Three Direct Repeats and a TATA-like Sequence Are Required for Regulated Expression of the Human Low Density Lipoprotein Receptor Gene*

(Received for publication, March 18, 1987)

Thomas C. Südhof§, Deneyes R. Van Der Westhuyzen‡, Joseph L. Goldstein‡, Michael S. Brown‡, and David W. Russell††

From the §Departments of Molecular Genetics and Internal Medicine, ‡Howard Hughes Medical Institute, University of Texas Health Science Center, Southwestern Medical School, Dallas, Texas 75235

The low density lipoprotein receptor is encoded by a “housekeeping” gene that is transcribed in most mammalian cells and is subject to negative feedback regulation by sterols. To determine the basis for this regulated expression, we performed a transfection analysis with hybrid genes containing up to 6500 base pairs of 5'-flanking DNA from the low density lipoprotein receptor gene fused to the coding region of the bacterial chloramphenicol acetyltransferase gene. These studies identified a 177-base pair fragment of 5'-flanking DNA that is sufficient for expression as well as negative regulation by sterols. The positive elements within this region were further defined by analysis of a series of 15 mutations in which overlapping 10-base pair segments were scrambled by site-specific mutagenesis. These studies identified the positive elements as three imperfect direct repeats of 16 base pairs and a TATA-like sequence. The three repeats contain a sequence that is homologous to the consensus DNA sequence recognized by transcription factor Sp1.

Mammalian cells in culture satisfy their cholesterol requirements through two metabolic pathways (1). Preferentially, they obtain cholesterol through the receptor-mediated endocytosis and lysosomal hydrolysis of plasma low density lipoprotein (LDL). When this source of cholesterol is not available, as in subjects with homozygous familial hypercholesterolemia whose cells lack LDL receptors, the cells activate a multienzyme pathway that manufactures cholesterol from acetate. Both sources of cholesterol are subject to negative feedback control. When excess cholesterol accumulates within the cell, the cells repress the synthesis of new LDL receptors and at least two enzymes in the cholesterol biosynthetic pathway, 3-hydroxy-3-methylglutaryl CoA (HMG-CoA) synthase and HMG-CoA reductase (2). Recent studies employing cloned genes for these three proteins have shown that this negative feedback regulation is mediated at the transcriptional level and is determined by sequences in the 5'-flanking regions of the genes (3-7).

The previous studies (4) relied upon the HSVTK promoter to provide positive transcription signals. Although these studies defined one element that may play a role in sterol-mediated repression, they did not test other regions of the LDL receptor promoter for signals conferring either positive or negative regulation. For this purpose, we have now studied transcription of transfected genes in which the LDL receptor promoter must supply all positive and negative functions. Through use of a modification of the linker scanning mutagenesis technique (9), we have identified a 177-bp fragment of the receptor gene that contains all of the detectable signals for positive expression as well as for negative regulation by sterols. This fragment contains the three direct repeats, each of which is required for mRNA expression. It also contains the two TATA-like sequences, only one of which is required for expression. The direct repeats are homologous to the consensus sequence recognized by the eukaryotic transcription factor Sp1 (Ref. 10), suggesting that this protein may play a role in the expression of the LDL receptor gene.

EXPERIMENTAL PROCEDURES

Materials

$^{32}$P-ATP (>5000 Ci/mmol) was obtained from ICN. Enzymes used in plasmid constructions were obtained from New England Biolabs and Boehringer Mannheim. Reversed transcriptase was purchased from Life Sciences (catalog no. AMV 907). G418 sulfate (Geneticin) was purchased from Gibco Laboratories. Plasmid pSVO-CAT (12) was kindly provided by Dr. Bruce Howard (National Institutes of Health, Bethesda, MD). Newborn calf lipoprotein-deficient serum (d > 1.215...
g/ml was prepared by ultracentrifugation (13). Oligonucleotides were synthesized on an Applied Biosystems model 380A DNA synthesizer.

**Plasmid Constructions**

**LDL Receptor Promoter-CAT Genes**—A series of plasmids was constructed by standard techniques of genetic engineering (14). These plasmids contained fragments of the LDL receptor promoter, extending for various distances in the 5’ direction and terminating at position -58, linked to the bacterial gene encoding chloramphenicol acetyltransferase (CAT) (Fig. 2). The LDL receptor fragments were inserted into the unique HindIII site of pSVO-CAT, a recombinant plasmid that contains the β-lactamase gene, the origin of replication from the pBR322 plasmid, and an ampicillin resistance gene. The resulting plasmids were excised and cloned into the pSVO-CAT backbone. The resulting plasmid constructs were used to transfect CHO-K1 cells, and the resulting CAT activity was measured 24 h after transfection (15). A subclone expressing the CAT enzyme activity using a primer extension assay for mRNA levels (see below). All the cloning junctions were verified by DNA sequence analysis and restriction endonuclease mapping.

