Sequence Conservation and Structural Organization of the Glycerol-3-Phosphate Dehydrogenase Promoter in Mice and Humans

BABETTE GWYNN, KIMBERLY A. LYFORD, AND EDWARD H. BIRKENMEIER*

The Jackson Laboratory, 600 Main Street, Bar Harbor, Maine 04609

Received 13 April 1990/Accepted 19 July 1990

Cloned segments of the mouse glycerol-3-phosphate dehydrogenase (GPDH) gene, *Gdc-1*, were used to screen a human library. Human clones obtained spanned 25 kilobases of genomic DNA containing the human GPDH gene, *GPD1*. The 4 kb of sequence obtained from the 5'-flanking region and first exon of *GPD1* was compared with the corresponding mouse sequence. Both sequences share a *HindIII* site located in what has proven to be the highly conserved 3' untranslated region of an upstream gene of unknown function, *D15Kz1*. The 3.6-kilobase segment of mouse DNA located between *D15Kz1* and *Gdc-1* was provisionally termed the GPDH promoter. Alignment of the mouse promoter with the corresponding human sequence revealed two conserved domains. An upstream distal promoter region is approximately 900 base pairs in length. A downstream or proximal promoter region consists of approximately 300 base pairs immediately upstream of a TATA-like box and contains the fat-specific elements 1 and 2. Analysis of the chromatin structure of the *Gdc-1* promoter revealed four DNase I-hypersensitive sites. They were present in DNA of liver and brown fat, in which GPDH expression is high, but were absent in DNA of spleen, in which GPDH expression is low. Methylation studies of the promoter showed it to be heavily methylated in sperm. However, the DNA from each adult somatic tissue had a unique distribution of nonmethylated sites and could easily be identified by its methylation pattern. These data suggest a structural model of the promoter that explains how *Gdc-1* expression is differentially regulated in many types of cells.

The murine isozymes of glycerol-3-phosphate dehydrogenase (GPDH) are encoded by two structural genes, *Gdc-1* and *Gdc-2*, located on chromosomes 15 and 9, respectively (28). *Gdc-2* is expressed at low levels in certain fetal tissues such as brain (29). *Gdc-1* is almost exclusively expressed in adult mouse tissues, but its level of expression varies over 2 orders of magnitude, depending on cell type (33). The level of *Gdc-1* mRNA in each tissue is determined in part by a complex interaction of genetic, environmental, and hormonal factors. There are at least two regulatory loci unlinked to *Gdc-1* that control the GPDH mRNA levels in the cerebellum during weeks 2 and 3 of postnatal development (27). In addition, the development of high levels of GPDH expression in Bergmann glial cells of the adult mouse cerebellum appears to depend on a sustained interaction of the glial cells with the adjacent Purkinje cells (13, 14). Environmental factors also increase GPDH mRNA levels in cold-adapted adult brown adipose tissue which is innervated by sympathetic nerves (33, 36). In 3T3 preadipocyte cell lines, hormonal stimulation is required to initiate a differentiation program that appears similar to that observed during postnatal development of epididymal adipose tissue (7, 11, 12). These examples show that although GPDH is expressed in most adult tissues, its level of expression, developmental pattern, and genetic mechanisms of regulation are not identical in all cell types. It seems likely that glial and adipose cells as well as many other cell types may each have their own unique mechanisms for controlling expression of *Gdc-1*.

Because the expression of GPDH displays characteristics of both ubiquitously expressed housekeeping genes and highly expressed cell-type-specific genes, it was of particular interest to sequence the *Gdc-1* promoter and determine whether it had structural characteristics in common with the

promoters of either class of genes. A limitation of such analyses is that the actual size of the promoter and the location of distant enhancer sequences is difficult to determine with sequence data alone. To better define sequences critical to gene expression, we cloned the related human gene, *GPD1*, assigned to human chromosome 12 (26). We compared the mouse and human promoters to locate sequences conserved in evolution and therefore likely to represent important regulatory domains. In this report, we identify four conserved regions of 5'-flanking sequences, locate the major DNase I-hypersensitive sites within or adjacent to evolutionarily conserved regions of the mouse and human promoters, and determine the methylation status of the *Gdc-1* promoter in several mouse tissues.

In the Discussion, these results are used to generate a model of the GPDH promoter in both mice and humans. This model predicts how the structure of the *Gdc-1* promoter may explain many of the interesting characteristics of *Gdc-1* expression observed in specific cell types both during development and in the adult animal.

MATERIALS AND METHODS

Cloning of *GPD1*. Two probes from the mouse *Gdc-1* locus were used to screen a genomic library. One probe was the partial cDNA clone F12, and the other probe was a 0.8-kilobase (kb) genomic *HindIII* fragment containing the entire second exon (20). The human library was provided by Thomas Maniatis and was constructed by using *AluI* and *HaeIII* partial digests of human DNA, *EcoRI* linkers, and the Charon 4A vector. Plaque hybridizations of bacteriophage were performed essentially as described by Benton and Davis (3).

DNA sequence determination and comparisons. Restriction fragments of cloned human DNA were subcloned in M13mp18 and M13mp19 phage and sequenced by the dideoxy-chain termination method (34). Comparisons of the

* Corresponding author.

mouse and human sequences were made by using an IBM personal computer AT and the MicroGenie alignment program (Beckman Instruments, Inc.) with the following parameters: minmatch 7, minper 75, loopen 9, and maxdist 500. Additional analyses were performed by using a MicroVax II (Digital Equipment Corp.) and the T-Matrix program as previously described (38, 40).

Isolation of nuclei. Nuclei were freshly prepared from adult BALB/cByJ mice. Tissues were homogenized in a Dounce glass homogenizer with 10 ml of ice-cold buffer A (60 mM KCl, 15 mM NaCl, 0.15 mM spermine, 0.5 mM spermidine, 15 mM 2-mercaptoethanol, 15 mM Tris hydrochloride [pH 7.4], 0.25 M sucrose, 0.2 mM phenylmethylsulfonyl fluoride [PMSF]) per g of tissue. PMSF was stored in ethanol at 0.1 M and added to buffer A immediately before use. Tissues were homogenized with 12 strokes of pestle B and then 12 strokes of pestle A. The homogenate was layered over a 10-ml cushion of buffer A with 0.75 M sucrose in a 30-ml Corex centrifuge tube. Nuclei were then pelleted in a Sorvall centrifuge, using an HB-4 swinging-bucket rotor spun at 10,000 rpm for 10 min at 4°C. Nuclei were resuspended in 0.1 to 1.0 times the original volume of buffer A and stored on ice.

