Transcriptional regulation of the rat fatty acid synthase gene: identification and functional analysis of positive and negative effectors of basal transcription

Babak OSKOUIAN, Vangipuram S. RANGAN and Stuart SMITH*

Children's Hospital Oakland Research Institute, Oakland, CA 94609-1809, U.S.A.

The gene for fatty acid synthase (FAS), which contains both GC-rich sequences and a TATA box in its promoter region, is expressed in a tissue-specific manner in response to developmental, nutritional and hormonal signals. Here we report the identification of sequence elements in the 5'-flanking region responsible for modulation of basal promoter activity. Transient transfection of H4IIE hepatoma cells and 3T3-30A5 preadipocytes with plasmids containing the chloramphenicol acetyltransferase gene driven by FAS promoter sequences of different lengths revealed that two regions between nucleotides –249 and –30 contain elements capable of enhancing transcription. One of these positive regulatory elements was localized to nucleotides –241/–236 using DNase I footprinting, electrophoretic mobility-shift assays and mutagenesis. The sequence element is a typical GC box and the nuclear protein binding to

this region appears immunochemically indistinguishable from Sp1. The second positive regulatory element, an inverted CCAAT box, was localized to nucleotides –98/–92 by electrophoretic mobility-shift assays and mutagenesis. A putative negative regulatory element, initially identified by reporter gene transfection experiments, was localized between nucleotides –319 and –301 by DNase I footprinting, electrophoretic mobility-shift assays and deletion mutagenesis; this region consists of 78 % G residues. In conclusion, initiation of FAS transcription from a single start site is enhanced by the presence of an adjacent TATA motif, an inverted CCAAT box and an upstream binding site for the transcription factor Sp1; further modulation of transcription is achieved through complex interactions between these promoter elements and an upstream negative regulatory element.

INTRODUCTION

In animals, the enzymes required for the biosynthesis of fatty acids *de novo* from malonyl-CoA are integrated into a single polypeptide chain that in turn is encoded by a single-copy gene so that expression of these enzymes is automatically co-ordinated [1,2]. The cellular concentration of this multifunctional protein is regulated in a tissue-specific manner by a variety of nutritional, hormonal and developmental factors. Thus, in response to starvation, animals decrease their hepatic fatty acid synthase (FAS) concentration and restore it after the ingestion of a high-carbohydrate diet [3,4]. During the development of fetal lung, FAS is induced just before birth in epithelial type II cells, where it plays an important role in the elaboration of the phospholipid component of lung surfactant [5]. In the maturation of preadipocytes [6] and in the differentiation of mammary epithelial cells [7], induction of FAS is an essential feature of the conversion of these cells into major sites of lipogenesis. In all of these physiological processes changes in the cellular concentration of the multifunctional protein are linked to changes in the concentration of its mRNA and in many instances there is strong evidence indicating that these changes result from alterations in the rate of transcription of the gene.

Despite the central role played by the lipogenic pathway in energy homeostasis and its unique role in certain specialized tissues, little is known about how transcriptional regulation of the FAS gene is achieved. It is likely that the key to understanding this process will be found in the DNA sequences flanking the transcribed portion of the gene, and these sequences have been

identified for the rat [1] and goose [8] genes. Inspection of the 5'-flanking sequences of these FAS genes has revealed possible recognition sites for a number of transcription factors, including Sp1, AP-2 and CCAAT/enhancer-binding protein (C/EBP), as well as putative binding sites for the tri-iodothyronine and glucocorticoid receptors [1,8]. However, experimental evidence for the functional involvement of any of these sites is lacking. We report here a detailed analysis of the transcriptional activity of various parts of the 5'-flanking sequence of the FAS gene and identification of regions proximal to the promoter that appear to be important in modulating basal expression.

EXPERIMENTAL

Materials

The sources of most materials used in this study have been reported previously [9]. Five polyclonal rabbit antibodies raised against the C-terminal region of various transcriptional factors were purchased from Santa Cruz Biotechnology (Santa Cruz, CA, U.S.A.): anti-(human Sp1), anti-(mouse Egr-1), anti-[rat CCAAT/enhancer-binding protein (C/EBP α)], anti-[human upstream stimulatory factor (USF)] and anti-(human AP-2). All antibody preparations cross-react with the corresponding rat, mouse and human transcription factors. DNase I and consensus oligonucleotide for cAMP-response-element-binding protein (CREB) (5'-AGAGATTGCCTGACGTCAGAGAGCTAG-3') were purchased from Promega (Madison, WI, U.S.A.). Other oligonucleotides were synthesized by Operon Technologies (Alameda, CA, U.S.A.).

Abbreviations used: FAS, fatty acid synthase; CAT, chloramphenicol acetyltransferase; TK, thymidine kinase; *Egr-1*, early growth response 1 factor (also known as Krox 24, zif268, TIS-8 and NGF1-A); USF, upstream stimulatory factor (also known as major late transcription factor); C/EBP, CCAAT/enhancer-binding protein; CRE, cAMP-response element; CREB, CRE-binding protein.

* To whom correspondence should be addressed.

Cell cultures

Rat H4IIE hepatoma cells were cultured in minimum essential medium containing Earl's balanced-salt solution and non-essential amino acids supplemented with 10 % fetal bovine serum and 10 % bovine calf serum. The mouse preadipocyte cell line 3T3-30A5 was obtained from Dr. Ki-Han Kim (Purdue University, IN, U.S.A.) and cultured in Dulbecco's modified Eagle's medium (DME-H21) containing 10 % fetal bovine serum. Adipocyte differentiation was induced by nutrient deprivation followed by insulin treatment [10]. Cells were grown at 37 °C under 5 % CO₂.

Transient transfections

3T3-30A5 and H4IIE cells were grown to 60–70 % confluency on 60 mm² dishes and were transfected with 10 µg of plasmid DNA per plate by the calcium phosphate precipitation method [11] as described previously [9]. Transfection efficiency was normalized by co-transfection of the cells with 2–5 µg of a β-galactosidase expression plasmid, pCH110 (Pharmacia LKB, Uppsala, Sweden) together with the DNA of interest. The lowest β-galactosidase activity measured in each experiment was arbitrarily assigned the value of unity.