**Scramble Mutations in LDL Receptor Promoter-CAT Genes**—To construct the series of 15 promoter mutations diagrammed in Fig. 4, a 1.8-kilobase EcoRI-PstI fragment containing the LDL receptor promoter was excised from plasmid pDLR-CAT 234 (Fig. 2) and cloned into the bacteriophage M13 mp19 vector (15). Site-specific mutagenesis was performed on recombinant single-stranded M13 DNA using the single-primer method of Zoller and Smith (16). Mutagenic oligonucleotides of 40-42 bases in length were synthesized so that a single base change to be scrambled was located in either the sense or antisense orientation relative to the 10-bp sequence of the oligonucleotide. The nucleotide sequence of the 10-bp scramble was designed to retain the base composition of the wild-type sequence and to contain novel NsiI and/or SphI restriction sites. Thus, all scramble mutations had very similar sequences and could easily be identified by diagnostic restriction mapping. After annealing and extension with the large fragment of DNA polymerase I in the presence of bacteriophage T4 DNA ligase, the double-stranded M13 DNA was transformed into *E. coli* TG1 cells. Plaques containing the desired mutation were identified by hybridization with the radiolabeled mutagenic oligonucleotide, subjected to one round of plaque purification, and then sequenced by the methods of Sanger et al. (17). The single-stranded DNA was rendered double-stranded by primer extension (14), and the HindIII fragment containing the mutation was excised and cloned into the pSVO-CAT backbone. The resulting plasmid was characterized by restriction mapping with NsiI or SphI (see above) and DNA sequencing and then assigned a name according to the 10-bp sequence scrambled, e.g., pDLR-CAT -234/-219 harbors an LDL receptor promoter fragment (Fig. 2) in which the normal 10-bp sequence between -228 and -219 (GGGTTAAAAG) has been scrambled (Fig. 1).

**DNA Transfection and G418 Selection**

Stock cultures of CHO-K1 cells were grown in monolayers at 37 °C in 5-7% CO₂ in medium A (Ham’s F-12 medium containing 17 mm Hepes at pH 7.4, 2 mm glutamine, 100 units/ml penicillin, and 100 μg/ml streptomycin) supplemented with 10% (v/v) fetal calf serum. Cells were seeded at 5 × 10⁶ cells/100-mm dish in medium B (Dulbecco’s modified Eagle’s medium containing 17 mm Hepes, 34 μg/ml proline, 100 units/ml penicillin, and 100 μg/ml streptomycin) supplemented with 10% fetal calf serum. On the following day, the cells were transfected with the calcium phosphate coprecipitation technique (18) with one or two test plasmids (7.5 μg of pDLR-CAT with or without 2.5 μg of pHSVTK-CAT) together with 0.5 μg of pSVα-Neo. The cells were incubated with the DNA for 5 h and then exposed to 20% (v/v) glycerol in medium B for 4 min. Thereafter, the cells were incubated in medium B supplemented with 10% fetal calf serum for 24 h and then switched to the same medium containing 700 μg/ml G418. Selection with G418 was maintained until stable-resistant colonies could be discerned (2-5 weeks). Resistant colonies were pooled (150–600/transfection), expanded in mass culture, and screened for CAT enzyme activity (12). A subclone expressing the highest level of CAT enzyme activity was then examined for regulatory activity using a primer extension assay for mRNA levels (see below).

