One Factor Recognizes the Liver-Specific Enhancers in α_1-Antitrypsin and Transthyretin Genes

DENNIS R. GRAYSON, ROBERT H. COSTA,
KLEANTHIS G. XANTHOPOULOS, JAMES E. DARNELL

Four different regulatory sites required for transcriptional stimulation by the enhancers of two unrelated liver-specific genes α_1-antitrypsin and transthyretin appear to bind the same nuclear protein that is found mainly in the liver. Such proteins may provide a basis for a coordinated, hepatocyte-specific control of gene transcription.

THE REGULATION OF GENE EXPRESsion in mammalian cells is an area of intense research activity. Nuclear proteins that are present in a wide variety of cell types and that play a role in expression of RNA polymerase II transcription units have been discovered, several have been purified, and the genes encoding these generally distributed factors already are or are soon to be cloned. An emerging rule regarding the regulatory elements controlling widely expressed viral genes (*1–3*) and certain cellular genes [histones (*4, 5*), human thymidine kinase (*6*), metallothionein (*7*), and heat shock (*8*)] is that multiple DNA regions are required for the expression of each gene. The functionally defined DNA elements are recognized by different proteins found in many cell types, and this interaction is required for active transcription of these genes. The resulting interactions between protein and DNA play a role in mediating the transcriptional activity of these genes.

Cell-specific, coordinated control of groups of genes that are expressed only or mainly in a single tissue or a single cell type raises three important questions: Do these genes also require multiple protein factors for expression? If so, are any of the factors limited in cell distribution, perhaps to a single differentiated cell type? Finally, do any such cell-specific factors regulate several genes coordinately?

To answer these questions we have been analyzing mouse genes transcribed only (or mainly) in hepatocytes (*9–12*). Upstream enhancer sequences as well as sequences proximal to the RNA start site in two of these genes, α_1-antitrypsin (α1-AT) and transthyretin (TTR), are required for cell-specific expression in a human hepatoma (Hep G2) cell line (*13, 14*). In the 100-nucleotide-long enhancer region of the TTR gene, three DNA regions are required for full activity. Two of these sequences represent independent protein-binding sites that interact with a factor that is liver-specific, while a third region binds a widely distributed factor (*13*). Here we show that the three DNA regions within the α1-AT enhancer, each of which is required for function (*14*), are recognized by factors that are shared between these two enhancers. Four of these sites compete for the same liver-specific protein factor, whereas two sites compete for a factor in all of the extracts tested.

A diagram of the relevant regulatory regions of the two genes TTR and α1-AT is shown (Fig. 1A). Both these genes are expressed at high rates in the liver at the same period during mouse development and are not transcribed in the majority of other cell types in the animal (*11*). The proteins encoded by the two genes are secreted into the blood, but the protein sequences and their associated functions are unrelated (*15, 16*). For maximal cell-specific transcription, both genes require sequences within the first 150 to 200 nucleotides 5′ to the RNA start site (promoter proximal elements) and more distant cell-specific enhancer elements (*13, 14*). Both the enhancer and promoter proximal sequences are required for maximal transcription in hepatoma cells, but neither enhancer nor promoter proximal sequence are active in HeLa cells (*13, 14*). The necessary protein-binding regions in the TTR enhancer are labeled 1, 2, and 3 and those in the α1-AT enhancer are labeled A, B, and C. The sequences of the oligonucleotides are given in Fig. 1B and are discussed below.