Stable transfections

H4IIE and 3T3-30A5 cell lines with stably integrated –350/+68 and –249/+68 FAS–chloramphenicol acetyltransferase (CAT) genes were obtained by separately co-transfecting the cell lines with plasmids containing these genes together with the linearized pSV-Neo plasmid. Stable clones were selected by their resistance to the antibiotic G418, and approx. 50 colonies were pooled [12]. The relative number of copies of reporter genes integrated into chromosomal DNA was assessed by Southern-blot analysis [13].

Enzyme assays

A radiochemical enzyme assay was used for the determination of CAT activity. Reaction mixtures contained 0.25 M Tris/HCl buffer, pH 7.8, 12 µM D-threo-[dichloroacetyl-1-¹⁴C]chloramphenicol (50 Ci/mol) and 1 mM acetyl-CoA in a final volume of 100 µl. Each mixture was incubated at 37 °C for 1–4 h, depending on the level of CAT activity, and the product isolated and quantified [9]. β-Galactosidase activity was determined spectrophotometrically [13].

Preparation of nuclear extracts

Nuclear extracts from H4IIE hepatoma and 3T3-30A5 cells (both before and after differentiation into adipocytes) were prepared essentially as described by Dignam et al. [14] except that the protease inhibitors leupeptin (1 µg/ml) and pepstatin (1 µg/ml) were used in all buffers, in addition to PMSF (1 mM).

Electrophoretic mobility-shift assays

DNA fragments from the promoter region of the FAS gene were 5'-end-labelled with [γ-³²P]ATP (3 Ci/µmol) using T4 polynucleotide kinase [13]. Nuclear extracts were first incubated at either 20 °C or 0 °C for 20 min in 20 µl of reaction mixture containing 20 mM Hepes, pH 7.9, 1 mM dithiothreitol, 60 mM KCl, 0.5 mM EDTA, 5 % glycerol and 20–60 µg/ml double-stranded poly(dI-dC). Approx. 1 pmol of radiolabelled DNA probe (10000–50000 c.p.m./pmol) was then added, incubation was continued for a further 30 min at 20 °C or 0 °C, and the DNA–protein complexes were identified [9].

Supershift assays

These assays were routinely performed in two different ways. Either the antibodies were added to the nuclear extracts before the addition of the radioactive probe or the probe was allowed to interact with the nuclear extracts and then the antibodies were added. No differences were observed in the electrophoretic profiles obtained by the two procedures. The reaction conditions were the same as described for electrophoretic mobility-shift assays.

DNase I footprinting assays

Various amounts of nuclear extract, or BSA, were preincubated with 20–60 µg/ml double-stranded poly(dI-dC) at 0 °C for 20 min in 20 mM Hepes, pH 7.9, containing 60 mM KCl, 1 mM dithiothreitol, 0.5 mM EDTA and 5 % glycerol contained in a final volume of 50 µl. The end-labelled probes were then added and the mixtures were incubated for a further 40 min at 0 °C. An equal volume of a solution containing 10 mM MgCl₂, 10 mM CaCl₂ and 0.2 unit of DNase I was then added, the digestion was allowed to proceed for 3 min at 20 °C and the products were identified [9]. Footprint location was ascertained by establishing the position of G + A nucleotides using chemical cleavage sequencing reactions of the same radiolabelled fragment [15].

Construction of plasmids containing a series of chimaeric FAS-reporter genes

Plasmids containing various FAS–CAT fusion genes were generated by engineering a series of nested deletions from the FAS portion of the DNA using exonuclease III and mung bean nuclease [16] as described previously [9].

Construction of the FAS–CAT fusion plasmids carrying mutations

Mutations were introduced in selected regions of the FAS–CAT fusion plasmids using *in vitro* site-directed mutagenesis as described by Kunkel et al. [17].

Construction of FAS–thymidine kinase (TK) plasmids

Plasmid –124/–30-FAS–TK was constructed as follows: plasmid –124/+68 was digested with *Hind*III and *Dra*I, and the *Hind*III site was blunt-ended by treatment with the Klenow fragment of DNA polymerase I. This fragment was then ligated to pTK–CAT which had been digested with *Xba*I and similarly treated with the Klenow enzyme. The orientation of the FAS insert was determined by DNA nucleotide sequencing.

The –249/–121-FAS–TK plasmid was constructed by isolating a *Hind*III–*Sma*I restriction fragment from the FAS gene that encompassed nucleotides –249 to –121 and cloning it into the *Hind*III–*Bam*HI sites of plasmid pTK–CAT, after first blunting the vector *Bam*HI site using Klenow polymerase.

Plasmid –350/–268-FAS–TK was constructed by cutting plasmid –350/+68 first with *Ban*II, blunting the restriction site with mung bean nuclease and then cutting with *Hind*III. The 82 bp fragment produced was isolated and cloned into *Hind*III–*Bam*HI sites of plasmid pTK–CAT, after first blunting the vector *Bam*HI site using Klenow polymerase.

RESULTS

Preliminary characterization of promoter elements using the chimaeric reporter gene approach

In order to identify the likely location of regulatory elements in the FAS promoter, we first constructed a series of plasmids

Figure 1 Promoter activity of 5'-flanking sequences of the FAS gene assessed in transiently transfected cells

A series of 5'-deletions was engineered in the upstream region of the FAS gene and the constructs cloned in front of a promoterless CAT gene. H4IIE cells (■) and 3T3-30A5 preadipocytes (▨) were transiently transfected with an equal mass of plasmids containing the chimaeric reporter genes; CAT activity was assayed in cell extracts and the data were normalized according to the molar quantity of each plasmid transfected. Each value represents the mean $\pm$ S.D. at least six independent experiments, using different plasmid preparations. Transcriptional activities were normalized on the basis of molar quantity of reporter gene used in the transfection, and the highest activity, that observed in cells transfected with the -249/+68 construct, was arbitrarily set at 100%. The range of observed CAT activities corresponded to acetylation of 3–5% and 12–15% of the chloramphenicol substrate per h for H4IIE and 3T3-30A5 cells respectively. The difference in the mean values for the -350 and -249 constructs was statistically significant for 3T3 cells ($P < 0.01$) but not for H4IIE cells ($P < 0.2$).

