Junctional Sequences of Fetal T Cell Receptor β Chains Have Few N Regions

By Ann J. Feeney

From the Division of Immunology, Medical Biology Institute, La Jolla, California 92037

Summary

T cell receptors (TCRs) and immunoglobulins (Igs) derive a large fraction of their repertoire from diversity generated at the junctions of the V, D, and J coding segments. This diversity is derived both from the random deletion of nucleotides from the ends of coding regions and from the subsequent addition of non-templated N region nucleotides. While the vast majority of TCRs and Igs from adult mice have N regions, <5% of both TCR-γ/δ and Ig from fetal and neonatal mice have N regions. This study analyzed the ontogeny of junctional diversity of TCR-α/β. Genomic DNA or Cβ-primed cDNA was prepared from thymocytes of mice at varying stages in ontogeny, and the rearranged Vββ or Vββ sequences were amplified by polymerase chain reactions. Sequencing of the Vβ-Dβ-Jβ junctions showed few N regions early in ontogeny, although the fraction of sequences with N regions exceeded that previously reported for Ig and for TCR-γ/δ. N regions were found in 13% of Vβ junctional sequences from day 18-19 fetal thymocytes, 33% of sequences from newborn thymocytes, 76% of sequences from day 4 postnatal thymocytes, and 88% of sequences from 5-wk-old thymocytes. In addition, nonrandom usage of the Dβ and Jβ segments was observed in both fetal and adult TCR sequences. While the usage of each of the six Jβ2 segments was different, the same pattern of usage was seen regardless of whether Dβ1 or Dβ2 was used, suggesting that a factor controlling the rate of usage of each J segment is intrinsic to the J gene itself. Since TCRs derive so much of their diversity from N regions, the relative paucity of N regions in fetal α/β T cells would create a fetal TCR-α/β repertoire that would be quite different from, and smaller than, the adult repertoire. The lack of N regions might be predicted to limit the range of affinities of TCR-MHC + peptide interactions, which may have important consequences for positive and negative selection of fetal and newborn T cells.

The three types of lymphocyte antigen receptors (Ig, TCR-α/β, and TCR-γ/δ) are each heterodimers whose variable regions are encoded by somatically recombined gene segments. To obtain their vast repertoire, T cells and B cells use the same mechanisms of combinatorial and junctional diversity (1–3). Combinatorial diversity is derived from the large number of V, D, and J gene elements used to create the variable regions, and from the random association of the two chains to make a complete receptor. Additional diversity is generated at the junctions of these gene segments during somatic recombination by two mechanisms. First, there is deletion of a variable number of nucleotides from the ends of the coding segments. Second, there is addition of a variable number of N region nucleotides to the junctions of the IgH chain and all junctions of the TCR before ligation of the DNA. The N region nucleotides are non-templated and are thought to be added by the enzyme terminal deoxynucleotidyl transferase (TdT) (4–7). Junctional diversity is even more important in generating the total diversity in the TCR repertoire than in the Ig repertoire, since combinatorial association alone leads to a much smaller repertoire for T cells than B cells, and since all TCR junctions can have N regions, while light chains of Ig do not have N regions (3).

T cells are subdivided into two lineages on the basis of which type of TCR they use. T cells with TCR-γ/δ arise first in ontogeny, but are soon outnumbered by α/β T cells, the predominant T cell type in the adult (8). Fetal γ/δ T cells, and some of their adult progeny, do not have N regions in their receptors (9–11), and IgH chains from fetal and newborn B cells also are devoid of N regions (12, 13). Since adult levels of TdT are not reached in the thymus until ~1 wk of age (14, 15), and since α/β T cells are the most predominant T cell type in the mouse thymus by day 18 of gestation (8), this study addressed the question of whether α/β T cells, like γ/δ T cells and B cells, would express few N regions in their receptors early in ontogeny. The results show that this is the case, although the frequency of junctions with N regions is higher than is observed in TCR-γ/δ and Ig for age-matched mice. It has been proposed that the complemen-
tarity determining region 3 (CDR 3)-equivalent regions of the TCR-α/β (which are composed of all the junctions, all of the D region in the β chain, and part of the J region) are mainly responsible for interacting with antigenic peptides (3,16). The paucity of N regions in fetal/newborn α/β T cells will restrict the diversity in CDR3, and will therefore limit the fetal T cell repertoire.

