Characterization of the Complete Genomic Structure of the Human WNT-5A Gene, Functional Analysis of its Promoter, Chromosomal Mapping, and Expression in Early Human Embryogenesis*

(Received for publication, September 5, 1995)

Keith G. Danielson‡, J huma Pillarisetti, Isabelle R. Cohen‡, Bijan Sholehvar‡, Kay Huebner§, Ling-J im Ng‡, J hoon M. Nicholls, Kathryn S.E. Cheah§, and Renato V. Iozzo‡§**

From the ‡Department of Pathology, Anatomy and Cell Biology and the §Jefferson Cancer Institute, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, AND the Departments of *Biochemistry and ‡Pathology, Hong Kong University, 5 Sassoon Road, Hong Kong

We report the complete genomic organization of the human WNT-5A gene, which encodes a cysteine-rich growth factor involved in cell-cell signaling during growth and differentiation. The gene comprises five exons with the terminal exon coding for a large 3'-untranslated region of ~6.5 kilobase pairs and utilizes multiple polyadenylation signals to generate at least four discrete transcripts. We discovered a new leader exon interrupted by a 411-base pair intron that was retained in our original cDNA cloning. The promoter region was located in a GpC-rich island and harbored numerous cis-acting elements including several GC boxes and Sp1, AP1, and AP2 binding motifs. It lacked TATA or CAAT boxes typical of housekeeping and growth factor genes. In support of this, primer extension revealed two transcription start sites. Transient cell transfection assays showed functional promoter activity for the 3.9-kilobase pair 5'-flanking region. Interestingly, internal and 5' deletions revealed that the distal promoter was not required for full transcriptional activity and that the first 631 base pairs of WNT-5A harbored the strongest promoter activity. Using a panel of rodent-human hybrid DNAs carrying portions of chromosome 3p, we mapped the gene to 3p14.2-p21.1, between a constitutional and a familial renal cell carcinoma-associated translocation. In situ hybridization analyses of early human embryos at 28–42 days of gestation revealed that WNT-5A transcripts were not restricted to the developing brain and limbs but were also observed in the mesenchyme bordering the pharyngeal clefts and pouches and in the developing gonads and kidneys. The relatively high expression in the cecum epithelium and in the precursors of follicles and seminiferous tubules suggest a novel role for WNT-5A in germ-cell differentiation. This study provides the molecular basis for discerning the regulation of the WNT-5A gene and offers the opportunity to investigate genetic disorders linked to this important gene.

The highly conserved WNT genes belong to a growing family of signaling molecules and potential proto-oncogenes that are involved not only in mammary carcinogenesis but also in the regulation of pattern formation during embryogenesis and differentiation of cell lineages (1–5). In general, WNT genes encode 38–43-kDa glycoproteins with features typical of secreted growth factors, including a hydrophobic signal peptide, the absence of additional transmembrane domains, highly conserved cysteine residues, and the presence of N-linked glycosylation sites (5). WNT proteins are closely associated with the cell surface (6) and the extracellular matrix (7). Accordingly, the radius of action of the paracrine effect extends across a few cell lengths, consistent with a factor diffusing over short distances in the extracellular matrix (8). Some of the WNT proteins also possess an ability to induce cell proliferation via an autocrine or paracrine route. In this respect, they represent multivalent factors that are capable of governing developmental processes by acting on different cellular systems similarly to the fibroblast growth factor or transforming growth factor-β families (9). For example, the neural cell line PC12 can be phenotypically modified upon transfection with the mammalian WNT-1 cDNA (10), and these effects are associated with modulation of plakoglobin and E-cadherin and an accompanying modulation of calcium-dependent cell adhesion (10). In support of this study is the determination that armadillo, a Drosophila homologue of mammalian plakoglobin, is directly involved in the wg-mediated signal transducing pathway (11, 12).