containing CAT reporter genes under the transcriptional control of various upstream sequences derived from the FAS gene. All of these constructs have a common 3' end corresponding to FAS position +68 (where +1 is the start of transcription) but they carry different lengths of upstream sequences from the FAS gene. The plasmids were transiently transfected into a rat hepatoma cell line (H4IIE) and a mouse preadipocyte cell line (3T3-30A5). Cells were harvested after 48 h and CAT activities in the cytosols were determined. The longest construct used contained 1604 bp of DNA originating upstream of the FAS transcription start site. Progressive 5'-deletions of the upstream FAS sequence up to position -249 resulted in a corresponding increase in CAT activity in transfected H4IIE cells (Figure 1). In these cells maximum CAT activity was observed when the CAT fusion construct carried 249 bp of FAS upstream sequence. Activity was at least 5-fold higher in cells transfected with the -249 construct than in those transfected with the -1604 construct. This finding suggested that negative regulatory elements might be present upstream of position -249. Qualitatively similar results were obtained when the same series of FAS-CAT reporter genes were introduced into 3T3-30A5 preadipocytes, i.e. the basal promoter activity increased as progressive deletions were made from position -1604 to -249. In this case, however, the overall increase was less than 2-fold. When 5'-deletions were extended to nucleotides -149 and -124, a substantial fall in CAT activity was observed in both transfected H4IIE and 3T3-30A5 cells. This finding is suggestive of the presence of positive regulatory elements located between nucleotides -249 and -124. Further deletion of nucleotides to the -62 position resulted in still lower CAT activity. No activity was detectable in

Figure 2 DNase I footprint analysis of the FAS promoter region

Radiolabelled probes corresponding to FAS nt -350/-23 (A), -249/-23 (B) or -149/-23 (C) were incubated with H4IIE nuclear extract, then digested with DNase I as described in the Experimental section. Left and centre, lane 1, no extract; lanes 2, 3 and 4, 10, 20 and 30 μ g of nuclear extract respectively. Right, lane 1, no extract; lanes 2 and 3, 25 and 50 μ g respectively.

cells transfected with the vector pCAT-Basic that lacked inserted FAS sequences. Our results observed with the 5'-deletion constructs are qualitatively similar to those reported earlier by Wolf et al. [18] using H4IIE cells and Moustaid et al. [19] using 3T3-L1 cells who also found that inclusion of certain sequences derived from the upstream promoter region of the FAS gene lowered expression of the reporter gene. On the other hand, possibly because of inclusion of a different selection of upstream FAS sequences in the chimaeric constructs, a subsequent study reported by Moustaid et al. [20] did not give any indication of the possible presence of an upstream negative regulatory element(s).

Use of DNase I footprinting to identify upstream regions of the FAS promoter capable of binding nuclear proteins

The results of the FAS-CAT reporter gene 5'-deletion experiments suggested that both positive and negative regulatory elements might be located within the first 350 nt upstream of the transcriptional start site. The possible presence of regulatory elements in this region was further investigated by DNase I footprinting. Using 5'-end-labelled probes representing nt -350/-23, -249/-23 and -149/-23, three regions between -321 and -284, -243 and -214 and -74 and -52 were clearly protected from DNase I digestion when increasing amounts of H4IIE nuclear extract were included in the assay (Figure 2). The -350/-23 and -149/-23 probes were also used in DNase I footprinting assays with nuclear extracts prepared from 3T3-30A5 cells, before and after differentiation, with the same result (not shown). Under the conditions used for footprinting, protection from DNase I digestion resulting from binding of nuclear proteins to the TATA box (nucleotides -32

Table 1 Assessment of the effect of FAS nt —350 to —249 on promoter activity in stably transfected cells

H4IIE cells and 3T3-30A5 preadipocytes were stably transfected with the —350/+68 and —249/+68 FAS-CAT. Preadipocytes were either uninduced or induced to differentiate for 4 (intermediate) and 10 days (adipocyte) before assay of CAT activity. Each value represents the mean $\pm$ S.D. for four independent experiments. The difference in CAT activity for cells transfected with the —350/+68 and —249/+68 constructs was statistically significant for both the H4IIE cells ($P < 0.002$) and 3T3-30A5 cells ($P < 0.03$) at all stages of differentiation.

Constructs used	Chloramphenicol acetylated (%/h)			
	3T3-30A5			
	H4IIE	Preadipocyte	Intermediate	Adipocyte
—350/+68	8.3 $\pm$ 0.43	1.3 $\pm$ 0.03	3.4 $\pm$ 1.23	7.2 $\pm$ 1.26
—249/+68	15.8 $\pm$ 1.12	5.3 $\pm$ 0.92	13.2 $\pm$ 2.7	27.0 $\pm$ 2.6

to —27) is not evident possibly because the motif is too close to the 3' end of the probe.

Functional analysis of the DNase I footprinted regions of the proximal promoter region of the FAS gene

In subsequent studies we used several different strategies to characterize sequences within the first 350 nt of the FAS promoter that had been identified as possibly containing regulatory elements.

(i) The —321/—284 footprinted region

The location of this footprint coincided with the tentative location of a negative regulatory element(s), based on the initial 5'-deletion experiments with the chimaeric reporter genes.

Although H4IIE cells or 3T3-30A4 preadipocytes transiently transfected with several different plasmid preparations containing the —249/+68 construct consistently expressed higher CAT activity than those cells transiently transfected with six different plasmid preparations containing the —350/+68 construct, the difference in activity was small and statistically significant only in the case of the 3T3-30A5 cells. In order to evaluate the possible presence of a negative regulatory element between nt —350 and —249 of the FAS gene, we prepared clones of the cell lines containing stably integrated —350/+68 and —259/+68 FAS-CAT genes. These stably transfected cells showed markedly higher CAT expression when driven by the —249/+68 region than the —350/+68 region of the FAS promoter; furthermore the difference was maintained as the 3T3-30A5 preadipocytes differentiated into mature adipocytes (Table 1). The difference could not be attributed to a markedly different frequency of integration of the two constructs into the host DNA, since quantitative Southern-blot analysis revealed that the group of clones carrying the —249/+68 construct contained only about 20% more copies of the CAT gene than those carrying the —350/+68 construct (details not shown). These results strengthened the circumstantial evidence suggesting that a DNA element capable of down-regulating transcription is present between nt —350 and —249 of the FAS promoter.

