Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 232–236, January 1987
Immunology

Multiple nuclear proteins bind upstream sequences in the promoter region of a T-cell receptor β-chain variable-region gene: Evidence for tissue specificity

(transcriptional regulation/enhancer/DNase I protection assay)

HANS DIETER ROYER AND ELLIS L. REINHERZ

Laboratory of Immunobiology, Dana–Farber Cancer Institute, and Departments of Pathology and Medicine, Harvard Medical School, Boston, MA 02115

Communicated by Francis D. Moore, September 12, 1986 (received for review August 4, 1986)

ABSTRACT **DNA–nuclear protein binding interactions were examined in the promoter region of a representative T-cell receptor Ti β-chain variable-region gene by means of electrophoretic mobility-shift and DNase I-protection analysis. Within 175 bases upstream of the transcription initiation site, four protected regions ("footprints") were identified on the coding strand, at nucleotides −46 to −68 (I), −72 to −92 (II), −113 to −134 (III), and −136 to −175 (IV). Nuclear proteins (0.6 M NaCl fraction from a heparin-Sepharose column chromatography of nuclear extracts) of a variety of cell types produced footprints I, III, and IV and a fifth footprint (beyond nucleotide −200). In contrast, footprint II was produced only by T-cell extracts, although nuclear extracts of a transformed B-lymphoblastoid line produced a partial footprint in this region. Furthermore, footprint analysis of the noncoding strand showed that a continuous region of protection corresponding to the entire region of footprints I and II was generated, along with a DNase I-hypersensitive site, by nuclear proteins derived from T cells but not other cell types. Footprint I has the sequence structure of many well-defined protein–DNA binding sites. Nucleotide sequences in the region of footprint II bore no homology to known sequences, whereas those in the areas of footprints III and IV were similar to motifs within immunoglobulin and other enhancers. These findings may have implications for the tissue specificity of human Ti β-chain gene transcription.**

The human T-cell receptor for antigen and major histocompatibility complex comprises five polypeptide chains, collectively termed the T3–Ti complex. Whereas the disulfide-linked 49-kDa (α) and 43-kDa (β) Ti subunits are polymorphic and confer ligand specificity through their individual immunoglobulin-like variable (V) domains, the three 20- to 25-kDa T3 subunits (γ, δ, and ε) are themselves monomorphic, noncovalently associated subunits presumably involved in signal transduction (1). During development, bone marrow precursor cells, which lack the T3–Ti structure, are induced to differentiate within the thymic microenvironment into lymphocytes bearing the T-cell receptor (2). This differentiative process is rigorously ordered: Ti β gene transcription is activated first and only subsequently Ti α gene transcription (3).

While it is obvious that rearrangements of individual V, diversity (D), and joining (J) segments of Ti β and α genes occur prior to transcriptional activation of the complete α and β V (V_α and V_β) genes, the molecular basis of subsequent α and β gene expression is not understood (3, 4). Because promoter recognition by RNA polymerase is a key step in regulation of transcription and requires auxiliary proteins to direct selective and accurate transcription (5), we have begun to study protein–DNA interactions between nuclear proteins and the DNA sequences of a representative V_β gene promoter region.

MATERIALS AND METHODS

Genomic DNA Cloning. A human T-cell genomic library from the REX tumor line was constructed essentially as described by Maniatis *et al.* (6), using λ Charon 30 *Bam*HI arms and size-selected [20-kilobase (kb)-long] fragments from a partial *Mbo* I digest of T-cell (REX) DNA. The library was screened with a REX Ti β constant-region (C_β) ^{32}P-labeled probe isolated and described previously (3). Clone λβREX contained a rearranged $C_\beta 1$ gene and was used for subcloning the REX Ti V_β gene.

λβREX was digested with *Bgl* II and *Xma* I, and the fragments were ligated to *Xma* I-cleaved pUC8 vector. Transformants were screened with a REX C_β DNA probe. The clone $pV_g\beta$REX contained the 1.5-kb *Bgl* II–*Xma* I/*Sma* I fragment plus 2-kb *Bgl* II–*Xma* I fragment located upstream in λβREX. For DNA sequence analysis, electrophoretic gel mobility shift, and DNase I protection analysis, DNA fragments were prepared from $pV_g\beta$REX by cutting with the appropriate restriction endonucleases, followed by preparative gel electrophoresis in polyacrylamide gels and electroelution of individual fragments.

