Title

The identification of transcriptional regulatory regions of the human apolipoprotein E gene and the characterization of an apo-E gene enhancer element and its associated factors

Author

Chang, David J

Publication Date

1989

Peer reviewed|Thesis/dissertation

The Identification of Transcriptional Regulatory Regions of the Human Apolipoprotein E Gene and the Characterization of an Apo-E Gene Enhancer Element and Its Associated Factors

by

David J. Chang

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Physiology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

ACKNOWLEDGMENTS

I would like to give my deepest thanks to my thesis advisor, Dr. John M. Taylor for his guidance and support, not only during my graduate studies but also for the entire period I have been at the Gladstone Foundation Laboratories. I would also like to thank the members of my dissertation committee, Dr. Keith R. Yamamoto and Dr. Richard M. Myers for spending the time reviewing the dissertation and for helpful advice. Special thanks are given to Dr. Young-Ki Paik, for his collaboration and friendship throughout the project. I am grateful to the members of the Taylor lab, both past and present, Dr. Stephen J. Lauer, Dr. Trond P. Leren, Dr. Catherine A. Reardon, David A. Walker, Dr. Geoff J. Howlett, Ellen D. Taxman, Bahareh Sahami, and Dr. Nabil A. Elshourbagy for technical support and many helpful discussions. I also thank Dr. Brian J. McCarthy and Dr. Beatriz Levy and members of their laboratories for their advice and encouragement. I am indebted to the scientific, administrative, and clerical staffs at the Gladstone Foundation Laboratories for all their help and support. I thank the UCSF Department of Physiology for giving me the opportunity to work for a doctorate degree. Most of all, I thank my parents, my sisters and their families, and my good friends at church for all the support and encouragement they have given me during this long and often difficult period. Finally, I thank the Lord for His constant guidance and neverending care.

The Identification of Transcriptional Regulatory Regions of the Human Apolipoprotein E Gene and the Characterization of an Apo-E Gene Enhancer Element and Its Associated Factors

David J. Chang

ABSTRACT

Apolipoprotein (Apo-) E, a 34 kD protein which is associated with all major classes of lipoproteins, is produced in a wide variety of tissues. The purpose of this study was to examine the regions that regulate the transcription of the human apo-E gene. To define DNA elements within the apo-E gene and its 5' flanking sequence that affect promoter activity or contain enhancer activity, apo-E gene fragments were fused to the bacterial reporter gene, chloramphenicol acetyltransferase (CAT). The chimeric apo-E gene/CAT gene recombinants were analyzed by transient transfections in cultured cell lines or by *in vitro* transcription assays using HeLa cell nuclear extracts. Nuclear extracts prepared from various cell lines also were employed for DNase I footprinting assays and gel retention assays to detect the binding of specific protein factors to the regulatory elements.

The proximal 383 nucleotides of the 5' flanking sequence permitted the transient expression of the apo-E gene in several cell lines. This region contained two elements with enhancer-like activity, located at nucleotides -193 to -124 (URE1) and at nucleotides -366 to -246 (URE2). A third enhancer element was identified in the first intron, within nucleotides +44 to +262 (IRE1). At least five protein-binding regions were detected in the 5' flanking sequence and in the first intron. Three of these sequences are recognized by the transcription factor, Sp1. They are located at nucleotides -59 to -45, where a GC box element resides, and at nucleotides -161 to -141 and -184 to -173, which are within the enhancer element URE1. The binding of Sp1 to the -161 to -141 sequence was found to be critical for the manifestaton of URE1 enhancer

activity. Furthermore, a 30-bp oligonucleotide which spans this protein-binding sequence was, by itself, able to act as an enhancer. The sequence, termed PET, was also recognized by a 55 kD protein, purified in this study, which may affect its transcriptional activity. Thus the control of apo-E gene expression is the result of complex interactions of multiple regulatory elements and transcription factors.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS iii

ABSTRACT iv

TABLE OF CONTENTS vi

LIST OF TABLES viii

LIST OF FIGURES ix

I. INTRODUCTION 1
- A. Lipoproteins : Structures and Classes 1
- B. Apolipoprotein E : Structure and Function 3
- C. The Molecular Biology of Apolipoprotein E 5
- D. Eukaryotic Gene Expression 6
- E. Rationale 9

II. METHODS AND MATERIALS 10
- A. Construction of Recombinant Plasmids 10
  - 1. Analysis of the human apolipoprotein E promoter 10
  - 2. Analysis of the enhancer activity of the human apo-E gene fragments 11
- B. Transient Gene Transfer of Recombinant Plasmids 11
- C. Maintenance of Tissue Culture Cells 12
- D. Chloramphenicol Acetyltransferase and β-Galactosidase Assays 13
- E. Preparation and Analysis of RNA 13
  - 1. Isolation of total cellular RNA 13
  - 2. Dot blot analysis 14
  - 3. Primer extension analysis 14
- F. Isolation and Preparation of Nuclear Extracts 15
- G. DNase I Footprint Assay 16
  - 1. Preparation of labelled fragments 16
  - 2. Footprinting assay 16
- H. Gel Retention Assay 17
- I. *In Vitro* Transcription Assay 18
- J. Synthesis of Oligonucleotides 18
- K. Site-Directed Mutagenesis 19
- L. Purification of Nuclear Protein Factors 20
  - 1. Gel filtration chromatography 20
  - 2. DNA affinity column chromatography 20
  - 3. SDS polyacrylamide gel electrophoresis 22
- M. Materials 22

III. RESULTS 24
- A. Identification and Characterization of Transcriptional Regulatory Elements 24

1. Expression of Apo-E/CAT hybrid genes in cultured cells 24
2. Identification of 5' flanking sequences necessary for gene expression 24
3. Identification of apo-E gene enhancer elements 30
4. Characterization of apo-E gene enhancer elements 36
5. Characterization of URE1 enhancer function 43

B. Detection of Protein Factors Binding to Upstream Regulatory Sequences 47
1. DNase I footprint analysis of upstream sequences 47
2. Cell-type specificity of the DNA-binding proteins 52
3. DNase I footprint analysis of URE1 deletion mutants 57

C. Identification and Characterization of the PET-Binding Proteins 62
1. Gel retention analysis of PET 62
2. Identity of the PET-binding proteins 70
3. Isolation of the PET-binding proteins 89

D. Functional Analysis of Regulatory Elements by *In Vitro* Transcription 94
1. Template and competition studies 94
2. Deletion and substitution studies 99
3. Studies with the Δ GC box template 104

E. Detection of Protein Binding in IRE1 109

IV. DISCUSSION 117

A. Multiple Regulatory Elements Control Human Apo-E Gene Expression 117

B. PET-Binding Proteins and Enhancer Activity 124

C. Cell-Specificity and Inducibility of Apo-E Gene Expression 128

V. REFERENCES 132

LIST OF TABLES

Table		page
I	Relative chloramphenicol acetyltransferase (CAT) activities directed by the 5' flanking sequences of the human apo-E gene in different cultured cell lines.....	27
II	Enhancer activity of apo-E gene fragments in different cultured cell lines.....	44
III	Sp1 binding sequence of PET and the substituted nucleotides of the mutant PET oligonucleotides.....	75
IV	Enhancer activity of selected oligonucleotides in CHO cells.....	88

LIST OF FIGURES

Figure		page
1	Construction of apo-E promoter/CAT expression recombinants	26
2	Promoter activity of deletion fragments of the 5' flanking region of the apo-E gene	29
3	Construction of apo-E enhancer test recombinants	32
4	Enhancement of the apo-E gene fragments on TK promoter-directed CAT expression (pTK1)	35
5	Dot blot analysis of apo-E enhancer/TK1 recombinants	38
6	Primer extension analysis of transcripts generated by apo-E enhancer/TK1 constructs	40
7	Effect of position on apo-E enhancer element activity in CHO cells	42
8	Analysis of site-directed deletion constructs of URE1	46
9	Enhancer activity of the PET sequence	49
10	DNase I footprint analysis of apo-E gene fragments by HepG2 cell nuclear proteins	51
11	DNase I footprint analysis of apo-E gene fragments by CHO cell nuclear proteins	54
12	DNase I footprint analysis of the proximal 5' flanking sequence of the apo-E gene by nuclear proteins from various cell lines	56
13	Gel retention analysis of -103 to -83 oligonucleotide	59
14	DNase I footprint analysis of URE1 deletion constructs by HeLa nuclear proteins	61
15	Gel retention analysis of PET with various nuclear extracts	64
16	Oligonucleotides used in competition gel retention analysis	67
17	Gel retention analysis of PET with oligonucleotide competitors and HeLa nuclear proteins	69
18	Gel retention analysis of PET in different KCl	

	concentrations	72
19	Gel retention analysis of PET in different poly (dI-dC)/ poly (dI-dC) concentrations	74
20	Gel retention analysis of PET and mutant PET	77
21	Gel retention analysis of PET with pure Sp1	80
22	DNase I footprint analysis of the proximal 5' flanking sequence of the apo-E gene with pure Sp1	82
23	Gel retention analysis of -185 to -169 oligonucleotide	85
24	Characterization of band c protein by gel retention analysis with various nuclear extracts	87
25	Gel retention analysis of PET with various oligonucleotide competitors for the identification of band c protein	91
26	Gel retention analysis with pure AP-2	93
27	Gel retention analysis of PET with Sephacryl S-300 fractions	96
28	Affinity-column purification of 55 kD protein	98
29	*In vitro* transcription analysis of different apo-E promoter/CAT gene templates	101
30	*In vitro* transcription analysis in presence of oligonucleotide competitors	103
31	*In vitro* transcription analysis of deletion and substitution constructs of the -383 apo-E promoter/ CAT gene template	106
32	*In vitro* transcription analysis of Δ GC box template with oligonucleotide competitors	108
33	Reconstitution *in vitro* transcription analysis of the Δ GC box template	111
34	DNase I footprint analysis of IRE1	114
35	Gel retention analysis of +161 to +197 oligonucleotide with various nuclear extracts	116
36	Identification of regulatory elements and protein-binding sequences in the proximal 5' flanking region and in the first intron of the human apo-E gene	119

I. INTRODUCTION

A. Lipoproteins : Structures and Classes

Lipid transport in the circulation is undertaken by large macromolecular complexes known as lipoproteins (reviewed in 1, 2). In these complexes, a core of nonpolar lipids, such as triglycerides and cholesterol esters, is surrounded by a mixed monolayer of polar lipids, such as phospholipids and free cholesterol, and specific proteins known as apolipoproteins. Thus, lipoprotein particles may be viewed as micellar structures, in which the hydrophilic components are arranged at the surface and the hydrophobic components form a core at the center of each particle (3). Lipoproteins are in a state of dynamic equilibrium with each other, as well as with other tissue and membrane components within the body. However, the water-insoluble components require specific transfer factors to move between the lipid surfaces (4). Normal plasma lipoproteins are spherical when visualized under the electron microscope, although some newly synthesized particles may appear more like disks (5). A certain proportion of neutral lipid, either triglycerides, or cholesterol esters which are converted from free cholesterol through the action of the enzyme lecithin:cholesterol acyltransferase (LCAT), seem to be needed for lipoproteins to assume a spherical appearance (6).

Lipoproteins are divided into five major classes which differ by size, charge, and the composition of both lipids and apolipoproteins. These classes are most commonly separated according to their density, as determined by ultracentrifugation in potassium bromide (KBr) solutions (7). The largest and least dense of the lipoprotein particles (0.93 g/ml) are chylomicrons which consist of 98-99% lipid and only 1-2% apolipoprotein in weight. Chylomicrons which are synthesized by the mucosal cells in the small intestine (8) contain

triglyceride as their major lipid constituent and the triglyceride composition reflects the content of dietary fatty acids (9). Their components consist of apolipoprotein (apo-) A-I, apo-A-IV, and apo-B-48. They are secreted by the small intestine into the mesenteric lymph, and then they enter the circulation. In the lymph and the circulating plasma, chylomicrons acquire apo-C-I, apo-C-II, apo-C-III and apo-E on their surface. The core triglycerides are hydrolyzed by the enzyme lipoprotein lipase (10), which is located on the endothelial surface of blood capillaries, with the released free fatty acids taken up by cells. The chylomicron remnant particles then are cleared by an apparent receptor-mediated endocytosis mechanism in the liver (11).

The second class of lipoproteins, the very low density lipoproteins (VLDL, density = 0.93 - 1.006 g/ml), are synthesized by the parenchymal cells in the liver (12), and contain approximately 90% lipid by weight, of which about two-thirds is represented by triglycerides. The triglycerides present in VLDL are those synthesized endogenously in the liver largely during the fasting state (13). The major protein constituents are the apo-C proteins, apo-B-100, and apo-E. During circulation, VLDL are metabolized by lipoprotein lipase, which hydrolyzes the bulk of the triglycerides, forming VLDL remnant particles. VLDL remnants with densities of 1.006 - 1.019 g/ml are referred to as intermediate density lipoproteins (IDL). IDL particles either can be cleared in the liver by receptor-mediated endocytosis, or can be processed with further loss of triglycerides and also of apo-E to form low density lipoproteins (LDL). LDL have densities of 1.019 - 1.063 g/ml and contain cholesterol esters as their major lipid component. Almost all of the protein in LDL consists of apo-B-100, which serves as the ligand for the binding of LDL to the cell-surface LDL receptor (14, 15). The LDL receptor is present in almost all cells and represents the means by which cells obtain exogenous cholesterol for

membrane synthesis, or in certain specialized cells, for bile acid and steroid hormone synthesis (reviewed in 14).

The final class of lipoproteins, the high density lipoproteins (HDL) are heterogeneous in their composition and their densities range from 1.063 - 1.210 g/ml. These particles are approximately one half lipid and one half protein by weight, with phospholipids and apo-AI being the most abundant of their respective components. HDL can be either newly synthesized in the liver and the intestine (16) or formed from the metabolism of chylomicrons (17). In circulation, HDL are able to pick up free cholesterol from peripheral tissues, which is then esterified by the enzyme, LCAT. The cholesterol ester-enriched HDL are then cleared mainly in the liver by interacting with hepatic receptors, thus completing a process known as reverse cholesterol transport (18).

B. Apolipoprotein E : Structure and Function

Apolipoprotein E (reviewed in 19,20) is a protein constituent of chylomicrons, VLDL and a subclass of HDL which participate in cholesterol redistribution among cells. It is a single chain polypeptide of 299 amino acids with a molecular size of 34,000 (21), and is glycosylated at a single site at threonine residue 194 (22). The major physiological role for apo-E in lipoprotein metabolism is its mediation of the cellular uptake of cholesterol by serving as a ligand for the LDL receptor (23, 24). Apo-E also is believed to mediate the uptake of chylomicron remnants via a postulated hepatic receptor mechanism in the liver (25).

In the human population, apo-E is present as three major isoforms, referred to as E2, E3, and E4, which are the products of three alleles at a single gene locus (26). These isoforms are the result of single amino acid substitutions, usually involving cysteine and arginine residues in the primary

structure of apo-E (21). Apolipoprotein E2, which is defective in receptor binding, is the most common form of apo-E associated with type III hyperlipoproteinemia, which can lead to premature atherosclerotic disease (27). Because of this pathological consequence, the receptor binding domain of apo-E has been mapped in detail. Different experimental approaches have all indicated that the basic amino acids, arginine, lysine, and histidine between residues 140 and 160 are important for apo-E to show normal receptor binding (28-30). The basic residues of this region appear to be involved in ionic interactions with the acidic residues of the ligand-binding domain of the LDL receptor (31). The most common E2 variant is a cysteine to arginine substitution at residue 158, which results in an apo-E protein with less than two percent of the normal receptor binding activity (32).

Because it is a constituent of various lipoprotein classes, and because it serves as the ligand for receptors which mediate the cellular uptake of these lipoproteins, apo-E is involved in metabolic pathways that direct the transport of cholesterol and other lipids among cells of different organs or cells within an organ. Apo-E becomes a component of chylomicrons following secretion by the small intestine and is responsible for mediating chylomicron remnant clearance by serving as the ligand for a postulated hepatic receptor (25). VLDL remnants and IDL which contain lipids synthesized in the liver are taken up by the LDL receptors in extrahepatic cells and this process is mediated primarily by apo-E (20). Apo-E also is involved in the transport of cholesterol from peripheral tissues to the liver by a reverse cholesterol transport process that utilizes HDL. Cholesterol-loaded cells such as macrophages can release their cholesterol to HDL particles in the interstitial fluid. As the HDL accumulate cholesterol, they acquire apo-E. The cholesterol-

rich HDL with apo-E are then taken up mainly in the liver by the LDL receptor (33).

Apo-E also can participate in the redistribution of lipid and cholesterol from cells with excess lipid to others which require lipid. During the regeneration of an injured peripheral nerve, apo-E is produced at high levels by resident and recruited macrophages (34). Later in the regeneration process, the macrophages can secrete cholesterol which is taken up by the LDL receptors on the regenerating nerve axons via the mediation of apo-E.

Apo-E has been postulated to have functions unrelated to lipid transport. The high levels of apo-E found in injured or regenerating nerves lead to the proposal that apo-E may be a neurotrophic growth factor (35). Apo-E may play a role in vascular smooth muscle cell growth through its ability to bind heparin and heparin-like glycosaminoglycans, which are potent inhibitors of smooth muscle cell proliferation and motility (36, 37). This arrest in the growth of smooth muscle cells can promote them to differentiate. Finally, apo-E may be involved in both the stimulation and suppression of T-lymphocyte activation and proliferation (38, 39).

C. The Molecular Biology of Apolipoprotein E

The human apo-E gene has been mapped to chromosome 19 (40) and is linked to the apo-CI gene and an apo-CI' pseudogene (41). The apo-E gene is 3597 nucleotides in length and contains four exons separated by three introns (42). The locations of the exon-intron junctions were identified by comparing the gene sequence with the apo-E mRNA sequence which is 1163 bp in length (43). The first intron occurs in the 5'-noncoding region; the second intron occurs in the codon for glycine at position -4 of the signal peptide region; and the third intron occurs in the codon for arginine at position +61 of the mature protein. The primary translation product from the mRNA is 317 amino acids in

length with the 18 amino-terminal amino acids serving as a signal peptide (21). The proximal 5' flanking region of the apo-E gene contains a TATA box element at nucleotide positions -33 to -27. In addition, there are four *Alu*I family repeat sequences associated with the gene, one at each end of the gene and two in the second intron (42).

The expression of the apo-E gene has been detected in a number of different cell types in various organs. The largest quantity of apo-E is produced by the liver, with the hepatic parenchymal cells being primarily responsible for apo-E production (44). Apo-E mRNA concentration in the brain is approximately one-third that of the liver (45) and astrocytes are the cell type responsible for producing apo-E (46, 47). Macrophages have been found to produce large quantities of apo-E (48), although in their precursor monocyte form, the apo-E gene is not expressed (49). Macrophage apo-E mRNA levels can be stimulated by cholesterol loading (50) and suppressed by treatment with bacterial endotoxin (51). Thus, macrophage apo-E gene expression can be influenced by differentiation and by external agents. Growth-arrested smooth muscle cells (52) and ovarian granulosa cells (53) have been shown to produce apo-E. In addition, recent studies have demonstrated that transfection of the human apo-E gene into various cultured mammalian cells, some of which do not express their endogenous apo-E gene, can lead to the synthesis of apo-E mRNA and the production of biologically active apo-E protein (54, 55). Therefore, a wide variety of cells are either expressing or possess the capability to express the apo-E gene.

D. Eukaryotic Gene Expression

The transcriptional regulation of eukaryotic genes has been one of the most intensely studied areas in the field of molecular biology. The genes transcribed by RNA polymerase II (56) are modulated by the interaction of

sequence-specific DNA-binding proteins (57) with *cis*-acting recognition sites (58, 59) located usually upstream of the transcription start site. Using a combination of *in vitro* mutagenesis and DNA-mediated gene transfer studies, two types of *cis*-acting regulatory sequences, promoters and enhancers, have been identified. Promoters are located immediately upstream from the start site of transcription and are typically about 100 bp in length (60). They are required for accurate and efficient initiation of transcription and act in a position dependent fashion. Promoters can be divided into proximal and distal elements. The proximal elements include the transcription cap site and the TATA box element which is located 30 bp upstream of the start site and is involved in fixing the site of initiation (61). The TATA box is recognized and bound by the transcription factor, TFIID (62). The distal elements, also referred to as upstream promoter elements (59), consist of one or more sequence elements usually ranging from 8 to 12 base pairs in length, each of which is recognized by one or more distinct protein factors, which alter the rate of transcription.

Unlike promoters, enhancers can be located hundreds and even thousands of base pairs away from the transcription start site. They can be positioned either 5' or 3' of the start site and can operate in an orientation-independent manner (63). Enhancers do not possess the ability to initiate transcription but they increase the rate of transcription from promoters. Strong viral enhancers such as the SV40 enhancer (59, 63) are able to stimulate the activities of their associated promoters by as much as one hundred fold. Like promoters, enhancers interact with sequence-specific DNA-binding proteins. Recent studies have shown that different sequence motifs within the SV40 enhancer are recognized by distinct protein factors (64-66). The operational difference between enhancers and promoters may be

a consequence of the arrangement and number of transcription factor recognition elements in a particular sequence, rather than some fundamental difference in the mechanism by which these elements act (63). In fact, some of the sequence elements in enhancers and promoters are interchangeable (67), which can make the distinction between these two regulatory sequences difficult.

Some enhancers, such as many viral enhancers, are constitutively active in most cell types, whereas other enhancers can be either tissue-specific (68) or can be induced by various external agents such as growth factors (69) and steroid hormones (70). Inducible enhancers can be activated by a number of different mechanisms, including the activation of a positive regulatory factor or the inactivation of a negative regulatory factor, or both actions (59). This induction effect is understood best for the steroid hormone receptor proteins (70) where the binding of the hormone to the receptor allows the activating protein to bind to the enhancer or responsive element, resulting in the stimulation (or the suppression) of transcription.

In recent years, the molecular cloning of the genes that encode transcription factors has allowed a functional analysis of these proteins. Studies of the primary amino acid sequence have led to the identification of structural motifs at the DNA-binding domain. A common motif found in bacterial repressor proteins and several eukaryotic DNA-binding proteins consists of two α helices bridged by a β turn, and is referred as the "helix-turn-helix" motif (71). Another motif termed the "zinc finger" (72) is characterized by cysteine or histidine residues that form a tetrahedral complex with a zinc ion. This structure, which was originally found in the 5S gene transcription factor, TFIIIA (73), is also found in the glucocorticoid (74) and estrogen (75) receptor proteins, as well as in transcription factor, Sp1 (76). Although it is

not a structural motif of the DNA-binding domain, the "leucine zipper" motif (77) enables the formation of homodimeric or heterodimeric complexes between two proteins, resulting in the juxtaposition of regions required for an association with DNA. This motif can be found in transcription factor, C/EBP (78), as well as the *c-jun*/AP-1 family of protein factors (79, 80).

The domains responsible for the stimulation of transcription (often referred to as the activation domains) have been identified in several transcription factors. In the yeast transcription factors, GAL4 and GCN4, the activation domains consist of regions of amino acids with significant negative charges that are capable of forming amphipathic α-helices (81, 82). In Sp1, four independent activation domains have been identified, including two glutamine-rich segments (83). Although the exact nature of the mechanism by which the transcription factors bring about a stimulation of transcription remains unknown, interactions involving their activation domains with RNA polymerase II itself (84) or with some component at the transcription initiation complex (such as the TATA box-binding TFIID factor) have been postulated (83, 85).

E. Thesis Research Rationale

The key role which apo-E plays in lipid transport and the numerous functions exhibited by this protein have prompted an investigation of the regulation of human apo-E gene expression. The aims of this study are to identify the sequences in the proximal 5' flanking region as well as within the gene that regulate apo-E gene transcription; to determine if any of these regulatory elements can function as enhancers; to examine the protein factors associated with these regulatory elements; and to establish the roles of these proteins in the transcription of the apo-E gene.

II. METHODS AND MATERIALS

A. Construction of Recombinant Plasmids

1. Analysis of the human apolipoprotein E promoter

For the analysis of the promoter activity of the 5' flanking region of the human apo-E gene, the plasmid vector pLS1, which had been derived from the vector pTE1 (86), was used. The pLS1 vector contains the following elements: (a) a 56-bp polylinker sequence containing restriction endonuclease sites for *Eco*RI, *Cla*I, *Hind*III, *Xba*I, *Bgl*II, *Pst*I, and *Sac*I; (b) a 1634-bp fragment containing the coding sequences of the bacterial chloramphenicol acetyltransferase (CAT) gene linked to simian virus 40 (SV40) splice junctions and polyadenylation signals; (c) a 3390-bp pBR 322 fragment containing the plasmid replication origin (ori) and the ampicillin resistance (amp) gene. A *Bgl*II-*Sac*I fragment encompassing nucleotides -651 to +73 from the 5' flanking region, the first exon, and a portion of the first intron of the human apo-E gene was inserted into pLS1 immediately upstream of the CAT gene to generate the parent recombinant plasmid, pHAE-CAT1 (Fig. 1).

The generation of plasmids with progressive deletions at the 5' end was achieved by *Bal*31 nuclease digestion. Briefly, pHAE-CAT1 was digested with *Bgl*II and treated with *Bal*31 for different lengths of time. The *Bal*31 reactions were stopped by adding 20mM EGTA, heating the solution to 65 ^{o}C for 15 minutes and precipitating with ethanol. The resulting ends were ligated with *Bgl*II linkers under standard conditions. The ligation products which had various lengths of the 5' flanking sequence, were excised by digestion with *Bgl*II and *Sac*I and purified on a 5% polyacrylamide gel. The fragments were then ligated into the pLS1 vector. The precise end points of these deletion mutants were determined by DNA sequencing using the method of Maxam and Gilbert (87) or dideoxynucleotide chain termination (88).

2. Analysis of the enhancer activity of human apo-E gene fragments

To determine whether regions of the human apo-E gene contained enhancer activity, the plasmid vectors, pTK1, pTK10, and $pA_{10}CAT_2$ were employed. The pTK1 vector (known as pTE1 in reference 68) contains all the elements in pLS1 but also includes immediately downstream from the polylinker sequence, a 598-bp fragment from the pBR322 *Bam*HI site (375) to the *Nru*I site (972), followed by a 160-bp segment corresponding to the promoter sequence (nucleotides -109 to +51) of the herpes simplex virus thymidine kinase (TK) gene. The pTK10 vector is identical to pTK1 except that it lacks the 598-bp pBR322 sequence. In both of these plasmids, the expression of the CAT gene is under the control of the TK promoter. In pTK1, fragments of the apo-E gene were inserted into the *Bgl*II site located approximately 600 bp upstream of the TK promoter, whereas in pTK10, the inserts were placed immediately 5' to the TK promoter. Initially, a 1,000-bp fragment of the apo-E gene corresponding to nucleotides -651 to +357 was inserted in both the normal and the reverse orientation relative to the TK promoter in pTK1 and pTK10 (Fig. 3). Subfragments within the 1,000-bp insert were generated by either restriction endonuclease or *Bal*31 exonuclease digestion followed by the addition of *Bgl*II linkers. These subfragments were then inserted into pTK1 and pTK10 in both orientations.

To examine the potential enhancer activity of apo-E gene fragments when positioned either 5' or 3' to the CAT gene, the plasmid $pA_{10}CAT_2$ was used. The CAT gene in this plasmid was under the control of the enhancerless SV40 promoter (89). Apo-E gene fragments that contained *Bgl*II linkers at each end were inserted at the *Bgl*II site 5' to the CAT gene or at the *Bam*HI site 3' to the CAT gene (see Figure 7).

B. Transient Gene Transfer of Recombinant Plasmids

Cultured cell lines were seeded in 75-cm culture flasks at a density of 4-6 x 10^6 cells per flask twenty four hours prior to transfection. The cells were transfected with 15 μg of CAT gene recombinant DNA and 5 μg of a vector containing the β-galactosidase gene driven by the Rous sarcoma virus (RSV) promoter (89) using a modification of the calcium phosphate coprecipitation method (90). Briefly, to prepare the coprecipitate, 100 μl of the DNA solution was mixed with 500 μl of 274 mM NaCl, 1 mM KCl, 1 mM NaH_2PO_4, 0.2% dextrose, and 42 mM HEPES pH 7.05. While air was being bubbled through this solution, 400 μl of 0.3 M $CaCl_2$ was added dropwise. The solution was allowed to stand at room temperature for 30 minutes, vortexed vigorously for 30 seconds, and then added to the cells in 10 ml of culture media. Sixteen hours after transfection, cells were subjected to a 15% glycerol treatment for 2 minutes, followed by 20 to 24 hours incubation in fresh culture media before being harvested.