Materials and Methods

Mice. BALB/c mice were bred at Medical Biology Institute. Fetal thymuses were obtained from timed pregnancies, with day 0 being the day that a vaginal plug was observed. The newborn mice were used within 24 h of birth. At least five mice from a litter were pooled for the fetal and newborn samples. Each litter was prepared as a separate sample. Young adult thymus RNA was obtained from a pool of two 5-wk-old mice, and young adult spleen RNA was prepared from a pool of two different 5-wk-old mice.

RNA and cDNA Preparation. Total cellular RNA was extracted from thymuses by the guanidinium thiocyanate/cesium chloride method (17), and CDNA was prepared as previously described (12, 18) with a Cβ primer. DNA was prepared by the freeze-thaw method previously described (12).

PCR Primers. All oligonucleotide primers were made by Genosys (The Woodlands, TX). Initially, a β primer (5'-CGACAAGCTTTGGTATCG-) that ended three nucleotides of Co was used. Later experiments used a Cβ primer (5'-GGGG-AGCAGGGTG). Primers were designed to amplify segments of the V, D, and J regions where the majority of the nucleotides were complementary to a sequence in the third exon of Coβ was used. Later experiments used a β primer (5'-GGGATATCG-), and the internal primer was Coβ primer described above. For Vβ5, the primer listed above was the external primer, and the internal primer was Cβ primer described above. The Jβ2.6 primer was Cβ primer described above.

PCR Amplification and Sequencing. PCR, amplification, cloning, and sequencing was performed as previously described (12, 18) with the following exceptions. Annealing temperature for the PCR was 60–64°C. For genomic DNA, 35 cycles of amplification were performed using the external Vβ8 primer and the Jβ2.6 primer. 15 µl of that primary PCR was directly added to a secondary PCR with the internal Vβ8 primer and the Jβ2.6 primer, and was amplified for 30 cycles. The amplified DNA was then processed as previously described (12).

Results

Analysis of Functional Sequences and Choice of PCR Primers. Analysis of junctional sequences requires knowledge of the complete germline sequence of the V, D, or J gene segment up to the heptamer of the recombination signal sequence (RSS). The germline sequences of the Dβ and Jβ segments are known (20–22), but not all Vβ genes are sequenced. The three-member Vβ8 gene family has been sequenced (19), and in used in ~25% of TCR T cells with the permission of the TCR repertoire.

N region nucleotides were defined as any nucleotides that could be encoded by P regions (11) and therefore is only an estimate of the TCR T cell repertoire.

Figure 1. Sequences of Vβ-Dβ-Jβ junctional regions from thymocytes of 5-wk-old mice. Junctional antibodies that could be encoded by P regions (11) are listed under the Pβ, Pα, or Pj columns. Nucleotides that could be encoded by either of the two adjacent germ-line genes (including P dinucleotides as extensions of the germline gene sequence) are listed in parentheses in the more 5' location. The 3' end of the Vβ7 sequence is the consensus from the sequences that we obtained in this study, and therefore is only an estimate of the germline sequence. N region nucleotides are written in bold type. Sequences that result in the Vβ region being out-of-frame with the Jb region are indicated by an asterisk after the sequence number. D region sequences that cannot be assigned to Dβ1 or Dβ2 are indicated by a double asterisk. The number of nucleotides deleted from the Jβ sequence proper (i.e., not including P region dinucleotides) are listed in the right column. 21 of 24 sequences (88%) had N region nucleotides at least one of the two junctions.
could not have been accounted for by any of the germline gene segments. This is a minimal estimate of N region nucleotides, since it is possible that a nucleotide could be added by RTT that would coincidently be identical to a germline gene segment. This is a minimal estimate of N region nucleotides.