The first WNT-5A gene was identified in the mouse where multiple transcripts ranging from 4.6 to >8 kb encode a cysteine-rich protein of 379 amino acids whose temporal expression correlates with spatial patterning and morphogenesis (13). The Drosophila (14, 15), frog (16), axolotl (17), and zebrafish (18) ortholog of the mammalian WNT-5A have been cloned, and their expression is also developmentally regulated. Of note, in embryos of Xenopus laevis, overexpression of WNT-5A leads to complex deformations that are distinct from those achieved by expression of other WNT genes (19). These findings suggest that WNT-5A has a potential activity in modifying the morphogenetic movements of tissues (19). During mouse embryonic development, the expression of WNT-5A gene follows a complex spatial and temporal pattern; it is associated with several distinct sites including the central nervous system, the facial processes, and the posterior region of the fetus; in the adult, expression is restricted to brain, heart, and lungs (13). WNT-5A

*This work was supported in part by National Institutes of Health Grants RO1 CA39481 and RO1 CA47282 (to R. V. I.) and RO1 CA51083 (to K. H.), by grants from the Hong Kong University Committee for Conference and Research Grants (to J. M. N. and K. S. E. C.), and by the Arthritis and Rheumatism Council (to K. S. E. C.). 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank[7]EMBL Data Bank with accession number(s) U39837.

**To whom correspondence and reprint requests should be addressed: Dept. of Pathology, Anatomy and Cell Biology, Rm. 249, Jefferson University, 1020 Locust St., Philadelphia, PA 19107. Tel.: 215-955-2208; Fax: 215-923-7969.

† The abbreviations used are: kb, kilobase pair(s); PCR, polymerase chain reaction; RT, reverse transcriptase; CAT, chloramphenicol acetyltransferase; bp, base pair(s).
is expressed at a low level in the mammary gland of virgin animals, but expression is increased during pregnancy, with a peak in early pregnancy followed by a rapid decline to undetectable levels (20, 21). Moreover, WNT-5A expression is up-regulated during retinoic acid-induced neural differentiation of P19 embryonal carcinoma cells (22). Two recent independent reports showed a lack of WNT-5A transforming activity based upon its inability to induce morphological transformation or altered growth characteristics onto C57MG mouse mammary epithelial cells (23, 24). The results of a screening of 10 different WNT members (24) suggest that this family of growth factors should be divided into three distinct groups: (i) a highly transforming group comprising WNT-1, WNT-3A, and WNT-7A, (ii) a moderately transforming group comprising WNT-2, WNT-5B, and WNT-7B, and (iii) a nontransforming group comprising WNT-4, WNT-5A, and WNT-6. Overall, these data demonstrate that WNT genes have distinct effects on cell growth and that they should not be regarded as functionally equivalent (24).

We have recently cloned the human WNT-5A cDNA (25) and discovered an aberrant expression of the gene in several human tumors, including lung and breast carcinomas and melanomas (26). To facilitate the study of human WNT-5A gene and to understand its transcriptional control, we have characterized the entire genomic organization of WNT-5A, established functional promoter activity for the 5'-flanking region in transient cell transfection assays, determined a finer chromosomal location for the gene, and investigated its expression in early human embryogenesis. This knowledge should provide the molecular basis for understanding the genetic control and fine tuning of this important gene during developmental processes and should allow the discovery of genetic diseases linked to this signaling molecule.

**EXPERIMENTAL PROCEDURES**

Materials—All the reagents were of molecular biology grade. Radiolabeled nucleotides (α-32P)dCTP (3000 C/mmol), (α-35S)dATP, (5000 CI/mmol), and (32P)dCTP (1000 C/mmol), and (35S)chloramphenicol (100 μCi/ml) were obtained from Amersham Corp. Restriction enzymes were purchased from Promega Corporation (Madison, WI). PCR products were obtained from Perkin-Elmer and Invitrogen Corp. (San Diego, CA).

