Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 3551–3554, May 1988
Immunology

A conserved sequence in the T-cell receptor β-chain promoter region

(transcription initiation/RNase protection/decanucleotide)

STEVEN J. ANDERSON, HUBERT S. CHOU, AND DENNIS Y. LOH

Howard Hughes Medical Institute, Departments of Medicine, Microbiology, and Immunology, Division of Allergy and Immunology, Washington University Medical School, St. Louis, MO 63110

Communicated by David M. Kipnis, January 11, 1988

ABSTRACT **The antigen-specific receptors of T and B lymphocytes are distinct, though structurally related, molecules. During development, lymphoid cells assemble functional variable (V) region genes for each receptor chain from separate multimember gene families by somatic DNA rearrangements of individual germ-line segments. Transcription may play a role in regulating the tissue and stage specificity of these rearrangements by controlling the accessibility of germ-line loci to the recombinational machinery. Immunoglobulin V-region genes are transcribed from tissue-specific promoters that have been well characterized. We report here the characterization of 14 T-cell receptor β-chain V-region gene promoters. Sequence analysis indicates that these promoters do not contain the conserved octamer that is located upstream of all immunoglobulin genes. However, a unique decanucleotide sequence, not present in immunoglobulin genes, is conserved in the promoter region of murine and human V_β genes. We identify this sequence as a potential regulatory element, based on its position, conservation, and sequence homology to sites known to bind transcription-activating factors. The possibility that the distinct structures of immunoglobulin and T-cell receptor gene promoters may contribute to the tissue-specific rearrangement and expression of receptor gene families is discussed.**

The tissue- and stage-specific expression of appropriate antigen receptors is critical to lymphocyte function. Mature T cells, which recognize antigen in the context of major histocompatibility complex products, express a cell-surface receptor comprised of disulfide-linked α and β chains. In contrast, B-cell receptors are immunoglobulin molecules consisting of heavy (H) and either κ or λ light (L) chains. Each of these chains contains an amino-terminal variable (V) region and a carboxyl-terminal constant (C) region. During lymphocyte differentiation, germ-line V, diversity (D), and joining (J) gene segments are somatically recombined to form the V-region genes of each receptor chain (1, 2). Assembly of functional immunoglobulin V-region genes is limited to precursor B cells, whereas functional T-cell receptor (TCR) V-region genes are assembled only in developing T cells. Studies on the recombination of transfected substrates, however, suggest that a common "recombinase" may mediate the assembly of all immunoglobulin and TCR V-region gene segments (3). One model to explain the specificity of rearrangement proposes that the accessibility of various gene segments to the common recombinase is modulated differently in each cell type (4). This accessibility is reflected in the nuclease sensitivity and often the transcriptional activity of the target segment. Although the precise relationship between transcription and accessibility is unclear, recent evidence suggests that transcription may be a prerequisite of recombinational activation (5). Thus, specific DNA sites and trans-acting factors that activate tissue-specific transcription may control locus accessibility for recombination. A knowledge of the differences between T- and B-cell V-region gene promoters may shed light on the mechanisms by which developmental signals determine receptor expression.

To investigate tissue-specific transcription and rearrangement in T cells, we have examined the promoters of 14 murine TCR β-chain V-region genes, representing 11 V_β subfamilies. Transcription initiation sites were mapped by using a sensitive RNase protection assay, and the nucleotide sequences of the 5′ flanking regions of these genes were determined. These analyses reveal a unique decanucleotide sequence located a short distance upstream of the transcription start site in most V_β genes. This sequence is similar to the cAMP-responsive element (6) and the AP-1 binding site (7). The position and conservation of this element suggest a potential role in V_β promoter function.