**Sterol Regulation Experiments**

Pooled or cloned cell lines were seeded at 2 × 10⁶ cells/100-mm dish on day 0 in medium A supplemented with 10% fetal calf serum. In the standard protocol, on day 2 the cells were washed with 5 ml of phosphate-buffered saline and fed with 8 ml of medium A containing 10% calf lipoprotein-deficient serum in place of whole fetal calf serum. This medium contained either no additions (induction medium) or a mixture of cholesterol and 25-hydroxycholesterol in a ratio of 20:1 added in 4–26 μl of ethanol (suppression medium). On day 3 after transfection, the medium was replaced for 2 h in induction or suppression medium, the cells from 12 dishes were then lysed with 1 ml of cell lysis buffer containing 6.25 g/liter lauryl sarcosine, 9.25 g/liter sodium citrate, and 0.7% (v/v) β-mercaptoethanol, and the total RNA was purified by centrifugation through CsCl (19). The RNA pellet was dissolved in buffer containing 10 mm Tris chloride and 1 mm EDTA at pH 8.0 (TE buffer), precipitated with ethanol, and then quantified by absorbance at 260 nm. Approximately 20–40 μg of total RNA were obtained from each 100-mm dish of cells.

**Primer Extension Assays**

To detect transcripts containing CAT sequences (derived from pDLR-CAT and/or pHSVTK-CAT), we used an oligonucleotide primer of 40 nucleotides complementary to bases 400–439 of the CAT mRNA (20). Transcripts from the neo gene (conferring G418 resistance) were detected with a primer of 37 nucleotides complementary to bases 1407–1443 of the transposon Tn5 (21). Endogenous hamster TK mRNA was detected with a 43-nucleotide-long primer complementary to bases 198–240 of the mRNA (22). Endogenous hamster HDL receptor mRNA was measured with a 44-nucleotide-long primer complementary to bases 41–80 of the mRNA (23). Endogenous hamster LDL receptor mRNA was detected with an oligonucleotide primer of 36 nucleotides complementary to an mRNA sequence encoded by exon 4 of the hamster gene.

Each oligonucleotide was 5'-end-labeled to a specific activity of >5000 Ci/mmol with [γ-32P]ATP and T4 polynucleotide kinase. Primers for the neo gene and the endogenous TK gene were diluted with unlabeled oligonucleotide to obtain a signal that was approximately equal in intensity to that from the test plasmid. The labeled primers (1-2 μl of a 5–10 × 10⁶ A₅₅₀ units/ml solution) were coprecipitated with 20 μg of total RNA in ethanol and resuspended in 10 μl of TE buffer and 0.27 M KCl. Hybridization was carried out for 15 min at 88 °C, after which the samples were centrifuged for 5 at 4 °C. A solution (24 μl) containing 17 units of reverse transcriptase (Life Sciences, St. Petersburg, FL), 20 mm Tris chloride (pH 8.7), 10 mm MgCl₂, 10 mm dithiothreitol, 0.4 mm dNTPs, and 0.25 μg/ml actinomycin D was added to each tube. The samples were incubated at 42 °C, diluted to 200 μl with TE buffer, extracted with 200 μl of phenol-chloroform, and ethanol precipitated. Samples were resuspended in 8 μl of TE buffer, and 12 μl of a formamide-dye solution was added. Following hybridization at 90 °C for 45 min, the samples were cooled on ice, the samples were subjected to electrophoresis for 2–3 h at 300 V on 5% polyacrylamide, 8 μt ures gels. The gels were fixed in 10% and then 1% trichloroacetic acid for 9 min each before drying in a heated vacuum dryer. ³²P-Labeled HaeIII-digested 6x174 DNA or MspI-digested pBR322 DNA was employed as molecular weight standards. The dried gels were used to expose Kodak XAR-5 film with intensifying screens at ~70 °C. Densitometry was performed on a model GS 300 scanning densitometer from Hoefer Scientific Instruments.

**RESULTS**

We constructed a series of three plasmids in which 5’ flanking sequences of the LDL receptor gene were fused to the bacterial CAT gene (Fig. 2). In describing the LDL receptor promoter, nucleotide +1 is assigned to the A of the translation initiation codon (ATG). This convention is employed because multiple transcription start sites located between positions -93 and -79 have been identified, and thus +1 cannot be used to refer to a single site of transcription initiation (Ref. 8 and Fig. 1). The LDL receptor sequences in each of the three plasmids in Fig. 2 terminated at the same 3' position (A) which is located within the transcribed region of the gene. Their 5' ends extended to different positions upstream (~6500, ~1563, and ~234). The plasmids were introduced into CHO cells by calcium phosphate-mediated transfection together with a second plasmid containing the

T. C. Südhof, J. L. Goldstein, M. S. Brown, and D. W. Russell, unpublished observations.
gene for neomycin (G418) resistance linked to the SV40 early region promoter. Permanent G418-resistant colonies were selected, and pools of approximately 150–600 colonies were assayed for expression of LDL receptor-CAT mRNA by primer extension. The cells were incubated for 24 h either in the absence of sterols (induction medium) or in the presence of a mixture of cholesterol and 25-hydroxycholesterol (suppression medium). This sterol mixture was used because it is more potent than cholesterol alone in suppressing the LDL receptor as well as other sterol-repressed genes.

