A Regulatory Element in the ApoCIII Promoter That Directs Hepatic Specific Transcription Binds to Proteins in Expressing and Nonexpressing Cell Types

(Received for publication, March 23, 1989)

Todd Leff, Karen Reue, Agustin Melian, Harry Culver, and Jan L. Breslow
From the Laboratory of Biochemical Genetics and Metabolism, The Rockefeller University, New York, New York 10021

To better understand the mechanisms that determine cell type-specific gene expression, we have examined the transcriptional activity of a 13-nucleotide long sequence element, designated C3P, located in the promoter of the apoCIII gene. We demonstrate that this element is required for high levels of apoCIII gene expression in hepatic cells and is sufficient to determine hepatic specific expression when introduced into a heterologous promoter. A protein was identified in hepatic cell nuclear extracts, designated AF-1, that binds to this sequence and is presumably responsible for its transcriptional activity in hepatic cells. Even though the C3P element is not active in HeLa cells, a protein with C3P binding specificity was identified in HeLa cell nuclear extracts. While the HeLa protein is similar to the hepatic AF-1 in its binding specificity and relative abundance, it has approximately twice the molecular weight of the hepatic protein, indicating that they are different proteins or different forms of the same protein. A variety of murine tissue types, including those that do not express the apoCIII gene, were found to contain C3P binding proteins. We conclude that the cell type-specific activity of the C3P element is not due to the absence of C3P binding proteins in nonexpressing cells but is the result of qualitative differences in C3P binding proteins in different cell types.

The mechanisms that determine tissue-specific patterns of gene expression are not well understood. The regulation of hepatic specific transcription of the apolipoprotein CIII (apoCIII) gene provides a model for the examination of these mechanisms. The apoCIII gene belongs to the apolipoprotein family. These genes code for the major protein components of the lipoprotein transport system that distributes cholesterol and triglycerides throughout the body (reviewed in Refs. 1-3). In vivo, the apoCIII gene is expressed primarily in the liver and to a small extent in the intestine (reviewed in Ref. 1). This cell type-specific pattern of expression is preserved in tissue culture cell lines, where transfected copies of the gene are expressed in HepG2 (hepatic) cells but not in HeLa (epithelial) cells (4). We have recently shown that the apoCIII gene promoter contains at least three regions that together determine transcriptional activity in transfected HepG2 cells (4). The proximal positive element contains a sequence motif (CAGGTGACCTTTG) that is found in the promoters of several apolipoprotein genes (4).

In the current report, we demonstrate that this sequence motif, designated C3P, is essential for high levels of transcription in HepG2 cells and is sufficient to determine hepatic specific expression when inserted into a heterologous promoter. We have also identified a protein in hepatic cells, designated AF-1 (apolipoprotein factor), that binds to the C3P element and is presumably responsible for its activity in hepatic cell types. Surprisingly, a protein was observed in HeLa cell nuclear extracts with the same binding specificity. While the HeLa and hepatic proteins show the same DNA binding specificity and appear to be present in equivalent amounts, gel filtration chromatography demonstrated that the two proteins have distinct molecular weights, indicating that they are different proteins or different forms of the same protein. These results suggest that the differential activity of the C3P element in these two cell types is caused by qualitative differences in their respective C3P binding proteins. Nuclear extracts prepared from a variety of mouse tissues, including liver, were found to contain a C3P binding protein of similar size and relative abundance. We conclude that the cell type-specific activity of the C3P element is due to qualitative differences in the C3P binding proteins in different cell types rather than their absence in nonexpressing cells.

