Isolation and characterization of the human stearoyl-CoA desaturase gene promoter: requirement of a conserved CCAAT cis-element

Lin ZHANG*1, Lan GE*, Tai TRAN*2, Kurt STENN*† and Stephen M. PROUTY*3


Stearoyl-CoA desaturase is the rate-limiting enzyme in the production of mono-unsaturated fatty acids. We have recently cloned and characterized the human Scd CDNA and SCD (the stearoyl-CoA desaturase structural gene) on chromosome 10, as well as the non-transcribed pseudogene on chromosome 17. In order to further define SCD regulation and function, we have isolated and characterized the promoter of the structural gene. Screening of chromosome-10-specific libraries resulted in the isolation of 4.1 kb of SCD sequence upstream of the translation start site. Binding sites for transcription factors critical for mouse Scd1 and Scd2 promoter activity, such as sterol-regulated-element-binding protein and nuclear factor Y, were present in the human SCD promoter (Scd is the mouse stearoyl-CoA desaturase gene). Deletion analysis in HaCaT keratinocytes identified a critical region for promoter activity between nts 496–609 upstream of the translation start site. Site-directed mutagenesis of binding sites in this region identified the CCAAT box as the critical cis-element for SCD promoter activity. An electrophoretic mobility-shift assay confirmed that this element binds nuclear proteins from HaCaT keratinocytes. The polyunsaturated-fatty-acid (PUFA) response element, previously identified in the promoters of mouse Scd1 and Scd2, was found to be conserved in the human SCD promoter, and contained the critical CCAAT cis-element. A minimal promoter construct including this region was responsive to fatty acids, with oleate and linoleate decreasing transcription and stearate increasing it. These studies indicate that CCAAT-box-binding proteins activate SCD transcription in cultured keratinocytes and that fatty acids modulate transcription, most likely through the conserved PUFA response element.

Key words: lipid metabolism, monosaturated fatty acids, palmitoleate, transcription.

INTRODUCTION

Stearoyl-CoA desaturase is a lipid-metabolic enzyme that catalyses the insertion of a double bond between carbon atoms 9 and 10 in the saturated fatty acids (SFAs) palmitoyl-CoA and stearoyl-CoA to generate the mono-unsaturated fatty acids (MUFAs) palmityloyl-CoA and oleoyl-CoA respectively. MUFAs play important roles in many processes, including energy metabolism (triaclyglycerol storage in adipose tissue) [1], membrane fluidity [2], anti-oxidation [3], signal transduction [4,5], apoptosis [6] and senescence [7]. Moreover, a critical role in differentiation is evidenced by the lack of development of the sebaceous and meibomian glands in the asebia mutant mouse, due to loss of Scd1 function [8]. Altered stearoyl-CoA desaturase gene expression is associated with various diseases such as cancer [9], obesity [10], and aging [11]. The pleiotropic nature of MUFAs requires complex regulation of stearoyl-CoA desaturase genes that includes diverse signals such as insulin [12], carbohydrate [13,14], fatty acids [15] and temperature [16]. Furthermore, tight regulation of the rat Scd protein is indicated by a 3-4 h half-life, due to the action of an integral membrane-bound protease in the endoplasmic reticulum [17].

In order to gain an insight into the complex regulation of SCD genes, nucleotide sequences have been isolated from many organisms, including plants, bacteria, yeasts, fish, cows, rats, mice [18], fruitflies (Drosophila melanaster) [19], nematode worms (Caenorhabditis elegans) [20], hamsters [21], sheep [22] and humans [23]. At the levels of nucleotide and amino acid sequences, the stearoyl-CoA desaturase gene family exhibits both similarities and differences across the phylogenetic Kingdoms [24]. The similarities are (1) the high levels of sequence identity of the open reading frames (ORF) among the orthologous genes, (2) the presence of conserved amino acids (and their spacing) involved in catalysis of the A desaturation reaction, and (3) the presence of a long 3′-untranslated region (3′-UTR) in the vertebrate genes. The major difference is the variable number of stearoyl-CoA desaturase genes between different species. Those organisms with a single structural gene include cyanobacteria [25], the yeast Saccharomyces cerevisiae [26], sheep [22] and human [23], whereas those with more than one structural gene include the yeast Mortierella alpina [27], D. melanaster [19], C. elegans [20], mouse [28,29] and rat [30]. The reason for the multiple stearoyl-CoA desaturase genes is unclear, although there is evidence to suggest that stearoyl-CoA desaturase proteins

Abbreviations used: AP-2, activator protein-2; RFX-1, X-box-binding regulatory factor-1; FES2, fat-specific element-2; SCD, stearoyl-CoA desaturase; Scd, human stearoyl-CoA desaturase gene; Scdl, mouse stearoyl-CoA desaturase gene (‘stearoyl-CoA desaturase’ indicates non-human and non-murine genes, as well as cDNA, mRNA and protein from any species); EMSA, electrophoretic-mobility-shift assay; LFA, unsaturated fatty acid; MUFAs, mono-unsaturated fatty acids; PUFA, polyunsaturated fatty acid; SFA, saturated fatty acid; ORF, open reading frame; UTR, untranslated region; SREBP, sterol-regulated-element-binding protein; NF, nuclear factor; CDS, coding sequence; Sp1, stimulatory protein 1, all sequences in the present paper are numbered relative to the translation start site, designated +1.