Reduction in N Region Use in cDNA-derived Junctional Sequences from Fetal and Newborn Thymocytes. RNA from thymocytes derived from mice at varying stages of ontogeny was converted into cDNA with a Cβ primer. The cDNA was then amplified by PCR with a Vβ and a Cβ primer, and the amplified DNA was cloned and sequenced. Figure 1 shows the sequences derived from thymocytes of 5-wk-old mice, where 88% of the sequences contained N regions (N region-containing sequences are hereafter referred to as N*). Likewise, 82% of the sequences from splenocytes from 5-wk-old mice were N* (data not shown). In Figure 2, the sequences derived from day 18 fetal thymocytes are displayed, and 16% of those have N regions. 10% of the sequences from day 19 fetal thymocytes are N* (Fig. 3). Figure 4 shows representative cDNA junctional sequences from four separate litters of mice <24 h old. 33% of the 83 total sequences are N*. There was some litter-to-litter variation: the four samples contained an average of 23, 33, 36, and 44% N* sequences. 76% of the sequences derived from thymocytes of 4-d-old mice contained N regions (Fig. 5). All of the results are summarized in Fig. 6. Analysis of only the sequences that had N regions showed a trend towards increasing the number of N nucleotides per N* sequence with time.

Reduction in N Region Use in Genomic DNA-derived Junctional Sequences from Newborn Thymocytes. Amplification of genomic DNA ensures that each cell is equally represented in the analysis, whereas analysis of RNA is subject to the possibility of cell-to-cell variation in the amount of mRNA (25). To address this issue of whether the analysis of RNA was representative of the cellular population in the thymus, we also analyzed sequences derived from genomic DNA. These sequences were amplified with a 3'PCR primer complementary to the frequently used Jβ2.6 segment (see Fig. 9). The decision to sample only that fraction of rearranged genomic sequences using Jβ2.6 was necessitated by the lack of a consensus Jβ sequence, and by the long intron between Jβ and Cβ, which precludes use of a Cβ 3' primer. 12% of the sequences derived from fetal day 18 DNA had N regions (Fig. 7). Two preparations of newborn genomic thymus DNA were made, and, as with the newborn RNA samples, variation between samples was observed. One sample (PCR 161-2, 163-2, and 165-3) contained 13% N regions, and the other (PCR 135-1 and 140-2) contained 29% N regions. Since this time point is the stage at which N regions begin to be seen, some sample-to-sample variation might be expected at this time. Overall, 22% of the newborn sequences contained N regions (Fig. 8).

Genomic DNA contains both productive and nonproductive alleles, and both are amplified. 28% of the sequences derived from genomic DNA were out-of-frame, whereas only 4% of the sequences from RNA were out-of-frame. The percentage of N regions was lower in the out-of-frame than in the in-frame rearrangements. Comparison of only in-frame rearrangements reveals that the percentage of N regions is very similar in both RNA and DNA. Thus, on the level of the expressed allele, there does not appear to be any skewing
of the analysis by amplifying RNA vs. DNA. Both data sets lead to the conclusion that the percentage of N+ sequences in fetal and newborn TCR-α/β is greatly reduced as compared with adult TCR-α/β.

Nonrandom D and J Usage. The number of sequences that use each of the Jβ segments in association with either of the D segments is shown in Fig. 9. Only sequences derived from Cβ-primed cDNA were used, since all of the sequences derived from DNA were generated using a Jβ2.6 primer. Since the genomic organization of the D/J region is Dβ1, Jβ1-6, Dβ2, Jβ2.1-6, the Jβ1 segments can only use Dβ1, whereas the Jβ2 segments can join to either Dβ segment (26). There are approximately equal frequencies of Dβ1-Jβ1, Dβ1-Jβ2, and Dβ2-Jβ2 rearrangements, resulting in the use of Dβ1 twice as often as Dβ2. All, three D region reading frames are used to a similar extent. These data, obtained from >200 unselected junctional sequences, confirm and expand previous analyses of T cell clones and hybridomas (20, 27, 28). There is significant nonrandom usage of Jβ segments. It is striking that the pattern of usage of the six Jβ2 segments is the same regardless of the Dβ segment used, as discussed below.

Discussion
The results presented here show that TCR β chains have a low frequency of N regions early in ontogeny. It has previ-
ously been shown that <5% of TCR-γ/δ from fetal and newborn mice have N regions, and we have recently shown that IgH junctions from newborn liver and spleen DNA are essentially devoid of N regions (10-13). Thus, the paucity of N regions early in ontogeny is a general phenomenon in all lymphocytes. The level of TdT, the enzyme thought to add N region nucleotides, is very low in fetal and newborn thymus, and does not reach adult levels until 1 wk after birth (14, 15). TdT is also not observed in fetal liver, the site of fetal B cell lymphopoiesis (15). Thus, the late appearance of TdT is the most likely explanation for the low frequency of N regions in fetal/neonatal TCR and Ig.

