Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 9799-9803, November 1991
Biochemistry

δ, a transcription factor that binds to downstream elements in several polymerase II promoters, is a functionally versatile zinc finger protein

(ribosomal protein gene/intragenic element/histidine tract)

NARAYANAN HARIHARAN*, DAWN E. KELLEY, AND ROBERT P. PERRY†

Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111

Contributed by Robert P. Perry, August 7, 1991

ABSTRACT **The promoters of several eukaryotic genes transcribed by RNA polymerase II contain elements located downstream of the transcriptional start site. To gain insight into how these elements function in the formation of an active transcription complex, we have cloned and sequenced the cDNA that encodes δ, a protein that binds to critical downstream promoter elements in the mouse ribosomal protein rpL30 and rpL32 genes. Our results revealed that the δ protein contains four C-terminal zinc fingers, which are essential for its DNA binding capability and a very unusual N-terminal domain that includes stretches of 11 consecutive negatively charged amino acids and 12 consecutive histidines. The sequence of the δ protein was found to be essentially identical to a concurrently cloned human transcription factor that acts both positively and negatively in the context of immunoglobulin enhancers and a viral promoter. Our structural modeling of this protein indicates properties that could endow it with exquisite functional versatility.**

The transcription of eukaryotic genes by RNA polymerase II is regulated by an interplay of sequence-specific activator or repressor factors and essentially nonspecific general factors and chromatin components (1). The specific factors recognize sequence modules or elements that are situated both in promoters—i.e., regions near the transcriptional start site—and in enhancers—i.e., regions distant from the start site. Specificity in transcriptional regulation is determined by the composition and organization of these elements and the availability of the factors that interact with them.

A considerable effort is currently being made to understand the way in which sequence-specific and general factors interact during formation of active transcription complexes (2, 3). In most mechanistic studies of promoter function, major attention has been given to factors that bind to elements located upstream of the transcriptional start site. However, it has become increasingly evident that a large number of both cellular and viral genes utilize elements that are located downstream of the start site (refs. 4 and 5 and references cited therein). In many cases, the genes that utilize these downstream elements lack a conventional TATA box in the −20 to −30 region. The binding of factors to such downstream promoter elements presents a particular problem that is not encountered by upstream element binding factors—namely, potential interference with the transit of RNA polymerase II, perhaps requiring a reversible alteration in its activation capacity. Although some progress has been made in the purification and preliminary characterization of these factors (5, 6), to date there is no sequence or structural information that might illuminate their functional properties. Therefore, we undertook to clone and sequence the cDNA that encodes a protein termed δ, which binds to critical downstream elements in the genes encoding mouse ribosomal proteins L30 and L32 (rpL30 and rpL32) (Fig. 1*A*) and to downstream elements of several other genes, as inferred from sequence comparisons of functionally important factor binding sites (Fig. 1*B*) and a direct competition assay (10).‡

At the completion of this study, we learned that the δ sequence is almost identical to that of a human transcription factor that can act either negatively or positively in the context of the immunoglobulin κ 3′ and heavy-chain enhancers (13) and the P5 promoter of adeno-associated virus (14). The properties of this protein apparently suit it to function with downstream promoter elements, as well as to perform both activating and repressive functions. Our results have revealed that δ is a zinc finger protein with unusual structural features that could endow it with exquisite functional versatility.

MATERIALS AND METHODS

The λgt11 cDNA library from which we isolated the δ clone was a gift from Murray Brilliant and John Gardner of this Institute. To construct this library, they used a cDNA synthesis kit (Pharmacia) based on the RNase H method of Gubler and Hoffman (15) to generate oligo(dT)-primed, *Eco*RI linker-terminated cDNA from poly(A)$^+$ RNA isolated from the brains of 5-day-old C57BL/6J mice. We initially screened this library by the method of Vinson *et al.* (16) with a ≥10-copy multimer of a double-stranded synthetic oligonucleotide representing the +3 to +41 region of rpL30 joined to an *Eco*RI linker. The oligonucleotide

```
AATTCTCGCTCCCCGGCCATCTTGGCGGCTGGTGTTGGTG
     GAGCGAGGGGCCGGTAGAACCGCCGACCACAACCACTTAA
```

is centered about the δ factor binding site (underlined), which was determined by methylation interference analysis (4) and DNase I footprinting of purified δ protein (D.E.K. and R.P.P., unpublished data). Purified phage plaques were positively screened with this probe and negatively screened with a multimerized segment containing a mutant δ site and with a battery of multimeric probes that contain binding sites for nuclear factors other than δ. In the δ site mutant (gl), the sequence CGGCCA was replaced by ATTTAC (4). The other negative screen probes represented 35- to 45-base-pair (bp) segments containing the binding sites for the rpL30-α, rpL30-γ, and rpL32-β factors (4) and for factors that bind to the −20 to −30 and cap site regions of rpS16 (17).

