Genomic Structure and Promoter Analysis of the Human obese Gene*

Da-Wei Gong, Sheng Bi, Richard E. Pratley,* and Bruce D. Weintraub

From the Molecular and Cellular Endocrinology Branch and SPhoenix Epidemiology and Clinical Research Branch, NIDDK, National Institutes of Health, Bethesda, Maryland 20892

The human gene encoding the homolog of the mouse obese (ob) gene was isolated and partially characterized. The human ob gene consists of three exons and two introns and spans about 18 kilobase pairs (kb), encoding a 3.5-kb cDNA. A 3-kb 5’-flanking region of the gene was cloned and transient transfection assay with luciferase reporter confirmed the promoter activity in differentiated F442-A adipocytes. Potential regulatory elements are discussed in this report.

Obesity is a common complex disease that increases the risk of hypertension and diabetes. Both genetic and environmental factors play a role in the development and maintenance of human obesity (for review, see Refs. 1-4). The multifactorial nature of obesity makes the genetically obese animal a useful model to dissect the pathogenesis of the disease, and several different models of obesity have been established in rodents (5). The mouse with the obese mutation (ob) is obese and develops diabetes, and is therefore a model for related human diseases (6). Recently, the mouse ob gene and its human homolog were cloned; a premature stop codon mutation or the total absence of mRNA results in the obesity in the ob/ob mice. Structural analysis of the ob protein sequence suggests the protein may be secreted as a circulating factor to regulate body weight (6).

The relationship of the ob gene product to human obesity is still not clear. No mutations were found in the coding region of human ob cDNA in obese people in a recent study (7). The ob mRNA was found to be increased in ob/ob mice with a nonsense mutation (6) and in the obese people (7), suggesting that the ob gene expression might be transcriptionally regulated. Thus, we are interested in how the ob gene expression is regulated in normal and obese subjects in order to understand the pathogenesis of obesity and the biology of the adipocyte. As a first step, we report here the isolation the human ob gene including its promoter region.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom all correspondence should be addressed: Bldg. 10/Rm. 8D14, 10 Center Dr. MSC 1758, Bethesda, MD 20892-1758. Tel.: 301-496-4152; Fax: 301-496-1649.

MATERIALS AND METHODS

Oligonucleotides—All the oligonucleotides were made on an Applied Biosystems model 392 DNA synthesizer. The sequence and position of the oligonucleotides are shown in Fig. 4.

Human mRNA—Fat tissue was obtained from normal Pima Indians by percutaneous needle biopsy under local anesthesia with an NIH-approved clinical protocol. mRNAs from three individuals were extracted using Trizol reagent (Life Technologies, Inc.), and equal parts of mRNA from each person were pooled for subsequent reverse transcriptions.

Analysis of the 5’-Untranslated Region— Primer extension was performed to identify the start site of the transcription. 5 μg of total RNA from human fat or yeast RNA was reverse transcribed with end-labeled primers p166 or p360 under the reaction conditions recommended by the supplier of reverse transcribe SuperScript H+ (Life Technologies, Inc.). The reaction product was analyzed on 6% acrylamide denaturing gel. 5’- Rapid amplification of cDNA end (RACE)1 system (Life Technologies, Inc.) was used to obtain the sequence of 5’-untranslated region (UTR). First strand cDNA was synthesized from human fat mRNA using an antisense primer p148 corresponding to the amino acid positions 159 to 152 of the Ob protein. Following dC tailing, two rounds of PCR amplification were performed using the anchor primer with p148 and then a nested primer, p146. The amplified fragments from the RACEs were blunt-ended with Klenow fragment, phosphorylated, subcloned into pKS/Smal vector, and subsequently sequenced.

Isolation of Human Genomic Clone—Only the coding cDNA sequence of human ob gene is available from the GenBank (accession no. U18915 (1995)). Based on the sequence, a pair of primers (p145 and 146) were designed to screen a human genomic P1 library (Genome Systems, Inc.) by a PCR-based cloning strategy (8). The P1 plasmid DNA was digested with appropriate restriction enzymes and subcloned into pKS+ (Stratagene) vectors for further analysis. DNA was sequenced by the dyeoxy chain termination method with Sequenase 2.0 (U. S. Biochemical Corp.) or the fluorescence dye termination cycle sequencing method (Applied Biosystems).

