Structure and Expression of the Human Angiotensinogen Gene

IDENTIFICATION OF A UNIQUE AND HIGHLY ACTIVE PROMOTER

Akiyoshi Fukamizu, Shigeru Takahashi, Min Seok Seo, Masazumi Tada, Keiji Tanimoto, Sayuri Uehara, and Kazuo Murakami

From the Institute of Applied Biochemistry, University of Tsukuba, Tsukuba Science City, Ibaraki 305, Japan

We have isolated the human angiotensinogen gene from a genomic library and determined the exon-intron junction sequences. The gene is 12 kilobases long and consists of five exons interrupted by four introns, as a single copy in the human genome. Of particular interest are the positions of the introns in the human angiotensinogen gene which are identical to those in the highly homologous human α1-antitrypsin and α1-antichymotrypsin genes, as well as rat and mouse angiotensinogen genes. Northern blot analysis showed that human hepatoma cells (HepG2) produce a large amount of angiotensinogen mRNA but not human glioma cells (T98G). To assay the promoter activity, the 1.3-kilobase genomic fragment containing the 5'-flanking region, first exon, and a part of first intron at positions -1222 to +44 was fused upstream to the chloramphenical acetyltransferase gene, then transfected into HepG2 and T98G cells. The gene sequence was active only in HepG2 cells, suggesting the presence of a functional promoter. Analysis of deletion mutants demonstrated that the 76-base pairs region from -82 to +44 containing the TATA box and first exon is the minimal promoter, whose activity is as high as that of the SV40 enhancer-promoter. Since the basal expression of the human angiotensinogen gene is much higher in HepG2 than T98G cells, these results may reflect cell-specific differences in the gene transcription.

Human angiotensinogen is a hormone precursor of M, 61,400. The mature form consists of a single polypeptide chain of 452 amino acid residues with 14% carbohydrate content (1, 2). Angiotensinogen is synthesized primarily by the liver and released into the circulation, where it is cleaved by renin (EC 3.4.23.15), an aminopeptidase, to generate a decapeptide, angiotensin I. It is subsequently converted to a multifunctional octapeptide, angiotensin II by angiotensin-converting enzyme.

One striking effect of angiotensin II is to cause arteriolar vasoconstriction (3). The renin-angiotensin system therefore plays an important role in the regulation of blood pressure and in the pathogenesis of hypertension (4). In addition to this peripheral effect, angiotensin II in the central nervous system (5) has important actions including stimulation of thirst, alteration of sympathetic outflow, and release of several pituitary hormones such as vasopressin, corticotropin, and prolactin. Still other important roles of angiotensin II are its effect on stimulation of new vessel formation (6) and reproductive functions such as follicular development and ovulation (7).

Although a number of in vivo studies with rats (8-13) revealed that angiotensin production is regulated by several factors including steroid hormones, as well as by certain pathophysiological conditions, the regulatory mechanisms of the human angiotensinogen gene expression have not been clarified. In the present study, in order to understand the basal regulation of the gene expression, we determined the structure of the human angiotensinogen gene and identified the minimal promoter region required for the cell-specific expression of the gene.

EXPERIMENTAL PROCEDURES

Materials—Restriction enzymes and DNA-modifying enzymes were purchased from Takara Shuzo, Nippon Gene, Toyobo, and New England Biolabs. BglII and HindIII linkers were obtained from Takara Shuzo. [a-32P]dCTP (400 and 300 Ci/mmole) was purchased from Amersham Corp. [32P]Chloramphenicol (50-80 Ci/mole) was from DuPont-New England Nuclear.

Screening of Human Genomic Library—Human genomic DNA was obtained from placenta as described (14). A genomic library was constructed in λ phage vector Charon 28 from a partial Sau3A1 digest of genomic DNA using standard methods (15). Approximately, 6 X 10^6 phage plaques were screened with probe A as described below. Hybridization was carried out at 65 °C for 16 h in 5 X SSPE (1 X SSPE = 0.15 M NaCl, 0.01 M NaH2PO4, H2O, 1 mm EDTA, pH 7.4), 1 X Denhardt's solution (0.02% bovine serum albumin, 0.02% polyvinylpyrrolidone, 0.02% Ficoll), 0.1% SDS, 100 μg/ml denatured salmon sperm DNA, and 5 X 10^5 cpm/ml labeled probe A. After hybridization, filters were washed twice with 2 X SSC (1 X SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0) at room temperature, twice with 2 X SSC and 0.1% SDS at 65 °C for 30 min, and twice with 0.1 X SSC and 0.1% SDS at 65 °C for 30 min. Dried filters were subjected to autoradiography at -70 °C with an intensifying screen.