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.

Abbreviation: rp, ribosomal protein.
*Present address: Bristol-Myers Squibb Pharmaceutical Research Institute, P.O. Box 4000, Princeton, NJ 08543-4000.
†To whom reprint requests should be addressed.
‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. M74590).

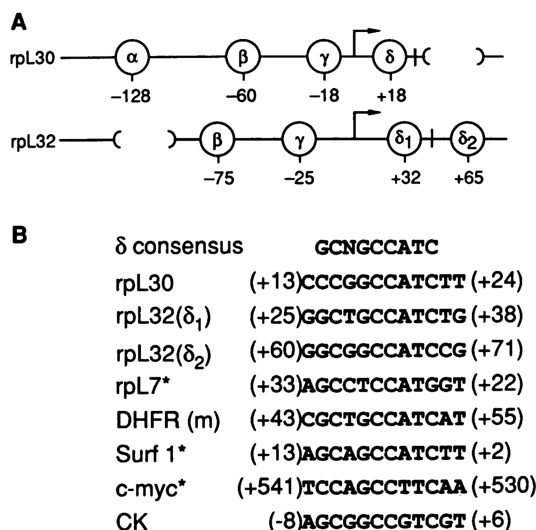

FIG. 1. Occurrence of intragenic δ factor binding sites in transcriptional regulatory elements. (*A*) Schematic representation of the rpL30 and rpL32 promoter regions showing the location of functionally important modules and the corresponding binding sites for a set of discrete nuclear factors designated α, β, γ, and δ (4). Arrows indicate transcriptional start sites. δ_1 and δ_2 are separate δ factor sites. Parentheses indicate functionally important modules for which a factor binding site has not yet been identified. (*B*) δ binding sites in various genes. Numbers in parentheses give the location of the nucleotides preceding and following the indicated sequence. For entries marked with an asterisk, the antisense strand is given; for all others, the sense strand is given. Data references are as follows: rpL30 and rpL32 (δ_1) (4), rpL32 (δ_2) (S. Chung and R.P.P., unpublished data), rpL7 (7), mouse dihydrofolate reductase [DHFR (m)] (8), Surf 1 (9), c-*myc* (10, 11), and creatine kinase (CK) (12).

An *Eco*RI fragment containing the cDNA sequence was excised from the selected λgt11 phage and inserted into the *Eco*RI site of either the transcription vector pGEM-4Z (TC cDNA construct) or a derivative of the transcription–translation plasmid βG-λ3 (18) (TL cDNA construct). The βG-λ3 plasmid contains the 5′ untranslated region and the first two codons of the human β-globin gene inserted between the T7 promoter and the *Eco*RI site of a composite pGEM vector. The *Sma* I, *Hin*cII, and *Sac* I truncation mutants were constructed by standard procedures. Transcription of these constructs with T7 or SP6 polymerase (Promega) in the presence of m^{7}GpppG and translation of the TL transcripts in a 10-μl reaction mixture containing rabbit reticulocyte lysate (Promega) were carried out according to the manufacturer's protocols. Translation products synthesized in the presence of [^{35}S]methionine were electrophoresed on SDS/10% polyacrylamide gels (19) together with a set of ^{14}C protein molecular weight standards (Bethesda Research Laboratories) and were visualized by autoradiography.

Gel mobility-shift assays were carried out as described by Singh *et al.* (20) with 8 μg of nuclear extract (17) or 1 μl of the *in vitro* translation reaction mixture, and 0.1–0.5 ng of ^{32}P-end-labeled oligonucleotide probe (+3 to +41 rpL30 wild-type or gl mutant segments). Methylation interference analysis of rpL30 sense and antisense strands was performed as described (4).

Methods for the extraction and purification of DNA and cytoplasmic poly(A)$^+$ mRNA and for Southern and Northern blot analyses were cited in a previous publication (21). Five-microgram samples of DNA and RNA were used for each analysis. The blots were given two successive 15-min stringent washes (68°C; 0.015 M NaCl/1% SDS) before being exposed for autoradiography.