EXPERIMENTAL PROCEDURES

The construction ~821WT contains the apoCIII promoter sequences from ~821 to +22 inserted into the CAT expression vector pX76 (see Ref. 4 for construction). The mutant apoCIII template ~821Xh was constructed by the insertion of a XhoI linker into a repaired BamHI site at ~84 in the wild-type apoCIII promoter. This resulted in the loss of 5 base pairs from ~82 to ~78 and their replacement by the 5-base pair XhoI linker. The vector pXT (see Fig. 3) is derived from the CAT expression vector pCT (5) and contains the adenovirus major late promoter sequences from ~50 to +33 (stippled box) linked directly to 60 nucleotides of the SV40 early region noncoding leader sequence (SV40 sequences 5235 to 5175) (stippled box). The heterologous promoter plasmids pX76, pX76, pX76, and pX76, pX76 were constructed by inserting a double-stranded oligonucleotide (oligo-WT, see below) in both orientations 60 nucleotides upstream from the start of major late transcription. Plasmids were

1. A. Melian, K. Reue, J. L. Breslow, and T. Leff, submitted for publication.

* This research was supported in part by a grant from the American Heart Association, New York City affiliate (to T. L.), an Eli Lilly Life Sciences Contact grant (to T. L.), Grant HL53714 from the National Institutes of Health (to J. L. B.), and by general support from the Pew Trust. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Investigator of the American Heart Association, New York City affiliate. To whom correspondence should be addressed: Rockefeller University, Box H29, 1230 York Ave., New York, NY 10021.
‡ Supported by a participating laboratory fellowship from the American Heart Association, New York City affiliate. Present address: Lipid Research Laboratory, V.A. Wadsworth Med. Center, Bldg. 115, Rm. 312, Los Angeles, CA 90024.
§ Fellow of the Stanley J. Sarnoff Society for Cardiovascular Research.
‖ Established Investigator of the American Heart Association.
prepared by double banding in cesium chloride and introduced into cells by the calcium phosphate precipitation method as described (4).

To avoid possible interactions between apoCIII regulatory elements and other transcriptional regulatory elements, co-transfection of at least two different plasmid preparations. CAT activity was determined by the method of Gorman et al. (6).

Nuclear extracts were prepared from tissue culture cells according to Dignam et al. (7) and from tissue essentially as described by Gorski et al. (8). Human liver tissue was obtained from the Liver Tissue Procurement and Distribution System at the University of Minnesota.

DNA binding reactions for the gel mobility shift assay were carried out in a total volume of 10 ml containing 60 mM KCl, 20 mM Hepes (pH 7.9), 4% Ficoll, 1 mM MgCl₂, and 1 μg of poly(dI-dC) (Pharmacia LKB Biotechnology Inc.). In addition, each reaction contained 10,000 cpm of the wild-type and mutant templates described above were transfected into HepG2 and HeLa cells. Neither the wild-type nor mutant constructions were expressed in HeLa cells. This observation that the C3P element is active in hepatic but not in HeLa cells is shown. Constructions containing wild-type or mutant apoCIII promoter sequences (~821 to +22, with respect to the start site of apoCIII transcription) cloned into the CAT expression vector pKT (4), were transfected into HepG2 and HeLa cells. CAT activity is presented relative to the wild-type construction and represents the averages of six experiments. Open boxes indicate previously described (4) and negative transcriptional elements in the apoCIII promoter. Boxed sequences represent nucleotides that differ between the wild-type and mutant templates.

**FIG. 1.** The C3P element is required for hepatic expression of the apoCIII gene. Transcriptional activity of wild-type (WT) and Xh mutant (Xh) apoCIII promoters in transfected HepG2 and HeLa cells is shown. Constructions containing wild-type or mutant apoCIII promoter sequences (~821 to +22, with respect to the start site of apoCIII transcription) cloned into the CAT expression vector pKT (4), were transfected into HepG2 and HeLa cells. CAT activity is presented relative to the wild-type construction and represents the averages of six experiments. Open boxes indicate previously described (4) positive and negative transcriptional elements in the apoCIII promoter. Boxed sequences represent nucleotides that differ between the wild-type and mutant templates.

To characterize the role of the C3P element in determining transcriptional activity of the apoCIII gene, a small mutation in this element was introduced that replaced the five nucleotides between ~84 and ~78 with an 8-nucleotide XhoI linker (referred to as the Xh mutant). The CAT expression vector pKT containing wild-type or mutant apoCIII promoter sequences from ~821 to +22 was transfected into HepG2 and HeLa cells. The Xh mutation caused an 8-fold decrease in the transcriptional activity of the apoCIII promoter in HepG2 cells (Fig. 1), indicating that C3P is required for normal hepatic expression of the apoCIII gene. Neither the wild-type nor mutant constructions were expressed in HeLa cells.