Isolation of Nuclear Proteins and Heparin-Sepharose Chromatography. Nuclear proteins derived from various cell lines were isolated as described by Wu (7). Cells were homogenized in solution I, and nuclei were recovered by centrifugation. For extraction of nuclear protein, nuclei were suspended in solution II (contains 400 mM NaCl) and incubated at 4°C for 30 min with continuous stirring to prevent aggregation. Nuclear extracts were clarified by ultracentrifugation at 39,000 rpm for 60 min at 4°C in a Sorvall TFT65.13 rotor. Nuclear extracts were dialyzed for at least 5 hr against solution III. To enrich for DNA binding proteins, nuclear extracts were chromatographed on heparin-Sepharose (Pharmacia). A protocol outlined by Wu was used (8). Extracts were loaded twice in chromatography buffer (CB)/0.1 M NaCl, washed with CB/0.1 M NaCl, and fractionated by elution with CB/0.3 M NaCl, CB/0.6 M NaCl, and CB/1 M NaCl. For further use, 0.3 M and 0.6 M NaCl fractions were dialyzed for at least 5 hr against solution III, using collodion bags (Schleicher & Schuell).

Electrophoretic Gel Mobility-Shift and DNase I Protection Analysis. Protocols described by Singh *et al.* (9) were followed. For gel mobility-shift experiments, DNA fragments were labeled at the 3′ end with the Klenow fragment of DNA polymerase (New England Biolabs) or the 5′ end with polynucleotide kinase (New England Biolabs). Individual

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: V, variable; D, diversity; J, joining; C, constant; kb, kilobase(s); bp, base pair(s).

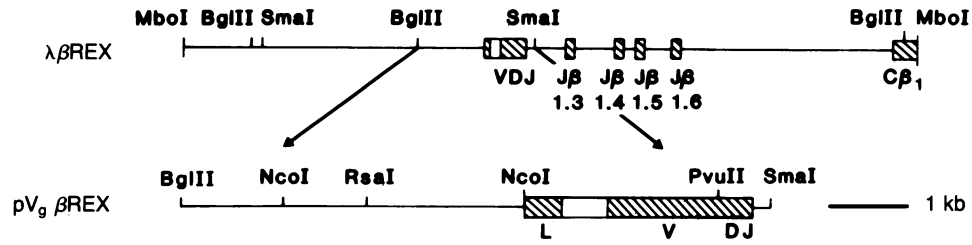

FIG. 1. Restriction maps of clones and fragments. λβREX is the clone containing the rearranged V_β gene in λ Charon 30. $pV_g\beta REX$ is a subcloned piece of λβREX, containing the rearranged V_β gene. V600 (*Nco* I–*Nco* I) and V400 (*Rsa* I–*Nco* I) are fragments, derived from $pV_g\beta REX$, that contain upstream sequences. L, leader.

fragments were incubated in binding buffer, as described, for 25 min at room temperature and loaded on a low-salt polyacrylamide gel (4% total monomer, 30:1 acrylamide/*N*, *N*'-methylenebisacrylamide ratio). Electrophoresis was at 11 V/cm for 90 min at room temperature with buffer recirculation. Gels were dried and exposed to x-ray film overnight.

For DNase I protection analysis, DNA fragments labeled at one end were incubated with 50 μg of nuclear extract protein in binding buffer (9) in the presence of 2.5 mM $MgCl_2$ and 5 μg of poly(dI-dC)·poly(dI-dC) for 25 min at room temperature, digested with DNase I (0.1–0.5 μg; New England Biolabs) for 2.5 min at room temperature, phenol-extracted, chloroform-extracted, and ethanol-precipitated twice. Samples were heated to 90°C in loading buffer and applied to an 8% sequencing gel as described by Maniatis *et al.* (10).

Sequence Determination and Comparison. DNA sequence analysis was done by the chemical degradation method (11). Release 42 of GenBank* was scanned for sequence similarities, using a local homology program (12). Local sequence similarities and alignments were investigated by use of a standard dynamic programming algorithm (13).