DNA sequencing was done with a Sequenase kit (United States Biochemical) and by the Maxam–Gilbert chemical degradation reaction. In some cases, inosine triphosphate was substituted for GTP in the dideoxynucleotide reactions in order to minimize gel compression problems.

RESULTS

Identification of the δ Factor Clone. An oligonucleotide containing the rpL30-δ factor binding site was multimerized and used to screen a λgt11 mouse cDNA expression library by selective binding analysis (16, 20). From plates containing a total of ≈3 × 10^5 plaques, we isolated several phage clones that exhibited binding to the δ site. A plaque-purified recombinant that tested uniformly positive for the δ site and uniformly negative for other rp gene binding sites (rpL30-α and -γ, rpL32-β, and two rpS16 sites) and for a nonbinding δ site mutant [gl (4)] was selected for further analysis. A cDNA insert of ≈1.9 kilobases (kb) was excised from this recombinant and inserted into the *Eco*RI sites of two different plasmid vectors, TL and TC. In both vectors, the inserts were flanked by T7 and SP6 transcriptional promoters. Vector TL contains, in addition, the translational initiation region of the human β-globin gene interposed between the T7 promoter and the *Eco*RI insertion site. This vector enables translatable RNA to be produced from any cDNA with an in-phase open reading frame, including those that lack translational start sites. In contrast, RNA transcripts of vector TC will direct protein synthesis only when a functional start site is contained within the cDNA sequence.

To verify the identity of the putative δ factor clone, we transcribed the TL vector constructs with T7 polymerase and translated the resultant transcripts in a reticulocyte lysate. The protein was assayed for DNA binding activity by gel retardation analysis with probes corresponding to both normal and gl mutant δ motifs. A parallel ^{35}S-labeled protein sample was analyzed by SDS/polyacrylamide gel electrophoresis. Translation of the TL cDNA transcripts produced a protein with an apparent molecular mass of 60–65 kDa (Fig. 2*A*, lane 1). When mixed with the δ probe, this protein formed a complex that migrated slightly slower than that formed by the natural nuclear protein (Fig. 2*B*, lane 3 vs. lane 2). Like the natural δ factor, the *in vitro*-synthesized protein did not bind to the gl mutant probe (data not shown).

As an additional check on the specificity of the cDNA clone, we compared the methylation interference footprint of the *in vitro* synthesized protein with that of the natural nuclear protein (Fig. 2*C*). The two proteins produced the same footprint, which was essentially identical to that obtained in earlier studies with nuclear extracts. These results confirmed our identification of the δ factor clone and encouraged us to determine the sequence of the 1.9-kb cDNA.

Sequence of the δ cDNA and Its Encoded Protein. The *Eco*RI-bounded cDNA insert comprised 1898 bp with a single long open reading frame extending from the 5′ end to position 1285 (Fig. 3). A protein initiating at the β-globin start site of the TL transcript would contain 430 amino acids (four from the β-globin sequence and *Eco*RI site and 426 from the δ cDNA). Since the mobility of the binding complex formed by the *in vitro* synthesized protein was slightly slower than that formed by the natural protein, we suspected that the in-phase AUG codon at positions 43–45, which occurs in a favorable context (22), might be the true initiation codon for this protein. To examine this possibility, we carried out a transcription–translation experiment with a plasmid in which the cDNA sequence was oriented downstream of the SP6 promoter in the TC vector. Translation of SP6 transcripts from this vector produced a protein that had a slightly faster mobility than that produced from TL transcripts (Fig. 2*A*, lane 2), consistent with a difference of 16 amino acids in the encoded proteins. Moreover, complexes between the δ probe and TC-directed protein migrated identically to those observed with the natural nuclear protein (Fig. 2*B*, lane 4 vs. lane 2). These results

