Immunity, Vol. 14, 21–32, January, 2001, Copyright ©2001 by Cell Press

The Shaping of the T Cell Repertoire

Margarida Correia-Neves,*‡ Caroline Waltzinger,* Diane Mathis,*§ and Christophe Benoist*†§
*Institut de Génétique et de Biologie Moléculaire et Cellulaire (CNRS/INSERM/ULP)
1 rue Laurent Fries
67404 Strasbourg
France

Summary

By combining a TCRβ transgene with a TCRα minilocus comprised of a single V and two J gene segments, we engineered a mouse line exhibiting ample but focused TCR diversity, restricted to CDR3α. Using single-cell PCR and high-throughput sequencing, we have exploited this system to scrutinize T cell repertoire selection and evolution. Some striking observations emerged: (1) thymic selection produces a repertoire that is very "bumpy," with marked overrepresentation of a subset of sequences; (2) MHC class I– and class II–restricted TCRs can be distinguished by minute, single-residue changes in CDR3α; and (3) homeostatic expansion and survival in the periphery can markedly remold the postselection repertoire, likely reflecting variability in the potential of cells displaying different TCRs to respond to homeostatic cues.

Introduction

Essentially all cells in multicellular organisms exchange signals, typically monomolecular and structurally invariant. T lymphocytes differ, however, in that the principal cues governing their differentiation and homeostasis are received through a somatically variable receptor that recognizes variable ligands composed of a diversity of small peptides complexed with polymorphic MHC molecules. This mode of cell:cell communication imposes important constraints that predicate some critical features of T cell behavior: MHC restriction (or the matching of an individual's complements of T cells and MHC molecules) and positive selection (or the learning process by which T cells maturing in the thymus acquire this matched repertoire) (Fink and Bevan, 1978; Zinkernagel and Doherty, 1997). Furthermore, MHC molecules may also mold the T cell repertoire postthymically, as these molecules seem necessary for the survival and homeostasis of naive cells in the peripheral lymphoid organs (Goldrath and Bevan, 1999a).

The TCR is composed of an α and a β chain, each encoded by variable, randomly rearranged V and (D)J segments. Variability is concentrated in the CDRs; CDR1 and CDR2 are encoded within the V segments themselves, while CDR3's much greater variability stems from its position at the junction of the V(D)J elements, compounded by their imprecise juxtaposition, with variable deletion and random addition of a few nucleotides at the joined ends. Given the number of V and J segments that randomly combine and the impact of nucleotide deletions and additions, the theoretical diversity of the mouse or human αβ T cell repertoire has been estimated at about 10^{14} (Davis and Bjorkman, 1988). This number represents a "raw" repertoire, before its streamlining by selection for self-MHC compatibility and self-tolerance. The postselection diversity of the peripheral repertoire is what conditions the individual's ability to mount appropriate immune responses or, conversely, to succumb to autoimmunity. Thus, an appreciation of the true extent of this diversity is important.

Defining the T cell repertoire is a daunting task because of its magnitude and of the complexities inherent in such a combinatorial structure. All TCR variable regions can be used by both MHC class I– and class II–reactive TCRs, even though there are preferential associations (Kappler et al., 1987; Jameson et al., 1990; Sim et al., 1996). TCR transgenic (tg) approaches have been very powerful for monitoring the fate of cells expressing TCRs of defined restriction and specificity, thereby allowing us to visualize selection processes (von Boehmer, 1990), but such strategies generally permit the assessment of only a single receptor at a time. PCR-based analyses that resolve TCRs on the basis of CDR3 lengths have addressed the bulk repertoire and have been particularly useful in highlighting recurrent amplifications in the context of immune responses (Pannetier et al., 1995). Some brute-force sequencing studies have attempted to describe the diversity of the VDJ joins of single Vβ regions in thymocytes or in peripheral $CD4^+$ and $CD8^+$ T cells (Candeias et al., 1991; Arstila et al., 1999), but the results have revealed only subtle correlates, probably because of the compensating influences of Vα elements. The difficulty has been side stepped with single-chain tg mice expressing the β chain of an antigen-specific TCR (Sant'Angelo et al., 1997, 1998). A significant proportion of T cells in these animals is reactive to the cognate peptide, and these cells frequently use TCRα chains that are the same as or similar to the parental clone. In addition, the presence or absence of the parental sequence motif can be used as a marker in differentiating thymocytes.

In the present report, we have attempted to obtain a broad but manageable view of the selection and evolution of the T cell repertoire by engineering a mouse line in which T lymphocytes are polyclonal but TCR diversity is focused. This was achieved with a TCRα minilocus transgene composed of a single Vα region and two Jα elements separated by an artificial rearrangement substrate. This transgene was combined with a mutation that inactivates the endogenous TCRα locus and with a TCRβ transgene that, due to the dictates of allelic exclusion, prevents the rearrangement of endogenous

†To whom correspondence should be addressed (e-mail: cb@igbmc.u-strasbg.fr [C. B.], dm@igbmc.u-strasbg.fr [D. M.]).
‡Present address: Instituto de Biologia Molecular e Celular, Porto, 4150, Portugal.
§Present address: Section on Immunology and Immunogenetics, Joslin Diabetes Center, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, One Joslin Place, Boston, Massachusetts 02215.

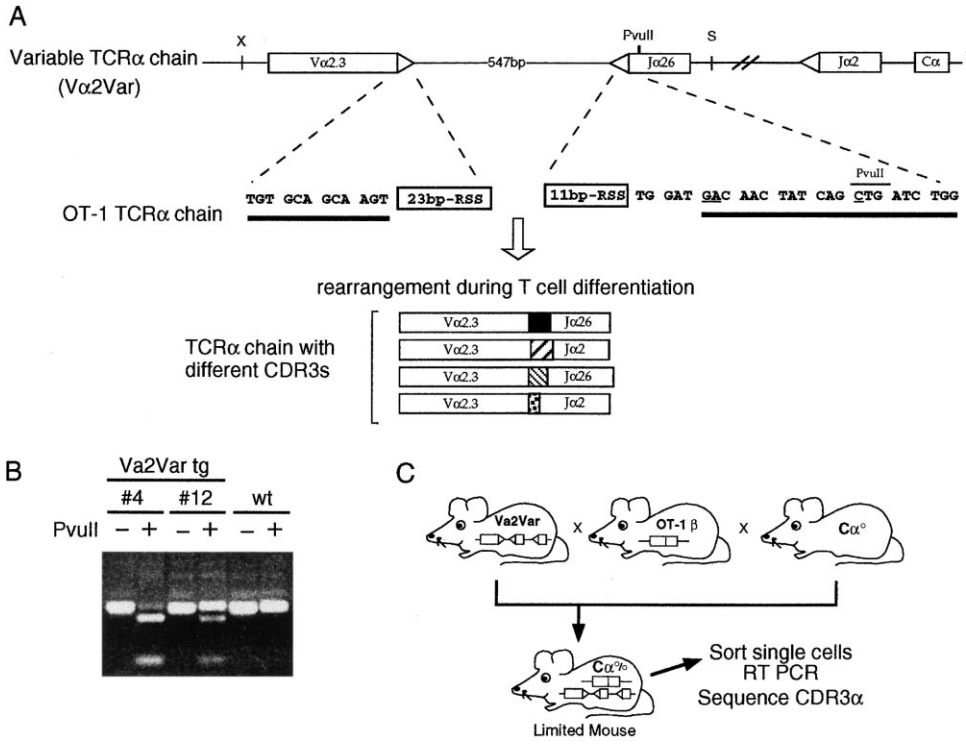


Figure 1. The TCRα Minilocus Used to Generate Mice with Limited TCR Diversity

(A) Va2Var construct. The artificial rearrangement substrate is composed of the Vα2.3 and Jα26 elements from the OT-1 TCRα chain, separated by a stuffer sequence flanked by RSS signals (see Experimental Procedures). This fragment is placed between the XmaI and SacII sites of the pTCRαCass vector (Kouskoff et al., 1993). Rearrangement will lead to multiple CDR3α variants of the same VJ combination.

(B) Total thymic RNA from two independent tg founder lines was amplified by RT–PCR with Vα2.3- and Jα26-specific primers, and the products were electrophoresed with or without prior digestion by PvuII, which distinguishes tg from endogenous transcripts; the band corresponds to amplification of the rearranged VJ segment.

(C) The transgenes present in the Limited mice, combining the Va2Var transgene, the OT-1 TCRβ chain transgene (Correia-Neves et al., 1999), and the TCRα knockout mutation.

TCRβ components. We expected that in the resulting "Limited" mice a repertoire of joining region sequences would be generated that was broad enough that T cells with different MHC restrictions and diverse antigen specificities would be produced but that was constrained enough that influences on repertoire selection would not be too obscured. What proportion of α:β TCR pairs can actually be selected? How much is the repertoire shaped by selection in the thymus versus preferential amplification and maintenance in the periphery? Starting from an MHC class I–restricted TCR, what alterations would permit it to recognize class II molecules?