Genomic Clones and DNA Sequencing—A lambda phage phix-FIX II library (Clontech) prepared from partial Mol digestion of human WI-38 lung fibroblast DNA was used for screening. Approximately, 5 × 10^9 plaques were screened using two cDNA probes, a 1.4-kb fragment derived from the 5'plaques were screened using two c-DNA probes, a 1.4-kb fragment from the 3'end of the cloned 1.2-kb fragment from the 3'end of the 3'end of the cDNA (25) and a 1.5-kb fragment from the 3'-untranslated region (25). Both probes were labeled to high specific activity by the random priming method (27) utilizing a kit from Stratagene. The filters (Duralase from Stratagene) were hybridized separately for each probe overnight at 65 °C in a solution containing 5 × SSC (1 × SSC = 0.15 M NaCl, 15 mM sodium citrate), 1 × Denhardt’s reagent, 0.1% SDS, and 100 μg/ml denatured salmon sperm DNA. After washing the filters at high stringency (0.5 × SSC, 0.1% SDS at 65 °C), the filters were exposed to Kodak XAR-5 film at −70 °C in a cassette containing intensifying screens and subsequently eleven positive clones were identified. Two strongly positive phage clones designated WI and W10 were selected and purified by sequential plaque screening for further characterization. DNA isolated from each phage clone was characterized by restriction enzyme digestion and Southern blotting analysis and hybridizing fragments containing exons were subcloned into pBluescript II KS (Stratagene). Plasmid DNA was sequenced by the automated sequencing system (Applied Biosystems) utilizing T3 or T7 primers in pBluescript or synthetic oligonucleotide primers. Computer analysis of DNA sequences was done using the GCG or PC/Gene computer programs as described previously (28).

RNA Isolation and Analysis—RNA was isolated from confluent monolayer of human embryonic fibroblasts or MGe-63 osteosarcoma cells by a modified method of Chomczynski and Sacchi (29) using a commercially available "Tri" reagent (Molecular Research Corp., Cincinnati, OH). RNA was dissolved in Formozol (Molecular Research Corp.) and stored at −70 °C. Approximately 10–15 μg of RNA was separated by electrophoresis in 1% agarose gels containing 0.66% formaldehyde (30), transferred to nylon membranes, and hybridized to 32P-labeled DNA probes, as indicated. Hybridization was done overnight at 42 °C in 50% formamide, 5 × SSC, 25 μg/ml sodium phosphate (pH 6.8), 0.5% SDS, 10 × Denhardt’s reagent, and 250 μg/ml denatured salmon sperm DNA. After hybridization, the blots were washed in 2 × SSC, 0.1% SDS at 50 °C before exposure to film.

**Prime Extension and Reverse Transcriptase Polymerase Chain Reaction**—The transcript start site of the human WNT-5A gene was analyzed by primer extension. Total RNA from fetal human fibroblasts (cell line CRL 1262) and human osteosarcoma cells (cell line MG-63) were used for screening. Amplifying "Tri" reagents. Approximately 1 μg of RNA was used as a template for a specific 21-base oligonucleotide (AAACGATGGACGAGGAAATTCG) corresponding to the inverse complement of +2 to +22 of the original WNT-5A cDNA sequence (25). The primer (10 pmol) was end-labeled with γ-32PdATP and T4 polydeoxykinase (Promega) for 10 min at 37 °C. The reaction mix was heated to 90 °C to inactivate the enzyme and passed over a Sephadex G-25 spin column (5 Prime—3 Prime, Inc., Boulder, CO) to remove unincorporated ATP. Annealing of the primer (0.1 pmol) to the RNA was allowed to occur in 10 μl of Superscript (RNase-free) reverse transcriptase (Life Technologies, Inc.) buffer (50 mM Tris-HCl, pH 8.3, 40 mM KCl, 6 mM MgCl2, and 1 mM dithiothreitol) for 20 min at 58 °C followed by slow cooling to room temperature for 10 min. Extension of the primer was done by addition of reverse transcriptase (50 U) to a solution containing 50 μl of Superscript buffer, 0.5 mM each of dATP, dCTP, dGTP, and dTTP, and 200 units of Superscript reverse transcriptase. Incubation was done at 42 °C for 30 min. The reaction mixture was precipitated with 0.3 M sodium acetate (pH 5.2) and 3 volumes of 100% ethanol. After centrifugation the pellet was washed in 70% ethanol, dried at room temperature, and allowed to dissolve in loading buffer (Promega). The samples were heated at 90 °C for 10 min prior to loading on a 6% sequencing gel.