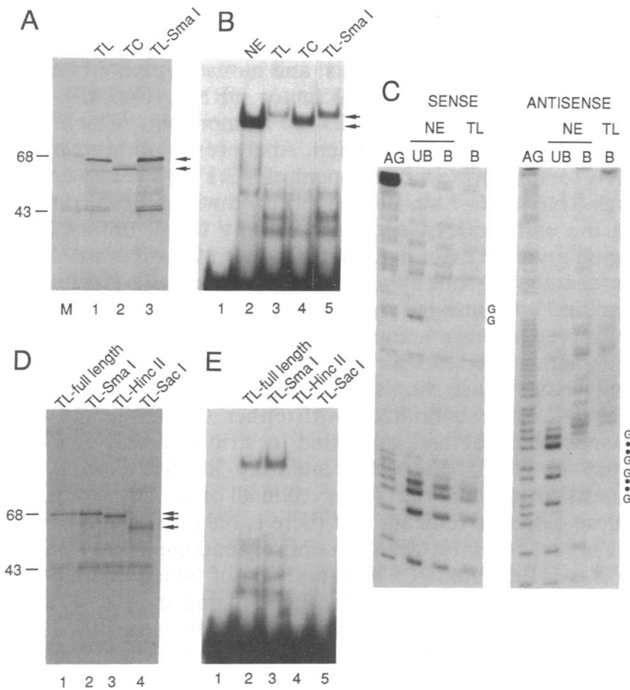

FIG. 2. Characterization of transcription–translation products of the δ factor cDNA and various truncation derivatives. The δ factor cDNA and derivatives of this cDNA that were cleaved at the indicated restriction sites were inserted into the TL or TC vector, transcribed with T7 or SP6 RNA polymerase, and translated in a reticulocyte lysate system. (*A* and *D*) ^{35}S-labeled translation products electrophoresed on SDS/polyacrylamide gels. Numbers on the left indicate size of markers in kDa. Arrows on the right indicate positions of the relevant translation products. (*B* and *E*) Gel retardation analyses of unlabeled translation products and nuclear extract protein from S194 cells (NE). No protein was added to the probe in lane 1. (*C*) Comparative methylation interference analysis of TL cDNA translation product and nuclear extract protein. UB, unbound fragment; B, bound fragment; AG, G+A sequencing reaction. G contact residues are indicated on the right.

strongly indicate that the AUG codon at positions 43–45 is indeed the true initiation codon and that we have, in fact, cloned the entire coding sequence of the δ factor.

The molecular weight of the δ protein calculated from the amino acid sequence is 44,557. This is significantly smaller than the apparent molecular mass of ≈60 kDa estimated from the mobility on SDS/polyacrylamide gels relative to a standard set of globular protein markers. Because of this discrepancy, we decided to verify that the termination codon at positions 1285–1287 was truly functional. For this purpose, we made T7 transcripts of TL cDNA that was truncated at the *Sma* I site at position 1314 and examined the mobility and binding properties of the protein encoded by these transcripts. No differences were observed between the proteins produced by normal and *Sma* I-truncated TL plasmids (Fig. 2*A*, lanes 1 and 3; Fig. 2*B*, lanes 3 and 5), indicating that the termination codon at positions 1285–1287 was indeed functional. Thus, the protein clearly has an anomalously slow electrophoretic mobility, presumably because of its very unusual properties (see below).

Examination of the δ protein sequence (Fig. 3) revealed four zinc fingers of the Cys-Cys-His-His variety in the C-terminal region. The N-terminal portion is extremely acidic. Within the first 53 residues there are 14 glutamic acids and 7 aspartic acids, 11 of which occur in a single stretch. Perhaps the most bizarre feature of all is a stretch of 12 consecutive histidines. Also notable are the extensive stretches of glycine and/or alanine. The possible significance of these unusual features will be discussed later.