1 Present address: R&D Department, N-3, Sigma-Aldrich/Biotechnology, P.O. Box 14508, St. Louis, MO, 63178, U.S.A.
2 Present address: UMDNJ–NJMS, Student Affairs, 185 South Orange Avenue, Newark, NJ, 07103, U.S.A.
3 To whom correspondence should be sent, at present address: Drug Discovery, R. W. Johnson Pharmaceutical Research Institute, Welsh & McKean Roads, P.O. Box 776, Spring House, PA 19477, U.S.A. (e-mail sprouty1@prius.jnj.com).

The nucleotide sequence reported in this paper has been submitted to the DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under the accession number AF320307.
in *C. elegans* [20] and mouse [31] have different substrate specificities. This hypothesis is further supported by the finding of tissue-specific expression of the rodent stearoyl-CoA desaturase homologues [28–30,32], whereas stearoyl-CoA desaturase in human [23] and sheep [22] is ubiquitously expressed. Further species-specific regulatory mechanisms are indicated by the use of tandem polyadenylation signals by human *SCD* to generate two alternate mRNA transcripts [23], each encoding the identical enzyme, whereas the rodent *SCD* genes each generate a single mRNA transcript [28–30].

Owing to the diverse nature of the processes in which stearoyl-CoA desaturase functions, the regulation of stearoyl-CoA desaturase gene expression involves both transcriptional and post-transcriptional mechanisms [15,33]. Studies of transcriptional regulation have focused on functional analysis of promoters, these having been isolated from yeast *OLE1* [18], mouse *Scd1* [28], mouse *Scd2* [29], rat stearoyl-CoA desaturase 1 [30] and sheep stearoyl-CoA desaturase [22]. The mouse *Scd1* and *Scd2* promoters have been extensively analysed and cis-elements identified that regulate transcription [15,34], particularly through the transcription factors sterol-regulated-element-binding protein (SREBP) and nuclear factor (NF) Y [33].

In order to understand the role of *SCD* in human health and disease, we recently isolated the full-length human stearoyl-CoA desaturase cDNA and *SCD* structural gene, demonstrating considerable homology with the mouse *Scd* genes, particularly with respect to the ORF, intron–exon junctions and the long 3′-UTR [23]. Unlike the mouse and the rat, which have at least two alternate mRNA transcripts [23], each encoding the identical enzyme, whereas the rodent *SCD* genes each generate a single mRNA transcript [28–30].

**EXPERIMENTAL**

**Genomic fragment isolation**

Isolation of the λ-phage encompassing the 5′ end of the *SCD* gene (4 kb of promoter, exon 1, intron 1, exon 2 and part of intron 2) from a chromosome-10-specific genomic library was previously obtained using standard techniques [23]. The 5.6 kb insert was cloned as an *EcoRI* fragment into pBlueScript (Stratagene, La Jolla, CA, U.S.A.) (clone 19) and subsequently used to generate both nuclease-protection-assy probes and luciferase-reporter-assy constructs as described below.

**Luciferase reporter construction**

To generate p(−3348/−150)*SCD*-luc (c1), clone 19 was digested with *PstI* and *SacII* and blunted with T4 polymerase. The 3.2 kb fragment containing only the 5′ end of the gene was cloned into a blunted *HindIII* site of the luciferase vector pGL3-basic (Promega, Madison, WI, U.S.A.). p(−2328/−150)*SCD*-luc (c2) and p(−1295/−150)*SCD*-luc (c4) were generated by digesting the clone 19 with *TaqI*/SacII and *KpnI*/SacII respectively, and the resulting fragments were subcloned into a blunted *HindIII* site of pGL3-basic. p(−1537/−150)*SCD*-luc (c3) and p(−882/−150)*SCD*-luc (c5) were subsequently generated by digesting c1 with *SacI* and *Smal* respectively. The resulting fragments were re-ligated using T4 ligase. p(−461/−150)*SCD*-luc (c6) was generated by digesting c4 with *EcoR1* and c2 with *NcoI* and the resulting fragment was ligated to blunted *Neol*/*SacII* sites of pGL3-basic. p(−270/−150)*SCD*-luc (c7) was created by digesting c4 with *Hinfl* and *Neol*. The resulting fragment was subcloned into *Neol*/*HindIII* digested blunted pGL3-basic. p(−753/−150)*SCD*-luc (c8), p(−609/−150)*SCD*-luc (c9) and p(−496/−150)*SCD*-luc (c10) constructs were assembled by PCR using one of the three specific sense oligonucleotides (bp 753 to −732, c8), 5′-GTTCCACACTGTGT-CTCTGAGA-3′; (−609 to −688, c9) 5′-GATGGCCGGGC-AGAGGCCCCAGCG-3′; (−496 to −474, c10), 5′-GGGAC-CCGCAAGAGAAGTTGCGC-3′; and a common antisense oligonucleotide (−166 to −145) 5′-CCGGCGGCGGTTTGGAG-GTCCCCG-3′. All PCR reactions were conducted with proof-reading Turbo *PFU* DNA polymerase (Stratagene), and the original c5 construct was used as template. Amplification products were phosphorylated with T4 kinase and subcloned into the *HindIII*-blunted site of pGL3-basic. The sequence of the PCR products were examined for accuracy using the dyeoxy-nucleotide-chain-termination method.