Results

Our approach to investigating the T lymphocyte repertoire was to generate tg mice in which T cells expressed a diversity of TCRs but with only focused variability. The structure of the diverse TCRs could then be correlated with MHC class specificity, MHC restriction, and antigenic peptide specificity. We started from the well-studied OT-1 TCR, K^b restricted, specific for the OVA257-264 peptide, and composed of Vα2.3Jα26 and Vβ5Jβ2.6 chains (Kelly et al., 1993; Hogquist et al., 1994). The TCRα gene sequence was converted into an artificial rearrangement substrate by inserting canonical rearrangement motifs (RSS) separated by a 550 bp spacer (Figure 1A). This construct does not encode a viable TCR as such but, after VJ rearrangement, should generate a number of variants differing in the junctional CDR3 region. The sequences were chosen such that a direct rearrangement would be out of frame, and so, nucleotide additions or deletions are necessary to generate a functional product. (Regenerating the OT-1 TCRα itself would require a five-base deletion of the J segment.) A silent mutation introduced a novel PvuII site into the Jα region to permit us to distinguish transcripts emanating from the normal germline elements. This DNA fragment was placed into the TCRαCass vector that drives expression in trangenic mice under the normal TCRα regulatory elements and provides an additional Jα2 segment (Kouskoff et al., 1993). This other Jα segment was intended as a control to guard against the results being overly biased by preferential regeneration of the original OT-1 VJ combination.

Several tg founder lines were obtained after injection of this construct into B6×SJL F2 embryos, identified by Southern blot analysis (data not shown). A few lines that harbored only a single transgene integrant were selected to avoid later problems with the cells having rearranged several independent genes. The ability of the transgene to rearrange and be expressed in T cells

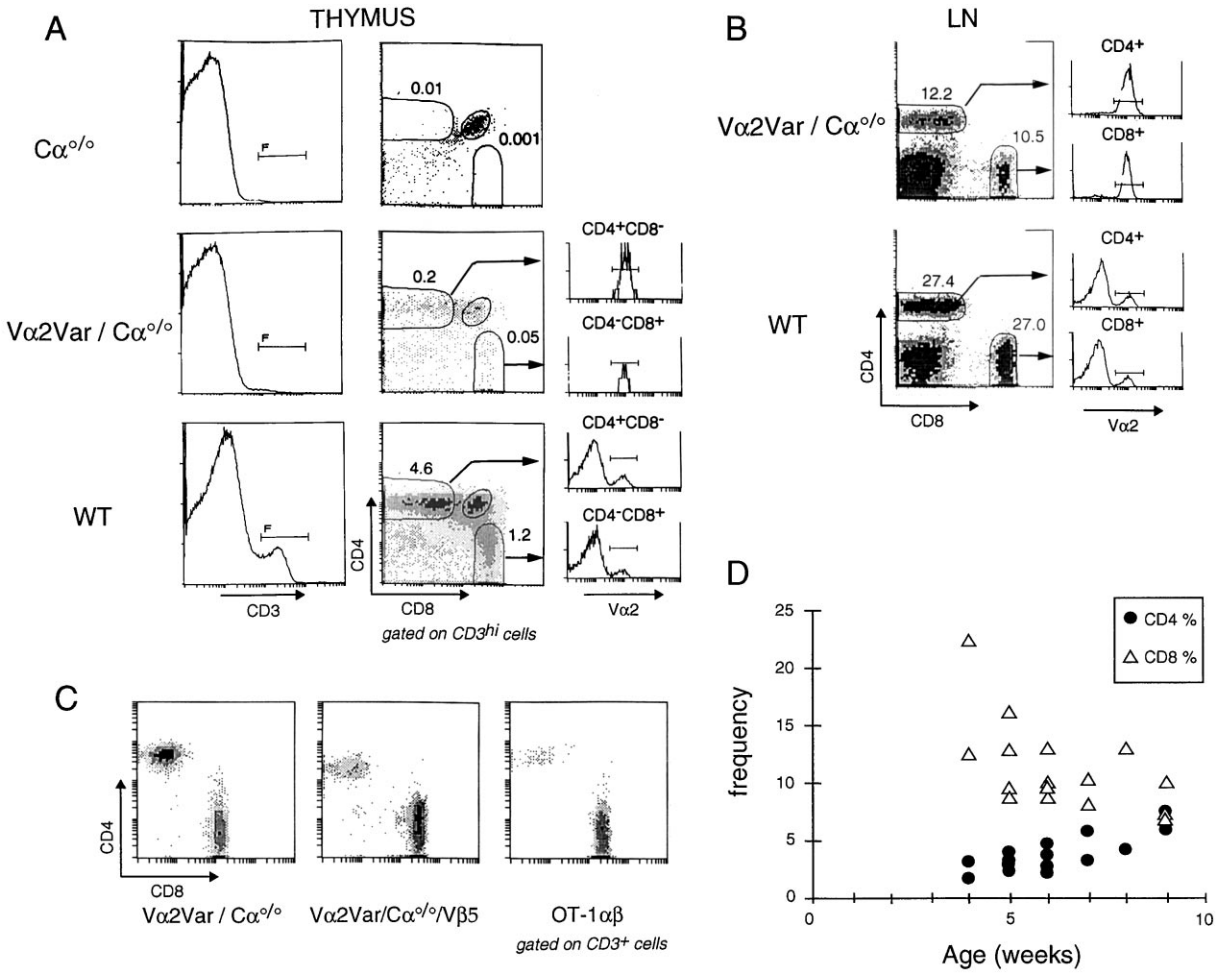


Figure 2. The Va2Var Minilocus Is Functional and Compensates a TCRα Deficiency

(A) Thymic expression. In this representative experiment, thymi from three mice were compared by four-color flow cytometry: a homozygous TCRα knockout ($C\alpha^{o/o}$), a Va2Var transgenic (also $C\alpha^{o/o}$ homozygous), and a wild-type reference mouse. The CD3 histogram (left panels) identifies the $CD3^{hi}$ cells, whose CD4/CD8 profile is shown (middle panels); the staining of gated $CD4^+CD8^-$ and $CD4^-CD8^+$ SPs (when present) for the transgene-encoded Vα2 is shown in the right panel. (Staining in the wild-type sample corresponds to expression of the normal Vα2 family members.)

(B) Presence of T cells of both $CD4^+$ and $CD8^+$ phenotypes in lymph nodes from Va2Var/$C\alpha^{0/0}$ mice.

(C) CD4/CD8 profiles of $CD3^+$ lymph node cells from Va2Var/$C\alpha^{0/0}$ mice, with or without the Vβ5 OT-1 transgene, compared with a tg mouse expressing both OT-1 α and β chain genes.

(D) Evolution of the frequency of $CD4^+$ and $CD8^+$ cells in lymph nodes of young Limited (Va2Var/$C\alpha^{0/0}V\beta5^+$) mice.

was evaluated by RT–PCR analysis of thymus RNA, using primers in the Vα2.3 and Cα regions. Digestion of the amplification products by PvuII revealed the diagnostic shortening of the fragment in some founder mice but not in negative littermates (Figure 1B). This indicated that the artificial rearrangement substrate was indeed functioning. Animals with this Va2Var transgene were then bred appropriately to introduce the OT-1 TCRβ chain transgene (Correia-Neves et al., 1999) and the TCRα knockout mutation (Cα°; Philpott et al., 1992). The resulting animals (hereafter referred to as Limited mice) have T cells whose TCRs exhibit only limited diversity: (1) the TCRα chains derive only from the transgene, with its single Vα and two Jα segments, and therefore diversity is restricted to CDR3; and (2) the transgene-encoded OT-1 TCRβ chain is expressed in >95% of cells, preventing rearrangement of any endogenous β genes, whose expression was undetectable by direct staining (data not shown). The experimental strategy was to then examine cells from different T cell populations by single-cell sorting, RT–PCR amplification, and high-throughput sequencing of the CDR3α stretch. Large databases of such sequences were compiled and analyzed.

We first asked where transgene expression could be detected and what types of T cells could be selected. As previously described (Philpott et al., 1992), no mature $CD4^+8^-$ or $CD4^-8^+$ αβ thymocytes were selected in the thymus of $C\alpha^{0/0}$ controls, in the absence of a functional endogenous TCRα locus (Figure 2A, top). Introduction of the Va2Var transgene promoted selection of distinct populations of $CD4^+CD8^-CD3^+$ and $CD4^-CD8^+CD3^+$ cells (Figure 2A). As expected, all expressed Vα2 (Figure 2A, right). The number of mature CD4 and CD8 lineage

cells averaged 0.5% and 0.2% of total thymocytes, respectively, or ~1/20th the number observed in wild-type control littermates (Figure 2A, bottom). This frequency is quite similar to that of T cells expressing any given single Vα.

On the other hand, consistent with the fact that the Va2Var transgene requires rearrangement for expression, we did not detect the aberrant levels of TCR in $CD4^-CD8^-$ cells that are usually observed with prerearranged TCR transgenes. Activation and maturation markers (CD69, heat shock antigen, and peanut agglutinin) were also normal in the populations undergoing positive selection (data not shown).

Since mature cells of both the CD4 and CD8 single-positive (SP) phenotype were found, the TCRs generated from the artificial substrate could apparently associate with both MHC class I and II molecules. These cells were exported to peripheral lymphoid organs (Figure 2B), resulting in sizeable T cell populations of both phenotypes, with a rough equivalence of $CD4^+$ and $CD8^+$ cells. This ratio was altered somewhat when the OT-1 TCRβ chain transgene was introduced, completing the Limited mouse (Va2Var, Vβ5, and $C\alpha^o$ traits; Figure 2C), with an increase in the relative proportion of $CD8^+$ cells. This increase was expected, as the OT-1 TCR is class I restricted. Interestingly, the proportion of $CD4^+$ cells changed with the age of the Limited mice (Figure 2D). Thus, the combination of transgenes in the Limited mouse seemed operational, generating TCRs with sufficient variability to allow positive selection in the thymus and maintenance in the periphery, directed by either MHC class I or class II molecules.