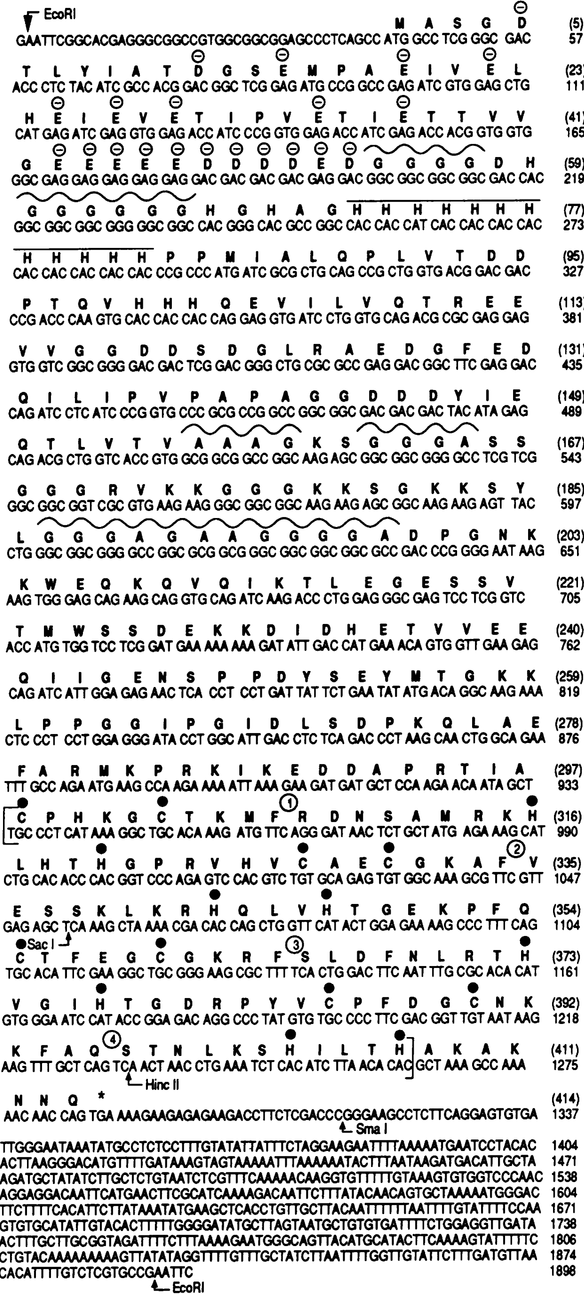

FIG. 3. Sequence of the δ factor cDNA. Negatively charged amino acids in the N-terminal region are indicated by encircled minus signs. The long histidine tract is overlined with a straight line. Tracts of four or more glycines or glycines and alanines are overlined with wavy lines. The four zinc fingers are enclosed by brackets with the cysteine- and histidine-defining residues indicated by dots. Only relevant restriction sites are shown.

A search of the sequences deposited in the GenBank and EMBL compilations did not reveal any entries with overall similarity to the δ protein. Of course, zinc fingers occur in many other DNA binding proteins and numerous examples with some similarity in this region were noted. Of the various zinc finger sequences in the data banks, that in the REX-1 protein (23) was the most similar to the δ sequence. In the region encompassing the four zinc fingers, these two proteins exhibit 76% identity and 85% similarity. However, they are totally dissimilar outside of this region.

The Zinc Finger Domain Is Essential for DNA Binding. To examine the importance of the zinc fingers for DNA binding, we truncated the TL cDNA at either the *Sac* I site at position

1054, which eliminates two-thirds of the zinc finger region, or the *Hin*cII site at position 1232, which amputates half of the C-terminal finger. The *in vitro* synthesized protein produced by the transcription and translation of these truncated plasmids (Fig. 2*D*, lanes 3 and 4) was incapable of binding to the δ probe (Fig. 2*E*, lanes 4 and 5), thus attesting to the importance of the zinc finger domain for DNA binding. Indeed, disruption of only the C-terminal finger motif was sufficient to abolish the DNA binding properties of the δ protein.

Widespread Occurrence of the δ Factor. A survey of nuclear extracts from a variety of cell types and animal species has indicated that the δ protein is present in many different types of tissue and that it is highly conserved among mammalian species. As shown in Fig. 4*A*, the retardation complexes observed with nuclear extracts from cells of mouse lymphoid, monkey kidney, and human cervical origin are indistinguishable from each other. Northern blot analysis of cytoplasmic poly(A)$^+$ mRNA from a variety of sources—i.e., mouse early-, mid-, and late-stage lymphoid cells; mouse fibroblasts; human epithelioid cells; and human lymphoid cells—showed similar patterns of δ factor mRNA (Fig. 4*B*). The predominant species is a 2.4-kb component, which was evident in all cell types studied. Also present in most of the RNA samples were a component of ≈2.1 kb and one or two large RNAs of 6–7 kb. Blots hybridized with probes containing the entire cDNA sequence or only the 3′ untranslated region revealed essentially the same set of components. The fact that the 3′ probe hybridizes with human mRNA indicates that the 3′ untranslated region of δ factor mRNA is also highly conserved among mammals. The strong conservation among mammalian species was also evident from a Southern blot analysis of mouse, human, monkey, and deer DNA. When the blots were hybridized with either the full-length or a 3′-specific probe and submitted to stringent washes (68°C; 0.015 M NaCl) (Fig. 4*C*), intensely labeled *Bam*HI and *Hin*dIII fragments were observed in all of the DNA samples. These data also indicate that there is only a single δ gene in the mouse, which is largely or entirely encompassed by a pair of *Bam*HI or *Hin*dIII fragments, one of which contains the cloned portion of the 3′ untranslated region and the other the remainder of the cDNA sequence. Since the same probes and hybridization stringencies were used for both the Southern and Northern blot analyses, it seems reasonable to conclude that the multiple forms of δ mRNA are the result of alternative or incomplete RNA processing.