Cells transfected with each of the three LDL receptor-CAT plasmids produced a fusion mRNA that initiated in the receptor cap region as determined by primer extension with a labeled oligonucleotide specific for the CAT coding sequence (Fig. 3A). To quantify this mRNA, we compared the intensity of the primer-extended fragment with the intensity of the fragment produced by primer extension upon the mRNA derived from the co-transfected neo gene (ratio b/a in Fig. 3A). The expression of this gene is driven by the constitutive SV40 early region promoter and does not respond to sterols. When the transfected CHO cells were grown in the presence of sterols, the amount of mRNA transcribed from the various LDL receptor-CAT genes was suppressed by 50–83% relative to the amount of neo mRNA (Fig. 3A).

Using the same mRNA samples as those in Fig. 3A, we estimated the amount of LDL receptor mRNA derived from the endogenous hamster receptor gene (Fig. 3B). For this purpose we used an oligonucleotide primer that is complementary to mRNA sequences encoded by exon 4 of the hamster gene that are located about 575 nucleotides 3' to the cap site. The use of such a remote oligonucleotide primer was necessary because we have only cloned sequences corresponding to exons 4–18 of the hamster LDL receptor gene. In extending over the large distance separating the primer and the 5' end of the mRNA, the reverse transcriptase encountered several strong stop sites. As a result, a family of primer-extended products was generated (Fig. 3B, far right lane). The most abundant extension products are marked by the black dots in Fig. 3B. In the presence of sterols, all of these primer-extended products were reduced in amount. On the other hand, the primer-extended product corresponding to mRNA derived from the transfected neo gene was not suppressed (Fig. 3B).

Fig. 4 shows an experiment designed to compare the sensitivity of the transfected LDL receptor-CAT promoter and the endogenous LDL receptor promoter to increasing concentrations of sterols. For this purpose, the construct that extended
A. Transfected Receptor Gene

pLDLR-CAT

Sterols

| Sterol | 234 | 234 | 1563 | 6500 |
|--------|-----|-----|------|------|
| (a) Neo | - | + | - | + |
| (b) CAT | - | + | - | + |

| Ratio b/a | 0.5 | 0.2 | 0.8 | 0.4 | 0.2 | 0.3 | 0.4 | 0.7 | 0.1 |
|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| % Suppression | 60% | 50% | 75% | 50% | 75% | 83% |

B. Endogenous Receptor Gene

LDLR 603

Sterols

| Sterol | 234 | 234 | 1563 | 6500 |
|--------|-----|-----|------|------|
| (a) LDLR  | - | + | - | + |
| (b) Neo | - | + | - | + |

| Ratio c/a | 2.4 | 0.6 | 2.7 | 0.2 | 5.9 | 0.4 | 0.7 | 2.3 | 1.1 |
|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| % Suppression | 75% | 81% | 76% | 75% |

Fig. 3. Sterol-mediated suppression of transfected and endogenous LDL receptor promoters in CHO cells. Panel A, pooled CHO cells (150 – 600 colonies) co-transfected with pSV3-Neo and the indicated pLDLR-CAT plasmid were set up for experiments as described under “Experimental Procedures.” In the case of pLDLR-CAT-234, two different pools of cells transfected on different days were studied. All cells were incubated for 20 h in the absence or presence of 10 μg/ml cholesterol plus 0.5 μg/ml 25-hydroxycholesterol, after which total RNA was isolated from 12 dishes of cells, and an aliquot (20 μg) was used as a template in primer extension assays. Each assay tube contained 32P-labeled oligonucleotides specific for the transfected neomycin resistance gene (driven by the SV40 promoter) and the CAT gene (driven by the LDL receptor promoter). The lanes on the far right show the primer extension products obtained when the CAT-specific or neomycin (Neo)-specific primers were used alone. The gel was exposed to x-ray film for 72 h. For quantitation, the amounts of neomycin (a) and CAT(b) primer extension products were estimated by densitometry, and a ratio (b/a) of CAT-specific to neomycin-specific product was calculated. Percent suppression was determined from this ratio. Panel B, the same RNA samples from Panel A were subjected to primer extension analysis using an oligonucleotide derived from exon 4 of the hamster LDL receptor gene plus the neomycin (Neo)-specific oligonucleotide. The lanes on the far right show the result obtained with the hamster LDL receptor primer alone. The black dots represent four products derived from the endogenous LDL receptor mRNA. The longest extension product (575 nucleotides, designated c) represents full length extension to the mRNA cap site. The three shorter bands represent strong-stop sequences encountered by the reverse transcriptase enzyme. Quantitation of “percent suppression” was determined as described under Panel A. The gel was exposed to x-ray film for 48 h. For Panels A and B, the positions to which DNA fragments of known size migrated are indicated on the right in nucleotides (nt).