MATERIALS AND METHODS

Construction of Genomic Clones. Isolation of genomic clones containing $V_\beta 11$, $V_\beta 12$, $V_\beta 8.2$, and $V_\beta 8.3$ has been described (8) [V_β nomenclature according to Barth *et al.* (9) as described in Behlke *et al.* (10)]. Additional clones containing $V_\beta 1$, $V_\beta 4$, $V_\beta 5.1$, $V_\beta 5.2$, $V_\beta 6$, $V_\beta 7$, $V_\beta 8.1$, $V_\beta 9$, $V_\beta 10$, and $V_\beta 16$ were obtained from a C57BL/6 genomic library (*Mbo* I partial digest in EMBL3) by screening with previously described V-region-specific probes (11). Appropriate DNA segments were subcloned into plasmid vectors pBS (Bluescribe, Stratagene, San Diego, CA) or pUC19, and relevant nucleotide sequences were determined by the method of Maxam and Gilbert (12).

Determination of Transcription Initiation Sites. Transcription start sites were determined by an RNase protection assay. DNA segments spanning the protein-coding and 5′ untranslated portions of the first exon from each gene were subcloned into the pBS expression vector, from which templates were generated by linearizing the plasmid at unique upstream sites. Antisense RNA was synthesized *in vitro* according to the manufacturer's protocols. Total cellular RNA was obtained from C57BL/6 thymus or T-cell hybridomas by the guanidinium isothiocyanate method (13). RNase protection was carried out essentially according to the method of Melton *et al.* (14). Antisense RNA probe (≈1 × 10^5 cpm) was added to 50 μg of test RNA and hybridized at 45°C for >12 hr. Samples were digested for 60 min at 30°C with 12 μg of RNase A and 0.6 μg of RNase T1 and separated on 7 M urea/polyacrylamide gels, and protected products were visualized by autoradiography. Primer extension assays were performed as described by Leonard *et al.*

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: TCR, T-cell antigen receptor; V, variable; C, constant; D, diversity; J, joining; H, heavy; L, light.

(15) by using a synthetic oligonucleotide primer, 5′ CTGTTTCCACAAGACATAGG 3′, complementary to 20 nucleotides at the 3′ end of the first exon of $V_\beta 1$.

RESULTS

Transcription Initiation Sites. We have determined the transcription start sites for 13 murine V_β genes by using RNase protection of probes designed to span the first exon of each gene. Autoradiographs and diagrams of respective probes from four experiments are shown in Fig. 1. Although the lengths of the first exons are variable between genes, two sizes of approximately 70 or 120 nucleotides predominate. In 10 of the genes analyzed, the most abundant transcripts are initiated within 100 base pairs (bp) of the translation start codon. Only two genes, $V_\beta 4$ and $V_\beta 6$, initiate most transcripts significantly farther upstream, at sites approximately 750 and 260 bp, respectively, 5′ to the protein-initiating ATG. These unusually long first exons correlate with the lengths of previously isolated cDNA clones of these genes, the longest of which contain 686 and 214 bp of 5′ untranslated sequence for $V_\beta 4$ and $V_\beta 6$, respectively (data not shown). The initiation sites of the prominent transcripts of the genes analyzed are summarized in Fig. 2.

Additional Transcriptional Patterns. The RNase protection data indicate that some V_β genes, such as $V_\beta 1$ and $V_\beta 12$, utilize multiple start sites (Fig. 1). Apart from artifacts generated by RNase digestion, the presence of multiple transcript sizes was confirmed for the $V_\beta 1$ gene by a primer extension assay (Fig. 1). Similarly, in addition to the transcripts initiating proximally, we identified transcripts arising from distant upstream sites for $V_\beta 1$, $V_\beta 5.1$, and $V_\beta 11$ (Fig. 1 and data not shown). Some transcripts containing $V_\beta 16$ extend at least 350 bp 5′ to the protein-initiating methionine codon and include the recombination recognition sequences (heptamer and nonamer) flanking the 3′ end of the nearest upstream gene, $V_\beta 4$ (Fig. 1). These transcripts may initiate from the normal promoter of the upstream $V_\beta 4$ gene. We have previously reported a related pattern of transcription from the tandemly arrayed $V_\beta 5.1$ and $V_\beta 8.2$ genes, as determined by cDNA sequences (17). RNase protection assays using thymus RNA confirm that >95% of transcripts encoding $V_\beta 8.2$ are initiated from the $V_\beta 5.1$ promoter and contain a correctly spliced leader exon derived from the upstream gene (data not shown).