No obvious conserved sequence exists in the protein-binding sites of the α1-AT or of the TTR enhancer or promoter regions. In spite of the apparent lack of sequence similarity, sites 2 and 3 in the TTR enhancer region cross-compete in protein-binding or gel-shift assays (*13*). Therefore, we determined whether any of the binding sites (A, B, or C) in the α1-AT enhancer region were related to each other. Protein-binding experiments with both methylation-protection and methylation-interference assays (*14*) had defined 15 to 20 nucleotide-binding sites in the regions designated A, B, and C. Therefore, we synthesized oligonucleotides (25 to 30 bp from each strand) that corresponded to the binding sites in the A, B, and C regions (Fig. 1B). Each of these double-stranded oligonucleotides was end-labeled and incubated with nuclear extracts of liver or Hep G2 cells (Fig. 2, left). Each oligonucleotide bound to a presumed factor or factors and exhibited one or more slowly migrating or "shifted" bands after gel electrophoresis and autoradiography (*17, 18*). The bands were specific for the particular

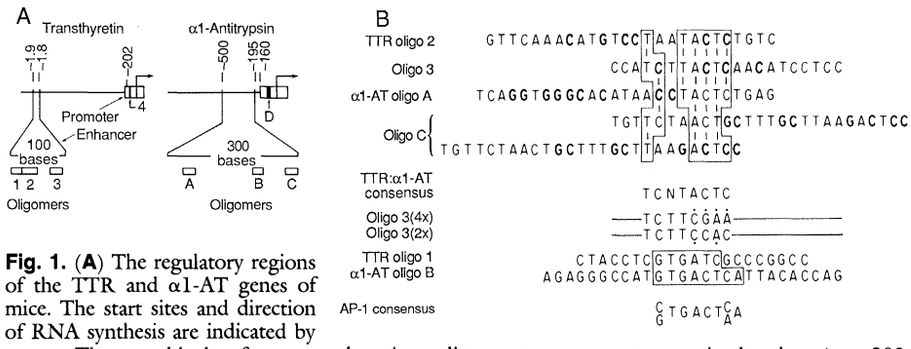

Fig. 1. (**A**) The regulatory regions of the TTR and α1-AT genes of mice. The start sites and direction of RNA synthesis are indicated by arrows. The open blocks of sequence show immediate upstream promoter proximal regions (to −202 nt in TTR and −160 nt in α1-AT) that serve a hepatoma cell–specific promoter function resulting in low-level expression. Not shown are the TATA consensus sequences located at position −30 nt relative to the RNA start sites. Sequence blocks further upstream (~−1.8 to −1.9 kb for TTR and −190 to −500 bp in α1-AT) serve a hepatoma cell–specific enhancer function resulting in a 10- to 30-fold transcriptional induction (*13, 14*). Double-stranded oligonucleotides 1, 2, and 3 of TTR and A, B, and C of α1-AT were constructed from each gene sequence to include 25- to 28-nucleotide stretches to which liver and hepatoma nuclear proteins were shown to bind (*13, 14*). These DNA regions were necessary for full enhancer activity. (**B**) Proposed consensus sequence for liver-specific and common binding sites within the TTR and α1-AT enhancers. The first five lines are the sequences of oligonucleotides used. The alignment maximizes overlap in a proposed binding site of a liver-specific factor (sixth line). Oligomer C is shown in two alignments since base contacts have been observed in two areas of this oligonucleotide region. Boldface letters represent sites where methylation protection and interference experiments indicated protein binding (*13, 14*). Within the consensus region, two mutant sequences of oligomer 3 were constructed [lines 7 and 8; oligo 3(4×) and oligo 3(2×)]. The dots indicate the altered bases within the sequence. The bottom three lines show the relation to the AP-1 consensus (*19, 20*) of the sequences within the TTR oligomer 1 and the α1-AT oligomer B.

The Rockefeller University, New York, NY 10021.

oligonucleotide because unlabeled homologous oligonucleotide in a 50-fold molar excess was capable of interrupting the formation of the different gel-shift bands. This assay allowed us to compare the specific nuclear-binding proteins recognizing the various α1-AT oligonucleotides.

In cross-competition experiments unlabeled oligomer A interrupted the band shift of labeled oligomer C and vice versa (Fig. 2, left side, left and right). However, when oligomer B was labeled and used in a gel-shift assay neither unlabeled oligomer A nor C competed the specific shifted band (Fig. 2, left, center). The unlabeled oligomer B did, however, reduce the formation of gel-shift bands by labeled oligomers A or C but did not block these bands completely. Thus the pattern previously seen in the TTR gene was also observed with the α1-AT enhancer: multiple binding sites existed within an enhancer region, at least two of which were related.