The sequence specificity of the DNA-protein interaction associated with the footprinted putative negative regulatory element was further investigated by electrophoretic mobility-shift assays. Incubation of a radiolabelled probe encompassing sequences between —350 and —268 with nuclear extracts obtained from either H4IIE cells or 3T3-30A5 preadipocytes and

Figure 3 Analysis of the specificity of protein binding to the FAS promoter region —325 to —267 using the electrophoretic mobility-shift assay

(A) The nucleotide sequence of a segment of the FAS promoter region is presented together with bars denoting the regions represented in the synthetic oligonucleotides used in competition studies. (B) A DNA fragment corresponding to FAS sequences —350/—268 was radiolabelled at position —350 and used as probe with nuclear extracts from H4IIE and 3T3-30A5 cells. Various unlabelled oligonucleotides, at 100-fold molar excess, were tested for their ability to compete with the labelled probe. Lane 1, free probe alone; lanes 2–10, nuclear extract from H4IIE cells included; lanes 11 and 12, nuclear extract included from 3T3-30A5 cells before and after differentiation respectively; lane 2, no competitive oligonucleotide included; lane 3, oligonucleotide no. 1 included; lane 4, oligonucleotide no. 2 included; lane 5, oligonucleotide no. 4 included; lane 6, oligonucleotide no. 3 included; lane 7, oligonucleotide corresponding to the CRE site included; lane 8, antibodies against Sp1 included; lane 9, antibodies against *Egr-1* included; lane 10, antibodies against AP-2 included.

adipocytes resulted in the formation of three specific DNA-protein complexes that exhibited slower mobilities than did the free probe (Figure 3B, lanes 2, 11 and 12 respectively). Formation of all complexes was prevented when the nuclear extract was first heated at 65 °C for 10 min. To narrow down further the specific nucleotides involved in complex-formation, we assessed the ability of several oligonucleotides representing different segments of the region between nt —325 and —278 to compete with the radioactive —350/—268 probe for the available nuclear proteins (Figure 3A). In a control experiment, an oligonucleotide containing the consensus sequence for factor CREB did not compete with the radiolabelled probe (Figure 3B, lane 7). FAS oligonucleotides no. 1 (nt —325 to —311) and no. 3 (nt —302 to —279) were also poor competitors (Figure 3B, lanes 3 and 6

Table 2 Effect of mutation of the G-rich region, nt -300 to -319 , on FAS promoter activity

The G-rich sequence between -300 and -319 , CCGCGGGGGTGGGGGGTGA, was mutated to GGCGACGCTGTCCAAGGTGG in the $-350/+68$ FAS-CAT construct. Both the original and mutated $-350/+68$ FAS-CAT constructs as well as the $-249/+68$ FAS-CAT construct were used to transfect 3T3-30A5 preadipocytes, and CAT activity was measured 48 h later. Each value represents the mean $\pm$ S.D. for four independent experiments.

Constructs used	Chloramphenicol acetylated (%/h)
$-350/+68$	5.19 ± 0.46
$-350/+68$ mutant	7.47 ± 0.5
$-249/+68$	9.03 ± 0.89

respectively). On the other hand, FAS oligonucleotide no. 2 (nt -310 to -290) eliminated formation of two of the three radiolabelled complexes (Figure 3B, lane 4) and FAS oligonucleotide no. 4 (nt -317 to -299) blocked formation of all three radiolabelled complexes (Figure 3B, lane 5). The most effective inhibitor, oligonucleotide no. 4, contained 18 of the 25 nt protected in the DNase I footprinting experiment whereas the second most effective inhibitor, oligonucleotide no. 2, contained 14 nt from the 3' end of the footprinted region. The ineffective inhibitors, oligonucleotides no. 1 and no. 3, contained relatively shorter sequences of the footprinted region from the 5' and 3' ends respectively. These experiments strongly suggested that protein(s) present in nuclei of H4IIE hepatoma cells and 3T3-30A4 cells, both before and after differentiation, bind to a recognition site within a G-rich sequence, GCG₅TG₇TGA, of the FAS gene. Based on this evidence it seemed reasonable to conclude that binding of the *trans*-acting factor(s) to this sequence may be responsible for the apparent downregulation of transcription observed when nt -350 to -249 were included in the transfected FAS-CAT reporter genes.

Evidence in support of this hypothesis was sought by analysing the effect on transcription produced by replacing the G-rich region between -301 and -319 with an unrelated sequence. Transfection of 3T3-30A5 preadipocytes with a $-350/+68$ FAS-CAT plasmid carrying the mutation resulted in significantly increased reporter gene expression compared with that observed using the unmutated construct (Table 2). This elevated level of transcription was comparable with that obtained with $-249/+68$ FAS-CAT plasmid which also lacks the G-rich sequence. These results clearly indicate that sequences between -319 and -301 down-regulate the FAS gene transcription.

(ii) The $-243/-214$ footprinted region

The DNase I protection experiments had revealed the presence of a footprint, extending from nt -243 to -214 , within the region ($-249/-124$) initially suspected, on the basis of the 5'-deletion study, of containing a positive regulatory element(s). Further characterization of this putative regulatory element was sought using electrophoretic mobility-shift assays. When a radiolabelled fragment encompassing the region between -249 and -121 was mixed with nuclear proteins derived from either H4IIE or 3T3-30A5 cells, two major DNA-protein complexes were observed in both cases, suggesting that a nuclear protein(s) interacting with this region is present in both preadipocyte (results not shown) and hepatoma cell lines (Figure 4, lane 2). Synthetic oligonucleotides spanning the footprinted region could prevent the formation of DNA-protein complexes formed by the interaction of H4IIE nuclear proteins with the $-249/-178$

Figure 4 Characterization of proteins binding to the FAS promoter region $-249/-121$