**RNase protection analysis**

Total RNA was extracted, using RNAzol (Tel-Test Inc., Friendswood, TX, U.S.A.), from human hair-follicle biopsies and the human keratinocyte cell line, HaCaT [35]. Total RNA from human brain and mRNA from human liver were purchased from Clontech, Palo Alto, CA, U.S.A. A *SCD* fragment was amplified with PCR, using clone 19 as a template with the following primers: sense (−275 to −256), 5′-GCCATCACTCTCTC-GCACT-3′; antisense (+7 to +27), 5′-ATCGCTCCGTGAGC-AAGTGGGC-3′. This resulted in a 302-bp fragment (nucleotides −275 to +27) which was subcloned into the TA cloning vector pCR2.1 (Invitrogen, Carlsbad, CA, U.S.A.). The plasmid (NPA1-pCR2.1) was further digested with *XbaI* and *EcoRV* to release unnecessary vector sequence between the insert and the T7 promoter and re-ligated with T4 ligase. The final construct (NPA2-pCR2.1) was linearized by *HindIII* and a labelled antisense RNA (400 bp) was synthesized from the T7 promoter of the plasmid using T7 RNA polymerase (Ambion, Austin, TX, U.S.A.). Following synthesis, the full-length probe was gel-purified. Total RNA from human hair follicles (15 μg), human brain (10 μg), and HaCaT keratinocytes (2, 5 and 10 μg) and mRNA from human liver (5 μg) was hybridized to the labelled transcript for 16 h at 42 °C. The RNA samples were digested with an RNase mixture (Ambion), and the sizes of the protected RNA were determined by electrophoresis on a denaturing 5%- (w/v)-polyacrylamide gel, using a DNA sequencing ladder generated from NPA2-pCR2.1.

**Nuclear extracts and electrophoretic-mobility-shift assays (EMMASs)**

Nuclear extracts of HaCaT keratinocytes were prepared using standard procedures [36]. Protein concentrations of the extract
were determined with BCA assay (Pierce, Rockford, IL, USA). Electrophoretic mobility shift assays were performed with a Gel Shift assay system according to the manufacturer’s instructions (Promega, Madison, WI, U.S.A.). Briefly, hybridization was performed at room temperature for 20 min in a volume of 20 μl consisting of 0.1 pmol of double-stranded CCAAT oligonucleotide probe (70000 c.p.m.) and 5 μg of nuclear extract. The gel shift mixture was incubated for 20 min at room temperature, and samples were immediately electrophoresed at 150 V for 3 h on a non-denaturing 5% polyacrylamide gel using 0.5 x TBE running buffer (34 mM Tris/borate/0.75 mM EDTA) pre-run at 100 V for 100 min. The gel was then dried for autoradiography. For competition studies, radiolabeled DNA competitor was added as a 100-fold molar excess and preincubated with the nuclear extract at room temperature for 5–10 min. The ²²P-labelled DNA was then added to the mixture and incubated at room temperature. The competitor oligonucleotides used for EMSAs were as follows: activator protein-2 (AP-2) (sense, –595 to –575), 5’-GCCAAGCAGGCGGTTGGAAGAG-3’; AP-2 (antisense), 5’-TCCTTCCACCCCGGCTGAGCAGC-3’; stimulatory protein 1 (Sp1) (sense, –559 to –539), 5’-AACAGAGGGGAGGGAGCGAGA-3’; Sp1 (antisense), 5’-TCGCTCCCCCTCCCCCTCAGTT-3’; CCAAT (sense, –509 to –489), 5’-GGCGCCAGAGCAATGGGGAGCAG-3’; CCAAT (antisense), 5’-CCGGTTGCCATGGCTCGGCG-3’.

Transfection and dual luciferase assay
HaCaT keratinocytes were used as the recipient cells for transient transfection assays. On the day prior to transfection, cells were seeded into six-well plates at 400000 cells/well. Transient transfections were performed with 2 μg of test plasmids using SuperFect Reagent (Qiagen, Valencia, CA, U.S.A.) according to the manufacturer’s protocol. As a control for transfection efficiency and cell density, HaCaT cells were then transfected with a 50 ng of Renilla luciferase control reporter vector driven by the thymidine kinase promoter (pRL-TK). At 48 h post-transfection, adherent cells were washed once with 3 ml of PBS, dissolved in 500 μl of cell culture lysis reagent (Promega, Madison, WI, USA), and harvested after 30 min at room temperature. Luciferase assays were performed immediately using a Dual-Luciferase Reporter Assay System kit (Promega, Madison, WI, USA), and measurements were taken using a TD-20/20 luminometer (Turner Designs, Inc., Sunnyvale, CA, U.S.A.). The promoter activity was reflected by the ratio of firefly/Renilla luciferase for each construct. Duplicate or triplicate readings were taken for each assay, and all assays were replicated a minimum of three times over a minimum of two independent experiments. Results of statistical tests are indicated in the Table and Figure legends.