C. Maintenance of Tissue Culture Cells

All tissue culture media were supplemented with 100 units/ml of penicillin and 100 μg/ml of streptomycin. HeLa (Human Epitheloid Carcinoma), L (Mouse Fibroblast), Cos7 (African Green Monkey Kidney), BaGL (Human Fibroblast), J774.2 (Mouse Macrophage), U937 (Human Monocyte), and XC (Rat Pancreatic Carcinoma) cells were grown in Dulbecco's modified Eagle's medium with 10% fetal bovine serum. HepG2 (Human Hepatoma), Fu_5AH (Rat Hepatoma), and THP-1 (Human Monocyte-like) cells were grown in modified Eagle's medium with 10% fetal bovine serum. CaCo-2 (Human Colonic Carcinoma) cells were grown in modified Eagle's medium with 15% fetal bovine serum. CHO (Chinese Hamster Ovary) cells were grown in F-12 and Dulbecco's modified Eagle's medium with 10% fetal bovine serum. HIT (Hamster Pancreatic Endocrine) cells were grown in Dulbecco's modified Eagle's medium with 2.5% fetal bovine serum and 15% horse serum.

D. Chloramphenicol Acetyltransferase and β-Galactosidase Assays

Transfected cells (at about 80% confluency) were harvested approximately 36 to 40 hours after transfection by scraping, washed twice with phosphate buffered saline, and resuspended in 200 µl of CAT assay buffer (0.25 M Tris-HCl, pH 7.8, 10% glycerol, 5 mM dithiothreitol). Cell lysates were prepared by three cycles of freezing (-70 °C) and thawing (65 °C), incubation at 65 °C for 10 minutes, and centrifugation at 12,000 rpm for 15 minutes. Supernatant fluids were collected and quantified for protein by the method of Bradford using bovine serum albumin (BSA) as a standard (91). The CAT assays (92) were performed in a final reaction volume of 200 ml, and contained 50 to 100 µg of protein, 0.2 µCi of [^{14}C] chloramphenicol (specific activity, 50-57 mCi/mmol), and 5.0 mM acetyl-coenzyme A (lithium salt). The reactions were allowed to proceed for 15 minutes to 2 hours at 37 °C. For reactions longer than 1 hour, more acetyl-coenzyme A (restoring its concentration to at least 5.0 mM) was added to preserve the linearity of the assay. The samples were extracted with 1 ml ethyl acetate, the solution was dried down and the residue redissolved in 20 µl ethyl acetate and analyzed by ascending thin layer chromatography on silica gel plates using chloroform:methanol (95:5, v/v). The chromatograms were autoradiographed and the CAT activities were quantitated by counting the radioactive regions by liquid scintillation spectrometry.

The β-galactosidase assay which was used to monitor transfection efficiency for each dish was performed as described (68), with 50 µg of protein, 0.5 mM o-nitrophenyl-β-d-galactopyranoside in 600 µl of total reaction volume.

E. Preparation and Analysis of RNA

1. Isolation of total cellular RNA

Total cellular RNA was isolated from cultured cells according to the procedure of Chirgwin et al (93). Briefly, the medium is aspirated and 1 ml of 4M guanidium thiocyanate is added to the plates. The solution is passed several times through a 23-gauge needle to shear the DNA, overlaid onto 2 ml of 5.7 M CsCl, and centrifuged at 25,000 rpm at 15 ^{o}C for 18 hours in a Beckman SW55 rotor. The pellet is rinsed with 70% ethanol and the RNA is dissolved in TE (25 mM Tris-HCl, 10 mM EDTA, pH 8.0).

2. Dot blot analysis

Dot blot analysis of RNA was performed with a template manifold apparatus (Schleicher & Schuell, Keene, NH) to assure uniform dot size. Total cellular RNA was applied in three different amounts (1.0, 2.0, and 3.0 μg) and supplemented with yeast RNA to provide a final total amount of 3 μg of RNA/sample. The RNA samples were denatured by adjusting them to 1.0 M formaldehyde, 0.9 M NaCl, 0.09 M sodium citrate, and heating them at 55 ^{o}C for 15 minutes. The samples were diluted into 20 volumes of 3 M NaCl containing 0.3 M trisodium citrate and applied to nitrocellulose filters under a gentle vacuum. The sample wells were then washed with additional diluent, baked at 80 ^{o}C for 2 hours under vacuum and hybridized to ^{32}P-labelled probes.

3. Primer extension analysis

Primer extension analysis of RNA derived from gene transfer experiments or from *in vitro* transcription was performed as described (86, 94). 20 μg of total cellular RNA or *in vitro* synthesized RNA was suspended in 9 μl of reverse transcriptase buffer (100 mM Tris-HCl pH 8.0, 100 mM KCl, 16 mM $MgCl_2$, 80 mM dithiothreitol (DTT)). The RNA was annealed then to a 5' end-labelled (using T4 polynucleotide kinase and [γ-^{32}P]ATP) 20-mer single stranded oligonucleotide primer corresponding to nucleotides 15-34 of the coding sequence of the CAT gene (0.5 - 2 ng, 5 X 10^5 - 1 X 10^6 cpm per reaction

sample) for 30 minutes at 65 °C. The reaction samples were then brought to room temperature, diluted with 10 µl of a buffer containing 1 mM deoxynucleotide triphosphates (dNTPs), 20 mM $MgCl_2$, 20 units avian myeloblastosis virus (AMV) reverse transcriptase and incubated for 1 hour at 42 °C. The primer extension products were precipitated with ethanol, washed in 70% ethanol, vacuum dried, suspended in 80% formamide dye and loaded onto 6% polyacrylamide-8 M urea sequencing gels. The relative amounts of the primary transcripts on autoradiograms were measured by densitometric scanning.

F. Isolation and Preparation of Nuclear Extracts

Nuclear extracts were prepared from cultured cell lines with a modification (95) of the method described by Dignam et al. (96). Monolayer cultures grown in 245 mm tissue culture plates at 80% confluence or suspension cultures grown in 6L round-bottom flasks at a density of 3-5 x 10^5 cells/ml were harvested by centrifugation at 4 °C for 10 minutes at 3000 rpm in a Beckman JS-5.2 rotor. Pelleted cells were suspended in ice-cold phosphate buffered saline, centrifuged as above and the packed cell volume was determined. The cells were then suspended in five packed cell volumes of buffer A (10 mM hydroxyethyl piperazineethanesulfonic acid (HEPES) pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, 0.5 mM phenylmethylsulfonyl fluoride (PMSF)) and incubated in ice for 10 minutes. The cells were then centrifuged as above, suspended in two packed cell volumes of buffer A, and lysed by 8-10 strokes with the tight pestle (B) of a dounce homogenizer. The homogenate was centrifuged at 4 °C for 30 seconds at 12,000 rpm in a Sorvall HB-4 rotor and the crude nuclei were pelleted and retained.

The nuclei pellet was resuspended in buffer B (20 mM HEPES pH 7.9, 25% glycerol (v/v), 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM ethylenediamine

tetraacetate (EDTA), 0.5 mM DTT, 0.5 mM PMSF) at a volume of 3 ml/10^9 cells and gently stirred in ice for 30 minutes. The suspension was then centrifuged at 2 ^oC for 30 minutes at 12,000 rpm (HB-4) and the supernatant was collected. When supernatant volume exceeded 10 ml, solid ammonium sulfate (0.33 g/ml) was added, and the sample was rocked for 20 minutes on a "Nutator", centrifuged at 25,000 rpm for 20 minutes at 2 ^oC in a Beckman SW 55 rotor, and the pellet was redissolved in 1.0 ml dialysis buffer C/10^9 cells. Buffer C consisted of 20 mM HEPES pH 7.9, 20% glycerol (v/v), 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF. The extract was dialyzed twice for 90 minutes against > 200 volumes of dialysis buffer and the protein concentration was determined by the Bradford method using BSA as a standard. The transcription extract was quick-frozen in small aliquots and stored at -70 ^oC.

G. DNase I Footprint Assay

1. Preparation of labelled fragments

DNA fragments which span the apo-E gene regulatory elements were either 5' end-labelled by T4 polynucleotide kinase or 3' end-labelled by Klenow DNA polymerase catalyzed fill-in reaction. The double end-labelled fragments were then digested with an appropriate restriction endonuclease to generate subfragments that contained label on only one strand. These fragments were then isolated by preparative polyacrylamide gel electrophoresis and purified by the Nensorb columns (New England Nuclear, Boston MA) and dried under vacuum before use in footprint assays.

2. Footprinting assay

The footprinting reactions were performed in a 50 µl volume in a buffer containing 10 mM HEPES pH 7.9, 50 mM KCl, 0.1 mM EDTA, 0.25 mM DTT, 0.25 mM PMSF and 10% glycerol. Various amounts of nuclear extracts purified from different cell culture lines were preincubated with 1 µg of poly (dI-dC)/poly

(dI-dC) in the reaction mixture for 20 minutes at room temperature. Then 2-5 ng of single end-labelled DNA fragment was added and the mixture was incubated for 40 minutes on ice. DNase I digestion was initiated by the addition of 50 µl of a 10 mM $MgCl_2$, 5 mM $CaCl_2$ buffer and allowed to proceed at room temperature for 1 minute. The amount of DNase I was adjusted empirically to obtain even digestion patterns. Usually, 10 ng DNase I was added in the no-extract control reactions whereas 250 ng and 500 ng DNase I were added for samples containing 36 µg and 72 µg of nuclear extract, respectively. The digestion was terminated by adding 100 µl "stop buffer" (100 mM Tris-HCl pH 8.0, 100 mM NaCl, 1% sodium dodecyl sulfate (SDS), 10 mM EDTA, 100 µg/ml proteinase K, 50 µg/ml tRNA). The reactions were incubated at 37 ^{o}C for 15 minutes, extracted with equal volume of phenol:chloroform (1:1), and precipitated with 3 volumes of ethanol. In order to localize the nucleotides that were protected from DNase I digestion, the single end-labelled fragments were subjected to the G and/or G+A reactions of the Maxam-Gilbert chemical cleavage sequencing reactions (87). All reaction products were analyzed on 8% polyacrylamide-8M urea sequencing gels.

H. Gel Retention Assay

Gel retention assay was performed essentially as described in Carthew et al (97). 0.5 ng of 5' end-labeled DNA fragment or double-stranded oligonucleotide was incubated in the footprint assay buffer with 10 µg nuclear extract and various amounts of poly (dI-dC)/poly (dI-dC) (usually 4 µg), at 30 ^{o}C for 30 minutes. In competition experiments, excess unlabelled DNA fragment or oligonucleotide was allowed to preincubate with the nuclear extract for 10 minutes at room temperature before the labelled probe was added. After incubation, the samples were loaded on a 5% polyacrylamide gel (acrylamide:bisacrylamide, 40:0.5) and electrophoresed in 0.5 X TBE buffer

(44.5 mM Tris-HCl pH 8.3, 44.5 mM boric acid, 0.5 mM EDTA) at 18 V/cm after the gel had been preelectrophoresed for 1 hour at 13 V/cm. The gel was then dried and autoradiographed.

I. In Vitro Transcription Assay

The circular DNA templates used for *in vitro* transcription reactions were the apo-E promoter/CAT recombinant plasmids derived from pLS 1 and pTK1 which had been described previously (in sections A1 and A2). In addition, the plasmid pRSV-CAT containing the Rous sarcoma virus (RSV) promoter immediately upstream of the CAT gene (89) was included in the reactions as an internal control. Standard transcription reactions (20 or 40 µl volume) generally contained 100 ng of each DNA template and 70-100 µg of transcriptionally active crude HeLa nuclear extract in a final buffer concentration of 10 mM HEPES pH 7.9, 50 mM KCl, 6 mM $MgCl_2$, 10% glycerol, 0.1 mM EDTA, 0.25 mM DTT, 0.6 mM nucleotide triphosphate (NTP) mixture, 50 µg/ml poly (dI-dC)/poly (dI-dC), and 1000 units/ml RNasin ribonuclease inhibitor. For reconstitution experiments, 5 or 10 ng of 90% pure Sp1, 50 ng of purified 55 kD protein, or 300 ng of an affinity column purified fraction was added to 100 µg of HeLa nuclear extract which had been depleted of relevant transcription factors. Reactions were incubated at 30 ^{o}C for 2 hours and stopped by the addition of 160 or 180 µl of STOP buffer (100 mM Tris-HCl pH 7.5, 2% SDS, 20 mM EDTA, 50 µg/ml tRNA). Each sample then was extracted twice with an equal volume of phenol: chloroform (1:1) and precipitated with 3 volumes of ethanol. The RNA pellets were resuspended in 9 µl reverse transcriptase buffer and subjected to primer extension analysis as described previously (in section E3).

J. Synthesis of Oligonucleotides

All synthetic oligonucleotides used for deletion mutant construction, for competition gel retention assays and for affinity column preparation were made on the Applied Biosystems 380B synthesizer. Purification of the oligonucleotides was either by preparative gel electrophoresis or by passage through the Oligonucleotide Purification Cartridge system (Applied Biosystems, Foster City, CA).

K. Site-Directed Mutagenesis

DNA fragments containing portions of the human apo-E gene promoter and other regulatory elements were subcloned in M13 for use as the template in oligonucleotide-directed mutagenesis. Mutant single-stranded oligonucleotides containing deletions or base substitutions within and around transcription regulatory sequences were annealed to the relevant templates at 55 ^{o}C for 5 minutes and then extended by the double primer method using the M13 universal primer as described by Zoller and Smith (98). The double-stranded phages were transformed in TG-1 cells and phage plaque-lifting was performed after an overnight incubation. After baking the filters for 2 hours at 80 ^{o}C, the plaques containing mutants were selected by plaque hybridization using 5' end-labelled mutant oligonucleotides as probes. Hybridization was performed in a sealed plastic bag at 65 ^{o}C for 2 hours in 6X SSC (0.9 M NaCl, 0.09 M sodium citrate), 0.25% instant milk, 0.1% SDS, 4 mM sodium pyrophosphate. The filters were washed at 65 - 80 ^{o}C in 6X SSC. Positive plaques were identified by autoradiography and DNA preparations were made. The mutant clones were confirmed by sequencing using the dideoxy chain termination method (88). Because the yield of the double-stranded, replicative form (RF) of M13 DNA was low, mutant phages were made double-stranded by annealing to the M13 universal primer for 5 minutes at 55 ^{o}C and extended by the action of the Klenow fragment of DNA polymerase I for 30 minutes at room temperature.

The double-stranded DNAs were then digested with appropriate restriction endonucleases to generate fragments which were then subcloned into either the pLS1 or the pTK1 vectors and analyzed by CAT assays or by *in vitro* transcription assays.

L. Purification of Nuclear Protein Factors

Nuclear extracts were prepared from HeLa cells grown as a suspension culture in 6L round bottom flasks in Jokliks minimal Eagle's medium with 5% calf serum. The scheme for the purification of protein factors binding to apo-E gene regulatory sequences was similar to that described by Tjian and co-workers for the isolation of several of the transcription factors (65, 66, 99, 100).

1. Gel filtration chromatography

A 100-ml Sephacryl S-300 column was equilibrated in nuclear dialysis buffer C and < 5 ml of HeLa nuclear extract (5-10 mg/ml) was applied to the column for each run. The column was run at 5 ml/hour with buffer C and 2 ml-fractions were collected. Each fraction was assayed for total protein concentration by the Bradford assay and for DNA binding activity by the gel retention and/or the DNase I footprint analyses. Fractions which contained these DNA binding proteins were pooled and subjected a DNA affinity column chromatography.

2. DNA affinity column chromatography

The preparation of double-stranded oligonucleotides, the coupling of the DNA to the Sepharose column and the running of the column was essentially as described by Kadonaga and Tjian (100). Two oligonucleotides corresponding to the two strands from nucleotides -169 to -140 in the human apo-E gene 5' flanking region, with 5' protruding GATC ends were annealed in a buffer consisting of 67 mM Tris-HCl pH 7.6, 13 mM $MgCl_2$, 6.7 mM DTT, 1.3 mM

spermidine and 1.3 mM EDTA. The annealing reaction involved successive incubations at 88 °C for 2 minutes, 65 °C for 10 minutes, 37 °C for 10 minutes, and room temperature for 5 minutes. The 5' ends of the annealed oligonucleotides were then phosphorylated by the T4 polynucleotide kinase reaction with 20 mM ATP containing 10 μCi of [γ-^{32}P] ATP as the substrate. The DNA was then ligated using T4 ligase at 16 °C for 20-24 hours to form concatemers of the the basic oligonucleotide unit ranging from 10-mers to 40-mers.

The DNA oligomers then were attached covalently to cyanogen bromide (CNBr) activated-Sepharose 4B resin which had previously been washed in 0.1M HCl. The coupling reaction involved approximately 200 μg of the DNA mixed with 3.5 ml settled volume of the resin in 10 mM potassium phosphate pH 8.0 buffer for 16 hours at room temperature. The coupled resin was collected on a sintered-glass funnel and washed successively with water and with 1M ethanolamine-HCl pH 8.0. The resin was then suspended in 1M ethanolamine and shaken for 6 hours at room temperature to inactivate the unreacted CNBr-derivatized Sepharose. The resin was then washed successively in 10 mM potassium phosphate pH 8.0, 1 M potassium phosphate pH 8.0, water, and storage buffer which consists of 10 mM Tris-HCl pH 7.6, 0.3 M NaCl, 1 mM EDTA, and 0.02% (wt/vol) sodium azide. The resin was stored in the buffer at 4 °C until use.

The affinity chromatography procedure generally employed 1 ml bed volume of the DNA coupled Sepharose column equilibrated in dialysis buffer C. Crude HeLa nuclear extract or Sephacryl S-300 fractions (< 10 mg total protein) were incubated with variable amounts of poly (dI-dC)/poly (dI-dC) in ice for 10 minutes before being loaded onto the column. The mixture was allowed to flow through the affinity column by gravity (approximately 15-20 ml/hour flow

rate) and then the column was washed with buffer C. To elute the bound proteins, either buffer C with 0.8 M KCl or with a gradient of 0.1 - 0.8 M KCl in 0.1 M increments was used. For second and third affinity column passages, the high salt eluted fractions were diluted to 0.1 M KCl with buffer C without KCl, mixed with additional poly (dI-dC)/poly (dI-dC) and reapplied to the affinity column. The affinity column were regenerated by washing with 5 mM Tris-HCl pH 7.6, 2.5 M NaCl, 0.5 mM EDTA followed by a wash with the storage buffer.

3. SDS polyacrylamide gel electrophoresis

8 or 10% polyacrylamide gel electrophoresis of the purified proteins in the presence of SDS was done according to the method of Laemmli (101). The gels were visualized by silver staining according to Oakley et al (102).

M. Materials

Cell culture media and nutrients were obtained from GIBCO (Grand Island, NY). [14C]chloramphenicol (50-57 mCi/mmol) was purchased from New England Nuclear (Boston, MA). Acetyl-coenzyme A (lithium salt) and o-nitrophenyl-β-d-galactopyranoside were purchased from Pharmacia Biochemicals (Piscataway, NJ) and Sigma Chemical Co. (St. Louis, MO), respectively. All restriction endonucleases, Bal31 nuclease, T4 DNA ligase, linkers, and Klenow fragments of DNA polymerase I were purchased from either New England Biolabs (Beverly, MA) or Boehringer Mannheim Biochemicals (Indianapolis, IN). T4 polynucleotide kinase was from U.S. Biochemical (Cleveland, OH). RNasin was purchased from Promega Biotec (Madison, WI). AMV reverse transcriptase was from Life Sciences Inc. (St. Petersburg, FL). DNase I was from Cooper Biomedical (Malvern, PA). Alpha amanitin was from Boehringer Mannheim Biochemicals (Indianapolis, IN). $\alpha[^{32}P]$dCTP (3000 Ci/mmol) and γ-$[^{32}P]$ATP (>5000 Ci/mmol) were purchased from Amersham (Arlington Heights, IL). NTP, dNTP, poly (dI-dC)/poly (dI-dC),

Sephacryl S-300, and CNBr-activated Sepharose 4B were purchased from Pharmacia Biochemicals (Piscataway, NJ). Unless stated otherwise, procedures involving recombinant DNA, enzymes, and reagents were used under the conditions recommended by the suppliers.

III. RESULTS

A. Identification and Characterization of Transcriptional Regulatory Elements

1. Expression of Apo-E/CAT hybrid genes in cultured cells

It had been shown previously using human apo-E genomic cosmid clones that the regulatory sequences necessary for the transcription of the apo-E gene are located within 0.7 kb of the 5' flanking region and 1.6 kb of the 3' flanking region (54). To identify the boundaries of the promoter region of the apo-E gene, a *Bgl*II-*Sac*I fragment of the gene (spanning nucleotides -651 to +73) was inserted into the pLS1 vector at the 5' end of the bacterial CAT gene which was employed as a general marker for promoter activity (Fig. 1). The resulting plasmid, pHAE-CAT1, was transfected into different cultured mammalian cell lines and the CAT activity in the cell extract was measured. As a control for differences in transfection efficiency, the various cell lines were also transfected with the plasmid, pRSV-CAT. The apo-E promoter-driven CAT activity in each cell line was determined relative to the CAT activity generated by the Rous sarcoma virus (RSV) promoter. The results, summarized in Table I, show that the apo-E promoter activity was strongest in CHO cells, with HepG2, L, and CaCo-2 cells also showing relatively strong activities. CHO and HepG2 cells were used for subsequent CAT assay analyses of the 5' flanking regulatory elements.

2. Identification of 5' flanking sequences necessary for gene expression

To identify specific sequences in the 5' flanking region that are significant for human apo-E gene expression, portions of the 5' ends of this region were progressively deleted by *Bal*31 nuclease treatment. The structures resulting from the deletions are outlined in Fig. 2, as are their activities. In both CHO and HepG2 cells, deletion of the 268 nucleotides between -651 and

Figure 1 Construction of apo-E promoter/CAT expression recombinants.

A 724-bp fragment (-651 to +73) of the human apo-E gene sequence was excised by *Bgl*II-*Sac*I digestion of the genomic clone pUCE4 (42) and ligated into plasmid pLS1 to produce pHAE-CAT1. This plasmid (pLS1) was derived from pTK1 by the removal of the TK promoter by *Bgl*II-*Sac*I digestion.

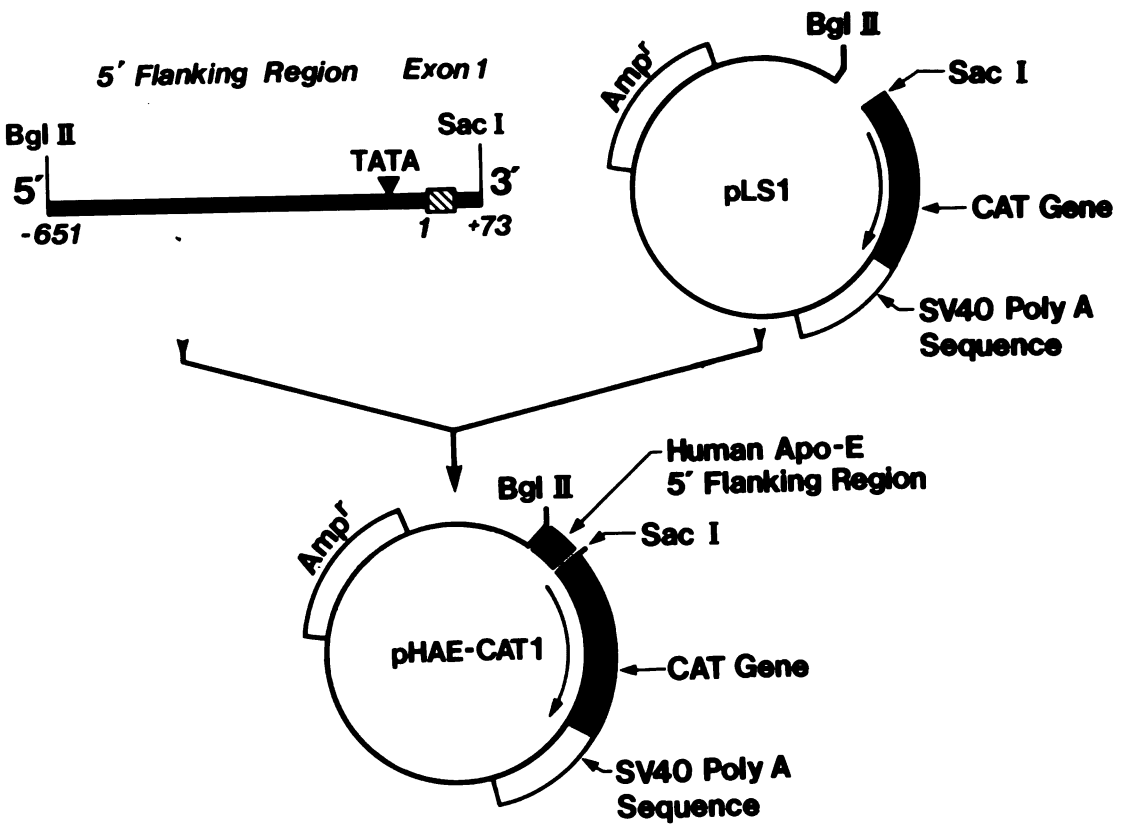

5′ Flanking Region
Exon 1
Bgl II
Sac I
TATA
5′
3′
-651
1
+73
Bgl II
Ampr
Sac I
pLS1
CAT Gene
SV40 Poly A
Sequence
Human Apo-E
5′ Flanking Region
Bgl II
Sac I
Ampr
pHAE-CAT1
CAT Gene
SV40 Poly A
Sequence

Table I

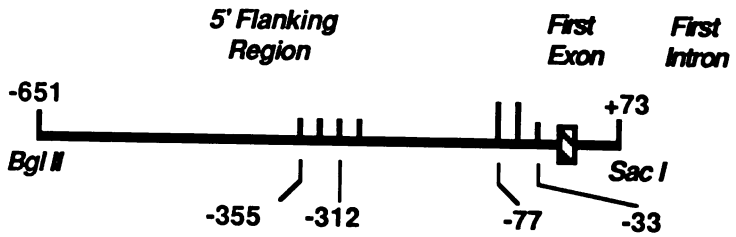
Relative chloramphenicol acetyltransferase (CAT) activities directed by the 5'-flanking sequences of the human apo-E gene in different cultured cell lines

For each cell type, the relative enzyme activity (normalized to β-galactosidase activity) is expressed as a percentage of that directed by the RSV enhancer (i.e. pRSVCAT = 100%). Results shown represent the mean ±S.D. from three independent experiments. Cells were transfected with 15 μg of pHAE-CAT1 and examined for CAT activity. One hundred mg of cell lysate was incubated in the presence of acetyl-coenzyme A and [14C]chloramphenicol for 15 minutes as described previously (92). A longer period of incubation (2 hours) was required for the assay of HeLa, BaGL, 3T3, U-937, J774.2, and Fu5AH cell lysates because of their low activity. The incubation times chosen here were at the middle of the linear curve for CAT reaction.

Cell line	Relative activity (n=3)
	%
CHO	48.5 ± 3.8
HepG2	28.1 ± 4.7
L	27.3 ± 2.9
CaCo-2	22.2 ± 5.3
3T3	8.3 ± 2.5
HeLa	5.6 ± 1.6
BaGL	5.2 ± 2.1
U-937	3.4 ± 0.3
J774.2	1.0 ± 0.4
Fu5AH	1.0 ± 0.2

Figure 2 Promoter activity of deletion fragments of the 5'-flanking region of the apo-E gene.

Apolipoprotein E promoter-CAT gene recombinants are shown. The numbers to each side of the solid bars indicate the positions in base pairs relative to the transcription initiation site. The hatched box represents the first exon. The vertical bars above and the numbers beneath the uppermost bar indicate the position of each pair of repeated sequences as described in Ref. 42. The values in the columns on the right represent the means of results that were obtained from two separate experiments. The relative CAT activities of the different deletion constructs are expressed as a percentage of that of pHAE-CAT1 (-651 to +73). To obtain relative CAT activities, the specific activity of each construct was normalized to the β-galactosidase activity before comparison (i.e. normalized CAT activity = CAT specific activity/β-galactosidase activity).