DNase I hypersensitivity assay. Bovine pancreatic DNase I (Miles Laboratories) was prepared in buffer A (without PMSF and with 30 mM MgCl₂) at a concentration of 1 mg/ml. Twelve 1:1 serial dilutions were made by using the same buffer and stored on ice. For digestion, 50 µl of each DNase I dilution was added to 450 µl of resuspended nuclei and incubated at room temperature for 15 min. After incubation, the solution was transferred to tubes containing 125 µl of a solution with 0.4 M EDTA (pH 7.4), 4% sodium dodecyl sulfate and 4 mg of proteinase K (Bethesda Research Laboratories, Inc.) per ml. The mixture was vortexed vigorously and incubated at 55°C for 2 to 3 h. DNA was purified by one phenol extraction and one chloroform extraction, followed by precipitation in 70% ethanol. DNA was resuspended in 10 mM Tris hydrochloride (pH 7.4)–1 mM EDTA. Then 5 µg of each DNA was digested with *Hind*III, electrophoresed in a 0.8% agarose gel, and blotted onto Zetabind nylon filters (CUNO, Inc.) by the procedure of Southern (37). Probes were made from the cloned DNAs by the random hexamer labeling method and hybridized to the filters as previously described (18).

Methylation analysis. Nuclei were isolated from tissues of adult male BALB/cByJ mice, and the DNA was purified by the procedures outlined above except that buffer A did not contain PMSF. Then 5 µg of each DNA was digested with *Hind*III, followed by a second digestion with either a methylation-sensitive or a methylation-insensitive restriction enzyme. The DNA was then electrophoresed in agarose gels, Southern blotted, and hybridized as described above.

Nucleotide sequence accession number. The nucleotide sequence data reported in this paper have been given the GenBank accession number M36917.

RESULTS

Cloning of *GPD1*. We screened 7.8×10^5 recombinant phage from a human genomic library, using two probes from the mouse *Gdc-1* locus (20). One probe was a 0.8-kb mouse genomic *Hind*III fragment containing the entire second exon of *Gdc-1* together with flanking intron sequences. The second probe, F12, was a partial cDNA clone of 0.9 kb from the untranslated 3' end of the mouse transcript. Both of these probes detected single bands on Southern blots of human DNA digested with either *Hind*III or *Eco*RI (data not

FIG. 1. Restriction map of human GPDH clones 1 and 3 showing area of overlap. Restriction sites are abbreviated as follows: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; K, *Kpn*I. One of the *Taq*I sites abbreviated T and labeled with an asterisk is polymorphic in human DNA samples and is present in the clone sequenced. Symbols: —→, location of the *GPD1* coding region and the direction of transcription on the composite map of the *GPD1* locus; ▨, region for which sequence was determined. Fragments A to E are the five *Hind*III fragments subcloned into pBR322.

shown). Four positive human clones were obtained from the library. Clones 1 and 2 hybridized exclusively to the exon 2 probe, and clones 3 and 4 hybridized exclusively to the cDNA probe. Restriction enzyme analysis indicated that each pair of clones was similar (data not shown). Therefore, only clones 1 and 3 were further characterized.

The overlapping restriction maps of clones 1 and 3 were used to generate a composite map covering approximately 25 kb of the human genome (Fig. 1). The location of the 5' end of the approximately 7.5-kb *GPD1* gene was determined by DNA sequence analysis (see below), and the 3' end was localized by hybridizing the mouse cDNA clone to restriction fragments of clone 3. The entire locus was subcloned into pBR322 by using the five *Hind*III fragments labeled A through E.

Fragments A, B, D, and E hybridized strongly to labeled human genomic DNA, indicating that they contained highly repetitive sequences (data not shown). Fragments A, D, and E hybridized specifically to a cloned *Alu* repetitive sequence, BLUR 8. When fragment C was used as a hybridization probe for Southern blots of human genomic DNA digested with *Hind*III or *Eco*RI, a single band predicted from the restriction map was detected. However, *Taq*I digestion produced a restriction fragment length polymorphism (data not shown). Fragments 2.6 and 6.2 kb in length were present in all of the human DNAs examined. However, in 2 of the 12 DNA samples, the 6.2-kb fragment was present in half copy number and an additional 8.8-kb fragment of similar intensity was found. Restriction enzyme analysis of the *GPD1* clones revealed that the *Taq*I restriction fragment length polymorphism results from the presence or absence of a single *Taq*I site that is located in fragment B of the cloned DNA (Fig. 1).

Sequence of the *GPD1* promoter region. We determined the sequence of the 5'-flanking region and the first exon of the *GPD1* gene. This 4,004-base-pair (bp) region was then compared with the corresponding 4,208-bp region of the *Gdc-1* gene (20). The sequence comparison begins 5' to the GPDH structural gene at a conserved *Hind*III site located at the 5' end of human fragment B (Fig. 2). A region of 78% sequence homology between the two species extends downstream for 558 bp in human DNA and 570 bp in mouse DNA. It ends

FIG. 2. Sequence comparison of the upstream region of the human and mouse GPDH promoters. The human sequence is printed above the mouse sequence, and identity between the two species is indicated by a vertical line. The comparison begins at a conserved *Hind*III site located 3,964 bp upstream from the translation initiation codon in the human gene and 4,166 bp upstream from the same codon in the mouse gene. Note that this *Hind*III site is located in the highly conserved 3' untranslated region of the mouse gene *D15Kz1*. The polyadenylation signal of this gene is underlined and labeled Poly(A), and the end of the gene is indicated by an asterisk. The start and end of the GPDH DPR are indicated by arrows. Two mouse B1 repetitive elements of 164 and 110 bp, respectively, in length have been deleted in the figure. The location of these elements is indicated (Δ). The sequence comparison ends just before the start of the conserved PPR.

just 3' to the polyadenylation consensus sequence AAT AAA. In the mouse genome, this region represents part of the 3' untranslated region of another gene named *D15Kz1* (21). *D15Kz1* appears to encode a protein of unknown function which has no obvious relationship to GPDH. It is not known whether this gene is also expressed in human tissues. For the purpose of this report, we will assume that

the 3' end of this upstream gene delineates the beginning of the GPDH promoter region.

