The EMBO Journal vol.10 no.12 pp.3647–3654, 1991

A special repertoire of α:β T cells in neonatal mice

Molly Bogue[1], Serge Candéias[1,2], Christophe Benoist and Diane Mathis

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de Médecine, 67085 Strasbourg Cedex, France

[1]Authors contributed equally to this study

[2]Present address: National Jewish Center for Immunology and Respiratory Medicine, 1400 Jackson Street, Denver, CO 80206, USA

Communicated by D.Mathis

According to several functional criteria, the mature thymocytes of neonatal and adult mice are distinctly different. We wondered whether these differences in function might have a structural correlate: do neonates have a distinct repertoire of α:β T cells? In this study, we have exploited the power of polymerase chain reaction technology to generate large numbers of T cell receptor sequences from sorted thymocyte populations from newborn and adult mice. The newborn-derived sequences show very few N nucleotide additions, usually the major source of diversity in T cell receptors. Most interestingly, the paucity of N insertions appears to be exaggerated by selection events that operate during T cell differentiation in the thymus. The significance of these results is largely: (i) that they parallel recent findings on the B cell repertoire in neonates, raising questions about the reactivities specified by such a special repertoire; and (ii) that they suggest a means to 'tag' T cells exported perinatally, allowing one to test the premise that autoreactive T cells derive preferentially from the newborn repertoire.

Key words: repertoire selection/sequence diversity/T cell differentiation/T cell receptor/thymus

Introduction

Thymocytes undergo a programmed series of differentiation events before emerging into the periphery as competent effector T cells (for reviews, see Boyd and Hugo, 1991; Nikolic-Zugic, 1991). These events can be monitored by assaying expression of the T cell antigen receptor (TCR) along with the CD4 and CD8 co-receptor molecules. The most immature thymocytes do not display surface TCRs and are found within the $CD4^-8^-$ (DN) population. They progress through a $CD4^+8^+$ (DP) stage, during which they express their receptors, initially at rather low levels. Eventually, they differentiate into $CD4^+8^-$ or $CD4^-8^+$ (SP) cells that express high levels of surface TCR and are the direct precursors of the MHC class II- or class I-restricted effector cells populating the periphery.

Several experiments, some performed over twenty years ago, have suggested that the mature thymocytes of newborn and adult mice are somehow different. First, it was discovered that neonatal, but not adult, thymocytes proliferate vigorously and manufacture Il-2 when challenged *in vitro* with syngeneic spleen cells (Howe *et al.*, 1970; von Boehmer and Byrd, 1972; Lattime and Stutman, 1986). Second, it was found that adult, but not neonatal, SP thymocytes have been 'purged' of certain self-reactive specificities. In adult mice expressing the class II E complex and one of various retroviral-encoded 'superantigens', almost all cells displaying particular TCR β-chain variable regions (Vβs) are absent from the peripheral repertoire, having been expunged in the thymus at or just before the SP stage (e.g. Kappler *et al.*, 1987a, 1988; Bill *et al.*, 1988; MacDonald *et al.*, 1988; Pullen *et al.*, 1988). However, in newborn mice of the same strains, these potentially self-reactive cells remain part of the repertoire (Schneider *et al.*, 1989; Smith *et al.*, 1989; Ceredig, 1990; Jones *et al.*, 1990). Third, it has been reported that neonatal SP thymocytes are a proliferating population of blast cells *in vivo*, and that they expand in the presence of IL-2 and IL-7 *in vitro*. Adult SP thymocytes do not exhibit these features (Ceredig, 1990; Ceredig and Waltzinger, 1990). And finally, it has recently been demonstrated that SCID mice—which have a defect in the machinery involved in TCR and Ig gene rearrangement—are provoked to produce significant levels of host-derived serum immunoglobulins when populated with neonatal, but not adult, thymocytes (Riggs *et al.*, 1991).

Intrigued by these observations, we wondered whether neonatal mice might have a special repertoire of α:β T cells. Thus, we have made use of polymerase chain reaction (PCR) technology to compare the sequences of TCRs expressed in thymocyte populations prepared from newborn and adult mice. The comparisons indicate that neonatal mice do have a special α:β T cell repertoire that is shaped noticeably by selection events in the thymus.

Results

Sequencing strategy

Our initial approach to determining whether neonatal mice employ a distinct set of α:β TCRs was to compare V_β usage in SP thymocytes from newborn and adult animals. Thymuses from C57Bl/6 (B6) × SJL F_1 mice were disrupted and stained with differentially conjugated anti-CD4 and anti-CD8 monoclonal antibodies (mAbs) together with one of a variety of specific anti-V_β reagents. No striking differences were observed in the relative frequencies of SP thymocytes carrying V_βs 2, 4, 6, 8, 14, or 17a (data not shown).

A second approach was to focus on the junctional diversity associated with one particular β-chain variable region. V_β17a was chosen because we had already generated a large database characterizing its diversity in various adult T lymphocyte populations: $CD4^+$ versus $CD8^+$ lymph

© Oxford University Press

node cells, peripheral versus thymic CD4 single-positives, SP versus DP thymocytes, cells from mice expressing or not expressing the E complex (Candéias *et al.*, 1991a). According to these data, the diversity associated with V_β17a, like that characteristic of any TCR β-chain, is multifaceted. There are contributions from the particular J_β and D_β segments used and from the frame of the D_β segment utilized. There can be exonucleolytic chewing of the 3′ end of V, both ends of D, and the 5′ end of J. So-called P nucleotides are sometimes present at undigested V, D or J termini, almost certainly as a by-product of the normal recombination process. Finally, so-called N nucleotides are often inserted at the V−D or D−J joints by a template-independent polymerase, probably terminal deoxynucleotidyl transferase (TdT). Any one of these facets might be distinct in neonatal mice; this can best be revealed through sequence analysis.

Our sequencing strategy relied on the power of PCR technology. Thymocyte suspensions were prepared from newborn or adult mice and stained with differentially conjugated anti-CD4, -CD8 and V_β17a mAbs. The triple-stained cells were visualized by cytofluorimetry, as exemplified in Figure 1, and were electronically sorted to produce populations of 1000 to 100 000 cells. RNA was isolated from the sorted populations, was converted to cDNA, and the TCR cDNAs subject to PCR amplification using primers from the constant and variable regions. The amplified material was cloned into an M13 vector and multiple clones from each mouse sequenced. The use of amplifications from two to six independent animals for each cell population precluded PCR artefacts; elaborate controls were also performed to rule-out the amplification of contaminant sequences (see Materials and methods).

V_β17a$^+$ TCRs from neonatal SP thymocytes have little N nucleotide addition

Figure 2 presents the junctional region sequences of V_β17a$^+$ TCRs from CD4$^+$8$^-$ thymocytes prepared from neonatal and adult mice. One feature is strikingly dissimilar in the two data sets: the sequences derived from newborn animals show very little diversity resulting from N nucleotide addition, while the adult-derived sequences are rich in this source of variability. The average number of N nucleotides per sequence is only 0.3 in the former case and is 3.0 in the latter.

The data in Figure 2 were subject to a variety of analyses in an attempt to reveal other, perhaps more subtle, differences between the two sets of sequences. As illustrated in Figure 3, there is a slight shift in the average length of the junctional region: the absence of N additions in the TCRs of newborn mice often results in shorter CDR3 segments. Although this shift is rather small (one amino acid on average), it could have a significant influence on the structure of the antigen-recognizing CDR3 loop (Chothia *et al.*, 1988; Davis and Bjorkman, 1988; Claverie *et al.*, 1989).

All other analyses, including plots of D_β usage, extent of exonucleolytic chewing, amino acid composition and amino acid diversity at each position, failed to reveal any substantial differences between the two sequence sets (not shown).

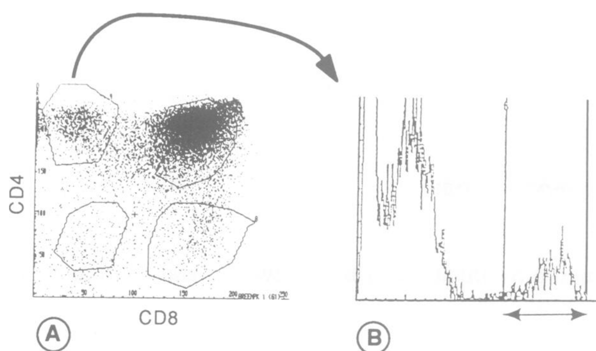

Fig. 1. Cytofluorimetry of triple-stained cells. Thymocytes were stained with anti-CD4, -CD8 and -V_β17a reagents as described in Materials and methods. Results from a typical experiment aimed at sorting CD4$^+$CD8$^-$ V_β17a$^+$ cells are shown. The cytogram (**A**) shows the separation of thymocyte populations based on the expression of CD4 and CD8. The CD4$^+$CD8$^-$ cells were gated as indicated and further analysed according to V_β17a expression. The histogram (**B**) represents the fluorescence profile for V_β17a negative and positive (double-headed arrow) cells in the CD4$^+$CD8$^-$ population. Cells that were CD4$^+$CD8$^-$ and V_β17a$^+$ were collected and RNA prepared immediately.