Transfections and Reporter Assays—Mouse F442-A preadipocyte was a generous gift from H. Green (Harvard Medical School). The preadipocyte was induced with insulin and triiodothyronine for 14 days into adipocyte (9) before transfection. Briefly, the cells were transfected with 2 μg of plasmid and 20 μl of Lipofectamine (Life Technologies, Inc.) for 12 h in 1 ml of OptiMEM (Life Technologies, Inc.), after which 2 ml of fresh 10% serum/Dulbecco’s modified Eagle’s medium was added, and incubated another 48 h. The 5’ deletions were made from pGL3/3kb(+) by religation after appropriate restriction enzyme digestion and confirmed by sequencing. The cells were lysed with cell culture lysis buffer (Promega) and assayed for chemiluminescence. Protein was quantitated with the Bradford assay kit (Bio-Rad) to normalize the luciferase activities.

Human IMAGE consortium cDNA clones (LLNL) were obtained from Research Genetics, Inc.

RESULTS AND DISCUSSION

Transcription Start Site Determination—the 5’-UTR of the human ob mRNA is not known. Primer extension was performed to obtain the distance from the primer to the transcription start site. As shown in Figs. 1 and 4, with the two independent primers p166 and p360, the major extended product extended an additional 51 nucleotides from the protein-coding signal ATG. To obtain the sequence of the 5’-UTR, 5’-RACE was performed. A single band of about 400 base pairs was obtained after two rounds of PCR using first primer P146 and a nested primer P146 with the anchor primer (data not shown). The nucleotide sequence analysis revealed that the downstream sequence of fragments was identical to the coding se-
The Human ob Gene

**Fig. 1.** Primer extension analysis on 5′-end of human ob gene. A, 5 μg of human fat total RNA (lanes 1 and 3) or yeast RNA (lanes 2 and 4) was subjected to primer extension with primers p166 (lanes 1 and 2) and p360 (lanes 3 and 4). The length of major extended products, determined by DNA ladders, is shown by arrows. B, schematic summary of the result with numbers referring to the start of exon 1.

**Fig. 2.** Genomic structure of the human ob gene. The position of the three exons is indicated by boxes. The coding region is depicted in black, while the untranslated regions are represented by open boxes. The restrictions sites used in subcloning and mapping are shown. The genomic fragments subcloned into KS vectors are outlined in bold. An enlargement of the promoter region shows the relative positions of repetitive sequences and putative regulatory elements. The IMAGE consortium clones were aligned with exon 3.

**Fig. 3.** Transfection assay of the human ob gene promoter. 2 μg of plasmid was transfected into differentiated F-442A adipocyte in a six-well plate and assayed for luciferase activities. Deletions of the 5′-flanking region were constructed at the indicated restriction site. The vector pGL3-Basic has no promoter. Luciferase activities were normalized by protein content and are shown as percentages of that of pGL3.3kb(+) promoter. Each bar represents the mean and standard deviation of three independent transfections.

Genomic Structure of Human ob Gene—The 5′-most probe p194 derived from 5′-RACE (see Fig. 4) was used to identify the corresponding genomic fragment. A 8-kb genomic fragment C2066 was obtained, and sequence analysis revealed a 29-bp perfect match of the genomic DNA to the cDNA, indicating that the fragment contains 29 bp of the first exon.

Using end-labeled probes p143 and p148, the genomic fragments containing exon 2 and 3 were cloned, which encoded the Ob protein (Figs. 2 and 4). The sizes of the first and second introns were about 10 and 1.9 kb, as determined by restriction enzyme mapping and PCR, separately (Fig. 2; data not shown).