There is little (45%) sequence similarity between the 1,434 bp of human DNA and the 1,433 bp of mouse DNA in the initial segment of the GPDH promoter. Then, a region of relatively strong (71%) sequence homology begins, extending for 908 bp in human DNA and for 888 bp in mouse DNA.

```

-3044  CACTTGTGGCTGGAGTGAGCCATAGATGCCCGTGATCCCTTAGGTTTGACGAACCTGAG
      |||||  ||      |||||  ||      ||      ||      ||      ||      ||
-3249  GCTTGTGACCTCAGCAATGCCCT  TG  GGATAGTACCTACCCACCTCCATCATGAA
      |||||  ||      |||||  ||      ||      ||      ||      ||      ||
-2984  CCGCCTTTGGAGGAGTTTCACTTCTTCCCCAGGGCTTCTGTTTAGATGTGGGCAGGATCT
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-3194  CAGACAAGGCAGTGGCACTGCTGGTGTCTCCCTCGTCTTGAATCTACCAGGTACAAC
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2924  GGACGACCAAAGACATTTTGTCTTTTGTGCTGTACTCTTTTTTTTTTTTTTTTTTA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-3134  TGCTGGCTGGA  ACAAGCTGTCCTTTGGCATCTG  TGGGC  CCCATTGAGCATTGATGA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2864  GACGGAGTCTCACTGTTGCCAGGCTGGAGTGCAGTAGTGC  GAT  CTTGGCTCACTGCA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-3078  AACTG  GGCTCTGTTCTCCCTAGGGTGCATGCCATCCAGCTGTTGCCAAGGCCCGGTG
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2806  AGCTCTGCCTCCTGGGTTACGCCATTCTCTGCCTCAGCCTCCCAGTAGCTGGGACTA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-3019  ACAGCAGGTTTCTCTGTT  TC  CCATACAAACCCATCAACCTCATGCTTTTCAGGTTCCA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2746  CAGGCGCCTGCCACCATGCCCGCAATTTTGTATTTTGTAGTAGAGCAGGTAACCG
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2961  CGGACAGC  CCCAAGATAGCAGG  GTTCTCTGTGTTGATTCCTGTGTATGT  AACCA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2686  TGTTAGCCAGGATGGTCTCGAT  CTCCTGACCTCGTGATCCGCCCCCTCGGTCTCCCA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2905  AGGCTGACCGTAACTTACTATGCAACTG  CCTCCTTCTCCCTAGTGAGATTACAGAT
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2628  AAGTGC  TGGGA  TTACAGGTGTGAACCACCGTGCCAGCCTGTGCCTGCTACTCTAAG
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2846  GTGTGCCTTTTCAGCGCCGATATAGAAGTCTGTTCTTTTGTGTTTTTTTCAAG
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2571  GTTCCGCTT  TCATGGTTCAAATAAGAGTTCAGTTTTTGTGCTGTTTGTAGATGGAG
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2786  ACAGGGTTTCTCTGTGTAGCCCTAGCTGTCCTGGAACCTCACTTTGTAGACCAGGCTGGCC
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2512  TTGTGTTCTTGTTCCTCCAGGCTGGAGTGCAATGGCGCGTTCTCGGCTTACTGCAACCTCC
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2716  TCGAACTC  AGAAATCC  GCCTGCCTCTGCCTCCGAGTGCTGGGATTAAAG  GCGTGC
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2452  GCCTCCAGGTTCAAGTGATTTTCCTG  CCTCAGCCTCCTGAGTAGCTGG  TTCTTAAGA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2670  GCC  ACCATGCCAGCTAAGAAGTCTGTTCTTAAAC  CCAGTGTGATGGCCGCTTAGT
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2394  GCACCATGGAAGTTCT  GTGTTGTATCCTCATACTGTAGACAAC  TAGCCCTCCTT  G
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2612  CCTGTGTTGAGTAGATAGGGTGCCTAGCTGGCTGCTCAGACATCAGAGATGCTCCTTGG
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2337  GCAGAGTTTGGAA  GACAGC  TCAGGAGGGATTCCCATAGCAATCACTTGAAACTCGCC
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2552  GAAAAGCAGGCAATGCCTGCAGTTCTACCTGCTTCCACCAGGGGCTCACAACCTCCTAC
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2279  TGCAGTTCTGCCTGCTTCCCACTTGGAGGCT  CATGACTTTTTTCTTTGTCTATTATGCA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2492  TGCACTCTGCCCAAGAAAGTGGAACATTGGATGAGCCACTGCCACATCCTGTAGTCA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2220  AGAATGGGGTTTGAATATAGGATCAGAG  TACAACCACTTTCCAAAGTTGTAGTTAAA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2432  GCTGAAAAATCTTACTCATCAGTACATTGATACATGCCTTAACTCTCCACGCTAGGGAGG
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2161  AATTATCCTTTGCCTGTTTG  GATGGGAACATTCTAGTATTGCCTGTTCAATATATGTA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||
-2372  CAGAGCGGGAGGGTCTCTGAGTTCCAGGCCAGGATTACAAAGGCACTATTTTACA
      |||||  ||      |||||  ||      |||||  ||      ||      ||      ||

```

FIG. 2—Continued

We defined this as the distal promoter region (DPR). Four segments of the human DPR had 85 to 90% identity with the corresponding mouse sequence. A subcloned *Hae*III fragment containing the initial 511 bp of the human DPR with three of the four aforementioned segments was hybridized to Southern blots of human, mouse, and rat DNAs digested with either *Hind*III, *Eco*RI, or *Bam*HI. Each lane of human DNA had a single band of the size expected from the restriction map of the cloned gene (data not shown). Although the blot was washed at approximately 90% stringency, the mouse and rat DNAs also contained a single easily discernible band. These Southern blot data suggest that this region of the DPR is single copy in the human genome and hybridizes to a closely related single-copy sequence in the mouse and rat genomes. Although this raises

the possibility that the DPR represents an independent gene upstream from GPDH, sequence analysis revealed no CCAAT or TATA boxes, no polyadenylation signals, and no open reading frames in the DPR.

The 3' end of the DPR is defined by a sudden drop to 52% similarity between the human and mouse sequences. No sequence conservation is apparent for the following 790 and 1,032 bp, respectively. Another region of higher sequence similarity (70%), termed the proximal promoter region (PPR), comprises the next 224 bp of mouse DNA and extends to the *Gdc-1* transcription initiation site. The human and mouse sequence alignment for the PPR and first exon is shown in Fig. 3. A mouse TATA-like box sequence, TTAAAT, located 28 bp upstream from the transcription start site is also present in the human promoter. Remark-

```

-2102   TCAAACCCACCAAAAAATCCAACCACTGTTCTACTCCGGAAGT   TGGG   GT   TTATC
        |||   |   ||||   |   ||||   |   ||   |   ||   |   ||   |
-2312   AAAAAGTGA   ATAAAAATTTTACCCACTG   AGCAGCTGGGGGTCTGCTGAGTAGTGTGTGCT
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-2047   AGACT   TGAGAA   GGGAAAGCAGAGAGG   AGTG   GTTG   GGCCGATAACTAGGAGT
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-2254   TGTCTGCTGCCAATCTGAGAACATCCTGGCTTCTCTGAGTTGAGAGGGTTAAAAAAACT

                                -----> Start DPR
-1996   GGAGAAAATTGATGTCAA           GGACATTGAAATGGCATCTGTTTTTACCCAGGCTA
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-2194   GTAATCACTGGGTCTAGAGAGAGCGGTGACTGAGCTGGCATCTGTTTTTACCCAGGCTA
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1943   ATTACATGTGAATGAATGGCTGATGAACCATTTAATTCAAGC   GAAATCTACAGTT
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-2134   ATTGCACGTGAATGAGTGGCCGATGA   ATCATCAATTCCAAGCAGATACGCACAGCC
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1884   TGCCTCTGGGAAAAGAGCTCTAGTGATACAGATAACTTCTGTGATGAGCCGATTCTGAA
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-2077   TG   CTCTTG   AGAG   GCTTTGGTGGCAGACAACTTCCTGT   TAA   CGA

-1824   TCTCTGCACCTCTTTATTCCCATGTAAATGAGCACATTCTTAG   TCTCCACACTATCT
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-2031   GCATTCTCTTTATTCTCGTGAATGGCCACGTGCTTTGTCTCTCCACACAGCCT

-1766   AAAATGATTGTAGGGGCAAAATGAGAGAGGAAGCATTAGGCAAAAACAAGGGGTTAT
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1976   AATAGGA   TGTAGGGGCAAA   GAGAGGCGGGAGTATTTAAGCAAAAACAA   GGATCCGT
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1706   TTTATGCATAAACATCCATTCTTTGGGGCTCCCATCTTTGTGCCAGGGACAGAAGCATC
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1920   TTCAGGCATAAATGCCATTCTTTAGGGTTCTTATCATATGTGCCAGGGACAGGAACATC
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1646   CTTTCTCTCCATCACCCGTAATCAGTGCTCCAGCTGGCAGCCTGGCACTGGAGCCTTC
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1860   CTTTCTCTTCCACCCCTCGTAATCAGCGCTCCAGCTGGCGGCTGGCACCAGGACCTTC
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1586   ATCTCAGTGCCAGTGCAATAAAGGAGTGGTCTGTGGGAGGCTGTTTCTGTATTAGGTATA
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1800   ATCTCTGTGCCAGTGTGATAAAGGAGAGGCTGTGGGAGGCTAGCTCTGCATCAGGCAGA
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1526   ACATTGTGCTAAGAAAAACAGCCAACACTACTGATTTTCATGTCCGTTTCCCTATGGGAAATT
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1740   A   TCTGATAGAAGAAAAACAACTAGCTATTGA   TTGATTTCAAGTATTCTCCCATCATT
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1466   ACTGGC   CATGGAAG   GAGAATATGGTGGTTTTTCTGGGAGTCCCCGGAACCTTAG
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1682   AGAGACTGCCTGGCCGCGAGACCATGATAGGCTGCCT   CTGTGTAATCCCC   GAGCTGAG
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1409   AAAGCCATTCTCCCCAGTCTTGACCTCCTGCCTGCTTACTCCTACCCATGGGTTGCA
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1624   AAAAAGACATGCTTCCAATCCTGACCTCCTACCCACTT   CTCCTCGCCGCTGGTGGCATT
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1351   CACAGACTCCCCGTGTAGCCCAAGTGCTTGTACCAGTTGCTGGCCAGCAACAGCCAGGG
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1565   GGTGACACTCCCCCTTAACCTAGTGCTGATTA   CTG   TCCTGGCCAG   GGCAGCCTGGG
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1292   GGATGG           CAACACCCACCTAGCCGTCTGCCAAGCCAGGAGGTGAGCATGAGCAC
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1508   GGGTGGGTGGCCAACATGC   CATTATGGTGTGCCAAGCCAGGAGGTCTG   ACATGAGCAT
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1237   ACAGTGGTGTCTGCCCTCTAGG   CCAAGGTGCA   CAGTGTTAGG   AG   CCTCTGTCTCC
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
-1450   ACAGTGGCGCTGCCCTGTGGGTCCAAGGTGCAGCACTGGAGGGCAAGAAGCCCTGGTGC
        |||   |||   |||   |||   |||   |||   |||   |||   |||   |||