Other TCRs from neonatal SP thymocytes lack N insertions

To determine whether the under-representation of N nucleotides in neonatal V_β17a$^+$ TCRs is a peculiarity of that particular variable region, we generated sequences from receptors carrying other V_βs and V_αs. Figure 4 lists the junctional region sequences of V_β6$^+$, V_β10$^+$ and V_α1$^+$ TCRs from the CD4$^+$8$^-$ thymocytes of newborn and adult mice. Again, the newborn-derived sequences show a paucity of N insertions. The average number of N nucleotides per sequence is 0.1 for $V_\beta(6+10)^+$ and 0.3 for V_α1$^+$ TCRs from neonates versus 3.4 for $V_\beta(6+10)^+$ and 2.6 for V_α1$^+$ TCRs from adults.

The lack of N region diversity is not due to selection during the DP to SP transition

The relative lack of N region diversity in neonatal TCRs could result from repertoire selection events that take place in the thymus, either positive selection for sequences that carry no or few N nucleotides or negative selection against sequences that carry more. Positive selection to achieve MHC restriction is considered by most to take place during the DP to SP thymocyte transition, coincident with an up-regulation of surface TCR levels (Guidos *et al.*, 1990; Ohashi *et al.*, 1990; Borgulya *et al.*, 1991; Robey *et al.*, 1991; Shortman *et al.*, 1991). Negative selection of some self-reactive T cells also occurs around this time (Hengartner *et al.*, 1988; Pircher *et al.*, 1989; Guidos *et al.*, 1990; Hugo *et al.*, 1991; Shortman *et al.*, 1991); C.Benoist, unpubished; K.Signorelli, unpublished) although others appear to be eliminated somewhat earlier (Kisielow *et al.*, 1988; Sha *et al.*, 1988; Berg *et al.*, 1989; Pircher *et al.*, 1989; White *et al.*, 1989). Thus, we decided to compare N region diversity in the CD4$^+$8$^+$ thymocyte populations from newborn and adult mice.

Figure 5 shows a clear dichotomy in the number of N additions characteristic of V_β17a$^+$ TCRs from DP thymocytes from the two sets of mice. There is on average 0.4 N nucleotides per sequence for newborn animals versus 2.6 for adults; only 1.5% of the newborn set have >2 N nucleotides versus 32% of the adult set. This dichotomy is highly reminiscent of the one observed with SP thymocytes, arguing against a role for selection forces that operate during these two stages.

$V_{\beta}17a^{+}$ TCRs from neonatal $CD4^{-}8^{+}3^{-/lo}$ thymocytes show substantial N nucleotide addition

With the aim of generalizing our findings, we also sorted $CD4^{-}8^{+}$ T cells from a neonatal thymus and sequenced some of their TCRs. Surprisingly, we discovered substantially more N nucleotide addition than was seen in TCRs from the equivalent $CD4^{+}CD8^{-}$ population: 32 of 63 sequences (51%) had N insertions, several as long as five nucleotides. An explanation for this unexpected result might lie in the heterogeneity of the $CD4^{-}8^{+}$ thymocyte population. As has been described (Bluestone *et al.*, 1987; Crispe *et al.*, 1987; Fowlkes *et al.*, 1987; Paterson and Williams, 1987) there are, in fact, two major components to this population: a mature subset expressing high levels of CD3 and TCR $\alpha{:}\beta$ and an immature subset displaying low or undetectable levels of these two markers. Therefore, we refined our analysis by sequencing $V_{\beta}17a^{+}$ TCRs from the $CD3^{-/lo}$ and $CD3^{+}$ components of the $CD4^{-}CD8^{+}$ population.

Figure 6 reveals that N nucleotide addition to $V_{\beta}17a^{+}$ TCRs is restricted to the $CD4^{-}8^{+}3^{-/lo}$ subset of neonatal thymocytes. That the number of N nucleotides characteristic of this immature population differs significantly from that characterizing the more mature DP and SP populations is illustrated in Figure 8 (panel c versus b and a), which graphically emphasizes the difference in frequency of N additions that are greater than two nucleotides and thus encode at least one additional amino acid.

Defective V_{β} 17b TCR transcripts also show substantial N insertions

The observation that immature thymocytes from neonatal mice show substantially more N insertions than mature thymocytes could have two explanations: (i) it may be that the difference reflects waves of dissimilar T cell precursors populating the thymus; (ii) it may be that the difference reflects positive or negative repertoire selection events. To distinguish between these explanations we compared the number of N nucleotides in the $V_{\beta}17a$ and $V_{\beta}17b$ transcripts of B6 × SJL F_1 mice. $V_{\beta}17a$ mRNA (transcribed from the SJL chromosome) is translated and gives rise to cell-surface protein; $V_{\beta}17b$ transcripts

CD4+CD8- Vβ17a

NEWBORN

V	P	N	P	D	P	N	P	J	
GCCAGCAGTCTG				ACAGGG		A		AAGTC	Jβ1.1
GCCAGCAGTCTG				ACAGGGG				CACAG	
GCCAGCAGTCTG				GGGA				CAAAC	Jβ1.2
GCCAGCAGTC				CAG				CAAAC	
GCCAGCAGTCTGTA	TA			AG				AACTC	
GCCAGCAGT		T		GGGACA				CAAAC	
GCCAGCAGTCTG				ACAG				CAAAC	
GCCAGCAGTCTG				GGGGC				CTCCG	
GCCAGCAGTCT				CA				CAAAC	
GCCAGCAGTC				ACAGGGGGC	G			CAACC	Jβ1.5
GCCAGCAGTCTG				ACA		A		ATAAT	Jβ1.6
GCCAGCAGTCTG				GGAC				ATAAT	
GCCAGCAGTCTG				GACAGG				TATAA	
GCCAGCAGTCTG				ACA			AA	TTCCT	
GCCAGCAGTCTGTA				CTGGGGG				ATGCT	Jβ2.1
GCCAGCAGT		T		GGACTGGG				TATGC	
GCCAGCAGT				GAC				AACTA	
GCCAGCAGTCTG				ACA		A		ATGCT	
GCCAGCAGTCTGTA				C				AACTA	
GCCAGCAGTCTGTA	T			TGGGG				CTATG	
GCCAGCAGTCTGTA	T			ACTGGGGG				TATGC	
GCCAGCAGTCTG				GAC				AACTA	
GCCAGCAGTCTGT		T		T				TATGC	
GCCAGCAGTCT		C		C				ACTAT	
GCCAGCAG				GACTGGG				AACTA	
GCCAGCAGTCTGTA	T			GGACAGGGG				GCTGA	
GCCAGCAGTC			C	GGGACTGGGGGGGC	G			CTATG	
GCCAGCAG				GGACAGG				AACTA	
GCCAGCAGTCTG				GAC				AACTA	
GCCAGCAGTCTGT				GGACAGGG				TATGC	
GCCAGCAGTCTG				GG				TGCTG	
GCCAGCAGTCTG				GGGGGC			A	TAACT	
GCCAGCAGTC								ACACC	Jβ2.2
GCCAGCAGTCTGTA				CTGGGGG			CT	AGTGC	Jβ2.3
GCCAGCAGTCTGTA	T			CAGG				GTGCA	
GCCAGCAGTCTG				ACAGGG				GCAGA	
GCCAGCAGTCTG				G				GTGCA	
GCCAGCAGTC				GGGACAGG				GCAGA	
GCCAGCAGTC			C	GGGACAGG		T		TGCAG	
GCCAGCAGTCTG				GACTGG				GCAGA	
GCCAGCAGTCTGTA	TA			AGGG				CACCT	Jβ2.4
GCCAGCAGTCTG				CTGGGG				CAAAA	
GCCAGCAGTCTGTA				CTGGGG				AAACA	
GCCAGCAGTCTG				GGGGGC				CAAAA	
GCCAGCAGTCTG				GAC				AGTCA	
GCCAGCAGTCT				ACTGGGG				AAAAC	
GCCAGCAGTCTG				GGGG				GTCAA	
GCCAGCAGTCTGTA				CAGGG				ACACC	
GCCAGCAGTC				GACAGG				GTCAA	
GCCAGCAGTCTGT				GG				GACAC	Jβ2.5
GCCAGCAGTCTG				GACAGG				GACAC	
GCCAGCAGTC				CTGG				CCAAG	
GCCAGCAGTC			C	GGGAC				ACCAA	
GCCAGCAGTCTG				GACA			TT	AACCA	
GCCAGCAGTCTG				GGGG		T		TGAGC	Jβ2.6
GCCAGCAGTC				AGGG				TGAGC	
GCCAGCAGTCTGTA	TA							CCTAT	
GCCAGCAGTC				ACTGGGGGGG		GATT		TATGA	
GCCAGCAGTCTGTA				CAG				CTCCT	
GCCAGCAGTCTGTA	T			ACAGGGG		T		TGAGC	
GCCAGCAGTCTGT				GGGG				TGAGC	
GCCAGCAGTCTG				GACTGGGGG				TATGA	
GCCAGCAGTCTG				ACTGGGGGG				TATGA	
GCCAGCAGTC			C	GGGACA		CTATGG		TGAGC	
GCCAGCAGTCTGT				CA				TATGA	
GCCAGCAGTCTG				ACAGG				CTATG	