	Relative Activity (%)	
	HEPG2	CHO
-651 to +73	100	100
-556 to +73	104	91
-383 to +73	90	87
-212 to +73	52	46
-81 to +73	12	12
-39 to +73	5	3
-14 to +73	< 0.5	< 0.5

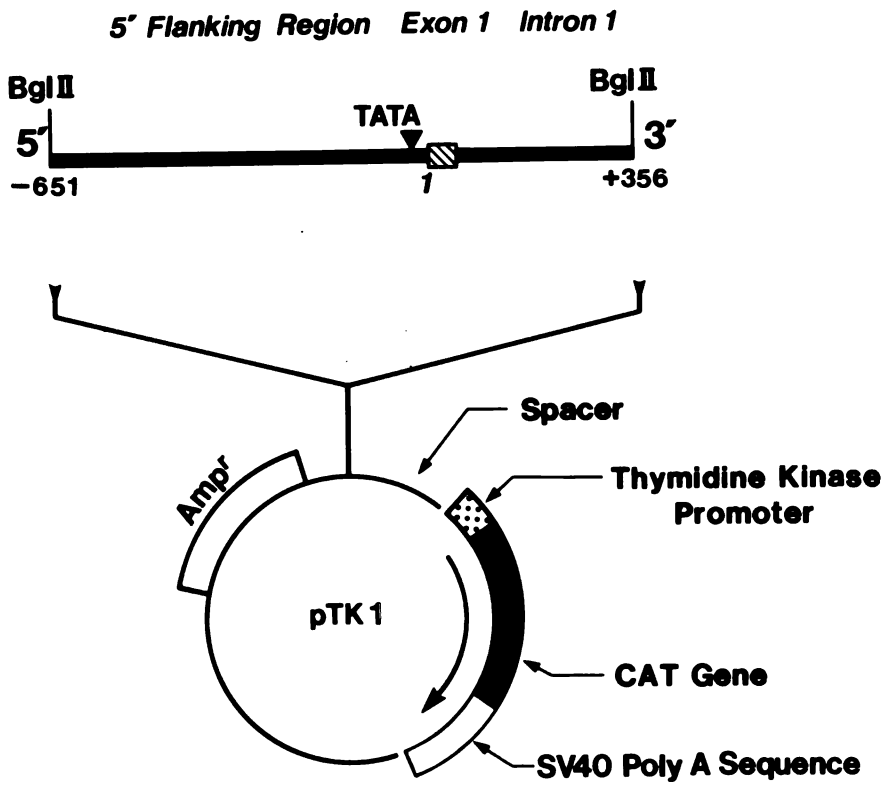
-383, which contains part of an *Alu*I sequence, had little effect on apo-E promoter-directed CAT activity. Deletion of the region between nucleotides -383 and -212 resulted in almost a twofold reduction in CAT activity. Deletion of the sequence between nucleotides -212 and -81 resulted in about a fourfold reduction in CAT activity in both cell types. The removal of the region between nucleotides -81 and -39 resulted in another 2.5- to 4-fold reduction in promoter activity. Since this region contains a consensus GC box element at positions -54 to -45, this result suggests that this element is a functional component of the apo-E promoter complex. Deletion of the region from nucleotides -39 to -14, which contains a consensus TATA box element (-33 to -27), resulted in an additional 6- to 10-fold decrease in activity. The results from these deletion studies with the chimeric CAT gene suggest that, in addition to the TATA box, at least three different domains within the proximal 383 nucleotides of the 5' flanking region of the human apo-E gene are involved in the regulation of its transcription.

3. Identification of apo-E gene enhancer elements

To determine whether the regulatory elements in the 5' flanking region of the apo-E gene as well as the intragenic sequences can function as transcriptional enhancers, the pTK1 and pTK10 vectors were employed. In these vectors, the expression of the CAT gene is controlled by the thymidine kinase (TK) promoter of the herpes simplex virus. The promoter consists of nucleotides -109 to +50 of the TK gene and contains two Sp1 binding sites and one CTF binding site, which allows the promoter to be active in most mammalian cultured cells (94). In pTK1, putative enhancer elements were inserted upstream of the TK promoter, separated by a 600 bp spacer (Fig. 3), whereas in pTK10, the enhancer insertion site is 5' adjacent to the TK promoter.

Figure 3 Construction of apo-E enhancer test recombinants.

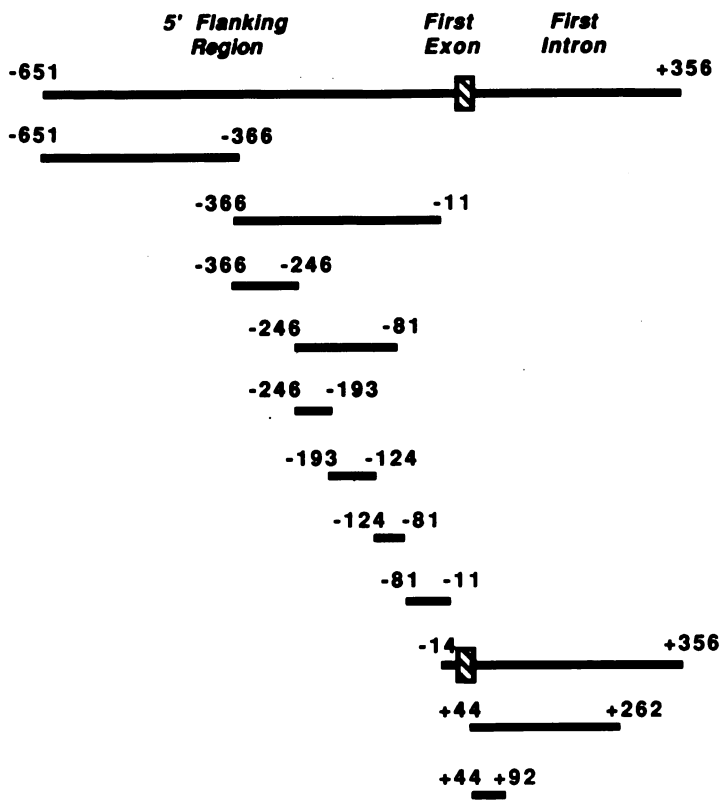
A 1001-bp fragment of the human apo-E gene was excised by *P*II and ligated into the *Bgl*II site of the vector, pTK1, which is located 617 bp upstream of the TK promoter, producing two plasmids in which the apo-E fragment is in different orientations relative to the TK promoter.


5′ Flanking Region
Exon 1
Intron 1
BglII
BglII
TATA
5′
3′
−651
1
+356
Spacer
Thymidine Kinase
Promoter
Ampr
pTK1
CAT Gene
SV40 Poly A Sequence

First, a 1 kb fragment spanning nucleotides -651 to +356 (including the first exon and 312 nucleotides of the first intron) was ligated into pTK1 (Fig. 3) and the CAT activity of the resulting construct was determined. In both CHO and HepG2 cells, this fragment stimulated the activity of the TK promoter by about fivefold in both orientations relative to the control pTK1 vector (Fig. 4). Then, to delineate enhancer sequences within this 1 kb fragment, a series of smaller fragments were generated by restriction endonuclease digestion and they were inserted into the pTK1 vector and tested for CAT activity. The results shown in Fig. 4 suggest that there are three elements with enhancer-like properties. A fragment spanning nucleotides -246 to -81 was able to stimulate TK promoter activity by as much as ninefold. Subdivision of this fragment revealed a 69-bp segment between nucleotides -193 and -124 which showed enhancer activity, although it was consistently less than that of the entire fragment. This "core" sequence was termed upstream regulatory element (URE) 1. Another element termed upstream regulatory element (URE) 2 which demonstrated about a fourfold stimulation of the TK promoter is located between nucleotides -366 and -246. The third enhancer element which also exhibited a fourfold stimulation was found within the first intron in a 219 bp segment between residues +44 and +262 and was termed intron regulatory element (IRE) 1. Subdivisions of URE2- and IRE1-containing fragments did not yield segments containing enhancer activities. Two apo-E gene fragments, one containing both URE1 and URE2 (nucleotides -366 to -11) and the other containing IRE1 (nucleotides -14 to +356) were inserted into pTK10 to examine the effect of distance on the activities of the apo-E enhancer elements. The observed enhancement was essentially identical to that observed with the fragments in pTK1, suggesting that the actions of the apo-E gene enhancer elements are independent of distance under these conditions.

Figure 4 Enhancement of the apo-E gene fragments on TK promoter-directed CAT expression (pTK1).

The arrows indicate the orientation of the inserts adjacent to CAT coding sequences and the first exon. "-" denotes "not determined". The numbers on each side of the solid bars indicate the position in base pairs relative to the transcription start site. The hatched box represents the first exon. The enhancement of each construct above the level of expression found in the control pTK1 is expressed as -fold by comparison of their specific activities. The values represent the mean ± S.D. obtained from three experiments, except those indicated by asterisks, which were from four independent experiments.



	Relative Activity			
	CHO Cells (- fold)		HepG2 Cells	
	→	←	→	←
-651 to +356	5.6 ± 2.1	4.5 ± 1.6	4.3 ± 1.2	3.9 ± 0.9
-651 to -366	1.1 ± 0.5	–	1.5 ± 0.7	–
-366 to -11	5.1 ± 1.8	3.5 ± 1.1	7.3 ± 1.1	5.7 ± 1.8
-366 to -246	4.2 ± 1.5*	4.2 ± 0.7*	4.9 ± 2.6*	3.5 ± 1.4*
-246 to -81	9.4 ± 2.6*	6.7 ± 1.2*	6.8 ± 2.1*	5.1 ± 0.7*
-246 to -193	1.4 ± 0.4	–	1.2 ± 0.7	–
-193 to -124	5.4 ± 1.1	4.3 ± 0.8	4.7 ± 1.3	5.1 ± 0.7
-124 to -81	0.9 ± 0.5	–	1.5 ± 0.6	–
-81 to -11	1.5 ± 0.7*	–	1.8 ± 0.1*	–
-14 to +356	2.8 ± 0.3	2.2 ± 0.4	3.1 ± 0.6	–
+44 to +262	4.8 ± 1.4*	3.5 ± 0.3*	3.3 ± 0.6*	3.8 ± 0.9*
+44 to +92	1.7 ± 0.5	1.3 ± 0.7	–	–

In order to verify that the stimulation of TK promoter-driven CAT activity by the apo-E gene enhancer elements was indeed reflecting events at the RNA level, total cellular RNA was isolated from CHO cells transfected with several of the apo-E enhancer/pTK1 recombinants. Dot blot analysis of total RNA hybridized to a ^{32}P-labelled CAT gene probe revealed that an IRE1-containing pTK1 vector produced transcript signals approximately 3.5-fold higher than those produced by pTK1 itself, whereas a URE1-containing vector stimulated transcription by approximately fivefold (Fig. 5). To determine if the dot blot RNA signals represented correctly initiated mRNA transcripts, primer extension analysis was performed. Using a primer corresponding to nucleotides 15 to 34 of the coding sequence of the CAT gene, pTK1 was able to generate a 134-bp transcript which indicated correct transcription initiation (Fig. 6). The URE1-containing pTK1 vector also generated a 134-bp transcript band, but at an intensity sixfold that of the pTK1 vector itself (Fig. 6). Thus the results with the RNA analyses were consistent with the enhancer test results obtained from the CAT activity assays. In addition, the apo-E promoter/CAT construct, pHAECAT1, was able to generate a 150-bp transcript band, indicating correct initiation from the apo-E gene transcription initiation site (Fig. 6).

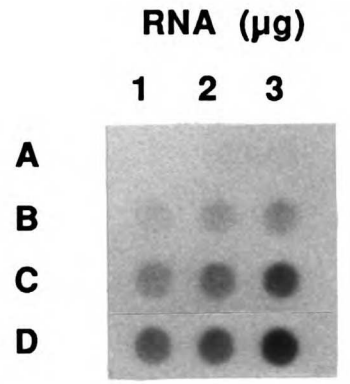
4. Characterization of apo-E gene enhancer elements

To determine whether the apo-E gene enhancer elements can act when positioned at the 3' end of the CAT gene and with a different promoter, another test vector, $pA_{10}CAT_2$, which contained an enhancerless SV40 promoter (89) was employed. The three enhancer elements were inserted either immediately 5' to the SV40 promoter or at the 3' end of the CAT gene and were transfected into CHO cells. The results shown in Fig. 7 indicate that URE1 and IRE1 were able to enhance SV40 promoter activity from both the 5' and the 3' positions. For IRE1, the effect from the 3' position was about half that from the 5'

Figure 5 Dot blot analysis of apo-E-enhancer/TK1 recombinants.

Total cellular RNA from transfected CHO cells was examined by the dot blot hybridization method. ^{32}P-labelled CAT gene probe was employed. (A) Autoradiogram of dot blot analysis. A: control RNA from CHO cells not transfected with DNA. B: RNA from pTK1 transfected cells. C: RNA from cells transfected IRE1-containing (+44 to +262) pTK1 vector. D: RNA from cells transfected with URE1-containing (-246 to -81) pTK1 vector. (B) Plot of the autoradiogram shown above scanned by quantitative densitometry. The letter codes are as above.

A



B

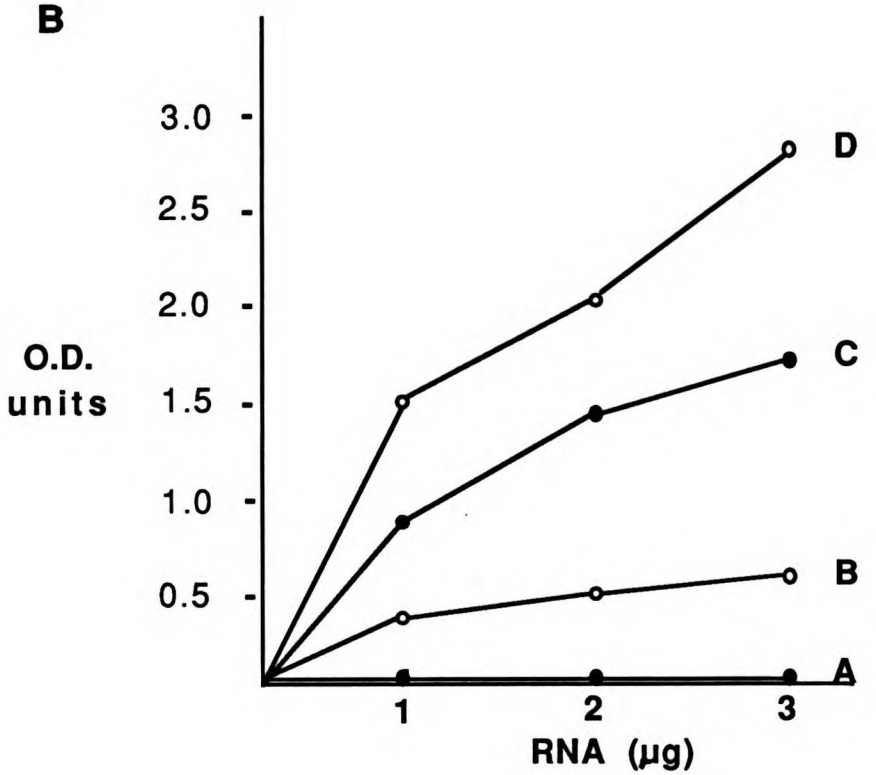


Figure 6 Primer extension analysis of transcripts generated by apo-E enhancer/TK1 constructs.

Total cellular RNA from transfected CHO cells was primed with an oligomer corresponding to nucleotides 15-34 of the protein coding sequence of the CAT gene (86). The positions of the primer extension products of RNA initiating at the transcription start sites from the apo-E promoter (apo-Ep, 150 bp in length) and the TK promoter (TKp, 134 bp) are indicated. Lane 1, pSVE-CAT containing SV40 enhancer that is positioned 600 bp upstream of pHAE-CAT1; lane 2, pHAECAT1; lane 3, control RNA with no transfected DNA; lane 4, pSVTK-CAT containing SV40 enhancer that is positioned 600 bp upstream of pTK1; lane 5, pTK1; lane 6, URE1-containing (-246 to -81) pTK1 vector.

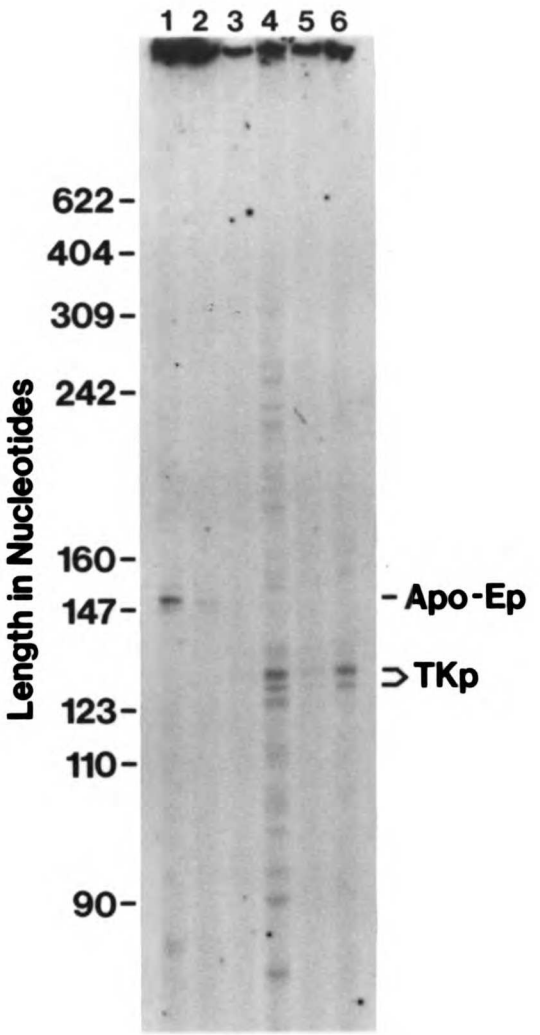
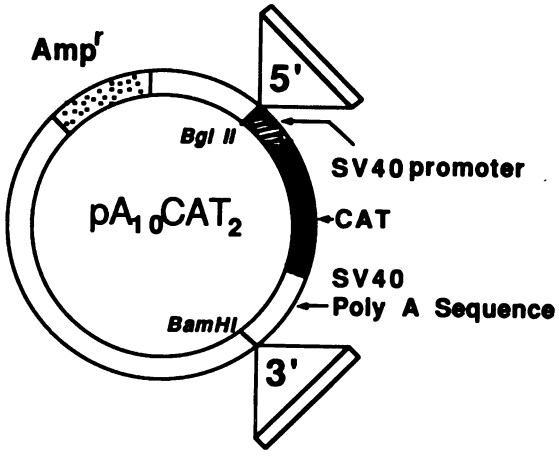

1 2 3 4 5 6
622
404
309
242
160
147
123
110
90
Length in Nucleotides
Apo-Ep
TKp

Figure 7 Effect of position on apo-E enhancer element activity in CHO cells.

The $pA_{10}CAT_2$ vector was digested with either *Bgl*II or *Bam*HI and ligated with one of three DNA fragments as listed in the table at the right of the panel. All constructs tested in this experiment have inserts in the correct orientation. Enhancer activity is expressed as -fold by comparison of specific activity of each construct over that of the control vector $pA_{10}CAT_2$. Values represent the mean $\pm$S.D. (n=3).



Enhancer Activity			
Relative Position	URE 2 *(-366/-246)*	URE 1 *(-246/-81)*	IRE 1 *(+44/+262)*
	-fold	-fold	-fold
5'	1.7 ± 0.2	4.9 ± 1.6	3.8 ± 0.5
3'	1.3 ± 0.1	3.0 ± 0.4	2.1 ± 0.1

position. URE2 possessed little activity at either position, suggesting that it may be influenced by the associated promoter, perhaps through specific interactions of *trans*-acting factors.

To test whether the apo-E enhancer sequences could exhibit cell-type specificity, the TK promoter/CAT constructs containing URE1, URE2, and IRE1, as well as the DNA segment spanning nucleotides -81 to -15 as a negative control, were transfected individually into four different cultured cell lines (Table II). URE1 exhibited the strongest enhancer activity in all the cell lines tested, and URE2 was also active in these cell lines. IRE1 showed equivalent activity in CHO and HepG2 cells, and little, if any, activity in HeLa and L cells. The URE1-containing plasmid was tested also in several other cell lines and the enhancer activities were as follows: Cos7 cells, 3.2-fold; XC cells, 5.4-fold; HIT cells, 4.6-fold; BaGL cells, 2.8-fold.

5. Characterization of URE1 enhancer function

Because it exhibited the strongest enhancer activity in all the cell lines that were tested and because it had been localized to a definite DNA segment, the URE1 enhancer element was studied in greater detail. First, a series of deletion constructs were made from a DNA fragment spanning nucleotides -246 to -81 using site-directed mutagenesis. These modified fragments were inserted into the pTK1 vector, and the resulting recombinants were transfected into CHO cells and tested for their abilities to stimulate TK promoter-directed CAT activities. Figure 8 shows the results of this experiment, in which the wild type -246 to -81 fragment stimulated TK promoter activity by about sevenfold and it was assigned a relative activity of 100%. Within URE1, deletions between nucleotides -165 and -155 and between nucleotides -154 and -144 resulted in the reduction the enhancer activity to 53% and 31% of the activity of the wild type fragment, respectively. Outside of

Table II

Enhancer activity of apolipoprotein E gene fragments in different cultured cell lines

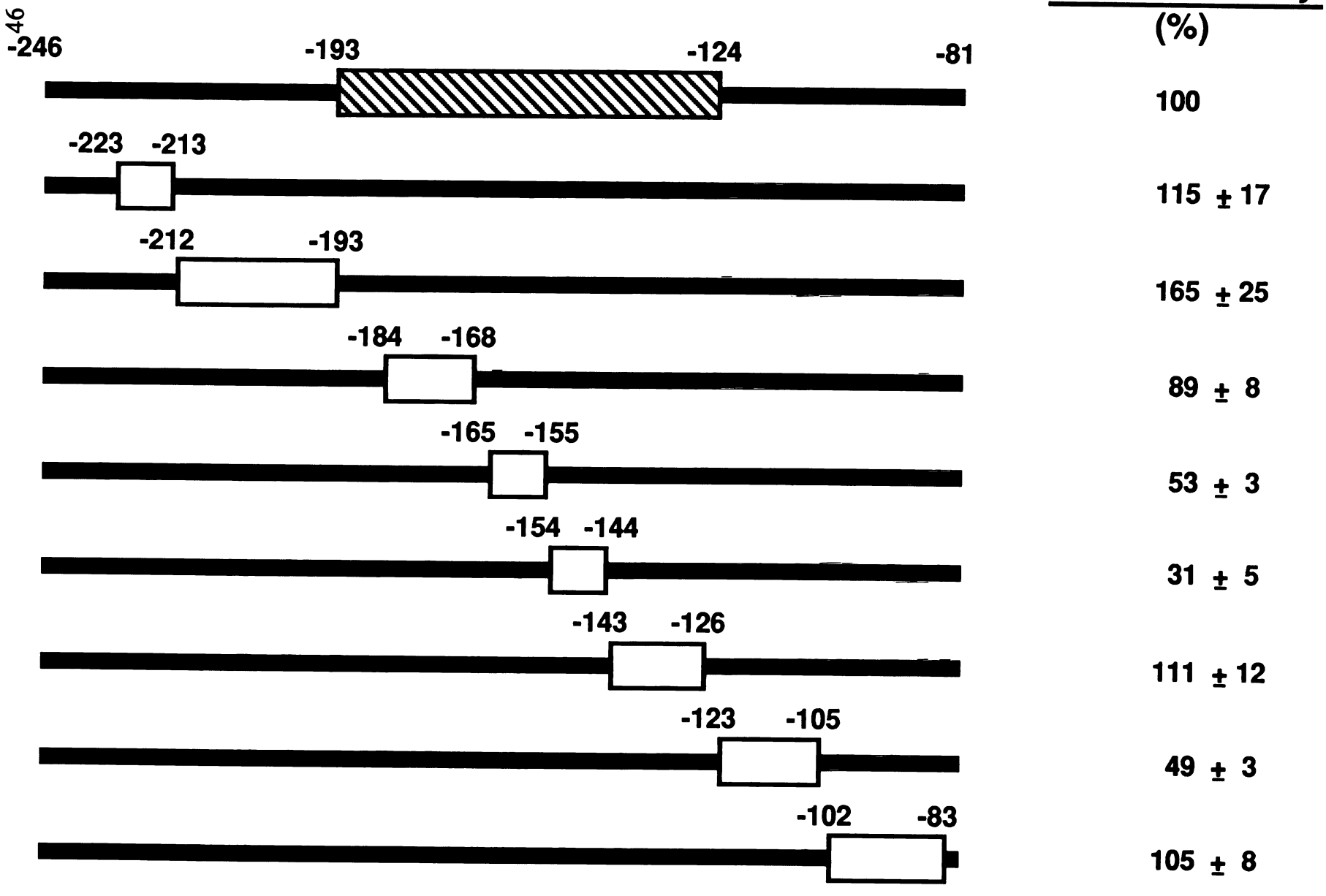
Results represent the mean ±S.D. from four independent experiments that were corrected individually for β-galactosidase activity. The apo-E gene fragments were inserted in pTK1

Plasmid	Fragment	Relative activity			
		CHO	HepG2	HeLa	L
		-fold			
pHAEN-1	-366 to -246 (URE2)	4.2 ± 1.5	4.9 ± 2.6	2.5 ± 0.5	2.2 ± 0.1
pHAEN-2	-246 to -81 (URE1)	9.4 ± 2.6	6.9 ± 2.1	5.0 ± 1.8	7.3 ± 2.9
pHAEN-3	-81 to -11	1.5 ± 0.4	1.8 ± 0.1	1.3 ± 0.2	1.4 ± 0.5
pHAEN-4	+44 to +262 (IRE1)	4.8 ± 1.4	3.3 ± 0.6	0.7 ± 0.3	1.7 ± 0.2
pTK1 (control)		1.0	1.0	1.0	1.0

Relative activity is expressed as -fold stimulation over the control, pTK1. All enhancer constructs used for this experiment have the 5' to 3' orientation with respect to the test TK promoter

Figure 8 Analysis of site-directed deletion constructs of URE1.

Enhancer activity of the deletion mutant constructs were determined. The hatched box in the top wild-type construct represents URE1. The numbers on each side of the open boxes in the mutant constructs represent the boundaries of the nucleotides that had been deleted. The relative CAT activities of the different constructs are expressed as a percentage of that of the wild-type -246 to -81 construct. The values represent the mean ±S.D. obtained from three separate experiments.



Relative Activity (%)

	100
	115 ± 17
	165 ± 25
	89 ± 8
	53 ± 3
	31 ± 5
	111 ± 12
	49 ± 3
	105 ± 8

URE1, the -212 to -193 deletion resulted in a 1.6-fold increase in enhancer activity, suggesting the removal of a negative element. Also, deleting nucleotides -123 to -105 showed a twofold decrease in activity, which may explain why the stimulation of promoter activity by URE1 is consistently less than that by the -246 to -81 fragment.

The enhancer test assay with the URE1 deletion mutants indicated that a region within URE1 may be important for this element to demonstrate enhancer activity. To determine if the "critical" region itself could function as an enhancer element, a double-stranded 30-bp oligonucleotide extending from bases -169 to -140 was synthesized and inserted into the pTK1 vector. When transfected into CHO cells, this sequence stimulated TK promoter activity by three- to fourfold in both orientations (Fig. 9). The degree of stimulation was approximately the same as that with URE1 and one-half that observed for the -246 to -81 fragment. This 30-bp enhancer sequence was termed PET (Positive Element for Transcription).