The Preselection Repertoire

To analyze the TCR diversity generated in preselection populations, we determined the CDR3α sequences of TCRs from immature thymocytes of Limited mice. Small, resting $CD4^+CD8^+$ (double-positive, DP) cells expressing a low but detectable level of surface CD3 (Figure 3A) were examined in order to ensure that the analysis focused on products of the Va2Var transgene able to pair with the Vβ5 chain but without biases induced by positive selection. Cells from three mice were sorted into six independent pools. (In this analysis, RT–PCR and sequencing were performed on independent pools of sorted DPs, rather than on single cells as in all of the subsequent experiments, because the relative rarity of Va2Var rearrangements in DP cells prevented us from obtaining large numbers of sequences by the single-cell approach.) RNA was prepared from each pool, the sequences derived from the Va2Var transgene amplified with primers located in the Vα2 and Cα segments, and the products cloned into pBS plasmid for sequencing. Two hundred and eighty-three independent clones were sequenced, of which an illustrative 95 derived from two independent pools of a single mouse are presented in Figure 3B; the complete set of sequences can be found as supplemental data at http://www.immunity.com/cgi/content/full/14/1/21/DC1.

Clearly, rearrangement of the Va2Var transgene was capable of generating a diversity of sequences. Only 17/283 were out of frame; this low number was most likely attributable to the reduced stability of untranslated mRNA. The remaining 266 sequences represented 212 individual entities with few repeats, most of these being intrapool and thus likely to correspond to duplicate cloning of the same sequence. The in-frame sequences encoded 164 distinct CDR3α motifs, none of which were present in more than 3% of the sequences. The OT-1-derived Jα26 segment was used slightly less frequently than the Jα2 segment. Variable numbers of deletions and N nucleotide insertions (on both the V and J side) were observed (Figure 3B, bottom), with a distribution quite similar to what is seen in naturally rearranged TCRα chain genes (average deletion of two to three bases on either the V or J segment; average addition of three N nucleotides). A few sequences repeated between different mice were present, but these represented a minority.

The Postselection Repertoire

To investigate the impact of thymic selection and peripheral homeostasis, we next determined the sequences of TCR CDR3s expressed by sorted postselection T cells, analyzing large numbers of cells to permit a valid view of the repertoire. This analysis was performed by single-cell sorting and amplification in order to obtain reliable information on repeat frequencies (which is less accurately determined by sorting of and cDNA amplification from pools of cells, where repeats can be artifactual) and to avoid scrambled sequences generated by "chimeric PCR" amplifications from bulk samples. The four populations studied were (1) mature $CD4^-CD8^+$ and (2) $CD4^+CD8^-$ thymocytes and (3) $CD4^+$ and (4) $CD8^+$ lymphocytes from pooled inguinal and axillary lymph nodes. All cells were sorted with strict gates on $V_\beta5^{hi}$ cells in order to eliminate cells that might be expressing an additional V_β chain through escape from allelic exclusion. Limited mice were analyzed at 5–6 weeks of age, a time when the $CD4^+$ and $CD8^+$ populations appear to have stabilized. This early time was also chosen to avoid complications due to the mtv-8/mtv-9 superantigens, which affect $V_\beta5$-expressing cells in older mice (Dunger et al., 1996). Single cells from three mice were sorted into wells of 96-well plates, and Vα2 transcripts were amplified from each cell. PCR products were purified and sequenced directly on high-throughput sequencers. The frequency of positive RT–PCR reactions from single cells was quite high (30%–80% positive wells in various experiments), such that the data sets are not likely to represent preferential amplification of rare sequences and do give an accurate representation of the repertoire in these mice. The complete sequence data from ~200 cells of each population are shown as supplemental data and are presented in compiled form in Figure 4A.

Several points can be made from this set of sequences. First, the repertoire of these selected and fully mature populations is markedly less diverse than that of preselection DPs. Of the 133 sequences seen in the DP sample, only 23 are also detected in postselection cells. Among the 785 sequences obtained in selected cells, there are 216 distinct CDR3α motifs, whose distribution is clearly nonrandom, 33 sequences being found five times or more in the compiled data sets. Some sequences are represented 40 times or more; certain of these were also found in immature DPs but only at a low frequency, indicating that they do not correspond

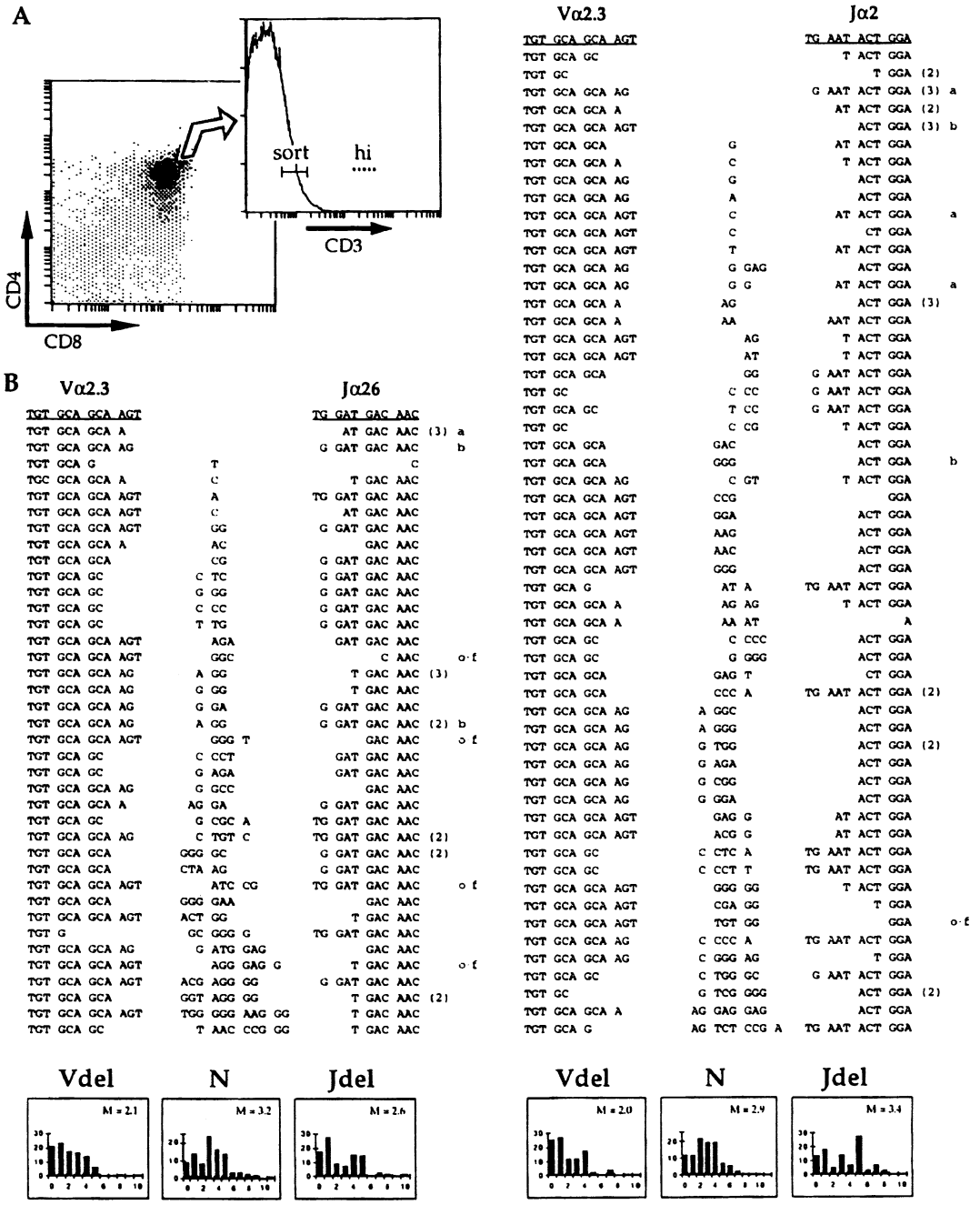
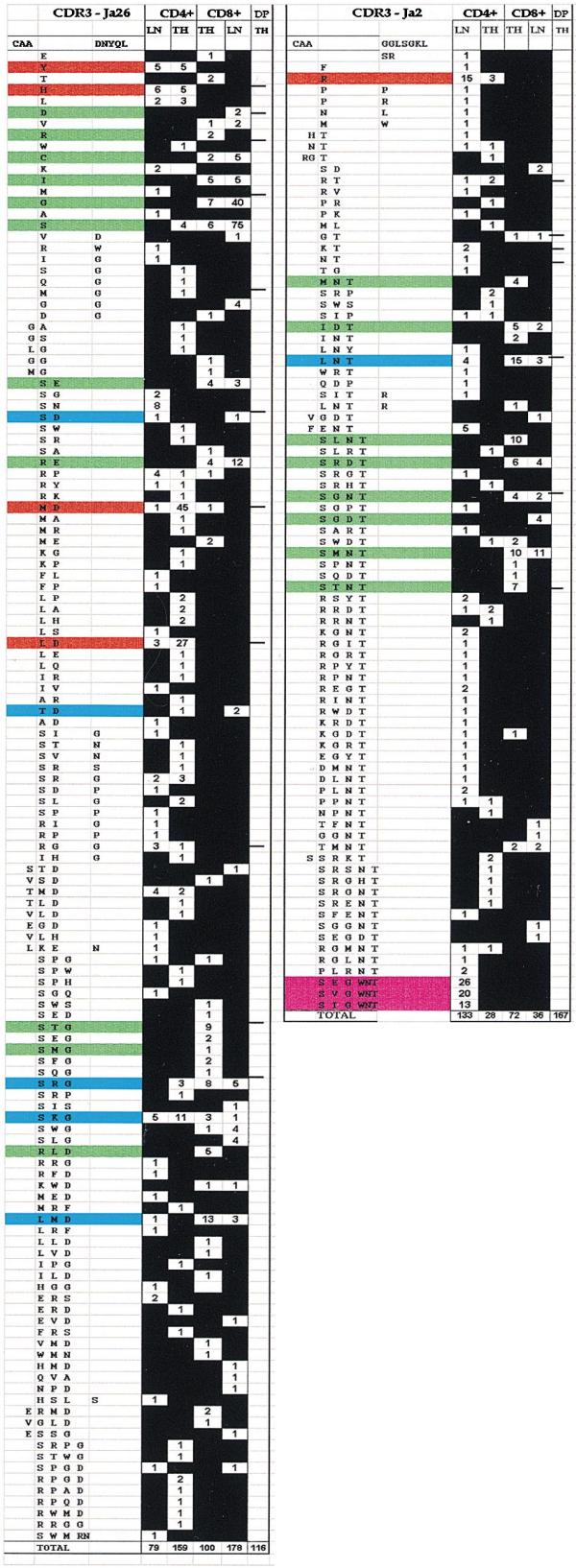