Therefore, we next compared each of the sites in the two enhancers. The experiments indicated that the two related TTR oligomers, 2 and 3, did in fact cross-compete for binding with the two similar α1-AT oligomers, A and C (Fig. 2, right). All of these sites therefore may bind a single liver-specific protein. Likewise the two sites that bind the factor present in all cells, oligomer 1 of TTR and oligomer B of α1-AT, also showed cross-competition.

A sample of these results showing the quantitative competing ability of various oligonucleotides is presented in Fig. 2, right. For example, with labeled α1-AT oligomer A an equal competition by α1-AT oligomer C was found (Fig. 2, right, a), but a more potent inhibition of complex formation by either TTR oligomer 2 or 3 was noted (Fig. 2, right, c). Furthermore, labeled oligomer 2 was competed equally well by unlabeled TTR oligomer 2 or 3 and not as effectively by either unlabeled α1-AT oligomer A or C (Fig. 2, right, b and d). (Five to ten times as much α1-AT oligomer A is required to achieve an equal competition as with unlabeled TTR oligomer 2.) Neither labeled TTR oligomer 2 nor labeled α1-AT oligomer A was competed by a 20-fold molar excess of TTR oligomer 1 (Fig. 2, right, b, c, and d). These results establish a family of related binding sites (TTR oligomers 2 and 3; α1-AT oligomers A and C with a hierarchy of binding affinities: 2, 3 > A, C). The other two sites under consideration, TTR oligomer 1 and α1-AT oligomer B also showed cross-competition with each other, with α1-AT oligomer B having a five- to tenfold greater competing capacity for both labeled oligomers (Fig. 2, right, e).

We next determined whether the nuclear proteins that bound to any of the enhancer sites of α1-AT were limited in their cellular distribution. The factors that bound to oligomers 2 and 3 of TTR had previously been shown to be more prevalent in liver than other cells, whereas oligomer 1 of TTR bound to nuclear proteins extracted from all cells tested (*13*). Extracts of nuclei from cultured HeLa and Hep G2 cells and extracts of nuclei from rat liver, brain, and spleen were each used in separate gel-shift experiments. All of the extracts contained nuclear protein capable of interacting with the labeled element B of α1-AT, producing a major gel-shift band that could be competed with the unlabeled homologous oligomer (Fig. 3, right). Within the footprint region of oligomer B (*14*) is the sequence GTCACTA, which is the reported AP-1 consensus sequence (*19, 20*). Thus the AP-1 protein, which is involved in the stimulation of several viral and cellular enhancers, is a likely candidate for binding to oligomer B of α1-AT.

In contrast to the results obtained with oligomer B of α1-AT, when labeled oligomer A was used in assays with various nuclear extracts, liver and Hep G2 nuclear proteins produced strong gel-shift bands, while all the other extracts produced very faint bands (Fig. 3, left panel). The same results were obtained with oligomer C (data not shown). Thus the protein(s) that interact with oligomer B are present in many cell types, but the proteins that recognize oligomers A and C are much more prevalent in cells of hepatic origin.

The results described so far indicate that within the various regions of protein interaction in the TTR and α1-AT enhancers there appear to be two families of sequences. The sequences of α1-AT oligomer B showed an exact match with the consensus sequence for the widely distributed transcriptional factor AP-1 (*19, 20*). Therefore, one of these sequence families is served by a previously recognized factor. TTR oligomer 1 contains a similar sequence and on the basis of this