Nucleotide Sequence Analysis. In general, analysis of V_β transcription initiation sites indicates that the transcriptional promoter is located a short distance 5′ to the protein-coding region of each gene. To examine the structure of these promoters, the nucleotide sequences of the 5′ flanking regions of 14 V_β genes, representing 11 V_β subfamilies, were determined. Approximately 150 nucleotides of sequence from each of these genes are presented in Fig. 2. Potential "TATA" boxes, consisting of (A+T)-rich sequences positioned 20–35 bases 5′ to the major transcription start sites, can be identified in each gene (Fig. 2). However, the poorly conserved, noncanonical sequences of the putative TATA box regions, in conjunction with multiple start sites in some cases, suggest that V_β gene promoters may not contain functional TATA boxes. No consensus CCAAT elements or Sp1 binding sites could be identified within this region. Moreover, the octamer sequence conserved in all immunoglobulin genes is not found in the TCR V-region promoters.

Since the protein-coding portions of murine V_β genes share only limited homology (9, 11), it is not surprising to find that the promoter regions of genes from different subfamilies show marked sequence diversity. However, a distinct motif consisting of a decanucleotide with consensus sequence AGTGA$^{TG}_{CA}$TCA is found in 13 of the 14 V_β genes examined (Fig. 2). This sequence is, in part, an inverted repeat, with similarity to the cAMP-responsive element (6) and the AP-1 binding site (7). In 12 of these 13 genes, this sequence lies 5′ to the putative TATA box, ranging from 40 to 75 bp upstream of the major transcription initiation site. No additional regions of extensive homology are shared between members of different V_β subfamilies.