Figure 3. Rearrangement of the TCRα Minilocus Generates CDR3α Diversity in Thymic Precursor Cells

Preselection DP thymocytes from Limited mice were sorted as expressing low levels of CD3 (two independent pools from three different mice). The sequence of the CDR3α joining region was determined after RT–PCR and cloning (283 sequences). All sequences from a single mouse are illustrated here (for complete sequences, see supplemental data), separated into those utilizing Jα26 and Jα2. In parentheses are the number of times individual sequences were observed in the data set, denoted "a" when observed in two independent pools from the same mouse and "b" when present in different animals. o-f, out of frame. Histograms at the bottom indicate the numbers of nucleotide deletions on the V segment, N additions, and deletions on the J segment as a result of the join.

to favored rearrangements (supplemental data; M. C.-N. et al., submitted). These frequent sequences are "public," (i.e., they were found with surprisingly reproducible frequencies in all three mice examined) (Table 1). Second, there are clear distinctions among the repeated sequences. They are not distributed evenly between the CD4$^+$ and CD8$^+$ compartments, most being found in only one of the lineages, consistent with restriction by either MHC class I or II molecules. This point is illustrated in the graph of Figure 4B, which represents the frequency of the CDR3α sequences most commonly observed in the different populations. The sequences found at the top are essentially exclusive to the CD4$^+$ lineage—those on the bottom to the CD8$^+$ lineage. How-

A


CDR3 - Ja26
CD4+
CD8+
DP
LN
TH
TH
LN
TH
CAA
DNYQL
TOTAL
79
159
100
178
116
CDR3 - Ja2
CD4+
CD8+
DP
LN
TH
TH
LN
TH
CAA
GGLSGKL
SEGWNT
SVGWNT
SIGWNT
TOTAL
133
28
72
36
167

B

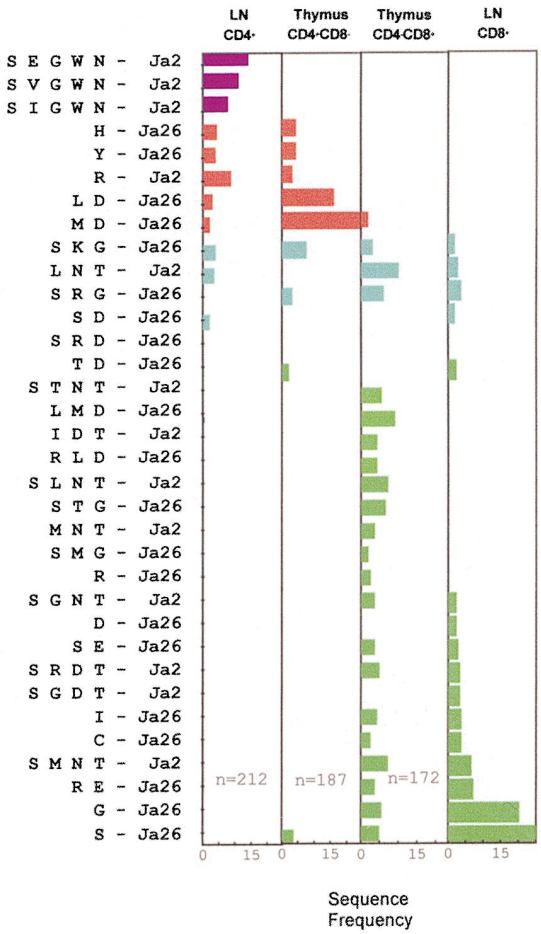

LN CD4-
Thymus CD4-CD8-
Thymus CD4-CD8-
LN CD8-
S E G W N - Ja2
S V G W N - Ja2
S I G W N - Ja2
H - Ja26
Y - Ja26
R - Ja2
L D - Ja26
M D - Ja26
S K G - Ja26
L N T - Ja2
S R G - Ja26
S D - Ja26
S R D - Ja26
T D - Ja26
S T N T - Ja2
L M D - Ja26
I D T - Ja2
R L D - Ja26
S L N T - Ja2
S T G - Ja26
M N T - Ja2
S M G - Ja26
R - Ja26
S G N T - Ja2
D - Ja26
S E - Ja26
S R D T - Ja2
S G D T - Ja2
I - Ja26
C - Ja26
S M N T - Ja2
R E - Ja26
G - Ja26
S - Ja26
n=212
n=187
n=172
0
15
Sequence Frequency

Table 1. Intermouse Reproducibility of Repeat Sequence Frequencies

	$CD4^+$ $CD8^-$			$CD4^-$ $CD8^+$		
Mouse number	57	63	91	57	63	91
CAA**MD**DNYQL	8	9	28	0	1	0
CAA**LD**DNYQL	6	5	16	0	0	0
CAA**SKG**DNYQL	1	0	10	0	0	0
CAA**H**DNYQL	1	1	3	0	0	0
CAA**LMD**DNYQL	0	0	0	1	1	11
CAA**SLNT**GGLSGKL	0	0	1	1	2	7
CAA**LNT**GGLSGKL	0	0	0	6	5	4
CAA**SRG**DNYQL	0	1	2	2	1	5
Total sequenced	41	33	113	31	43	98

ever, a few of the sequences occur in both compartments. The third point that comes out clearly from Figure 4B is that extensive changes in frequencies have occurred between the thymocyte populations and their peripheral counterparts. For example, the G-Jα26 and S-Jα26 sequences make up, together, 54% of the peripheral repertoire of $CD8^+$ cells but are found in only 8% of thymocytes. Conversely, some sequences abundant among mature thymocytes are essentially absent from the peripheral cells (for example, MD-Jα26 or several of the sequences specific to the $CD8^+$ lineage).

These marked divergences between the thymic and peripheral repertoires could be due to expansion of particular clonotypes induced by an immune response or to favored homeostatic expansion and survival as a consequence of interaction with self-MHC:peptide complexes (Goldrath and Bevan, 1999a). If the latter, one might expect to observe the peculiar phenotypic change in late activation markers that has been described for T cell populations homeostatically expanding (Goldrath and Bevan, 1999b). This was indeed the case (Figure 5); $CD8^+$ cells from Limited mice exhibited an upregulation of CD44 without the concomitant downregulation of CD62L normally observed with activated cells. On $CD4^+$ cells, however, both markers were altered, and CD69 was also expressed, suggesting that homeostatic expansion may not have the same phenotypic consequences in $CD4^+$ and $CD8^+$ cells.

As discussed in detail below, the sequences in Figure 4 that reflect recognition of MHC class I versus II molecules are only subtly different—single amino acid replacements being enough to switch the class specificity. We wondered whether these were intrinsic properties dominantly imparted by the CDR3α sequence and what could be the impact of TCRβ chain elements. Therefore, we analyzed CDR3α sequences from mice lacking the OT-1 Vβ5 transgene, in which products of the Va2Var

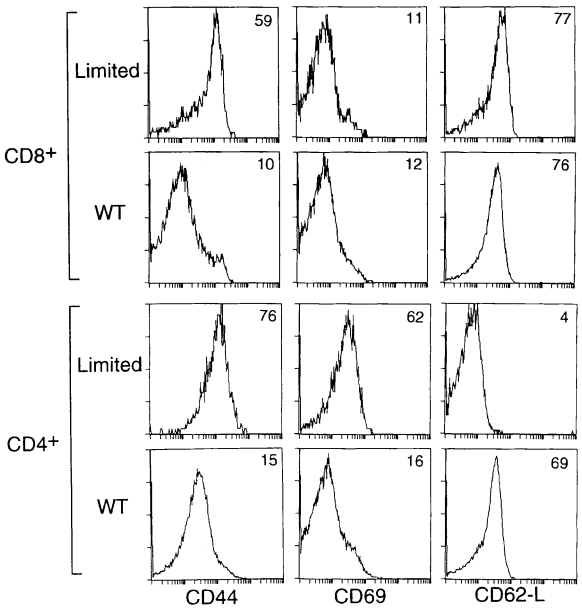


Figure 5. Activation Markers in Lymph Node Cells of Limited Mice

Histograms of gated $CD4^+$ and $CD8^+$ lymph node cells from a Limited or a control mouse, stained for CD44 and CD62L late activation markers and for CD69, an early marker.

transgene were associated with a wide range of Vβ elements. As shown in Figure 6, the correlations described above were largely absent in this case. Interestingly, several of the sequences found prominently in V$\beta5^+$ mice were again observed in this data set but with differing lineage specificity. These results indicate that the MHC class distinctions imparted by CDR3α are conditioned by the Vβ contribution.