```

FIG. 2—Continued

ably, 33 of the next 34 bp upstream from this A+T-rich region are identical in the two species. A potential mouse Sp1-binding site, GGGGCGGAGC, located 60 bp upstream from the transcription start site has two mismatches with the corresponding human sequence AGGGCGGGC. Neither promoter has a CCAAT box, although there is strong sequence homology continuing upstream of the mouse Sp1 site.

In addition to the sequences described above, the human and mouse PPRs contain sets of homologies known as fat-specific elements (FSE) 1 and 2. The FSEs were originally found by comparing the 5'-flanking sequences of the three mouse genes *Gdc-1*, *Adn*, and *Ap2* (19, 32). These genes have similar patterns of differentiation-dependent

expression in adipocytes (8, 9, 11), and at least some of the FSEs appear to be involved in regulating this expression (10, 15, 43). The FSE1 consensus sequence is GGC(A/T)CTGG TCA(G/T)G; the two previously identified FSE1 in the *Gdc-1* promoter (19) are shown in Fig. 3. Although it differs at three positions from the consensus sequence, the upstream FSE1 (AGCACTG-TCAGC) is identical in mouse and human cells. The other mouse FSE1 (GGCTGAGGCCAGG), which is inverted relative to the gene, also deviates at three positions from the consensus sequence. However, the analogous human FSE1 in GPD1 (GGCAGAGACTAGG) differs from the mouse sequence at three positions.

Single FSE2 were originally identified in the mouse *Gdc-1* (ACTCGAGAGGAAAA) and *Ap2* (ACTC-AGAGGAAAA)

```

-1182  ACTCTCCAGATCAGCCTGGAACAGAGCTTCCCTCGTCTTAGACAAACAGATGGCCGTG
      || ||||| | ||||| ||| || ||||| ||||| ||||| |||||
-1390  GCTTTCAGAGTGACTTGGAAACCCGAG TT CCTCTCTTAGTCAAACAGATGTCTGTG

                                     End DPR <-----
-1122  GCGGATGCAGAGCCCTGAACTGCTGAGGGTGTCTTCCATTCTCGCCCTTCCCCAAA A
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
-1332  GGTGCCATGGAGCCC AAAGTGTGGATATGGGCTGCCTTCTCCCCCTCCCCAAAGC

-1063  AACATACAC CAACTTACCCCCAATCAC AGCATCTGAGAC CACAAAGCCCCACGG
      ||| ||| || ||||| ||||| ||||| ||||| ||||| |||||
-1273  TACACACCCACACCGGCATCAGAGACCACAGAGCTGCGGAGACATGGCTACCCCTGAGTT

-1008  ACACA GCTGCCCCGGGTACTGCCTGCCTTCCCCTTTTCTGGCCTCTGCAAAGTAG CCA
      || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
-1213  TTAGAGGCTGCTGTGTAAGTGGCTTCTCTGGCCCTGTCTAGAGACTGTAACAGGACTC

-950   TTCACCTGATGAAGCCCTGG GGCAGT ACCTTGGAAACCCCTGAGAGGAGGAGTTGAA
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
-1153  TACAGCTGTTTAATTTAAGGTCCTGTGGGATAAGCAGAGCCTGAGAGAAGAAATTGAA

-892   GGAGGCTCAAAGCCTTTCAAGGACAGTGTGCACTCTGCCTGCCTCTATTCTCTCTGTC
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
-1093  GGGAGATC AGGCTTATCCAGGAGCATGTG CTGGCCACCACC TTCTGCCCT TC

-832   CTCTTAAGCCAGAGTCTACAACAAAATTGTGTACCGCAATTCTCAGAGCTCTCTAATT
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
-1041  CTGAAGC CTGAAGGAGTGTGGGTGGGCCACGTGTACCAGTGGC CTT

-772   TAACCTTTCCACCTTCCCTAGAACACATTTTGGGTGGGGATCATCTGAGAGGCCCATG
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
-993   TGACCTT CCTTCCCT TTAGAGACCTTCAAACGGAGATTATCTGAGGTGAAGGGG

-712   GATGAGTCCCCAAGATGACCTCTGTGGGTGACCAAGACAATCTTGTCTTTTGGAGCACCT
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
-939   CAACCTGTGGGTCCCTGAC CTGCAAGTAACCAAGATGACCTTTTCT C CCG

                                     164 bp B1

-652   GGAGTCTCTCATCTTTTGTAGCT GGGCTGT CCTCAAAGCAGGCCCTTACCTCTTAT
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
-726   AGATATTTAATCCTTGGTCGGGTGGGGTATACACCAAGCAGCCCTTACTGTGTAT

-594   GTGGCCT CATCTTTAGAGGTGGGAAGTGGGCCACATCCCCACCCAGCCCTGCAGT
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
-666   GTGGTATTGTAGTATTATCCCTAGAAG GTGCTCACATCCCTTGATCCTCTACCCAG

-536   GGAGCTAGCGTGGGCCTGCTGGAGGACAGAACTGTGGAACAGGGTCTGTGAGGATATGT
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
-608   AGCCCGTGATGGAGCTACCATCATGGGACATC ATGGGAGTGAACAGCTGAGGATCTGG

-476   GGCAGGGTTGGCGGGGGATAGGGGGTCATGGAGGAGCCTACCTCTGAAAAGCATTCCCC
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
-549   GGGAGGCACATGCTAGCTAGCGAGAGCGAT GAGGCGTTTGCCCTTGAAAA CATCGCTG

-416   TAGCTGCTACTGAA
      ||||| ||
-491   TTGCTGACGCTCCCT
      |||||
                                     110 bp B1

