **269**, 417–418.
5. Zinkernagel, R.M., Callahan, G.N., Althage, A., Cooper, S., Klein, P.A. & Klein, J. (1978) *J. exp. Med.* **147**, 882–896.
6. Anderson, G., Moore, N.C., Owen, J.J.T. & Jenkinson, E.J. (1996) *A. Rev. Immun.* **14**, 73–99.
7. Anderson, G. & Jenkinson, E.J. (1997) *Immun. Today* **18**, 363–364.
8. Fort, M.M. & Pardoll, D.M. (1996) *Cel. Immun.* **171**, 74–79.
9. Laufer, T.M., DeKoning, J., Markowitz, J.S., Lo, D. & Glimcher, L.H. (1996) *Nature* **383**, 81–85.
10. Capone, M., Romagnoli, P., Beermann, F., MacDonald, H.R. & van Meerwijk, J.P. (2001) *Blood*, **97**:1336-1342.
11. Brocker, T., Riedinger, M. & Karjalainen, K. (1997) *J. exp. Med.* **185**, 541–550.
12. Goverman, J., Brabb, T., Huseby, E.S., & Farr, A.G. (1997) *Immun. Today* **18**, 204–208.
13. Nieto, M.-A., González-Fernández, A., López-Rivas, A., Díaz-Espada, F. & Gambón-Deza, F. (1990) *J. Immun.* **145**, 1364–1368.
14. van Meerwijk, J.P.M., Marguerat, S., Lees, R.K., Germain, R.M., Fowlkes, B.J. & MacDonald, H.R. (1997) *J. exp. Med.* **185**, 377–383.
15. Surh, C.D. & Sprent, J. (1994) *Nature* **372**, 100–103.
16. Murphy, K.M., Heimberger, A.B. & Loh, D.Y. (1990) *Science* **250**, 1720–1723.
17. Ashton-Rickardt, Ph.G., Bandeira, A., Delaney, J.R., van Kaer, L., Pircher, H.-P., Zinkernagel, R.M. & Tonegawa, S. (1994) *Cell* **76**, 651–663.
18. Sebzda, E., Wallace, V.A., Mayer, J., Yeung, R.S.M., Mak, T.W. & Ohashi, P. (1994) *Science* **263**, 1615–1618.
19. Fukui, Y., Ishimoto, T., Utsuyama, M., Gytoku, T., Koga, T., Nakao, K., Hirokawa, K., Katsuki, M. & Sasazuki, T. (1997) *Immunity* **6**, 401–410.
20. Cook, J.R., Wormstall, E.-M., Hornell, T., Russell, J., Connolly, J.M. & Hansen, T.H. (1997) *Immunity* **7**, 233–241.