Subcloning of Genomic DNA Fragments—The 3.5-kb EcoRI/BamHI, 5.9-kb BamHI, and 9.5-kb BglII fragments from the recombinant phAG-1 were subcloned into the polylinker region of plasmid pUC19 and were designated phAG35EB, phAG90E, and phAG95G, respectively. The 2.7-kb BamHI fragment from phAG95G was inserted into the BamHI site of pUC19 to generate phAG27B.

Preparation of Probes—A 1293-bp EcoRI fragment, designated probe A, was isolated from phAG3 (1) and used for plaque screening. A 1293-bp EcoRI fragment (the 5'-end probe, probe B) was excised from phAG35EB and used for Southern blot analysis. A 413-bp Rsal fragment (the exon 2 probe, probe C) was from phAG27B and used for Southern and Northern blot analyses. These probes were labeled with [α-32P]dCTP by the random primer method (16).

Southern Blot Analysis—Human placenta DNA (10 μg) and the cloned bacteriophage DNA (0.3 pg) were digested with a variety of restriction enzymes and Southern blot analysis was performed as described above.
Human Angiotensinogen Gene Expression

restriction enzymes, electrophoresed on 0.7% agarose gels, and transferred to GeneScreen Plus (nylon) membrane (Du Pont-New England Nuclear) as described (17). Hybridization was conducted at 65 °C for 16 h in 1 M NaCl, 1% SDS, 10% dextran sulfate, 100 μg/ml denatured salmon sperm DNA, and 1 × 10^6 cpm/ml labeled probe B or C. Filters were washed twice with 2 × SSC at room temperature for 5 min, twice with 2 × SSC and 1% SDS at 65 °C for 30 min, and twice with 0.1 × SSC at room temperature for 30 min. Filters were then subjected to autoradiography at −70 °C with an intensifying screen.

Preparation of RNA and Northern Blot Analysis—Total RNA was isolated from human liver, HepG2, and T98G cell lines by the guanidinium-cesium chloride method (18). Total RNA was denatured with 1 M glyoxal and 50% dimethyl sulfoxide (19), electrophoresed on a 1.2% agarose gel, and transferred to GeneScreen Plus membrane. Hybridization was carried out at 60 °C for 16 h in 1 × SSC and 1% SDS, 50% dextran sulfate, 100 μg/ml denatured labeled salmon sperm DNA, and 1 × 10^6 cpm/ml labeled probe C. The filter was washed twice with 2 × SSC at room temperature for 5 min, twice with 2 × SSC and 1% SDS at 60 °C for 30 min, and twice with 0.1 × SSC at room temperature. The filter was then subjected to autoradiography at −70 °C with an intensifying screen.

DNA Sequence Analysis—DNA fragments carrying exons and the 5' and 3'-flanking regions were identified by restriction mapping. After subcloning into M13 or pUC19 vectors, the sequence of exon/exon/intron boundaries, and flanking regions were determined by the dideoxynucleotide chain termination method (20).

Plasmid Constructions—pSV2cat (21) was digested with AccI, made blunt-ended by treatment with Escherichia coli DNA polymerase I (Klenow subunit) and ligated to BglII linker. The DNA was then digested with BglII and BamHI to generate a 2159-bp fragment. pUC19 was digested with HindIII, converted to blunt ends by treatment with E. coli DNA polymerase I (Klenow subunit) and ligated to HindIII/BglII linker. This vector DNA was digested with HindIII and BamHI to construct pUCSV2cat.

The vector and the BglII/BamHI fragment from pSV2cat were ligated to construct pUCSV2cat.

A 1651-bp HindIII/BamHI fragment derived from pSV2cat was inserted into HindIII/BamHI sites of pUC19 to construct pUC0cat. A 1288-bp Apal fragment from phAG35EB was treated with T4 DNA polymerase I to make blunt ends and was then ligated to HindIII linker. This fragment was digested with HindIII and ligated into HindIII site of pUC0cat. Using this procedure, insertions representing both orientations were obtained to construct pUChAG(-1222)cat (sense orientation) and pUChAG(-1222)catRv (reverse orientation). Deletion mutants were produced by digestion with Exonuclease III as described (22), and the resulting end points were determined by sequencing.