RESULTS

Cloning of the Rearranged $V_{\beta REX}$ Gene. To obtain an active Ti V_β gene and its 5' flanking region, we constructed a REX genomic library in λ Charon 30 bacteriophage and screened the library with a C_β cDNA probe. The REX human line is of T lineage and expresses the T3-associated Ti α/β heterodimer (2). A genomic clone termed λβREX was isolated and shown by restriction mapping to contain the rearranged V_β gene. The organization of the λβREX insert is outlined in Fig. 1. The *Bgl* II–*Sma* I fragment was subcloned in pUC8 to generate the clone $pV_g\beta REX$. A restriction map of the insert containing the V_β gene is shown in Fig. 1. Regions of this insert, termed V600 and V400, were utilized to study putative DNA–protein interactions by electrophoretic mobility shift (9, 14, 15) and DNase protection assay (9, 16).

Analysis of DNA Binding Activity of Nuclear Extracts by Gel Mobility Shift. As putative regulatory sequences of the Ti β gene could reside 5' of the initiation codon (ATG), we isolated a 600-base-pair (bp) *Nco* I fragment (V600) from $pV_g\beta REX$ whose 3' *Nco* I site contains the ATG initiation codon. Subsequently, V600 was 3'-end-labeled with the Klenow fragment of DNA polymerase and incubated with various amounts (1–5 μg) of nuclear extract protein derived from REX as described in *Materials and Methods*. As a set of controls, we utilized two DNA fragments derived from the $J_{\beta 1}$–$C_{\beta 1}$ intron present in λβREX, which had itself been subcloned into pUC8. These fragments were produced by digestion of a 350-bp *Mae* I fragment digested with *Rsa* I or with *Scr*fI. Fig. 2 shows the results of a representative mobility-shift assay that employed each of these end-labeled fragments in the presence or absence of extract. Lane a shows the mobility of the free V600 fragment in the absence of added nuclear extract. Incubation of V600 prior to electrophoresis with 1 μg of nuclear extract derived from the 0.6 M NaCl fraction eluted from the heparin-Sepharose column resulted in a reduction in intensity of the free V600 band and a heterogeneous shift in the position of V600 (lane b). With addition of a greater amount of extract (5 μg), virtually all of the labeled V600 fragment was retarded at the top of the gel (lane c). In contrast, neither of the two J_β–C_β intron fragments (lanes d and f) was shifted in its mobility by addition of 5 μg of the same REX nuclear extract (lanes e and g).

Detection of DNA Binding Sites by DNase Protection Analysis. The above results suggested that one or more DNA-binding proteins interact in the region upstream of V_β. To identify and localize DNA–protein interactions, we performed DNase protection analysis. In this method the DNA fragment of interest, labeled at a single end, is incubated in the presence or absence of nuclear extract and then partially digested with DNase I, and the resulting DNA fragments are separated by electrophoresis in a sequencing gel. A typical pattern of DNase I digestion of the 3'-end-labeled coding strand of the V400 (Fig. 1) upstream V-region fragment in the absence of nuclear extract is shown in Fig. 3 (lanes a and d). Addition of REX nuclear extract derived from 0.3 M and 0.6 M NaCl fractions from the heparin-Sepharose column resulted in regions of protection from DNase I digestion—i.e., "footprints." In the case of addition of the 0.6 M fraction, five distinct footprints were evident (lane c, footprints I–V). In contrast, the 0.3 M extract yielded only three of these five protected regions in the same fragment (lane b, footprints I, IV, and V). Footprint V was more prominent with the 0.3 M NaCl extract than with 0.6 M NaCl extract. That this

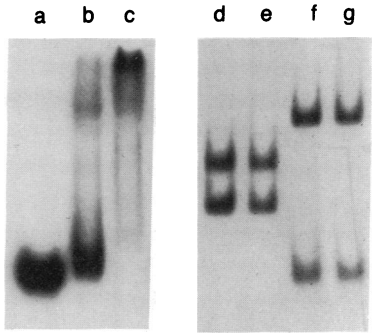

FIG. 2. Mobility-shift assay with end-labeled DNA fragments derived from the REX V_β gene. Lanes: a, V600 alone; b, V600 plus 1 μg of nuclear extract protein; c, V600 plus 5 μg of extract; d, *Mae* I–*Rsa* I fragment alone; e, *Mae* I–*Rsa* I fragment plus 5 μg of extract; f, *Mae* I–*Scr*fI fragment alone; g, *Mae* I–*Scr*fI fragment plus 5 μg of extract.