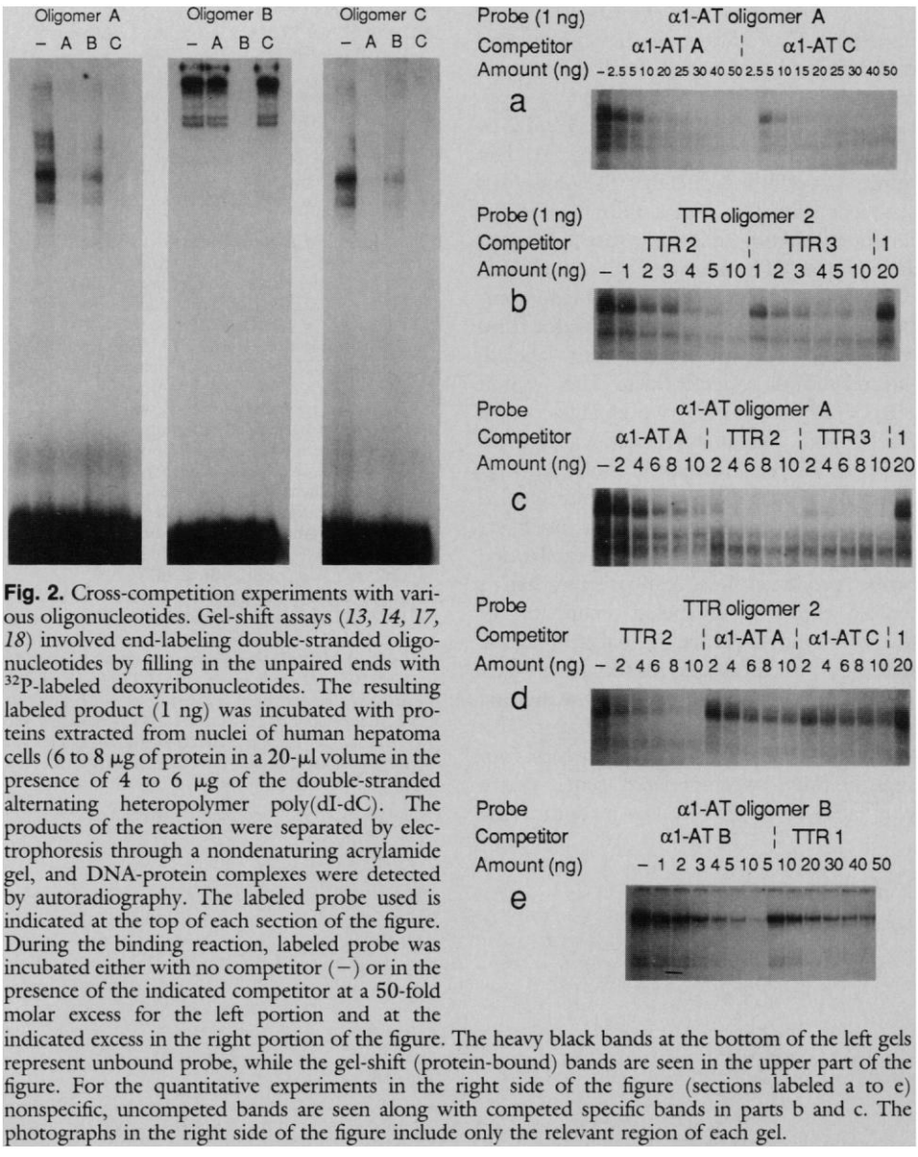

Fig. 2. Cross-competition experiments with various oligonucleotides. Gel-shift assays (*13, 14, 17, 18*) involved end-labeling double-stranded oligonucleotides by filling in the unpaired ends with ^{32}P-labeled deoxyribonucleotides. The resulting labeled product (1 ng) was incubated with proteins extracted from nuclei of human hepatoma cells (6 to 8 μg of protein in a 20-μl volume in the presence of 4 to 6 μg of the double-stranded alternating heteropolymer poly(dI-dC). The products of the reaction were separated by electrophoresis through a nondenaturing acrylamide gel, and DNA-protein complexes were detected by autoradiography. The labeled probe used is indicated at the top of each section of the figure. During the binding reaction, labeled probe was incubated either with no competitor (–) or in the presence of the indicated competitor at a 50-fold molar excess for the left portion and at the indicated excess in the right portion of the figure. The heavy black bands at the bottom of the left gels represent unbound probe, while the gel-shift (protein-bound) bands are seen in the upper part of the figure. For the quantitative experiments in the right side of the figure (sections labeled a to e) nonspecific, uncompeted bands are seen along with competed specific bands in parts b and c. The photographs in the right side of the figure include only the relevant region of each gel.