Mutagenesis
Mutation in the binding sites for transcription factors Sp1 and CCAAT were introduced into the plasmid p9, containing the 0.5 kb SCD promoter subcloned into the blunted HindIII site of pGL-3 basic, using the Quick Change Site-Directed Mutagenesis kit (Stratagene). Two double-stranded oligonucleotide primers containing the desired mutations are as follows (underlined base-pairs denote mutant substitutions): Sp1 (–566 to –532), 5’-AAGGAGAAACCAGAGAAAGGGGAGCGGAGGAGCTG-3’, and CCAAT (–517 to –483), 5’-AGCAGATTGCCGGAGAAATATTGGCAGGCAGAC-3’.

Cell culture
All reagents were from Gibco BRL (Rockville, MD, U.S.A.) unless otherwise stated. HaCaT keratinocytes were routinely cultured in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal-calf serum (Sigma, St Louis, MO, U.S.A.), 2 mM l-glutamine and 50 units/ml penicillin + streptomycin. The cells were incubated in humidified incubators at 37 °C and 5%, CO₂. For fatty acid treatments, HaCaT keratinocytes were exposed to fatty acid–BSA conjugate (100 μM) for 24 h prior to harvest. Stock solutions (2.75 mM) of fatty acid–BSA solutions were prepared as follows: 8.5 mg of steartate (Sigma) was dissolved in 1 ml of 0.1 M NaOH, then heated to 90 °C for dissolution. This solution was added dropwise to 6 ml of 10% fatty-acid-free BSA (Sigma) with stirring at 37 °C. This solution was then adjusted to 10 ml with distilled water, filtered through a 0.22 pm-pore-size filter, and stored at 4 °C. The same procedure was followed for oleate (Sigma); however, heating was kept to a minimum. The BSA solution used for controls was prepared following above procedure, but without added fatty acids. Linoleate–BSA conjugate was purchased from Sigma and reconstituted to 3 mM with distilled water. All stock solutions were diluted to 100 μM in cell-culture media for application to cells.

DNA sequencing
The DNA sequence in both directions was determined by Howard Hughes Medical Institute Biopolymer/Keck Foundation Bio-technology Resource Laboratory (Yale University, New Haven, CT, U.S.A.) using PerkinElmer Biosystems (Foster City, CA, U.S.A.) model 337 gel DNA sequencer and model 3700 capillary DNA sequencer.

Sequence comparison of the 5’ ends of the human and mouse SCD genes
The mouse Scd1 sequence used for alignment was obtained from GenBank* and used as follows: the promoter, 5’-UTR and proximal coding sequence (CDS) (1–885; M21280) had a G inserted at nt 698 in order to conform to the published sequence [28,37] to generate a 886-nt sequence. Mouse Scd2 sequence used for alignment was obtained from GenBank* and used as follows: the promoter (nts 1–587; M26269) was joined to the 5’-UTR proximal CDS of the mRNA (nts 1–324; M26720) to generate a 911-nt sequence. A 910-nt section of human SCD sequence was used for alignment, consisting of 887 nts 5’ to the ATG start site, plus 24 nts of ORF (nts 3127–4126; AF320307). The three sequences were aligned using the ClustalW algorithm from the Vector NTI sequence analysis software. In the alignment window, various subregions (as described in the Results section) were highlighted and percentage identity calculated for the three-way and pairwise comparisons. Pairwise comparisons were made by removing one of the three sequences from the three-way alignment, then recalculating the percentage identity between the remaining pair, a procedure that preserves the original alignment.

RESULTS
Characterization of the 5’-region of the SCD gene
Previously we have shown that there are two SCD loci in human genome, with one being the functional gene present on chromosome 10, and the other a pseudogene located on chromosome 17 [23]. To obtain the upstream sequence of the functional SCD
gene, chromosome-10-specific libraries were screened with a 1.9 kb fragment (from clone 34) containing the 5’ end of the Scd cDNA [23]. One of the genomic clones obtained from this screen had a 5.6 kb insert [23]. Analysis of the insert indicated that there was approx. 4.1 kb upstream of the ATG translation start site.

Putative transcription-factor binding sites were identified using the MatInspector Professional algorithm with the Transfac database v4.1 [38] using core similarity setting of 0.75 and matrix similarity setting of 0.85. Many classes of transcription-factor binding sites were identified, consistent with the pleiotropic function of SCD genes (Table 1, Figure 1).