Cell Culture, DNA Transfection, and Chloramphenicol Acetyltransferase Assay—Human hepatoma cells (HepG2) and human glioma (T98G) cells were maintained in minimum essential medium (GIBCO) containing 10% fetal bovine serum and nonessential amino acids.

One day before transfection, 1 × 10^6 and 0.5 × 10^6 cells for HepG2 and T98G cells, respectively, were plated on 60-mm dishes. The cell medium was changed 3 h prior to transfection. A calcium phosphate coprecipitate containing 3 μg of DNA was added to the cultured cell lines. Cells were incubated for 6 h, washed, and incubated for 36 h in fresh medium. Cells were collected and cell extracts were prepared by freezing and thawing as described (21). The extracts were heated at 60 °C for 10 min, and the precipitate was removed by centrifugation. The protein concentration was determined using bovine serum albumin as a standard (23). The reaction mixture contained 140 mM Tris-HCl, pH 7.8, 0.2 μCi of [3H]chloramphenicol, 4 mM acetyl-coenzyme A (Pharmacia LKB Biotechnology Inc.), and 40 μg of cell extract in a final volume of 150 μl. The mixture was incubated at 37 °C for 1 h and then extracted with cold ethyl acetate. The solution was dried, and the pellet redissolved in 15 μl of ethyl acetate. The labeled chloramphenicol and acetylated derivatives were separated by as-

![Diagram](http://example.com/diagram.png)

**Fig. 1.** Structure of the human angiotensinogen gene. A, partial restriction map and relevant restriction sites of λhAG-1 carrying the human angiotensinogen gene are displayed. The five exons are represented by solid bars. Restriction sites are indicated as follows: A, Apal; B, BamHI; G, BglII; L, ApaLI. B, Southern blot analysis. Human genomic DNA (lanes 1, 2, and 3) and λhAG-1 (lanes 2, 5, and 6) were digested with ApaLI/BglII (lanes 1 and 2), BglII (lanes 3 and 5), and BamHI (lanes 4 and 6). Probe B was used for lanes 1 and 2; probe C was used for lanes 3-6. λDNA digested with HindIII served as a size marker (sizes in kilobases).
FIG. 2. Nucleotide sequence of the 5'-flanking region, all five exons, and the 3'-flanking region of the human angiotensinogen gene. The nucleotide sequence (second line) is shown with the deduced amino acid sequence (first line). Nucleotides are numbered at the left with the putative initiation site of transcription underlined by dashed lines. Filled and open circles indicate possible glucocorticoid and estrogen responsive elements, filled and open squares indicate the consensus sequences for RNA polymerase III promoters and the heat shock sequence is boxed with a solid line. Arrowheads indicate the positions where the four introns are inserted in the gene. The termination codon is marked by asterisks. TATA box and the putative polyadenylation signals are underlined by dotted lines.
RESULTS

Isolation and Characterization of the Human Angiotensinogen Gene—A human genomic DNA library (6 x 10⁶ phages) was screened with probe A, resulting in isolation of one positive clone, XhAG-1. This clone, containing an insert of 16 kb, was characterized by a combination of restriction enzymes and Southern blot analyses. Human genomic and the cloned angiotensinogen gene have not rearranged during cloning procedures. This result also suggests that there is a single gene encoding angiotensinogen in the human genome.

Structure of the Human Angiotensinogen Gene—Nucleotide sequences of all exon-intron junctions have been determined. The human angiotensinogen gene is composed of five exons separated by four introns, which are identified by comparison with the cDNA sequence (1). As in other eukaryotic genes, the introns begin with the nucleotides GT and end with AG (24). The nucleotide sequence of all exons and their flanking sequences are in capital letters; intron sequences are in lower-case letters. hAAT, human α₁-antitrypsin; hACT, human (Y₁-antichymotrypsin; hANG, rANG, and mANG are human, rat, and mouse angiotensinogen, respectively.