A radiolabelled DNA fragment representing FAS nt $-249/-121$ was used as a probe in electrophoretic mobility-shift assays. Lane 1, free probe alone; lane 2, $10 \mu\text{g}$ of H4IIE nuclear extract; lane 3, $10 \mu\text{g}$ of H4IIE nuclear extract with anti-Sp1 antibodies; lanes 4 and 5, nuclear extract with anti-*Egr-1* antibodies; lane 6, purified human Sp1; lane 7, purified Sp1 and anti-Sp1 antibodies. In the experiments illustrated in lanes 3, 4 and 7 the probe was mixed with nuclear proteins before addition of antibodies whereas in the experiment shown in lane 5 the antibodies were mixed with nuclear proteins before the addition of the probe. The anti-(human Sp1) and anti-(mouse *Egr-1*) antibody preparations employed in these experiments cross-react with the homologous rat transcription factors (see the Experimental section).

probe (results not shown), suggesting that the observed complexes probably result from the specific interaction of nuclear proteins with nucleotides present in the footprinted region. Formation of all complexes was prevented by prior heating of the nuclear extract for 10 min at 65°C .

The footprinted region extending from nt -243 to -214 , which is characterized by an abundance of G and C residues, coincides with two potential Sp1-binding sites containing the asymmetric GGGCGG hexanucleotide core (GC box). One site, GC box 1, is located at nt -225 to -220 , the other, GC box 2, at nt -241 to -236 . We therefore sought additional experimental evidence to evaluate the possibility that the well-characterized transcriptional activator Sp1 might be responsible for formation of one or more of the complexes with the $-249/-121$ probe. Indeed one of the two major complexes exhibited an electrophoretic mobility identical with that formed by purified human Sp1 (Figure 4, lane 6). Furthermore, in the presence of anti-(human Sp1) polyclonal antibodies, the mobility of this complex was retarded. This 'supershift' was observed whether the complex was formed from the nuclear extract or from purified Sp1 (Figure 4, lanes 3 and 7). In contrast, antibodies raised against *Egr-1*, which also binds to GC-rich sequences [21], did not alter the mobility of any of the observed complexes (Figure 4, lanes 4 and 5). These results offer strong evidence that footprinting of the FAS gene between nt -243 and -214 results, at least in part, from Sp1 binding in this region and are in accord with the results of the FAS promoter deletion studies indicating the presence of a positive regulatory element(s) between nt -249 and -121 . The ability of sequences between nt -243 and -214 to activate transcription from the FAS promoter was further evaluated by mutating the Sp1-like sequences between -225 and -220 (GC

Table 3 Effect of mutations in putative regulatory elements on the FAS promoter activity

The original or mutated $-249/+68$ FAS-CAT constructs were transfected into H4IIE cells and CAT activity assessed 48 h later. Where indicated, GC boxes 1 (-225 to -220) and/or 2 (-241 to -236) were mutated from GGCGG to ATATC; in the GC box 1 and 2 deletion, nt from -240 to -220 were removed. In the inverted CCAAT box mutation, nt -99 to -92 (CATTGGCC) were changed to GTCCAAGG and in the E box mutation nt -69 to -61 (AGCCCATGT) were changed to TAAAAACCC. Each value represents the mean $\pm$ S.D. for four independent experiments.

Constructs used	Chloramphenicol acetylated (%/h)
$-249/+68$ FAS-CAT	5.5 ± 0.29
GC box 1 mutation	9.2 ± 1.51
GC box 2 mutation	2.5 ± 0.59
GC boxes 1 and 2 mutations	2.0 ± 0.66
$-240/-220$ deletion	1.3 ± 0.07
CCAAT box mutation	2.6 ± 0.42
E box mutation	5.1 ± 0.84

Table 4 Effect of various regions from the 5'-flanking sequence of the FAS gene on activity of the TK promoter

FAS DNA sequences corresponding to nt $-350/-268$, $-249/-121$ and $-124/-30$ were cloned in front of the TK promoter in the plasmid TK-CAT and used to transfect H4IIE cells and 3T3-30A5 preadipocytes. CAT activities were determined in cell extracts and compared with activities in extracts prepared from cells transfected with the TK-CAT control plasmid. Each value represents the mean $\pm$ S.D. for at least four to six independent experiments.

Construct	Induction of CAT activity (fold)	
	H4IIE cells	3T3-30A5 preadipocytes
$-350/-268$ TK-CAT	1.2 ± 0.2	2.3 ± 0.1
$-249/-121$ TK-CAT	3.6 ± 0.4	2.2 ± 0.3
$-124/-30$ TK-CAT	2.6 ± 0.3	5.2 ± 0.2

box 1) and -241 and -236 (GC box 2), in the context of the $-249/+68$ FAS-CAT fusion construct. H4IIE cells were transfected with either the original or the mutated constructs and the CAT activities compared (Table 3). Mutation of sequences in GC box 2 resulted in about 50% lower promoter activity, but surprisingly mutation of GC box 1 increased the promoter activity about 2-fold. The construct carrying mutations in both GC boxes also yielded about 50% lower CAT activity compared with that of the original construct. When both GC boxes were removed by deleting nucleotides between -240 and -220 , promoter activity was reduced approx. 70%. Thus the results of the mutational analysis are also consistent with the presence of a positive regulatory element within the $-243/-214$ footprinted region.

Finally, the ability of these sequences to function in the context of a heterologous promoter was evaluated by cloning a fragment encompassing FAS sequences $-249/-121$ into a CAT expression vector driven by the basal promoter of the TK gene. H4IIE and 3T3-30A5 cells when transfected with the plasmid $-249/-121$ TK-CAT exhibited elevated CAT activities (3.6- and 2.2-fold respectively), compared with cells transfected with TK-CAT control plasmid (Table 4), confirming the presence of a positive regulatory element in this region.

Figure 5 Analysis of proteins binding to the FAS promoter region $-149/-23$ using the electrophoretic mobility-shift assay

A radiolabelled DNA fragment representing FAS nt $-149/-23$ was incubated with H4IIE nuclear extract and DNA-protein complexes were detected by the electrophoretic mobility-shift assay. Lane 1, free probe alone; lanes 2–4, 10 μ g of nuclear extract; lanes 3 and 4, 100- and 200-fold excess of an unlabelled oligonucleotide corresponding to the FAS promoter sequence $-73/-52$; lane 5, 100-fold excess of an unlabelled oligonucleotide corresponding to the FAS promoter sequence $-242/-225$; lane 6, anti-(rat C/EBP α) antibody included.