To identify the proteins that interact with the C3P element, the wild-type and mutant templates described above were analyzed by the gel mobility shift and DNase I footprinting assays. A wild-type fragment of the apoCIII promoter (~110 to +22) shows a single major shifted protein band on a mobility shift gel which does not appear when a template containing the Xh mutation is used (Fig. 2A, compare lanes 1 and 2). These results demonstrate that the binding of a protein, designated AF-1 (apolipoprotein factor), correlates with the transcriptional activity of the C3P element. This protein is probably the trans-acting factor that determines the transcriptional activity of the C3P element. The binding of AF-1 to the wild-type template can be competed by the addition of an unlabeled oligonucleotide representing the C3P element but not by the addition of a similar oligonucleotide containing the Xh mutation (Fig. 2A, compare lanes 3 and 4). In addition, an apoCIII template extending to ~82 bound the protein, while a template extending to only ~77 did not (Fig. 2A, lanes 5 and 6). These results indicate that the sequence between ~82 and ~77, which includes half of the C3P sequence, is required for AF-1 binding.

To precisely map the regions of the apoCIII promoter that interact with DNA binding proteins, we carried out DNase I footprinting analysis of a fragment of the apoCIII promoter with proteins from HepG2 nuclear extracts. The results of these experiments indicated that the region between positions ~12 and ~71 is protected by a DNA binding protein (Fig. 2B). This protection can be competed by the wild-type oligonucleotide (Fig. 2B, lanes 2 and 4) and does not occur on a template containing the Xh mutation (Fig. 2B, lanes 9–11), indicating that the protein detected in the gel mobility shift assay is the same protein responsible for the ~86 to ~71 footprint.

The inactivity of the apoCIII promoter in HeLa cells makes it difficult to determine whether the C3P element has any effect, positive or negative, on transcription in this cell line. To address this question and to directly evaluate the cell type specificity of the C3P element, we analyzed its effect on transcription in a heterologous promotor construct that is expressed equally well in HepG2 and HeLa cells. An oligonucleotide representing the C3P sequence was inserted 60 nucleotides upstream from the transcriptional start site of the adenovirus major late promotor in the CAT expression vector pXT. The presence of this element caused a 6-fold stimulation of pXT expression in HepG2 cells but had no significant effect on expression in HeLa cells (Fig. 3). These results confirm that the C3P sequence is a hepatic specific transcriptional regulatory element.

Sequence elements similar to C3P appear in the promoters of several apolipoprotein genes (4, 10), including apoB. The observation that the C3P element is active in hepatic but not HeLa cells (Fig. 3) is supported by our recent results demon-
The C3P element is sufficient to determine hepatic specific transcriptional activity. Constructions containing the C3P element inserted in the forward (pXT.WT) or reverse (pXT.WTr) orientations upstream of the adenovirus major late promoter in the CAT expression vector pXT were transfected into HepG2 and HeLa cells. Transcriptional activity is presented relative to the activity of the pXT construction (no insert) in each cell type and represents the average of six experiments.

The transcriptional specificity of the C3P element must be due to differences in the protein factors that interact with this sequence in hepatic and nonhepatic cells. One possibility is that HeLa cells do not contain a C3P binding protein. To test this hypothesis DNase I footprinting was carried out with HeLa cell nuclear extracts. Surprisingly, a protein is present in HeLa cells that gives a qualitatively identical footprint to that observed with either HepG2 or mouse liver nuclear extracts (Fig. 4A). The presence of a protein in HeLa cells that specifically interacts with the C3P element was confirmed by the mobility shift assay using oligonucleotides representing the wild-type and Xh mutant sequences as templates (Fig. 4B, lanes 1–6). These results, combined with the activity measurement (Fig. 1), suggest that while proteins in both HeLa and HepG2 cells interact with the C3P element, only the hepatic protein can activate transcription.