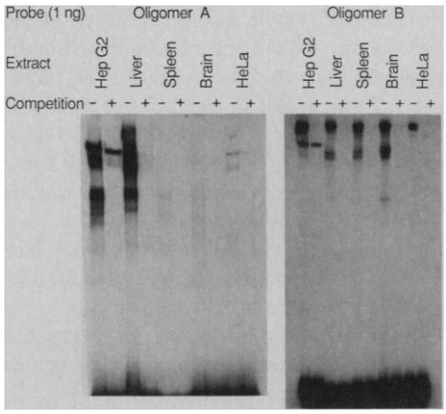

Fig. 3. Tissue or cell distribution of proteins that bind α1-AT oligomers A and B. Nuclear extracts from the indicated cells or tissues were exposed to either labeled oligomer A or B without (−) or with (+) a 50-fold molar excess of unlabeled homologous oligomer. The cell and tissue extracts were prepared as described (*13, 14*).

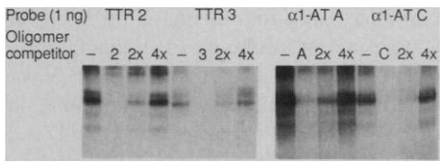

Fig. 4. Effect of mutations in oligomer 3 of TTR in DNA-binding assays. Gel-shift experiments were performed as in Figs. 2 and 3. Labeled TTR oligomer 2 or 3 or labeled α1-AT oligomers A or C were competed in gel-shift assays by a tenfold excess of homologous unlabeled oligonucleotide or by a 100-fold excess of either of the two mutant oligonucleotides [oligo 3(2×) or oligo 3(4×); see Fig. 1 for sequence].

weaker competition for the same factor as α1-AT oligomer B. Both the DNA regions containing α1-AT oligomer B and TTR oligomer 1 have been found to bind by a footprint assay (*21*) in the same regions where we identified sites of methylation protection (*13, 14*). Thus both these sites do bind and function through the AP-1 protein (Fig. 1B). Computer searches for similarity in the regions of TTR oligomers 2 and 3 and α1-AT oligomers A and C of α1-AT revealed several weak homologies. The sequence alignment that gave the maximal overlap within the regions that had earlier (*13, 14*) been shown by chemical tests (methylation protection and methylation interference) to interact with protein is shown in Fig. 1. An interrupted consensus (two bases, skip one base, five bases) was present: five or six bases of the seven conserved sites matched in all four binding sites.

To test whether this alignment might indeed indicate important bases for protein binding, two separate mutants in TTR oligonucleotide 3 were constructed: oligo 3(2×) had the most highly conserved A and T changed to C and A and oligo 3(4×) had the conserved ACTC changed to CGAA (Fig. 1, nucleotides indicated by dots). The remaining 27 or 25 bases in the oligonucleotide remained unchanged. Both the mutant oligonucleotides were then tested in separate competitions with oligomers A and C of α1-AT or 2 and 3 of TTR (Fig. 4). The results strongly indicate that the consensus sequence shown in Fig. 1 must have some functional significance for protein binding at least for TTR oligomer 3. The mutant oligo 3(4×) was completely unable to compete at a hundredfold molar excess for band formation with any of the four labeled, interrelated oligonucleotides. The mutant oligo 3(2×) with only two changes out of 30 bp had a greatly reduced capacity to compete for band formation in the gel-shift assay with each one of the four labeled oligonucleotides. A hundredfold molar excess of oligo 3(2×) compared to the labeled probe produced less competition than a tenfold excess of unlabeled competitor of the wild-type sequence. Thus we propose that the consensus sequence shown in Fig. 1 (TCNTACTC) is at least part of the binding site for a liver-specific enhancer factor. This liver-specific factor, which binds at least four sites in these two unrelated genes, seems very likely to be involved in mediating the activity of these two cell-specific enhancers.