Discussion

The data presented above provide a broad view of the generation and evolution of the T lymphocyte repertoire, encompassing T cell selection in the thymus and colonization of the peripheral lymphoid organs. We have taken as an approach a modification of the "single-chain TCR transgenic" strategy that has already proven so powerful (Jorgensen et al., 1992a; Sant'Angelo et al., 1998), further constraining diversity by limiting the variability of the second TCR chain. The combination of transgenes in the Limited mouse performed as had been hoped for, yielding a naive repertoire broad enough for analytical significance but of a manageable size such that impor-

Figure 4. CDR3α Sequences in Positively Selected Cells in Limited Mice

(A) Single cells from thymus and lymph node (TH and LN) of Limited mice were sorted (as $V_\beta5^{hi}$) and their CDR3α sequence determined. The data from 785 cells are grouped according to CDR3α size, with the number of occurrences in each population shown. For clarity, not all amino acids are written out; those shared by most sequences are shown at the top. Sequences that were found repeatedly in different mice and compartments (>12 occurrences in the combined data set) are highlighted. Red, sequences specific to the $CD4^+$ lineage; green, sequences specific to the $CD8^+$ lineage; blue, sequences found for both; and purple, sequences exclusively found in peripheral $CD4^+$ lymphocytes and never in the thymus. The sequences also found (single occurrences) in preselection thymic DPs are shown.

(B) Schematic representation of the frequencies of repeated sequences in the four subpopulations. For clarity, only the distinguishing amino acids are listed (e.g., Y-Jα26 denotes a CAAYDNYQL sequence); refer to Figure 4A or supplemental data for complete sequence. Color coding as in (A).

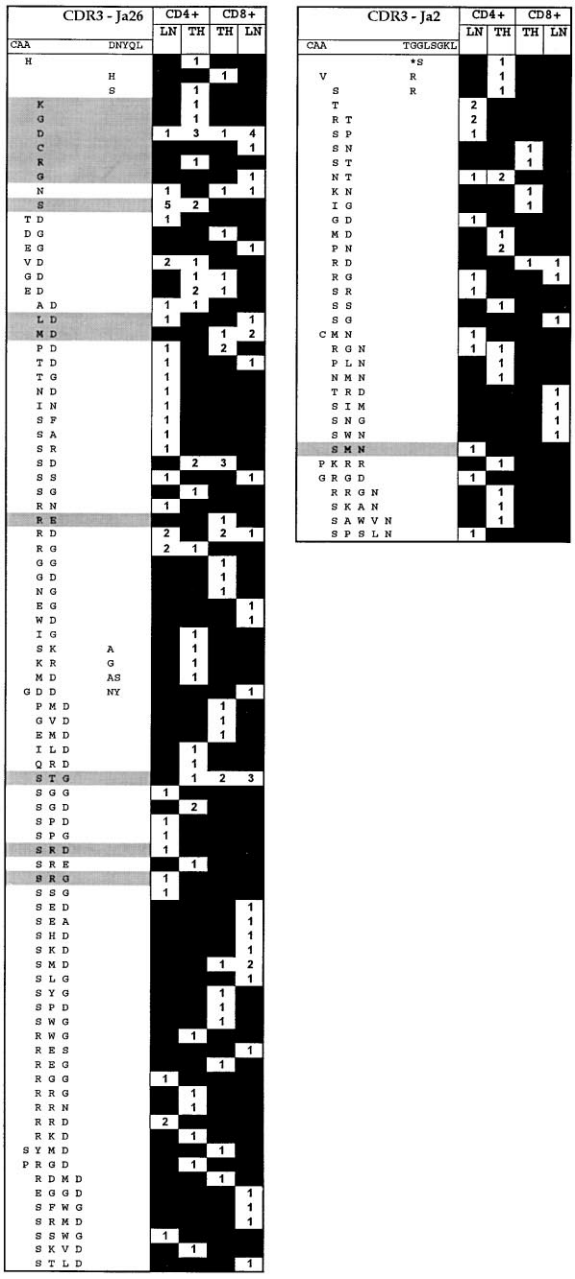


Figure 6. CDRα Sequences in Va2Var Mice Devoid of the TCRβ Transgene

Sequences from CD4$^+$ and CD8$^+$ lymph node cells of three Va2Var/Cα$^{0/0}$ mice, from 175 sorted single cells. Sequences and their relative distribution are represented as in Figure 4A. Those sequences previously identified in the analysis of Limited mice are highlighted.

tant influences could be discerned. A wide array of sequences was generated by the rearrangement process finalized in DP thymocytes, but only a limited subset was maintained in the mature thymocyte pools. The combination of Vα2 and Vβ5 regions encoded by the transgenes allowed sufficient flexibility for recognition of either MHC class I or II molecules. Both Jα segments were utilized, and the conclusions apply to sequences using either of them, ruling out biases due to preferential pairing and reconstitution of the parental OT-1 TCR. The experimental strategy relied on TCRα gene transcript amplification and sequencing from sorted single cells; the frequencies of individual sequences can thus be relied on. The observed frequency distribution, with many repeated sequences in the postselection repertoire, allows one to follow the fate of T cells expressing a number of different sequences in a manner that begins to be significant. This point is bolstered by the fact that these frequent sequences are shared by all three mice analyzed (Table 1) and are thus public specificities. This broad view of T cell repertoire formation and flux leads to novel conclusions along three different lines.

Repertoire Selection in the Thymus

Combined positive and negative selection in the thymus results in a T cell repertoire that is very different from the "raw" repertoire generated by random TCR gene rearrangements. We have demonstrated that, starting from a rather "flat" collection in which very few sequences are repeated, these selection processes result in a repertoire of mature thymocytes expressing a number of unique sequences but dominated by many instances of multiply reiterated sequences. These reiterations make up an impressive proportion of the mature thymocyte repertoire in the Limited mouse; 63% of CD4$^+$CD8$^-$ and 69% of CD4$^-$CD8$^+$ cells express one of these sequences, defined as occurring 12 times or more in the data set (supplemental data); the single most abundant accounts for 25% of CD4$^+$CD8$^-$ thymocytes. This dominance is even more striking when one realizes that many of these repeats are very closely related in sequence (Figure 4B) (for example, SLNT-Jα2, STNT-Jα2, SGNT-Jα2, and SMNT-Jα2). These repeats are not due to favored molecular rearrangements, as they are not particularly frequent in the preselection repertoire of DP cells and because they are encoded by different degenerate nucleic acid sequences (supplemental data). They are also very unlikely to reflect trivial PCR contamination artifacts, as drastic measures were taken during single-cell sorting and PCR to prevent and diagnose contamination problems.

Rough estimates suggest that these repeated sequences are present at frequencies in the 1/10^6 range—a marked departure from the estimated frequency of 1/10^{14} for any randomly generated TCR sequence combination. This highly nonuniform repertoire is consistent with and enhances previous observations (Sant'Angelo et al., 1998). Employing a single-chain TCRβ tg mouse line, it was found that the "parental" CDR3α sequence occurred at high frequency in a sequence set derived by amplifying cDNA from sorted CD4$^+$ cells for the Vα2Jα11 segments of the original T cell clone. Our findings are different in that most of the repeated sequences we found diverge from the TCR of origin; they use two different J regions, and the TCRs they encode engage different MHC molecules. Thus, the caveat of a specially favored selection of the parental sequences does not apply.

We cannot formally rule out that these repeats reflect preferential proliferative amplification of a subset of mature thymocytes in the Limited tg mice. However, BrdU incorporation experiments have shown that the prolifer-

ation of $CD4^+CD8^-$ and $CD4^-CD8^+$ cells from Limited and normal mice is about the same (data not shown). Since cell division in normal mice seems to correspond to short rounds of cycling immediately prior to export (Ernst et al., 1995; Penit and Vasseur, 1997), it would be unlikely to account for the numbers of repeated sequences we observe here. Furthermore, elevated numbers of these repeated sequences are found in the earliest transitional populations in the thymus ($CD3^{hi}$ DPs and $CD4^+CD8^{int}$) (M. C.-N. et al., submitted). Thus, the data suggest that it is the positive selection process itself that generates this "bumpy" repertoire. DP cells expressing particular CDR3α sequences might have quite different probabilities of being selected rather than having a simple yes/no choice, and these probabilities may result in a spectrum of frequencies in the repertoire.