![Fig. 2—continued](image)

![Fig. 3. Exon-intron junction sequences of hAAT, hACT, hANG, rANG, and mANG genes. Exon sequences are in capital letters; intron sequences are in lower-case letters. hAAT, human α₁-antitrypsin; hACT, human (Y₁-antichymotrypsin; hANG, rANG, and mANG are human, rat, and mouse angiotensinogen, respectively.](image)
regions is shown in Fig. 2. The nucleotide numbering was designated by assigning the proposed transcription initiation site (1) as nucleotide 1.

The first exon contains 36 bp of untranslated nucleotides that are present in the 5′ region of the transcript (1). The second exon consists of 889 nucleotides in which the signal peptide and angiotensin I are encoded. The third and fourth exons are composed of 268 and 165 nucleotides, respectively. The last exon, exon 5, consists of 796 nucleotides in which two possible polyadenylation signals are present.

Comparison of the exon sequences with the cDNA sequence (1) revealed two differences. In the protein coding region, a T to C transition at position 842 changes a methionine codon in the cDNA to a threonine codon in the genomic DNA. Another is an A to C transversion at position 2054 in the 963-nucleotide cDNA coding for human al-antichymotrypsin genes with respect to the number and position of the introns (27). Therefore, we compared the intron positions of the human angiotensinogen gene with those of the human α1-antitrypsin (28) and human α1-antichymotrypsin genes with respect to the number and position of the introns (27). Therefore, we compared the intron positions of the human angiotensinogen gene with those of the human α1-antitrypsin (28) and human α1-antichymotrypsin genes (29), as well as rat (27) and mouse (26) angiotensinogen genes. All these genes have three features in common. 1) They consist of five exons and four introns. 2) The first intron is located in the 5′-untranslated portion of the genes. 3) The second, third, and fourth introns are in homologous positions and are interrupted in the same manner (Fig. 3). These similarities in structure are consistent with a hypothesis that the serine protease inhibitor gene family evolved from a common ancestral gene through a series of gene duplications, insertions, and deletions (30).

Analysis of the Nucleotide Sequence of the 5′-Flanking Region—Sequence analysis of the 5′-flanking region revealed a number of interesting sequence motifs (Fig. 2). The putative promoter region contains a classical TATA box (24) at positions −31 to −34. This analysis did not exceed the background level for pUCOcat which was less than 1% relative to pUCSV2cat. This complete inactivity of all deletion mutants in T98G cells was not due to a lower transfection efficiency, since expression directed by pUCSV2cat was higher in these cells than in HepG2 cells. In HepG2 cells, pUCHAG(−1222)cat elicited a high level of chloramphenicol acetyltransferase activity. In contrast, pUCHAG(−1222)catRv failed to express the enzyme activity, indicating that this expression is dependent upon the orientation of the gene promoter. Deletion from −1222 to −33 resulted in no significant change in the promoter activity. The removal of the region between −16 and +44 containing exon 1 from pUCHAG(−1222)cat resulted in a 95% reduction of the promoter activity, showing that this domain represents a region of functional importance for the efficient gene expression. Together with the fact that all deletion mutants are not transcribed in T98G cells, these findings suggest that the 76-bp region from −32 to +44 containing the TATA box and exon 1 is the minimal promoter required for cell-specific expression of the angiotensinogen gene at the basal level, a 1.3-kb ApaI fragment (Fig. 1A) of the gene containing 1222 bp of the 5′-flanking region, 36 bp of exon 1 (Fig. 2), and 8 bp of intron 1 (Fig. 3) at positions −1222 to +44 was inserted at a unique HindIII site of pUC0cat. pUCHAG(−1222)cat has the 1.3-kb fragment in the sense orientation to the chloramphenicol acetyltransferase gene, and pUCHAG(−1222)catRv has the sequence in the reverse orientation. A series of deletion mutants extending from −1222 to −33 and from +44 to −16 were constructed. These constructions were then introduced into HepG2 and T98G cells (Fig. 5).