Different immunoglobulin genes are thought to share similar enhancer factors but those genes are, of course, ancestrally related (*22*). We believe this is the first known case of a single factor or factors acting in a cell-specific fashion on apparently unrelated but coordinately expressed genes of mammals (*23*).

REFERENCES AND NOTES

1. S. McKnight and R. Tjian, *Cell* **46**, 795 (1986).
2. M. Sawadogo and R. G. Roeder, *ibid.* **43**, 165 (1985).
3. P. F. Johnson, W. H. Landschulz, B. J. Graves, S. McKnight, *Genes Dev.* **1**, 133 (1987).
4. H. L. Sive, N. Heintz, R. G. Roeder, *Mol. Cell. Biol.* **6**, 3329 (1986).
5. L. Dailey, S. M. Hanly, R. G. Roeder, N. Heintz, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7241 (1986).
6. J. A. Kreidberg and T. J. Kelly, *Mol. Cell. Biol.* **6**, 2903 (1986).
7. R. I. Richards, A. Heguy, M. Karin, *Cell* **37**, 263 (1984).
8. W. D. Morgan *et al.*, *Mol. Cell. Biol.* **7**, 1129 (1987).
9. E. Derman *et al.*, *Cell* **23**, 731 (1981).
10. D. F. Powell, J. M. Freidman, A. J. Oulette, K. S. Krauter, J. E. Darnell, *J. Mol. Biol.* **179**, 21 (1984).
11. R. H. Costa, E. Lai, J. E. Darnell, *Mol. Cell. Biol.* **6**, 4697 (1986).
12. L. E. Babiss, R. H. Herbst, A. L. Bennet, J. E. Darnell, *Genes Dev.* **1**, 256 (1987).
13. R. H. Costa, E. Lai, D. R. Grayson, J. E. Darnell, *Mol. Cell. Biol.* **8**, 81 (1988).
14. D. R. Grayson, R. H. Costa, K. G. Xanthopoulos, J. E. Darnell, *ibid.*, in press.
15. A. Raz and D. S. Goodman, *J. Biol. Chem.* **244**, 3230 (1969).
16. M. Laskowski and I. Kato, *Annu. Rev. Biochem.* **49**, 593 (1980).
17. M. Freid and D. M. Crothers, *Nucleic Acids Res.* **9**, 6505 (1981).
18. M. M. Garner and A. Revzin, *ibid.*, p. 3047.
19. W. Lee, P. Mitchell, R. Tjian, *Cell* **49**, 741 (1987).
20. P. Angel *et al.*, *ibid.*, p. 729.
21. D. Bohmann *et al.*, *Science* **238**, 1386 (1987).
22. M. Lenardo, J. W. Pierce, D. Baltimore, *ibid.* **236**, 1573 (1987).
23. During the review of this report we learned from G. Crabtree of his results indicating the possibility of binding site in the α- and β-fibrinogen regulatory region that also binds a liver-specific factor. The sequence bound in those two genes is also present in other genes expressed in hepatocytes but is different from the binding site we describe. [G. Courtois, J. G. Morgan, L. A. Campbell, G. Fourel, G. R. Crabtree, *Science* **238**, 688 (1987)].

24. Supported by PHS grants CA 160006-14 and CA 18213-11 from NIH, by grant NP-624Q from the American Cancer Society, by National Institute of General Medical Sciences postdoctoral fellowships GM 1066-02 (D.R.G.) and GM 11947 (R.H.C.), and by a postdoctoral fellowship from the Swedish Research Council (B-PD 8826-100) (K.G.X.). We thank C. Chorny for technical assistance, E. Paulson for his critical review of the manuscript, and A. English for the preparation of this manuscript.

8 October 1987; accepted 18 December 1987