The percentage of N regions observed in TCR β chains from newborn mice is much higher than in TCR-γ/δ and IgH from newborn mice. By postnatal day 4, the percentage of N regions in TCR β chains is almost at the adult level. In contrast, only 29% of IgH junctional sequences from day 9 postnatal mice contain N regions (A.J. Feeney, unpublished results). There are two alternative explanations for the earlier appearance of N region nucleotides in α/β T cells as compared with γ/δ T cells and Ig. The first is that TdT could be differentially regulated in γ/δ T cells, α/β T cells, and B cells. The second explanation is that the intrinsic rate of N region addition could be equally low in all lymphocytes. However, α/β T cells undergo both positive selection for reactivity with self-MHC, as well as negative selection (toler-
ance) in the thymus (29, 30). The diversity afforded to TCRs by N regions is likely to be important in generating the range of receptor affinities able to successfully undergo positive and negative selection, and emerge as mature T cells. Thus, if the lack of N regions restricts the ability of many TCRs to have sufficient affinity to interact with self-MHC (plus peptide) for positive selection, then the selection process could skew the percentage of TCR-α/β with N regions from the rate with which they were initially generated. In support of this latter hypothesis is the fact that there are fewer N regions in the nonproductive (8% N+) than the productive (28% N+) rearrangements in the sequences derived from newborn DNA.

TCRs have significantly less diversity generated by combinatorial usage of germline gene elements than do Igs, since there are fewer germline gene segments. However, this is compensated for by extensive diversity in CDR3-equivalent regions. This difference in the generation of diversity of T and B cell repertoires may have been selected for on the basis of the different ways in which T and B cells interact with antigen. Igs bind to antigens directly, and all IgH CDRs can be involved in antigen binding (31, 32). In contrast, α/β T cells are known to respond to antigenic peptides that are thought to be presented in the groove of class I or class II MHC (33, 34). On the basis of sequence comparisons between TCRs and Igs, it has been proposed that the CDR3 regions of the α and β chains of the TCR bind to the peptide, while the other CDRs bind to MHC (3, 16). This model
In the thymus (24), and since TdT levels are low at this time in ontogeny, it is likely that this phenomenon will hold true for other TCR-α/β as well. It is unlikely that the minor thymocyte subpopulation of CD4-CD8- αβ+ cells (which are greatly enriched in Vβ8.2 expression) has skewed our analysis since these cells arise late in ontogeny, and therefore would be absent from our fetal and newborn thymocyte samples (38-40).

In addition to the absence of N regions in fetal γδ T cells, it has been suggested that there was less "exonucleolytic nibbling" of coding regions in junctional sequences from fetal γδ T cells than in adult γδ T cells (11). The sequences shown here show a slightly increasing number of nucleotides deleted from Vβ8 with age; however, this difference was not significant (p = 0.135). Further data should clarify the significance of this trend.

Other parameters that could affect the size and composition of the neonatal and adult TCR-α/β repertoire were also analyzed. There was wide variation in the frequency of use of the 12 Jβ segments, and this pattern was consistent throughout ontogeny. Strikingly, although the six Jβ2 segments were used with very different frequencies, the pattern of usage was the same regardless of whether Dβ1 or Dβ2 was used, as shown in Fig. 9. These observations strongly suggest that each Jβ segment itself controls the frequency with which it is used, and rejects any hypothesis based on relative proximity of Dβ and Jβ segments. It has been shown that the frequency with which recombination occurs in a transfected recombination substrate is related to the exact sequence of the heptamer or nonamer of the RSS (41). The pattern of Jβ usage observed here is most easily explained by proposing that the RSS (or other closely flanking sequences) 5′ of each