B. Detection of Protein Factors Binding to Upstream Regulatory Sequences

1. DNase I footprint analysis of upstream sequences

Since *cis*-acting DNA sequences required for gene regulation are very often sites for the binding of nuclear protein factors (59), the apo-E gene regulatory sequences were tested for the presence of sequence-specific protein binding activities by the DNase I footprinting technique (103). First, a DNA fragment spanning nucleotides -366 to -11 was end-labelled at position -11 on the 5' strand and incubated with two different concentrations of nuclear extract prepared from HepG2 cells. The fragment was subjected then to partial cleavage with pancreatic DNase I. A region between nucleotides -165 and -144 was observed to be protected from DNase I digestion, which indicated specific binding of nuclear protein(s) to that sequence (Fig. 10). This footprint lies not

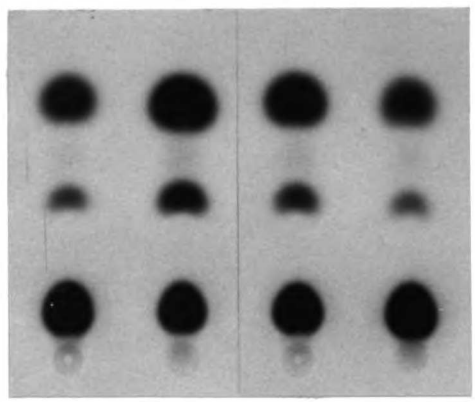
Figure 9 Enhancer activity of the PET sequence.

(A) A double-stranded 30 bp-oligonucleotide (-169 to -140), termed, PET, was synthesized and inserted into the pTK1 vector. (B) CAT enzyme activities of PET/TK1 constructs. The arrows represent the orientation of the inserted PET sequence relative to the TK promoter. CTL represents a TK1 vector containing the -81 to -15 fragment. The enhancement of the constructs above the level of expression found in pTK1 is expressed as -fold by comparison of their specific activities. The values represent the mean ±S.D. obtained from four separate experiments.

A

-169 AAAGACCTCTATGCCCCACCTCCTTCCTCC -140
TTTCTGGAGATACGGGGTGGAGGAAGGAGG

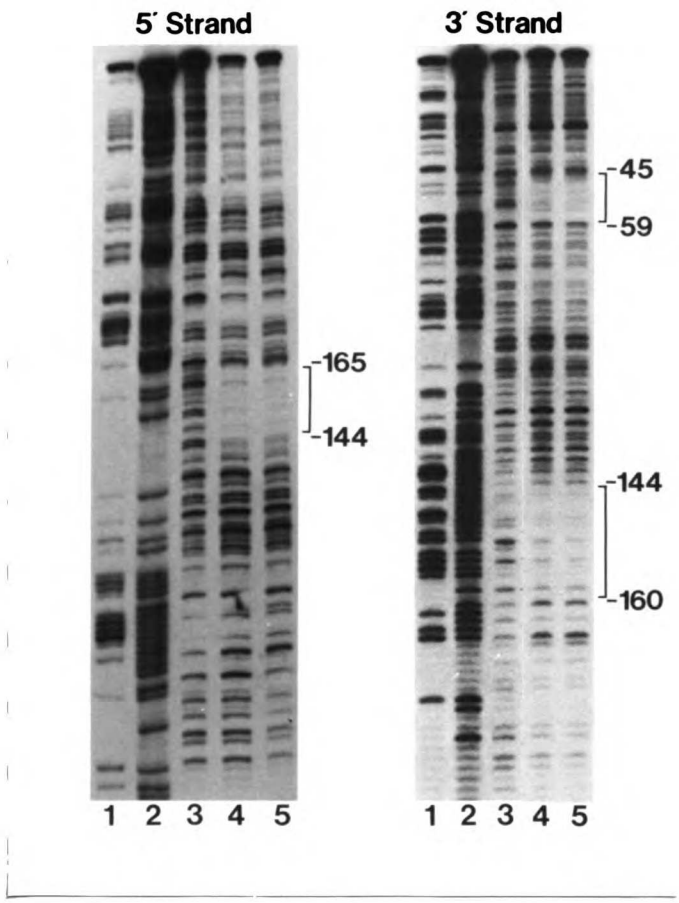
B



	CTL	PET →	PET ←	TK1
-fold	1.3	4.1	3.0	1.0
	±0.1	±1.3	±0.4	

Figure 10 DNase I footprint analysis of apo-E gene fragments by HepG2 nuclear proteins.

Apolipoprotein E DNA probes were 3' end-labelled at nucleotide -11 for the 5' strand, and at nucleotide -366 for the 3' strand. Lanes 1 and 2, G and G+A reactions of Maxam and Gilbert sequencing reactions; lane 3, no nuclear extract; lane 4, 36 μg of nuclear extract; lane 5, 72 μg of nuclear extract. Sequences protected from DNase I digestion are indicated by the brackets, with the nucleotide positions at the boundaries indicated on the right.


5′ Strand
3′ Strand
165
144
45
59
144
160
1 2 3 4 5
1 2 3 4 5

only within URE1 but also within the PET sequence (nucleotides -169 to -140), which had been shown to exhibit enhancer activity. When the corresponding 3' strand was end-labelled and tested for DNase I footprinting, a similar protected sequence (nucleotides -160 to -144) was observed (Fig. 10). Also protected from DNase I digestion on the 3' strand were nucleotides -59 to -45 where the proximal GC box element (nucleotides -54 to -45) resides and is probably bound to the transcription factor, Sp1 (104). On the other hand, no sequence within the URE2 enhancer element displayed protein binding with HepG2 cell nuclear extracts under the conditions employed.

2. Cell-type specificity of the DNA-binding proteins

Because most of the CAT assay experiments to identify apo-E gene promoter and enhancer elements were performed in CHO cells, DNase I footprinting analysis was also performed with nuclear extracts prepared from CHO cells using either a DNA fragment spanning nucleotides-366 to -11 or a fragment spanning nucleotides -212 to -11. Within URE1, nucleotides -161 to -141 on the 5' strand and nucleotides -160 to -144 on the 3' strand were found to be protected from DNase I digestion (Fig. 11). The binding of Sp1 (or an Sp1-like protein) to the proximal GC box (nucleotides -59 to -45) was observed on both strands (Fig. 11). Similar to HepG2 extracts, no footprints were detected within URE2. To examine further the cell-type specificity of the footprint detected within URE1, the short DNA fragment (-212 to -11), end-labelled on the 5' strand, was incubated with nuclear extracts from HepG2, HeLa and L cells and then subjected to DNase I digestion. Figure 12 shows that extracts from all cell lines protected the sequence spanning nucleotides -161 to -141. The slight shift of this footprint sequence with HepG2 extracts (from -165 to -144 to -161 to -141) was probably due to the fact that the original footprint using the longer DNA fragment was located near the top of the sequencing gel, making

Figure 11 DNase I footprint analysis of apo-E gene fragments by CHO nuclear proteins.

Apolipoprotein E DNA probes were 5' end-labelled at nucleotide -366 or -212 for the 5' strand (by kinase reaction) and 3' end-labelled at nucleotide -366 for the 3' strand (by Klenow reaction). Lanes 1 and 2, G and G+A reactions of Maxam and Gilbert sequencing reactions; lane 3; no nuclear extract; lane 4, 36 μg of nuclear extract; lane 5, 72 μg of nuclear extract. Sequences protected from DNase I digestion are indicated by the brackets, with the nucleotide positions at the boundaries indicated on the right.

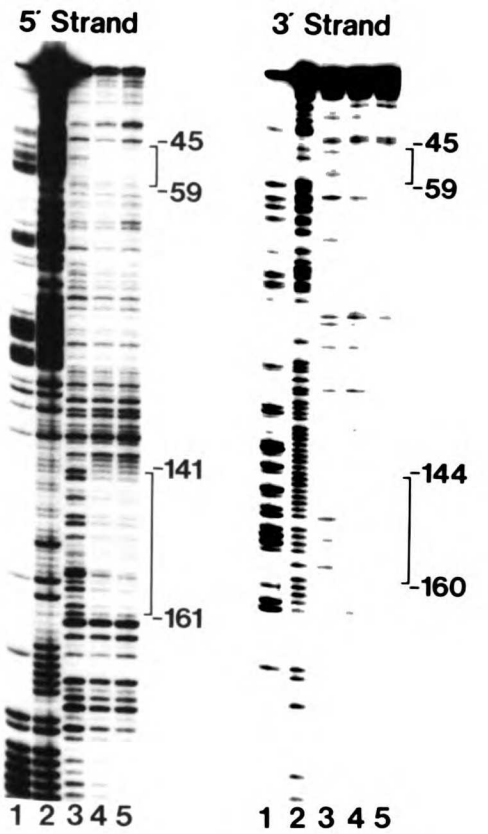
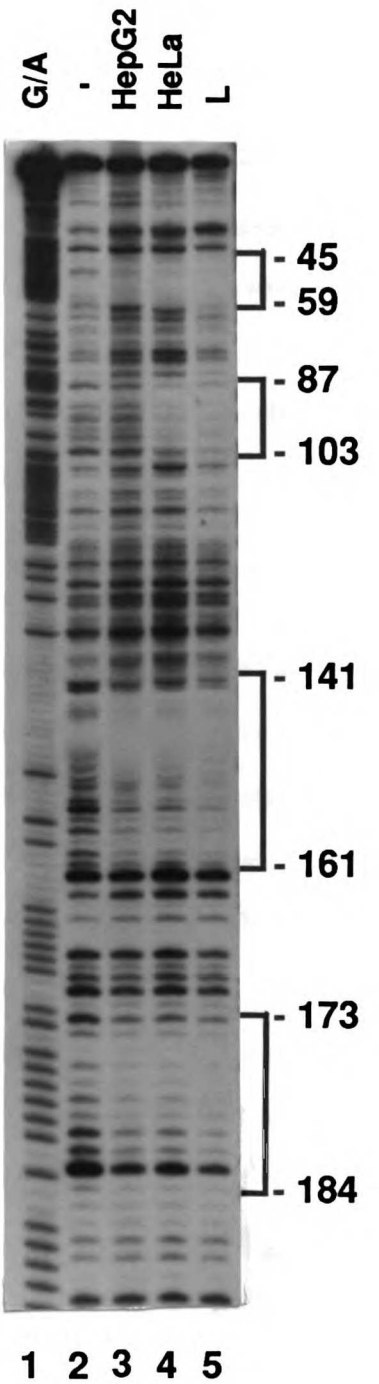

5′ Strand
3′ Strand
-45
-59
-141
-161
-45
-59
-144
-160
1 2 3 4 5
1 2 3 4 5

Figure 12 DNase I footprint analysis of the proximal 5' flanking sequence of the apo-E gene by nuclear proteins from various cell lines.

DNase I footprint assay was performed on apo-E DNA that had been 5' end-labelled at position -212 on the 5' strand. Lane 1, G+A sequence ladder; lane 2, no nuclear extract; lanes 3-5, 36 μg of nuclear extract from HepG2 (lane 3), HeLa (lane 4), and L (lane 5) cells. Sequences protected from DNase I digestion are indicated by the brackets, with the nucleotide positions at the boundaries indicated on the right.


G/A
-
HepG2
HeLa
L
45
59
87
103
141
161
173
184
1 2 3 4 5

the accurate reading of the nucleotides difficult. The GC box footprint at bases -59 to -45 was observed in all the extracts. A third footprint spanning nucleotides -184 to -173 was also detected in extracts from the three cell lines (Fig. 12). Although this segment lies within URE1, the enhancer test-CAT assay experiment demonstrated that the deletion of this sequence had no effect on the enhancer activity of the -246 to -81 fragment (Fig. 8). Finally, a fourth footprint, spanning nucleotides -103 to -87 was observed in extracts from HeLa and L cells but not in HepG2 cells (Fig. 12). Closer examination of the DNase I footprint experiment with CHO extracts revealed that the two new footprints were also present at the same locations (Fig. 11).

To determine if the lack of the -103 to -87 footprint with HepG2 cell nuclear extracts indicates the absence of this binding protein or its presence in low amounts, a gel retention assay, which is more sensitive than the DNase I footprinting assay for detecting DNA-binding proteins, was performed using a synthetic double-stranded oligonucleotide spanning bases -103 to -83 that had been end-labelled. Incubation of this probe with the different nuclear extracts revealed a single retarded band in all the lanes (Fig. 13). The intensity of the retarded band with HepG2 cell nuclear extract suggests that the protein is present in this cell line but at a level that is substantially lower than the level in the other cell lines examined and would explain why a footprint at this binding site could not be detected.

3. DNase I footprint analysis of URE1 deletion mutants

Since the enhancer activity in URE1 appeared to be associated with the protein binding region within the element, DNase I footprint assays were performed on the deletion constructs of the -246 to -81 fragment in which enhancer activity was affected (Fig. 8). The results are shown in Figure 14. As expected, the two deletion constructs between bases -165 and -155 and between

Figure 13 Gel retention analysis of -103 to -83 oligonucleotide.

Gel retention analysis was performed on a synthetic, double-stranded oligonucleotide corresponding to nucleotides -103 to -83 which had been 5' end-labelled. Lane 1, no extract; lane 2, 10 μg of HepG2 nuclear extract; lane 3, 10 μg of CHO nuclear extract; lane 4, 10 μg of HeLa nuclear extract; lane 5, 10 μg of L nuclear extract.

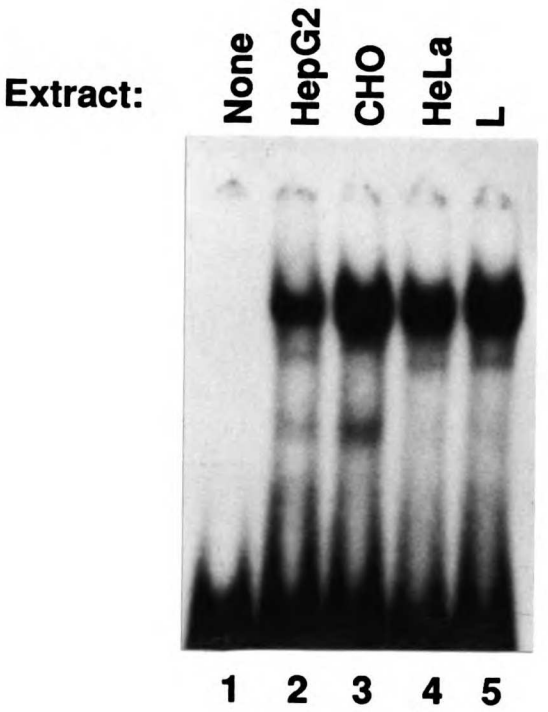
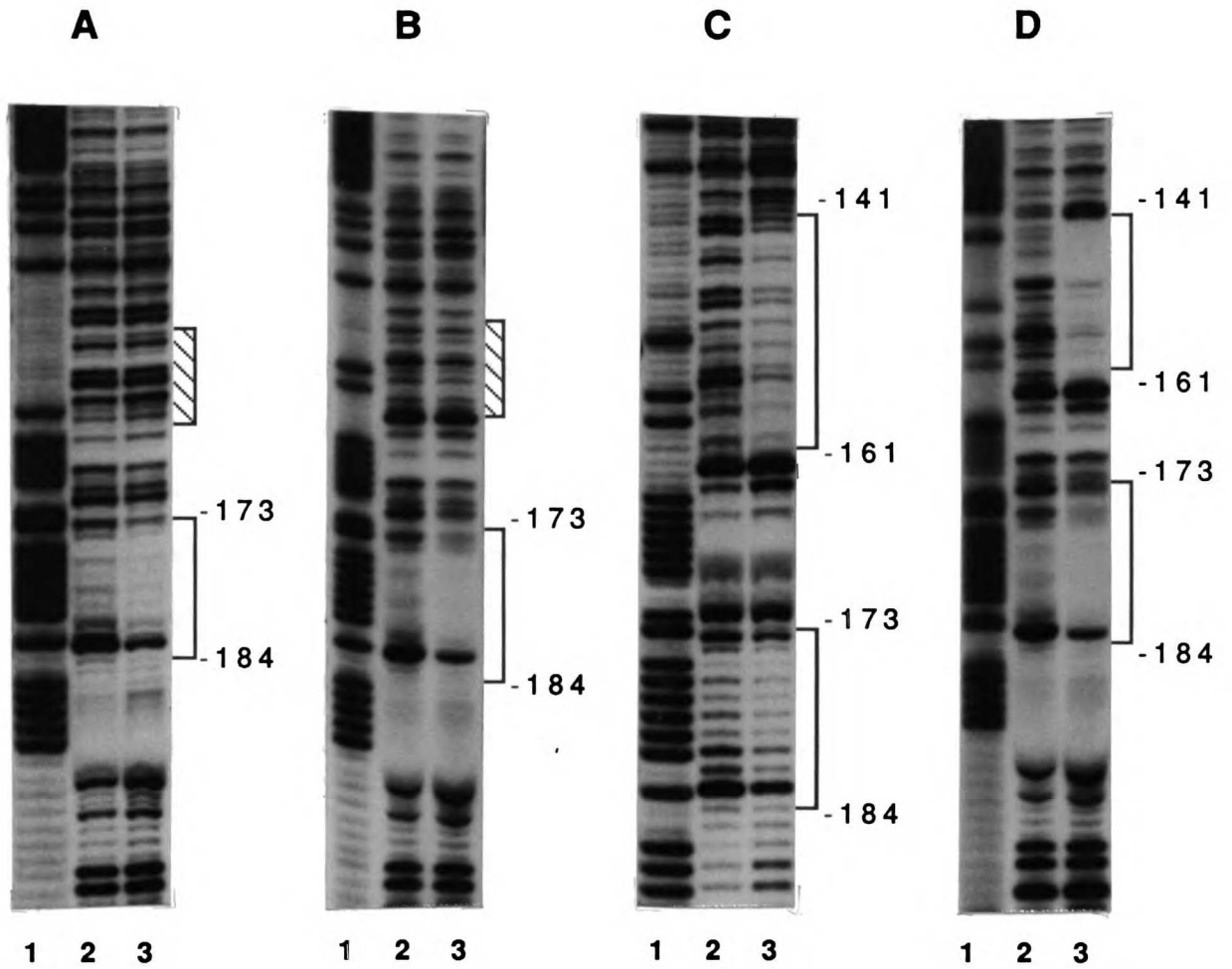

Extract:
None
HepG2
CHO
HeLa
L
1
2
3
4
5

Figure 14 DNase I footprint analysis of URE1 deletion constructs by HeLa nuclear proteins.

Apo-E DNAs with deletions at various regions around and within the URE1 element were 5' end-labelled at position -246 on the 5' strand and subject to DNase I footprint assay. (A) Construct with deletion at nucleotides -165 to -155. (B) Construct with deletion at nucleotides -154 to -144. (C) Construct with deletion at nucleotides -212 to -193. (D) Construct with deletion at nucleotides -123 to -105. For each panel, lane 1 is the G+A sequence ladder, lane 2 is with no extract, and lane 3 is with 36 μg of HeLa nuclear extract. Open brackets surround the sequences protected from DNase I digestion and the nucleotide positions at the boundaries are indicated on the right. Striped brackets surround where the -161 to -141 footprint sequence should be located.


A
B
C
D
173
184
173
184
141
161
173
184
141
161
173
184
1 2 3
1 2 3
1 2 3
1 2 3

bases -154 and -144 no longer displayed the -161 to -141 footprint (Fig. 14 A, B) and no new protected regions were generated. However, in the -212 to -193 and in the -123 to -105 deletions, no changes in the location or intensity of the -161 to -141 footprint were observed (Fig. 14 C, D). Thus the mechanism by which these sequences affect URE1 enhancer activity is unclear. These sequences may be bound by regulatory proteins which cannot be detected under the conditions employed or they could somehow mediate the interaction between the protein(s) at the -161 to -141 sequence and the transcription initiation complex.

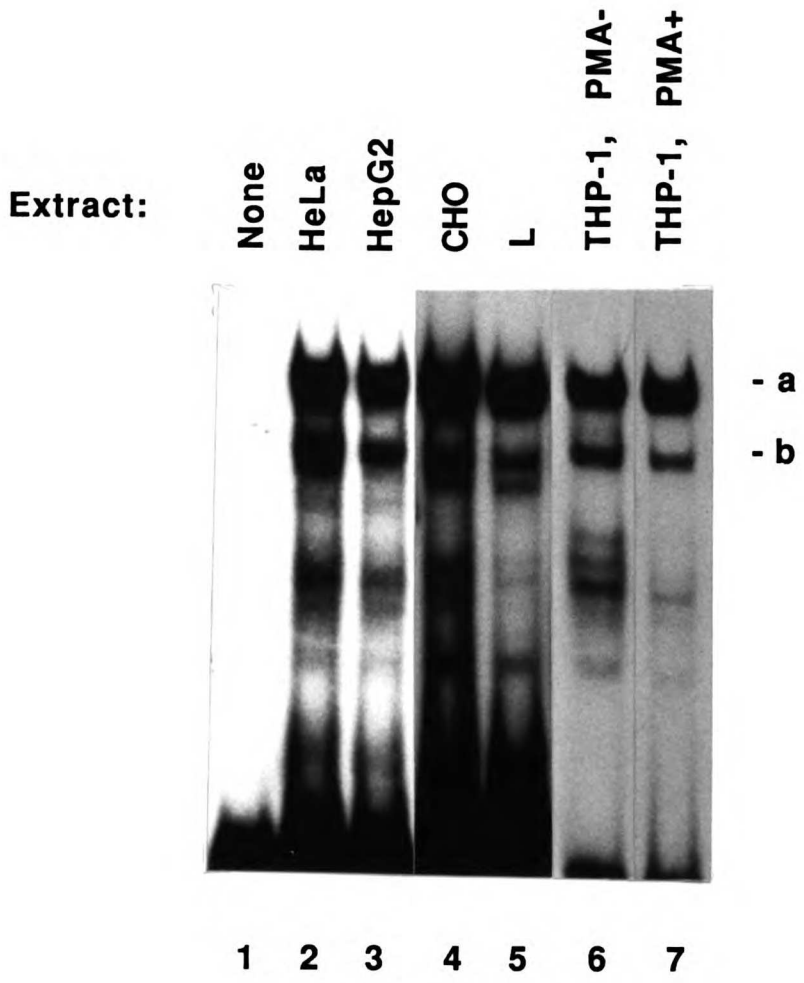
C. Identification and Characterization of the PET-Binding Proteins

1. Gel retention analysis of PET

The synthetic oligonucleotide PET, which spanned bases -169 to -140 and represented the "core" enhancer element within URE1, was analyzed by the gel retention assay to determine if more than one type of protein might be binding to this region and if the sequence might be divided into several protein binding subdomains. When ^{32}P-labelled PET was incubated with nuclear extracts from HepG2, HeLa, CHO and L cells, very similar gel mobility shift patterns were obtained (Fig. 15). Although the intensity varied somewhat, all the extracts generated two prominent retarded bands at the top (referred to as a and b) and a few minor bands at the bottom. Two sets of nuclear extracts that were prepared from the human monocytic cell line, THP-1, that had been treated either with or without the phorbol ester, phorbol 12-myristate 13-acetate (PMA, which induces differentiation of these cells into macrophages) were also examined for PET-binding proteins by gel retention. Once again, the identical retention pattern was observed for both treated and untreated cells (Fig. 15). These experiments suggest that multiple ubiquitous

Figure 15 Gel retention analysis of PET with various nuclear extracts.

5' end-labelled PET was analyzed for the binding proteins by the gel retention assay. In all lanes with nuclear extracts, 10 μg of protein was used. Lane 1, no extract; lane 2, HeLa nuclear extract; lane 3, HepG2 nuclear extract; lane 4, CHO nuclear extract; lane 5, L nuclear extract; lane 6, THP-1 nuclear extract; lane 7, Phorbol 12-myristate 13-acetate (PMA, 160 nM)-treated THP-1 nuclear extract. Different intensities of the bands represent different exposure periods for the autoradiograms and do not represent quantitative differences. The a and b marks indicated on the right refer to the two prominent retained bands.


Extract:
None
HeLa
HepG2
CHO
L
THP-1, PMA-
THP-1, PMA+
- a
- b
1 2 3 4 5 6 7

proteins or multiple forms of an ubiquitous protein are able to bind to the PET sequence.

To study which of the nucleotides within PET are important for the binding of the proteins, a series of double-stranded oligonucleotides that contained various portions of PET was synthesized (Fig. 16). These oligonucleotides then were used as competitors for the end-labelled PET in gel retention assays using HeLa cell nuclear extracts (Fig. 17). When a 100-fold molar excess unlabelled PET (oligonucleotide A in Fig. 16) was added to the reaction, all the retarded bands were reduced substantially in intensity, indicating that the cold PET oligonucleotide was competing with the labelled PET probe for the binding proteins. Although this high amount of competitor oligonucleotide was added to make sure that the binding proteins were removed, usually, a 20-fold molar excess of cold PET oligonucleotide was sufficient to eliminate the retarded bands. When three overlapping oligonucleotides covering the 5' portion, the middle, and the 3' portion of PET were used as competitors (oligonucleotides B, C, D, Fig. 16), only oligonucleotide C was able to compete out the retarded bands, (Fig. 17, lane 5) which suggests that this portion contains the recognition sequences for all the binding proteins. Interestingly, all the competitors were able to generate a novel retarded band (referred to as c) which was not observed in reactions without competitors, and suggested the binding of yet another protein to PET. It would appear that the oligonucleotide competitors somehow were able to activate the band c protein(s) (perhaps by modification) such that it (they) now can bind to the PET sequence. Alternatively, the competitors may have disrupted a "repressor" factor which had prevented the band c protein from binding to PET.

Figure 16 Oligonucleotides used in competition gel retention analysis.

The nucleotides between positions -171 and -138 on the 5' strand are shown and the numbers above the bars represent the positions at the boundaries of the protein-binding region. A, B, C, and D represent four synthetic double-stranded oligonucleotides; the bars represent the sequences spanned by each oligonucleotide; the numbers under the bars represent the start and end positions of each oligonucleotide. The A oligonucleotide is the same as PET.

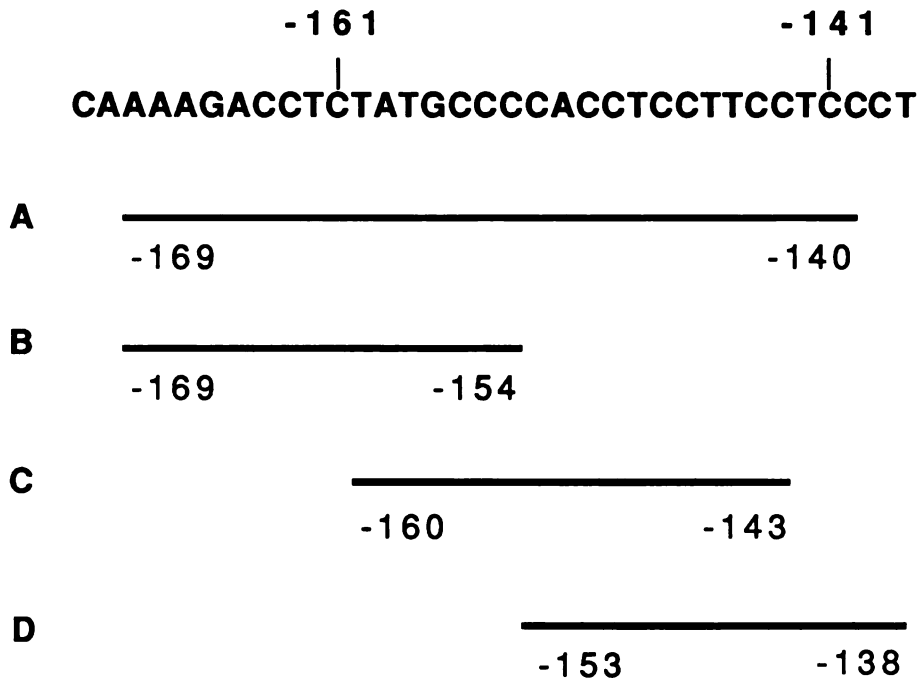
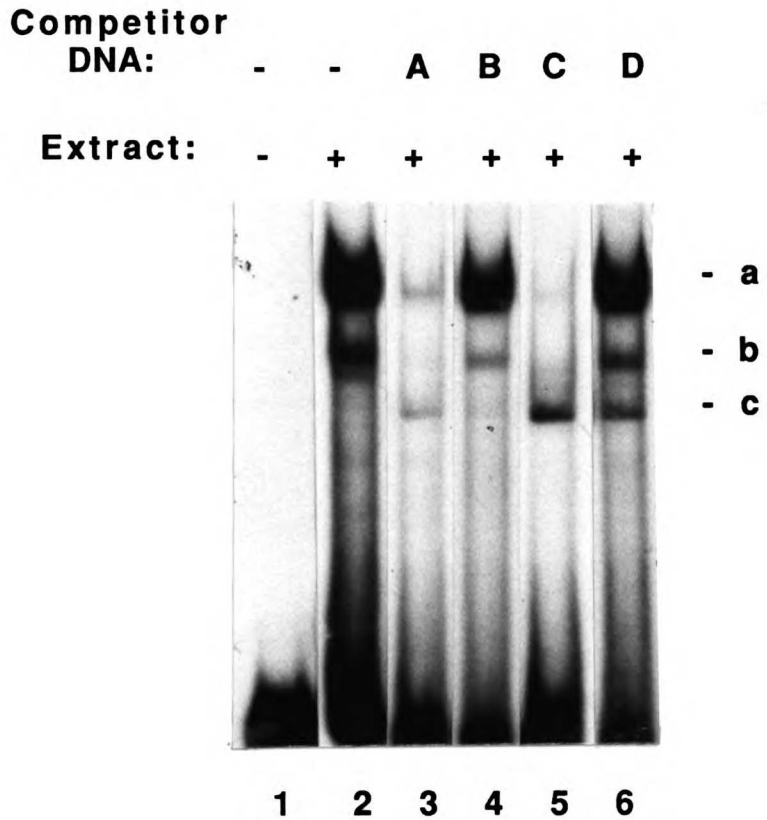

-161
-141
CAAAAGACCTCTATGCCCCACCTCCTTCCTCCCT
A
-169
-140
B
-169
-154
C
-160
-143
D
-153
-138

Figure 17 Gel retention analysis of PET with oligonucleotide competitors and HeLa nuclear proteins.