Extrapolating from the particular context of the Limited mouse, the implication is that a repertoire selected in the thymus contains, for any VαJαVβJβ combination, two categories of sequences: some present at low frequencies and others present in many cells. We propose that this stems from a probabilistic aspect of positive selection. Depending on the affinity/avidity for self-MHC, some TCRs lead to highly probable selection, while others confer only a limited probability. These overselected TCRs may reflect positive selection on abundant MHC:peptide complexes—a prediction consistent with the loss of "canonical" responses when moth cytochrome c–specific T cells were selected on experimentally introduced neopeptides displayed on thymic epithelial cells (Nakano et al., 1997). It would be interesting to test this notion using the various "single-peptide" mice that have been derived over the last few years.

The dominant selection of a relatively limited number of TCRs cannot be without consequences on immune function. It is tempting to speculate that preferentially selected repeats underlie the canonical motifs that often characterize the responses to particular peptides, such as the moth cytochrome c 88–103 and myelin basic protein 1–11 peptides (Acha-Orbea et al., 1988; Jorgensen et al., 1992b). The presence of grossly overrepresented TCRs in the repertoire might also explain the immunodominance of particular peptides in complex proteins (Bousso et al., 1998; Wallace et al., 2000)—immunodominance that cannot always be explained by processing considerations or by enhanced binding to MHC molecules (e.g., Vitiello et al., 1996).

Repertoire Molding in the Periphery

Once shaped by selection in the thymus, the T cell repertoire is subject to further molding in the peripheral lymphoid organs, graphically illustrated in Figure 4B. The dominant sequences still occur, but their relative prevalence changes considerably—some are virtually absent from the peripheral pool; others become markedly more frequent. These general patterns were observed reproducibly in three individual mice (Table 1; supplemental data). It is possible that some of the thymus/periphery differences are due to preferential export from the thymus, but it seems more likely that they represent differential homeostatic expansion and maintenance in the periphery. It has become clear in the past few years that T lymphocytes require continued engagement by MHC molecules to survive in the periphery as well as for homeostatic expansion (Goldrath and Bevan, 1999a). Since selection in the thymus is strongly influenced by abundant self-peptides and since the relative contribution of these peptides is likely to be somewhat different in cells of the peripheral organs, one would expect the peripheral peptidic self to exert new pressures on the repertoire. Thus, the profound changes we observed between the mature thymic repertoire and its peripheral descendant probably reflect secondary molding by an altered complement of self-peptides (Ernst et al., 1999), differentially modulating the expansion and survival of individual T cells.

It is unlikely that these changes are due to amplification during background-level immune responses (1) because they were detected in three different mice sampled at spaced times, (2) because some correspond to sequences already observed in the context of an OVA-specific immune response in a mouse colony located in another hemisphere (Kelly et al., 1993), and (3) because most of those repeats that were expanded in the periphery correspond to sequences already overselected in the thymus. In other words, the peripheral interactions continue but remold selective processes begun in the thymus. It is also unlikely that the results are influenced by superantigens, as the reported effects of mtv-8 and mtv-9 on $V_\beta5^+$ cells are only apparent in mice between 15 and 30 weeks of age (Dunger et al., 1996).

Limited mice have a low thymocyte output because T cells in these animals rely on the rearrangement of a single Vα substrate. Hence, there is more homeostatic expansion than would normally be observed, and a good proportion of the $CD8^+$ cells exhibit the particular "homeostatic expansion" phenotype described by Goldrath and Bevan (1999b), with disjuncted expression of the CD44 and CD62L late activation markers. The implication from our data is that the degree of homeostatic expansion can vary widely for cells with different TCRs. These data extend recent reports (Ernst et al., 1999; Ferreira et al., 2000) showing that cells expressing different tg TCRs were variably responsive to homeostatic cues. It is not known at present how much of a role homeostatic expansion or differential survival plays in a normal immune system, although it is suggestive that one of the amplified sequences is precisely that of OT-1 itself (S-Jα26), already found repeatedly in normal mice. It will be important to determine how much of the repertoire remolding we observed also occurs in the context of a "full" immune system.

A special mention must be made of three peculiar sequences found repetitively in the peripheral $CD4^+$ compartment of all mice (Figure 4). They are very similar to each other and have uncharacteristically long CDR3s with very conserved nucleotide sequences (nine N additions identical or highly similar in all cases and in the different mice; see supplemental data). They are conspicuously absent from the over 1000 sequences derived from various thymic populations analyzed in the course of this project (this paper and M. C.-N. et al., submitted). It is tempting to speculate that these sequences reflect extrathymic differentiation, perhaps relying on an unusual form of rearrangement related to the DIR motifs observed in immunoglobulin rearrangements (Tuaillon and Capra, 1998).

In conclusion, the analyses performed on the Limited mouse line clearly demonstrate that the repertoire of naive T cells available to participate in an immune response is far less diverse and flat in frequency than previously thought and that it is molded in a similarly important manner by thymic selection and peripheral expansion/maintenance forces. In the future, these mice should prove valuable for investigating the fine specificity of TCR/MHC:peptide interactions.

MHC Class Specificity of the TCR

The cardinal observation is that when diversity is limited to such an extent there are very clear differences between sequences selected into the $CD4^+$ versus $CD8^+$ lineages—by inference, primarily restricted by MHC class II versus class I molecules (these assignments are confirmed in M. C.-N. et al., submitted). Consistent patterns emerge from the data compiled in Figure 4A. For example, all CDR3s of LX-Jα26 type are class II restricted, irrespective of the nature of "x". There are also CDR3 length biases: eight-residue CDR3s are more frequent in $CD4^+$ cells, and ten-residue CDR3s are more frequent in $CD8^+$ cells. However, no overriding rule emerges, and minute sequence differences can switch MHC class preference. The clearest example can be found in short CDR3s of the form X-Jα26. Bulkier Y, H, L, and K residues at the X position are only found in $CD4^+$ cells, while the smaller S, G, I, C, D, or V residues occur in $CD8^+$ cells. There are many other examples in the data of Figure 4 of single amino acid differences distinguishing MHC class I– and class II–restricted TCRs. Thus, much as TCR/MHC:peptide interactions are exquisitely sensitive to variations in peptide sequence, similar discrimination applies to TCR recognition of MHC class. Generalizing from their crystal structure of the A^k-restricted D10 TCR, Reinherz and coworkers have proposed that the TCR binds MHC class II:peptide complexes in a spatially distinct more orthogonal orientation than they bind MHC class I:peptide complexes (Reinherz et al., 1999). Such a dichotomy is difficult to reconcile with the overlapping patterns and the effects of minor differences observed here. Our results are more in line with Wilson's proposal that the D10 structure represents one extreme in a spectrum of subtly different conformations (Wilson, 1999). This view is also consistent with the great diversity of TCR/MHC contacts that has been observed with the various structures (Garboczi et al., 1996; Garcia et al., 1996; Ding et al., 1998; Teng et al., 1998; Reinherz et al., 1999). It seems then that MHC class specificity is imparted by a variable combination of contacts leading to congruence, rather than by differential docking and conserved TCR/MHC contacts.

Interestingly, the amino acid positions that condition MHC class specificity in our system are those that interact far more with peptide than with the MHC molecule α helices, at least in the crystal structures that have been solved so far (Garboczi et al., 1996; Garcia et al., 1996; Ding et al., 1998; Teng et al., 1998; Reinherz et al., 1999). For example, in the corresponding positions of the crystal structure of Reinherz et al. (1999), amino acids 99–101 engage in 16 contacts with peptide but only three with MHC molecule atoms. The implication is that MHC class specificity can be determined by the peptides presented by MHC molecules during positive selection in the thymus, rather than by contacts with the MHC molecules themselves. A determining influence for peptide interactions is also consistent with results from studies on Vα families used preferentially by $CD4^+$ or $CD8^+$ cells; the positions found to be crucial for these biases, 51 for Vα3 and 30 for Vα2 (Sim et al., 1996; Correia-Neves et al., 1999), participate in CDR1 and CDR2 but are centrally located in the crystal structures and are likely to engage primarily in peptide contacts.

These data also support the notion that MHC class specificity is not solely imparted by the CDR1 and CDR2 regions of the TCR, as results from some past studies (Sim et al., 1996, 1998) are often taken to indicate, but that CDR3 also plays a major role (Correia-Neves et al., 1999). In addition, the very fact that the Vα2Vβ5 TCR of the Limited mice allows balanced selection into both the $CD4^+$ and $CD8^+$ compartments indicates that with a fixed CDR1 and 2 there is still great leeway for CDR3 to fix MHC class preference. It should be kept in mind that the experiments of Sim et al. (1996, 1998) focused on CDR1 and 2 differences and thus could only identify influences within these regions; in the present experiments, we have concentrated on CDR3, as CDR1 and 2 are fixed; all three CDRs must have a role to play.

A few sequences were found in both $CD4^+$ and $CD8^+$ cells from Limited mice in the thymus or periphery (eg SKG-Jα26 or LNT-Jα2). In hindsight, it may not be too surprising that such sequences exist. There must be a degree of evolutionary selection for general congruence between the TCR and MHC molecules, and several authors (Blackman et al., 1986; Gilfillan et al., 1994; Gavin and Bevan, 1995) have demonstrated that the unselected repertoire has a high level of overall compatibility with any given MHC molecule; in some estimates, up to 30% of unselected TCR sequences can interact with an MHC molecule encountered in the thymus. Thus, one might expect in a Limited mouse expressing A^b and K^b/D^b that up to 10% of sequences could engage both MHC classes, at least before negative selection takes its toll. In fact, the proportion of such cells we observed was rather lower, and most of the shared sequences showed distinct class bias. However, this number may be an underestimate, since we have found that, when the Limited mouse is rendered MHC class II deficient, some of the sequences normally picked up only in $CD4^+$ cells are also detected in $CD8^+$ cells, e.g., the H-Jα26 sequence (M. C.-N. et al., submitted).