*National Institutes of Health (1983) Genetic Sequence Databank: GenBank (Research Systems Div., Bolt, Beranek, and Newman, Cambridge, MA), Tape Release 42.0.

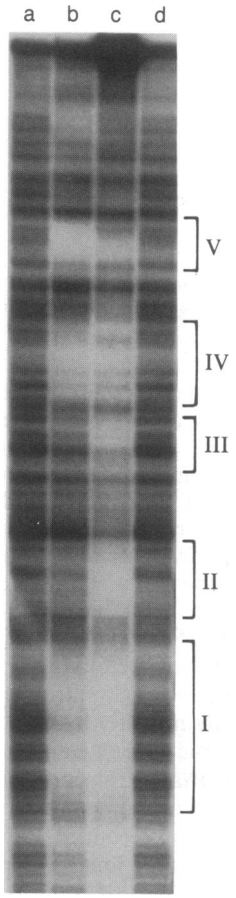

FIG. 3. Analysis of DNase protection of V_β upstream sequence V400 by REX nuclear proteins. Lanes: a and d, no extract; b, 0.3 M NaCl heparin-Sepharose fraction of REX nuclear extract; c, 0.6 M NaCl heparin-Sepharose fraction of REX nuclear extract.

difference is a reflection of intrinsic affinity for a given DNA sequence is likely but not certain.

Comparison of Protected Regions Generated from Various Nuclear Extracts. To determine whether the REX nuclear extract components that produced the regions of protection were restricted to cells of the T lineage, we obtained nuclear extracts from both lymphoid and nonlymphoid cell lines. Fig. 4*A* shows the patterns of footprints on the V400 coding strand generated with nuclear extracts from the T-lineage tumor REX (lanes b and h), the Epstein–Barr virus-transformed B-lymphoblastoid line Laz 509 (lanes c and i), the carcinoma cell line CX1 (lanes d and j), the erythroleukemia line K562 (lanes e and k), a mixture of peripheral blood T and B cells (lane f), and the monkey kidney cell line COS (lane g). In the case of REX, Laz 509, CX1, and K562, both 0.3 M and 0.6 M NaCl extracts were utilized. For comparison, lanes a and l show the DNase I pattern in the absence of extract.

Consistent with the above results, five distinct footprints are evident in the 600 mM REX nuclear extract. In addition, three major and one minor DNase I-hypersensitive sites are noted (arrows) in the pattern obtained with the 600 mM NaCl REX extract. While the major hypersensitive sites flanking footprint V are present in the patterns of all nuclear extracts, the hypersensitive site between footprints III and IV is apparently unique to the patterns seen with the 600 mM REX extract and, to a lesser extent, the 600 mM extract from peripheral blood lymphocytes (>70% T lymphocytes) (lanes b and f, respectively). In addition, footprints I and V are identical for extracts from all cells, whereas footprints III and IV, as defined by the pattern given by the 600 mM REX extract, are observed with extracts from all cells, but with minor variations unique to each cell type (Fig. 4*A*). Perhaps more important, footprint II is restricted to the lymphoid cells REX (lane b), peripheral blood lymphocytes (lane f; the footprint is less prominent), and Laz 509 (lane c; the footprint has a different pattern). The pattern obtained with nuclear extract from a second T-lineage tumor, HPB-ALL, was identical to the pattern obtained with REX (data not shown).