(iii) The $-74/-52$ footprinted region

The earlier DNase I protection experiments using a probe spanning the region $-149/-23$ had revealed the presence of a footprint within the $-149/-62$ region, originally suspected on the basis of the 5'-deletion experiments of possibly containing a positive regulatory element. When the same probe was used in an electrophoretic mobility-shift assay, three DNA-protein complexes were detected. Inclusion of 100-fold molar excess of an unlabelled oligonucleotide representing a remote region of the FAS promoter, $-242/-225$, had no effect on the formation of the radiolabelled DNA-protein complex (Figure 5, lane 5). However, in the presence of an excess of an unlabelled $-73/-52$ oligonucleotide, formation of one of three radiolabelled complexes was completely eliminated (Figure 5, lanes 3 and 4). Inspection of the nucleotide sequence of the $-73/-52$ region revealed the presence of a sequence CATGTG between nt -65 and -60 that matches the E-box consensus sequence CANNTG. Since E-box-binding proteins such as USF are known to function as positive regulators of transcription, we examined the possibility that the presence of this putative E box might account for both the footprint at $-74/-52$ and the transcription activation effect localized to the $-149/-62$ region by the 5'-deletion experiments. However, this possibility was discounted by the finding that mutation of nt $-69/-61$, in the context of the $-249/+68$ FAS-CAT reporter gene construct, did not affect CAT expression, either in H4IIE cells (Table 3) or 3T3-30A5 preadipocytes (details not shown). The possibility that the $-74/-52$ footprint may mark the location of the insulin-response element is discussed later in the text.

Functional analysis of the inverted CCAAT box at $-92/-98$

Inspection of the nucleotide sequence of the FAS promoter had previously identified an inverted CCAAT box at $-92/-98$ [1].

Figure 6 Analysis of proteins binding to the inverted CCAAT box region of the FAS promoter using the electrophoretic mobility-shift assay

A radiolabelled DNA fragment representing FAS nt $-106/-85$ was incubated with H4IIE nuclear extract, and DNA-protein complexes were detected by the electrophoretic mobility-shift assay. Lane 1, free probe alone; lanes 2–4, 10 μ g of H4IIE nuclear extract included; lane 3, 100-fold excess of an unlabelled oligonucleotide containing the wild-type sequence from $-106/-85$ included; lane 4, 100-fold molar excess of an unlabelled oligonucleotide containing the inverted CCAAT box mutation (Table 3) included.

Since we were unable to find evidence for a DNA-protein interaction in this region using DNase I footprinting we utilized the more sensitive electrophoretic mobility-shift assay. Indeed when the region $-106/-85$ was used as a probe in the presence of H4IIE nuclear extract, formation of a DNA-protein complex was observed (Figure 6, lane 2). Production of this radiolabelled complex was eliminated in the presence of a 100-fold molar excess of unlabelled $-106/-85$ oligonucleotide but not in the presence of the unlabelled oligonucleotide in which the inverted CCAAT box had been mutated (Figure 6, lanes 3 and 4 respectively); the result of this experiment indicated that formation of the DNA-protein complex is attributable to a specific interaction of a nuclear protein with the inverted CCAAT box region. To assess the possible functional significance of this interaction, we mutated the inverted CCAAT box within the $-249/+68$ FAS-CAT construct and examined the effect on promoter activity using H4IIE cells. Introduction of the mutation lowered the promoter activity by approx. 50%, indicating that these sequences are required for optimal promoter activity (Table 3).

Finally, to determine whether the positive regulatory activity associated with this region of the FAS gene could function in the context of a heterologous promoter, we constructed a plasmid carrying FAS nucleotides -124 to -30 cloned upstream of the TK promoter in the plasmid TK-CAT. H4IIE and 3T3-30A5 preadipocytes transfected with this plasmid exhibited 2.6- and 5.2-fold higher CAT activity respectively than cells transfected with the TK-CAT control plasmid (Table 4). The result confirmed the presence of a positive regulatory element (or elements) in the region $-124/-30$ that can function out of context of the FAS gene.

DISCUSSION

FAS is encoded by a single-copy gene of approx. 20 kb [1] in the rat and approx. 50 kb in the goose [8]. Transcription is initiated

at the same site in lipogenic tissues such as liver, lactating mammary gland and lung, 27 nt downstream from an unusual TATA box having a T in the second position: TTTAAAT in the rat [1], TTAAATA in the goose [8]. Inspection of the 5'-flanking sequence had revealed the presence of potential binding sites for a number of transcription factors in addition to the TATA motif [1,8] but direct experimental evidence for the functionality of these sequence elements had been lacking. The objective of this study therefore was to identify by functional analysis sequence elements in the proximal promoter region of the rat FAS gene responsible for modulation of the basal promoter activity and to look for nuclear proteins that specifically recognize and bind to these sequences. A summary of our findings is illustrated in Figure 7.

One region, encompassing nt $-249/-121$, when coupled to either FAS or a heterologous promoter, increased expression of the reporter gene, and a positive regulatory element was localized within this region at nt $-243/-214$ by DNA-protein-binding experiments and mutational analysis. This sequence element consists of 87% G+C residues and includes two closely positioned GC boxes, potential binding sites for transcription factors of the Sp1 type. Indeed gel mobility-shift and supershift experiments showed that a protein present in nuclear extracts that interacted with this region was indistinguishable from authentic Sp1 (Figure 4) and oligonucleotides representative of both GC box 1 and 2 were capable of binding authentic Sp1 (results not shown). The observation that mutation of GC box 2 in the context of the $-249/+68$ FAS-CAT reporter gene led to decreased CAT expression is also consistent with the presence of a functional positive regulatory element at this site. However, the functional significance of GC box 1 is less clear, since the introduction of mutations at this site actually increased expression of the reporter gene in the context of the $-249/+68$ FAS-CAT construct. Possibly binding of an Sp1-like protein to this site is unproductive and hinders binding to the functionally important site nearby in GC box 2 or perhaps the extent of modulation of transcription is determined by a direct interaction between the SP1 bound to GC box 2 and a second protein bound to GC box 1.