5' end-labelled PET was incubated with no extract (lane 1) or with 10 μg of HeLa nuclear extract (lanes 2-7). In the lanes with extract, the following oligonucleotide competitors (100-fold molar excess) were included: lane 2, no competitor; lane 3, oligonucleotide A (PET); lane 4, oligonucleotide B; lane 5, oligonucleotide C; lane 6, oligonucleotide D. Refer to figure 16 for the nucleotides spanned by each oligonucleotide competitor. a and b refer to the two upper retained band. c refers to the novel retained band which appear in the presence of competitors.


Competitor DNA: - - A B C D
Extract: - + + + + +
- a
- b
- c
1 2 3 4 5 6

To study the specificity of the PET-binding proteins, gel retention analyses were performed in different salt concentrations and in the presence of different concentrations of the carrier DNA, poly (dI-dC)/poly (dI-dC). Figure 18 shows that as the concentration of KCl increased, the intensity of bands a and b decreased, whereas the bottom diffuse bands remained unchanged. Most noticeable was the appearance of the novel band c which was quite prominent at 200 mM KCl. The results of the poly (dI-dC)/poly (dI-dC) titration experiment are shown on Figure 19. In the absence of the carrier DNA, band c appears along with another intense new retarded band (indicated by the arrow), probably caused by a non-specific DNA-binding protein. With increasing concentrations of poly (dI-dC)/poly (dI-dC), the novel bands and the bottom bands disappeared and bands a and b became most prominent. These two experiments indicated that the protein(s) representing the two retarded bands, a and b, are more sequence-specific but less tightly bound to PET than the other protein(s). Also, there was tremendous variation in the intensities of the bottom diffuse bands from experiment to experiment, suggesting that these bands were caused by non-specific proteins binding to the PET sequence.

2. Identity of the PET-binding proteins

A close examination of the PET element revealed that it contained a sequence spanning bases -157 to -148 in which nine of ten nucleotide positions correspond to the consensus sequence for Sp1 binding (Table III). To determine if Sp1 might be one of the PET-binding proteins, an oligonucleotide spanning the proximal GC box element in the apo-E promoter (nucleotides -59 to -45) was synthesized and used as a competitor in the gel retention assay with labelled PET. This GC box oligonucleotide was able to eliminate bands a and b (Fig. 20, lane 2), suggesting that these bands were caused by the binding of Sp1

Figure 18 Gel retention analysis of PET in different KCl concentrations.

5' end-labelled PET was incubated with 10 μg of HeLa nuclear extract in 50 mM KCl (lane 1), 100 mM KCl (lane 2), 150 mM KCl (lane 3), or 200 mM KCl (lane 4). Markings on the right are as described above.

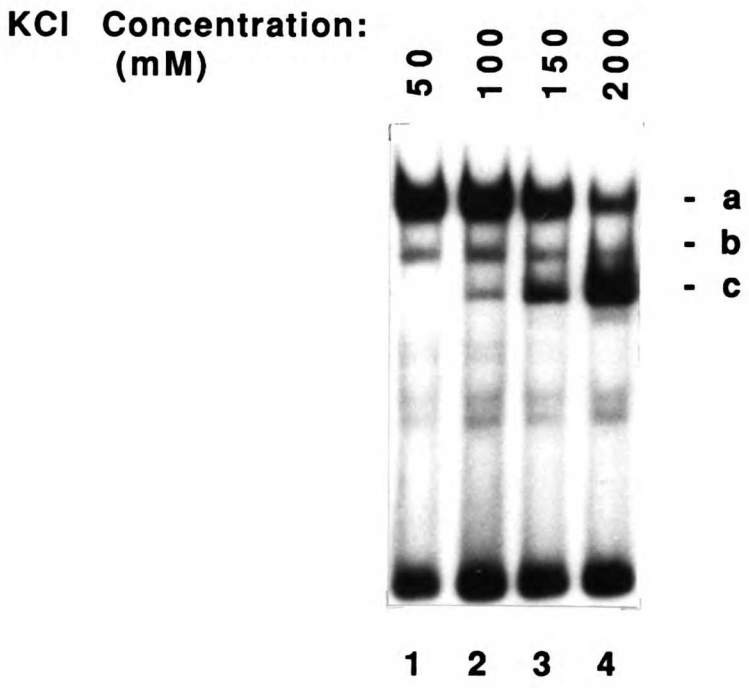

KCl Concentration:
(mM)
50
100
150
200
- a
- b
- c
1
2
3
4

Figure 19 Gel retention analysis of PET in different poly (dI-dC)/poly (dI-dC) concentrations.

5' end-labelled PET was analyzed by the gel retention assay with HeLa nuclear extract (10 μg) under different conditions. Lane 1, no poly (dI-dC)/poly (dI-dC) added. Lanes 2-6: poly (dI-dC)/poly (dI-dC) added: lane 2, 0.1 μg; lane 3, 0.4 μg; lane 4, 1 μg; lane 5, 4 μg; lane 6, 9 μg. Lane 7, 4 μg poly (dI-dC)/poly (dI-dC) with 100-fold molar excess of proximal GC box (-59 to -45)-spanning oligonucleotide. Lane 8, no extract.

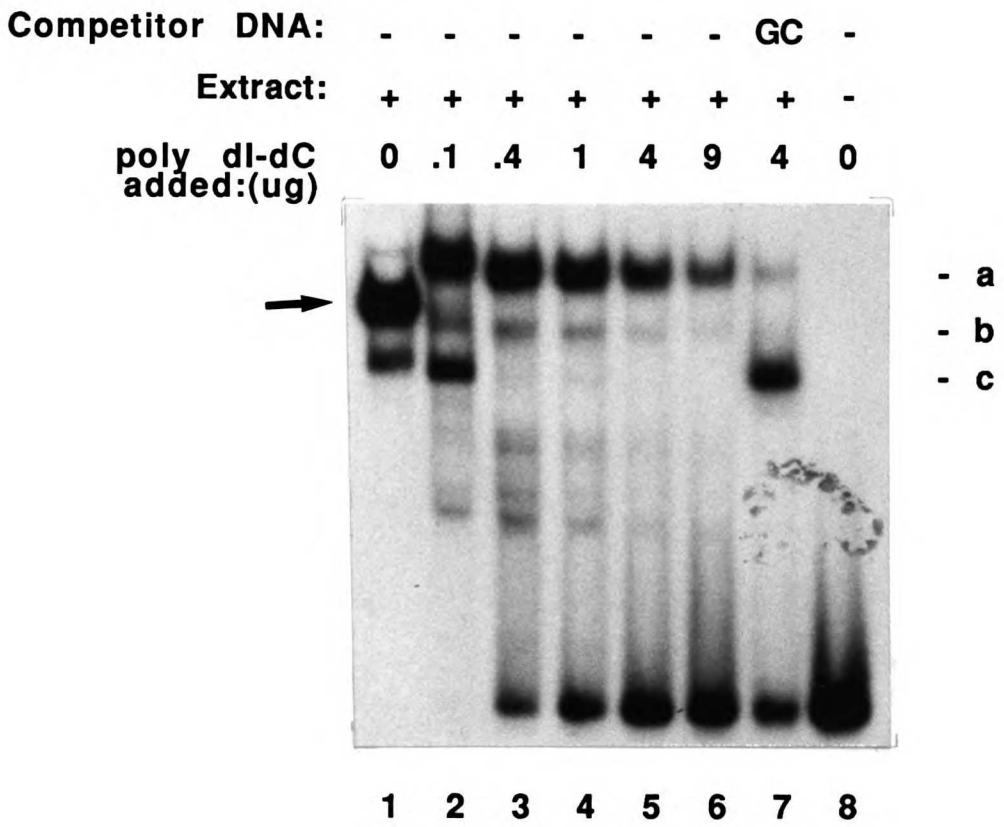

Competitor DNA: - - - - - - GC -
Extract: + + + + + + + -
poly dI-dC added:(ug) 0 .1 .4 1 4 9 4 0
- a
- b
- c
1 2 3 4 5 6 7 8

Table III

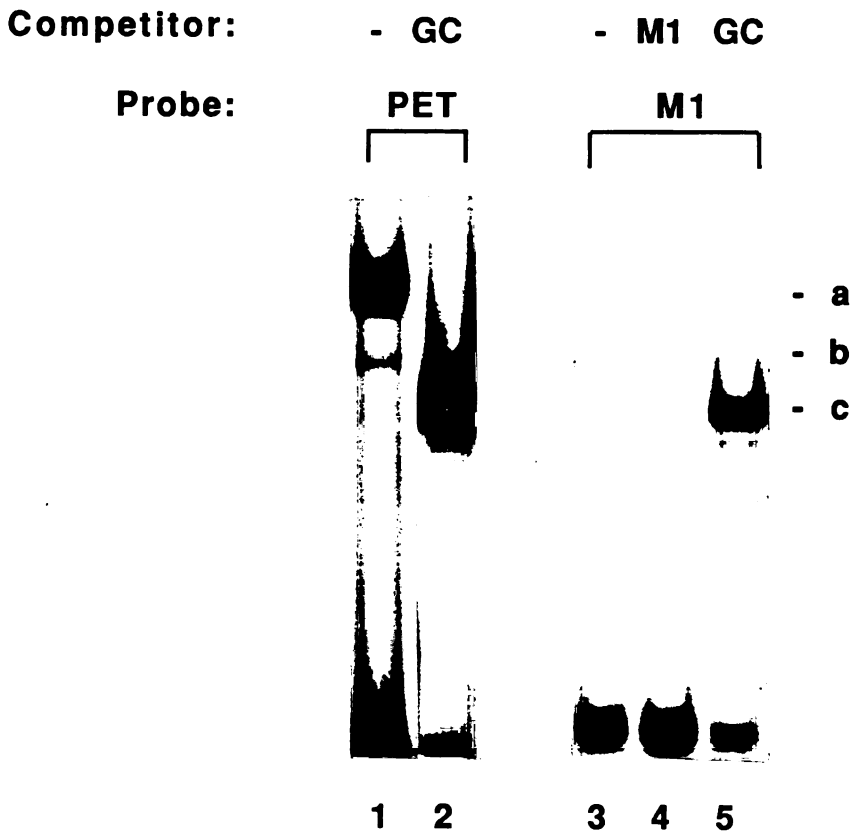
Sp1 binding sequences in the apo-E promoter and the substituted nucleotides of the mutant PET oligonucleotides

Overall Consensus	5'	G G G G C G G G G C T A A T A A T	3'
Proximal GC Box -54 to -45		G G G G C G G G A C	
PET Element -148 to -157		G A G G T G G G G C	
Mutant PET M1		G A G T T G G G G C	
Mutant PET M2		G A G G T T G G G C	
Mutant PET M3		G A G G T T T G G C	
Mutant PET M4		G A G T T T T T G C	
-181 to -172		G G G G G G T G G T	

For the proximal GC box element, and the -181 to -172 sequence, nucleotides on the 5' strand are shown while for the PET element and the mutant PET sequences, nucleotides on the 3' strand are shown. The asterisks represent the positions where bases substitutions were introduced in the mutant PET oligonucleotides.

Figure 20 Gel retention analysis of PET and mutant PET.

5' end-labelled PET (lanes 1-2) or mutant PET with a C to A substitution at base position -151 (lanes 3-5), known as mutant M1 in Table III, were analyzed by the gel retention assay in standard conditions with HeLa nuclear extract. Lane 1, 10 μg HeLa extract; lane 2, 10 μg HeLa extract and 100-fold molar excess of proximal GC box (-59 to -45) oligonucleotide; lane 3, 10 μg HeLa extract; lane 4, no extract; lane 5, 10 μg HeLa extract and 100-fold molar excess of proximal GC box oligonucleotide.


Competitor:
- GC
- M1 GC
Probe:
PET
M1
- a
- b
- c
1 2
3 4 5

or a Sp1-like protein to PET. The appearance of a prominent band c was detected also. To test further for the possibility of Sp1 binding to PET, a mutant PET oligonucleotide was synthesized, with a single substitution at a nucleotide in the GC box homology portion (C to A at nucleotide -151, mutant M1 in Table III), known to be critical for recognition by Sp1 (104). When the mutant PET was labelled and analyzed by gel retention, bands a and b were barely detectable (Fig. 20, lane 3). However, this mutant PET probe was still capable of binding to the protein(s) responsible for band c (Fig. 20, lane 5).

To demonstrate definitively that Sp1 binds to PET, a >90% pure preparation of Sp1 was obtained from Drs. Stephen Jackson and Robert Tjian. Gel retention analysis of PET with pure Sp1 revealed the identical a and b gel retention bands, as well as some minor bands (Fig. 21, lanes 1, 4). The proximal GC box oligonucleotide completely obliterated the Sp1 bands (Fig. 21, lane 5), which demonstrated that Sp1 does bind to the proximal GC box element. Furthermore, band c was not detected, suggesting that this band was caused by the binding of a protein in the crude nuclear extract to the PET sequence, and was not an artifact caused by the addition of excess unlabelled GC box oligonucleotide. To confirm the gel retention results, a DNase I footprint assay was performed with the -212 to -15 fragment using the pure preparation of Sp1. The results demonstrate that the -161 to -141 footprint was identical in size and position to that obtained with crude nuclear extracts and that Sp1 alone was capable of protecting this sequence from digestion (Fig. 22). The pure Sp1 preparation was able to produce the footprint between nucleotides -184 to -173, suggesting that this sequence also was recognized by Sp1. To test this possibility further, a double-stranded oligonucleotide corresponding to nucleotides -185 to -169 was synthesized, labelled and tested by the gel retention analysis to identify the binding protein(s). Once again, the two

Figure 21 Gel retention analysis of PET with pure Sp1.

5' end-labelled PET was analyzed by the gel retention assay under standard conditions. Lane 1, 10 μg HeLa nuclear extract; lane 2, 10 μg HeLa extract with proximal GC box oligonucleotide; lane 3, 10 μg HeLa extract with oligonucleotide A (PET); lane 4, 10 ng of > 90 pure Sp1; lane 5, 10 ng of pure Sp1 with proximal GC box oligonucleotide.

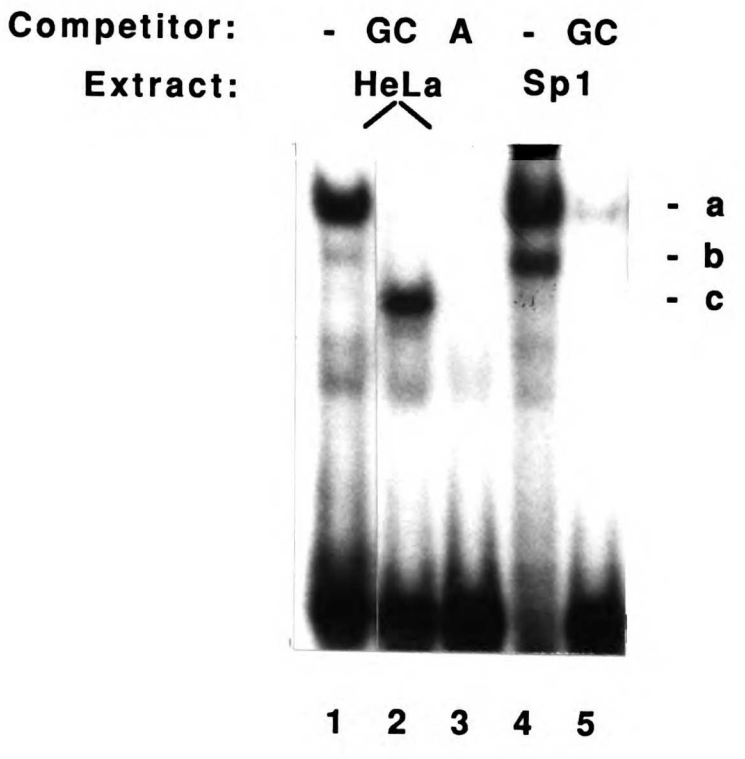
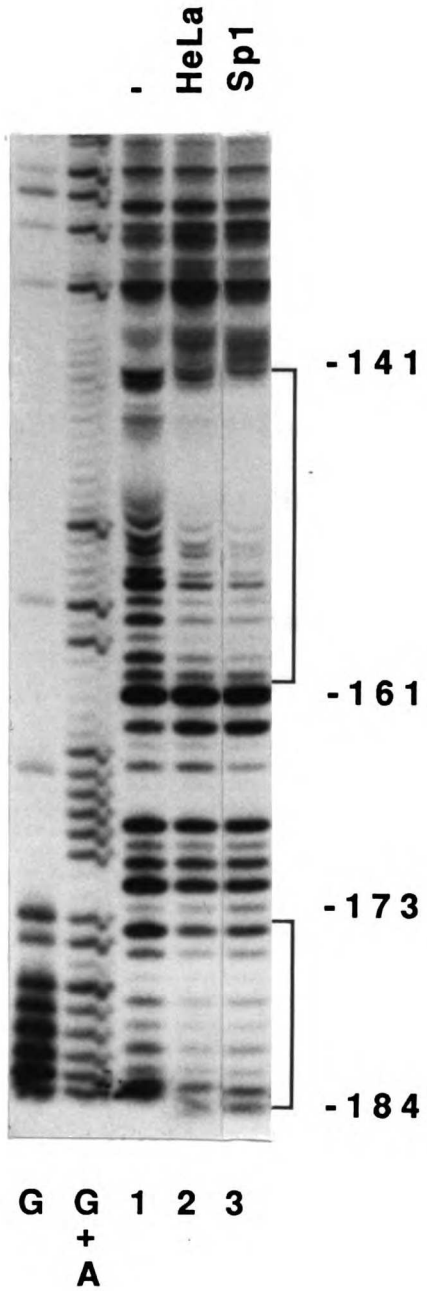

Competitor: - GC A - GC
Extract: HeLa Sp1
- a
- b
- c
1 2 3 4 5

Figure 22 DNase I footprint analysis of the proximal 5' flanking sequence of the apo-E gene with pure Sp1.

DNase I footprint analysis was performed on apo-E DNA that had been 5' end-labelled at position -212 on the 5' strand. The G and G+A lanes represent the respective Maxam and Gilbert sequencing reactions. Lane 1, no nuclear extract; lane 2, 36 μg of HeLa nuclear extract; lane 3, 20 ng of pure Sp1. Sequences protected from DNase I digestion are indicated by the brackets, with the nucleotide positions at the boundaries indicated on the right.


-
HeLa
Sp1
-141
-161
-173
-184
G
G + A
1
2
3

retarded bands, a and b characteristic of Sp1 binding were detected along with several minor bands (Fig. 23).

To assess the possible requirement of Sp1 to produce the enhancer activity of the PET element, mutant PET oligonucleotide M1 (Table III) which bound Sp1 very poorly but still showed strong binding for the band c protein (Fig. 20, lane 5) was inserted into the pTK1 vector and tested for enhancer activity. The results are shown in Table IV. The PET element, once again, stimulated TK promoter activity by fourfold. However, the modified PET displayed no enhancer activity and actually reduced TK promoter activity slightly. This indicated that Sp1-binding was absolutely necessary for the PET element to exhibit enhancer activity.

Although Sp1 appeared to be the dominant nuclear protein that bound to the -161 to -141 sequence, and confer enhancer activity to the PET sequence, the identity and significance of the protein responsible for the formation of band c in gel retention analyses was unknown. Thus a series of gel retention assays was performed with labelled PET in the presence of the GC box oligonucleotide in order to detect band c. First, when nuclear extracts from different cell lines were tested, band c was detected in HeLa, HepG2, and PMA-treated and untreated THP-1 cells but not in CHO and L cells (Fig. 24). Since URE1 functioned as an enhancer in CHO and L cells (Table II) and the enhancer activity of the PET sequence had been detected in CHO cells (Fig.9, Table IV), the absence of the band c protein in these cell lines suggests that this protein is not required for URE1 and PET enhancer activity. Next, several double-stranded oligonucleotides were added in the reactions in 100-fold molar excess in order to examine competition against band c. Of the three oligonucleotides containing portions of the PET sequence (Fig. 16), only oligonucleotide C (-160 to -143) showed a significant reduction in the intensity

Figure 23 Gel retention analysis of -185 to -169 oligonucleotide.

Gel retention analysis was performed on a synthetic, double-stranded oligonucleotide corresponding to nucleotides -185 to -169, which had been 5' end-labelled. Lane 1, no extract; lane 2, 10 μg of HeLa nuclear extract; lane 3; 10 μg of HeLa extract with 100-fold molar excess of unlabelled -185 to -169 oligonucleotide; lane 4; 10 μg of HeLa extract with 100-fold excess of proximal GC box oligonucleotide.

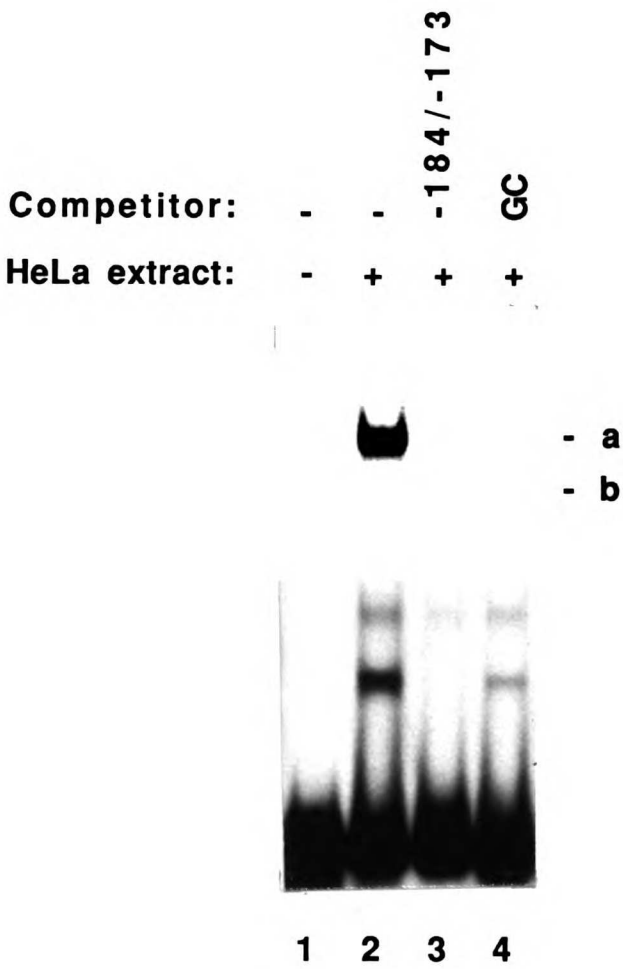

Competitor: - - -184/-173 GC
HeLa extract: - + + +
- a
- b
1 2 3 4

Figure 24 Characterization of band c protein by gel retention analysis with various nuclear extracts.

Gel retention analysis was performed with the 5' end-labelled PET oligonucleotide in the presence of 100-fold molar excess of the proximal GC box oligonucleotide. All lanes were incubated with 10 mg of nuclear extracts from different cell lines. Lane 1, HeLa cells; lane 2, HepG2 cells; lane 3, THP-1 cells; lane 4 PMA-treated THP-1 cells; lane 5, CHO cells; lane 6, mouse L cells.

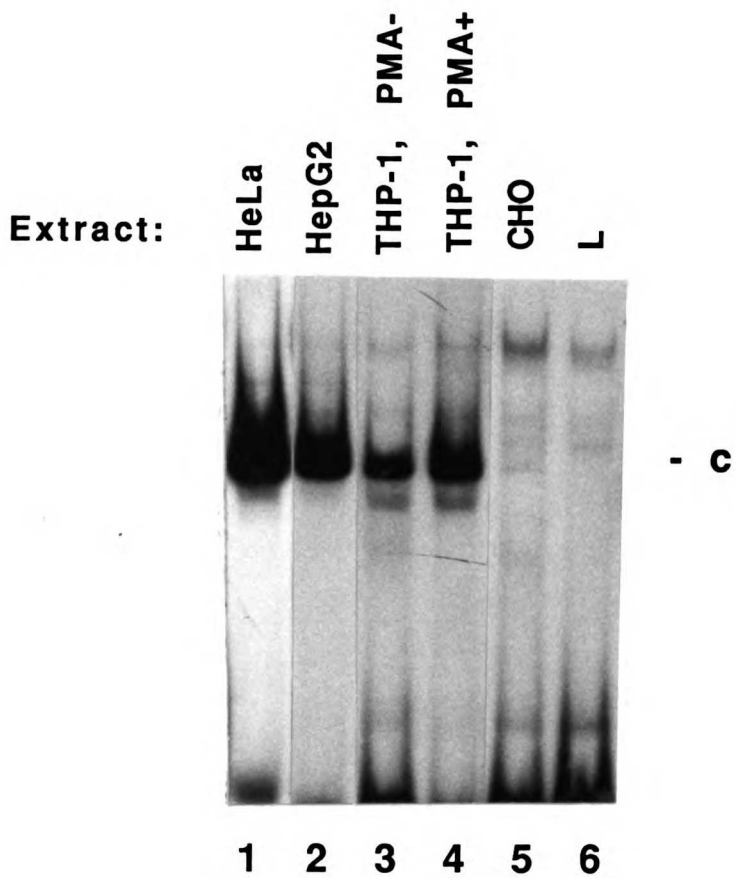

Extract:
HeLa
HepG2
THP-1, PMA-
THP-1, PMA+
CHO
L
- c
1 2 3 4 5 6

Table IV

Enhancer activity of selected oligonucleotides in CHO cells

DNA FRAGMENT	ENHANCEMENT -fold
pTK1	1.0
PET	4.0 ± 0.4
Mutant PET M1	0.7 ± 0.1

The results represent the mean ± S.D. from two independent experiments. Enhancement is expressed as -fold stimulation over the control, pTK1. All the inserts were positioned in the normal orientation relative to the TK promoter.

of band c (Fig. 25, lane 4). This suggests that the recognition sequence for the band c protein either is close to or overlaps the GC box homology region (-157 to -148) within PET. To further localize the recognition sequence, modified PET oligonucleotides in which one, two or four nucleotides between bases -155 and -151 had been substituted (Table III, mutants M1, M2, M3 and M4) were tested as competitors against band c. The results (Fig. 25, lanes 6-9) showed that all the mutant oligonucleotides were able to compete out the retarded band. This finding indicates that either the "critical" nucleotides required for band c protein recognition lie outside of bases -155 to -151, or this protein binds to some feature in PET and in the -160 to -143 oligonucleotide which is independent of the actual nucleotide sequence. When oligonucleotides corresponding to the binding sequence of several known transcription factors were tested, only an AP-2-binding oligonucleotide from the SV40 enhancer element (66) was able to compete out band c (Fig. 25, lane 11). To determine if the band c protein might be AP-2, a pure preparation of this transcription factor was obtained (generous gift of Drs. Trevor Williams and Robert Tjian), and assayed by gel retention. The results (Fig. 26) showed that AP-2 bound poorly to PET and formed a retarded band which migrated slightly slower than band c. Therefore, although the band c protein is able to bind to an AP-2 binding sequence, it is definitely not AP-2. Furthermore, recent reports have stated that AP-2 is not detected in HepG2 cells (105, 106), which contrasts with the finding of band c in HepG2 cell nuclear extracts (Fig. 24, lane 2).