To further characterize the C3P binding proteins in these two cell types, nuclear extracts were chromatographed on a Superose 12 gel filtration column (Fig. 5A) and the molecular weight of the C3P binding activity determined by comparison with the elution of standard proteins (Fig. 5B). The C3P binding activity present in the HeLa extract eluted at a molecular mass of 240 ± 48 kDa, while the HepG2 activity eluted at 133 ± 28 kDa (Fig. 5 and Table I). HepG2 cells also contained a minor species with an apparent molecular mass of 240 ± 48 kDa (visible in the gel mobility shift pattern shown in the center panel marked HepG2, Fig. 5). This was not peculiar to HepG2 cells, as nuclear extracts of human liver tissue showed the same pattern (data not shown). These results suggest that HepG2 and HeLa cells contain different C3P binding proteins or substantially different forms of the same protein. The relative concentrations of C3P binding activity in HeLa and HepG2 cells (as measured by the gel mobility shift assay) were found to be roughly equivalent (Table I). Together, these results suggest that the differential activity of the C3P element is due to qualitative rather than quantitative differences in the C3P binding proteins found in these two cell types.

The C3P binding proteins found in HepG2 and HeLa cells could represent products of distinct genes that share DNA sequence specificity for binding but have different transcriptional activities. It is also possible, since the molecular weight of the HeLa protein is about twice that of the HepG2 protein,
FIG. 4. C3P binding proteins are present in a variety of cell and tissue types. A, DNase I footprinting analysis of C3P binding proteins in hepatic and HeLa cell nuclear extracts. DNA templates isolated from wild-type apoCIII promoter and labeled as in Fig. 2 were incubated with mouse liver (Liu), HeLa cell (HeLa), or HepG2 (HpG2) nuclear extracts as indicated and analyzed by DNase I footprinting as described under "Experimental Procedures." Lane 1 is a G+A sequencing ladder of the same template. Boxed nucleotides represent sequences protected from DNase activity in each of the three extracts. B, gel mobility shift assay of C3P binding activity in nuclear extracts prepared from a variety of cell and tissue types. Labeled DNA templates (0.2 ng) were either oligo-WT (WT) or oligo-Xh (Xh) as indicated. Extracts were prepared from HeLa cells (HeLa), mouse liver (Liu), HepG2 (HpG2), rat hepatoma cells RH7777 (7777), mouse spleen tissue (Splen), mouse kidney tissue (Kid), human colonic cell line CaCo-2 (CaCo), mouse proximal intestine tissue (Int Prx), and mouse distal intestine tissue (Int Dst).

that the same protein is present in both cell types but appears only as a dimer in HeLa cells. In this case, the 240-kDa species seen in the HepG2 extract (Fig. 5A, HepG2 gel mobility shift panel) would suggest that a fraction of AF-1 in HepG2 cells is also present as a dimer. The HeLa and hepatic proteins could be products of the same gene but differentially modified to influence their ability to form dimers and their transcriptional activity. These modifications could include phosphorylation or glycosylation, both of which have been shown to occur on transcription factors (for examples see Refs. 11 and 12). Finally, the size differences could be artifacts of extract preparation, and the active forms of the protein in HeLa and HepG2 cells might be identical. In this case, their activities could be influenced by their interaction with proteins unique to that cell type. Purification and characterization of C3P binding proteins from HepG2 and HeLa cells will resolve these different possibilities.

To determine if C3P binding activity is present in other tissues, nuclear extracts were prepared from a variety of cell and tissue types and assayed by the gel mobility shift assay using wild-type and Xh mutant templates. The cell and tissue types tested include those that express the apoCIII gene (liver and intestine) and those that do not (spleen and kidney). The results shown in Fig. 4B demonstrate that every extract tested contained a protein that bound to the wild-type but not to the mutant sequence. In addition, the relative quantity of C3P binding activity in each extract was found to be roughly equivalent (Table I). The C3P binding activity in each of these extracts was further analyzed by DNase I footprinting on the CIII promoter template. Each extract protected the...
same region of the template as the liver AF-1 protein (Fig. 6), suggesting that each tissue contains an equivalent C3P binding activity.