ADULT

V	P	N	P	D	P	N	P	J	
GCCAGCAGTCTGTA				CAGGGG				AAACA	Jβ1.1
GCCAGCAGTC			C	GGGACAGGG		A	G	CAAAC	
GCCAGCAGTCTGTA	T		CC	GGGACA				AAACA	
GCCAGCAGTCTGT		TC		CAGGGG		TT		CACAG	
GCCAGCAGTCT		T		AGGGGGC	GC	G		ACACA	
GCCAGCAGTCTGTA	T			GGGA		GAGGG		GAAGT	
GCCAGCAGTC				CAGGGGGC	G	G		CACAG	
GCCAGCAG			CC	GGGACAGGGG		T		GAAGT	
GCCAGCAGTCT				CAGGGG		AG		ACACA	
GCCAGCAGTCT		TC		ACAGGGG		TGG		AAACA	
GCCAGCAG		CCAGTGGCCCA		ACAGGGG		AAG		AGTCT	
GCCAGCA		CCAA		AGGGG		A		AAACA	
GCCAGCAG		A		ACAGGGGG		TG		ACTCC	Jβ1.2
GCCAGCAG		CCC		GACAGGGGG				AAACT	
GCCAGCAGTC				GGGACA		CCTT		TTCTG	Jβ1.3
GCCAGCAGTCTGTA	TA	T		GGGGG		T		TTCTG	
GCCAGCAGTCTG		GG		GGGAC				CTGGA	
GCCAGCAGTCTGT				CAGG			AA	TTTCC	Jβ1.4
GCCAGCAGTCTGTA								TTCCT	Jβ1.6
GCCAGCAGTCTGTA	T			GGGACA				TATAA	
GCCAGCAGTCTGTA	TA	A		AG				ATTCG	
GCCAGCAGTC		CCAC		GGGG		TT		TGCTG	Jβ2.1
GCCAGCAGTC				CAGGG		AGGAGGG		ACTAT	
GCCAGCAGT		T		GGACAG				ACTAT	
GCCAGCAGTCTGTA	T	TCT		GGGG				ACTAT	
GCCAGCAGTCTGTA	T	GC		CAGGGG		A		AACTA	
GCCAGCAGTCTGT		CTC		CAG		T		CTATG	
GCCAGCACT				ACTGGG		C		ACTAT	
GCCAGCAGTCTG		CA		ACA		T		AACAC	Jβ2.2
GCCAGCAGTCTG		GA		GGGA		G		GTGCA	Jβ2.3
GCCAGCAGTC		CTC		GACTGGGGG		TC		CGCTG	
GCCAGCAGTCT		C		ACAGGG		TT		TGCAG	
GCCAGCAGTCT		CCCC		GACTGGGGGG		ATGG		GAAAC	
GCCAGCAGTC		CCA	C	GGGACAGGG				TGCAG	
GCCAGCAGTC		CCG	CC	GGGACAG			CT	AGTGC	
GCCAGCAGTCT		CCG		CAGG		ACT		AGTGC	
GCCAGCAGTCTGTA	TA	T		GGGGG		TTTCTG		GAAAT	
GCCAGCAGTC				CT		CA	T	AGTGC	
GCCAGCAGTCTG				GGACTGGGGGGGC	GC	A		CAGAA	
GCCAGCAGTCTGTA				CAG		TGT		AGAAA	
GCCAGCAGTCTGT		CT		CTGGGGG		AG		AGTGC	
GCCAGCAGTCTGTA	T	T		TGGGGGGG		GA		TGCAA	
GCCAGCAG				GGGACTGGG		ATT		TCAAA	Jβ2.4
GCCAGCAGTCT				CAGGG				GTCAA	
GCCAGCAGTCT		AAG		TGGG				AACAC	
GCCAGCAGTCTGT		C		CAGGG		CTG		ACACC	
GCCAGCAGTCTG				GGGG		TC		CAAAA	
GCCAGCAGTCTG				GGGGG		A		CAAAA	
GCCAGCAGTCT		TGA		GGGACAGGG				AGACA	Jβ2.5
GCCAGCAGTCTGTA	TA	AC		GGGGGG		A		AAGAC	
GCCAGCAGTCTGT		GG		CAG				CAAGA	
GCCAGCAGTC		C		GACA		AAA		CAAGA	
GCCAGCAGTCT				AGG		C	AG	CTCCT	Jβ2.6
GCCAGCAGTCTG		AA		TGGGGGGGC	GC			ATGAG	
GCCAGCAGTCTGT		TG		CAGGG		AG		ATGAG	
GCCAGCAGTC		CTT	CC	GGGACAGGG		TT		CTATG	
GCCAGCAGTCTG				GG		TCT		CTCCT	
GCCAGCAGTCT		TG		GGGACA		A		CTCCT	
GCCAGCAGTG				GGGGGC	G	A		GAGCA	
GCCAGCAGTCT		T	CC	GGGACA		A		TCCTA	
GCCAGCAGTCT		CC		GACTGGGGGGGC	GC	AGA		GAGCA	
GCCAGCAGTCTGTA	T	GC		GAC		GGATT		ACTAT	
GCCAGCAGTCTG		GAGG		AGGGG		TC		TGAGC	
GCCAGCAGTCTG				GACA		T		CTCCT	
GCCAGCAGTCTGTA								CTCCT	
GCCAGCAG								CTCCT	
GCCAGCAGTCTG				ACAGGGG		T		TGAGC	
GCCAGCAGTCTG		G		CTG		AT		TATGA	

Fig. 2. Junctional region sequences of $V_{\beta}17a$ TCRs from $CD4^{+}CD8^{-}$ thymocytes of newborn (**left panel**) and adult (**right panel**) mice. Sequences were obtained from sorted $CD4^{+}CD8^{-}V_{\beta}17a^{+}$ populations individually derived from the thymuses of four newborn and three adult SJL mice. Germline sequences are known for $V_{\beta}17a$ (Cazenave *et al.*, 1990), D_{β} (Kavaler *et al.*, 1984), and J_{β} (Gascoigne *et al.*, 1984; Malissen *et al.*, 1984). Nucleotides in each sequence are assigned to V, D, or J segments that comprise the junctional region of the rearranged TCR gene. Mono- or dinucleotides representing inverted repeats of undigested termini are grouped as P nucleotides according to Lafaille *et al.*, 1989. Although the addition of more than two P nucleotides at a single TCR joint has been reported in SCID mice (Schuler *et al.*, 1991), possible additions of three P nucleotides were observed so infrequently in these analyses that no more than two P nucleotides at a single joint were assigned. Nucleotides that cannot be assigned to another category are template independent and are thus designated N nucleotides. Only the 3′ end of the V and the 5′ of the J segments are shown. Sequences are grouped according to their J usage, as shown at the right of each panel. Protein sequences corresponding to the adult sequence set are to appear elsewhere (Candéias *et al.*, 1991). The nucleotide sequences are shown here to permit direct comparison with the newborn sequences.

Fig. 3. CDR3 loop lengths. From the newborn and adult sequences presented in Figure 2, the length of the CDR3 segment was determined based on the alignment of TCR variable regions with immunoglobulin variable structures (McCormack *et al.*, 1991). The exact position of CDR3 was defined on the basis of V_H CDR3, e.g. between positions 95 and 107 (residue numbering and V_H CDR3 definition according to Kabat *et al.*, 1987). For each population, the average length is indicated by a bar.

(contributed by the B6 chromosome) carry a point mutation specifying a stop-codon and so do not engender cell-surface protein (Wade *et al.*, 1988). Thus, T cells expressing $V_\beta17b$ transcripts are unselectable, and a difference in the number of N nucleotides inserted into the $V_\beta17b$ versus $V_\beta17a$ transcripts of neonatal mice would argue for the influence of selection forces and against a difference in the rearrangement machinery of successive precursor waves.

Figure 7 lists $V_\beta17b$ sequences from the $CD4^+8^+$ thymocytes prepared from neonatal B6 ×SJL mice; it was not technically feasible to obtain a similar set of sequences from $CD4^+8^-$ thymocytes (see Materials and methods). These data are to be compared with the $V_\beta17a$ sequences from DP and SP cells already presented in Figures 5 and 2, respectively. Examination of the three data sets reveals that the defective $V_\beta17b$ transcripts show significantly more N region diversity than do the normal $V_\beta17a$ transcripts, particularly those from fully mature SP cells. This point is illustrated graphically in Figure 8 (d versus b and a), where the comparison of sequences carrying greater than two N nucleotides is most striking: $\leq 3\%$ of $V_\beta17a$ sequences versus 16% of $V_\beta17b$.

Discussion

Two major points emerge from our analysis of junctional region sequences from the $\alpha{:}\beta$ TCRs of neonatal and adult mice. First, it is clear that the newborn-derived sequences carry far fewer N nucleotide additions, and second, this paucity of N insertions is exaggerated by selection events that operate during T cell differentiation in the thymus. Both of these points merit further comment.

Reduced N nucleotide addition

Our observation that the receptors on neonatal $\alpha{:}\beta$ T cells show only very limited N region diversity—and parallel observations made simultaneously by others (Feeney, 1991; McCormack *et al.*, 1991)—are reminiscent of recent findings on the B cell repertoire of newborn animals. Several reports have established that the Ig heavy chain transcripts from fetal and neonatal mice have few if any N nucleotide additions while adult IgH transcripts almost always have N additions, often as long as 12 nucleotides (Gu *et al.*, 1990; Carlsson and Holmberg, 1990; Feeney, 1990; Meek, 1990; Bangs *et al.*, 1991).

The $\gamma{:}\delta$ T cell repertoire appears similarly dichotomous. It was recognized several years ago that adult δ chain transcripts carry substantially more N nucleotide additions than the corresponding fetal transcripts (Elliott *et al.*, 1988). This observation has more recently been extended to adult versus fetal/newborn γ chain transcripts (Lafaille *et al.*, 1989; Aguilar and Balmont, 1991). However, it is important to note for $\gamma{:}\delta$ T cells (though not for $\alpha{:}\beta$ T cells nor for B cells) that the diminished level of N addition to the joints of perinatal TCRs is coupled with a near-complete absence of exonuclease digestion. Consequently, for any particular $V_\gamma-J_\gamma$ or $V_\delta-J_\delta$ combination, the junctional regions are essentially homogeneous in perinatal animals.