3. Isolation of the PET-binding proteins

To determine its functional significance by the assays that are described below, the band c protein was purified from HeLa cells. HeLa cells were employed because the band c protein is abundant in this cell line (Fig. 24, lane 1), and because the purification of several transcription factors had been done

Figure 25 Gel retention analysis of PET with various oligonucleotide competitors for the identification of band c protein.

Gel retention analysis was performed on 5' end-labelled PET under standard conditions. All lanes contained 10 μg of HeLa nuclear extract. Lanes 2-14 also contained 100-fold molar excess of the proximal GC box oligonucleotide (-59 to -45). Lane 1, HeLa extract only; lane 2, with GC box oligonucleotide; lanes 3-6: with GC box oligonucleotide and PET-overlapping oligonucleotides (Fig. 16); lane 3, B; lane 4, C; lane 5, D; lanes 6-10: with GC box oligonucleotide and mutant PET oligonucleotides (Table III); lane 6, mutant M1; lane 7, mutant M2; lane 8, mutant M3; lane 9, mutant M4; lanes 10-14: with GC box oligonucleotide and synthetic oligonucleotides with the recognition sequence for : lane 10, AP-1 (from SV40 enhancer element, ref. 65); lane 11, AP-2 (from SV40 enhancer element, ref. 66); lane 12, NF-A1 (from human immunoglobulin κ light chain promoter, ref. 67); lane 13, CTF (from HSV-TK promoter, ref. 94); lane 14, Sp1 (from SV40 promoter, ref. 104). All synthetic oligonucleotides used as competitors were added in 100-fold molar excess over the end-labelled PET probe.

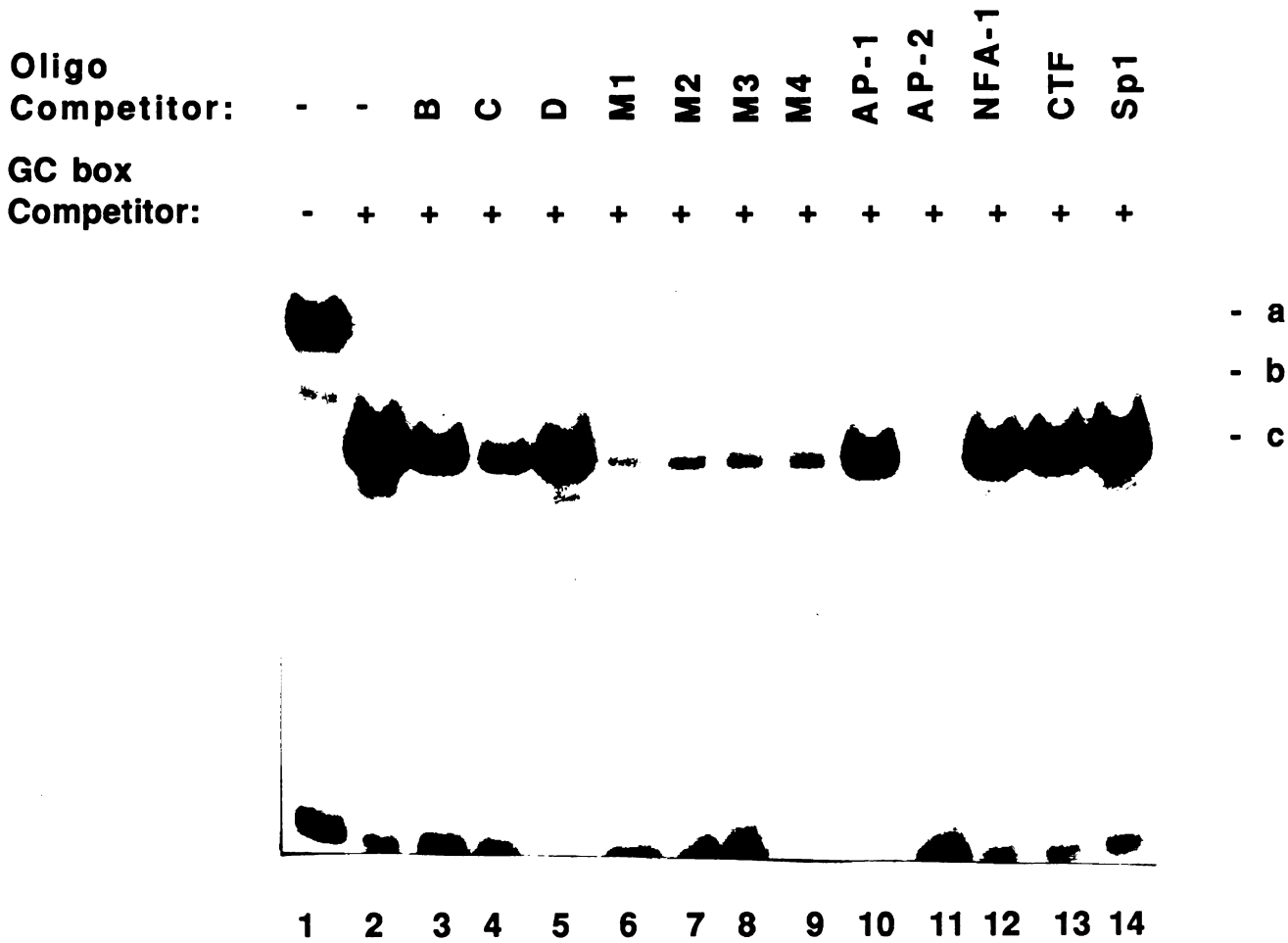
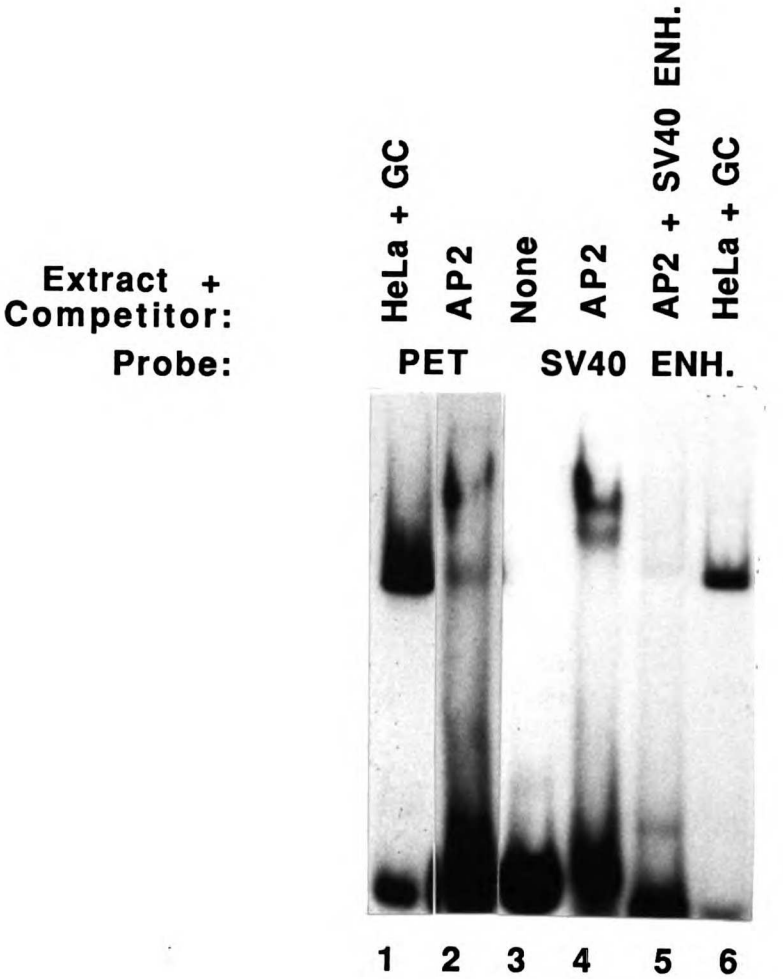

Oligo Competitor: - - B C D M1 M2 M3 M4 AP-1 AP-2 NFA-1 CTF Sp1
GC box Competitor: - + + + + + + + + + + + + +
- a
- b
- c
1 2 3 4 5 6 7 8 9 10 11 12 13 14

Figure 26 Gel retention analysis with pure AP-2.

Gel retention analysis was performed with 5' end-labelled PET (lanes 1-2) or with 5' end-labelled AP-2 binding-oligonucleotide from the SV40 enhancer (lanes 3-6). Lane 1, 10 μg HeLa extract and 100-fold molar excess GC box oligonucleotide; lane 2, 10 ng pure AP-2; lane 3, no extract; lane 4, 10 ng pure AP-2; lane 5, 10 ng pure AP-2 and 100-fold molar excess AP-2 binding oligonucleotide; lane 6, 10 μg HeLa extract and 100-fold molar excess GC box oligonucleotide.


Extract +
Competitor:
Probe:
HeLa + GC
AP2
None
AP2
AP2 + SV40 ENH.
HeLa + GC
PET
SV40 ENH.
1 2 3 4 5 6

using HeLa cells (65, 66, 99, 105, 110). Nuclear extracts from HeLa cells grown in suspension culture were applied to a Sephacryl S-300 gel filtration column and the fractions were assayed for PET-binding activity by the gel retention analysis. Figure 27A shows the results of the assay performed on one set of S-300 fractions. The prominent appearance of the lower retarded bands in all the fractions, suggested that they were caused by an abundant and non-specific DNA-binding protein. The identical set of S-300 fractions was also tested by gel retention in the presence of the GC box oligonucleotide in order to visualize band c (Fig. 27B). The peak fractions of Sp1 overlapped those of the novel protein and these fractions were pooled and applied over a DNA affinity column consisting of concatemerized PET oligonucleotide. The column then was washed and the bound material eluted with stepwise increments of KCl from 0.2 M to 0.8 M concentration. The elution fractions were assayed by gel retention analysis with PET and the peak fractions of PET-binding activity were pooled, diluted and reapplied to the affinity column. After three consecutive passages over the affinity column, the purified proteins were visualized by SDS polyacrylamide gel electrophoresis. A number of bands were detected, including Sp1 which showed up as 105 kD and 95 kD polypeptides (Fig. 28, lane 2). This fraction then was incubated with the proximal GC box oligonucleotide and re-applied to the affinity column. This time, SDS gel electrophoresis of the eluted fraction revealed a 55 kD species which corresponds, most likely, to the novel PET-binding protein which produces band c (Fig. 28, lane 3).

D. Functional Analysis of Regulatory Elements by In Vitro Transcription

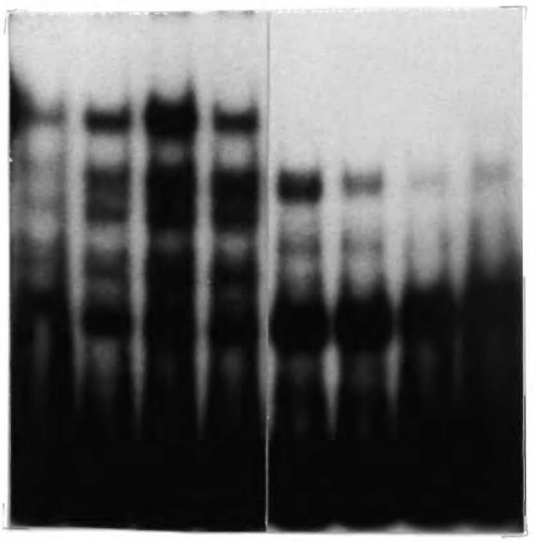
1. Template and competition studies

In vitro transcription assays were performed to examine the significance of the protein-binding sequences in the 5' flanking region on

Figure 27 Gel retention analysis of PET with Sephacryl S-300 fractions.

HeLa nuclear extract (approximately 50 mg) was loaded onto a 100 ml bed volume-Sephacryl S-300 column and passaged at a flow rate of 5 ml/hr. Fractions were not collected until the dead volume (15 ml) had passed through and then 2 ml-fractions were collected. The protein concentration of each fraction was determined by the Bradford assay (91). (A) Gel retention analysis with PET was performed, using 4 μg of protein from each fraction in the presence of 50 ng/μl poly (dI-dC)/poly (dI-dC). (B) S-300 fractions were analyzed by gel retention as above except 100-fold molar excess of the proximal GC box oligonucleotide was included in each lane in order to detect band c.

A


- a
- b
10 12 14 16 18 20 22 24
Fraction number

B

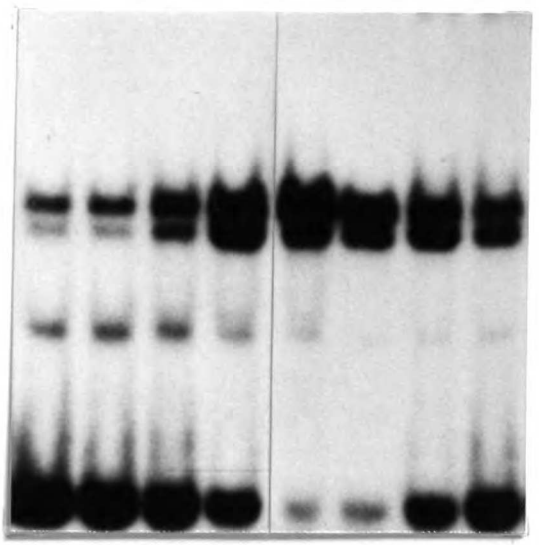
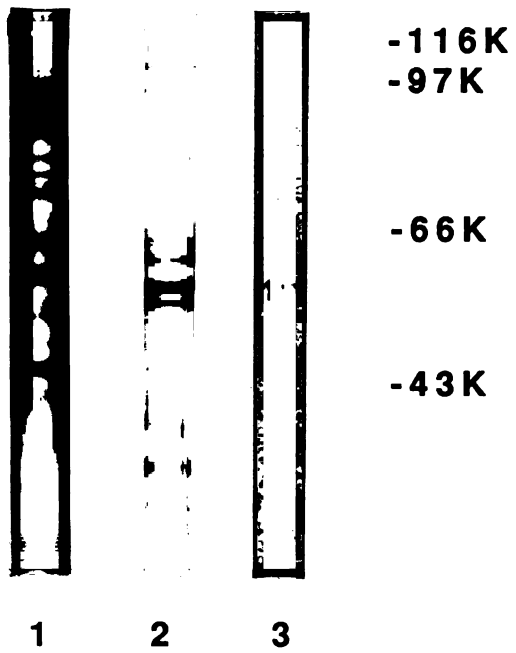

- c
10 12 14 16 18 20 22 24
Fraction number

Figure 28 Affinity-column purification of 55 kD protein.

SDS-polyacrylamide (10%) gel electrophoresis was performed on fractions at different stages of purification and visualized by silver staining. Lane 1, 2 μg of protein from pooled Sephacryl S-300 fractions with PET-binding activity; lane 2, 100 ng of protein eluted from third affinity column passage; lane 3, 20 ng of protein from the fourth affinity column passage which was preceded by incubation with the GC box oligonucleotide. The numbers on the right represent molecular sizes in kilodaltons, corresponding to the molecular size markers.


-116K
-97K
-66K
-43K
1
2
3

apo-E gene transcription, as well as to confirm some of the results obtained from the gene transfer experiments. Three apo-E promoter/CAT gene constructs which contained different lengths of the apo-E promoter (383, 212 and 81, Fig. 2) were incubated with transcriptionally active HeLa cell nuclear extracts to generate transcripts which were examined by primer extension. The pRSVCAT plasmid which contained nucleotides -494 to +30 of the RSV promoter next to the CAT gene (89) was included in each reaction as an internal control of transcriptional efficiency. There was virtually no difference in the intensity of the apo-E promoter-generated transcript band between the -383 and the -212 template (Fig. 29, lanes 1, 3). However, the -81 template which contains only the proximal GC box element generated transcripts at a level that was fourfold lower than that of the two other apo-E promoter/CAT templates (Fig. 29, lane 2). This latter result is in agreement with the results from the transfection experiment using the same promoter-deletion constructs (Fig. 2) and would suggest that the protein-binding sequences between nucleotides -212 and -81 also play a significant role in the *in vitro* transcription of the apo-E gene.

In a separate experiment, the -212 apo-E promoter/CAT and the -81 apo-E promoter/CAT templates were incubated with active HeLa extract along with 100-fold molar excess of double-stranded oligonucleotides. Figure 30 shows that excess PET oligonucleotide was able to reduce transcription from both templates (lanes 2, 6), whereas a non-protein binding oligonucleotide had no effect (lanes 3, 7). These results demonstrate the functional importance of Sp1 binding not only to the GC box element but also to the PET sequence.

2. Deletion and substitution studies

Mutant -383 apo-E promoter/CAT templates in which protein-binding sequences had been either deleted or base substituted by site-directed

Figure 29 *In vitro* transcription analysis of different apo-E promoter/CAT gene templates.

RNA synthesized *in vitro* from 100 ng of apo-E promoter/CAT gene templates and from 100 ng of pRSVCAT template using 100 μg of transcriptionally active HeLa nuclear extracts was primed with a CAT gene oligomer (bases 15-34 of coding sequence). The positions of primer extension products of RNA initiating at the transcription start sites from the apo-E promoter (ApoEp, 150 bp in length) and the RSV promoter (RSVp, 107 bp) are indicated. Lane 1, template with 212 bases of the apo-E promoter next to the CAT gene; lane 2, template with 81 bases of the apo-E promoter; lane 3, template with 383 bases of the apo-E promoter. The numbers indicated on the right represent the length in nucleotides as determined from a pBR 322 size marker.

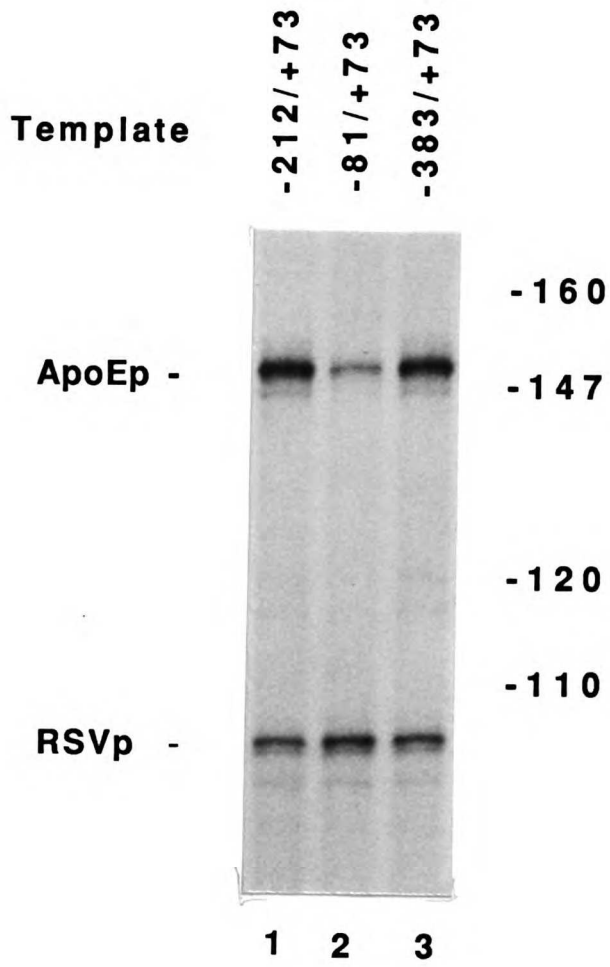
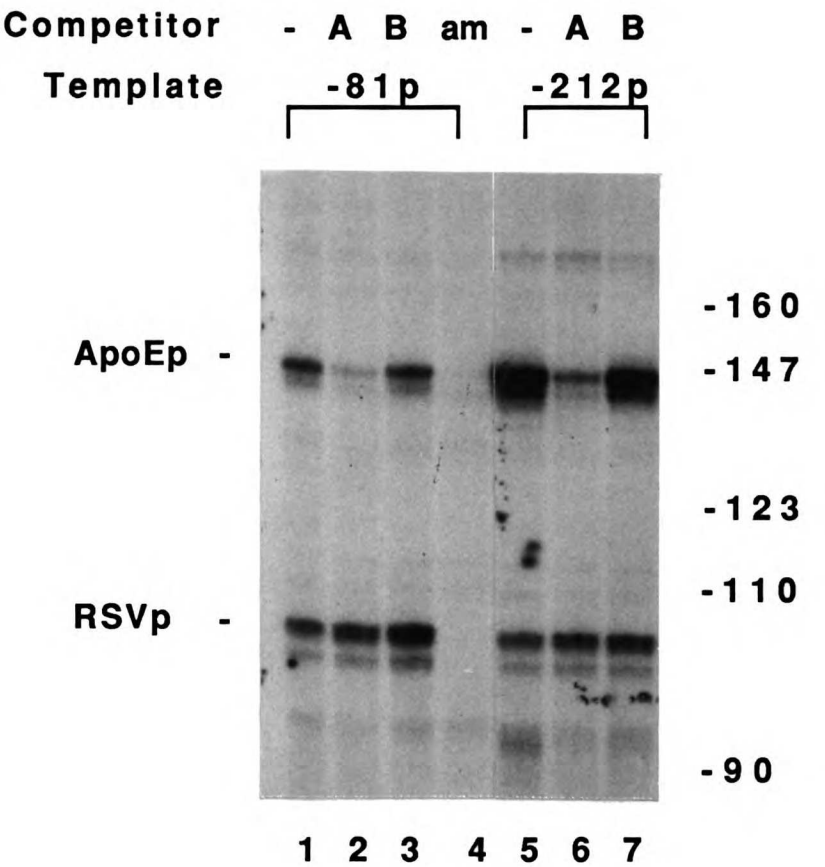

Template
-212/+73
-81/+73
-383/+73
ApoEp -
RSVp -
-160
-147
-120
-110
1
2
3

Figure 30 *In vitro* transcription analysis in presence of oligonucleotide competitors.

Reaction conditions, the positions of the primer extension products and the length markers are as described in the legend to figure 29. In lanes 1-4, the -81 apo-E promoter/CAT template was used and in lanes 5-7, the -212 apo-E promoter/CAT template was used. Lane 1, no competitor; lane 2, 60 ng of PET oligonucleotide (A in figure 16) added; lane 3, 60 ng of oligonucleotide B (figure 16) added; lane 4, 50 ng of the RNA polymerase II inhibitor, α-amanitin added; lane 5, no competitor; lane 6, 60 ng of PET oligonucleotide added; lane 7, 60 ng of oligonucleotide B added.


Competitor - A B am - A B
Template -81p -212p
ApoEp
RSVp
-160
-147
-123
-110
-90
1 2 3 4 5 6 7

mutagenesis were assayed by *in vitro* transcription. The most profound effect was observed when the -161 to -141 footprint sequence had been deleted, which caused a reduction in gene transcription rates of approximately sevenfold (Fig. 31, lane 4). An apo-E promoter construct with a C to A substitution at nucleotide position -151 which dramatically reduced the binding of Sp1 to the PET sequence (Fig. 20, lane 3) showed a threefold loss in transcriptional activity (Fig. 31, lane 5). The deletion of the footprint sequence spanning the proximal GC box element at -59 to -45 reduced the transcript level by approximately 60%, similar to the effect of removing the footprint sequence at -103 to -83 (Fig. 31, lanes 6, 7). In contrast, deletion of the Sp1-binding -184 to -173 sequence had no effect on transcriptional activity (Fig. 31, lane 3).

3. Studies with the Δ GC box template

In order to distinguish the transcriptional effect of Sp1 binding to the proximal GC box from its binding to the PET element, *in vitro* transcription experiments were performed using the -383 apo-E promoter/CAT template in which the GC box element had been deleted. In the first experiment, several of the double-stranded oligonucleotides were added as competitors. As expected, both the PET oligonucleotide and the GC box oligonucleotide reduced the transcript signal by approximately fourfold (Fig. 32, lanes 3, 4).

To determine the individual contributions by Sp1 and the 55 kD protein on the stimulatory activity of PET, a reconstitution experiment was performed with HeLa extracts that had been depleted of PET-binding proteins by adsorption to the PET-affinity Sepharose 4B resin. The apo-E promoter-generated transcripts from the depleted extract was approximately 30% that of the control HeLa extract (Fig. 33, lanes 1, 2). Purified protein fractions from HeLa cell nuclear extracts were assayed for their ability to stimulate the

Figure 31 *In vitro* transcription analysis of deletion and substitution constructs of the -383 apo-E promoter/CAT gene template.

Reaction conditions, the positions of the primer extension products and the length markers are as described in the legend to figure 28. Lane 1, wild-type -383 apo-E promoter template; lane 2, wild-type template with α-amanitin; lane 3, template with deletion at nucleotides -184 to -173; lane 4, template with deletion at nucleotides -161 to -141; lane 5, template with substitution at position -151 (C to A); lane 6, template with deletion at nucleotides -59 to -45; lane 7, template with deletion at nucleotides -103 to -83

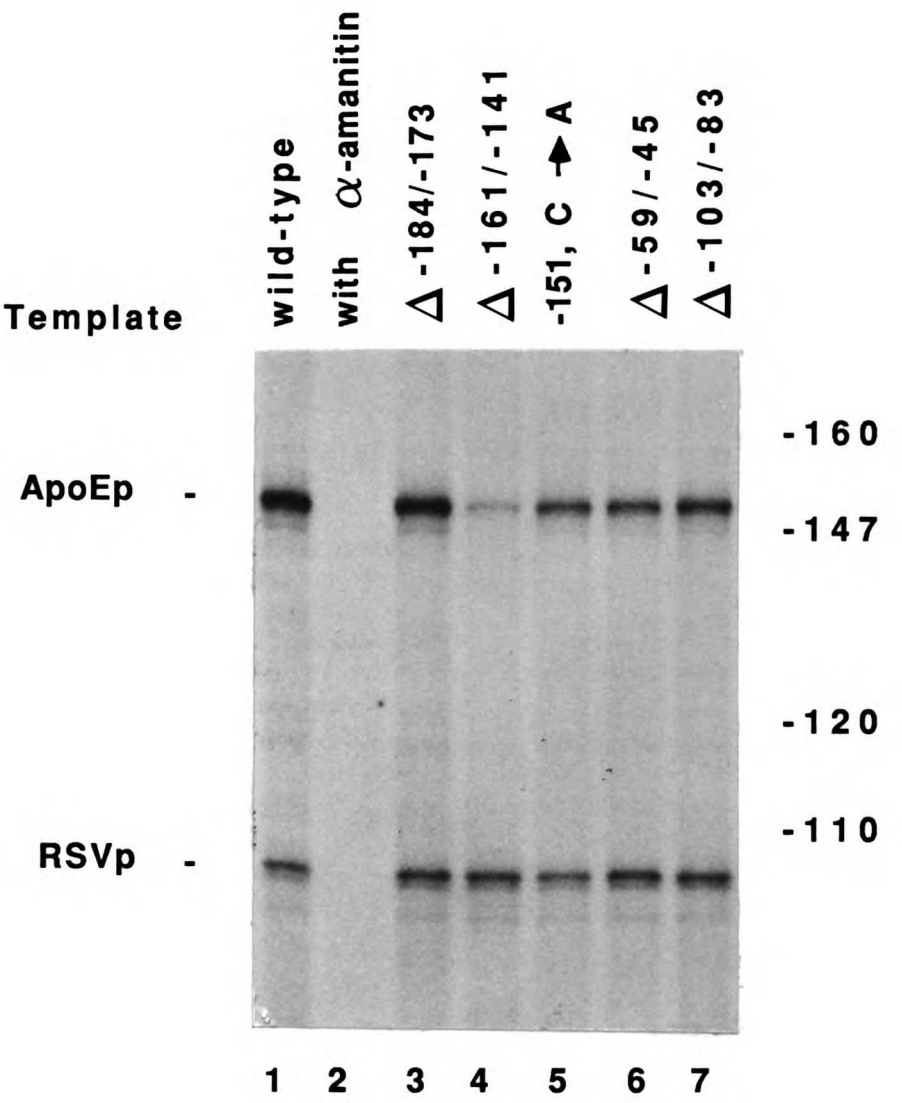
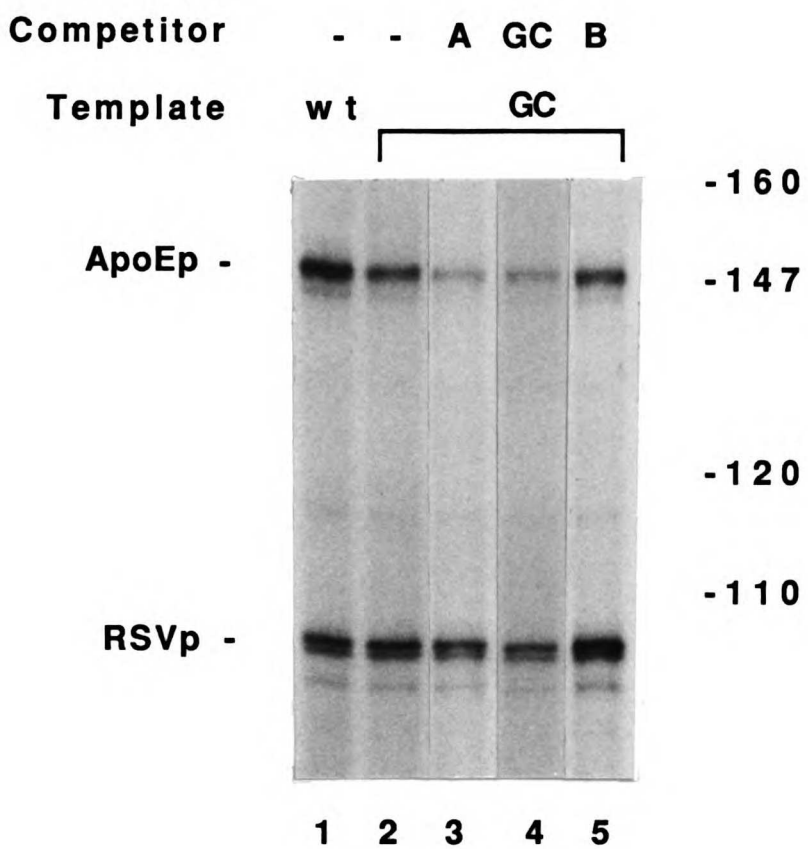

Template
wild-type
with α-amanitin
Δ -184/-173
Δ -161/-141
-151, C → A
Δ -59/-45
Δ -103/-83
ApoEp
RSVp
-160
-147
-120
-110
1 2 3 4 5 6 7

Figure 32 *In vitro* transcription analysis of Δ GC box template with oligonucleotide competitors.