### Identification of the transcription initiation site

Screening for Scd in a cDNA library from cultured human foreskin keratinocytes resulted in two clones which extend 5’ from the ATG translation initiation site, these being clone 34 to −235 and clone 45 to −229 [23] (arrowheads, Figure 1). Further definition of the transcriptional start site was carried out for cultured human keratinocytes, human hair-follicle keratinocytes (the major human skin cell type to express Scd mRNA in vivo; results not shown), human brain, and human liver. Primer extension and RNase protection were carried out with total RNA isolated from HaCaT keratinocytes, hair-follicle biopsies, brain, and mRNA from liver. Owing to the high GC content in this region, primer extension did not yield consistent results. For the RNase protection assays, a cRNA probe, consisting of 302 nts of SCD (excluding vector sequence) from −275 to +27 was used in the protection assay. A major band of 172 nts was detected in RNA isolated from liver and hair follicles, whereas a major band of 300 nts was seen in brain and HaCaT keratinocyte RNA (Figure 2). No bands of the size of the full-length probe (400 nts, including vector sequence) were detected, indicating lack of contaminating sense DNA template and that of HaCaT cells. This indicates that, for liver and hair follicles, the major site of transcription initiation is at −35 nts downstream of the proximal TATA box at position 1 (single-lined box in Figure 1) from the translation start site. This further extends to that of mouse Scd/ in 3T3-L1 cells [28]. In brain and cultured keratinocytes, the transcription start site is 5’ to that of hair follicles, as indicated by the 300-nt protected fragment in brain and HaCaT cell RNA and the termination of cDNA clones 34 and 45 in the region of the distal TATA box located at −234 (Figure 1). There are minor protected bands in brain, liver, and HaCaT cells, with a common band of 220 nts (Figure 2). This indicates the presence of a minor transcription start site at −193.

### Table 1 Putative transcription-factor binding sites*

<table>
<thead>
<tr>
<th>Transcription factor</th>
<th>Binding site</th>
<th>Position(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP-2</td>
<td>CCCS</td>
<td>−165†</td>
</tr>
<tr>
<td>CREB</td>
<td>TGAC</td>
<td>−1377, −1145</td>
</tr>
<tr>
<td>NF-1</td>
<td>TGGC</td>
<td>−1872, −1744, −1851, −1241, −827, −502, −459, −244, −194</td>
</tr>
<tr>
<td>NF-1</td>
<td>CCAA</td>
<td>−3961, −3898, −501, −458</td>
</tr>
<tr>
<td>Oct-6</td>
<td>AATT</td>
<td>−177</td>
</tr>
<tr>
<td>RXF-1</td>
<td>GYAA</td>
<td>−202, −495, −196</td>
</tr>
<tr>
<td>SP-1</td>
<td>GGGG</td>
<td>−1038, −930, −551, −314, −384</td>
</tr>
<tr>
<td>TATA-BP</td>
<td>TAAA</td>
<td>−234, −179</td>
</tr>
<tr>
<td>Tissue — specific</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HNF-3β</td>
<td>TRTT</td>
<td>−2730, −2424, −1958, −217</td>
</tr>
<tr>
<td>Whn</td>
<td>ACGC</td>
<td>−2290, −942, −858, −811, −615, −291, −155</td>
</tr>
<tr>
<td>Lipid metabolism</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C/EBP</td>
<td>GMMA</td>
<td>−2913, −1946, −1124, −440, −420, −219</td>
</tr>
<tr>
<td>PPAR</td>
<td>AAAG</td>
<td>−1166, −257</td>
</tr>
<tr>
<td>SREBP</td>
<td>TCAC</td>
<td>−4027, −3692, −3484, −3612, −2236, −2010, −5178</td>
</tr>
<tr>
<td>Steroid-hormone nuclear receptors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER</td>
<td>GTCA</td>
<td>−1738, −1359</td>
</tr>
<tr>
<td>RAR-α1</td>
<td>GTTC</td>
<td>−9060, −2228, −1904, −1762</td>
</tr>
<tr>
<td>T3R</td>
<td>GTTC</td>
<td>−2228</td>
</tr>
<tr>
<td>VDR</td>
<td>GAGG</td>
<td>−324, −222</td>
</tr>
<tr>
<td>Immune system</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRF-1</td>
<td>GAAA</td>
<td>−3612, −2521</td>
</tr>
<tr>
<td>NF-κB</td>
<td>GOGGA</td>
<td>−2556, −2525, −1047, −1041, −868, −862, −727, −281</td>
</tr>
<tr>
<td>c-Rel</td>
<td>TTCC</td>
<td>−1689, −220</td>
</tr>
<tr>
<td>STAT</td>
<td>GAAA</td>
<td>−1138</td>
</tr>
<tr>
<td>Proliferation/differentiation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP-1</td>
<td>TGAC</td>
<td>−3116, −1678, −454, −271, −204</td>
</tr>
<tr>
<td>c-Myc</td>
<td>CACG</td>
<td>−941, −291</td>
</tr>
<tr>
<td>N-Myc</td>
<td>CGTG</td>
<td>−292</td>
</tr>
</tbody>
</table>

* MatInspector Professional output from Transfac database v4.1 with core similarity at ≥ 0.75 and matrix similarity at ≥ 0.85.
† Core binding site, 5′ to 3′ direction (IUPAC nucleotide code).
‡ Numbering is relative to translation initiation site (+1).
§ Novel SRE identified in the promoters of Scd1 and Scd2.
Promoter analysis of human stearoyl-CoA desaturase gene