As in the comparison between HepG2 and HeLa cells, the differences in mobility of the C3P binding activity-DNA complexes in the various mouse tissue extracts suggest qualitative differences in their DNA binding proteins. However, in contrast to the HepG2 and HeLa observations, the molecular weights of the C3P binding proteins in each mouse tissue, as determined by the gel filtration chromatography, were similar, between 110,000 and 125,000 (Table I). The significance of the size difference between the human and mouse activities and between expressing and nonexpressing cell types is not understood and can only be investigated by purification and characterization of the proteins involved.

The presence of proteins in different cell types that share DNA binding specificity but differ in transcriptional activity has been observed in other systems. The trans-acting proteins OTF-1 and OTF-2 bind to the same octamer recognition sequence but are distinct proteins with different molecular weights and different activities (13-15). OTF-2 is found only in lymphocytes and activates transcription of some immunoglobulin genes by binding to the octamer sequence in their promoters. OTF-1 is found in a variety of other cell types and while it recognizes the same octamer sequence it does not activate transcription of immunoglobulin genes (13, 16).

The apoCIII gene is a member of the apolipoprotein gene family (3), the products of which are important components of the lipoprotein system that transports triglycerides and cholesterol throughout the body and regulates their levels in the serum (1, 2). Four members of this gene family (including apoCIII) contain C3P sequence homologies in their promoter regions (Table II). We have recently determined that AF-1 interacts with these apolipoprotein genes, and we believe that AF-1 is a major determinant of apolipoprotein gene expression. A search of the GenBank for homologies to the C3P sequence element (CAGGTGACCTTTG) revealed homologies in the promoters of several mammalian genes in addition to the apolipoprotein genes (Table II). The homologies identified in the transthyretin and α1-antitrypsin genes are notable in that they reside in transcriptionally active regions of the promoter (17, 18). Cross-competition experiments with an oligonucleotide containing the transthyretin homology indicated that AF-1 can bind to these sequences. A protein has

5 T. Leff and R. Costa, unpublished observations.
been identified from rat liver that interacts with the α-
antitrypsin homology (12), but its relationship to AF-1 has
not been determined. Other well characterized genes
expressed in the liver, including albumin and α-fetoprotein, do
not contain similar sequences (19–23), suggesting that AF-1
does not interact with these promoters. In addition, oligonu-
cleotides containing the binding site sequences of the tran-
scription factors HNF-1, C/EBP, and HNF-3 (17) did not
compete for AF-1 binding sites (data not shown). These
results indicate that AF-1 is a distinct protein from these
hepatic transcription factors.

Although the transthyretin and α1-antitrypsin genes are
expressed in the liver as are the apolipoprotein genes, C3P
sequence similarities were found in genes that are expressed
in a wide variety of tissues (Table II), including pituitary
(oxytocin-neurophysin), muscle (myoglobin), parotid (parotid
secretory protein), and intestine (fatty acid binding protein).
Although the significance of these sequence similarities is
difficult to evaluate, they suggest that AF-1 may have a
transcriptional role in a variety of tissue types. Purification
type-specific transcriptional activity and its role in regulating
brain Duda for help in preparing the manuscript.

and characterization of C3P binding proteins from HepG2
and HeLa cells from different tissues will clarify its cell
type-specific transcriptional activity and its role in regulating
tissue-specific expression of the apolipoprotein genes.

Acknowledgments—We would like to thank Dr. Jonathan Smith
for providing the pXT construction and Alexandra McNear and
Lorrain Duda for help in preparing the manuscript.