---

### Table: Vβ and Dβ1 Usage

| Vβ | Pd | Dβ1.1 | Pd | Jβ deleted |
|----|----|-------|----|------------|
| Vβ8.1 | GCC | AGC | AGT | GAT | CC | GGGACAGGGGGG | GC |
| 121-1B | --- | --- | --- | --- | CA | CGGG | GC |
| 121-3J | --- | --- | --- | --- | AG | 1.6 | 5 |
| 121-38K | --- | --- | --- | --- | A | GGGCA | T |
| 121-38C | --- | --- | --- | --- | C | AG(A) | 1.7 | 2+3 |
| 121-1H | --- | --- | --- | --- | AC | ACGGGG | 2.1 | 6 |
| Vβ8.2 | GCC | AGC | GGT | GAT | CC | GGGACAGGGGG | GC |
| 121-1C | --- | --- | --- | --- | CA | CCCC | CA |
| 121-3J | --- | --- | --- | --- | AG | 1.6 | 1 |
| 121-3H | --- | --- | --- | --- | A | GGGGAGG | AAAGG |
| 121-1BK | --- | --- | --- | --- | AG | 2.4 | 5 |
| 121-30E | --- | --- | --- | --- | A | GGGGAGG | AG | 2.6 | 0 |
| Vβ8.3 | GCC | AGC | ACT | GAT | CC | GGGACAGGGGG | GC |
| 121-1J | --- | --- | --- | --- | A | GGGCA | CC |
| 121-1K | --- | --- | --- | --- | CA | 2.6 | 0 |
| 121-3H | --- | --- | --- | --- | T | GGGCA(A) | 2.5 | 0-1 |
| 121-30E | --- | --- | --- | --- | A | GGGGAGG | AG | 1.1 | 8-9 |
| 121-1BK | --- | --- | --- | --- | T | GGGCA(G) | 2.1 | 0 |
| 121-30A | --- | --- | --- | --- | AGG | 2.3 | 3 |

---

**Figure 6.** Percent N region usage in sequences derived from thymic RNA from mice of varying ages. The total number of sequences analyzed for each time point is shown on the right.
Figure 7. Sequences of VßDß Jß junctional regions from DNA from day 18 fetal thymocytes. Sequences are displayed as described in the legend to Fig. 1. 3 of 24 sequences (12%) had N region nucleotides in at least one of the two junctions.

Figure 8. Sequences of VßDß Jß junctional regions from DNA from thymocytes of newborn mice <24 h old. Sequences are displayed as described in the legend to Fig. 1. 8 of 37 sequences (22%) had N region nucleotides in at least one of the two junctions.
Jβ segment controls the rate with which each Jβ is used. Significant variation in usage for the four Ig-Jh segments and for the Ig-Dh segments has also been observed (12, 13, 42). Interestingly, the sequence of the heptamer of the RSS of both Jβ2.2 and Ig-Jh1, GACTGTG, is not used by any other IgJh or Jβ segment (22, 43), and these two Js are the least frequently used J segments of IgH and TCRβ. These data are consistent with the hypothesis, proposed for IgVh genes (41), that the frequency with which all V, D, and J segments of all lymphocyte receptors are used may be controlled by their flanking RSS.

T cells and B cells use the same recombinational enzymes to somatically recombine their receptors (44). In addition, both lineages presumably use TdT to add N region nucleotides, and probably use the same exonuclease to delete nucleotides from the coding region. This study shows that N region addition is a developmentally regulated step in the generation of diversity of all lymphocytes. The result of this is that the immunological repertoire of T and B cells early in ontogeny is more restricted than, and differs from, the adult repertoire. The impact of this paucity of N region nucleotides will be more pronounced for αβ T cells than for B cells because T cells depend upon junctional diversity much more than combinatorial diversity to generate a large repertoire, and because N regions are in the CDR3-equivalent region of the TCR that is proposed to recognize antigenic peptides (3, 16).

The excellent technical assistance of Donna Thuerauf is very gratefully appreciated. I would also like to thank Drs. Donald Mosier, Susan Webb, and Jonathan Sprent for critical reviews of the manuscript. In addition, I am grateful to Dr. Daniel Gold for the use of the Cβ primer in the initial experiments, and for continuing helpful discussions.

This work was supported by grant IM-513 from the American Cancer Society and by grants AI-22871 and AI-24526 from the National Institutes for Health.