21. Jameson, S.C., Hogqvist, K.A. & Bevan, M. (1995) *A. Rev. Immun.* **13**, 93–126.
22. Alam, S.M., Travers, P.J., Wung, J.L., Nasholds, W., Redpath, S., Jameson, S.C., & Gascoigne, N.R.J. (1996) *Nature* **381**, 616–620.
23. Suzuki, H., Guinter, T.I., Koyasu, S., & Singer, A. 19(98) *Eur. J. Immunol.* **28**:3252–3258.
24. Legge, K.L., Min, B., Pack, C., Caprio, J., & Zaghoulani, H. (1999) *J. Immunol.* **162**:5738–5746.
25. Kersh, G.J. & Allen, P.M. (1996) *Nature* **380**, 495–498.
26. Mason, D. (1998) *Immun. Today*. **19**:395.
27. Janeway, C.A. Jr.. (1999) *Cur. Biol.* **9**:R342–R345.
28. Ezine, S. (1989) *Bull. Inst. Pasteur* **87**, 171–202.
29. Farr, A., DeRoos, P.C., Eastman, S. & Rudensky, A.Yu. (1996) *Eur. J. Immun.* **26**, 3185–3193.
30. Savino, W., Dardenne, M. & Carnaud, C. (1996) *Immun. Today* **17**, 97–98.
31. Merckenschlager, M., Graf, D. , Lovatt, M. , Bommhardt, U. , Zamoyska, R. & Fisher, A.G. (1997) *J. exp. Med.* **186**, 1149–1158.
32. Goldbetter, A. (1996) *Biochemical Oscillations and Cellular Rhythms* (Cambridge Univ. Press).
33. Nowak, M.A. & Bangham, C.R.M. (1996) *Science*. **272**, 74–79.
34. Detours, V., Mehr, R. & Perelson, A.S. (2000) *J. Immunol.* **164**:121–128.
35. Guidos, C.J., Danska, J.S., Fathman, C.G., & Weissman, I.L. (1990) *J. exp. Med.* **172**:835–845.
36. van Meerwijk, J.P. & Germain, R.M. (1993) *Science* **261**:911.
37. Lucas, B., Štefanová, I., Yasutomo, Y., Dautigny, N. & Germain, R.N. (1999) *Immunity* **10**:367–376.
38. Germain, R.N. & Štefanová, I. (1999) *A. Rev. Immun.* **17**:467–522.
39. Grossman, Z. & Singer, A. (1996) *Proc. Ntnl. Acad. Sci., U.S.A.* **93**, 14747–14752.
40. Egerton, M., Scollay, R. & Shortman, K. (1990) *Proc. ntnl. Acad. Sci., U.S.A.* **87**, 2579–2582.
41. Huesmann, M., Scott, B., Kisielow, P. & von Boehmer, H. (1991) *Cell* **66**, 533–540.
42. Scollay, R. & Godfrey, D.I. (1995) *Immun. Today* **16**, 268–273.
43. Ignatowicz, L., Kappler, J. & Marrack, P. (1996) *Cell* **84**, 521–529.

44. Fung-Leung, W.-P., Surh, Ch.D., Liljedahl, M., Pang, J., Leturcq, D., Peterson, P.A., Webb, S.R. & Karlsson, L. (1996) *Science* **271**, 1278–1281.
45. Tourne, S., Miyazaki, T, Wolf, P., Ploegh, H., Benoist, C. & Mathis, D. (1997) *Proc. Ntnl. Acad. Sci. U.S.A.* **94**, 9255–9260.
46. Tourne, S., Miyazaki, T, Oxenius, A., Klein, L., Fehr, T., Kyewski, B., Benoist, C. & Mathis, D. (1997) *Immunity* **7**, 187–195.
47. Surh, C.D., Lee, D.-S., Fung-Leung, W.-P., Karlsson, L. & Sprent, J. (1997) *Immunity* **7**, 209–219.
48. Gapin, L, Fukui, Y., Kanellopoulos, J., Sano, T., Casrouge, A., Malier, V., Beau-  
doing, E., Gautheret, D., Claverie, J.-M., Sasazuki, T. & Kourilsky, P. (1998) *J. exp. Med.* **187**:1871–1883.
49. Barton, G.M. & Rudensky, A.Y. (1999) *Science* **283**:67–70.
50. De Boer, R.J. & Perelson, A.S. (1993) *Proc. R. Soc. Lond. B Biol. Sci.* **252**:171-5.
51. Nemazee D. (1996) *Immunol Today*, **17**:25-9.
52. de Bont, E.S., Reilly, C.R., Lo, D., Glimcher, L.H., Laufer, T.M. 1999. *Int. Im-  
munol.* 11:1295-306
53. Anderson, G., Hare, K.J. & Jenkinson, E.J. (1999) *Immun. Today* **20**:463–468.
54. Nakagawa, T., Roth, W., Wong, P., Nelson, A., Farr, A., Deussing, J., Villadangos,  
J.A., Ploegh, H., Peters, C. & Rudensky, A. (1998) *Science* **280**:450–453.
55. Wong, P., Barton, G.M., Forbush, K.A. & Rudensky, A.Y. (2001) *J. exp. Med.*  
**193**:1179–1187.
56. Signorelli, K., Benoist, C. & Mathis, D. (1992) *Eur. J. Immunol.* **22**, 2487–2493.
57. Tarakhovsky, A., Kanner, S.B., Hombach, J., Ledbetter, J.A., Müller, W. Killeen,  
N. & Rajewsky, K. (1995) *Science* **269**:535–537.
58. Ohashi, P.S. (1996) *Curr. Opin. Immun.* **8**, 808–814.