As discussed by Davis and Bjorkmann (1988), the complexity of the Ig and TCR repertoires is enormously dependent on N region diversity, especially in the latter case where the combinatorial diversity is not so impressive. Thus, the marked reduction in N nucleotide addition to neonatal TCR and Ig transcripts is bound to have a significant influence on the newborn mouse's repertoire gestalt. One has to wonder whether the reduction is pure ontogenetic happenstance or whether it has a purpose. Certainly, there have been reports of increasing TdT activity through ontogeny (Gregoire *et al.*, 1979; Rothenberg and Triglia, 1983) although these deal with complex mixtures of lymphocytes and are therefore difficult to evaluate in the present context. Certainly, one can not argue with the fact that all of the sequencing studies, including our own, document an overall increase in the number and size of N additions with age. Yet, we have seen clear indications that the lack of N nucleotides in neonates is exaggerated by selection and therefore presumably with some purpose.

Selection on the level of N insertions

The initial indication that there might be selection on the degree of N region diversity in neonatal mice was the demonstration that all N nucleotides found in the $V_\beta17a^+$ TCRs from $CD4^-CD8^+$ thymocytes were confined to the $CD3^{-/lo}$ subset. This subset is generally considered to be rather immature, a poorly understood interjection between the DN and DP populations (Bluestone *et al.*, 1987; Crispe *et al.*, 1987; Fowlkes *et al.*, 1987; Paterson and Williams, 1987). Thus, progression through different stages of thymocyte differentiation in the neonate ($CD3^{-/lo}$ SP → DP → $CD3^+$ SP) is accompanied by a reduction in N insertions, particularly those which are >2 nucleotides long and therefore increase the CDR3 length by at least one amino acid (20% → 1.5% → 3%). This suggests that some type of selection may be occurring at the immature SP → DP transition or soon afterwards. There is, however, an alternative explanation: that DP cells represent the end of a first wave of thymocytes impoverished in TdT activity while immature SP cells lead off a second wave of cells that have ample TdT. Such an explanation is not consistent with the claim of Jotereau *et al.* (1987) that the initial seeding

CD4+CD8-

NEWBORN

Vβ6

Vβ6	P	N	P	D	P	N	P	J	
GCCAGCAGTATAG				GGGGC				GAAGT	Jβ1.1
GCCAGCAGTA				CAGGGG				CAAAC	
GCCAGCAGTAT				GACAG				CAAAC	Jβ1.2
GCCAGCAGTATAG				GACAGGGGC				TGGAA	Jβ1.3
GCCAGCAGTA								CCAAC	Jβ1.4
GCCAGCAGTATA				CAG				TATAA	Jβ1.5
GCCAGCAGTATAG				CTGGGGG				CTATG	Jβ2.1
GCCAGCAGTATAG				GGACA				TATGC	
GCCAGCAGTAT				GG				GTCAA	Jβ2.4
GCCAGCAGTATA				CTGGGGGGGC	G			CCAAG	Jβ2.5
GCCAGCAGTATAG				GGGGGGC	GC	G		AGACA	
GCCAGCAGTAT				GGACTG				CTATG	Jβ2.6
GCCAGCAGTATAG				ACTGGGGGG				CTATG	
GCCAGCAGTATAG				GGACAGGG				TATGA	
GCCAGCAG				ACTGGG				CTATG	
GCCAGCAGTATA		AT		AGGG				TATGA	

Vβ10

Vβ10	P	N	P	D	P	N	P	J	
GCCAGCAGCTA				C				AGTCA	Jβ2.4
GCCAGCAGC		C	CC	GGGACAGGGGG				CACCT	
GCCAGCAG				T				CAAGA	Jβ2.5
GCCAGCAGCTA				CTGGGGG				TATGA	Jβ2.6
GCCAGCAGC				GACTGGGG				TGAGC	

Vα1

	Vα1	P	N	P	J (PHE)	Jα
5H	GCAGTGAGC				AGAGGTTCAGCCTTAGGGAGGCTGCATTTT	LB2
8.1	GCAGCTAGTGAGC		C		TAATACAGGAAACTACAAATACGTCTTT	JA19
	GCAGCTAG		GG		ATTATGGGGGCAGTGGCAACAAGCTCATCTTT	TT11
	GCAGCTAG				GAACATGGGCTACAAACTTACCTTC	14T
	GCAGCTAGTG				CTGGAGCTAACACTGGAAAGCTCACGTTT	C5
	GCAGCTAGTG				CAGGAGGTGCAGATAGACTCACCTTT	JA80
	GCAGCTAGTGAGCA				GGGTTACCAGAACTTCTATTTT	JA65
A	GCAGT		A		ACAGGCAATACTGGAAAACTCATCTTT	TA57
B	GCAG				ACACAAATGCTTACAAAGTCATCTTT	TA13
C	GCAGCAAGC				ACAAATGCTTACAAAGTCATCTTT	TA13
D	GCAGTGAGC				AGTGGAGGCAGCAATTTACAAACTGACATTT	3DT
	GCAG		C		TACTGGAGGCAATAATAAGCTGACTTTT	2B4
E	GCAGTTAGT				GACACAAATGCTTACAAAGTCATCTTT	TA13

ADULT

Vβ6

Vβ6	P	N	P	D	P	N	P	J	
GCCAGCAGT				GGGACAG				ACACA	Jβ1.1
GCCAGCAGTA		AT	CC	GGGACAGGG		TGGG		TCCGA	Jβ1.2
GCCAGCAGTA		G		CAGGGG				TGGAA	Jβ1.3
GCCAGCAGT		C		CA		A		CCAAC	Jβ1.4
GCCAGCAGTATA				AG		TA		CCTAT	Jβ1.5
GCCAGCAGT		CCT		GACAG			T	AGTGC	Jβ2.3
GCCAGCAGTATA		A		ACTGGG				GTGCA	
GCCAGCAGTAT		CAT		CAGGG		AAA		AAAAC	Jβ2.4
GCCAGCAGT		CCCC		GGACTGGGGGG		TTC		GACAC	Jβ2.5
GCCAGCA		AT		GGGACAGGGGG		A		TATGA	Jβ2.6
GCCAGCAG				GGACTGGGG				CTCCT	
GCCAGCAG		ACA		GGGACTGGGGG		A		CTATG	

Vβ10

Vβ10	P	N	P	D	P	N	P	J	
GCCAGCAGC		C	CC	GGGACA		AA		AAACA	Jβ1.1
GCCAGCAGC		AT		GACAGGGG				ACTAT	Jβ2.1
GCCAGCA		CCCT		GACTGG		TCTG		GCTGA	
GCCAGCAGCT				GGGGGGC	G	G		CCGGG	Jβ2.2
GCCAGCAGC		CCAC		CTGGG		AG		AGACA	Jβ2.5
GCCAGCAG		TTC		GGACTGGGGGGG		G		TCCTA	Jβ2.6
GCCAGCAG		TCA		GACTGGGG		CGTTCGG		TGAGC	
GCCAGCAGCT		CCC		ACTGGGGGGGC			AG	CTCCT	
GCCAGCAGCT		T		AG		A		CTATG	
GCCAGCAGCT				GGGACA		A		CTCCT	

Vα1

	Vα1	P	N	P	J (PHE)	Jα
5H	GCAG		CG		GACACAAATGCTTACAAAGTCATCTTT	TA13
	GCAGTGAG		G		TACCAGGGAGGCAGAGCTCTGATATTT	A10
8.1	GCAGCTAG		AGG		GGGAGGCAGAGCTCTGATATTT	A10
	GCAGCTAGT				AATACAGGAGGTGCAGATAGACTCACCTTT	JA80
	GCAGCTAGTGAGCA		G		ACAGGCAATACCGGAAAACTCATCTTT	TA57
A	GCAGTGA		CT		TCTGGCAGCTGGCAACTCATCTTT	TA28
	GCAG		CAAGGGCG		TCTAATTACAACGTGCTTTACTTC	14.4
	GCAGTGA		CCC		CAAATTCTGGGACTTACCAGAGGTTT	39
	GCA		AAGGG		CAATAACAGAATCTTCTTT	MD13
C	GCAGCAAGCG				ACCAGGGAGGCAGAGCTCTGATATTT	A10
	GCAGCAAGCG		GGGGGG		ATTATGGGAGCAGTGGCAACAAGCTCATCTTT	JATT11
	GCAGCAAGC		AGT		ACAGGCAATACCGGAAAACTCATCTTT	TA57
	GCAGCAAGCG				ATGGGGGCAGTGGCAACAAGCTCATCTTT	TT11
	GCAGCAAGC		AT		AACTTCAAGTGGCCAGAAGCTGGTTTTT	JA84
	GCAGCAAGC		CCG		AATACAGGAGGTGCAGATAGACTCACCTTT	JA80
	GCAGC		CAC	TG	CAAATTCTGGGACTTACCAGAGGTTT	39
D	GCAGTGAGC		CTC		AATAACAATAACAGAATCTTCTTT	MD13
E	GCAGTTAGTGA		CGG		CACTGGAGCTAACACTGGAAAGCTCACGTTT	C5
F	GCAG		AGG		ACAACGTGCTTTACTTC	14.4