For RT-PCR, 5 μg of total RNA from human embryonic fibroblasts (CRL-1262) were incubated with an antisense primer 5′-TGGACGTCCACCCGTTGATGTCAC-3′ located in exon 3 in the presence of avian myeloblastosis virus reverse transcriptase according to the manufacturer’s recommendation (Invitrogen). For amplifying the DNA, 2 μl of RT reaction was incubated in a 50-μl volume containing 5 μl of 10 × PCR buffer, 1 μl of 100 μM of dNTPs, 1 μM of a sense primer 5′-AGCCCGAAGGTGGTTGAC-3′ hybridizing to the mRNA (Exomed, CA). Amplification was performed using the Perkin-Elmer Model 480 GeneAmp System in 96-well dishes or on line by Light Cycler System (Roche Molecular Biochemicals, Mannheim, Germany). Amplification was performed using a Stratagene Model 2400 thermal cycleer (90 °C for 10 min prior to loading a 6% sequencing gel.

For RT-PCR, 5 μg of total RNA from human embryonic fibroblasts (CRL-1262) were incubated with an antisense primer 5′-TGGACGTCCACCCGTTGATGTCAC-3′ located in exon 3 in the presence of avian myeloblastosis virus reverse transcriptase according to the manufacturer’s recommendation (Invitrogen). A nested deletion construct encompassing the 631 bp upstream of the leader exon was cloned upstream of the CAT gene within the pBS vector, the pBSCT (31). The pBSCT contained a promoterless CAT reporter gene within pBluescript and was generated as described before (31). A nested deletion construct encompassing the first construct of 2 kb of restriction enzyme HindIII. This resulted in an internal deletion of a 750-bp fragment that was subsequently ligated into the corresponding site of the pBSCT (32). Finally, we generated a 5′ deletion construct encompassing the 631 bp upstream of the leader exon. Briefly, the largest 3888-bp construct was partially digested with HindIII, producing three DNA fragments; two of them spanning from −3888 to −1382 and −1382 to −632 were discarded, whereas the remaining fragment, comprising a sequence between −631 and −1 bp relative to the major transcription start site, was isolated, purified, and subjected to self-ligation. All of these constructs were checked by sequencing both ends with either internal primers or primers based on the T3 or T7 promoter sequences.

For transient cell transfection, HEK 293 cells were cultivated in Dulbecco’s modified Eagle’s medium containing 10% fetal bovine serum, 4 × 10^−5 M L-glutamine, 100 units/ml penicillin at 37 °C in a humidified 5% CO2 atmosphere. Transient transfections of HEK 293 cells were performed by the calcium phosphate method as described previously (32). Briefly, cells were transfected with 5 μg of the psCT construct (32) using 20 μl of WNT-5A promoter/pBSCT plasmid along with 10 μg of pcDNAⅢ reporter vector. pcDNAⅢ control vector. The cells were cultured in Dulbecco’s modified Eagle’s medium for 12 h at 37 °C. After an additional 48 h incubation, the cells were rinsed twice with a phosphate-buffered saline solution (pH 7.4) and detached by incubation for 5 minutes with 0.25% trypsin and centrifuged at 100 × g for 5 minutes. The cell pellet was resuspended in 1 ml of Eagle’s minimal essential medium (Gibco BRL) supplemented with 10% fetal bovine serum and 100 units/ml penicillin. Isolated cells were seeded into 24-well plates at a density of 10,000 cells/well. After 24 h, the medium was replaced with fresh medium and the cells were incubated for 24 h. The medium was then harvested and assayed for β-galactosidase activity according to standard procedures (33). CAT assays were performed as described before (31). The acetylated proc...
ucts were separated in preslotted thin layer chromatography plates and subjected to autoradiography. The resultant autoradiograms were quantified by scanning laser densitometry and computer integration. Additional experimental details are provided in the text and legends to figures.