Table 1. Parameters used in the model.

Name	Definition
$p_{cN}, p_{c+}, p_{c-},$ $p_{mN}, p_{m+}, p_{m-}$	Probabilities of TCR/ligand interactions in the cortex ( $p_c$ ) or medulla ( $p_m$ ) leading, respectively, to null, positive or negative signals
$n_c, n_m$	Average number of distinct peptides per cortical and medullar APC, respectively
$r_c, r_m$	Average number of different APCs with which thymocytes interact in their way through the cortex and medulla, respectively
$z_c, z_m$	Minimum number of APCs with which thymocytes must interact positively in order to be positively selected in the cortex and medulla, respectively
$F_{cN}, F_{c+}, F_{c-},$ $F_{mN}, F_{m+}, F_{m-}$	Fractions of neglected, positively selected and negatively selected thymocytes, respectively in the cortex ( $F_c$ ) or medulla ( $F_m$ )
$F_N, F_+, F_-$	Total fractions of neglected, positively selected and negatively selected thymocytes, respectively

Table 2. Experimentally and theoretically derived estimates of cortical, medullary and total fractions of thymocytes that die of neglect and are positively and negatively selected(\*).

Thymic compartment	Fraction of cells Neglected	Fraction of cells Pos. Selected	Fraction of cells Neg. Selected
Cortex	$\frac{F_{cN}}{(0.75 - 0.90)}$	$\frac{F_{c+}}{(0.10 - 0.20)}$	$\frac{F_{c-}}{(< 0.05)}$
Medulla	$\frac{F_{mN}}{(< 0.25)}$	$\frac{F_{m+}}{(0.25 - 0.50)}$	$F_{m-}$ (0.50 - 0.70)‡
Total(†)	$F_N$ (0.85 - 0.90)	$F_+$ (0.02 - 0.05)	$F_-$ (0.05 - 0.10)

(\* )Theoretical estimates are underlined (see Appendix C).

(†)From references 40–42.

(‡)From reference 14 (see Appendix B).

Table 3. Predicted ranges for the average number of Ags per APC in the cortex and the medulla ( $n_c$  and  $n_m$ , respectively) for different values of the other parameters.

Thymic compartment	Positive signalling probability	Negative signalling probability	Minimum no. of pos. signalling rounds	Numbers of selecting Ags compatible with experimental data					
				$n_c$					
Cortex*	$p_{c+}$	$p_{c-}$	$z_c$	$r_c = 10$	$r_c = 50$	$r_c = 100$			
				$10^{-2}$	$\leq 10^{-4}$	1	1 – 2	< 1	< 1
						5	31 – 40	5 – 7	2 – 3
	10	157 – 190	14 – 16†			6 – 8			
	$10^{-1}$	$\leq 10^{-4}$	1	< 1	< 1	< 1			
			5	3 – 4	< 1	< 1			
10			15 – 18	1 – 2	< 1				
Medulla	$p_{m+}$	$p_{m-}$	$z_m$	$n_m$					
				$r_m = 10$	$r_m = 50$	$r_m = 100$			
				$\leq 0.2$	$10^{-3}$	1 to 10	70 – 120	14 – 25	10 – 12
				$\leq 0.2$	$10^{-4}$	1 to 10	693 – 1204	139 – 241	69 – 120

\*For  $r_c = 10$ ,  $p_{c+} = 10^{-2}$  and  $p_{c-} = 10^{-4}$  the maximum value of  $z_c$  compatible with experimental observations is  $z_c = 6$ , and the corresponding range for  $n_c$  is: 43–50. †Also, for  $r_c = 50$ ,  $p_{c+} = 10^{-2}$  and  $p_{c-} = 10^{-4}$  the maximum value of  $z_c$  compatible with experimental observations is  $z_c = 8$ , and the corresponding value for  $n_c$  is 10.