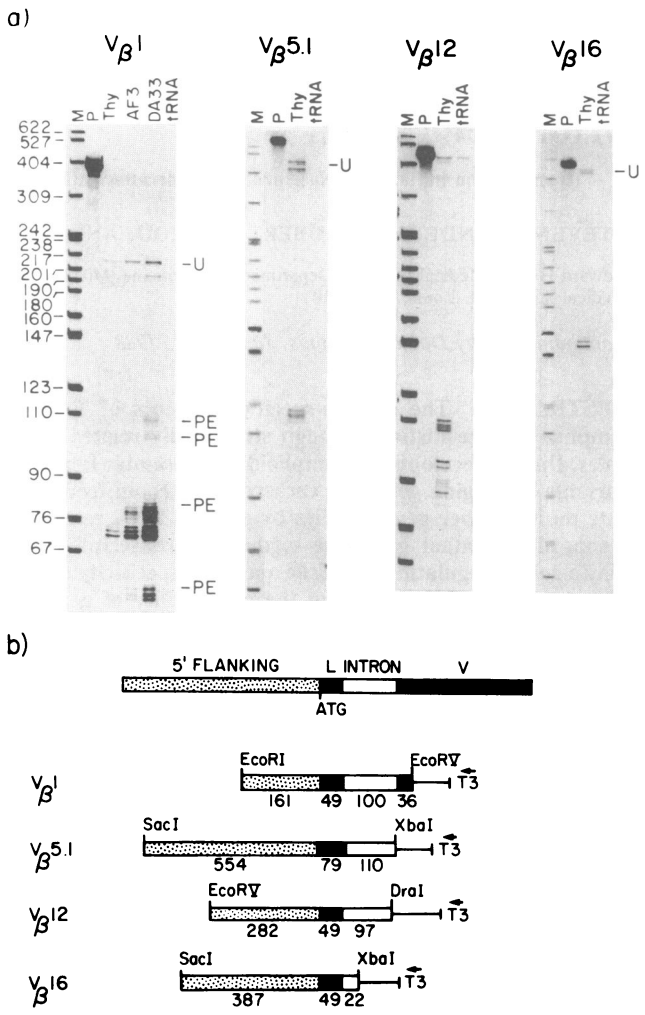

FIG. 1. Mapping of transcription initiation sites for four V_β genes by using RNase protection. (*a*) Autoradiographs of gels from RNase protection assays using probes for $V_\beta 1$, $V_\beta 5.1$, $V_\beta 12$, and $V_\beta 16$. The source of test RNA (C57BL/6 thymus, AF3 or DA33 T-cell hybridomas, or yeast tRNA) is as indicated. AF3 and DA33 are beef-insulin-specific T hybridomas that express one and two copies, respectively, of $V_\beta 1$ (16). Protected bands correspond to the first exon of the gene, since the portions of the probe upstream of the cap site and downstream of the exon–intron boundary are unprotected. The longest bands for $V_\beta 1$ and $V_\beta 16$ represent protection of the full-length probe except for intron and plasmid sequences. M, radiolabeled *Msp* I digest of pBR322, sizes in nucleotides; P, input probe; U, protected bands representing transcripts initiating at upstream sites; PE, protected bands that correspond to bands also visualized in primer extension assay (data not shown). (*b*) Structure of DNA templates for antisense RNA probes. The top diagram is a schematic representation of a typical germ-line V_β gene. The ATG indicates the initiator methionine codon. Restriction enzyme sites shown for each probe indicate the downstream site used for cloning into the Bluescribe polylinker and the upstream site used for linearizing template. Numbers below each diagram indicate the sizes in nucleotides of 5′ flanking, leader (L), intron, and V portions of the gene included in each probe. Vector sequences included in each probe are indicated by a straight line (—). Direction of *in vitro* RNA synthesis from the T3 promoter is indicated by the arrow.

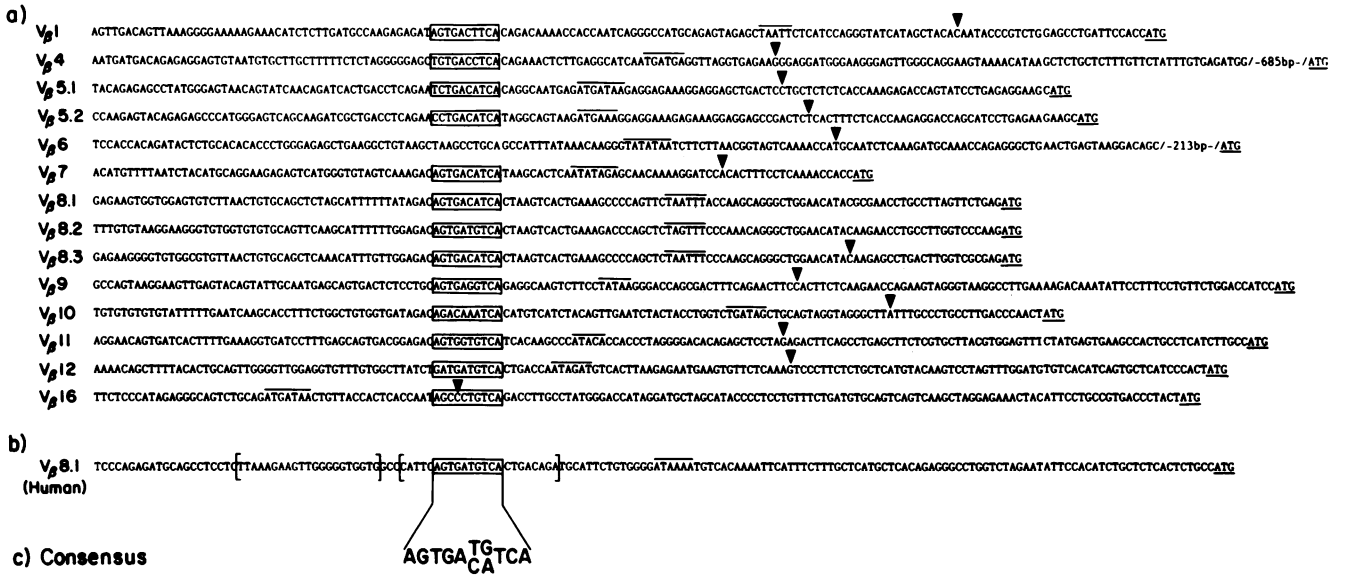

FIG. 2. Nucleotide sequences of promoter regions of TCR β-chain V-region genes. (*a*) Nucleotide sequences of the 5′ flanking portions of 14 murine V_β genes are presented. The methionine codon (ATG) for translation initiation is underlined. The initiation sites of the most prominent transcripts, as determined by RNase protection assays, are indicated by triangles. Putative TATA boxes are overlined. The conserved decanucleotide sequence (see text) is boxed. (*b*) Nucleotide sequence of the 5′ flanking portion of the human V_β *8.1* gene as reported by Siu *et al.* (18). The brackets denote coding-strand regions I and II reported by Royer and Reinherz (19) to be protected by nuclear extracts in DNase footprint analysis. (*c*) Consensus sequence of the conserved decanucleotide determined from murine and human genes.