Fig. 4. Junctional region sequences of $V_\beta6$ (**upper**), $V_\beta10$ (**middle**), and $V_\alpha1$ (**lower**) TCRs from $CD4^+CD8^-$ thymocytes of newborn (**left panel**) and adult (**right panel**) mice. Sequences were obtained from sorted $CD4^+CD8^-$ populations from the thymuses of two newborn and two adult B6 × SJL F_1 mice. The following germline sequences are known: $V_\beta10$ (Hirama *et al.*, 1991), $V_\alpha1$-5H (Winoto *et al.*, 1985), $V_\alpha1$-8.1 (Hirama *et al.*, 1991), and several J_α segments: A10, 39, TT11 (Malissen *et al.*, 1988; M.Malissen, personal communication), JA19, JA65, JA80, JA84, JATT11 (Winoto *et al.*, 1985). Additional sequences were assigned with a high level of confidence through extensive analysis of numerous sequences along with the availability of non-germline sequence data for $V_\beta6$ (Smith *et al.*, 1990) and additional J_α segments: TA28, TA57 (Arden *et al.*, 1985), 2B4, C5, LB2, MD13 (Becker *et al.*, 1985), TA13 (Wilson *et al.*, 1988); 14.4 (Winoto *et al.*, 1986) and 3DT (Yague *et al.*, 1988). Nucleotides from $V_\beta6$ and $V_\beta10$ sequences were assigned as described in the legend to Figure 2. Since V_α genes do not contain D regions, nucleotides in $V_\alpha1$ sequences were assigned as V, J, P or N. Members of the $V_\alpha1$ family which have not been previously described are designated A to F. J regions are aligned at the conserved phenylalanine (PHE) residue; nucleotide differences from published J_α segments are underlined. It should be noted that the paucity of P nucleotides in the newborn and adult $V_\alpha1$ sequences may not reflect the actual frequency of P nucleotide addition in these populations due to the constraints in assigning V and J regions without known germline sequence data. The newborn $V_\beta6$ sequences that are shown were chosen randomly from a larger set.

of a fetal thymus by lymphocyte precursors gives rise to all of the $Thy1^+$ thymocytes present until ~7 days after birth.

Nonetheless, we sought a completely independent indication that the level of N insertion is subject to selection in neonates. This was obtained by comparing defective $V_\beta17b$ transcripts with their expressed $V_\beta17a$ counterparts. The former had more N nucleotides per sequence and more sequences with >2 nucleotides. Thus, we conclude that the paucity of N insertions in neonatal $\alpha:\beta$ TCRs is exaggerated by selection forces operating during thymocyte differentiation. This may entail either positive selection for those cells displaying TCRs with no or few N nucleotides or negative selection against those cells expressing TCRs with more N additions.

These observations are, once again, reminiscent of recent findings on the B cell repertoire. L.Carlsson, C.Overmo and D.Holmberg (submitted for publication) compared N nucleotide insertions in the non-productive versus productive IgH rearrangements of neonates and discovered significantly more inserts in the former. However, this was true for the 7183 but not the J558 family, raising the interesting possibility that this type of selection may operate only on a subset of variable regions.

There has been much discussion about positive selection of fetal $\gamma:\delta$ T cells expressing TCR transcripts essentially devoid of N nucleotides (Lafaille *et al.*, 1990; Aguilar and Belmont, 1991). But, as noted in the previous section, these TCR transcripts also show essentially no exonuclease nibbling, resulting in near-identical TCR junctional regions and quasi-monoclonal T cells. It is very likely that the selection forces on these T cells are unrelated to those influencing neonatal $\alpha:\beta$ T cells and B cells.

Indeed, the mechanics of selection on N region diversity at the time of the immature SP → DP transition appear rather enigmatic. First of all, this selection appears to operate on cells expressing almost undetectable levels of CD3 and $\alpha:\beta$ TCR. Secondly, it is difficult to envisage a mechanism by which such a diverse set of receptors could be selected. Size

CD4+CD8+ Vβ17a

NEWBORN

V	P	N	P	D	P	N	P	J	
GCCAGCAGTCTG				G		T		AAACA	Jβ1.1
GCCAGCAG				ACAGGG				CACAG	
GCCAGCAGTCTG		GA		GACAG				CAAAC	Jβ1.2
GCCAGCAGTC				GGGACAG			TG	CAAAC	
GCCAGCAGTCTGTA	T			GGAC				ACTCC	
GCCAGCAGTC			C	GGGACAGG				CAAAC	
GCCAGCAG				ACAGGG				TGGAA	Jβ1.3
GCCAGCAGTC			C	GGGACA				TTCTG	
GCCAGCAGTCTG				AC				TTCCA	Jβ1.4
GCCAGCAG				GACTGGGGGGGC				TGAGC	Jβ2.1
GCCAGCAGTCTG				GGGGGGC	G			CTATG	
GCCAGCAGTCTGTA				CTGG				AACTA	
GCCAGCAGTC		A		GGAC				AACTA	
GCCAGCAGTC			C	GGGACAGGGGG				CTGAG	
GCCAGCAG				GACAGGG				TATGC	
GCCAGCAG			CC	GGGACTGGGGG				TATGC	
GCCAGCAGTCTGT				GGGAC				AACTA	
GCCAGCAGTCTGTA				TGGGG				ACTAT	
GCCAGCAGTCTG				ACAGGG				AACAC	Jβ2.2
GCCAGCAGTCT			C	GGGACAGGG		AG		CACCG	
GCCAGCAGTCTG				GGGGGC				GCAGA	Jβ2.3
GCCAGCAGACTG				GGGGG				AAACA	Jβ2.4
GCCAGCA		T		A				AGTCA	
GCCAGCAGTCTGTA	TA		C	GGGACTGGGGG			CT	AGTCA	
GCCAGCAGTC				ACTGGGGGGGC	G	A		TCAAA	
GCCAGCAGTC				GGGACT				ACCAA	Jβ2.5
GCCAGCAGTCTGT				GGGGGG				AAGAC	
GCCAGCAGACTG				GGG				CAAGA	
GCCAGCAGTCTGT		G		GGGACTGGG				CTATG	Jβ2.6
GCCAGCAGTC			C	GGGACTGGGGGG				TGAGC	
GCCAGCAGTCTG				ACAGGGGG				CTCCT	
GCCAGCAGTC			C	GGGACAGGGGG		GGCGC		GAGCA	
GCCAGCAGTCTGT				GGGGGC	G			CTATG	

ADULT

V	P	N	P	D	P	N	P	J	
GCCAGCAG		CC		CAGGG				AACAC	Jβ1.1
GCCAGCAG				AGGGGG		GGGCGCG		AACTC	Jβ1.2
GCCAGCAGTCTG				GGGACA		A		CAAAC	
GCCAGCAGTC		C		ACAGGGGG		A		CAACG	Jβ1.4
GCCAGCAGTC		C		CAGGGG				CAACG	
GCCAGCA		CCAT		GGGACA		CC		ATAAT	Jβ1.6
GCCAGCA		T		TGGGGGG		AAGGG		TGCTG	Jβ2.1
GCCAGCAGTC		C		GGACTGGGGG		A		ATGCT	
GCCAGCAGTCTG				GAC		GC		TAACT	
GCCAGCAGTCTG		C		CAGGG			TA	TAACT	
GCCAGCAG			C	GGGACAG		ATT	CT	AGTGC	Jβ2.3
GCCAGCAGTC		G	CC	GGGACAGGG		TTG		GTGCA	
GCCAGCA		A		A				AGTGC	
GCCAGCAGTC			C	GGGAC		CA		GCAGA	
GCCAGCAGTCTG				GGACTGGGG		AGGGC		TCAAA	Jβ2.4
GCCAGCAGTCTGT				GGGGG		T	CT	AGTCA	
GCCAGCAG		CGCATTA		ACTGGG		AAAC		CCTTG	
GCCAGCAG				ACTGG		AA		AGTCA	
GCCAGCAGTCTGT			C	GGGACAGGG				AAACA	
GCCAGCAGTCTGTA	TA			AGGGGGC				TCAAA	
GCCAGCAGTCTG				GGGG		A	T	AGTCA	
GCCAGCAGTC		CAA		TGGGGG		A		TCAAA	
GCCAGCAGTCTGT				CTGGGGGG				AACCA	Jβ2.5
GCCAGCAG		AC		ACAGGGG		AG		ACCAA	
GCCAGCAGTC		A		A				CAAGA	
GCCAGCAGTCT		CAT	CC	GGGACAGGGGGC		TT		GACAC	
GCCAGCAG			CC	GGGACTGGGGG				AACCA	
GCCAGCAGTCTGTA		CCCT		GGGACTGGGG		T		CCAAG	
GCCAGCAGTCTGT				CAGGG		C		ACCAA	
GCCAGCAG		A		GGGACTGGGGG			T	AACCA	
GCCAGCAG		A		GGGACTGGGG				CCTAT	Jβ2.6
GCCAGCAGTCTG				ACAG		AT		TATGA	
GCCAGCAGTC		C		GACTGG		A	AG	CTCCT	
GCCAGCAGTCT		CG		CAG		CA		TATGA	

Fig. 5. Junctional region sequences of $V_{\beta}17a$ TCRs from $CD4^+CD8^+$ thymocytes of newborn (**left panel**) and adult (**right panel**) mice. Sequences were obtained from sorted $CD4^+CD8^+$ thymocytes from the same newborn and adult mice (B6 × SJL F_1) as in Figure 4. Nucleotide assignments were made as described in the legend to Figure 2. The newborn sequences that are shown were chosen randomly from a larger set.

of the CDR3 segment might be the determining feature, but the CDR3s of neonatal TCRs are on average only one amino acid shorter than those of adult TCRs. A more subtle feature might instead be determinant, e.g. the alignment of particular amino acid pairs in the CDR3 loop. These hypotheses are testable in transgenic mouse system.