In T98G cells, chloramphenicol acetyltransferase transcription did not exceed the background level for pUC0cat which was less than 1% relative to pUCSV2cat. This complete inactivity of all deletion mutants in T98G cells was not due to a lower transfection efficiency, since expression directed by pUCSV2cat was higher in these cells than in HepG2 cells. In HepG2 cells, pUCHAG(−1222)catRv elicited a high level of chloramphenicol acetyltransferase activity. In contrast, pUCHAG(−1222)catRv failed to express the enzyme activity, indicating that this expression is dependent upon the orientation of the gene promoter. Deletion from −1222 to −33 resulted in no significant change in the promoter activity. The removal of the region between −16 and +44 containing exon 1 from pUCHAG(−1222)cat resulted in a 95% reduction of the promoter activity, showing that this domain represents a region of functional importance for the efficient gene expression. Together with the fact that all deletion mutants are not transcribed in T98G cells, these findings suggest that the 76-bp region from −32 to +44 containing the TATA box and exon 1 is the minimal promoter required for cell-specific expression of the angiotensinogen gene.
expression of the human angiotensinogen gene in this transient assay.

**DISCUSSION**

Nucleotide sequence and Southern blot analyses revealed that the human angiotensinogen gene consists of five exons and four introns distributed over 12 kb and that it exists as a single gene in the human genome.

Northern blot analysis showed that the human angiotensinogen mRNA is detected in human hepatoma cells (HepG2) but not in human glioma cells (T98G), implying that the expression of the human angiotensinogen gene is regulated in a cell-specific manner. As a first approach for elucidating the cell-specific regulation of the gene expression at the basal level, we have identified the minimal promoter of the human angiotensinogen gene by deletion analysis using transient chloramphenicol acetyltransferase assay.

It has been shown that the upstream elements of genes transcribed by RNA polymerase II are required for maximum levels of gene expression and are involved in tissue- or cell-specific expression of genes (40). In most cases thus far examined, the loss of transcriptional activities has been shown to be associated with progressive deletions of the 5'-flanking region. Unexpectedly, our results indicated that successive deletions of the upstream region from -1222 to -106 have little effect on the high level of expression of the human angiotensinogen-chloramphenicol acetyltransferase hybrid genes in HepG2 cells. It was even more surprising to find that the high level of chloramphenicol acetyltransferase expression is still observed in a minimum gene construct, pUChAG(-32)cat, containing the TATA box (at positions -31 to -24), exon 1 (at positions +1 to +36), and a part of intron 1 (at positions +37 to +44). On the other hand, the same deletion mutant did not express the chloramphenicol acetyltransferase activity in T98G cells. Further analysis showed that deletion of the region between -16 and +44 containing exon 1 leads to significant reduction of the chloramphenicol acetyltransferase transcriptional activity even though the TATA box and 1.2-kb upstream sequences are retained.

Our results can raise two questions: why is the expression of pUChAG(-32)cat as efficient as pUCSV2cat in HepG2 cells, and why is it suppressed in T98G cells? One possible explanation for these questions is that a cell-specific enhancer protein or a positive regulatory factor may exist, which can act on sequences within the short regulatory region only in HepG2 cells. However, no sequence homology could be found between consensus sequences of any known viral enhancer and the region from -32 to +44. Another is that a cell-specific activating protein can stimulate transcription by facilitating the formation of the preinitiation transcription complex. In any case, it is unique that the basal expression of the human angiotensinogen gene is directed only by the 76-bp sequences at positions -32 to +44.

Although we found a number of tentative regulatory sequences, including glucocorticoid, estrogen, acute phase,
cAMP, and heat shock-responsive elements in the 5’-flanking region of the human angiotensinogen gene, it should be emphasized that none of the above sequence motifs have yet been shown to function in the expression of the angiotensinogen gene at the basal level. However, this information will serve as a guide in studying the mechanisms that underlie the regulation of the gene transcription.

A more detailed characterization of the 76-bp regulatory region at the basal level and of the putative upstream elements at the regulatory level will be required for a more thorough understanding of the regulation of the human angiotensinogen gene. It would be also interesting to investigate how a protein factor(s) interacts with the short regulatory region and enhances the transcriptional activity.

Acknowledgment—We are especially grateful to Drs. Shigetada Nakanishi and Hiroaki Ohkubo (Kyoto University) for providing us with the human angiotensinogen cDNA (pHag3). We wish to thank Dr. Youji Mitsui (Agency of Industrial Science and Technology) for providing us with HepG2 cell line. We thank Dr. Miranda Mirfakhrai and Norma Foster for critical reading of the manuscript.

REFERENCES
Structure and expression of the human angiotensinogen gene. Identification of a unique and highly active promoter.
A Fukamizu, S Takahashi, M S Seo, M Tada, K Tanimoto, S Uehara and K Murakami