```

FIG. 2—Continued

promoters (19). The corresponding human FSE2 (GCCAG GAGAAAA; Fig. 3) has only 64% (9 of 14) homology to the *Gdc-1* sequence. However, another human FSE2-like sequence (GCCAAGGAAGAAGC) is present upstream at the beginning of the PPR. A similarly located mouse sequence (GCCAaacgAGGGAAGAAGC) is identical except for a 4-bp insert.

Downstream from the PPR is the first exon, which encodes the initial 14 amino acids at the N-terminal end of mouse and human GPDH (Fig. 3). As a result of a single base change in the third codon, the human enzyme has a serine residue instead of glycine. This conservative amino acid change is the only difference present in the portion of the enzyme encoded by the first exon. This amount of sequence conservation is not surprising, since the N-terminal region is part of the NAD-binding domain (20).

In summary, we have tentatively defined the human and

mouse GPDH promoter region as beginning at the 3' end of another highly conserved upstream gene and ending at the start of the first exon in the GPDH structural gene. Computer-generated DNA sequence alignments between mouse and human genes suggested that this promoter region contains two domains (DPR and PPR) that were highly conserved in evolution. Additional studies comparing the human and mouse promoter regions by using the T-MATRIX program (38) confirmed that the DPR and PPR are the major GPDH promoter elements present in both mouse and human cells (data not shown).

DNase I-hypersensitive sites. The 5'-flanking regions of many genes have tissue-specific chromatin structures related to binding of regulatory proteins (2, 4, 23, 24, 30). These structures are detected by the presence of regions that are highly sensitive to digestion by DNase I (42). We examined the DNase I sensitivity of the *Gdc-1* promoter in brown fat,

DNA isolated from nuclei treated with increasing concentrations of DNase I was digested with *Hind*III to generate a 4.6-kb promoter fragment. This fragment extends from the conserved *Hind*III site in the 3' untranslated region of the upstream gene (−4166) to the next *Hind*III site in the first intron (+480) of *Gdc-1* (20). Southern blots of the DNA were then hybridized to a 0.6-kb *Pst*I fragment (4.8/6 probe) or a 0.6-kb *Pst*I-*Hind*III fragment (4.8/5 probe) located at the 5' or 3' end, respectively, of the mouse 4.6-kb *Hind*III fragment (21). Figure 4 shows Southern blot analyses of brown fat DNA in which four major DNase I-hypersensitive sites were detected. The furthest upstream site is located near position −3174 relative to the *Gdc-1* initiation codon. This is approximately 422 bp downstream from the 3' end of the *D15Kz1* gene. No obvious sequence similarity was detected between mouse and human genes in this region of the GPDH promoter (see above). The next hypersensitive site was located in the region surrounding position −2534. The original computer-generated sequence alignments of the entire promoter region had failed to detect significant sequence conservation within this promoter region upstream of the DPR. However, alignments of only this region with the human sequence located a common *Pst*I site flanked by sequences with strong (86%) homology. The conserved sequence, GCCTGCAGTTCTACCTGCTT-CCACCAGGGG GCTCA, extends from −2538 to −2504 in the mouse promoter, and the corresponding human sequence, GCCTGCA

A third major hypersensitive site was present within the DPR between -1541 and -1466. The fourth and last hypersensitive site was more difficult to precisely locate because it was near the 3' end of the 4.6-kb *Hind*III fragment. It appears to be within a region extending from +5 to +241 which includes part of the first exon and the first half of the first intron. Interestingly, sequence analysis of the first human intron revealed the presence of a sequence with 86% (32 of 37) homology to the mouse sequence surrounding position +241. The mouse intron sequence from +232 to +268 is GCCCGCTGGGCGCTGCACCTGTTCAACCA GGAAGG, and the corresponding human sequence from +213 to +249 is GCCCTCCTGGCAGCTGCACCTGTTCA CCATGGGAAGG. Taken together, these results suggest that there may be sequences within the transcribed gene that are important in regulating the expression of GPDH.

Additional studies using nuclei from liver identified DNase I-hypersensitive sites that were very similar to those found in brown fat except that the furthest upstream site was not observed (data not shown). In contrast, nuclei from spleen appeared to have no discernible hypersensitive sites at all. Since spleen has a relatively low level of GPDH activity, these data are consistent with the possibility that there is a correlation between the presence of hypersensitive sites and the level of *Gdc-1* expression in certain tissues.

FIG. 4. Location of DNase I-hypersensitive sites in the *Gdc-1* promoter. Nuclei isolated from brown fat were digested with increasing amounts of DNase I. The purified DNA was then digested with *Hind*III to generate the intact 4.6-kb promoter fragment. (A) Inverted Southern blot of the DNA hybridized to clone 4.8/6, located at the 5' end of the 4.6-kb promoter fragment; (B) similar blot (in opposite orientation) that was hybridized to the 4.8/5 clone from the 3' end of the same promoter fragment. The linear map between the two panels represents the 4.6-kb *Hind*III promoter fragment, and arrows indicate the approximate locations of the hypersensitive sites in the *Gdc-1* promoter. Pr, Location of the subcloned fragments used as probes; G, the upstream gene *D15Kz1*; □, DPR2; □, DPR and PPR; ■, the first exon of GPDH. Horizontal arrows below the blots indicate increasing digestion by DNase I. Vertical arrows in the columns with the size standards indicate the direction of DNA migration during electrophoresis.

Extent and pattern of methylation of the *Gdc-1* promoter.

Since hypomethylation of specific DNA sequences may correlate with both the presence of DNase I-hypersensitive sites and the level of gene expression in certain tissues (1, 4, 6, 16, 17, 25, 35, 39, 41), we examined the methylation status of the *Gdc-1* promoter by using methylation-sensitive restriction enzymes whose recognition sequences were clustered within this G+C-rich region. The experimental approach was similar to that described above. DNA was isolated from adult mouse tissues, digested with *Hind*III to generate the 4.6-kb promoter fragment, and then digested with a second enzyme inhibited by methylation of cytosine at CpG dinucleotides. Southern blots were then prepared and hybridized to the 4.8/6 and 4.8/5 probes obtained from the 5' and 3' ends, respectively, of the promoter fragment.