The sequence analysis showed that the upstream sequence of the genomic fragment C2012 included exon 3 and encoded the C terminus of the Ob protein. A further analysis of the downstream sequence to nucleotide 2241 revealed 71% identity to mouse 3′-UTR (Bestfit, Genetics Computer Group, Inc.), indicating this region may be exonic. 3′-RACE was performed in order to isolate the entire human 3′-UTR. A 1.1-kb fragment was obtained after two-round PCR using primers first p180 and nest p197 with an anchor primer (data not shown). Sequencing revealed that the fragment was identical to the human genomic sequence and shared homology to mouse ob cDNA, indicating the fragment was a 3′-UTR of the human ob cDNA. The fragment apparently ended at the poly(A) sequence corresponding to genomic DNA (3066–3081), suggesting the ending might be a premature termination due to the annealing of oligo(dT)$_{18}$ in the reverse transcription. A BLAST search$^2$ using the human ob 3′-sequence from positions 2541 to 3081 found perfect matches to 5′-ends of two IMAGE consortium (LLNL) human ob cDNA clones (10). The clones were sequenced in full and found to be identical to the human ob genomic sequence (Fig. 2). Clones 139081 and 182874 were derived from a human placenta and breast library, respectively, consistent with our Northern analysis that mRNA ob is expressed in human fat and placenta (data not shown). Moreover, a putative termination signal AATAAA was found 20 bp before the mRNA polyadenylation sequence in both clones, indicating they are authentic 3′-ends of the human ob mRNA (Figs. 2 and 4). BLAST search found that the 3′-UTR of the human ob mRNA contains an Alu repetitive sequence from positions 3081 to 3370. The physiological function of the repeat, if any, remains to be explored. Taken together, the human ob gene was composed of three exons, encoding a 3.5-kb cDNA and spanning around 18 kb of the genome (Figs. 2 and 4). The sequence of the intron/exon junction followed the TG/GA rule for splicing.

Promoter Analysis—To confirm that the 5′-flanking region actually conferred promoter activity, different lengths of the genomic fragment with the first exon were subcloned into lu-

$^2$The BLAST data base is available through Gopher server at ncbi.nlm.nih.gov.
ciferase reporter vectors. Although the F442A cell started to express ob mRNA 3 days after induction (11), the transfection was performed in 14-day-induced F442A adipocytes that were 90% differentiated and expressing ob mRNA as detected by RT-PCR (data not shown). As shown in Fig. 3, compared with the promoterless pGL3-Basic vector and antisense 3-kb construct whose activities were almost undetectable, the sense reporter constructs showed remarkable luciferase activities, indicating that this 5'-flanking region contains all of the necessary elements to support basal transcription and serves as a positive regulatory element.

**Fig. 4. Nucleotide sequence of the human ob gene.** The sequences of all the exons (uppercase letters), the promoter region, and exon/intron junction (lowercase letters for intron) are shown. Nucleotide numbering starts with the +1 as defined by primer extension and 5'-RACE, and proceeds as indicated in the left margin, taking into account only exonic sequences. The deduced amino acid sequence and numbering are shown below the nucleotide sequence. The oligonucleotides used for PCR or probe are underlined. The putative TATA core promoter at -228 bp is bold, and potential transcription regulatory sites are double-underlined. C/EBP, CCAAT/enhancer-binding protein; GRE, glucocorticoid response element; CREB, cAMP response element-binding protein. The putative polyadenylation signal is boxed.
promoter. Moreover, the 300-bp fragment was apparently more active than the 3-kb fragment (Fig. 3), indicating the remote region might be inhibitory for transcription in the F442A adipocytes.

BLAST search with the 3-kb promoter region reveals several features of the region. Two repetitive sequences MER11 (medium reiteration frequency; Ref. 12) and Alu were found from –2514 to –1545 and from –1364 to –1046. Repetitive elements have been involved in regulation of several genes, such as human CD8α (13), keratin 18 gene (14), and immunoglobulin κ light chain gene (15). Whether the above MER11 and Alu repeats are functional in regulating the ob expression remains to be studied.

A computer scan of the 3-kb 5′-flanking region with the TFD (Transcription factor data base, with 0 mismatch) from GenBank disclosed dozens of putative binding sites for known transcription factors. Of the notables are Sp-1 sites, cAMP response element, glucocorticoid response element, and CCAAT/enhancer-binding protein sites. The ectopic expression of CCAAT/enhancer-binding protein has been found to induce adipogenesis in F442A preadipocytes (16). The ob gene expression has been reported to respond to insulin (11, 17) and dexamethasone (18). Whether the above elements are involved in the regulation is under investigation.

Acknowledgments—We thank S. Cushman for help and encouragement, M. Rechler and G. Ooi for critical reading of the manuscript, and J. Woottton for pointing out the MER11 repeat. We are grateful for G. Poy and J. Liu for DNA sequencing.

REFERENCES