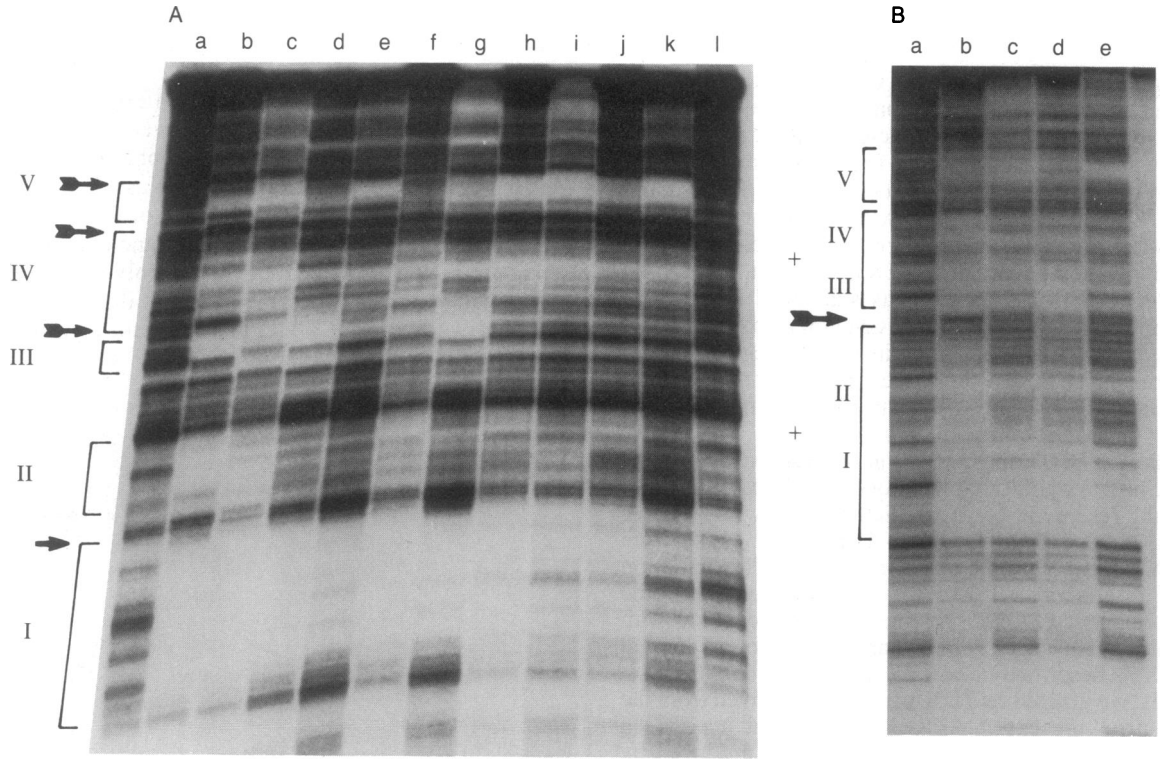

FIG. 4. Analysis of DNase protection of coding (*A*) and noncoding (*B*) strands by nuclear proteins isolated from various cell lines. (*A*) Lanes a and l: no extract. Lanes b–g: 0.6 M NaCl heparin-Sepharose fraction of nuclear proteins isolated from REX (lane b), Laz 509 (lane c), CX1 (lane d), K562 (lane e), peripheral blood lymphocytes (lane f), and COS (lane g). Lanes h–k: 0.3 M NaCl heparin-Sepharose fraction of nuclear proteins isolated from REX (lane h), Laz 509 (lane i), CX1 (lane j), and K562 (lane k). (*B*) Lane a: no extract. Lanes b–e: 0.6 M NaCl fraction from REX, Laz 509, CX1, and K562, respectively. The amount of nuclear protein used in each assay was 10 μg in both *A* and *B*.

In contrast, the 0.3 M NaCl nuclear extracts derived from many cell types gave rise to footprints I, IV, and V (Fig. 4*A* and data not shown).

Collectively, the above results suggest that a protein unique to cells of the T lineage might bind DNA in the region of footprint II. That a different pattern was obtained in a similar region of the V400 coding strand with 0.6 M nuclear extracts from Laz 509 implies that a related but distinct B-cell protein might bind to this area. To determine the protected regions on the complementary DNA strand, the 5′ end of the V400 noncoding strand was labeled with T4 polynucleotide kinase and subjected to footprint analysis (Fig. 4*B*). In comparison to the DNase I pattern generated in the absence of extract (lane a), multiple footprints resulted from addition of the 0.6 M NaCl heparin-Sepharose fraction of REX nuclear extracts (lane b). Based on comparison of DNase patterns of both strands with the corresponding nucleotide sequence (see below), it is possible to relate the footprints obtained on the noncoding strand to those on the coding strand. Fig. 4*B* shows the footprints corresponding to the I–V labeling scheme of the coding strand. The region corresponding to footprints I and II on the coding strand is likewise protected on the noncoding strand (lane b) and, in this case, is represented by a contiguous stretch of protection. A similar contiguous protected region corresponds to footprints III and IV. Note that the region of footprint V is also present on the noncoding strand.