The first methylation-sensitive enzyme examined was *Xho*I. The single *Xho*I site present in the promoter is located at position -209 within one of the FSE2 elements. Southern blots of DNA from brain, heart, kidney, liver, and spleen indicated that this *Xho*I site was unmethylated in the majority of cells from all tissues examined (data not shown). Thus, there was no obvious correlation of the various levels of gene expression present in these tissues with the degree of methylation at the *Xho*I site.

The next enzymes used were *Hpa*II and its methylation-insensitive isoschizomer *Msp*I. Interpretation of these results (Fig. 5) is complicated by the fact that there are eight

*Hpa*II-*Msp*I restriction sites within the promoter region. Figure 5A shows the results of digesting DNA from various tissues with *Hind*III and *Msp*I and hybridizing the Southern blot to the 4.8/5 probe. The two smallest visible fragments comigrated and were generated by the *Msp*I sites at positions +224 and -28. The additional large fragment detected by this probe extended from -1261 to -28. Figure 5B shows the results obtained with *Hpa*II. Of the tissues tested, brain, kidney, liver, lung, spleen, and epididymal fat looked remarkably similar. The bands that were present in the *Msp*I digests were also detected in the *Hpa*II digests, indicating lack of methylation in many cells at the upstream site (-1261) and either (or both) of the first two sites (+224 and -28). However, the presence of additional *Hpa*II-specific bands indicated that these sites were methylated in a significant number of cells in all tissues tested. Only the DNA from sperm and testis looked notably different. Figure 5C shows an autoradiogram exposed for a shorter length of time to demonstrate the relatively intense 4.6-kb *Hind*III fragment present in the DNA from sperm and testis. These results indicated that all eight of the *Hpa*II sites were methylated in most if not all of the male germ cells (the sperm cell preparations were believed to contain some epididymal cell contamination). Thus, although the somatic tissues other than testis had very similar patterns of hypomethylation at specific *Hpa*II sites, sperm DNA was heavily methylated at all of the *Hpa*II sites in the *Gdc-1* promoter.

FIG. 5. Tissue-specific methylation of *HpaII* sites in the *Gdc-1* promoter. Panels A to C are autoradiograms of Southern blots of DNA isolated from different adult mouse tissues. Each DNA was digested with *HindIII* and then further digested with either *MspI* or *HpaII* as indicated. Tissue DNAs are abbreviated as follows: B, brown fat; K, kidney; Li, liver; Lu, lung; Sp, sperm; S, spleen; T, testis; and E, epididymal fat. Panel C is a shorter exposure of the blot in panel B to demonstrate the relative intensity of the high-molecular-weight bands. The linear map drawn between panels B and C represents the 4.6-kb *HindIII* fragment containing the *Gdc-1* promoter (see legend to Fig. 4 for further explanation). The *MspI*-*HpaII* sites are indicated (+) and are approximately aligned with the corresponding bands in the autoradiograms. The location of the 4.8/5 clone used as a hybridization probe is also shown (Pr).

The five *HaeII* (PuGCGCPy) and nine *HhaI* (GCGC) methylation-sensitive sites in the 4.6-kb *HindIII* promoter fragment were also examined for the presence of methylated cytosine. Note that all *HaeII* sites contain a *HhaI* site. We examined DNA from spleen, kidney, liver, and epididymal fat pad in order to study tissues representing a wide range of metabolic functions and GPDH expression levels. The methylation patterns generated by Southern blots of *HhaI* and *HaeII* digests hybridized to the 4.8/6 and 4.8/5 probes are shown in Fig. 6A and B, respectively. The patterns seen with the 4.8/6 probe were relatively simple. Spleen, which has the lowest level of GPDH activity, showed a high degree of methylation at all *HhaI* and *HaeII* sites (note the relatively intense 4.6-kb band in Fig. 6A). The 4.8/6 probe also showed that the *HhaI* and *HaeII* sites upstream of the DPR were unmethylated in a fraction of the cells from each tissue. However, the fragments generated by cleavage at these sites produced only relatively subtle tissue specific differences. In contrast, the 4.8/5 probe detected dramatic tissue specific methylation patterns (Fig. 6B). The complexity of these patterns resulted from various combinations of methylated and nonmethylated sites within the DPR and first intron of *Gdc-1* as well as the single *HhaI*-*HaeII* site 5' of the PPR. Although the number of possible combinations of hybridizing fragments is large because of the number of *HhaI* and *HaeII* sites, it is easy to see that each tissue generated its own methylation-sensitive restriction pattern. These tissue-specific patterns were so pronounced that each tissue ap-

peared to be uniquely fingerprinted by the methylation of specific *HhaI* and *HaeII* sites.

DISCUSSION

The results presented here show conservation of particular domains in the murine and human GPDH promoters. A schematic diagram that summarizes and compares the structural organizations of the human GPD1 and mouse *Gdc-1* promoters is shown in Fig. 7. The comparison begins on the left with a conserved *HindIII* site. This site is located upstream of both GPDH genes and is within the 3' untranslated region of another highly conserved gene present in both species. This upstream gene has been thoroughly characterized in the mouse and was provisionally named *D15Kz1* (21). It has no apparent relationship to GPDH at the sequence level and shows no evidence of coordinate regulation with GPDH in several adult tissues and differentiating 3T3-F442A adipocyte cultures. Thus, *Gdc-1* and *D15Kz1* appear to be regulated independently even though they are in remarkably close proximity. This raises the interesting question of whether there are any intergenic sequences such as enhancers or suppressors that may affect the expression of each gene in different ways. Alternatively, there may be sequences between the two genes which function to insulate one from the effects of regulatory elements in the other.

Between the 3' end of *D15Kz1* and the first exon of *Gdc-1*, there are three blocks of DNA sequence that are highly

FIG. 6. Methylation of *Hha*I and *Hae*II sites in DNA isolated from tissues expressing different levels of GPDH. The DNAs were digested with *Hind*III, followed by digestion with either *Hha*I (·) or *Hae*II (+). Tissue DNAs are abbreviated as follows: K, kidney; S, spleen; L, liver; and E, epididymal fat. The linear map of the 4.6-kb promoter fragment is shown as described in the legend to Fig. 4. (A) Inverted Southern blot hybridized to the 4.8/6 probe (Pr, upper left) from the 5' end of the *Hind*III promoter fragment; (B) similar blot (in opposite orientation) hybridized to the 4.8/5 probe (Pr, lower right) from the 3' end of the *Hind*III promoter fragment. The vertical arrows in the columns containing the size standards indicate the direction of DNA migration during electrophoresis.

conserved in the mouse and human genes. The region furthest upstream from *Gdc-1* is shown as an open rectangle in Fig. 7. This is DPR2, which is 35 bp in length and is near or within a major DNase I-hypersensitive site present in the DNA of mouse liver and brown adipose tissue. Next is the large 888-bp DPR. The 908-bp human DPR is presented as a histogram to indicate the percent sequence identity with the corresponding segments of the mouse DPR. The DPR also contains a major DNase I-hypersensitive site near its 3' end in mouse liver and brown adipose tissue. Interestingly, the DPR2 and DPR hypersensitive sites are not detectable in spleen, which suggests a possible correlation between the low level of GPDH expression in spleen and the absence of these hypersensitive sites.