DISCUSSION

Potential Promoter Function. The data presented here reveal the distinct nature of the 5′ flanking regions of murine V_β genes. Members of different V_β subfamilies display marked heterogeneity in nucleotide sequence and position of transcription start sites. The location of initiation sites within a short distance of the protein-coding region of each gene, however, indicates that a transcriptional promoter is located within this region. These promoters, although diverse in sequence, must be functionally equivalent in driving the transcription of rearranged V-region genes. Immunoglobulin V-region promoters show similar sequence diversity but invariably contain a conserved octanucleotide that determines the tissue specificity of immunoglobulin transcription (20–23). This octamer is clearly not contained within the V_β promoter region. We have, however, identified a unique element shared by all but one of the murine V_β genes examined. The consensus sequence of this decanucleotide is AGTGA$^{TG}_{CA}$TCA. This limited region of homology contained within otherwise diverse 5′ sequences suggests that the decanucleotide is a conserved element in murine V_β promoters and might therefore be found in homologous genes of other species. A search of the published sequences of the human T-cell receptor V_β *8* subfamily, whose members are most closely related to the murine V_β *11* gene, reveals conservation of this decanucleotide sequence in all subfamily members (18). In the human V_β *8.1* gene this sequence is located 115 bp 5′ to the initiating ATG, placing it just upstream of the putative TATA box and transcription start site (Fig. 2). Further sequence comparison indicates no additional regions of significant homology between the human and murine V_β promoters, supporting a prediction that the conserved decanucleotide sequence might be involved in regulating transcription of these genes.

As is common with many regulatory protein-binding sites (24), the decanucleotide displays a degree of dyad symmetry. In fact, the most highly conserved portion of the decanucleotide is an inverted repeat with the sequence TGA--TCA. Interestingly, this decanucleotide also shows limited homology to other putative regulatory elements, including the palindrome TGACGTCA, contained within a small region required for cAMP responsiveness (6, 25), and the sequence TGACTCA, identified as the binding site of the transcription-activating protein AP-1 (7, 26). Although the cis- and trans-acting elements required for V_β transcription have not been identified, evidence of possible regulatory function for the decanucleotide sequence was derived from DNase footprint analysis of the human V_β *8.1* promoter area (19). These studies showed that the conserved sequence lies within a 23-bp region protected by nuclear extracts of lymphoid and nonlymphoid cells, indicating protein binding to this region. A lack of absolute conservation of this sequence in all V_β genes suggests flexibility in the protein recognition site. Although additional areas within the human V_β *8.1* promoter were protected in this analysis, these regions do not share significant homology with the murine genes (Fig. 2). The potential to be bound by a trans-acting factor further supports a functional role for the decanucleotide in V_β regulation.

V_β Gene Regulation. The role of promoter elements in directing the tissue specificity of TCR expression is unknown. Extensive studies on immunoglobulin promoters, however, have shown that maximal expression of these genes requires the conserved octamer sequence (21–23) and a tissue-specific nuclear factor, perhaps a posttranslationally modified form of a ubiquitous octamer-binding protein (27–31). The distinct structures of immunoglobulin and TCR V-region promoters suggest that transcription of these genes can be regulated by different mechanisms. One intriguing outcome of developmentally regulated transcription of V-region genes may lie in the control of rearrangement events. The absence of inappropriate V to D–J rearrangements despite the occurrence of inappropriate D–J joins [$D_\beta J_\beta$ in B cells and $D_H J_H$ in T cells] (32) suggests that V-segment rearrangement is tightly regulated in lymphocytes. This strict tissue specificity of V-segment rearrangement prevents the expression of inappropriate protein products that might interfere with the normal regulation of subsequent recombination events. As evidenced by recent transient recombination assays (33), an absolute requirement for transcription as a prerequisite to recombination has not been demonstrated.