All reactions were performed with the -383 apo-E promoter/ CAT gene template with a deletion at nucleotides -59 to -45, corresponding to the proximal GC box element except for the reaction in lane 1. Position and length marker are as described. Lane 1, wild-type -383 apo-E promoter template; lane 2, -383 apo-E promoter with proximal GC box deletion; lane 3, 60 ng of PET oligonucleotide added; lane 4, 60 ng of proximal GC box oligonucleotide added; lane 5, 60 ng of oligonucleotide B (fig. 16) added;


Competitor
- - A GC B
Template
w t
GC
-160
ApoEp -
-147
-120
-110
RSVp -
1 2 3 4 5

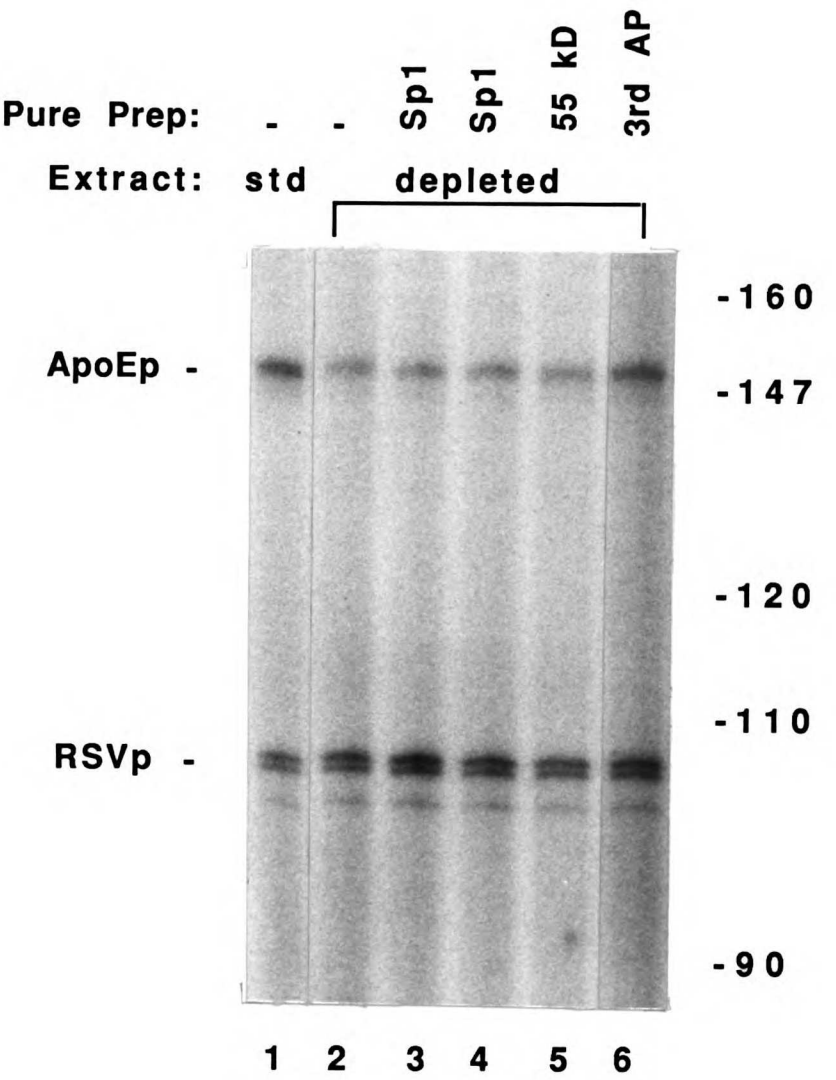
transcriptional activity of the depleted extract. Neither the pure preparation of Sp1 (Fig. 33, lanes 3, 4) nor the 55 kD protein (Fig. 33, lane 5) alone was able to increase the transcript signal. However, the affinity column-purified fraction containing both proteins, as well as a few other proteins (Fig. 28, lane 2), was able to generate a transcript band that was 2.5-fold that of the depleted extract, and about 75% that of the control extract (Fig. 33, lane 6). These results and several previous results suggest that Sp1 plays the key role in producing the enhancer activity of PET but that it may need to interact with another protein or be modified in some manner (both possible mechanisms could be achieved in the partially pure affinity column fraction) for the PET element to be functional. Alternatively, the activity of the pure Sp1 preparation may have been reduced partially by the isolation procedure. The activity which was detected in the depleted extracts may have been due to the action of non-PET, non-GC box elements, such as the -103 to -83 footprint region. The activity of the RSV control template was not affected by the depletion of PET-binding proteins.

E. Detection of Protein Binding in IRE1

A DNA fragment spanning IRE1 (nucleotides +44 to +262) was examined for protein binding by DNase I footprint analysis. Using HepG2 nuclear extract, a weakly protected region spanning nucleotides +166 to +195 was detected on the 5' strand (Fig. 34). Titration experiments with additional nuclear extract did not improve the clarity of the footprint. Thus the IRE1-binding protein(s) may be present in low concentrations in HepG2 nuclear extract, or may have very low binding affinities, or may require different conditions to optimize proper binding. To explore some of these possibilities, a double-stranded oligonucleotide covering nucleotides +161 to +197 was synthesized, labelled and analyzed by the gel retention assay with nuclear

Figure 33 Reconstitution *in vitro* transcription analysis of the Δ GC box template.

The -383 apo-E promoter/CAT gene template with the GC box deletion at nucleotides -59 to -45 was incubated either with 100 μg of standard HeLa nuclear extract (lane 1) or with 100 μg of HeLa extract that had been depleted of the PET-binding proteins by a 30 minute mixing with the PET-affinity Sepharose 4B resin (lanes 2-6). Various purified protein preparations were added to the depleted extract to attempt to reconstitute transcriptional activity. Lane 1, standard HeLa extract; lane 2, depleted HeLa extract only; lane 3, addition of 5 ng of pure Sp1; lane 4, addition of 10 ng of pure Sp1; lane 5, addition of 50 ng of purified 55 kD protein; lane 6, addition of 300 ng of protein eluted from third affinity column passage.


Pure Prep:
-
-
Sp1
Sp1
55 kD
3rd AP
Extract:
std
depleted
ApoEp -
RSVp -
-160
-147
-120
-110
-90
1
2
3
4
5
6

extracts from various cell lines. The results (Fig. 35) show that multiple retarded bands were detected and that the different extracts exhibited similar retention patterns even though the intensity of each individual band varied among the extracts. The functional significance of these proteins has not been determined.

Figure 34 DNase I footprint analysis of IRE1.

Apo-E DNA was 5' end-labelled on the 5' strand at position +44 and subjected to DNase I footprint analysis. The G and G+A lanes represent the respective Maxam and Gilbert sequencing reactions. Lane 1, no extract added; lane 2, 72 μg of HepG2 nuclear extract; lane 3, 144 μg of HepG2 nuclear extract. The sequences protected from DNase I digestion are indicated by the brackets with the numbers on the right indicating the nucleotide positions at the boundaries.

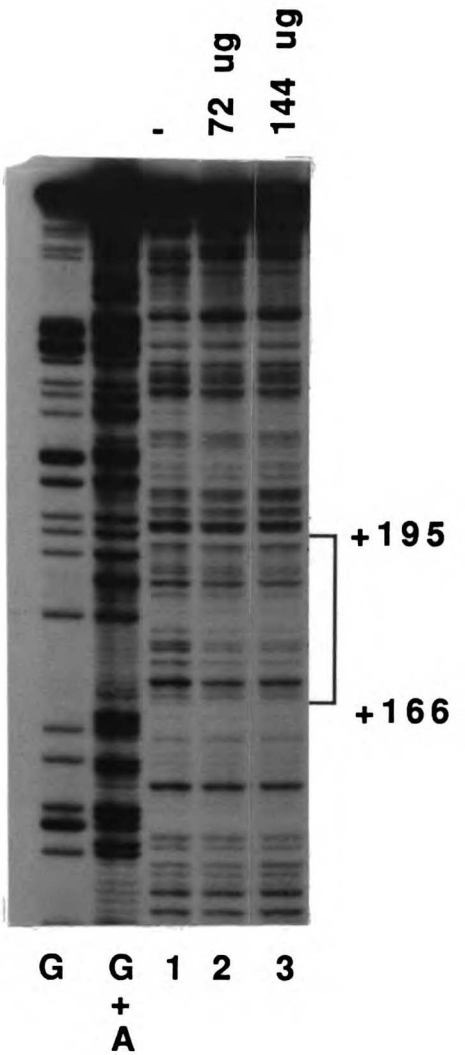
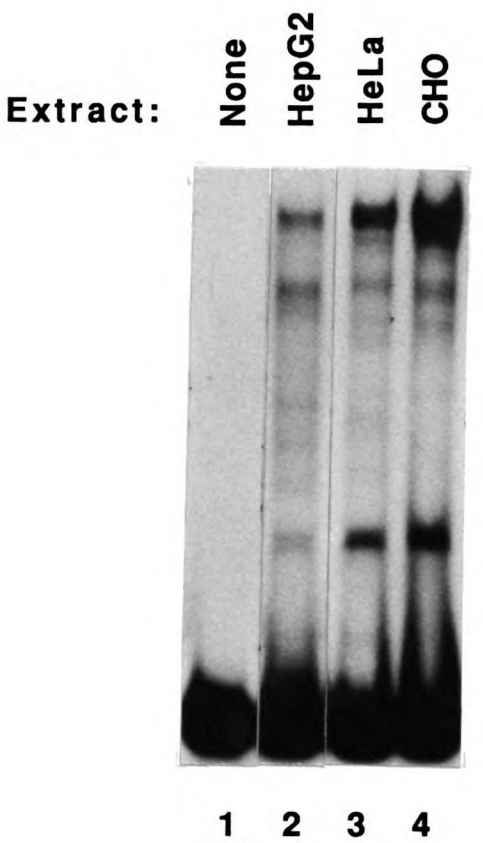

-
72 ug
144 ug
+195
+166
G
G + A
1
2
3

Figure 35 Gel retention analysis of +161 to +197 oligonucleotide with various nuclear extracts.

Gel retention analysis was performed on a synthetic double-stranded oligonucleotide corresponding to nucleotides +161 to +197 which had been 5' end-labelled. Lane 1, no extract; lane 2, 10 μg of HepG2 nuclear extract; lane 3, 10 μg of HeLa nuclear extract; lane 4, 10 μg of CHO nuclear extract.


Extract:
None
HepG2
HeLa
CHO
1
2
3
4

IV. DISCUSSION

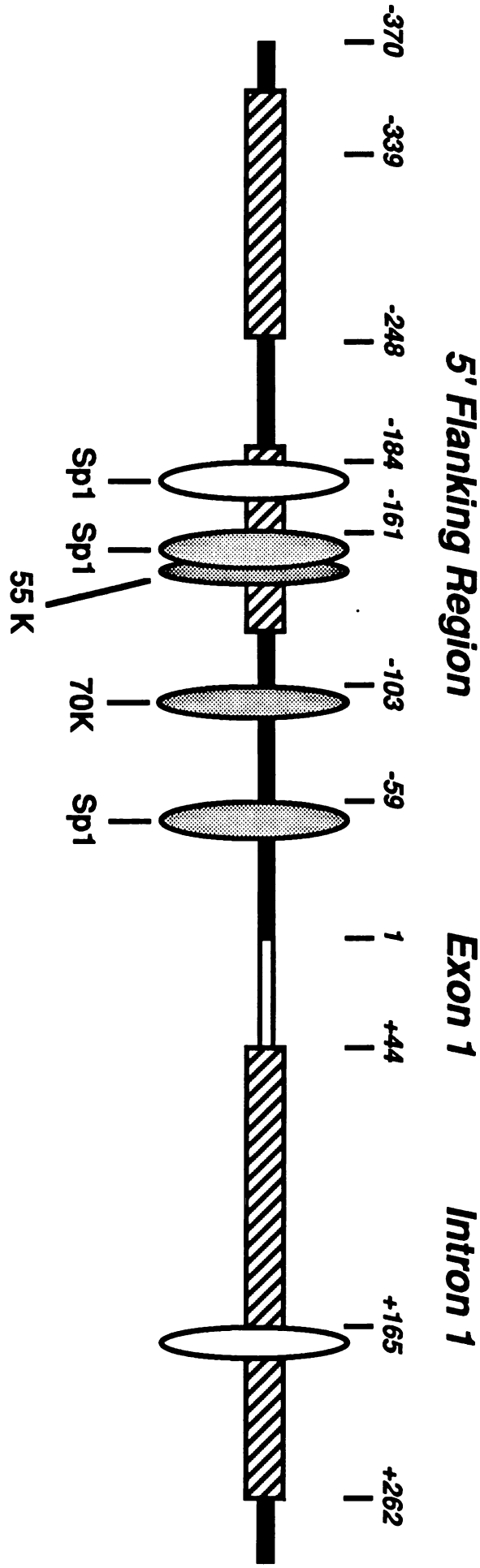
A. Multiple Regulatory Elements Control Human Apo-E Gene Expression

Several regions which regulate the transcription of the human apo-E gene were identified, both in the 5' flanking sequence and within the gene. These findings are summarized in Figure 36. The promoter deletion studies using apo-E promoter/CAT gene recombinant constructs revealed at least five domains in the region between nucleotides -383 and -15 that affect apo-E gene expression. For two of these domains, the functional elements could be deduced by examining the nucleotide sequence. When the TATA box element, located at nucleotides -33 to -27, was deleted, virtually no CAT activity was observed from the resulting construct (Fig. 2). This element, which is known to bind to the transcription factor, TFIID (62), usually represents the minimum promoter whereby a basal level of transcription can be observed (61). Between nucleotides -54 and -45 a consensus GC box element resides that stimulates gene transcription by binding to the transcription factor, Sp1 (104). Deletion of the GC box element resulted in a twofold loss in apo-E promoter activity as determined by *in vitro* transcription and a 2.5- to 4-fold loss in activity as determined from studies with transfected cultured cells. Within the -212 to -81 regulatory domain, three other protein-binding regions (-184 to -173, -161 to -141, -103 to -87) were identified by DNase I footprint analyses. The functional significance of each of these elements is discussed later.

In contrast to the transfection experiments, *in vitro* transcription did not show a difference in promoter activity between the -383 apo-E promoter/CAT template and the -212 apo-E promoter/CAT template (Figs. 2 and 29). It is possible that transcription regulatory effects from distant elements (e.g. more than 200 bases upstream from the initiation site) cannot be detected *in vitro* . It is also possible that during the preparation of nuclear extracts,

Figure 36 **Identification of regulatory elements and protein-binding sequences in the proximal 5' flanking region and the first intron of the human apo-E gene.**

The hatched boxes indicate the apo-E gene fragments that exhibited enhancer activity. The protein binding sites as determined by DNase I footprinting assays and gel retention assays are represented by ovals. Shaded ovals represent proteins demonstrated in this study which show transcription-regulating activity. Clear ovals represent proteins for which functional activities at their binding sites have not been detected. The numbers indicate the nucleotide sequence in base pairs relative to the apo-E transcription initiation site.


5' Flanking Region
Exon 1
Intron 1
-370
-339
-248
-184
-161
-103
-59
1
+44
+165
+262
Sp1
Sp1
55 K
70K
Sp1

some protein factors may have been degraded or inactivated, or they may not have been extracted, causing some regulatory effects to be missed in the *in vitro* transcription assays. On the other hand, in cells that have been transfected with the CAT gene plasmids, post-transcriptional effects, such as RNA or protein stability, could cause an increase or decrease in CAT enzyme activity which then may be misinterpreted as transcriptional effects. Alternatively, the discrepancy in results may be the result of using different cell lines for the two assays. Unfortunately, it was not possible to use only one cell line for both assays since HeLa cells which yielded transcriptionally active nuclear extracts yielded extremely low CAT enzyme activities upon transfection of appropriate constructs, whereas CHO and HepG2 cells produced strong CAT enzyme signals but relatively weak extracts for *in vitro* transcription. Thus, although an enhancer activity has been demonstrated within the -383 to -212 domain (discussed in the following paragraph), the significance of this region on apo-E gene transcription is not clear.

Three elements with enhancer properties, URE1 (-193 to -124), URE2 (-366 to -246), and IRE1 (+44 to +262) were identified from experiments with the enhancer test vectors, pTK1 and pTK10. URE1 is functional in both orientations relative to the promoter, from both long and short distances, and from being positioned either 5' or 3' to the promoter. Furthermore, these effects could be observed with different promoters. The enhancer action of URE1 was localized to the sequence at nucleotides -161 to -141 and was found to be associated directly with the recognition and binding of specific proteins to this sequence.

URE2, on the other hand, displayed enhancer activity that was both position- and promoter-dependent. DNase I footprint assays did not reveal any protein-binding sequences within this element, and *in vitro* transcription

assays showed no difference in transcriptional activity between a template that contained URE2 and one that did not. The failure to detect a footprint in URE2 may be attributed to very low amounts of the associated protein factor, or to the footprinting reaction condition not being optimal for the binding of this protein. However, if URE2 is not associated with a specific binding protein, the significance of its enhancer effect is questionable. Perhaps the insertion of URE2 into the pTK1 test plasmid is able to "open up" or arrange the DNA into a configuration which makes the TK promoter more accessible for the binding by transcription factors.

The intron enhancer element, IRE1, displayed orientation- and distance-independence, although some evidence of position dependence and cell-type specificity was detected (Fig.7, Table II). However, unlike URE1, the "core" sequence of enhancer activity was not determined conclusively. When the +44 to +262 fragment was subdivided into smaller pieces, none of the subfragments displayed enhancer activity. One possible explanation for this observation is that IRE1 enhancer activity requires some interaction to occur between distant segments within the fragment. One of these segments might be the protein-binding sequence between nucleotides +166 and +195 (Fig. 34). Interestingly, this footprint was less distinct and the boundaries were less defined than the footprints in the 5' flanking region. Neither increasing the amount of extract (to three times the normal amount) nor changing the incubation conditions (by adjusting temperature and KCl and $MgCl_2$ concentrations) improved the quality of the footprint. These observations suggested that the corresponding proteins were binding to this sequence with relatively low affinities. Gel retention analysis with the IRE1 footprint-spanning oligonucleotide showed two relatively distinct retarded bands and several diffuse bands, and the pattern was identical for extracts prepared from

HepG2, CHO, and HeLa cells (Fig. 35) even though the amount of protein within each cell type appeared to vary. Thus,there does not seem to be a simple correlation between the presence or absence of these protein factors in the different cell lines and the cell-specific enhancer activity of IRE1. However, it is possible that in HeLa cells, where IRE1 enhancer activity was not detected (Table II), the interaction between the +166 to +195 sequence-binding factors and the other target factor(s) within IRE1 cannot occur.

The protein-binding sequence between nucleotides -103 and -87 has been determined to be a positive transcription regulatory element which, when deleted, reduced the activity of the apo-E promoter by approximately twofold in *in vitro* transcription studies (Fig. 29). A fragment which spans this sequence (nucleotides -123 to -81) was shown not to have enhancer activity (Fig. 4). Thus this element, referred as upstream regulatory element 3 (URE3), apparently is unable to stimulate transcription from a distance. Furthermore, the deletion of URE3 did not affect the enhancer activity of the -246 to -81 fragment which contains both URE1 and URE3 (Fig. 8). The URE3-binding protein, which appears to be ubiquitous (Fig. 13), was partially purified from HeLa cells using Sephacryl S-300 gel filtration chromatography and DNA-affinity chromatography with concatemerized URE3 oligonucleotides. The identity of this protein is under investigation currently (T. Leren, personal communication).

Transcription-regulating activity was not detected for the Sp1-binding -184 to -173 sequence. Although this sequence lies within the URE1 element, its deletion had no effect either on the enhancer activity of the URE1 domain (Fig. 8) or on the transcriptional activity, *in vitro*, of the apo-E promoter (Fig. 31, lane 3). Why would a sequence that binds a known positive regulatory transcription factor fail to exhibit any functional activity? Perhaps, the

conditions under which DNase I footprint and gel retention analyses were performed "created" an Sp1-binding sequence which is not present *in vivo*. However, this possibility seems unlikely since other 5' flanking sequences (i.e. nucleotides -77 to -68, and -280 to -271) which have as much homology to the consensus GC box sequence as the -184 to -173 sequence did not display any Sp1 binding. One possibility is that Sp1-binding to this sequence is coupled with an induction event, perhaps involving another transcription factor, to regulate transcription. Another possibility is that this sequence becomes functionally significant only when some or all of the other 5' flanking regulatory sequences (such as the PET element or the proximal GC box element) are inactivated in some manner. Indeed, the residual enhancer activity detected in the constructs in which portions of the -161 to -141 footprint sequence had been deleted (Fig. 8) may be caused by the -184 to -173 sequence that may have become active when the neighboring Sp1-binding sequence had been eliminated.

The human apo-E gene is endowed with a complex array of regulatory elements including enhancers and protein-binding sequences. For a particular cell type which produces apo-E, the question arises as to whether all of these elements need to be functional at all times. It is possible that for a given physiological condition, only a few of the *cis*-acting sequences and their associated protein factors are involved directly in regulating the rate of transcription. During this time, the other elements and proteins may be completely silent or may play indirect regulatory roles by mediating the interactions between the "active" elements and/or proteins and the transcription initiation machinery. Data in this study which support this possibility include the finding that the 1-kb fragment which contains all three enhancer elements stimulated promoter activity to a level that was

similar to that of any single enhancer element (Fig. 4). Also, cell-free transcription assays indicated that the total activity of the apo-E promoter was not the sum of the activities of all the regulatory sequences, which were determined from the experiment involving deletion and substitution templates (Fig. 31). Under different physiological conditions, distinct combinations of regulatory elements and protein factors may form, leading to different levels of transcriptional activity.

B. PET-Binding Proteins and Enhancer Activity

A number of studies have shown that the PET sequence (-169 to -140), which includes the protein-binding nucleotides between -161 and -141, represents the most dominant positive regulatory element in apo-E gene expression. All assays which involve the deletion of this sequence or the prevention of protein-binding to this sequence exhibited reductions in transcription levels (Figs. 8, 30, 31, 32, Table IV). The PET element was found to possess enhancer activity that is roughly equal to that of URE1 but about one-half that of the -246 to -81 fragment. The difference might be attributed to the sequence at nucleotides -123 to -105 which, when deleted, decreased the enhancer activity of the large fragment by twofold. Since the deletion of this sequence did not seem to affect the protein-binding at -161 to -141 (Fig. 14 D), its transcription-regulating mechanism may be from mediating the interaction between the proteins at the enhancer element and those at the transcription initiation complex (107). On the other hand, regulatory proteins which bind to this sequence may not yet have been discovered. The -123 to -105 sequence has been studied also by Smith et al (108), who found it to be a region of negative activity that neutralized the apparent strong positive activity of the -103 to -87 footprint region. It is unclear why the two laboratories have come up with different observations concerning this

sequence. Perhaps this sequence exerts a positive effect on the activity of the PET region but a negative effect on the activity of the -103 to -87 sequence.

The PET sequence is recognized and bound specifically by transcription factor, Sp1. In fact, Sp1 alone was able to protect the entire -161 to -141 sequence from DNase I digestion (Fig. 22). Gel retention studies showed that Sp1 binding to PET gives rise to two DNA-protein complexes (Fig. 21, lane 4). These complex bands probably represent either monomer/dimer or dimer/tetramer forms of Sp1 interacting with the PET sequence. Tjian and his colleagues have provided much information about this transcription factor following its purification as a pair of polypeptides with sizes of 105 and 95 kD (99). The DNA-binding domain of Sp1 has been localized to a C-terminal 168 amino acid region that contains three contiguous Zn(II) finger motifs (76). Four other distinct domains have been identified as being important for the transcriptional activation function of this protein (83). Furthermore, Sp1 was found to be glycosylated with multiple O-linked N-acetylglucosamine residues (109) which may play a role in the mechanism of Sp1 action.

The transcription-stimulating function of URE1 and PET is dependent upon the binding of Sp1 to the -161 to -141 sequence. The prevention of Sp1 binding to this sequence resulted in the reduction or elimination of the enhancer activity of URE1 and PET, as determined by CAT assays (Fig. 8, Table IV), and in the reduction of apo-E promoter activity, as determined by *in vitro* transcription assays (Fig. 30-33). The observation that Sp1 is involved directly in the enhancer activity of the PET element was rather surprising since Sp1 and the GC box element have always been associated previously only with promoter activity. However, a recent study has shown that multiple GC box elements placed downstream of the transcription start site were able to stimulate the activity of an associated promoter (R. Tjian, personal

communication). Thus Sp1 can behave as an enhancer-binding protein. In fact, there have been numerous reports about the ability of individual nuclear factors to exhibit multiple activities. For example, transcription factor C/EBP recognizes not only the CCAAT sequence which is a common upstream promoter element, but also a common enhancer core element (77). Another CCAAT element-binding factor, CTF, is identical to the factor, NF-1, which is required for adenovirus DNA replication (110). The glucocorticoid receptor can confer either positive or negative regulation to genes which contain the respective glucocorticoid responsive elements (111).

The finding that Sp1 protects 21 bp of duplex DNA in the PET enhancer sequence, but only 12 bp at the upstream binding site (-184 to -173) and 15 bp at the proximal GC box, suggests different interactions of this transcription factor with its various recognition sequences. Each Sp1 binding site may be associated with different activities or influences on gene transcription rates. It is possible that slight differences in DNA-binding may allow the activation domains of Sp1 to be more accessible to interact with other transcription factors or with the transcription initiation complex. The joint modulation of transcription by Sp1 and other transcription factors have been suggested because Sp1 recognition sequences often are found near the binding sites for transcription factors such as CTF and AP-1 (65, 94).