DNA isolation, Southern Blotting, and Regional Chromosomal Mapping—Human WNT-5A Gene—Genomic DNA was isolated as described before (34). Tissue culture cells were centrifuged, and the resulting pellets were mixed with 600 μl of a solution containing 50 mM Tris (pH 8.0), 100 mM EDTA, 100 mM NaCl, 1% SDS, and 585 μg/ml Proteinase K (Boehringer Mannheim) and incubated at 55°C overnight with occasional mixing. The solution was then extracted once with phenol and again with phenol:chloroform (24:1), and the DNA was precipitated with 2 volumes of 100% ethanol. The DNA precipitate was removed with a capillary pipette and dissolved in 10 mM Tris, 0.1 mM EDTA, pH 8.0 (TE). Approximately 10 μg of genomic DNA was digested with HindIII, precipitated, and redissolved in TE, and DNA fragments were separated on a 0.7% Tris borate-EDTA (TBE) agarose gel. The DNA was transferred to nitrocellulose supported by polyester films (30) and hybridized to a WNT-5A cDNA 1.4-kb fragment (443 to 1904) generated by PCR. Probes labeled with [α-35S]dCTP were prepared by random priming using a kit from Stratagene (San Diego, CA). For regional chromosomal mapping, a rodent-human hybrid panel retaining partial chromosome 3 was used as described previously (35). Hybrid DNAs were tested for the presence of specific human WNT-5A restriction fragments by Southern blot hybridization using partial cDNA libraries as described above.

Human Fetal Material, Synthesis of Riboprobes, and In situ Hybridization Analysis—Human embryos were obtained from elective therapeutic terminations of pregnancy. Approval was obtained from the Ethical Committee of the Medical Faculty, the University of Hong Kong for work with fresh embryonic materials and archival specimens. The gestational ages of archival specimens were determined by assessing the centers of ossification and the histological characteristics of the stained sections. The crown-rump length was measured for intact fresh specimens. Five embryos, two at 28 days and one each at 33, 38, and 42 days, were studied. Archival samples of human embryos were previously fixed in 10% (v/v) neutral buffered formalin. Fresh embryos were fixed in 4% paraformaldehyde. Both sources of human embryonic tissue were analyzed by in situ hybridization. Sense and antisense riboprobes for in situ hybridization were generated from a recombinant plasmid containing a 498-bp DraI-SacI fragment of human WNT-5A cDNA (951-1448) subcloned into pBlueScript II KS (Stratagene). Synthesis of [α-35S]UTP-labeled riboprobes, in situ hybridization, autoradiography, and histological staining were carried out as described (36, 37). Single-stranded sense and antisense riboprobes labeled with [α-35S]UTP were generated using T3 and T7 RNA polymerase, respectively. As controls, hybridizations were also performed using sense riboprobes. Sections were counterstained with Harris hematoxylin and eosin and photographs were taken using Kodak Ektachrome ASA 64 film as described previously (36, 37).