Sequence of the 5' end of the human SCD gene. Numbering is relative to the translation initiation codon (+1). Nts —4152 to +27 are present in the GenBank Nucleotide Sequence Database under accession number AF320307; however, only residues —902 to +27 are shown. Amino acids are indicated using the single-letter code. Intron 1 is indicated in lower case. The start site of transcription in liver and hair-follicle tissue is boxed. A minor transcription site common to liver, brain and HaCaT keratinocytes is shown as a box bounded by a broken line. The FSE2 element is shown as a box bounded by a double line. Termination of Scd cDNA clones 34 (filled arrowhead) and 45 (open arrowhead) are indicated. The downward-pointing arrow indicates the 5'-most extent of RNase-protected fragment in brain and HaCaT keratinocytes. The core sequence of putative transcription-factor binding sites are single-underlined and notated above the sequence with rightward (>) or leftward (<) arrowheads indicating 5' to 3' direction of the element. Overlapping core sequences are double-underlined. SREBP binding site at fi517 is a novel SRE identified in mouse Scd1 and Scd2 promoters [33]. See Table 1 for core sequence of putative binding sites within the FSE2 element. Restriction-enzyme sites are indicated by tildes above the sequence. Bold letters and numbers indicate the 5' extent of the clones used for reporter assays.

Functional analysis of the SCD promoter

To determine which promoter sequences are sufficient to direct transcription, a 3.2 kb PstI – SacII fragment of the 5' flanking region was fused in both orientations to a promoterless luciferase reporter gene and examined for the ability to mediate basal transcription. In addition, progressive 3' deletions in the putative promoter region were also examined for their effect on reporter activity. The resulting constructs were transiently transfected into HaCaT keratinocytes, and cell extracts were assayed for luciferase activity at 48 h post-transfection. A total of seven different constructs were initially generated using convenient restriction-enzyme sites in the 3.2 kb flanking region (Figure 3). Significant transcriptional activity was contained with 5 constructs covering the region between fi3348 to fi882. Further deletion to fi461 resulted in complete loss of activity. SCD sequences fused in opposite transcriptional orientation did not induce any significant level of reporter activity. The highest promoter activity was obtained with construct c5, which corresponded to bp fi882 to fi150. These results demonstrated that, of the constructs tested, c5 contained the minimal required nucleotide sequence information for SCD transcription, and that essential cis-elements lay between fi882 and fi461.

To further delineate cis-elements responsible for human SCD gene expression, the c5 construct was deleted further to create three more constructs: c8, c9 and c10 (Figure 4; see also Figure 1). These constructs were transiently transfected into HaCaT keratinocytes, and cell extracts were assayed for luciferase activity at 48 h post-transfection. A total of seven different constructs were initially generated using convenient restriction-enzyme sites in the 3.2 kb flanking region (Figure 3). Significant transcriptional activity was contained with 5 constructs covering the region between —3348 to —882. Further deletion to —461 resulted in complete loss of activity. SCD sequences fused in opposite transcriptional orientation did not induce any significant level of reporter activity. The highest promoter activity was obtained with construct c5, which corresponded to bp —882 to —150. These results demonstrated that, of the constructs tested, c5 contained the minimal required nucleotide sequence information for SCD transcription, and that essential cis-elements lay between —882 and —461.
keratinocytes, in addition to the parental construct c5. Deletion of 129 nts from the 5'-end of the c5 construct (to generate c8) did not have any major effect on the promoter activity. Similarly, deletion of a further 144 nts, to generate c9, did not affect SCD promoter activity. However, deletion of a further 113 nts, to generate c10, completely eliminated all promoter activity as indicated by the reduced luciferase activity to background level (Figure 4). Taken together, these results narrowed the region that contains essential cis-elements required for SCD transcription in HaCaT keratinocytes to 609 to 496.

Examination of the sequences of this region revealed several potential transcription-factor binding sites (Figure 4, lower panel); AP-2 (−594), interferon-stimulated response element (−561), Sp1 (−551), X-box-binding regulatory factor-1 (RFX-1; −522, −495), SREBP (−517), NF-1 (−502) and NF-Y (−501). Considering the co-operative role of Sp1 and NF-Y in the regulation of the fatty acid synthase promoter [39], and NF-Y [33] in the regulation of mouse Scd genes, the effect of mutations in the Sp1 binding site and CCAAT box on SCD promoter activity was determined using site-directed mutagenesis. Using c9, the core binding site for Sp1 at −551, was changed from GGAG to AAAG (boxed in Figure 4) and the CCAAT box binding site for NF-Y at −501 was changed from CCAA to AAAA (boxed in Figure 4). The mutation in the CCAAT box also changes the core binding site for NF-1 (−501 on the non-coding strand) from TGGC to TAAC. It is unlikely that this NF-1 site is functional, as it contains the two most detrimental deviations (substitution at position 3 in the 5'-half-site of the palindrome and a shortened spacer of four nucleotides between half-sites), as indicated by a recent, more reliable prediction method for functional NF-1 sites [40]. Relative to the promoter activity of c9, mutation of the Sp1 site has no effect (Table 2). However, mutation of the CCAAT box completely removes SCD promoter activity (Table 2). These results, taken together with the requirement of the NF-Y-binding CCAAT element for mouse Scd1 transcription, strongly suggest that a critical cis-element for SCD promoter activity is the CCAAT box located at −501.