Only the REX extract resulted in generation of a hypersensitive site between footprints II and III (Fig. 4*B*, arrow). Further, although the DNA in region I was protected by extracts from Laz 509 (lane c), CX1, (lane d), and K562 (lane e) cells, the noncoding-strand region corresponding to footprint II on the coding strand was protected only by REX extract (lane b). Collectively, with the results of differences in footprint II generated by REX vs. Laz 509 extracts in the coding strand, the results with the noncoding strand support the view that a protein unique to T cells binds to this region of DNA.

DNA Sequence of the REX V_β Upstream Region. To determine the precise nucleotide sequence of the protected area of coding and noncoding strands in the V_β upstream region, V400 was labeled at either the 3′ or the 5′ end and subjected to the chemical degradation methods of Maxam and Gilbert (11). The nucleotide sequence of the coding strand is shown in Fig. 5 *Upper*. The positions of the "TATA box," putative transcription initiation (cap) site, and translation initiation codon (ATG) are also indicated (Fig. 5 *Lower*). There are no CpG DNA methylation sites within this potential regulatory domain. Computer homology search against the GenBank* primate and rodent files indicated that the region of footprints III and IV bore sequence similarities (score 8.86–12.56) (13) with a number of enhancer motifs. Similarities were found with the polyoma, simian virus 40 (SV40), hepatitis B, human insulin, bovine papilloma virus, and immunoglobulin enhancers. Even though the similarity values are not extremely high, the fact that all the most similar identified regions contained enhancers suggests functional significance. The region of footprint IV in the REX V_β promoter and sequences within the IgG enhancer and SV40 enhancer are compared in Fig. 6 (*Top* and *Middle*, respectively). The similarity with IgG enhancer includes a portion of the immunoglobulin-specific octamer (underlined) and additional sequences. Arrows indicate the presence of an inverted repeat (GTGTG) in the footprint IV sequence. Footprint I also contains an inverted repeat (TCAGTGA; Fig. 6 *Bottom*).

DISCUSSION

Protein-binding sites were characterized in the promoter region of a representative rearranged genomic T-cell receptor V_β gene by means of gel mobility shift, DNase I protection analysis, and nucleotide sequencing. For DNase protection analysis, nuclear extracts were first enriched by heparin-Sepharose chromatography, which selects for DNA-binding proteins on the basis of charge. On the coding strand, five discrete protected regions (footprints) were detected at nucleotides −46 to −68 (I), −72 to −92 (II), −113 to −134 (III), −136 to −175 (IV), and beyond −200 (V). The size of these footprints (≈20 bp) is consistent with areas protected by known DNA-binding proteins (5). In contrast, on the noncoding strand, the protected regions corresponding to footprints I and II and to footprints III and IV were fused. The sequences of the individual protected regions I–IV are not identical, which suggests that each footprint is produced by a different DNA-binding protein. Thus, multiple proteins bind to the promoter region of the REX Ti V_β gene, consistent with findings for well-studied genes in yeast systems (17, 18).

The REX Ti V_β promoter region shows no obvious general homology with known immunoglobulin genes. This is perhaps not surprising, since promoters are generally not conserved, even among immunoglobulin genes, with the excep-

```
           -164                  -144                  -124
GCAAGTGTGC ATCTCTATTT CACACCAATT ATAGTTGAGT TAATTCCTGC

                      -104                  -84
CTGATTCATC TCCCAGAGAT GCAGCCTCCT CTTAAAGAAG TTGGGGGTGG

           -64                   -44        ________-24
TGGCCCATTC AGTGATGTCA CTGACAGATG CATTCTCTGG GGATAAAATG

                         -4   +1              +17
TCACAAAATT CATTTCTTTG CTCATGCTCA CAGAGGGCCT GGTCTAGAAT

           +37                  +55
ATTCCACATC TGCTCTCACT CTGCCATG
```

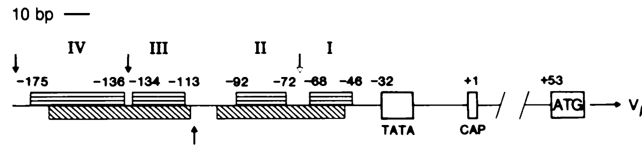

FIG. 5. (*Upper*) Nucleotide sequence of the coding strand of the V_β upstream region and protein binding sites. The location of the "TATA box" is shown by the overline and the putative transcription initiation (cap) site is located at position +1. (*Lower*) Schematic representation of the 5′ V_β gene and distribution of DNase I-hypersensitive sites (arrows) and of protein binding sites on the coding (horizontal lines) and noncoding (diagonal lines) strands.