REFERENCES
1. Breslow, J. L. (1988) Physiol. Rev. 68, 85–132
2. Mahley, R. W., Innerarity, T. L., Rall, S. C., Jr., and Weisgraber,
K. H. (1984) J. Lipid Res. 25, 1277–1294
3. Li, W. H., Tanimura, M., Luo, C. C., Datta, S., and Chan, L.
(1988) J. Lipid Res. 29, 245–271
4. Reue, K., Leff, T., and Breslow, J. L. (1988) J. Biol. Chem. 263,
6857–6864
5. Smith, J. D., Melián, A., Leff, T., and Breslow, J. L. (1988) J.
Biol. Chem. 263, 8300–8308
6. Gorman, C. M., Moffat, L. F., and Howard, B. H. (1982) Mol.
Cell. Biol. 2, 1044–1051
7. Dignam, J. D., Martin, P. L., Shastry, B. S., and Roeder, R. G.
(1983) Methods Enzymol. 101, 582–588
8. Gorski, K., Carneiro, M., and Schibler, U. (1986) Cell 47, 767–
776
9. Maxam, A. M., and Gilbert, W. (1980) Methods Enzymol. 65,
499–560
10. Das, H. K., Leff, T., and Breslow, J. L. (1988) J. Biol. Chem.
263, 11452–11458
11. Jackson, S. P., and Tjian, R. (1988) Cell 55, 125–133
12. Yamamoto, K. K., Gonzalez, G. A., Biggs, W. H., III, and Mont-
moy, M. R. (1988) Nature 334, 494–498
13. Scheidereit, C., Heguy, A., and Roeder, R. G. (1987) Cell 51,
783–793
14. Fletcher, C., Heintz, N., and Roeder, R. G. (1987) Cell 51,
773–781
15. Staudt, L. M., Singh, H., Sen, R., Wirth, T., Sharp, P. A., and
Baltimore, D. (1986) Nature 323, 640–643
16. Mizushima-Sugano, J., and Roeder, R. G. (1986) Proc. Natl.
Acad. Sci. U. S. A. 83, 8511–8515
17. Costa, R. H., Lai, E., Grayson, D. R., and Darnell, J. E., Jr.
(1988) Mol. Cell. Biol. 8, 81–90
18. Monaci, P., Nicosia, A., and Cortese, R. (1988) EMBO J. 7,
2075–2087
19. Courtiois, G., Morgan, J. G., Campbell, L. A., Foureul, G., and
Crabtree, G. R. (1987) Science 235, 688–692
20. Cereghini, S., Raymondjean, M., Carrance, A. G., Herbomel, P.,
and Yaniv, M. (1987) Cell 50, 627–638
21. Lichtsteiner, S., Wuarin, J., and Schibler, U. (1987) Cell 51,
963–973
22. Godbout, R., Ingram, R. S., and Tilghman, S. M. (1988) Mol.
Cell. Biol. 8, 1169–1178
23. Nahon, J.-L., Danan, J.-L., Poiret, M., Tratiner, I., Jose-Estanyol,
M., and Sala-Trepat, J.-M. (1987) J. Biol. Chem. 262, 12479–
12487
24. Karanthanasis, S. K., Zannis, V. I., and Breslow, J. L. (1983)
Proc. Natl. Acad. Sci. U. S. A. 80, 6147–6151
25. Das, H. K., Leff, T., and Breslow, J. L. (1988) J. Biol. Chem.
263, 11452–11458
26. Akaboshi, E. (1985) Gene (Amst.) 33, 241–249
27. Jenne, D., and Stanley, K. (1987) Biochemistry 26, 6735–6742
28. Sweater, D. A., Birkenmeier, E. H., Klisak, I. J., Zollman, S.,
Sparks, R. S., Mohandas, T., Lusis, A. J., and Gordon, J. I.
(1987) J. Biol. Chem. 262, 16060–16071
29. Sausville, E., Carney, D., and Battey, J. (1986) J. Biol.
Chem. 260, 10236–10241
30. Fowlkes, D. M., Mullis, N. T., Comeau, C. M., and Crabtree, G.
R. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 2313–2316
31. Shaw, P., and Schibler, U. (1986) J. Mol. Biol. 192, 567–576
32. Osumi, T., Ishii, N., Miyazawa, S., and Hashimoto, T. (1987) J.
Biol. Chem. 262, 8138–8143

Hepatic Specific Gene Expression 16137