**EMSA for CCAAT-box-binding proteins**

To determine whether the CCAAT box critical for SCD promoter activity was actually occupied by DNA-binding proteins present in nuclei of HaCaT keratinocytes, we prepared nuclear extracts and performed EMSAs (Figure 5). A specific gel shift was obtained when nuclear extract from HaCaT keratinocytes was assayed with radiolabelled oligonucleotide corresponding to the CCAAT-box-binding site at −501. The gel shifts were competed with non-radiolabelled oligonucleotide containing the CCAAT element (see Figure 4), but not with the oligonucleotides containing either AP-2 or Sp1 binding sites. There are a few lower bands that were partially competed by CCAAT oligonucleotide, but not by the AP-2 and Sp1 oligonucleotides. Since NF-Y is the major protein that recognizes the CCAAT box [41] and binds as a trimer, it is possible that the lower, partially competed, bands represent partially assembled NF-Y complexes.

**Regulation of the SCD promoter by fatty acids**

Previous work identified a region of the mouse Scd1 and Scd2 promoter that was responsible for down-regulation of promoter activity.
Promoter analysis of human stearoyl-CoA desaturase gene

Figure 4 Fine deletion analysis of the human SCD promoter

Upper left: panel of deletion constructs. Restriction sites and 5’ ends of primer sites are indicated. Upper right: promoter activity of constructs transfected into HaCaT keratinocytes. Constructs were transfected in duplicate, cell lysates harvested 48 h post-transfection, and readings performed in duplicate. Values are plotted as the means ± S.E.M. for four samples from two independent experiments, and represent the ratio of the relative light units (RLU) of firefly luciferase to Renilla luciferase. P < 0.0001, one-way ANOVA; P < 0.0001, Tukey-Kramer HSD test for all comparisons of c5, c8, c9 versus c10, c6. Lower panel: promoter region showing location of c9, c10, c6 (bold), location of transcription-factor binding sites (as annotated in Figure 1), nucleotides (boxed) subjected to site-directed mutagenesis in Table 2, and competitor oligonucleotides (lower case) used for EMSA in Figure 5.

Table 2 Site-directed mutagenesis of the SP-1 site and CCAAT box

<table>
<thead>
<tr>
<th>Construct</th>
<th>RLU (mean ± S.E.M) (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGL3 + pRL</td>
<td>0.05 ± 0.01 (3)</td>
</tr>
<tr>
<td>c9</td>
<td>31.46 ± 0.34 (4)</td>
</tr>
<tr>
<td>c9 (SP1)</td>
<td>28.18 ± 1.62 (5)</td>
</tr>
<tr>
<td>c9 (CCAT)</td>
<td>0.47 ± 0.03 (5)</td>
</tr>
</tbody>
</table>

To examine whether the minimal SCD promoter could respond to UFA, c9 was transfected into HaCaT keratinocytes, and the cells subsequently exposed to various fatty acids conjugated to BSA (Table 3). BSA alone had no effect on SCD promoter activity. In contrast, stearate up-regulated promoter activity, whereas oleate (a MUFA) and linoleate (a PUFA) down-regulated promoter activity to the same extent. It is notable that stearate and oleate are the substrate and product, respectively, of the Scd enzyme.

Comparison of the promoters of human SCD, mouse Scd1 and mouse Scd2

In order to compare possible functional similarities and differences in SCD gene regulation between human and mouse, a multiple sequence alignment of the 5’ ends of human SCD, mouse Scd1, and mouse Scd2 genes was performed using sequences approx. 900 nts upstream of the ATG translation start.
sites. Sub-regions of the aligned sequences were selected for further analysis, the largest of which is shown in Figure 6. This region (Scd1, −626 to +1; Scd2, −500 to +1; SCD, −565 to +1) has 41.8% identity, indicating a homologous relationship between the three SCD genes (Figure 7). When pairwise comparisons are made over the same region, SCD is more similar to Scd1 (59.7% identity) as compared with Scd2 (55.9% identity) (Figure 7). A comparison of the proximal regions of SCD (−297 to +1), Scd1 (−301 to +1), and Scd2 (−235 to +1) genes demonstrate 36.4% identity, indicating an increase in overall variability between the three sequences in this region (Figure 7).