RESULTS

Heterogeneity of the 5' End within the WNT-5A cDNA Clones and Determination of the Transcription Start Sites—We previously reported (25) a 483-bp sequence encoding the 5'-untranslated region of WNT-5A cDNA. We now have identified the sequence within two adjacent BamHI genomic fragments. A close examination of the sequence data detected possible donor and acceptor splice sites at positions 33 and 444, which would generate a 411-bp intervening sequence. Further analyses of the cDNA clones revealed that the 411-bp region was spliced out in three different clones and retained in the others (Fig. 1A). To determine whether the 411-bp fragment was intronic, we generated a PCR probe of 363 bp within this region and used it as a labeled probe in Northern blotting analyses of the same fibroblast RNA that was used in the cDNA library construction. The probe failed to hybridize to any transcript (not shown) under the same conditions in which we could detect several transcripts by using other portions of the WNT-5A gene (cf. Fig. 4). Therefore, it appears that the 411-bp fragment is a retained intron cloned from an incompletely spliced nuclear RNA present in the original cDNA library. Fig. 1A shows the position of the intron and the "correct" numbering of the cDNA.
RNA from skin fibroblasts (CRL-1262) and an antisense oligonucleotide between bp 356 and 373. The generation of WNT-5A cDNA was confirmed as control, by specific amplification of a 342-bp fragment between +52 and +373. The products of the RT reaction were hybridized to the 342-bp probe under stringent conditions by Northern blotting analysis. Fig. 1C shows that two bands of ~533 and 740 bp were detected, confirming the presence of two transcription start sites. The sizes of the fragments shown in Fig. 1C corresponded well to those detected by primer extension (Fig. 1B). These data also confirm that the 411-bp segment is an intron inasmuch as the size of the fragments in Fig. 1C would have been 944 and 1151 bp, respectively, if the intron were present in the cDNA. Furthermore, these findings indicate that the WNT-5A messages begin ~160 and 330 bp 5’ to the site predicted by our cDNA sequence (25).

Cloning of the Human WNT-5A Gene—The exon/intron boundaries of the human WNT-5A gene were established by cloning and sequencing various overlapping portions of genomic DNA that hybridized strongly with PCR-generated portions of the human cDNA (25). The complete WNT-5A gene was contained in two overlapping genomic clones (W1 and W10) and spanned about 22 kb of contiguous DNA (Fig. 2A). The gene comprised five distinct exons ranging in size between 134 and ~6800 bp. Both the 5’ and 3’ acceptor splice sites in each of the introns followed the GT-AG consensus sequence for eukaryotic genes (Fig. 2B). The 3’ exon-intron junctions split codons in three different phases, making it unlikely that alternative splicing of the primary transcript would occur because this would alter the reading frame. This is in agreement with the high degree of conservation across species of the primary protein structure in all the members of the WNT gene family (9). Exon 1 (306 bp) encoded most of the 5’-untranslated region, whereas exon 2 (134 bp) encoded the remaining 39 bp of the 5’-untranslated region plus the first 22 amino acids comprising the signal peptide and an additional 9 amino acid residues. As in the case of most mammalian WNT genes (38), the second intron was at an identical position; that is, the two guanosine residues of the codon for the second of the two conserved Trp residues are split by a phase II intron (Fig. 2B). Exons 3 (251 bp) and 4 (293 bp) coded for the next 83 and 98 amino acids respectively, whereas exon 5 (~6800 bp) contained the final 153 codons and an extensive 3’-untranslated region of ~6.5 kb.

The 5’-flanking Region of the Human WNT-5A Gene—A Complex Promoter Active in Transient Cell Transfection Assays—A BamHI genomic fragment of 4.8 kb was found to overlap with the 5’ end of the human WNT-5A cDNA. This fragment was subcloned in pBluescript and sequenced in total with either specific internal primers or by using the T3 and T7 promoter sequences and overlapping subclones. A 2.5-kb sequence is shown in Fig. 3A that includes the first exon and ~2 kb of 5’-flanking region. This DNA sequence contained neither conventional CAAT nor TATA boxes near the transcription start site, a finding that is in agreement with the presence of multiple transcription start sites as shown above. Of note, this region was enriched in GpC dinucleotides with an overall 60% G+ C content and peaks reaching 80% (Fig. 3B). The 5’ end and 5’-flanking region of the human WNT-5A gene were located in a CpG island, a microenvironment in which the number of CpG approximates 0.9, thus indicating the absence of CpG suppression. This region harbored seven copies of the hexanucleotide GGGCGG, the binding site for the zinc finger transcription factor Sp1, and three GC boxes. In addition, a collection of cognate transcription factor-binding sites, summarized in the legend to Fig. 3, were present in the proximal and distal region of the promoter. In summary, the overall organization of the 5’-flanking region strongly suggested that this was the functional promoter of the WNT-5A gene. These GpC-rich promoters are usually found in housekeeping genes, genes encoding growth factors, oncoproteins, or extracellular matrix proteins (39).