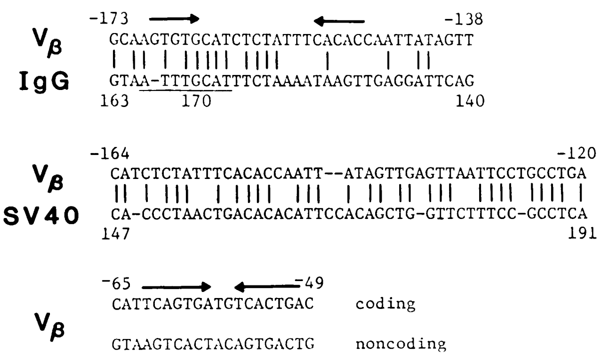

FIG. 6. Features of REX V_β promoter sequences. (*Top* and *Middle*) Comparison of the footprint IV region with IgG and SV40 enhancers. Dashes represent gaps inserted to maximize homology. Arrows indicate an inverted repeat. The immunoglobulin-specific octamer is underlined. (*Bottom*) Footprint I region.

tion of the 8-nucleotide octamer sequence ATTTGCAT (19, 20) and the TATA box. In this regard, it is of note that this octamer is itself not present in the REX V_β gene promoter. Likewise, neither Sp1 recognition sequence [5′ GNTGGGCGGRRY 3′, present in all Sp1-responsive promoters (5)] nor the 25-bp double-stranded segment of DNA protected by NF-1 [consensus sequence 5′ TGGMNNNNNGCCAA 3′, where M = A or C (aMino); ref. 21] is present. These results imply that the proteins binding to the REX V_β promoter region are distinct from those binding to these putative regulatory sequences.

Nevertheless, there is an obvious similarity of known immunoglobulin enhancers (22) with the sequences encompassed in footprints III and IV of the coding strands when short regions of sequence are compared. In particular, we noted that areas around and including some of the octamer sequence are present in footprint IV. This is not unexpected, given that enhancer regions appear themselves to be composed of short DNA sequences or motifs (23). To date, and in contrast with the findings for immunoglobulin heavy and light chains, no evidence for an enhancer region in the J_β–C_β intron of the Ti β-chain gene has been found (unpublished results). Thus, it is possible that the 5′ enhancer-like sequence described in this report represents an upstream transcriptional control element identified by DNase protection. Alternatively, the area might be involved in conferring transcriptional cell-type specificity. In this regard, deletion analysis of an immunoglobulin heavy-chain V gene promoter indicated that the conserved octamer 5′ of the TATA box is a functional component of a tissue-specific upstream promoter element (24). Further evidence for or against this notion must await analysis of *in vitro* transcription assays and *in vivo* transfection studies employing modified promoter regions. However, the striking correlation of footprint analysis and functional studies for known regulatory regions makes this a very likely possibility (25). Because T-cell receptors are not secreted, there is no requirement in the T-cell system (as there is in the case of the B-cell system) for high-level production of this gene product (i.e., no plasma-cell equivalent exists within the T-cell system). That the immunoglobulin enhancer in the intron of immunoglobulin genes evolved from a primordial upstream control element of this putative site is possible but certainly speculative.

Given the contiguity of footprints I and II and footprints III and IV on the noncoding strand, it is possible that each of two proteins within a set contact one another, resulting in cooperative binding. One example of such a cooperative binding interaction occurs in the human rRNA promoter, where one binding factor recruits a second factor to the template to collectively protect an extended region of sequence (26). The GTG-containing inverted repeats in nucleotide sequences within footprints I and IV (Fig. 6) are characteristics of protein-binding sequences in well-defined prokaryotic and eukaryotic systems (15, 27–29).

Tissue specificity of transcription is conferred upon immunoglobulin genes by enhancer elements as well as promoter elements, including the octamer sequence (30, 31). The existence of a binding site for an apparent T-lineage factor as shown herein might be, at least in part, responsible for tissue specificity of Ti β transcripts. It is also possible that T-cell-specific modification of more widespread factors could lead to differential binding. Within human intrathymic ontogeny, Ti β transcription begins in the late stage I–stage II compartment (3). Therefore, it will now be important to determine whether this putative T-lineage-specific DNA-binding protein or other factors are rate-limiting for further progression of the differentiative process.