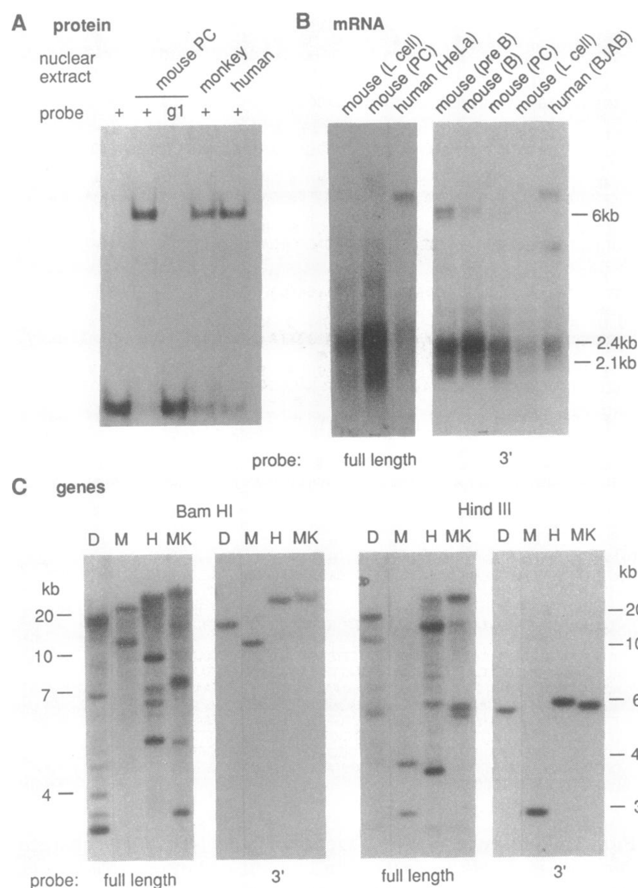

FIG. 4. Occurrence of δ factor protein, mRNA, and genes in cells from different tissues and different mammalian species. (*A*) Gel retardation analysis of nuclear extracts from S194 mouse plasmacytoma (PC) cells, monkey kidney (COS-7) cells, and human epithelioid (HeLa) cells. Probes are the same as in Fig. 2*A*. (*B*) Northern blot analysis of cytoplasmic poly(A)$^+$ mRNA from mouse fibroblasts (L cell), PC cells (S194), pre-B cells (3-1), and B cells (WEHI 231) and from human epithelioid cells (HeLa) and lymphoid cells (BJAB). Probes represented either the full-length cDNA (*Eco*RI fragment) or the 3′ untranslated region (*Sma* I/*Eco*RI fragment). Approximate sizes of RNAs are indicated on the right. Although the 6-kb components were not detected in the illustrated sample of L cell mRNA, these components were clearly evident in an mRNA sample from another L cell culture. (*C*) Southern blot analyses of *Bam*HI and *Hin*dIII digests of DNA from four mammalian species [D, deer (Muntjac cells); M, mouse (S194 cells); H, human (HeLa cells); MK, monkey kidney (COS) cells]. Probes are the same as in *B*. Fragment size scales based on coelectrophoresed marker fragments are shown on the right and left.

DISCUSSION

Using the selective binding procedure (16, 20) to screen a λgt11 mouse cDNA library, we have cloned the cDNA encoding the δ transcription factor. The authenticity of this cDNA clone was established by gel-retardation and methylation interference analyses of protein synthesized from transcripts of the cDNA. The importance of the δ factor for gene expression stems from the fact that its target DNA sequence is a critical intragenic promoter element in several ribosomal protein genes and in other genes that are expressed in a wide variety of tissues (Fig. 1). Expression of the rpL30 gene *in vivo* was decreased by more than an order of magnitude by a mutation in the δ element that abolishes its factor binding ability (4). Similar mutations in the pair of rpL32-δ elements or a 33-bp deletion of rpL32 exon 1 sequences that eliminated the δ1 binding site had essentially the same effect (ref. 24; S. Chung and R.P.P., unpublished data). Deletion of the δ factor binding site in the dihydrofolate reductase gene caused a 5-fold reduction of *in vitro* transcriptional activity (8). The fact that δ factor functions downstream from the transcriptional start site in genes that generally lack well-defined TATA boxes suggests that it may have a special role in the formation of the transcriptional initiation complex for this class of genes.