Implications

The observation that neonates have a special repertoire of $\alpha:\beta$ T cells is both provocative and of practical value.

It is provocative in view of the aforementioned parallel findings on the neonatal B cell repertoire: IgH transcripts from fetal or newborn mice also have little N region diversity, and this paucity seems also to be exaggerated by selection events. Interestingly, neonatal B cells have been reported to exhibit a distinct profile of reactivities. The antibodies which they produce tend to show multireactive, 'connective', and autoreactive specificities of generally low affinity (for reviews and references, see Hayakawa and Hardy, 1988; Casali and Notkins, 1989; Coutinho *et al.*, 1989; Kearney *et al.*, 1989). By analogy, one is led to question whether neonatal T cells might also display low-affinity, multireactive or autoreactive receptors—the shorter CDR3 loop perhaps reducing contacts with peptides in the groove of the MHC molecule. Any such cells might be more prone to recognize simple, recurring epitopes on bacteria and parasites, or might provide some form of non-specific help to encourage development of the B cell repertoire, or both. It is intriguing, then, that the $V_{\beta}11^+$ T cells which expand in neonatally thymectomized mice show an enhanced reactivity to the 65 kDa heat shock protein from *Mycobacterium bovis* (Iwasaki *et al.*, 1991).

Our observation that neonate TCRs have few N additions is of practical importance because it may allow one to 'tag' T lymphocytes produced perinatally, just as has been done recently for B lymphocytes (Gu *et al.*, 1990). T cells populating the periphery of newborn mice include potentially autoreactive cells that have escaped intrathymic clonal deletion. It has been hypothesized that the autoreactive T cells dormant in normal adult mice (e.g. MBP-reactive, islet cell-reactive, collagen-reactive) may derive from these early emigrants (Schneider *et al.*, 1989; Smith *et al.*, 1989; Jones *et al.*, 1990). Such a scenario would be consistent with the observation that neonatally thymectomized mice develop

NEWBORN — **CD4-CD8+ CD3+ Vβ17a**

V	P	N	P	D	P	N	P	J	
GCCAGCAGTCTGTA				CAGGG				CAAAC	Jβ1.1
GCCAGCAGTCTGTA	T			CAGGGGC				CACAG	
GCCAGCAGTCTG				GGACAG				TATAA	Jβ1.6
GCCAGCAGTC			C	GGGAC				ACTAT	Jβ2.1
GCCAGCAGTCTG				ACAGGGGC				GCTGA	
GCCAGCAGTC			C	GGGACTGGGGGGC	G			CTATG	
GCCAGCAGTCTG				GGACAGG				GTGCA	Jβ2.3
GCCAGCAG				GACTGGG				AGTGC	
GCCAGCA				CTGGGGGGC				TCAAA	Jβ2.4
GCCAGCAGTCTG				GGG				AGTCA	
GCCAGCAGTCTG				ACTGGGGGGC				AGACA	Jβ2.5
GCCAGCAGTC			CC	GGGACAGGG				GACAC	
GCCAGCAGTCTGTA								TGAGC	Jβ2.6
GCCAGCAGTCTGT				GGGGG			AG	CTCCT	

NEWBORN — **CD4-CD8+ CD3-/lo Vβ17a**

V	P	N	P	D	P	N	P	J	
GCCAGCAG		CCC		GGACAG			TG	CAAAC	Jβ1.2
GCCAGCAGT		T		CAG		AC		AACTC	
GCCAGCAGTC				AGGGGG				CCTAT	Jβ1.6
GCCAGCAGTC			C	GGGACTGGGGGG				AACTA	Jβ2.1
GCCAGCAGTC				GACTGGG				TAACT	
GCCAGCAGTCTGTA	TA	AC		ACA		A	A	TAACT	
GCCAGCAGTCTG				GACTGGGGGGGC				GCTGA	
GCCAGCAGTCTG				GG				CACCG	Jβ2.2
GCCAGCAGTC				ACAGGG				GTGCA	Jβ2.3
GCCAGCAGTCTG		CA		ACTGGGGGG				AGTGC	
GCCAGCAGTCTG				ACTGGG				GTGCA	
GCCAGCAGTC				AGG				GTGCA	
GCCAGCAGTCTG				GGGGGGC	GC			GCAGA	
GCCAGCAGTCTGT		CC		ACTGGG		CAGGG		AGTCA	Jβ2.4
GCCAGCAGTCTGT				TGGGGG				AGTCA	
GCCAGCAGTCT			CC	GGGACAGG				ACACC	
GCCAGCAGTCT		AA		ACTGGG				GTCAA	
GCCAGCAGTCTGT				GGG				AAAAC	
GCCAGCAGTCTG				GGGG				ACACC	
GCCAGCAGTC			C	GGGACTGGGGGGC		C		AAACA	
GCCAGCAGTCTG				GACAGG				GTCAA	
GCCAGCAGTC				GACTG				AGTCA	
GCCAGCAGTC		A	CC	GGGACA		A		TCAAA	
GCCAGCAG			CC	GGGACAGGGG				CAAGA	Jβ2.5
GCCAGCAGTCTG				GGGGG				CAAGA	
GCCAGCAGTC		C		GGACTGGGG		A	TT	AACCA	
GCCAGCAGTCT			C	GGGACAGG		AGG		GAGCA	Jβ2.6
GCCAGCAGTCTG		GC		CAGGGG		T		TATGA	
GCCAGCAGTCTGT		TG		CTGGGGG		AGGA		GAGCA	
GCCAGCAGTCTG				GGGG				TGAGC	
GCCAGCAGTCTG				GGACAGG				TGAGC	
GCCAGCAGTCTG				GACA				CCTAT	
GCCAGCAGTCTGTA	TA			GGC		AG		TATGA	
GCCAGCAGTCTGTA	TA			GGGACAG				CTCCT	
GCCAGCAGTCTGTA								TGAGC	

Fig. 6. Junctional region sequences of $V_{\beta}17a$ TCRs from $CD4^-CD8^+CD3^+$ (**upper panel**) and $CD4^-CD8^+CD3^{-/lo}$ (**lower panel**) thymocytes from newborn mice. Sequences were obtained from sorted $CD4^-CD8^+$ cells that were $CD3^+$ or $CD3^{-/lo}$ from four independent experiments representing six newborn B6 × SJL F_1 mice. See legend to Figure 2 for additional information.

NEWBORN **CD4+CD8+ Vβ17b**

V	P	N	P	D	P	N	P	J	
GCCAGCAGTCTG		AGC		CA				CAAAC	Jβ1.2
GCCAGCAGTCTG				GAC				ACTCC	
GCCAGCAGTCTGTA	T			GGACAGGG				CAAAC	
GCCAGCAGTCTGTA				GGG				CTCCG	
GCCAGCAG				GGACAGGG				CAAAC	
GCCAGCAGTCTGTA	T			GGGACA				TCTGG	Jβ1.3
GCCAGCAG				GGGA				TGGAA	
GCCAGCAGTCTGTA	T			ACA			A	TTCTG	
GCCAGCAGTCTGTA	T							TGGAA	
GCCAGCAGTCT		TC		ACAGGGGGC		TA		CAACG	Jβ1.4
GCCAGCAGTCTGT				GGGGGC	G			CCAAC	
GCCAGCAGTC			C	GGGACAGGG				CGAAA	
GCCAGCAGTCTG				ACAGGG				AACGA	
GCCAGCAGTCTGT				GACAGGGG				CAACC	Jβ1.5
GCCAGCAGTCT		A		GACA				AACAA	
GCCAGCAGTC				ACAGGG				ATTCG	Jβ1.6
GCCAGCAG				GGACTGG				TATGC	Jβ2.1
GCCAGCAGTCTG				GACTGGGGG				TATGC	
GCCAGCAGT		GT		GACTGGGGGG		TC		TAACT	
GCCAGCAGTCT				CTGGG				AACAC	Jβ2.2
GCCAGCAGTCTGTA	T			GGACT		CGAG		GAAAC	Jβ2.3
GCCAGCAGTC		C		GACTG				GAAAC	
GCCAGCAGTCTG				GACTG				GTGCA	
GCCAGCAGTCTGTA	T			AGGG				GTGCA	
GCCAGCAGTCTG				GGACAGG				GTGCA	
GCCAGCAGTCTG				ACTGGGG				GTCAA	Jβ2.4
GCCAGCAGTCTGT		GG		CTGGGGG		A	CT	AGTCA	
GCCAGCAGTC			CC	GGGACTGGGGG				GTCAA	
GCCAGCAG				GGGC				CAAAA	
GCCAGCAGTCTG		A		ACTGGG		CC		AACCA	Jβ2.5
GCCAGCAGTC		A		GGGGC				GACAC	
GCCAGCAGTCTG				GGACAGGGG				AGACA	
GCCAGCAGTCTGT		T		CAGGGGGC	GC	C		ACACC	
GCCAGCAGTCTG				GGACAGGGG				CCAAG	
GCCAGCAGTCTG				GACTGGGGGGG				CCAAG	
GCCAGCAGT				GACAGGGGG		GGGGGG		CAAGA	
GCCAGCAGT		GT		GACTGGGGG				AGACA	
GCCAGCAGT		T		GGGACTGGGGGGGC				GAACA	Jβ2.6
GCCAGCAGTCTG			CC	GGGACTGGGGGG				ATGAA	
GCCAGCA				CTGGGGGG				TGAAC	
GCCAGCAGTCTGTA		C		CTGGGGGGGC	G	G		GAACA	
GCCAGCAGTCTG				CTGGGGGGG				TGAAC	
GCCAGCAG			CC	GGGACT				TATGA	
GCCAGCAGTCTGT				GGGACTGGGGGGGC				CTATG	
GCCAGCAGT				GGACTGG			AG	CTCCT	
GCCAGCAGTCTGT				GGACTGGGG				CCTAT	
GCCAGCAGTCTGT		T		CTGGGGG		ACGG		GAACA	

Fig. 7. Junctional region sequences of V_β17b TCRs from CD4$^+$CD8$^+$ thymocytes of newborn mice. Sequences were obtained from CD4$^+$CD8$^+$ thymocytes from five newborn B6 × SJL F1 mice. Sorted thymocytes from three of these mice were also used for experiments shown in Figures 4, 5 and 6. The germline sequence of V_β17b is known (Cazenave *et al.*, 1990). Nucleotides were assigned as described in the legend to Figure 2.