A second region that increased transcription when coupled to either FAS or a heterologous promoter was associated with nt $-124/-30$, and several well-characterized transcription factors were evaluated as candidates interacting with this region of the gene. Studies reported by Moustaid et al. [20], as well as those presented here (Figure 2), have shown that nt $-74/-52$ are protected from DNase I digestion by nuclear proteins and the interaction at this site has been implicated in the positive regulation of transcription induced by insulin [20]. Since these authors implicated nt $-68/-60$ (GCCCCATGTG) in the rec-

Figure 7 Map of FAS promoter region

NRE, negative regulatory element; CCAAT, inverted CCAAT box; IRE, insulin-response element. All regulatory elements appear to be functionally operative in both H4IIE and 3T3-30A5 cells.

ognition of the *trans*-acting factor conveying insulin responsiveness [20], we investigated the possibility that this sequence element might also contribute to the enhanced transcription associated with the $-124/-30$ region of the gene. Within the $-68/-60$ region is a potential E box (CATGTG) very similar to the E boxes identified as carbohydrate-response elements in the pyruvate kinase [22] and S_{14} [23] genes, CCCGTG and CACGTG respectively. We considered USF, one of the transcription factors known to bind to E boxes, as a *trans*-acting factor that might mediate transcriptional activation of the FAS gene. However, whereas USF is known to be a heat-stable protein, the protein(s) binding to this region of the FAS promoter appears to be heat-labile. Furthermore none of the three protein-DNA complexes formed with H4IIE nuclear extract exhibited altered electrophoretic mobilities in the presence of antibodies that recognize USF (details not shown). Finally, since mutation of this putative E box, in the context of the $-249/+68$ FAS-CAT construct, did not affect expression of the reporter gene, it would seem that the presence of this sequence elements is unlikely to account for the enhanced basal transcription associated with nt $-124/-30$. Since completion of our study, however, Wang and Sul [24] have reported that USF1, which is greatly increased in the livers of starved-re-fed rats, can bind to the E box located at $-65/-60$; these authors have suggested that USF1 may play a role in mediating the stimulatory effect of insulin on transcription of the FAS gene.

One alternative candidate we considered that might contribute to the enhanced basal transcription associated with nt $-124/-30$ was a putative GC box at nt $-90/-85$. However, since we found that this region was not footprinted by nuclear extracts (Figure 2) and the electrophoretic mobilities of the DNA-protein complexes formed using nt $-149/-23$ as labelled probe were unaffected by anti-Sp1 antibodies (details not shown), we concluded that Sp1 is unlikely to be responsible for the enhanced transcription associated with nt -149 to -23 . Nevertheless a recent report by Bennett et al. [25] demonstrates that purified Sp1 is capable of binding to the putative GC box at nt $-90/-85$. These authors have suggested that Sp1 binding to this region, together with binding of the sterol-regulatory-element-binding protein to adjacent regions of the promoter, facilitates the regulation of FAS transcription by sterols. The reason for these apparently differing conclusions is unclear, and further investigation is required to define conditions that might influence the binding of Sp1 to this region of the gene. In our experiments, strong evidence was obtained indicating that the sequence in this region of the gene that most likely contributes to enhanced transcription is located at nt $-98/-93$. This sequence element, when inverted, matched the consensus motif for the family of CCAAT-box-binding proteins active in the transcription of eukaryotic genes [26]. The ability of this sequence element to bind to nuclear proteins was confirmed by electrophoretic mobility-shift assays. Furthermore the results of mutational analysis of the $-98/-93$ region were consistent with the presence of a positive regulatory element at this location. One well-characterized CCAAT-box-binding protein, C/EBP α , has been implicated in the expression of several genes involved in lipid metabolism, including acetyl-CoA carboxylase [27]. Nevertheless, several lines of evidence suggest that the *trans*-acting factor binding to the inverted CCAAT box of the FAS gene is not a member of the C/EBP family of proteins. First, the DNA-protein complexes observed in electrophoretic mobility-shift assays were not formed when the H4IIE nuclear extract was preheated for 10 min at 65 °C, conditions under which C/EBP proteins are stable. Secondly, the electrophoretic mobility of none of the three DNA-protein complexes formed was altered in the presence of

anti-C/EBP α antibodies (Figure 5, lane 6). Exactly the same result was obtained using fully differentiated 3T3-30A5 cells, which are known to be an abundant source of C/EBP α [26]. Thirdly, an unlabelled oligonucleotide containing the C/EBP consensus binding-site does not compete with a labelled oligonucleotide containing the FAS inverted CCAAT box for binding to recombinant C/EBP α ; conversely, an unlabelled oligonucleotide containing the FAS inverted CCAAT box does not compete with a labelled oligonucleotide containing the C/EBP consensus binding site for binding to recombinant C/EBP α but this DNA-protein complex formed is 'supershifted' by the anti-C/EBP α antibodies [9]. It seems likely then that an as yet unidentified factor, not C/EBP, interacts with the positive response element associated with the inverted CCAAT box. Recent work in our laboratory has also implicated this inverted CCAAT box as an essential element required for mediation of the transcriptional regulation by cAMP [9].