Next, we wished to determine whether this putative promoter could exhibit functional activity in transient cell transfection assays. To this end, three major CAT constructs harboring the 3.9-kb promoter region, a nested deletion comprised between ~1382 and ~632, and a 5’ deletion comprised between ~631 and ~1 relative to the major transcription start site were used (Fig. 4). In addition, SV40-driven and promoterless CAT constructs were used, and the values were normalized on β-galactosidase activity determined by co-transfecting SV40-driven β-galactosidase plasmids (32). All of the three constructs exhibited functional promoter activity, with the highest activity residing in the proximal promoter, the region contained within the 631 bp upstream of the transcription start site (Fig. 4C).

Interestingly, both the longest promoter construct and the internal deletion showed an approximate 40% decline in functional promoter activity (Fig. 4C, compare lanes 1 and 2 with
The Human WNT-5A Gene

The Human WNT-5A Gene

Fig. 1. Nucleotide sequence (A) and GC content (B) of the WNT-5A gene promoter. A shows ~2.5 kb of sequence including the 5’-flanking region, a portion of exon 1, and the beginning of intron 1. The major transcription start site, as detected by primer extension (Fig. 1G), is indicated by +1. The exon 1 is in italics and underlined, whereas the GT 5’-splicing junction is boldface. Consensus sequences for the various cognate cis-acting regulatory elements are bold underlined and labeled above the coding strand. The candidate cis-acting factors with minimal mismatching were identified by using the GCG and PCGENE package programs as reviewed before (60): ubiquitous zinc finger transcription factor (Sp1), metal responsive element binding factor (MRF-4), ubiquitous homodimer of J un or heterodimer of J un/F os or J un/AT F family (AP1), developmentally regulated dimer that binds to a palindromic sequence most abundant in neural crest cells (AP2), related to the members of Ets family and most abundant in B lymphocytes (Pu1), inducer of c-myc gene promoter (PuF), CCAAT enhancer binding protein, a regulator of cell growth and differentiation (C/EBP), a primary target of signal transduction induced by epidermal growth factor, phorbol ester and serum (PEA3), a transcription factor that interacts with AP1 in the regulation of SV40 gene expression (AP4), activator of c-fos gene expression and induced by c-sis and platelet derived growth factor (SIF), a transcription factor that binds to human immunodeficiency virus tar element and to the TATA box (UBP1). B shows the distribution of GC content of the 5’ end and upstream region of the human WNT-5A gene and a schematic diagram of the promoter organization. The seven Sp1 binding sites are represented by filled circles, whereas the three GC boxes are indicated by open diamonds.

Fig. 4. Functional activity of the human WNT-5A promoter using transient cell transfection assays. A is a schematic representation of the various constructs used to test promoter activity. The negative numbers indicate the 5’ end of the promoter fragment relative to the major transcription start site (+1). The relative position of the regulatory motifs is also schematically shown. In addition, transfections assays were performed with the SV40-driven CAT vectors. B shows a representative autoradiogram of CAT activity in transiently transfected HeLa cells. The two acetylated species (Ac) are shown for the SV40-CAT (lane 1), promoterless CAT (lane 2), the ~3888 bp construct (lanes 3–5), and the ~1382–~632 internal deletion (lane 6–8). C shows quantitation of CAT activity as determined by scanning densitometry of various autoradiograms. The values are the mean ± S.E. (n = 3) of individual experiments performed in triplicate and are expressed as a percentages of the total activity (arbitrarily set as 100%) of the construct with the highest activity, the ~631 bp promoter construct (lane 3). Lane 1, ~3888 CAT construct; lane 2, ~1382–~632 internal deletion construct; lane 4, promoterless CAT. The values were normalized on β-galactosidase activity that was concurrently transfected in all the assays to correct for transfection efficiency.