The δ factor has a molecular weight of 44,557. At its C terminus there are four zinc fingers that are essential for its DNA-binding properties. Removal of only a portion of the C-terminal-most finger is sufficient to abolish its capacity to bind a target sequence in the rpL30 gene.

The consensus core sequence for δ factor binding is GCNGCCATC. Minor deviations from this sequence are apparently tolerated in some genes and additional similarities in flanking sequences can sometimes be discerned (Fig. 1). Although a downstream element in the adenovirus major late promoter contains a similar sequence (GCGTCCATC), it appears to bind a protein other than δ. This protein has an apparent molecular mass of 40 kDa, as judged by UV-crosslinking data (6), whereas a UV-crosslinking analysis of affinity-purified δ protein (D.E.K. and R.P.P., unpublished data) indicates an apparent molecular mass of ≈60 kDa,

consistent with the size of the *in vitro* translation product of the δ cDNA transcripts. In the adenovirus major late promoter, the δ-like sequence may overlap with a higher-affinity site for another protein.

The N-terminal portion of the δ protein is very unusual. To the best of our knowledge, the stretches of 11 consecutive negatively charged amino acids and 12 consecutive histidines is unprecedented among previously characterized transcription factors. A structural model based on an algorithm developed by Ross and Gollub (25) indicates that these stretches could form two oppositely charged symmetrical helices separated by a highly flexible glycine-rich loop. The model further predicts that these helices are joined to a more N-terminal acidic amphiphilic helix via a β-sheet loop structure. Such features might provide this protein with an acidic activation domain that could be neutralized or modulated under certain sets of conditions. This property might be particularly useful for a protein that functions downstream of the transcriptional start site because it could allow alternative interactions with polymerase II general transcription factors (e.g., TFIIB) before and after transcription has commenced.

Similar modeling of the rest of the δ protein indicated a central region that is largely unstructured, consisting of extensive loop regions interspersed with short strongly positive and strongly negative helices. This unstructured region may contribute to the anomalously slow electrophoretic mobility of the δ protein. The C-terminal region was predicted to have a structure that contains alternating antiparallel β-sheets and α-helices, features that are demonstrably present in zinc fingers of the Cys-Cys-His-His type (26).

A comparison of the sizes of the principal mRNA component (≈2.4 kb) and the cDNA (1.9 kb) indicates that the cDNA clone may be lacking 0.3–0.4 kb of mRNA sequence, most likely at the 3′ end. A slightly shorter mRNA (≈2.1 kb), which is present in some mRNA samples, may represent a species with an alternative poly(A) site. The cloned portion of the 3′ untranslated region (UTR) contains several oligo(A) and oligo(T) tracts. The A+U-richness of the 3′ UTR (68% A+U) is in striking contrast to the C+G-richness of the 5′ portion of the mRNA (73% G+C in the first 640 nucleotides). The A+U-rich 3′ UTR, which is highly conserved among mammals, may play a role in regulating the metabolic stability and/or translational efficiency of the δ mRNA. A+U-rich motifs are known to be implicated in determining mRNA half-life (27) and cytoplasmic poly(A) addition (28).

At the completion of this study, we learned that cDNA clones encoding the human counterpart of the δ factor had been isolated in studies of the 3′ enhancer of the mouse κ immunoglobulin gene (13) and the P5 promoter of adeno-associated virus (14). This factor apparently corresponds to the factor termed common factor 1, which binds to the μE1 element in the intronic enhancer of the immunoglobulin heavy-chain genes and to upstream elements in the skeletal actin and c-*myc* genes (29). The 3′ κ, μE1, and P5 binding sites match the rpL30-δ binding site in 8 of 10, 10 of 10, and 7 of 10 positions, respectively. Thus, these findings further expand the number of genes known to be transcriptionally regulated by the δ factor. Interestingly, this factor seems to exert negative regulation on the 3′ κ enhancer (13), positive regulation on the intronic heavy-chain enhancer (ref. 29 and references therein), and both positive and negative regulation on the P5 promoter. Conceivably, the diversity of its functional roles may be related to its structural plasticity.