to position −1563 in the LDL receptor 5′ flanking region (pLDLR-CAT−1563 of Fig. 2) was used. As a control, another 32P-labeled oligonucleotide primer was employed to measure the amount of cellular mRNA for TK produced by the endogenous hamster TK gene. The addition of 10 μg/ml cholesterol plus 0.5 μg/ml 25-hydroxycholesterol produced a nearly complete suppression of the LDL receptor-CAT mRNA without affecting the level of endogenous TK mRNA (Fig. 4A). This concentration of sterols also suppressed the endogenous LDL receptor mRNA in the same cells (Fig. 4B, black dots).

We also measured the amount of mRNA for another hamster cholesterol-suppressed gene, HMG-CoA synthase, using a 32P-labeled oligonucleotide primer complementary to the synthase mRNA. This primer produced two extended products which reflect the existence of two species of synthase mRNA that differ in the presence or absence of a 59-bp optional exon in the 5′ untranslated region (7). Both of these transcripts were suppressed by cholesterol plus 25-hydroxycholesterol (Fig. 4B).

Fig. 5 summarizes in graphic form the quantitative results of the experiment shown in Fig. 4. These data show that the endogenous LDL receptor mRNA and the mRNA derived from the LDL receptor-CAT 1563 construct were suppressed in parallel by the sterol mixture. On the other hand, endoge-
nous HMG-CoA synthase mRNA was more sensitive to the sterols. Complete suppression of this latter mRNA occurred at 3 \( \mu \)g/ml cholesterol plus 0.15 \( \mu \)g/ml 25-hydroxycholesterol.

To determine the time course of induction of the mRNA derived from pLDLR-CAT 1563 following the removal of sterols from the medium, we performed the experiment shown in Fig. 6. Cells were maintained in suppression medium containing 10 \( \mu \)g/ml cholesterol and 0.5 \( \mu \)g/ml 25-hydroxycholesterol for 20 h and then switched to induction medium (no sterols) for varying time periods. Total cellular RNA was then isolated and subjected to primer extension analysis using oligonucleotides specific for the LDL receptor-CAT mRNA and the endogenous hamster TK mRNA. The amount of LDL receptor-CAT mRNA rose 4-fold by 2 h after sterol removal and reached a maximum (%fold) at 11 h (Fig. 6). As expected, there was no change in the level of the endogenous TK mRNA.

The experiments of Figs. 3–6 indicate that fragments of the LDL receptor promoter that include sequences from -234 to -58 are capable of driving expression of the CAT gene in a sterol-responsive manner. To delineate further the sequences within this 177-bp fragment that confer positive and negative regulation, we constructed and analyzed the series of 15 scramble mutations shown in Figs. 1 and 7. To avoid problems associated with gross deletions, overlapping 10-bp segments of the promoter fragment from pLDLR-CAT 234 were scrambled by site-directed mutagenesis (Fig. 1). Each mutant promoter was then transfected into CHO cells on two to five separate occasions, together with pSV3-Neo for G418 selection and pHSVTK-CAT as a transfection control. The latter plasmid contains TK promoter sequences (from -108 to +55) associated with gross deletions, overlapping 10-bp segments of the promoter fragment from pLDLR-CAT 234 were scrambled by site-directed mutagenesis (Fig. 1). Each mutant promoter was then transfected into CHO cells on two to five separate occasions, together with pSV3-Neo for G418 selection and pHSVTK-CAT as a transfection control. The latter plasmid contains TK promoter sequences (from -108 to +55) derived from the HSV genome fused to the CAT gene. The HSVTK promoter, which has been well characterized by McKnight and co-workers (9,25), does not respond to sterols (4). By comparing the amount of mRNA derived from the HSVTK-CAT gene to that from a transfected LDL receptor-CAT gene, we could estimate the relative promoter strengths of the different LDL receptor mutations as well as calculate the “percent suppression” obtained in the presence of sterols. The results from one primer extension analysis are shown in Fig. 7.