DISCUSSION

A sequence comparison of the 5′ end of human and mouse SCD genes indicates two regions of high homology, these being (1) the 5′-UTR and basal/proximal promoter (−297 to +1 in human) and (2) a region containing the PUFA-response element (−565 to −451 in human). Pairwise comparisons of these regions, particularly the proximal portion, indicate that human SCD is the homologue of mouse Scd1. The similarities include the 5′ UTR, the major transcription start site in human skin and liver relative to the proximal TATA box, and the presence and spacing of several putative transcription-factor binding sites. The NF-1 [43] and AP-2 [37] elements are implicated in transcription of several lipogenic genes, including Scd1 [46]. Deletion of the human SCD promoter, which leaves these elements intact but removes the upstream CCAAT box (−501), causes the loss of all transcriptional activity, indicating dispensability of the proximal promoter for basal transcription in HaCaT keratinocytes. This is consistent with a major transcription start site in HaCaT keratinocytes 5′ to these elements. However, these elements may have a more central role in human hair-follicle keratinocytes and liver, since the major transcription start site in these tissues is 3′ to these elements. Notably, Scd1 (expressed in liver) requires some of these sites for transcription, and the FSE2 element contains a conserved putative hepatocyte nuclear factor-3β binding site. Scd2 (expressed in brain) does not contain these proximal elements and initiates transcription from an upstream site, similar to transcription initiation of human SCD in brain. Thus the single SCD gene in human uses multiple transcription initiation sites that are tissue-specific and that reflect the transcription initiation patterns of the individual tissue-specific mouse Scd homologues.

The second region of high homology between human SCD and mouse Scd1/Scd2 is the region from −565 and −461 (in human) which contains the PUFA-response region, an element previously shown to mediate the down-regulation of mouse Scd...
Figure 6  Sequence alignment of the regulatory regions of human and mouse SCD genes

Multiple sequence alignment of 5' UTR and promoter regions of human SCD, mouse Scd1, and mouse Scd2 as described in the Experimental section. Numbering is relative to translational start site (+1). Grey boxes indicate sequence identity. Single-lined boxes with overlying filled arrowheads indicate start site of transcription for Scd1[28], Scd2[29], and in liver and hair follicle tissue for SCD. The double-lined box indicates the FSE2 element. Bold letters indicate the PUFA-response element [42]. Circled numbers indicate sub-regions used for comparisons in Figure 7. The core sequence of putative transcription-factor binding sites are annotated as in Figure 1. See Table 1 for core sequence of putative binding sites within the FSE2 element. The SREBP-binding site is the novel SRE element from mouse Scd1 and Scd2[33]. The AP-2 binding site in mouse Scd1 is defined in [37].

mRNA transcription in response to PUFAs [42]. The very high sequence identity is partly due to the presence and spacing of two CCAAT boxes and a novel binding site for SREBP. This SREBP-binding site represents a functional, non-canonical SRE first identified in the promoters of mouse Scd1 and Scd2 [33], and is located within the cis-acting PUFA-response region. The finding that SREBP activates Scd genes [33,47] and that SREBP mRNA and protein activation [49] are negatively regulated by PUFAs suggest that PUFAs down-regulate Scd transcription by interfering with SREBP function [33]. That the human SCD promoter activity is decreased by linoleate is consistent with this mechanism. The PUFA-response region also contains a CCAAT box, located 5 nts downstream of the SRE element. This CCAAT element is required for full activation of the mouse Scd1 and Scd2 promoters, and was shown to bind the NF-Y transcription factor [33]. The human SCD promoter is also dependent on this CCAAT site. Although the HaCaT cell proteins that bind to this CCAAT element were not characterized, based on the similarity to the mouse Scd promoters and the well-established SREBP-NF-Y interaction in various lipogenic genes [41], it is likely that NF-Y is involved in transcriptional activation of human SCD.

The MUFA oleate decreased human SCD transcription in HaCaT keratinocytes and decreased transcription of both Scd1 and Scd2 in CHO-K1 cells [50]. Oleate has been found to prevent the proteolytic activation of SREBP [51,52], and thereby decrease transcription from SRE-dependent promoters. In contrast with its effect in CHO-K1 cells, oleate does not down-regulate Scd1 in liver or adipose tissues [13,15], suggesting a cell-type- or tissue-specific effect of MUFA suppression for Scd1 [50]. The extent of these effects on MUFA regulation of human SCD remain to be tested. Stearate, a saturated-fatty-acid substrate of Scd, was found to increase transcription of the human SCD promoter, via the same minimal construct that is down-regulated by MUFA and PUFA. Conflicting data exist for the effect of stearate on SCD: no effect on Scd1 mRNA in mouse liver [13] or 3T3-L1 adipocytes [53], and increased Scd mRNA in pig adipose tissue [54] and enzyme activity in mouse liver [55]. It is possible that differential regulation of SCD transcription in response to stearate is dependent on the type of cell and/or tissue, as may be the case for oleate-mediated transcriptional suppression. In summary, we have isolated and characterized the human SCD promoter, and have demonstrated similarities as well as differences in sequence and activity as compared with mouse Scd.
promoters. The similarities most likely reflect the common function of SCD in regulating endogenous MUFA production, whereas the differences may be related to the different requirements for gene regulation in species with different numbers of functional SCD genes.

REFERENCES

24 Shanklin, J., Wilt, E. and Fox, B. G. (1994) Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase and are conserved in alkane hydroxylase and xylene monooxygenase. Biochemistry (Moscow) 33, 12787–12794


Received 28 February 2001/2 April 2001; accepted 10 May 2001

© 2001 Biochemical Society