This is manuscript no. 205 from the Medical Biology Institute.

Address correspondence to Ann J. Feeney, Division of Immunology, Medical Biology Institute, 11077 North Torrey Pines Road, La Jolla, CA 92037.

Received for publication 18 December 1990 and in revised form 6 March 1991.

References

1. Tonegawa, S. 1983. Somatic generation of antibody diversity. Nature (Lond.) 14:575.
2. Siu, G., M. Kronenberg, E. Strauss, R. Haars, T.W. Mak, and L. Hood. 1984. The structure, rearrangement and expression of Dα gene segments of the murine T-cell antigen receptor. Nature (Lond.) 311:344.
3. Davis, M.M., and P.J. Bjorkman. 1988. T-cell antigen receptor genes and T-cell recognition. Nature (Lond.) 334:395.
4. Alt, F.W., and D. Baltimore. 1982. Joining of immunoglobulin heavy chain gene segments: implications from a chromo-
5. Desiderio, S.V., G.D. Yancopoulos, M. Paskind, E. Thomas, M.A. Boss, N. Landau, F.W. Alt, and D. Baltimore. 1984. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxynucleotidyl transferase in B cells. Nature (Lond.) 311:752.
6. Landau, N.R., D.G. Schatz, M. Rosa, and D. Baltimore. 1987. Increased frequency of N-region insertion in a murine pre-B-cell line infected with a terminal deoxynucleotidyl transferase...
retroviral expression vector. Mol. Cell Biol. 7:3237.

7. Lieber, M.R., J.E. Hesse, K. Mizuuchi, and M. Gellert. 1988. Lymphoid V(D)J recombination: nucleotide insertion at signal joints as well as coding joints. Proc. Natl. Acad. Sci. USA. 85:8588.

8. Kubo, R.T., W. Born, J.W. Kappler, P. Marrack, and M. Pigeon. 1989. Characterization of a monoclonal antibody which detects all murine αβ T cell receptors. J. Immunol. 142:2736.

9. Asarnow, J.R., M.M. Davis, and Y. Chien. 1984. Localization of αβ T cell membrane determinants. Proc. Natl. Acad. Sci. USA. 81:2415.

10. Papiernik, M., and C. Pontoux. 1990. In vivo and in vitro recombination of murine TCRαβ chain genes. EMBO J. 9:2233.

11. Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald, and W.J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry. 18:5294.

12. Maguire, J.E., S.A. McCarthy, A. Singer, and D.S. Singer. 1990. Inverse correlation between steady-state RNA and cell surface T cell receptor levels. FASEB J. 4:3131.

13. Maguire, J.E., S.A. McCarthy, A. Singer, and D.S. Singer. 1990. Inverse correlation between steady-state RNA and cell surface T cell receptor levels. FASEB J. 4:3131.

14. Blackman, M.R., J. Kappler, and P. Marrack. 1990. The role of the T cell receptor in positive and negative selection of developing T cells. Science Wash. DC. 248:1335.

15. von Boehmen, H., and P. Kiesielow. 1990. Self-nonself discrimination by T cells. Science Wash. DC. 248:1369.

16. Davies, D.R., and H. Metzger. 1983. Structural basis of antibody function. Annu. Rev. Immunol 1:87.

17. Amit, A.G., R.A. Mariuzza, S.E.V. Phillips, and R.J. Poljak. 1986. Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. Science Wash. DC. 233:747.

18. Engel, L., and S.M. Hedrick. 1987. Structure of the human class I histocompatibility antigen, HLA-A2. Nature (Lond.) 329:506.

19. Sheehan, K.M., and P.H. Brodeur. 1989. Molecular cloning
of the primary IgH repertoire: A quantitative analysis of V\textsubscript{\textalpha} gene usage in adult mice. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:2313.

43. Sakano, H., R. Maki, Y. Kurosawa, W. Roeder, and S. Tonegawa. 1980. Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy chain genes. *Nature (Lond.)* 286:676.

44. Yancopoulos, G.D., T.K. Blackwell, H. Suh, L. Hood, and F.W. Alt. 1986. Introduced T cell receptor variable region gene segments recombine in pre-B cells: evidence that B and T cells use a common recombinase. *Cell.* 44:251.