The third highly conserved region is the PPR, which is immediately 5' to the first exon of the GPDH structural gene. In addition to a TATA-like box sequence, it contains FSE1 and FSE2. Although the FSEs in the GPDH promoter were conserved in the mouse and human genes, only one of the four elements had exactly the same sequence in both species. This is consistent with the sequence divergence seen among the FSEs found in several mouse genes, including the genes encoding GPDH, glutamine synthetase, stearyl coenzyme A desaturases 1 and 2, a serine protease named adipsin, and an adipocyte fatty acid-binding protein named aP2 (5, 19, 22, 31). The upstream FSE1 identical in the mouse and human genes appears twice in the mouse glutamine synthetase promoter (5), whereas the other, inverted

mouse FSE1 appears once. The mouse stearyl coenzyme A desaturase 1 gene (*Scd-1*) promoter (31) from -408 to -394 has a sequence (GCCAATGGAAAAGGC) with 80% homology to the upstream human FSE2. The mouse *Scd-2* promoter (22) from -137 to -123 has a sequence (GCCAATGAGAGCAGC) with 73% homology to this same human FSE2. It is unclear whether this divergence represents a degenerative recognition sequence or the existence of a separate binding protein(s) for each of the slightly different recognition sequences. The second possibility is supported by the fact that a protein complex containing the *fos* proto-oncogene gene product binds to the FSE2 in the aP2 promoter but does not bind to the related FSE2 in the *Gdc-1* promoter (10).

In addition to showing the conserved promoter domains described above, Fig. 7 shows the locations of methylation-sensitive restriction enzyme sites, CpG dinucleotides, and both G+C-rich (>66%) and A+T-rich (>66%) regions. It is the methylation-sensitive *Hha*I and *Hae*II sites located within the DPR and first intron of *Gdc-1* that generate the remarkable tissue-specific methylation patterns seen on Southern blots of DNA from adult mouse tissues. As is the case with tissue-specific genes, the *Gdc-1* promoter is highly methylated in sperm and testis. However, unlike most tissue-specific genes, the *Gdc-1* promoter has a unique hypomethylation pattern in each adult tissue that appears to represent a fingerprint of a tissue-specific developmental program. Perhaps the fingerprint is generated by interactions

FIG. 7. Summary of the structural organization of the GPDH promoter in mouse and human genes. Schematic diagrams of the human and mouse GPDH promoters are shown to illustrate the similar structures of the two promoters. Regions of high sequence similarity are indicated by open boxes. From left to right, these boxes represent *D15Kz1* (5' Gene), DPR2, DPR, PPR, and the first exon (E). In the human promoter, the DPR is presented as a histogram to indicate the percent sequence homology with the corresponding regions of the mouse DPR. The locations of human *Alu* repetitive sequences and mouse B1 repetitive sequences are indicated by arrows. Between the two maps, the regions in each promoter containing >66% G+C (■) or >66% A+T (□) are indicated. Above the human promoter and below the mouse promoter the locations of methylation-sensitive restriction enzyme sites are indicated by vertical bars and the locations and number of CpG dinucleotides are indicated by histograms.

between the DPR and regulatory proteins. This possibility is particularly interesting because it implies that although GPDH is expressed in most adult tissues, the gene may have a developmentally programmed pattern of demethylation similar to that of tissue-specific genes (6). We predict that if there are tissue- or cell-specific *trans*-acting factors interacting with the *Gdc-1* promoter, at least some of them will either bind to DNA sequences near the *HhaI* and *HaeII* sites or interact with other proteins already bound to those sites.

The results presented in this study suggest that the GPDH promoter contains three domains (DPR2, DPR, and PPR) conserved in evolution. It is likely that tissue-specific *trans*-acting factors binding to these regions modulate the absolute level of GPDH expression in many adult tissues. Now that the promoter region is well mapped, *GPD1* gene constructs that contain specific deletions of one or more promoter elements can be used to generate transgenic mice. Assay of both the developmental course and level of human enzyme expression in the transgenic mice will indicate which DNA sequences are required for gene expression in each cell type. For example, levels of gene expression resulting from the deletion of specific FSEs may indicate which tissues contain FSE-binding proteins and suggest the actual function of these proteins in each mouse tissue. In addition, examination of methylation patterns and DNase I-hypersensitive sites in the human transgene may further clarify the relationship between promoter structure and function.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant DK-34384 to Edward H. Birkenmeier from the National Institutes of Health.