RNA was isolated from CHO cells transfected with pLDLR-CAT 234 and pHSVTK-CAT and grown in the absence of sterols. When this RNA was subjected to primer extension analysis using CAT-specific and endogenous hamster TK-specific 32P-oligonucleotides, three products were visualized after autoradiography (Fig. 7, left lane of upper panel). Two products are derived from the transfected chimeric CAT genes; the 314-nucleotide band is from mRNA initiated at the correct cap site of the HSVTK-CAT gene, while the 290-nucleotide band is the product of the LDL receptor-CAT gene. The third product is a 260-nucleotide band that corresponds to the primer extension product from the endogenous TK mRNA. The addition of sterols to the medium resulted in an 83% reduction in the amount of the LDL receptor-CAT mRNA relative to the HSVTK-CAT mRNA (Fig. 7, second lane of upper panel). A similar reduction was calculated using the endogenous TK mRNA as a standard (data not shown). Experiments with the mutated LDL receptor-CAT genes yielded qualitatively similar results with respect to sterol regulation. Nine of the mutated promoters produced an mRNA whose transcription was suppressed in a normal manner in response to sterols. One mutation (-186/-177) responded less well to sterols, suppressing transcription of the LDL receptor-CAT mRNA by only 25% (Fig. 7). However, in this case the overall transcription of the mutant gene was
CAT products corresponding to mRNAs derived from the transfected LDL receptor-CAT gene (b) and from the HSVTK-CAT gene (a) as shown in Fig. 7. A value of 1.0 (dashed horizontal line) was assigned to the ratio observed when the normal LDL receptor promoter (pLDLR-CAT 234) was studied. The ratios obtained from the 15 different scramble mutations (Fig. 1) and their relative locations in the LDL receptor promoter are indicated by the heights and widths of the blocks, respectively, in the histogram. The data shown represent the average of two to five separate transfection experiments. A schematic of the normal LDL receptor promoter and its relevant landmarks is shown at the bottom of the figure.

substantially reduced, making it difficult to accurately measure suppression.

Many of the scramble mutations dramatically affected expression (Fig. 7). To compare relative transcription levels between the mutant genes, we assigned a value of 1.0 to the amount of mRNA transcribed from the transfected normal LDL receptor-CAT gene (b) and the HSVTK-CAT gene (a) as shown in Fig. 7. A value of 1.0 (dashed horizontal line) was assigned to the ratio observed when the normal LDL receptor promoter (pLDLR-CAT 234) was studied. The ratios obtained from the 15 different scramble mutations (Fig. 1) and their relative locations in the LDL receptor promoter are indicated by the heights and widths of the blocks, respectively, in the histogram. The data shown represent the average of two to five separate transfection experiments. A schematic of the normal LDL receptor promoter and its relevant landmarks is shown at the bottom of the figure.

DISCUSSION

The data presented in this paper define a minimal DNA region from the human LDL receptor gene to which expression and sterol-dependent regulation functions may be ascribed. When fused to a bacterial marker gene and transfected into CHO cells, as little as 177 bp of 5’ flanking DNA from the receptor gene (bp -234 to -58) directed the synthesis of a correctly initiated mRNA that was decreased in amount when sterols were added to the medium (Fig. 3). The amount of mRNA transcribed from this construct and its sterol response were identical to that obtained from the chimeric genes containing much larger amounts of 5’ flanking receptor DNA, indicating that no important transcription signals for regulation and expression had been deleted (Fig. 3). Titration experiments with the -1563 to -58 construct revealed that the response to sterols was equivalent for the transfected gene and the endogenous hamster LDL receptor gene (Fig. 5). The kinetics of induction of the chimeric mRNA were rapid; half-maximal expression was obtained 2 h after removal of sterols (Fig. 6).