## Figure legends.

Figure 1. The model of thymocyte selection. Cortical and medullary thymocytes interact on average with  $r_c$  and  $r_m$  APCs, respectively (in the figure,  $r_c = 2$  and  $r_m = 3$ ). Cortical and medullary APCs express in their membranes on average  $n_c$  and  $n_m$ , respectively, different self-MHC-peptide complexes (in the figure  $n_c = 3$  and  $n_m = 5$ ; these values are intended only for the sake of clarity). *APC*, antigen presenting stromal cell; *TCR*, T cell receptor; *DP*, double positive ( $CD4^+CD8^+$ ) thymocyte; *SP*, single positive ( $CD4^+CD8^-$  or  $CD4^-CD8^+$ ) thymocyte.

Figure 2. Assumed frequency distribution of a randomly generated thymocyte population with respect to the avidity of TCR/MHC-peptide-based thymocyte/APC interaction (left panels) and the affinity (right panels) of their TCRs for any MHC-peptide; the distributions for different Ags are further assumed to have similar means and widths. Left panels,  $T_1$ ,  $T_2$ , lower and upper avidity (excitation) thresholds for positive signalling of cortical and medullary thymocytes. Note that because of lower TCR density in cortical *vs* medullary thymocytes, the curve in the cortical compartment is shifted to the left compared to curve in the medullary one. Right panels,  $K_1^c$  and  $K_2^c$ ,  $K_1^m$  and  $K_2^m$ , are affinity thresholds for cortical and medullary thymocytes, respectively, corresponding to the generic avidity thresholds  $T_1$  and  $T_2$  and lower densities of TCRs in cortical *vs.* medullary cells.  $p_{c_N}$ ,  $p_{c_+}$  and  $p_{c_-}$  in the cortex and  $p_{m_N}$ ,  $p_{m_+}$  and  $p_{m_-}$  in the medulla are the probabilities of the three signalling events defined, respectively, by the affinity thresholds  $K_1^c$ ,  $K_2^c$  and  $K_1^m$ ,  $K_2^m$ .

Figure 3. Impact of model parameters on the frequencies of selected thymocytes. Dependence of the distinct fractions of dying and surviving thymocytes in the cortex (left panels) and the medulla (right panels), on the number of selecting Ags per APC ( $n_c$  in the cortex, and  $n_m$  in the medulla) and on the minimal number of positive signalling rounds that a thymocyte needs in order to be positively selected ( $z_c$  in the cortex, and  $z_m$  in the medulla). Dashed lines, plots of the expected frequencies of neglected thymocytes in the cortex and medulla ( $F_{cN}$  and  $F_{mN}$ , respectively); dotted lines, plots of the expected frequencies of positively selected thymocytes in the cortex and medulla ( $F_{c+}$  and  $F_{m+}$ , respectively); solid lines, plots of the expected frequencies of negatively

selected thymocytes in the cortex and medulla ( $F_{c-}$  and  $F_{m-}$ , respectively). Horizontal bands labelled d.a. (light gray), + (white) and - (darker gray) indicate, respectively, the theoretically estimated frequency ranges of default apoptosed, positively and negatively selected thymocytes (see Table 2). Ranges of admissible values for the numbers of selecting ligands ( $n_c$ , and  $n_m$ ) are those where the 3 curves fall within their corresponding horizontal bands. Notice that, as expected from Eqs. A4, the parameters  $z_c$  and  $z_m$  have no effect at all on negative selection (solid lines) in the cortex and the medulla, respectively. Other parameter values were as follows:  $r_c = r_m = 50$  APC,  $p_{c+} = 10^{-2}$ ,  $p_{c-} = 10^{-4}$ ,  $p_{m+} = 2 \cdot 10^{-2}$ ,  $p_{m-} = 10^{-3}$ . Note that for these parameter values the cases  $z_c = 1$ ,  $z_m = 1$  and  $z_m = 10$ , but not the case  $z_c = 10$ , have ranges of values of the number of ligands compatible with experimental data (indicated by a thick line at the bottom of the graph).

Figure 4. The ratio of positive to negative selected thymocyte fractions as a function of the number of selecting peptides per APC. Both in the cortex and the medulla this ratio has (in most cases) a biphasic behaviour. For increasing diversity of peptides per APC, this ratio first rises, in general from below 1, until a maximum is attained (in which the fraction of positively selected thymocytes can be as high as 20-fold or more the fraction of negatively selected thymocytes) and then starts to decrease steadily to values below 1. The fact that parameters  $z_c$  and  $z_m$  can have an important impact on that maximum is illustrated here by showing the ratio of the cortical and medullary fractions for two values of  $z_c$  and  $z_m$ , respectively. Horizontal lines correspond to a ratio of 1, i.e., when  $F_{c+} = F_{c-}$  and  $F_{m+} = F_{m-}$ . Gray bands indicate the curve regions that fit data in Table 2. Other parameters are as in fig. 3.