The possibility that the 5'-flanking region might contain a sequence element(s) that mediates a negative influence on transcription was first raised by results of the 5'-deletion experiments (Figure 1 and Table 1). Subsequent experiments focused specifically on the proximal promoter region up to -350 , although the possibility that an additional negative element is located further upstream cannot be discounted at this time. Using DNA-protein-binding assays we were able to narrow down the location of the negative regulatory element to the sequence GCG₅TG₇TGA. Mutation of this region increased expression of the reporter gene in the context of the $-350/+68$ FAS-CAT construct, an observation also consistent with the role of this region of the gene as a negative regulatory element. This sequence element shows some similarity to the so-called 'GT and GC boxes [28]' recognized by members of the Sp1 multigene family of transcription factors and to a promoter element of the chicken β -globin gene that contains a run of 16 G residues that reportedly binds a tissue-specific transcription factor BGP 1 [29]. However, transcription factors Sp1, *Egr-1* and AP-2 were ruled out as candidates interacting with this region by the following criteria. Purified human Sp1 protein did not interact with the radiolabelled $-350/-268$ probe (results not shown) and none of the DNA-protein complexes formed by the interaction of the $-350/-268$ probe and H4IIE nuclear extract could be supershifted by anti-Sp1, anti-*Egr-1* or anti-AP-2 antibodies (Figures 3B, lanes 8, 9 and 10). Identification and characterization of this nuclear protein with an unusual DNA-recognition sequence may be of help in understanding how transcription of the FAS gene is regulated. Surprisingly, sequences containing the putative negative element, when removed from the context of the FAS gene and coupled to the TK promoter, had no effect on transcription in the hepatoma cell line and differentiated 3T3-30A5 cells but exerted an enhancing effect (about 2.3-fold) on transcription in preadipocytes (Table 4). So it would appear that both the context of the surrounding DNA sequences and the differentiation state of the cell may influence the biological role of this sequence element.

In conclusion, although expression of the FAS gene is regulated in a tissue-specific manner, the promoter region of the FAS gene resembles that of a constitutively expressed housekeeping gene in that it contains a GC-rich region encompassing at least one authentic binding site for the transcriptional activator Sp1. In addition to the TATA-factor- and Sp1-binding sites, an inverted CCAAT box also plays the role of a transcriptional activator, and a sequence element further upstream functions as a down-regulator. Further work is required to elucidate the mechanism by which both positive and negative factors interact to determine the level of expression.

We are grateful to Mr. Dennis Pham for excellent technical assistance in the early stages of this project and Dr. Christopher Amy for preparing some of the fatty acid synthase constructs. We thank Dr. Ki-Han Kim for kindly providing the 3T3-30A5 cell line, Dr. Fattah Kashanchi and Dr. Ramin Sheikhattar for their helpful discussions and Dr. Andrzej Witkowski and Dr. Anil Joshi for critical reading of the manuscript. The research was supported by Grant DK16073 from the National Institutes of Health.

REFERENCES

- 1 Amy, C. M., Williams-Ahlf, B., Naggert, J. and Smith, S. (1990) *Biochem. J.* **271**, 675–679
- 2 Amy, C. M., Williams-Ahlf, B., Naggert, J. and Smith, S. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1105–1108
- 3 Volpe, J. J. and Vagelos, P. R. (1976) *Physiol. Rev.* **56**, 339–417
- 4 Goodridge, A. G. (1986) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **45**, 2399–2405
- 5 Rooney, S. A. (1989) *J. Physiol. (London)* **257**, L195–L201
- 6 Paulauskis, J. D. and Sul, H. S. (1989) *J. Biol. Chem.* **264**, 574–577
- 7 Smith, S. and Ryan, P. (1979) *J. Biol. Chem.* **254**, 8932–8936
- 8 Kameda, K. and Goodridge, A. G. (1991) *J. Biol. Chem.* **266**, 419–426
- 9 Rangan, V. S., Oskouian, B. and Smith, S. (1996) *J. Biol. Chem.*, **271**, 2307–2312
- 10 Park, K. and Kim, K. (1991) *J. Biol. Chem.* **266**, 12249–12256
- 11 Kingston, R. E., Chen, C. A. and Okayama, H. (1989) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D., Seidman, J. G., Smith, J. A. and Struhl, K., eds.), pp. 9.1.1–9.1.3, John Wiley & Sons, New York
- 12 Southern, P. J. and Berg, P. (1982) *J. Appl. Gen.* **1**, 327–341
- 13 Sambrook, J., Fritsch, E. F. and Maniatis, T. (eds.) (1989) in *Molecular Cloning: A Laboratory Manual*, pp. 16.66–16.67, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- 14 Dignam, J. D., Lebovitz, R. M. and Roeder, R. G. (1983) *Nucleic Acids Res.* **11**, 1475–1489
- 15 Maxam, A. M. and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560
- 16 Henikoff, S. (1984) *Gene* **28**, 351–359
- 17 Kunkel, T. A., Roberts, J. D. and Zakour, R. A. (1987) *Methods Enzymol.* **154**, 367–382
- 18 Wolf, S. S., Hofer, G., Beck, K.-F., Roder, K. and Schweizer, M. (1994) *Biochem. Biophys. Res. Commun.* **203**, 943–950
- 19 Moustaid, N., Sakamoto, K., Clarke, S., Beyer, R. S. and Sul, H. S. (1993) *Biochem. J.* **292**, 767–772
- 20 Moustaid, N., Beyer, R. S. and Sul, H. S. (1994) *J. Biol. Chem.* **269**, 5629–5634
- 21 Kriwacki, R. W., Schulz, S. C., Steitz, T. A. and Caradonna, J. P. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9759–9763
- 22 Bergot, M.-O., Diaz-Guerra, M.-J. M., Puzeat, N., Raymondjean, M. and Kahn, A. (1992) *Nucleic Acids Res.* **20**, 1871–1878
- 23 Shih, H.-M. and Towle, H. C. (1994) *J. Biol. Chem.* **269**, 9380–9387
- 24 Wang, D. and Sul, H. S. (1995) *J. Biol. Chem.* **270**, 28716–28722
- 25 Bennett, M. K., Lopez, J. M., Sanchez, H. B. and Osborne, T. F. (1995) *J. Biol. Chem.* **270**, 25578–25583
- 26 Birkenmeier, E. H., Gwynn, B., Howard, S., Jerry, J., Gordon, J. L., Landschulz, W. H. and McKnight, S. L. (1989) *Genes Dev.* **3**, 1146–1156
- 27 Tae, H.-J., Luo, X. and Kim, K.-H. (1994) *J. Biol. Chem.* **269**, 10475–10484
- 28 Hagen, G., Müller, S., Beato, M. and Suske, G. (1992) *Nucleic Acids Res.* **20**, 5519–5525
- 29 Lewis, C. D., Clark, S. P., Falsenfeld, G. and Gould, H. (1988) *Genes Dev.* **2**, 863–873