A correlation between these two activities, however, is well established (4). For example, some germ-line V_H segments, which are transcriptionally silent in T cells and mature B cells, are actively transcribed from their normal promoters in precursor cells undergoing V_H to D_H–J_H rearrangement (34). It is possible that one function of distinct factors that drive developing lymphocytes into either a T-cell or B-cell lineage is the differential activation of TCR or immunoglobulin germ-line gene segments, thereby generating unique accessibility to a common recombinase. Tissue-specific trans-acting factors generated in response to such extracellular differentiation signals could result in a battery of specific gene expression events. The similarity of the V_β decanucleotide to elements that contribute to cAMP responsiveness (25) and phorbol ester inducibility (7, 35) suggests that a protein kinase may be involved in the generation of tissue-specific factors that control V_β transcription. The ability of phorbol esters to augment β-chain expression in EL4 cells (36) supports this possibility. Identification of the factors that bind the V_β promoters and control transcription may shed light on the signals for lymphocyte differentiation.

Unrearranged V_β Transcription. The promoter and/or enhancer sequences required for V_β transcription have not yet been determined. Immunoglobulin V-region gene promoters are generally transcriptionally active only when juxtaposed with enhancer elements found within the J–C intron of H and L chains (37–39). Similar enhancer elements have not yet been identified within the β-chain locus, and the activity of V_β promoters independent of such elements has not been tested. Interestingly, two separate immunoglobulin gene promoters in a tandem array can be transcriptionally active in an enhancer-dependent fashion when transfected into B cells (40). A similar tandem array occurs naturally for the murine $V_\beta 5.1$ and $V_\beta 8.2$ genes. Transcription from the upstream $V_\beta 5.1$ gene following rearrangement of the $V_\beta 8.2$ gene indicates that activation is not limited to the promoter most proximal to the C region. Such activity from nearby upstream promoters after V-region rearrangement may be a general feature of the closely linked V_β genes, since similar transcription appears to occur with $V_\beta 4$ and $V_\beta 16$. Nuclear run-on assays, which would determine the extent to which upstream V_β genes are transcriptionally active following rearrangement, may shed light on the contribution of promoter and nonpromoter sequences to the regulated expression of V_β genes.

We thank Drs. David Chaplin, Stanley Korsmeyer, Marjorie Louie, and Timothy Ley and Mr. Mark Behlke for critical comments on the manuscript and Ms. Deanna London for typing of the manuscript. This research was supported by funds from the Howard Hughes Medical Institute (D.Y.L.) and by National Institutes of Health Research Service Awards AI07163 and GM07067 (to Washington University, S.J.A. and H.S.C.).