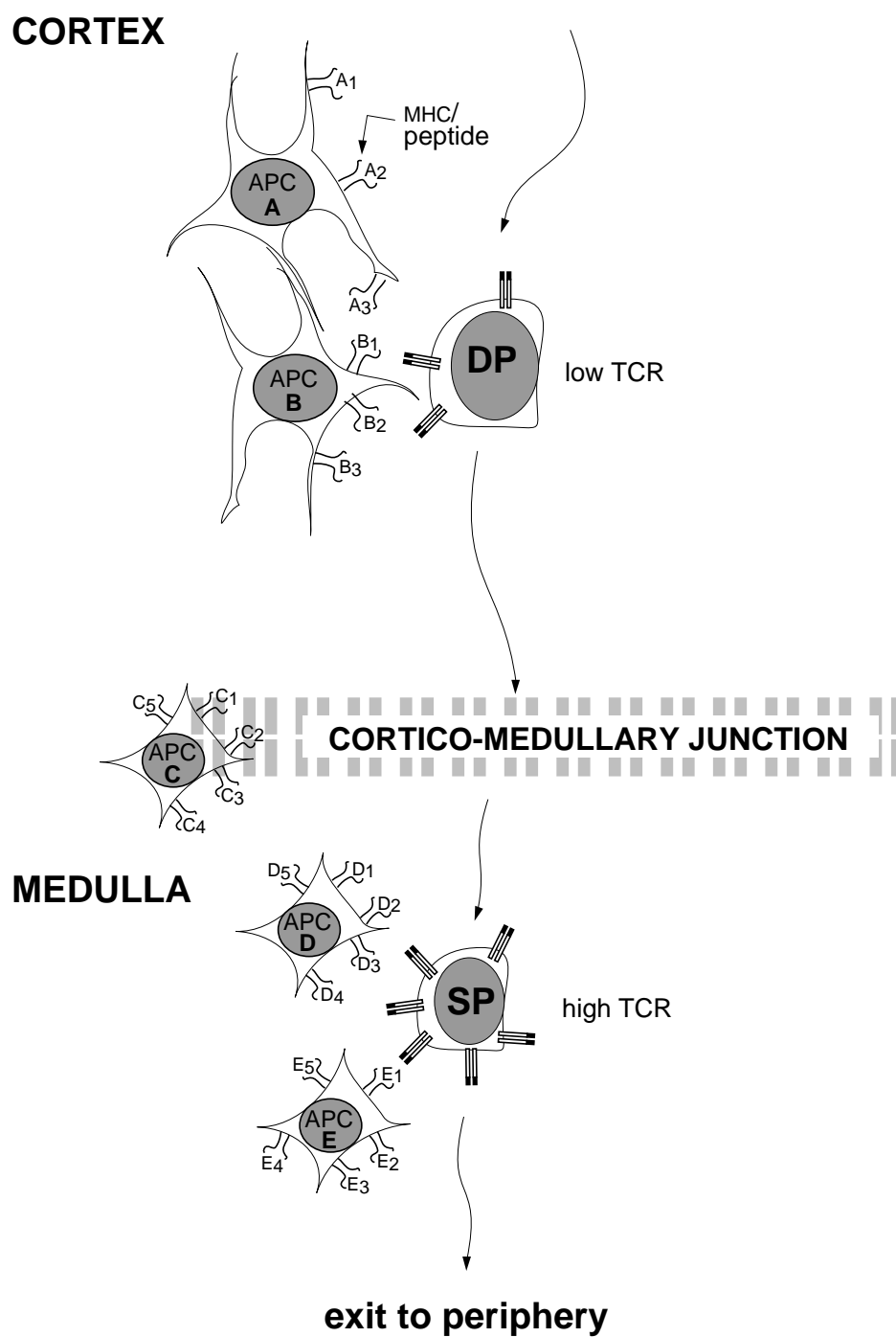


Fig. 1