In the several crystal structures of TCR/MHC:peptide complexes that have been published to date, the TCR Vα region has the most reproducible configuration and footprint, while there is far more divergence in the overall configuration of Vβ, which in some structures appeared to angle away from the MHC:peptide surface or to make very few contacts (reviewed in Wilson, 1999). One might have predicted then that the CDR3α sequence/MHC class preference correlations we observed in the T cell repertoire of the Limited mouse, with its fixed Vβ5 chain (Figure 4), might also have been seen, at least partially, in the repertoire of mice expressing the panoply of Vβ regions produced by rearrangement of endogenous TCRβ genes. But, this turned out not to be the case. Some of the more frequently repeated sequences were found again, but this time in the other mature T cell

compartment—for example, S-Jα26 was present only in CD8$^+$ cells displaying the OT-I Vβ5 region but only in CD4$^+$ cells when found in combination with diverse endogenously encoded Vβs.

Experimental Procedures

DNA Construct and Transgenesis

The construct for generating TCRα chains with CDR3 variability allowed recombination between Vα2.3 and Jα26 as well as Vα2.3 and Jα2 gene segments. The template containing Vα2.3 and Jα26 was made by PCR and introduced into the TCRα cassette vector (Kouskoff et al., 1993). This cassette contains the regulatory elements for expression of the α chain, the constant region of the TCRα chain as well as the Jα2 gene segment, downstream of the cloning sites for the variable TCR genes. The Vα2.3 and Jα26 gene segments were amplified from a plasmid containing the original OT-1 TCRα chain (clone 149.42 in Kelly et al., 1993). The Vα2.3-Jα26 junctional region from the original OT-1 TCRα chain was replaced by a rearrangement substrate containing an RSS downstream of Vα2.3 (CACAGTG.CTCTCCAGGCACCTGTAGCCCAT.ACCCAAACC) (Okazaki and Sakano, 1988) and the natural RSS from the Jα26 segment, separated by the 547 bp sequence naturally found upstream of Jα26 (Koop et al., 1992). A silent mutation was engineered in the Jα26 segment to create a novel PvuII site (TATCAGCTGATCTGG). The rearrangement substrate was assembled by PCR ligation with overlapping primers (Vα2.3-VαRSS-intersignal DNA-JαRSS-Jα26) and cloned into the TCRα chain expression cassette (Kouskoff et al., 1993) using the unique XmaI and SacII sites.

The construct was excised from plasmid DNA and injected into fertilized B6×SJL F2 eggs. Screening of tg founders was performed on Southern blots with a 2.8 kb probe originating from the TCRα cassette. Limited mice were generated by breeding mice carrying the transgene-encoded variable α chain with mice tg for the OT-1 TCRβ chain (Correia-Neves et al., 1999) and again with knockout mice deficient for the expression of the normal TCRα chain (Philpott et al., 1992), with subsequent intercrossing of progeny. All mice used in this work were homozygous for the H-2^b haplotype.

All animals were maintained in the conventional facility of the Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) under Ministère de l'Agriculture (Agrément 67227) and European Economic Community guidelines.

PCR Amplification to Test for the Rearrangement of the Variable α Chain

RT–PCR from lymph node RNA was performed with primers that specifically amplified cDNA containing a rearranged Vα2 gene—sense primer (Vα2), GTTCCCGGGATCGAGGAATGGACAAGATTC TG; reverse primer (TCRα constant), GGCCCCATTGCTCTTGGA ATC. A second PCR round was performed using nested primers that amplify Vα2 gene segments (sense primer, TCCATACGTTCAG TGTCCGATAAA) recombined with Jα26 (antisense primer, GGC TTTATAATTAGCTTGGTCCCAGAG). The resulting PCR products (200 bp) were digested with PvuII.

PCR Amplification and Cloning of Vα2$^+$ TCRα Chains from DP Thymocytes

Thymocytes were stained with anti-CD8α, anti-CD4 (Caltag Laboratories, Inc., Burlingame, CA), and KT3, specific for CD3 (Tomonari, 1988). Two independent pools containing 2×10^5 CD4$^+$CD8$^+$CD3^{low} thymocytes were purified from each of three individual mice. HeLa cells (10^6) were added to each sample as carrier, and RNA was prepared by NP-40 lysis. cDNA was synthesized using half of the RNA from each sample in 40 μl of reverse transcription buffer (50 mM Tris-HCl, 75 mM KCl, and 3 mM MgCl2). 10 μl of this reaction was used for the first round of PCR amplification. The PCR conditions used were the same as described previously (Correia-Neves et al., 1999), except that only 20 cycles were performed for each round, and the antisense primer for the second PCR round was distinct (TCTCGAATTCAGGCAGAGGGTGCTGTCC). The PCR products were purified by phenol/chloroform extraction, digested with XmaI and EcoRI restriction enzymes, purified on agarose gel, and cloned into pBS. The ligation was transfected into XL1-blue bacteria, and the resulting colonies were screened by hybridization with an oligo specific for the Vα2 segments (GGAAGATGGACGATT).

Single-Cell Sorting and RT-PCR

Thymocytes and lymphocytes were stained with anti-CD8α, anti-CD4 (Caltag Laboratories, Inc.), and MR9-4, specific for the Vβ5 family (Bill et al., 1990). Individual cells were deposited into wells of 96-well PCR plates containing 10 μl of RT buffer (50 mM Tris-HCl, 75 mM KCl, and 3 mM $MgCl_2$) with 2% Triton X-100, 1 μg BSA, 500 μM dNTP, 50 ng of oligo(dT)12-18 (Pharmacia), 8 U RNasin, and 30 U of Moloney murine leukemia virus reverse transcriptase (Life Technologies, Grand Island, NY). The plates were incubated for 90 min at 37°C. For the first PCR round, 10 μl of the resulting cDNA were amplified by adding 40 μl of Taq buffer (50 mM KCl, 10 mM Tris-HCl [pH 9.3], and 2.5 mM $MgCl_2$) containing 2.5 U Taq polymerase, 500 μM dNTP, 400 ng sense primer specific for the Vα2 family (CAGCAGCAGGTGAGACAAAGT), and antisense primer specific for the TCRα constant region (GTTTTGTCAGTGATGA ACGT) (3 min at 93°C; 35 cycles with 50°C annealing temperature). For the second PCR round, 2 μl of the first amplification product were amplified under the same conditions in 50 μl of Taq buffer containing 1U of Taq polymerase, 500 μM dNTP, 200 ng of sense AAGGCCCGGGTCTCTGACAGTCTGGGAAGGA, and AATCTGCAG CGGCACATTGATTTGGGA nested primers (22 cycles). The PCR products were purified by PEG (PEG 6000 20% and NaCl 2.5 M) precipitation followed by two washes in 75% ethanol. The PCR products were sequenced directly using one of the primers for the second round of PCR (AATCTGCAGCGGCACATTGATTTGGGA).

Painstaking precautions were taken against PCR contamination, with 16 control wells per 96-well PCR plate in which a mock droplet of PBS was sorted. Plates showing any trace of positive amplification in these control wells were discarded.

Acknowledgments

We thank Drs. F. Carbone, S. Chan, and D. Wiley for discussions and reagents, S. Vicaire, P. Brottier, and J. Weissenbach (Génoscope; Centre National de Séquencage) for sequencing, F. Jeanmougin for help with data analysis, and P. Gerber, P. Marchal, J. Hergueux, and C. Ebel for technical assistance. This work was supported by institute funds from the INSERM, the CNRS, the Hopital Universitaire de Strasbourg, and by a grant from the Association pour la Recherche contre le Cancer. M. C.-N. received fellowships from the Programa Gulbenkian de Doutoramento em Biologia e Medicina and the Fundacao Para a Ciencia e a Tecnologia.

Received August 31, 2000; revised December 11, 2000.

References

Acha-Orbea, H., Mitchell, D.J., Timmermann, L., Wraith, D.C., Tausch, G.S., Waldor, M.K., Zamvil, S.S., McDevitt, H.O., and Steinman, L. (1988). Limited heterogeneity of T cell receptors from lymphocytes mediating autoimmune encephalomyelitis allows specific immune intervention. Cell *54*, 263–273.

Arstila, T.P., Casrouge, A., Baron, V., Even, J., Kanellopoulos, J., and Kourilsky, P. (1999). A direct estimate of the human αβ T cell receptor diversity. Science *286*, 958–961.

Bill, J., Kanagawa, O., Linten, J., Utsunomiya, Y., and Palmer, E. (1990). Class I and class II MHC gene products differentially affect the fate of Vβ5 bearing thymocytes. J. Mol. Cell. Immunol. *4*, 269–280.

Blackman, M., Yague, J., Kubo, R., Gay, D., Coleclough, C., Palmer, E., Kappler, J., and Marrack, P. (1986). The T cell repertoire may be biased in favor of MHC recognition. Cell *47*, 349–357.

Bousso, P., Casrouge, A., Altman, J.D., Haury, M., Kanellopoulos, J., Abastado, J.P., and Kourilsky, P. (1998). Individual variations in the murine T cell response to a specific peptide reflect variability in naive repertoires. Immunity *9*, 169–178.