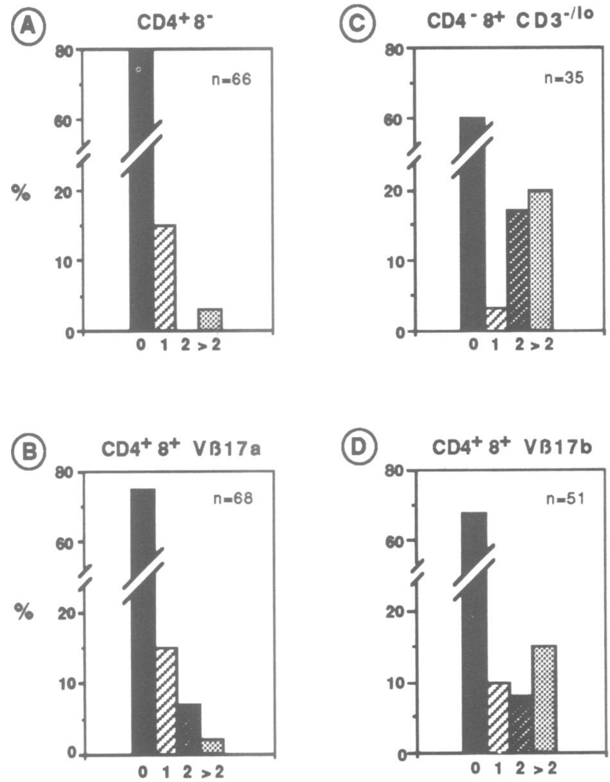

Fig. 8. N nucleotide additions in different thymocyte populations of newborn; CD4$^+$CD8$^-$ V_β17a (**A**), CD4$^+$CD8$^+$ V_β17a (**B**), CD4$^-$CD8$^+$CD3$^{-/lo}$ V_β17a (**C**), and CD4$^+$CD8$^+$ V_β17b (**D**). Newborn sequence data from Figures 2, 5, 6 and 7 are represented graphically according to the percentage of samples with 0, 1, 2 or >2 N nucleotide additions. Sample size (n) is shown and may be greater than the number of sequences shown in the corresponding figure. The graph of CD4$^+$CD8$^+$ V_β17b (**D**) includes data from 5 out of 73 CD4$^+$CD8$^+$ V_β17a sequences that were out-of-frame (not shown) since the resulting transcripts would also be non-functional and thus non-selectable.

multiple organ-specific autoimmune diseases (Tung *et al.*, 1987; Sakaguchi and Sakaguchi, 1990). It is now possible to test this hypothesis, although such an analysis will have to be performed at the population level since the normal adult repertoire does include some TCRs devoid of N insertions (Figure 2).

A major goal should now be to establish experimentally the precise functional consequences of a special $\alpha:\beta$ T cell repertoire in neonatal mice.

Materials and methods

Mice

SJL or B6 × SJL mice were maintained in our animal facility at Strasbourg. Neonatal mice were almost always taken within 24 h of birth. Adults were 5–8 weeks of age.

Thymocyte populations

Thymocytes were prepared from individual mice and were stained with a cocktail of mAbs: anti-V_β17a (KJ23; Kappler *et al.*, 1987b) followed by Texas red-tagged anti-mouse IgG, phycoerythrin-conjugated anti-CD4 (Becton-Dickenson) and fluorescein-conjugated anti-CD8 (Becton-Dickenson). Staining and subsequent sorting of the triple-labelled cells were performed as previously described (Lemeur *et al.*, 1985; Benoist and Mathis, 1989). Populations of 3×10^3 to 1×10^5 were routinely obtained and one or two million HeLa cells were added to each as a carrier.

In one set of experiments anti-V_β17 was replaced by an anti-CD3 mAb (KT-3; Tomonari, 1988).

PCR amplification and sequencing

Details of the amplification and sequencing procedures have recently been published for V_β17a$^+$ TCRs (Candéias *et al.*, 1991a) and for diverse TCRs (Candéias *et al.*, 1991b). For V_α1$^+$ TCRs, we followed the same procedures using the oligonucletoides OT68 [GTCCTGACCTCGCATGCC-AGCAGAGCCCAGAATTCCCTC] and NJ109 [CGGCACATTGATTTG-GGAGTC] for the first amplification and OT68 and NJ110 [TCTCGA-ATTCAGGCAGAGGGTGCTGTCC] for the second.

We would like to emphasize that elaborate precautions were taken to prevent sample contamination, a major problem of PCR-based techniques. All solutions were aliquoted, and aliquots used only once. Aside from the customary negative controls, a mock sample was processed along with each set of experimental samples: droplets of PBS without any cells were sorted and the entire procedure continued, including the screening of M13 plaques. This control ruled out contamination at any step along the way.

In all figures except Figure 7, only in-frame sequences are presented.

Screening for V_β17b transcripts

Since we found that untranslated transcripts are much rarer than translated ones (1:100), we were obliged to screen for clones carrying V_β17b amplification products. The oligonucleotide ACAGAGCTACAGTG permitted discrimination between the V_β17a and V_β17b products derived from B6 × SJL F_1 mice. Hybridization was performed at 40°C in 6× SSC/2× Denhardt's for 3 h and the filters were washed at 37°C or 40°C in 6× SSC for at least 1.5 h.

So few V_β17b clones were obtained from the DP population that we did not screen for clones from the much smaller SP population.

Acknowledgements

We would like to thank R.Ceredig for helpful discussions, D.Holmberg for unpublished information, P.Repis for maintaining mouse lines, S.Vicaire, C.Ebel, P.Rosay and J.Hergueux for excellent technical assistance, and P.Bohn (especially) for help with the sequencing. This work was supported by institutional funds from the Institute National de la Santé et de la Recherche

Médicale and the Centre National de la Recherche Scientifique and by a grant to D.M. and C.B. from the National Institutes of Health (USA) and the Association pour la Recherche sur le Cancer. S.C. received fellowships from the Ministère de la Recherche et de la Technologie and the ARC, and M.B. from the INSERM and the NIH.

References

Aguilar,L.K. and Belmont,J.W. (1991) *J. Immunol.*, **146**, 1348–1352.

Arden,B., Klotz,J.L., Siu,G. and Hood,L.E. (1985) *Nature*, **316**, 783–787.

Bangs,L.A., Sanz,I.E. and Teale,J.M. (1991) *J. Immunol.*, **146**, 1996–2004.

Becker,D.M., Patten,P., Chien,Y.-H., Yokota,T., Eshhar,Z., Giedlin,M., Gascoigne,N.R.J., Goodnow,C., Wolf,R., Arai,K.-I. and Davis,M.M. (1985) *Nature*, **317**, 430–434.

Benoist,C. and Mathis,D. (1989) *Cell*, **58**, 1027–1033.

Berg,L.J., Fazekas de St. Groth,B., Pullen,A.M. and Davis,M.M. (1989) *Nature*, **340**, 559–562.

Bill,J., Apple,V.B. and Palmer,E. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 9184–9189.

Bluestone,J.A., Pardoll,D., Sharrow,S.O. and Fowles,B.J. (1987) *Nature*, **326**, 82–84.

Borgulya,P., Kishi,H., Muller,U., Kirberg,J. and von Boehmer,H. (1991) *EMBO J.*, **10**, 913–918.

Boyd,R.L. and Hugo,P. (1991) *Immunol. Today*, **12**, 71–79.

Candéias,S., Waltzinger,C., Benoist,C. and Mathis,D. (1991a) *J. Exp. Med.*, in press.

Candéias,S., Katz,J., Benoist,C., Mathis,D. and Haskins,K. (1991b) *Proc. Natl Acad. Sci. USA*, **88**, 6167–6170.

Carlsson,L. and Holmberg,D. (1990) *Int. Immunol.*, **2**, 639–643.

Casali,P. and Notkins,A.L. (1989) *Annu. Rev. Immunol.*, **7**, 513–535.

Cazenave,P.-A., Marche,P.N., Jouvin-Marche,E., Voegtlé,D., Bonhomme,F., Bandeira,A. and Coutinho,A. (1990) *Cell*, **63**, 717–728.

Ceredig,R. (1990) *Int. Immunol.*, **2**, 859–867.