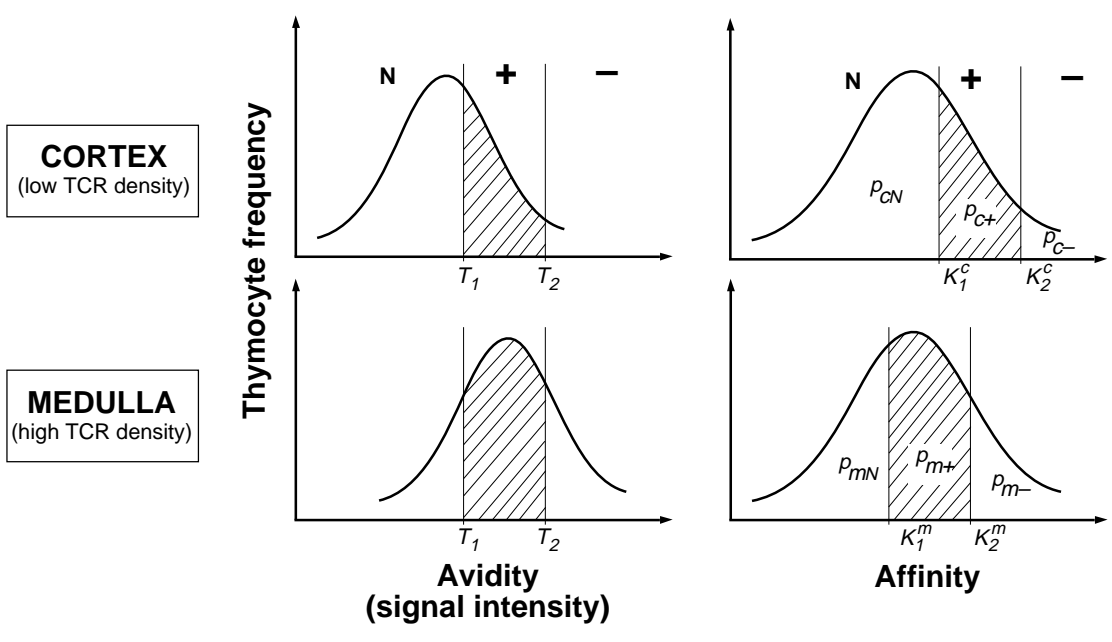


Fig. 2

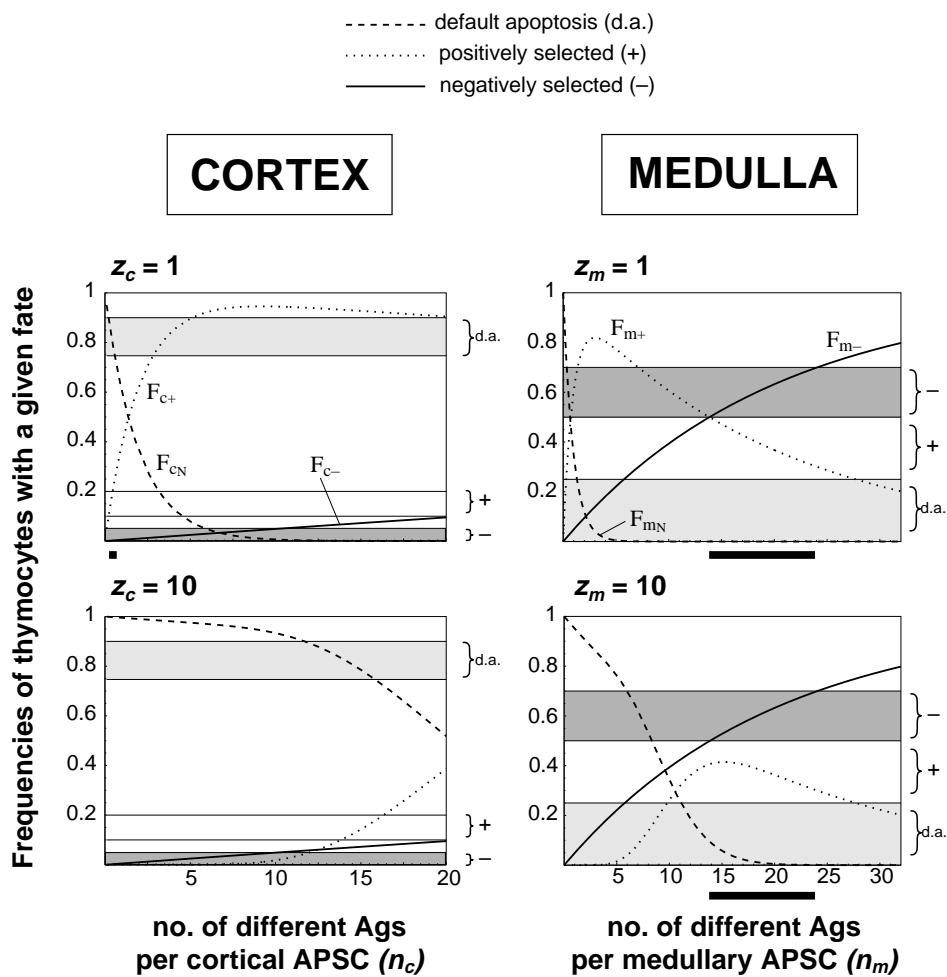