Although Sp1-binding is required for the enhancer activity of the PET element to be observed, it is not clear whether this transcription factor is the only factor involved in this activity. The reconstitution experiment has demonstrated that a pure preparation of Sp1, by itself, could not increase the level of transcription from nuclear extracts that had been depleted of the PET-binding proteins (Fig. 33, lanes 3, 4). However, a partially pure preparation eluted from three passages through the PET-affinity column (which contains

Sp1 and several other proteins, Fig. 28, lane 2) was able to restore some of the transcriptional activity (Fig. 33, lane 6). One possible explanation for this result may be that the pure Sp1 preparation had lost its functional activity due to the isolation procedure, even though the protein was still capable of binding to the PET sequence. On the other hand, Sp1 present in the partially pure fraction, was still transcriptionally active. A second possibility would be that another protein, present in the less homogeneous preparation, needs to interact with Sp1 in order for the regulatory activity of the PET element to be manifested. This protein might affect transcription directly by serving as a contact between Sp1 bound at the PET sequence and some component in the transcription initiation complex. Alternatively, the secondary protein may be involved in modifying Sp1 in a manner which allows the Sp1-PET complex to exhibit enhancer activity. This "second protein" possibility may explain why Sp1 binding to PET affects transcription differently than its binding to the proximal GC box element or to the -184 to -173 sequence. However, there is no information yet about whether the second protein actually exists.

Using the PET oligonucleotide affinity column, a 55 kD nuclear protein was purified. The characterization of this protein has shown that it is rather abundant and is found, so far, only in human cell lines (detected in HeLa, HepG2 and THP-1 cells but not in CHO and L cells). The absence of the 55 kD protein in CHO and L cells suggest that it is not absolutely required for the enhancer activity of URE1 and PET. This protein shows strong binding to poly (dI-dC)/poly (dI-dC), and when more than 0.4 μg of this carrier DNA was added (20 ng/μl concentration), the binding of the 55 kD protein to PET could not be detected in gel retention assays (Fig. 19). The addition of oligonucleotide competitors seems to release the 55 kD protein from poly (dI-dC)/poly (dI-dC) even at high concentrations of the polymer, enabling the protein to bind PET.

The 55 kD protein exhibits specificity in binding since it does not bind to the GC box element, nor to the recognition sequence of several transcription factors, but is capable of binding to a SV40 enhancer sequence which recognizes the transcription factor, AP-2 (66, 106). Experiments involving nuclear extracts from various cell lines (Fig. 24) and a pure preparation of AP-2 (Fig. 26) indicate that the 55 kD protein is not AP-2. When nucleotides within a five base segment of the PET sequence (-156 to -152, CCCCA), which contained the only homology with the AP-2 binding sequence, were substituted, no loss in 55 kD protein binding was detected (Fig. 25, lanes 6-9). It is not clear how the 55 kD protein exhibits sequence-specific binding to DNA since the exact nucleotide sequence that is recognized by this protein was not determined. However, the recognition site for the 55 kD protein appears to lie between nucleotides -160 and -143, suggesting that it overlaps the GC box homology region (-157 to -148). Gel retention assays have shown that the binding of Sp1 and the 55 kD protein appear to be mutually exclusive (Fig. 18-20). Therefore, the 55 kD protein might serve as a repressor or blocking factor which suppresses apo-E gene transcription by preventing the binding of Sp1 to PET. Since PET appears to be the most powerful positive element for the regulation of human apo-E gene expression, human cells may have devised this mechanism to modulate the activity of this element.

C. Cell-Specificity and Inducibility of Apo-E Gene Expression

The present study has identified numerous transcription-regulating and/or protein-binding elements, but none of them appears to confer cell-specific expression of the human apo-E gene. Since the apo-E gene is expressed in various diverse cell types (20), it is not surprising that no cell-specific transcription factor was identified in the promoter and enhancer regions. Transient gene transfer studies involving either the human apo-E

gene or the apo-E promoter attached to a reporter gene have shown efficient expression in virtually all the transfected cell lines (54, 112). This suggests that the apo-E promoter and enhancer elements in the transfected DNA are accessible to the protein factors necessary for the transcription of the gene. Considering the ubiquitous nature of the proteins binding to the PET sequence, to the -103 to -87 sequence, and to the IRE1 footprint sequence, it seems logical to surmise that the principal mechanism controlling cell-specific apo-E gene expression is a negative one. Chromatin structure (113, 114) may play a role in determining the activation or repression of the apo-E gene in that histones and non-histone proteins may act as barriers to prevent the binding of transcription factors to the regulatory elements. Perhaps, far upstream elements similar to the dominant control region (DCR) discovered for the human β-globin gene (115, 116) are somehow responsible for "opening up" the chromatin around the apo-E gene to allow transcription to occur. The investigation of these possibilities would require examining long stretches of DNA in a wide variety of cells and tissues. DNase I sensitivity studies have been used to examine the state of activation or inactivation of numerous genes in numerous cells (113). These experiments have also identified hypersensitivity sites which correspond to sequences which regulate the activation of the gene and can be located thousands of bases away from the gene (113, 115). Information concerning the cell-specific expression of the apo-E gene obtained from the chromatin studies then can be used in studies involving intact animals such as in transgenic mice (117).

The multitude of regulatory and protein-binding elements associated with the human apo-E gene may represent a system for the complex regulation of the expression of this gene. The role that apo-E plays in cholesterol transport would suggest that this lipid could influence the

regulation of apo-E transcription. Indeed, in cholesterol-loaded mouse peritoneal macrophages, an increase in apo-E mRNA synthesis was detected (50). Furthermore, macrophages near damaged nerve sheaths in rats secrete large amounts of apo-E, probably due to an increase in the intracellular cholesterol of these phagocytic cells (35). Attempts to detect possible cholesterol effects on the numerous regulatory elements and nuclear factors identified in this study have been unsuccessful (Y.K. Paik, personal communication). It is not understood yet whether this lack of a direct cholesterol effect is due to the inadequate loading of cholesterol to the cells being tested, or to the fact that cholesterol responsive elements of the apo-E gene are located further upstream or downstream of the regions being studied, or to the possibility that human apo-E gene expression simply is not responsive to cholesterol treatment. Recent findings have shown that apo-E mRNA levels are increased in rat ovarian granulosa cells that had been treated with cAMP or TPA (53). However, similar treatments on cultured cells that had been transfected with CAT gene vectors containing human apo-E promoter or enhancer elements showed no changes in CAT activity (Y.K. Paik, personal communication).

The only clear case of differential or induced expression of the human apo-E gene is observed in human blood monocytes. Although monocytes do not express the apo-E gene, upon differentiation into macrophages, they begin to synthesize high levels of apo-E mRNA (49). Phorbol esters are often used to induce differentiation, but the study mentioned previously indicate that neither TPA nor PMA seem to be involved directly in regulating apo-E gene transcription. To examine the expression of the apo-E gene in macrophages, studies are under way currently to introduce human apo-E gene sequences using retrovirus vectors (118) into mouse bone marrow hematopoietic stem

cells which are the precursors for all blood cells (119). Preliminary data with several apo-E gene-containing retrovirus constructs transfected in cultured cell lines have demonstrated that the deletion of the PET sequence or the first intron reduced the level of apo-E mRNA synthesis by tenfold and by fivefold, respectively (S. Lauer, personal communication).

Although the significance of the complex array of regulatory elements in the human apo-E gene remains unclear, a consequence of these multiple elements is that the human apo-E promoter is relatively strong in comparison to other mammalian and viral promoters. In gene transfer experiments with cultured cells using apolipoprotein gene promoters ligated to the CAT gene, the CAT enzyme activities with human apo-E promoter-containing constructs were generally four- to tenfold higher than the activities obtained with the human apo-CI-, B-, and AIV-promoter constructs as well as with the mouse apo-E promoter constructs (Y.K. Paik, personal communication). In the *in vitro* transcription experiments with HeLa nuclear extracts, the strength of the human apo-E promoter was demonstrated even more dramatically in that the generated transcripts were about twofold higher than those of the RSV promoter and almost tenfold higher than the transcripts produced by the HSV-TK promoter (Figs. 28-31 and unpublished observations). If the mechanisms that regulate its activity are understood better, the human apo-E promoter may be useful in studies where high levels of expression of a reporter gene are desired.

V. REFERENCES

1. Gotto, A.M. Jr., Pownall, H.J., and Havel, R.J. (1986) *Methods Enzymol.* **128**, 3-41.

2. Driscoll, D.M., and Getz, G.S. (1986) *Methods Enzymol.* **128**, 41-70.

3. Deckelbaum, R.J., Shipley, G.G., Small, D.M., Lees, R.S., and George, P.K. (1975) *Science* **190**, 392-394.

4. Tollefson, J.H., and Albers, J. (1986) *Methods Enzymol.* **129**, 797-816.

5. Hamilton, R.L., Williams, M., Fielding, C., and Havel, R.J. (1976) *J. Clin. Invest.* **58**, 667-680.

6. Fielding, C.J. (1986) *Methods Enzymol.* **129**, 783-797.

7. Redgrave, T.G., Roberts, D.C.K., and West, C.E. (1975) *Anal. Biochem.* **65**, 42-48.

8. Zilversmit, D.B. (1978) in *Disturbances in Lipid and Lipoprotein Metabolism* (Dietschy, J.M., Gotto, A.M. Jr., and Ontko, J.A., eds) pp. 69-81, American Physiological Society, Bethesda.

9. Fraser, R., Cliff, W.J., and Courtice, F.C. (1968) *Quart. J. Exptl. Physiol.* **53**, 390-398.

10. Fielding, C.J., and Fielding, P.E. (1976) *J. Lipid Res.* **17**, 419-423.

11. Angelin, B., Ravida, C.A., Innerarity, T.L., and Mahley, R.W. (1983) *J. Clin. Invest.* **71**, 816-831.

12. Jackson, R.L., Morrisett, J.D., and Gotto, A.M. (1976) *Physiol. Rev.* **56**, 259-316.

13. Eisenberg, S., and Levy, R.T. (1975) *Advan. Lip. Res.* **13**, 1-89.

14. Brown, M.S., and Goldstein, J.L. (1986) *Science* **232**, 34-47.

15. Goldstein, J.L., Brown, M.S., Anderson, R.G.W., Russell, D.W., and Schneider, W.J. (1985) *Annu. Rev. Cell Biol.* **1**, 1-39.

16. Hamilton, R.L. (1978) in *Disturbances in Lipid and Lipoprotein Metabolism* (Dietschy, J.M., Gotto, A.M. Jr., and Ontko, J.A. eds) pp. 155-171, American Physiological Society, Bethesda.

17. Havel, R.J., Kane, J.P., and Kashyap, M.D. (1973) *J. Clin. Invest.* **53**, 32-38.

18. Mahley, R.W. (1982) *Med. Clin. North Am.* **66**, 375-402.

19. Mahley, R.W. (1986) *Clin. Invest. Med.* **9**, 304-308.

20. Mahely, R.W. (1988) *Science* **240**, 622-630.

21. Rall, S.C. Jr., Weisgraber, K.H., and Mahley, R.W. (1982) *J. Biol. Chem.* **257**, 4171-4178.

22. Rall, S.C. Jr., Weisgraber, K.H., Innerarity, T.L., and Mahley, R.W. (1985) *Circulation* **72**, III-143.

23. Mahley, R.W., and Innerarity, T.L. (1983) *Biochim. Biophys. Acta* **737**, 197-222.

24. Mahley, R.W., Innerarity, T.L., Rall, S.C. Jr., and Weisgraber, K.H. (1984) *J. Lipid Res.* **25**, 1277-1294.

25. Brown, M.S., and Goldstein, J.L. (1983) *Annu. Rev. Biochem.* **52**, 223-261.

26. Zannis, V.I., Breslow, J.L., Utermann, G., Mahley, R.W., Weisgraber, K.H., Havel, R.J., Goldstein, J.L., Brown, M.S., Schonfeld, G., Hazzard, W.R., and Blum, C. (1982) *J. Lipid Res.* **23**, 911-914.

27. Mahley, R.W., and Angelin, B. (1984) *Adv. Intern. Med.* **29**, 385-411.

28. Innerarity, T.L., Friedlander, E.J., Rall, S.C. Jr., Weisgraber, K.H., and Mahley, R.W. (1983) *J. Biol. Chem.* **258**, 12341-12347.

29. Weisgraber, K.H., Innerarity, T.L., Harder, K.J., Mahley, R.W., Milne, R.W., Marcel, Y.L., and Sparrow, J.T. (1983) *J. Biol. Chem.* **258** 12348-12354.

30. Lalazar, A., Weisgraber, K.H., Rall, S.C. Jr., Giladi, H., Innerarity, T.L., Levanon, A.Z., Boyles, J.K., Anut, B., Gorecki, M., Mahley, R.W., and Vogel, T. (1988) *J. Biol. Chem.* **263**, 3542-3545.

31. Yamamoto, T., Davis, C.G., Brown, M.S., Schneider, W.J., Casey, M.L., Goldstein, J.L., and Russell, D.W. (1984) *Cell* **39**, 27-38.

32. Mahley, R.W., Innerarity, T.L., Rall, S.C. Jr., and Weisgraber, K.H. (1985) *Ann. N.Y. Acad. Sci.* **454**, 209-221.

33. Koo, C., Innerarity, T.L., and Mahley, R.W. (1985) *J. Biol. Chem.***260**, 11934-11943.

34. Boyles, J.K., Hui, D.Y., Weisgraber, K.H., Pitas, R.E., and Mahley, R.W. (1987) *Soc. Neurosci. Abstr.* **13**, 294.

35. Ignatius, M.J., Shooter, E.M., Pitas, R.E., and Mahley, R.W. (1986) *Soc. Neurosci. Abstr.* **12**, 277.

36. Millis, A.J.T., Hoyle, M., Reich, E., and Mann, D.M. (1985) *J. Biol. Chem.* **260**, 3754-3761.

37. Majack, R.A., Castle, C.K., Goodman, L.V., Weisgraber, K.H., Mahley, R.W., Shooter, E.M., and Gebicke-Haerter, P.J. (1988) *J. Cell Biol.* **107**, 1207-1215.

38. Cuthbert, J.A., and Lipsky, P.E. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4539-4543.

39. Hui, D.Y., Harmony, J.A.K., Innerarity, T.L., and Mahley, R.W. (1980) *J. Biol. Chem.* **255**, 11775-11781.

40. Das, H.K., McPherson, J., Bruns, G.A.P., Karanthanasis, S.K., and Breslow, J.L. (1985) *J. Biol. Chem.* **260**, 6240-6247.

41. Lauer, S.J., Walker, D., Elshourbagy, N.A., Reardon, C.A., Levy-Wilson, B., and Taylor, J.M. (1988) *J. Biol. Chem.* **263**, 7277-7286.

42. Paik, Y.-K., Chang, D.J., Reardon, C.A., Davies, G.E., Mahley, R.W., and Taylor, J.M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3445-3449.

43. Mclean, J.W., Elshourbagy, N.A., Chang, D.J., Mahley, R.W., and Taylor, J.M. (1984) *J. Biol. Chem.* **259**, 6498-6504.

44. Havel, R.J., Goldstein, J.L., and Brown, M.S. (1980) in *Metabolic Control and Disease* (Bondy, P.K., and Rosenberg, L.E., eds) pp. 393-494, W.B. Saunders Company, Philadelphia.

45. Elshourbagy, N.A., Liao, W.S., Mahley, R.W., and Taylor, J.M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 203-207.

46. Boyles, J., Pitas, R.E., Wilson, E., Mahley, R.W., and Taylor, J.M. (1985) *J. Clin. Invest.* **76**, 1501-1513.

47. Pitas, R.E., Boyles, J.K., Lee, S.H., Foss, D., and Mahley, R.W. (1987) *Biochim. Biophys. Acta* **917**, 148-161.

48. Basu, S.K., Brown, M.S., Ho, Y.K., Havel, R.J., and Goldstein, J.L. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7545-7549.

49. Basu, S.K., Ho, Y.K., Brown, M.S., Bilheimer, D.W., Anderson, R.G.W., and Goldstein, J.L. (1982) *J. Biol. Chem.* **257**, 9788-9795.

50. Mazzone, T., Gump, H., Diller, P., and Getz, G.S. (1987) *J. Biol. Chem.* **262**, 11657-11662.

51. Werb, Z., and Chin, J.R. (1983) *J. Biol. Chem.* **258**, 10642-10648.

52. Driscoll, D.M., and Getz, G.S. (1984) *J. Lipid Res.* **25**, 1368-1379.

53. Wyne, K.L., Schreiber, J.R., Larsen, A.L., and Getz, G.S. (1989) *J. Biol. Chem.* **264**, 981-989.

54. Reardon, C.A., Lau, Y.-F., Paik, Y.-K., Weisgraber, K.H., Mahley, R.W., and Taylor, J.M. (1986) *J. Biol. Chem.* **261**, 9858-9864.

55. Reardon, C.A., Paik, Y.-K., Chang, D.J., Davies, G.E., Mahley, R.W., Lau, Y.-F., and Taylor, J.M. (1986) *Methods Enzymol.* **128**, 811-823.

56. Allison, L.A., Moyle, M., Shales, M., and Ingles, C.J. (1985) *Cell* **42**, 599-610.

57. Jones, N.C., Rigby, P.W.J., and Ziff, E.B. (1988) *Genes Dev.* **2**, 267-281.

58. McKnight, S., and Tjian, R. (1986) *Cell* **46**, 795-805.

59. Maniatis, T., Goodbourn, S., and Fischer, J.A. (1987) *Science* **236**, 1237-1245.

60. Dynan, W.S., and Tjian, R. (1985) *Nature* **316**, 774-778.

61. Breathnatch, R., and Chambon, P. (1981) *Ann. Rev. Biochem.* **50**, 349-383.

62. Sawadogo, M., and Roeder, R.G. (1985) *Cell* **43**, 165-175.

63. Serfling, E., Jasin, M., and Schaffner, W. (1985) *Trends Genet.* **1**, 224-230.

64. Clark, L., Pollock, R.M., and Hay, R.T. (1988) *Genes Dev.* **2**, 991-1002.

65. Lee, W., Mitchell, P., and Tjian, R. (1987) *Cell* **49**, 741-752.

66. Mitchell, P.J., Wang, C., and Tjian, R. (1987) *Cell* **50**, 847-861.

67. Bohmann, D., Keller, W., Dale, T., Schuler, H.R., Tebb, G., and Mattaj, I.W. (1987) *Nature* **325**, 268-272.

68. Edlund, T., Walker, M.D., Barr, P.J., and Rutter, W.J. (1985) *Science* **230**, 912-916.

69. Treisman, R. (1985) *Cell* **42**, 889-902.

70. Yamamoto, K.R. (1985) *Annu. Rev. Genet.* **19**, 209-257.

71. Pabo, C.O., and Sauer, R.T. (1984) *Annu. Rev. Biochem.* **53**, 293-321.

72. Evans, R.M., and Hollenberg, S.M. (1988) *Cell* **52**, 1-3.

73. Miller, J., McLachlan, A.D., and Klug, A. (1985) *EMBO J.* **4**, 1609-1614.

74. Godowski, P.J., Rusconi, S., Miesfeld, R., and Yamamoto, K.R. (1987) *Nature* **325**, 365-368.

75. Kumar, V., Green, S., Staub, A., and Chambon, P. (1986) *EMBO J.* **5**, 2231-2236.

76. Kadonaga, J., Carner, K.C., Masiarz, F.R., and Tjian, R. (1987) *Cell* **51**, 1079-1090.

77. Landschultz, W.H., Johnson, P.F., and McKnight, S.L. (1988) *Science* **240**, 1759-1764.

78. Landschultz, W.H., Johnson, P.F., and McKnight, S.L. (1989) *Science*

243, 1681-1688.

79. Turner, R., and Tjian, R. (1989) *Science* **243**, 1689-1694.

80. Gentz, R., Rauscher, F.J. III, Abate, C., and Curran, T. (1989) *Science* **243**, 1695-1699.

81. Hope, I.A., and Struhl, K. (1986) *Cell* **46**, 885-894.

82. Ma, J., and Ptashne, M. (1987) *Cell* **48**, 847-853.

83. Courey, A.J., and Tjian, R. (1988) *Cell* **55**, 887-898.

84. Sigler, P.B. (1988) *Nature* **333**, 210-212.

85. Horikoshi, M., Carey, M.F., Kakidani, H., and Roeder, R.G. (1988) *Cell* **54**, 665-669.

86. Walker, M.D., Edlund, T., Boulet, A.M., and Rutter, W.J. (1983) *Nature* **306**, 557-561.

87. Maxam, A.M., and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560.

88. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467.

89. Gorman, C.M., Merlino, G.T., Willingham, M.C., Pastan, I., and Howard, B.H. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6777-6781.

90. Graham, F.L., and van der Eb, A.J. (1973) *Virology* **52**, 456-467.

91. Bradford, M.M. (1976) *Anal. Biochem.* **72**, 248-254.

92. Gorman, C.M., Moffat, L.F., and Howard, B.H. (1982) *Mol. Cell. Biol.* **2**, 1044-1051.

93. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J., and Rutter, W.J. (1979) *Biochemistry* **18**, 5294-5299.

94. Jones, K.A., Yamamoto, K.R., and Tjian, R. (1985) *Cell* **42**, 559-572.

95. Shapiro, D.J., Sharp, P.A., Wahli, W.W., and Keller, M.J. (1988) *DNA* **7**, 47-55.

96. Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983) *Nucleic Acids Res.* **11**, 1475-1489.

97. Carthew, R.W., Chodosh, L.A., and Sharp, P.A. (1985) *Cell* **43**, 439-448.

98. Zoller, M.J., and Smith, M. (1984) *DNA* **3**, 479-488.

99. Briggs, M.R., Kadonaga, J.T., Bell, S.P., and Tjian, R. (1986) *Science* **234**, 47-52.

100. Kadonaga, J.T., and Tjian, R. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5889-5893.

101. Laemmli, U.K. (1970) *Nature* **227**, 680-685.

102. Oakley, B.R., Kirsch, D.R., and Morris, N.R. (1980) *Anal. Biochem.* **105**, 361-363.

103. Galas, D.J., and Schmitz, A. (1978) *Nucleic Acids Res.* **5**, 3157-3170.

104. Kadonaga, J.T., Jones, K.A., and Tjian, R. (1986) *Trends Biochem. Sci.* **11**, 20-23.

105. Imagawa, M., Chiu, R., and Karin, M. (1987) *Cell* **51**, 251-260.

106. Williams, T., Admon, A., Luscher, B., and Tjian, R. (1988) *Genes Dev.* **2**, 1557-1569.

107. Sawadogo, M., and Roeder, R.G. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4394-4398.

108. Smith, J.D., Melian, A., Leff, T., and Breslow, J.L. (1988) *J. Biol. Chem.* **263**, 8300-8308.

109. Jackson, S.P., and Tjian, R. (1988) *Cell* **55**, 125-133.

110. Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J., and Tjian, R. (1987) *Cell* **48**, 79-89.

111. Sakai, D.D., Helms, S., Carlstedt-Duke, J., Gustafsson, J.-A., Rottman, R.M., and Yamamoto, K.R. (1988) *Genes Dev.* **2**, 1144-1154.

112. Paik, Y.-K., Chang, D.J., Reardon, C.A., Walker, M.D., Taxman, E., and Taylor, J.M. (1988) *J. Biol. Chem.* **263**, 13340-13349.

113. Weintraub, H. (1985) *Cell* **42**, 705-711.

114. Workman, J.L., and Roeder, R.G. (1987) *Cell* **51**, 613-622.

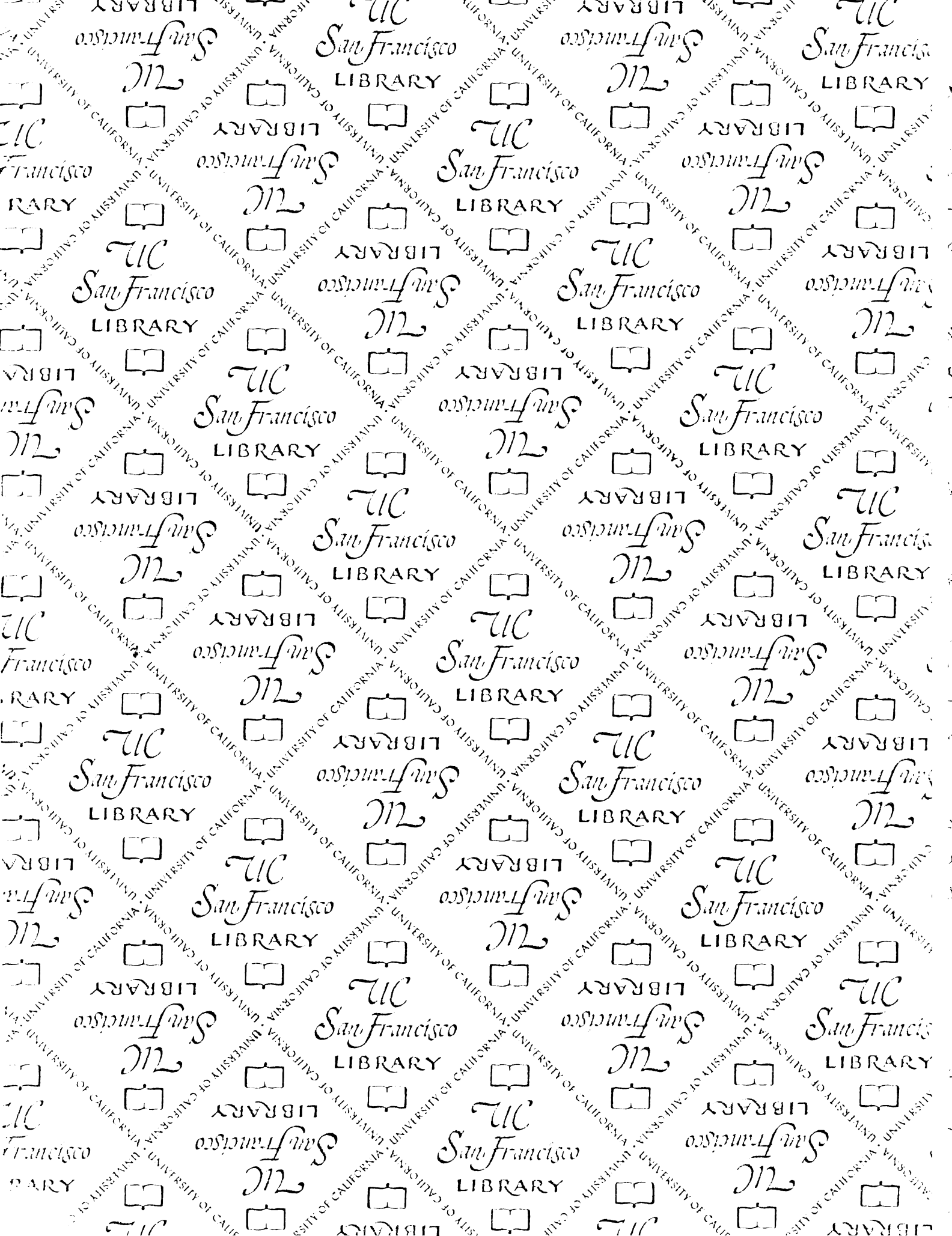
115. Grosveld, F., Blom van Assendelft, G., Greaves, D.R., and Kollius, G. (1987) *Cell* **51**, 975-985.

116. Blom van Assendelft, G., Hanscombe, O., Grosveld, F., and Greaves, D.R. (1989) Cell 56, 969-977.

117. Palmiter, R.D., and Brinster, R.L. (1986) *Annu. Rev. Genet.* **20**, 465-499.

118. Williams, D.A., Lemischka, I.R., Nathan, D.G., and Mulligan, R.C. (1984) *Nature* **310**, 476-480.

119. Quesenberry, P., and Levitt, L. (1979) *N. Engl. J. Med.* **301**, 755-760.


UC
San Francisco
LIBRARY
UNIVERSITY OF CALIFORNIA · UNIVERSITY OF CALIFORNIA


UC
San Francisco
LIBRARY
UNIVERSITY OF CALIFORNIA · UNIVERSITY OF CALIFORNIA