Ceredig,R. and Waltzinger,C. (1990) *Int. Immunol.*, **2**, 869–877.

Chothia,C., Boswell,D.R. and Lesk,A.M. (1988) *EMBO J.*, **7**, 3745–3755.

Claverie,J.M., Prochnicka-Chalufour,A. and Bougueleret,L. (1989) *Immunol. Today*, **10**, 10–14.

Coutinho,A., Bandeira,A., Pereira,P., Protnoi,D., Holmberg,D., Martinez,C. and Freitas,A.A. (1989) *Cold Spring Harbor Symp. Quant. Biol.*, **LIV**, 159–170.

Crispe,I.N., Shimonkevitz,R.P., Husmann,L.A., Kimura,J. and Allison,J.P. (1987) *J. Immunol.*, **139**, 3585–3589.

Davis,M.M. and Bjorkman,P.J. (1988) *Nature*, **334**, 395–402.

Elliott,J.F., Rock,E.P., Patten,P.A., Davis,M.M. and Chien,Y.-H. (1988) *Nature*, **331**, 627–631.

Feeney,A.J. (1990) *J. Exp. Med.*, **172**, 1377–1390.

Feeney,A.J. (1991) *J. Exp. Med.*, **174**, 115–124.

Fowlkes,B.J., Kruisbeek,M., Ton-That,H., Weston,M.A., Coligan,J.E., Schwartz,R.H. and Pardoll,D.M. (1987) *Nature*, **329**, 251–254.

Gascoigne,N.R.J., Chien,Y.-H., Becker,D.M., Kavaler,J. and Davis,M.M. (1984) *Nature*, **310**, 387–391.

Gregoire,K.E., Goldschneider,I., Barton,R.W. and Bollum,F.J. (1979) *J. Immunol.*, **123**, 1347–1352.

Gu,H., Förster,I. and Rajewsky,K. (1990) *EMBO J.*, **9**, 2133–2140.

Guidos,C.J., Danska,J.S., Fathman,C.G. and Weissman,I.l. (1990) *J. Exp. Med.*, **172**, 835–845.

Hayakawa,K. and Hardy,R.R. (1988) *Annu. Rev. Immunol.*, **6**, 197–218.

Hengartner,H., Odermatt,B., Schneider,R., Schreyer,M., Walle,G., MacDonald,H.R. and Zinkernagel,R.M. (1988) *Nature*, **336**, 388–390.

Hirama,T., Takeshita,S., Matsubayashi,Y., Iwashiro,M., Masuda,T., Kuribayashi,K., Yoshida,Y. and Yamagishi,H. (1991) *Eur. J. Immunol.*, **21**, 483–488.

Howe,M.L., Goldstein,A.L. and Battisto,J.R. (1970) *Proc. Natl Acad. Sci. USA*, **67**, 613–619.

Hugo,P., Boyd,R.L., Waanders,G.A., Petrie,H.T. and Scollay,R. (1991) *Int. Immunol.*, **3**, 265–272.

Iwasaki,A., Yoshikai,Y., Yuuki,H., Takimoto,H. and Nomoto,K. (1991) *Eur. J. Immunol.*, **21**, 597–603.

Jones,L.A., Chin,L.T., Merriam,G.R., Nelson,L.M. and Kruisbeck,A.M. (1990) *J. Exp. Med.*, **172**, 1277–1285.

Jotereau,F., Heuze,.F, Salomon-Vie,V. and Gascan,H. (1987) *J. Immunol.*, **138**, 1026–1030.

Kabat,E.A., Wu,T.T., Reid-Miller,M., Perry,H.M. and Gottesman,K.S. (1987) *Sequences of Proteins of Immunological Interest*. Department of Health and Human Services, Bethesda, MD, USA, 4th Edition.

Kappler,J.W., Roehm,N. and Marrack,P. (1987a) *Cell*, **49**, 273–280.

Kappler,J.W., Wade,T., White,J., Kushnir,E., Blackman,M., Bill,J., Roehm,N. and Marrack,P. (1987b) *Cell*, **49**, 263–271.

Kappler,J.W., Staerz,U., White,J. and Marrack,P.C. (1988) *Nature*, **332**, 35–40.

Kavaler,J., Davis,M.M. and Chien,Y.-H. (1984) *Nature*, **310**, 421–423.

Kearney,J.F., Vakil,M. and Solvason,N. (1989) *Cold Spring Harbor Symp. Quant. Biol.*, **LIV**, 203–207.

Kisielow,P., Bluthmann,H., Staerz,U.D., Steinmetz,M. and von Boehmer,H. (1988) *Nature*, **333**, 742–746.

Lafaille,J.J., DeCloux,.A, Bonneville,M., Takagaki,Y. and Tonegawa,S. (1989) *Cell*, **59**, 859–870.

Lafaille,J.J., Hass,W., Goutinho,A. and Tonegawa,S. (1990) *Immunol. Today*, **11**, 75–78.

Lattime,E.C. and Stutman,O. (1986) *J. Immunol.*, **136**, 2741–2746.

Lemeur,M., Gerlinger,P., Benoist,C. and Mathis,D. (1985) *Nature*, **316**, 38–42.

McCormack,W.T., Tjoekler,L.W., Carlson,L.M., Stella,G. and Thompson,C.B. (1991) *Proc. Natl Acad. Sci. USA*, in press.

MacDonald,H.R., Schneider,R., Lees,R.K., Howe,R.C., Acha-Orbea,H., Festenstein,H., Zinkernagel,R.M. and Hengartner,H. (1988) *Nature*, **332**, 40–45.

Malissen,M., Minard,K., Mjolsness,S., Kronenberg,M., Goverman,J., Hunkapiller,T., Prystowsky,M.B., Yoshikai,Y., Fitch,F., Mak,T.W. and Hood.L. (1984) *Cell*, **37**, 1101–1110.

Malissen,M., Trucy,J., Letourneur,F., Rebaí,N., Dunn,D.E., Fitch,F.W., Hood,L. and Malissen,B. (1988) *Cell*, **55**, 49–59.

Meek,K. (1990) *Science*, **250**, 820–823.

Nikolic-Zugic,J. (1991) *Immunol. Today*, **12**, 65–70.

Ohashi,P.S., Pircher,H., Burki,K., Zinkernagel,R.M. and Hengartner,H. (1990) *Nature*, **346**, 861–863.

Paterson,D.J. and Williams,A.F. (1987) *J. Exp. Med.*, **166**, 1603–1608.

Pircher,H., Burki,K., Lang,R., Hengartner,H. and Zinkernagel,R.M. (1989) *Nature*, **342**, 559–561.

Pullen,A.M., Marrack,P. and Kappler,J.W. (1988) *Nature*, **335**, 796–801.

Riggs,J.E., Stowers,R.S. and Mosier,D.E. (1991) *J. Exp. Med.*, **173**, 265–268.

Robey,E.A., Fowlkes,B.J., Gordon,J.W., Kioussis,D., von Boehmer,H., Ramsdell,F. and Axel,R. (1991) *Cell*, **64**, 99–107.

Rothenberg,E. and Triglia,D. (1983) *J. Immunol.*, **130**, 1627–1633.

Sakaguchi,S. and Sakaguchi,N. (1990) *J. Exp. Med.*, **172**, 537–545.

Schneider,R., Lees,R.K., Pedrazzini,T., Zinkernagel,R.M., Hengartner,H. and MacDonald,H.R. (1989) *J. Exp. Med.*, **169**, 2149–2158.

Schuler,W., Ruetsch,N.R., Amsler,M. and Bosma,M.J. (1991) *Eur. J. Immunol.*, **21**, 589–596.

Sha,W.C., Nelson,C.A., Newberry,R.D., Kranz,D.M., Russel,J.H. and Loh,D.Y. (1988) *Nature*, **336**, 73–75.

Shortman,K., Vremec,D. and Egerton,M. (1991) *J. Exp. Med.*, **173**, 323–332.

Smith,H., Chen,I.M., Kubo,R. and Tung,K.S.K. (1989) *Science*, **245**, 749–752.

Smith,L.R., Plaza,A., Singer,P.A. and Theofilopoulos,A.N. (1990) *J. Immunol.*, **144**, 3234–3237.

Tomonari,K. (1988) *Immunogenetics*, **28**, 455–458.

Tung,K.S.K., Smith,S., Teuscher,C., Cook,C. and Anderson,R.E. (1987) *Am. J. Pathol.*, **126**, 293–303.

von Boehmer,H. and Byrd,W.J. (1972) *Nature*, **235**, 50–52.

Wade,T., Bill,J., Marrack,P.C., Palmer,E. and Kappler,J.W. (1988) *J. Immunol.*, **141**, 2165–2167.

White,J., Herman,A., Pullen,A.M., Kubo,R., Kappler,J.W. and Marrack,P. (1989) *Cell*, **56**, 27–35.

Wilson,R.K., Lai,E., Concannon,P., Barth,R.K. and Hood,L.E. (1988) *Immunol. Rev.*, **101**, 149–172.

Winoto,A., Mjolsness,S. and Hood,L. (1985) *Nature*, **316**, 832–836.

Winoto,A., Urban,J.L., Lan,N.C., Goverman,J., Hood,L. and Hansburg,D. (1986) *Nature*, **324**, 679–682.

Yague,J., Blackman,M., Born,W., Marrack,P., Kappler,J. and Palmer,E. (1988) *Nucleic Acids Res.*, **16**, 11355–11364.

Received on July 23, 1991