Candeias, S., Waltzinger, C., Benoist, C., and Mathis, D. (1991). The V beta 17+ T cell repertoire: skewed J beta usage after thymic

selection; dissimilar CDR3s in CD4+ versus CD8+ cells. J. Exp. Med. *174*, 989–1000.

Correia-Neves, M., Waltzinger, C., Wurtz, J.M., Benoist, C., and Mathis, D. (1999). Amino acids specifying MHC class preference in TCR V alpha 2 regions. J. Immunol. *163*, 5471–5477.

Davis, M.M., and Bjorkman, P.J. (1988). T-cell antigen receptor genes and T-cell recognition. Nature *334*, 395–402.

Ding, Y.H., Smith, K.J., Garboczi, D.N., Utz, U., Biddison, W.E., and Wiley, D.C. (1998). Two human T cell receptors bind in a similar diagonal mode to the HLA-A2/Tax peptide complex using different TCR amino acids. Immunity. *8*, 403–411.

Dunger, A., Cunningham, J.M., Delaney, C.A., Lowe, J.E., Green, M.H.L., Bone, A.J., and Green, I.C. (1996). Tumor necrosis factor-α and interferon-γ inhibit insulin secretion and cause DNA damage in unweaned-rat islets. Diabetes *45*, 183–189.

Ernst, B., Surh, C.D., and Sprent, J. (1995). Thymic selection and cell division. J. Exp. Med. *182*, 961–971.

Ernst, B., Lee, D.S., Chang, J.M., Sprent, J., and Surh, C.D. (1999). The peptide ligands mediating positive selection in the thymus control T cell survival and homeostatic proliferation in the periphery. Immunity *11*, 173–181.

Ferreira, C., Barthlott, T., Garcia, S., Zamoyska, R., and Stockinger, B. (2000). Differential survival of naive CD4 and CD8 T cells. J. Immunol. *165*, 3689–3694.

Fink, P.J., and Bevan, M.J. (1978). H-2 antigens of the thymus determine lymphocyte specificity. J. Exp. Med. *148*, 766–774.

Garboczi, D.N., Ghosh, P., Utz, U., Fan, Q.R., Biddison, W.E., and Wiley, D.C. (1996). Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. Nature *384*, 134–141.

Garcia, K.C., Degano, M., Stanfiled, R.L., Brunmark, A., Jackson, M.R., Peterson, P.A., Teyton, T., and Wilson, I.A. (1996). An αβ T cell receptor structure at 2.5 A and its orientation in the TCR-MHC complex. Science *274*, 209–219.

Gavin, M.A., and Bevan, M.J. (1995). Increased peptide promiscuity provides a rationale for the lack of N regions in the neonatal T cell repertoire. Immunity *3*, 793–800.

Gilfillan, S., Waltzinger, C., Benoist, C., and Mathis, D. (1994). More efficient positive selection of thymocytes in mice lacking terminal deoxynucleotidyl transferase. Int. Immunol. *6*, 1681–1686.

Goldrath, A.W., and Bevan, M.J. (1999a). Selecting and maintaining a diverse T-cell repertoire. Nature *402*, 255–262.

Goldrath, A.W., and Bevan, M.J. (1999b). Low-affinity ligands for the TCR drive proliferation of mature CD8+ T cells in lymphopenic hosts. Immunity *11*, 183–190.

Hogquist, K.A., Jameson, S.C., Heath, W.R., Howard, J.L., Bevan, M.J., and Carbone, F.R. (1994). T cell receptor antagonist peptides induce positive selection. Cell *76*, 17–27.

Jameson, S.C., Kaye, J., and Gascoigne, N.R. (1990). A T cell receptor V alpha region selectively expressed in CD4+ cells. J. Immunol. *145*, 1324–1331.

Jorgensen, J.L., Esser, U., Fazekas de St. Groth, B., Reay, P.A., and Davis, M.M. (1992a). Mapping T-cell receptor-peptide contacts by variant peptide immunization of single-chain transgenics. Nature *355*, 224–230.

Jorgensen, J.L., Reay, P.A., Ehrich, E.W., and Davis, M.M. (1992b). Molecular components of T-cell recognition. Annu. Rev. Immunol. *10*, 835–873.

Kappler, J.W., Roehm, N., and Marrack, P. (1987). T cell tolerance by clonal elimination in the thymus. Cell *49*, 273–280.

Kelly, J.M., Sterry, S.J., Cose, S., Turner, S.J., Fecondo, J., Rodda, S., Fink, P.J., and Carbone, F.R. (1993). Identification of conserved T cell receptor CDR3 residues contacting known exposed peptide side chains from a major histocompatibility complex class I-bound determinant. Eur. J. Immunol. *23*, 3318–3326.

Koop, B.F., Wilson, R.K., Wang, K., Vernooij, B., Zallwer, D., Kuo, C.L., Seto, D., Toda, M., and Hood, L. (1992). Organization, structure, and function of 95kb of DNA spanning the murine T-cell receptor C alpha/C delta region. Genomics *13*, 1209–1230.

Kouskoff, V., Fehling, H.J., Lemeur, M., Benoist, C., and Mathis, D. (1993). A vector driving the expression of foreign cDNAs in the MHC class II-positive cells of transgenic mice. J. Immunol. Methods *166*, 287–291.

Nakano, N., Rooke, R., Benoist, C., and Mathis, D. (1997). Positive selection of T cells induced by viral delivery of neopeptides to the thymus. Science *275*, 678–683.

Okazaki, K., and Sakano, H. (1988). Thymocyte circular DNA excised from T cell receptor alpha-delta gene complex. EMBO J. *7*, 1669–1674.

Pannetier, C., Even, J., and Kourilsky, P. (1995). T-cell repertoire diversity and clonal expansions in normal and clinical samples. Immunol. Today *16*, 176–181.

Penit, C., and Vasseur, F. (1997). Expansion of mature thymocyte subsets before emigration to the periphery. J. Immunol. *159*, 4848–4856.

Philpott, K.L., Viney, J.L., Kay, G., Rastan, S., Gardiner, E.M., Chae, S., Hayday, A.C., and Owen, M.J. (1992). Lymphoid development in mice congenitally lacking T cell receptor αβ-expressing cells. Science *256*, 1448–1452.

Reinherz, E.L., Tan, K., Tang, L., Kern, P., Liu, J., Xiong, J., Hussey, R.E., Smolyar, A., Hare, B., Zhang, R., et al. (1999). The crystal structure of a T cell receptor in complex with peptide and MHC class II. Science *286*, 1913–1921.

Sant'Angelo, D.B., Waterbury, P.G., Cohen, B.E., Martin, W.D., van Kaer, L., Hayday, A.C., and Janeway, C.A., Jr. (1997). The imprint of intrathymic self-peptides on the mature T cell receptor repertoire. Immunity *7*, 517–524.

Sant'Angelo, D.B., Lucas, B., Waterbury, P.G., Cohen, B., Brabb, T., Goverman, J., Germain, R.N., and Janeway, C.A., Jr. (1998). A molecular map of T cell development. Immunity *9*, 179–186.

Sim, B.C., Zerva, L., Greene, M.I., and Gascoigne, N.R.J. (1996). Control of MHC restriction by TCR Valpha CDR1 and CDR2. Science *273*, 963–966.

Sim, B.C., Lo, D., and Gascoigne, N.R. (1998). Preferential expression of TCR V alpha regions in CD4/CD8 subsets: class discrimination or co-receptor recognition? Immunol. Today *19*, 276–282.

Teng, M.K., Smolyar, A., Tse, A.G., Liu, J.H., Liu, J., Hussey, R.E., Nathenson, S.G., Chang, H.C., Reinherz, E.L., and Wang, J.H. (1998). Identification of a common docking topology with substantial variation among different TCR-peptide-MHC complexes. Curr. Biol. *8*, 409–412.

Tomonari, K. (1988). A rat antibody against a structure functionally related to the mouse T-cell receptor/T3 complex. Immunogen. *28*, 455–458.

Tuaillon, N., and Capra, J.D. (1998). Use of D gene segments with irregular spacers in terminal deoxynucleotideyltransferase (TdT)+/+ and TdT-/- mice carrying a human Ig heavy chain transgenic minilocus. Proc. Natl. Acad. Sci. USA *95*, 1703–1708.

Vitiello, A., Yuan, L., Chesnut, R.W., Sidney, J., Southwood, S., Farness, P., Jackson, M.R., Peterson, P.A., and Sette, A. (1996). Immunodominance analysis of CTL responses to influenza PR8 virus reveals two new dominant and subdominant Kb-restricted epitopes. J. Immunol. *157*, 5555–5562.

von Boehmer, H. (1990). Developmental biology of T cells in T-cell receptor transgenic mice. Annu. Rev. Immunol. *8*, 531–556.

Wallace, M.E., Bryden, M., Cose, S.C., Coles, R.M., Schumacher, T.N., Brooks, A., and Carbone, F.R. (2000). Junctional biases in the naive TCR repertoire control the CTL response to an immunodominant determinant of HSV-1. Immunity *12*, 547–556.

Wilson, I.A. (1999). Perspectives: protein structure. Class-conscious TCR? Science *286*, 1867–1868.

Zinkernagel, R.M., and Doherty, P.C. (1997). The discovery of MHC restriction. Immunol. Today *18*, 14–17.