Fig. 3

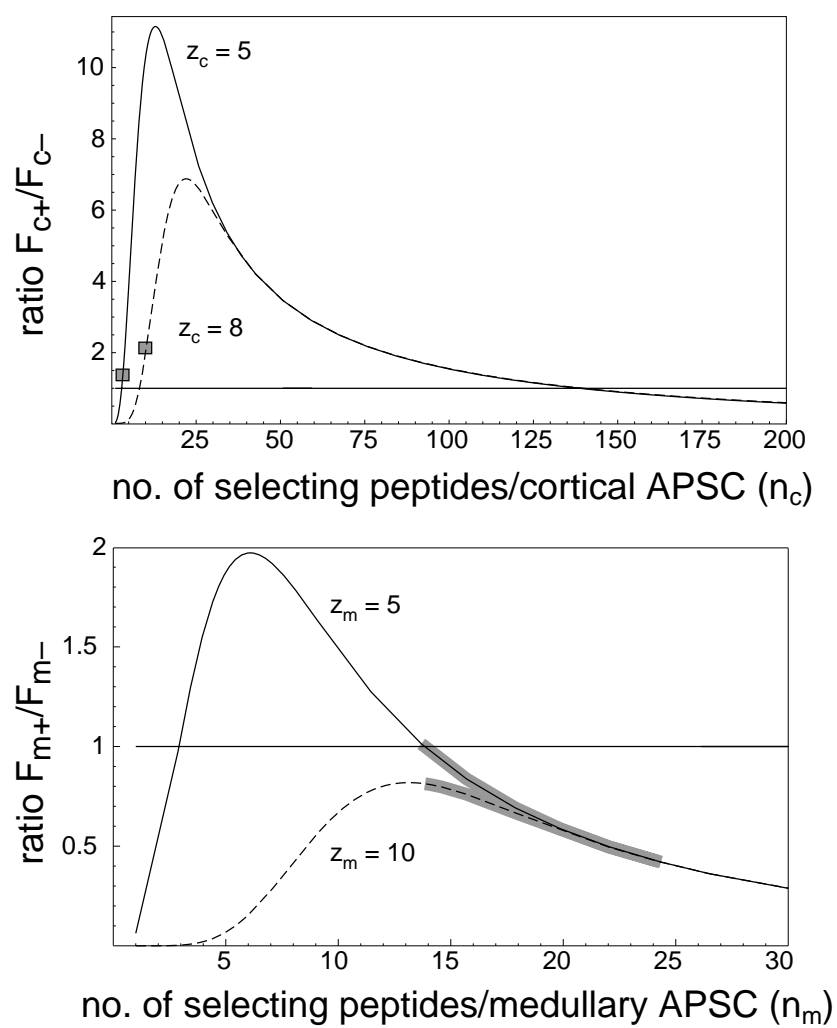


Fig. 4