As predicted from our Northern and Southern blot analyses, the sequence conservation between the mouse and human cDNAs is substantial. The sequences are 95% identical in both coding and untranslated regions. In the coding region, there are 47 third-base neutral substitutions, and 4 replacement substitutions, 2 discrete codon insertions, and 2 discrete codon deletions. The mouse and human proteins are therefore 98.6% identical. Given the widespread occurrence of this factor in various tissues and its possible functional diversity, we suspect that it plays a key role in the expression of many mammalian genes.

Note Added in Proof. Contrary to our earlier finding (4), recent studies by G. Safrany and R.P.P. (unpublished data) indicate that the activity of an otherwise intact rpL30 promoter is not affected by a mutation (gl) that abolishes δ factor binding.

The authors would like to thank Dr. Murray Brilliant for providing the λgt11 library, Dr. Tom Kaedesch for the transcription/translation vector, and Anita Cywinski for help with the DNA sequencing. We are also indebted to Drs. Eileen Jaffe and Alfred Zweidler for help with the protein–structure programs and to Drs. Michael Atchison and Tom Shenk for communicating their findings to us prior to publication. This research was supported by grants from the National Science Foundation (DMB-8918147) and the National Institutes of Health (AI-17330, CA-06927, RR-05539), and by an appropriation from the Commonwealth of Pennsylvania.

1. Sharp, P. A. (1991) *Nature (London)* **351**, 16–18.
2. Mitchell, P. J. & Tjian, R. (1989) *Science* **245**, 371–378.
3. Ptashne, M. & Gann, A. (1990) *Nature (London)* **346**, 329–331.
4. Hariharan, N., Kelley, D. E. & Perry, R. P. (1989) *Genes Dev.* **3**, 1789–1800.
5. Ayer, D. E. & Dynan, W. S. (1990) *Mol. Cell. Biol.* **10**, 3635–3645.
6. Cohen, R. B., Yang, L., Thompson, J. A. & Safer, B. (1988) *J. Biol. Chem.* **263**, 10377–10385.
7. Meyuhas, O. & Klein, A. (1990) *J. Biol. Chem.* **265**, 11465–11473.
8. Farnham, P. J. & Means, A. L. (1990) *Mol. Cell. Biol.* **10**, 1390–1398.
9. Lennard, A. & Fried, M. (1991) *Mol. Cell. Biol.* **11**, 1281–1294.
10. Atchison, M. L., Meyuhas, O. & Perry, R. P. (1989) *Mol. Cell. Biol.* **9**, 2067–2074.
11. Yang, J.-Q., Remmers, E. F. & Marcu, K. B. (1986) *EMBO J.* **5**, 3553–3562.
12. Mitchell, M. T. & Benfield, P. A. (1990) *J. Biol. Chem.* **265**, 8259–8267.
13. Park, K. & Atchison, M. L. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9804–9808.
14. Shi, Y., Seto, E., Chang, L.-S. & Shenk, T. (1991) *Cell* **67**, 377–388.
15. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25**, 263–269.
16. Vinson, C. R., LaMarco, K. L., Johnson, P. F., Landschulz, W. H. & McKnight, S. L. (1988) *Genes Dev.* **2**, 801–806.
17. Hariharan, H. & Perry, R. P. (1989) *Nucleic Acids Res.* **17**, 5323–5337.
18. Beckman, H., Su, L.-K. & Kadesch, T. (1990) *Genes Dev.* **4**, 167–179.
19. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
20. Singh, H., LeBowitz, J. H., Baldwin, A. S. & Sharp, P. A. (1988) *Cell* **52**, 415–423.
21. Kelley, D. E., Pollok, B. A., Atchison, M. L. & Perry, R. P. (1988) *Mol. Cell. Biol.* **8**, 930–937.
22. Kozak, M. (1989) *J. Cell Biol.* **108**, 229–241.
23. Hosler, B. A., LaRosa, G. J., Grippo, J. F. & Gudas, L. J. (1989) *Mol. Cell. Biol.* **9**, 5623–5629.
24. Moura-Neto, R., Dudov, K. P. & Perry, R. P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3997–4001.
25. Ross, A. M. & Gollub, E. E. (1988) *Nucleic Acids Res.* **16**, 1801–1812.
26. Pavletich, N. P. & Pabo, C. O. (1991) *Science* **252**, 809–817.
27. Shaw, G. & Kamen, R. (1986) *Cell* **46**, 659–667.
28. Wickens, M. (1990) *Trends Biochem. Sci.* **15**, 320–324.
29. Riggs, K. J., Merrell, K. T., Wilson, G. & Calame, K. (1991) *Mol. Cell. Biol.* **11**, 1765–1769.