The 177-bp fragment of receptor DNA in pLDLR-CAT 234 was small enough to allow a further delineation of transcriptionally important sequences by a form of saturation mutagenesis. The results from these studies revealed that all three of the 16-bp direct repeats are required for maximal expression (Fig. 8). In addition, the more 5’ of the two TATA-like sequences (TTGAAAT) was required. Surprisingly, a mutation that scrambled sequences in the 3’ TATA-like sequence (TGTAAT) led to a marked increase in transcription from the mutant gene (Fig. 8). The mechanism behind this promoter-up phenotype is at present not known, although we note that the mutation did not alter the start site of transcription, as an identical primer extension product was obtained with mRNA from cells transfected with this construct (Fig. 7). Neither of these TATA-like sequences match well the canonical sequence TATAAA derived from many other eukaryotic genes (26). Thus, it is conceivable that the more 3’ element may play a regulator role (as observed here) rather than serve as a signal for precise mRNA start site selection as observed in other genes. Future studies with the pLDLR-CAT -106/-100 construct should clarify the role of this sequence.

With respect to the direct repeats, mutations that alter repeat 1 decrease transcription to a slightly lesser extent (50-90%) than those that alter repeats 2 and 3 (80-95% decrease). These differences may be real, implying nonequivalence of the three repeats with respect to expression, or they may be a consequence of the exact sequences scrambled in a given repeat. In this light, it is notable that a mutation that alters few as 3 bp of a direct repeat leads to decreased transcription; mutation -203/-194 alters the 5’ three nucleotides of repeat 1 and reproducibly decreased mRNA synthesis by 50% (Fig. 8). These results imply that each direct repeat is recognized essentially in its entirety by a transcription factor or factors.

In considering which of the known transcription factors might interact with these sequences, we noted that the central core of the LDL receptor direct repeats shares sequence homology with the so-called “GC boxes” found in other eukaryotic RNA polymerase II promoters (10). Table I indicates that repeats 1, 2, and 3 have 8, 7, and 9 matches, respectively, with a 19-bp consensus GC box sequence. Tjian and colleagues (27) have recently isolated a protein termed transcription factor IIB (TFIIB) that has strong homology to repeated GC box elements.

| Repeat 1 | A A A A T C T C C T C T T G C |
| Repeat 2 | A A A A T C A C C C C A C T G C |
| Repeat 3 | A A A A T C T C C C C C C T G C |
| GC Box Consensus* | G C C C C C C C A T T |

*From Ref. 10.
factor Sp1 that stimulates transcription by binding to GC box sequences in several viral and cellular promoters (28). Initially, it was postulated that Sp1 had a recognition sequence of 10 nucleotides with a central hexanucleotide core consisting of CCGCCC that was invariant (28). This hexanucleotide is not found in any of the three direct repeats of the LDL receptor promoter (Table 1). More recently, two decanucleotide sequences that differ from the Sp1 consensus sequence by two positions in the hexanucleotide CCGCCC have been shown to bind Sp1 and activate transcription from the human immunodeficiency virus long terminal repeat promoter (29). This observation suggests that there is some flexibility in the hexanucleotide portion of the Sp1 consensus sequence and raises the possibility that the sequences in repeats 1 and 3 of the LDL receptor, which differ by 1 and 2 positions, respectively, from CCGCCC, are in fact Sp1 binding sites (Table I). This hypothesis is supported by the observation that the sequences in repeats 1 and 3 of the LDL receptor promoter are intimately related. If these sequences from digestion with DNase I (4). The insertion of a 42-bp fragment spanning these repeats into the promoter of the HSVTK gene results in protected regions spanning 20 bp each, which is similar in size to protection obtained when homogeneous Sp1 binds to its classical recognition site (27).

Multiple GC boxes and Sp1 binding sites have been demonstrated in the human immunodeficiency virus long terminal repeat (29), the SV40 early region (30), the dihydrofolate reductase promoter (31), and the human Harvey ras promoter (32), indicating that Sp1 may interact with repeated GC boxes in a synergistic manner. Consistent with this idea are the findings with the SV40 early region promoter that contains three GC boxes that are required for high level transcription. Disruption of any one of these three GC boxes reduces transcription by 50–90% (30). These results are similar to those reported here and provide further support for the idea that cooperative interactions between multiple binding sites for Sp1 or an Sp1-like protein may be involved in the transcription of mRNA from the LDL receptor promoter.

Our previous studies have indicated that direct repeats 2 and 3 of the human receptor gene function as a sterol regulatory element (4). The insertion of a 42-bp fragment spanning these two repeats into the promoter of the HSVTK gene conferred negative regulation by sterols on the expression of the chimeric gene (4). The present studies show that in addition to harboring a sterol regulatory element sequence, these two direct repeats also contain positive transcription signals. Furthermore, none of the scramble mutations analyzed in Fig. 7 produced nonregulated expression, indicating that the sterol regulatory and positive expression sequences of the LDL receptor promoter are intimately related. If these signals are the same, it is conceivable that competition for binding between a sterol repressor and Sp1 or an Sp1-like transcription factor to a direct repeat sequence may underlie the ability of sterols to repress transcription from the LDL receptor gene. Future studies centered around the purification of the proteins that interact with repeats 2 and 3 may provide support for this hypothesis.