1. Kronenberg, M., Siu, G., Hood, L. & Shastri, N. (1986) *Annu. Rev. Immunol.* **4,** 529–591.
2. Tonegawa, S. (1983) *Nature (London)* **302,** 575–581.
3. Yancopoulos, G. D., Blackwell, T. K., Suh, H., Hood, L. & Alt, F. W. (1986) *Cell* **44,** 251–259.
4. Alt, F. W., Blackwell, T. K., De Pinho, R. A., Reth, M. G. & Yancopoulos, G. D. (1986) *Immunol. Rev.* **89,** 5–30.
5. Blackwell, T. K., Moore, M. W., Yancopoulos, G. D., Suh, H., Lutzker, S., Selsing, E. & Alt, F. W. (1986) *Nature (London)* **324,** 585–589.
6. Montminy, M. R. & Bilezikjian, L. M. (1987) *Nature (London)* **328,** 175–178.
7. Lee, W., Mitchell, P. & Tjian, R. (1987) *Cell* **49,** 741–752.
8. Chou, H. S., Behlke, M. A., Godambe, S. A., Russell, J. H., Brooks, C. G. & Loh, D. Y. (1986) *EMBO J.* **5,** 2149–2155.
9. Barth, R. K., Kim, B. S., Lan, N. C., Hunkapiller, T., Lobieck, N., Winoto, A., Gershenfeld, H., Okada, C., Hansburg, D., Weissman, I. L. & Hood, L. (1985) *Nature (London)* **316,** 517–523.
10. Behlke, M. A., Chou, H. S., Huppi, K. & Loh, D. Y. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 767–771.
11. Behlke, M. A., Spinella, D. G., Chou, H. S., Sha, W., Hartl, D. L. & Loh, D. Y. (1985) *Science* **229,** 566–570.
12. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
13. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
14. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984) *Nucleic Acids Res.* **12,** 7035–7056.
15. Leonard, W. J., Depper, J. M., Kanehisa, M., Kronke, M., Peffer, N. J., Svetlik, P. B., Sullivan, M. & Greene, W. C. (1985) *Science* **230,** 633–639.
16. Spinella, D. G., Hansen, T. H., Walsh, W. D., Behlke, M. A., Tillinghast, J. P., Chou, H. S., Whiteley, P. J., Kapp, J. A., Pierce, C. W., Shevach, E. M. & Loh, D. Y. (1987) *J. Immunol.* **138,** 3991–3995.
17. Chou, H. S., Anderson, S. J., Louie, M. C., Godambe, S. A., Pozzi, M. R., Behlke, M. A., Huppi, K. & Loh, D. Y. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 1992–1996.
18. Siu, G., Strauss, E. C., Lai, E. & Hood, L. (1986) *J. Exp. Med.* **164,** 1600–1614.
19. Royer, H. D. & Reinherz, E. L. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 232–236.
20. Parslow, T. G., Blair, D. L., Murphy, W. J. & Granner, D. K. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2650–2654.
21. Falkner, F. G. & Zachau, H. G. (1984) *Nature (London)* **310,** 71–74.
22. Bergman, Y., Rice, D., Grosschedl, R. & Baltimore, D. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 7041–7045.
23. Mason, J. O., Williams, G. T. & Neuberger, M. S. (1985) *Cell* **41,** 479–487.
24. Sen, R. & Baltimore, D. (1986) *Cell* **46,** 705–716.
25. Montminy, M. R., Sevarino, K. A., Wagner, J. A., Mandel, G. & Goodman, R. H. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 6682–6686.
26. Lee, W., Haslinger, A., Karin, M. & Tjian, R. (1987) *Nature (London)* **325,** 368–372.
27. Sive, H. L. & Roeder, R. G. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 6382–6386.
28. Singh, H., Sen, R., Baltimore, D. & Sharp, P. A. (1986) *Nature (London)* **319,** 154–158.
29. Staudt, L. M., Singh, H., Sen, R., Wirth, T., Sharp, P. A. & Baltimore, D. (1986) *Nature (London)* **323,** 640–643.
30. Mocikat, R., Falkner, F. G., Mertz, R. & Zachau, H. G. (1985) *Nucleic Acids Res.* **14,** 8829–8844.
31. Mizushima-Sugano, J. & Roeder, R. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 8511–8515.
32. Alessandrini, A., Pierce, J. H., Baltimore, D. & Desiderio, S. V. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 1799–1803.
33. Hesse, J. E., Lieber, M. R., Gellert, M. & Mizuuchi, K. (1987) *Cell* **49,** 775–783.
34. Yancopoulos, G. D. & Alt, F. W. (1985) *Cell* **40,** 271–281.
35. Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R. J., Rahmsdorf, H. J., Jonat, C., Herrlich, P. & Karin, M. (1987) *Cell* **49,** 729–739.
36. Noonan, D. J., Isakov, N., Theofilopoulos, A. N., Dixon, F. J. & Altman, A. (1987) *Eur. J. Immunol.* **17,** 803–807.
37. Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. (1983) *Cell* **33,** 717–728.
38. Banerji, J., Olson, L. & Schaffner, W. (1983) *Cell* **33,** 729–740.
39. Queen, C. & Baltimore, D. (1983) *Cell* **33,** 741–748.
40. Atchison, M. L. & Perry, R. P. (1986) *Cell* **46,** 253–262.