We thank Dr. Temple F. Smith and Ms. Susan Tolman of the Molecular Biology Computer Research Resource for helpful comments and assistance with computer search, Dr. Temple F. Smith for informative discussions on protein-binding domains, and Dr. Don Diamond for his stimulating assessments. This work was supported by National Institutes of Health Grants 1RO1 CA40134 and 5RO1 AI21226.

1. Reinherz, E. L. Acuto, O., Fabbi, M., Bensussan, A., Milanese, C., Royer, H. D., Meuer, S. C. & Schlossman, S. F. (1984) *Immunol. Rev.* **81,** 95–129.
2. Acuto, O., Hussey, R. E., Fitzgerald, K. A., Protentis, J. P., Meuer, S. C., Schlossman, S. F. & Reinherz, E. L. (1983) *Cell* **34,** 717–726.
3. Royer, H. D., Ramarli, D., Acuto, O., Campen, T. J. & Reinherz, E. L. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 5510–5514.
4. Royer, H. D., Acuto, O., Fabbi, M., Tizard, R., Ramachandran, K., Smart, J. & Reinherz, E. L. (1984) *Cell* **39,** 261–266.
5. Dynan, W. S. & Tjian, R. (1985) *Nature (London)* **316,** 774–777.
6. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C. & Quon, D. (1978) *Cell* **15,** 687–701.
7. Wu, C. (1984) *Nature (London)* **311,** 81–84.
8. Wu, C. (1985) *Nature (London)* **317,** 84–87.
9. Singh, H., Sen, R., Baltimore, D. & Sharp, P. A. (1986) *Nature (London)* **319,** 154–157.
10. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
11. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
12. Smith, T. F. & Faulkner, D. V. (1986) *Comput. Appl. Biosci.*, in press.
13. Smith, T. F. & Waterman, M. S. (1985) *Nucleic Acids Res.* **13,** 645–656.
14. Fried, M. & Crothers, D. M. (1981) *Nucleic Acids Res.* **9,** 6505–6525.
15. Carthew, R. W., Chodosh, L. W. & Sharp, P. A. (1985) *Cell* **43,** 439–448.
16. Galas, D. & Schmitz, A. (1978) *Nucleic Acids Res.* **5,** 3157–3170.
17. Brent, R. & Ptashne, M. (1985) *Cell* **43,** 729–736.
18. Hope, I. A. & Struhl, K. (1985) *Cell* **43,** 177–188.
19. Parslow, T. G., Blair, D. L., Murphy, W. J. & Granner, D. K. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2650–2654.
20. Falkner, F. G. & Zachau, H. G. (1984) *Nature (London)* **310,** 71–74.
21. Henninghausen, L., Siebenlist, U., Danner, D., Leder, P., Rawlins, D., Rosenfeld, P. & Kelly, T., Jr. (1985) *Nature (London)* **314,** 289–292.
22. Banerji, J., Olson, L. & Schaffner, W. (1983) *Cell* **33,** 729–740.
23. Zenke, M., Grundstrom, T., Matthes, H., Wintzerith, M., Schatz, C., Wildeman, A. & Chambon, P. (1986) *EMBO J.* **5,** 387–397.
24. Mason, J. O., Williams, G. T. & Neuberger, M. S. (1985) *Cell* **41,** 479–487.
25. Davidson, I., Fromental, C., Augereau, P., Wildeman, A., Zenke, M. & Chambon, P. (1986) *Nature (London)* **323,** 544–548.
26. Learned, R. M., Learned, T. K., Haltiner, M. M. & Tjian, R. T. (1986) *Cell* **45,** 847–857.
27. Cheung, S., Arndt, K. & Lu, P. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 3665–3669.
28. Ptashne, M. (1980) *Cell* **19,** 1–11.
29. Sadler, J. R., Sasmor, H. & Betz, J. L. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 6785–6789.
30. Church, G. M., Ephrussi, A., Gilbert, W. & Tonegawa, S. (1985) *Nature (London)* **313,** 798–800.
31. Ephrussi, A., Church, G. M., Tonegawa, S. & Gilbert, W. (1985) *Science* **227,** 134–138.