Acknowledgments—We thank our colleagues Paul Dawson, Sandra Hofmann, Gregorio Gil, and Tim Osborne for suggestions and stimulating discussions. Amy Bolk, James Cali, and Daphne Davis provided excellent technical support.

REFERENCES

1. Brown, M. S., and Goldstein, J. L. (1986) Science 232, 34–47
2. Goldstein, J. L., and Brown, M. S. (1984) J. Lipid Res. 25, 1450–1461
3. Russell, D. W., Yamamoto, T., Schneider, W. J., Slaughter, C. J., Brown, M. S., and Goldstein, J. L. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 7501–7505
4. Sudhof, T. C., Brown, M. S., Russell, D. W., and Goldstein, J. L. (1987) Cell 48, 1061–1069
5. Luksey, K. L., Faust, J. R., Chin, D. J., Brown, M. S., and Goldstein, J. L. (1983) J. Biol. Chem. 258, 8462–8469
6. Osborne, T. F., Goldstein, J. L., and Brown, M. S. (1985) Cell 42, 203–212
7. Gil, G., Brown, M. S., and Goldstein, J. L. (1986) J. Biol. Chem. 261, 3717–3724
8. Sudhof, T. C., Goldstein, J. L., Brown, M. S., and Russell, D. W. (1985) Science 228, 815–822
9. McKnight, S. W., and Kingsbury, R. (1982) Science 217, 316–324
10. Kadonaga, J. T., Joes, K. A., and Tjian, R. (1986) Trends Biochem. Sci. 11, 26–33
11. Southern, P. J., and Berg, P. (1982) J. Mol. Appl. Genet. 1, 327–341
12. Gorman, C. M., Moffat, L. F., and Howard, B. H. (1982) Mol. Cell Biol. 2, 1044–1051
13. Goldstein, J. L., Basu, S. K., and Brown, M. S. (1983) Methods Enzymol. 98, 241–260
14. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, pp. 1–545, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
15. Messing, J. (1983) Methods Enzymol. 101, 20–78
16. Zoller, M. J., and Smith, M. (1984) DNA 3, 479–488
17. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., and Roe, B. A. (1980) J. Mol. Biol. 145, 161–178
18. van der Eb, A. J., and Graham, F. L. (1980) Methods Enzymol. 65, 826–839
19. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) Biochemistry 18, 5294–5299
20. Alton, N. K., and Vapnek, D. (1979) Nature 282, 864–869
21. Beck, E., Ludwig, G., Auerswald, E. A., Reiss, B., and Schaller, H. (1982) Gene (Amst.) 19, 327–336
22. Lewis, J. A. (1986) Mol. Cell. Biol. 6, 1998–2010
23. Gil, G., Goldstein, J. L., Slaughter, C. A., and Brown, M. S. (1986) J. Biol. Chem. 261, 3710–3716
24. Goldstein, J. L., and Brown, M. S. (1977) Annu. Rev. Biochem. 46, 897–930
25. McKnight, S. L. (1982) Cell 31, 365–365
26. Breathnach, R., and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383
27. Briggs, M. R., Kadonaga, J. T., Bell, S. P., and Tjian, R. (1986) Science 234, 47–52
28. Dyman, W. S., and Tjian, R. (1985) Nature 316, 774–778
29. Jones, K. A., Kadonaga, J. T., Luciw, P. A., and Tjian, R. (1986) Science 232, 755–759
30. Gidoni, D., Kadonaga, J. T., Barrera-Saldana, H., Takahashi, K., Chambon, P., and Tjian, R. (1985) Science 230, 511–517
31. Dyman, W. S., Sazer, S., Tjian, R., and Schimke, R. T. (1986) Nature 319, 246–248
32. Jahi, S., Kadonaga, J. T., Tjian, R., Brady, J. N., Merlino, G. T., and Pastan, I. (1986) Science 232, 1410–1413
33. Cato, A. C. B., Milsicek, R., Schutz, G., Arnemann, J., and Beato, M. (1986) EMBO J. 5, 2237–2240