LITERATURE CITED

- Arapinis, C., J. Elion, D. Labie, and R. Krishnamoorthy. 1986. Differences in DNase I sensitivity and methylation within the human beta-globin gene domain and correlation with expression. *Eur. J. Biochem.* **156**:123-129.
- Babiss, L. E., A. Bennett, J. M. Friedman, and J. E. Darnell, Jr. 1986. DNase I-hypersensitive sites in the 5'-flanking region of the rat serum albumin gene: correlation between chromatin structure and transcriptional activity. *Proc. Natl. Acad. Sci. USA* **83**:6504-6508.
- Benton, W. D., and R. W. Davis. 1977. Screening lambda gt recombinant clones by hybridization to single plaques in situ. *Science* **196**:180-182.
- Benvenisty, N., and L. Reshef. 1987. Developmental acquisition of DNase I sensitivity of the phosphoenolpyruvate carboxykinase (GTP) gene in rat liver. *Proc. Natl. Acad. Sci. USA* **84**:1132-1136.
- Bhandari, B., K. D. Beckwith, and R. E. Miller. 1988. Cloning, nucleotide sequence, and potential regulatory elements of the glutamine synthetase gene from murine 3T3-L1 adipocytes. *Proc. Natl. Acad. Sci. USA* **85**:5789-5793.
- Cedar, H. 1988. DNA methylation and gene activity. *Cell* **53**:3-4.
- Cook, J. R., and L. P. Kozak. 1982. *sn*-Glycerol-3-phosphate dehydrogenase gene expression during mouse adipocyte development *in vivo*. *Dev. Biol.* **92**:440-448.
- Cook, J. S., J. J. Lucas, E. Sibley, M. A. Bolanowski, R. J. Christy, T. J. Kelly, and M. D. Lane. 1988. Expression of the differentiation-induced gene for fatty acid-binding protein is activated by glucocorticoid and cAMP. *Proc. Natl. Acad. Sci. USA* **85**:2949-2953.
- Cook, K. S., C. R. Hunt, and B. M. Spiegelman. 1985. Developmentally regulated mRNAs in 3T3-adipocytes: analysis of transcriptional control. *J. Cell Biol.* **100**:514-520.
- Distel, R. J., H.-S. Ro, B. S. Rosen, D. L. Groves, and B. M. Spiegelman. 1987. Nucleoprotein complexes that regulate gene expression in adipocyte differentiation: direct participation of *c-fos*. *Cell* **49**:835-844.
- Djian, P., M. Phillips, and H. Green. 1985. The activation of specific gene transcription in the adipose conversion of 3T3 cells. *J. Cell. Physiol.* **124**:554-556.
- Dobson, D. E., D. L. Groves, and B. M. Spiegelman. 1987. Nucleotide sequence and hormonal regulation of mouse glycerophosphate dehydrogenase mRNA during adipocyte and muscle cell differentiation. *J. Biol. Chem.* **262**:1804-1809.
- Fisher, M. 1984. Neuronal influence on glial enzyme expression: evidence from mutant mouse cerebella. *Proc. Natl. Acad. Sci. USA* **81**:4414-4418.
- Fisher, M., and R. J. Mullen. 1988. Neuronal influence on glial enzyme expression: evidence from chimeric mouse cerebellum. *Neuron* **1**:151-157.
- Franza, Jr., B. R., F. J. Rauscher III, S. F. Josephs, and T. Curran. 1988. The fos complex-related antigens recognize sequence elements that contain AP-1 binding sites. *Science* **239**:1150-1153.
- Grainger, R. M., R. M. Hazard-Leonards, F. Samaha, L. M. Hougan, M. R. Lesk, and G. H. Thomsen. 1983. Is hypomethylation linked to activation of delta-crystallin genes during development? *Nature (London)* **306**:88-91.
- Groudine, M., R. Eisenman, and H. Weintraub. 1981. Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation. *Nature (London)* **292**:311-317.
- Heuckeroth, R. O., E. H. Birkenmeier, M. S. Levin, and J. I. Gordon. 1987. Analysis of the tissue-specific expression, developmental regulation, and linkage relationships of a rodent gene encoding heart fatty acid binding protein. *J. Biol. Chem.* **262**:9709-9717.
- Hunt, C. R., J. H.-S. Ro, D. E. Dobson, H. Y. Min, and B. M. Spiegelman. 1986. Adipocyte P2 gene: developmental expression and homology of 5'-flanking sequences among fat cell-specific genes. *Proc. Natl. Acad. Sci. USA* **83**:3786-3790.
- Ireland, R. C., M. A. Kotarski, L. A. Johnston, U. Stadler, E. Birkenmeier, and L. P. Kozak. 1986. Primary structure of the mouse glycerol-3-phosphate dehydrogenase gene. *J. Biol. Chem.* **261**:11779-11785.
- Johnston, L. A., M. A. Kotarski, D. J. Jerry, and L. P. Kozak. 1989. An ubiquitously expressed gene 3.5 kilobases upstream of the glycerol-3-phosphate dehydrogenase gene in mice. *Mol. Cell. Biol.* **9**:935-945.
- Kaestner, K. H., J. M. Ntambi, T. J. Kelly, Jr., and M. D. Lane. 1989. Differentiation-induced gene expression in 3T3-L1 preadipocytes. A second differentially expressed gene encoding stearoyl-CoA desaturase. *J. Biol. Chem.* **264**:14755-14761.
- Kaye, J. S., S. Pratt-Kaye, M. Bellard, G. Dretzen, F. Bellard, and P. Chambon. 1986. Steroid hormone dependence of four DNase I-hypersensitive regions located within the 7000-bp 5'-flanking segment of the ovalbumin gene. *EMBO J.* **5**:277-285.
- Kemper, B., P. D. Jackson, and G. Felsenfeld. 1987. Protein-binding sites within the 5' DNase I-hypersensitive region of the chicken alpha^D-globin gene. *Mol. Cell. Biol.* **7**:2059-2069.
- Keshet, I., J. Lieman-Hurwitz, and H. Cedar. 1986. DNA methylation affects the formation of active chromatin. *Cell* **44**:535-543.
- Kielty, C. M., S. Povey, and D. A. Hopkinson. 1982. Regulation of expression of liver specific enzymes. *Ann. Hum. Genet.* **46**:307-327.
- Kozak, L. P. 1985. Interacting genes control glycerol-3-phosphate dehydrogenase expression in developing cerebellum of the mouse. *Genetics* **110**:123-143.
- Kozak, L. P., D. L. Burkart, and J. P. Hjorth. 1982. Unlinked structural genes for the developmentally regulated isozymes of *sn*-glycerol-3-phosphate dehydrogenase in mice. *Dev. Genet.* **3**:1-6.
- Kozak, L. P., and J. T. Jensen. 1974. Genetic and developmental control of multiple forms of L-glycerol-3-phosphate dehydrogenase. *J. Biol. Chem.* **249**:7775-7781.
- Kunnath, L., and J. Locker. 1985. DNaseI sensitivity of the rat albumin and alpha-fetoprotein genes. *Nucleic Acids Res.* **13**:115-129.

31. **Ntambi, J. M., S. A. Buhrow, K. H. Kaestner, R. J. Christy, E. Sibley, T. J. Kelly, Jr., and M. D. Lane.** 1988. Differentiation-induced gene expression in 3T3-L1 adipocytes. Characterization of a differentially expressed gene encoding stearoyl-CoA desaturase. *J. Biol. Chem.* **263**:17291–17300.
32. **Phillips, M., P. Djian, and H. Green.** 1986. The nucleotide sequence of three genes participating in the adipose differentiation of 3T3 cells. *J. Biol. Chem.* **261**:10821–10827.
33. **Ratner, P. L., M. Fisher, D. Burkart, J. R. Cook, and L. P. Kozak.** 1981. The role of mRNA levels and cellular localization in controlling *sn*-glycerol-3-phosphate dehydrogenase expression in tissues of the mouse. *J. Biol. Chem.* **256**:3576–3579.
34. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
35. **Sheffery, M., R. A. Rifkind, and P. A. Marks.** 1982. Murine erythroleukemia cell differentiation: DNase I hypersensitivity and DNA methylation near the globin genes. *Proc. Natl. Acad. Sci. USA* **79**:1180–1184.
36. **Smith, R. E., and B. A. Horwitz.** 1969. Brown fat and thermogenesis. *Physiol. Rev.* **49**:330–425.
37. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503–517.
38. **Staden, R.** 1977. Sequence data handling by computer. *Nucleic Acids Res.* **4**:4037–4051.
39. **Stein, R., N. Sciaky-Gallili, A. Razin, and H. Cedar.** 1983. Pattern of methylation of two genes coding for housekeeping functions. *Proc. Natl. Acad. Sci. USA* **80**:2422–2426.
40. **Sweetser, D. A., E. H. Birkenmeier, I. J. Klisak, S. Zollman, R. S. Sparkes, T. Mohandas, A. J. Lusis, and J. I. Gordon.** 1987. The human and rodent intestinal fatty acid binding protein genes. *J. Biol. Chem.* **262**:16060–16071.
41. **Tratner, I., J.-L. Nahon, J. M. Sala-Trepat, and A. Venetianer.** 1987. Albumin and alpha-fetoprotein gene transcription in rat hepatoma cell lines is correlated with specific DNA hypomethylation and altered chromatin structure in the 5' region. *Mol. Cell. Biol.* **7**:1856–1864.
42. **Weisbrod, S.** 1982. Active chromatin. *Nature (London)* **297**:289–295.
43. **Yang, V. W., R. J. Christy, J. S. Cook, T. S. Kelly, and M. D. Lane.** 1989. Mechanism of regulation of the 422(aP2) gene by cAMP during preadipocyte differentiation. *Proc. Natl. Acad. Sci. USA* **86**:3629–3633.