This suggests that there are negative elements present in the region upstream of the ~631 bp construct. Nevertheless, the data indicate that the distal promoter region is not required for full transcriptional activity. This is consistent with the fact that the highest concentration of transcription factor-binding elements is within the proximal promoter region (Figs. 3A and 4A). In summary, these data demonstrate for the first time that the 5’-flanking region of the WNT-5A gene has features of a functional promoter that can drive the expression of a reporter gene in HeLa cells.

Characterization of the 3’ End of the WNT-5A Gene and Evidence of Multiple Polyadenylation Sites—Because our Northern blotting studies (26) indicated that the human WNT-5A mRNA should be larger than the cDNA sequence of 4.1 kb (25), we carried out further experiments to clarify the extent of the 3’ end of this gene. The results are shown in Fig. 5. Partial sequencing of several genomic fragments in the untranslated 3’ end of the gene revealed the presence of two AATAAA polyadenylation signals at position 4239 of the cDNA and further downstream near position 7900. These predicted transcript sizes were in reasonable agreement to the size of WNT-5A mRNAs detected by Northern blotting of total human skin fibroblast RNA. To map the 3’ end of the mRNA, we hybridized total human skin fibroblast RNA separately with four different genomic probes spanning different locations.
within the gene (Fig. 5). Hybridization of RNA by Northern blotting with probe A or B revealed the presence of at least four transcripts of approximately 4.4, 5.2, and 6.8 and a poorly resolved doublet of 8.6/9.2 kb, respectively. However, when probe C was used only the 6.8- and 8.6/9.2-kb transcripts were detected. In contrast, probe D did not hybridize with any transcripts of approximately 4.4, 5.2, and 6.8 and a poorly resolved doublet of 8.6/9.2 kb, respectively. However, when probe E was used only the 6.8- and 8.6/9.2-kb transcripts were detected in the mesenchyme surrounding the Wolffian and Müllerian ducts of the 33-day embryo (not shown).}

**FIG. 6.** Order of loci on 3p relative to the WNT-5A locus. The rodent-human hybrids carrying defined regions of 3p are listed across the top of the figure with probe or gene names listed vertically on the left. Presence of a specific probe in the respective hybrid is indicated by a plus sign (+); absence of a probe is indicated by a minus sign (-); a blank space indicates that the probehybrid combination was not tested. The portion of chromosome 3 retained in the 3p panel hybrids divides 3p21.2 to 3p13 into five intervals as indicated in the right-hand column (not all relevant hybrids are shown). Interval A2, defined on the telomeric side by the constitutional break present in hybrid SLF3 (61) (not shown) and on the centromeric side by the constitutional break in hybrid 3.5c13 (35), covers the proximal portion of the large 3p21 band, extending from 3p21.1 through 3p21.2; region B0 is bracketed by the constitutional break of the t(13)5 translocation and the t(3;8) break, represented in hybrids 3.5d3 and 5L8, respectively; interval B1a (p14.2) is delineated on the centromeric side by a de novo break in a hybrid not illustrated here (61), as are regions B1b and B2. Mapping of these other intervals and probes has been reported (61) and is included in the figure to show the WNT-5A relative location.

Translocation of 3p. Specifically, human-rodent and parental DNAs were digested with restriction enzyme HindIII, electrophoresed, and blotted to filters. Filters were then hybridized to the radiolabeled full-length WNT-5A cDNA probe, which was generated by PCR amplification. A strongly hybridizing human fragment of ~10 kb was present in hybrids retaining a specific region of chromosome 3 between 3p14.2 and 3p21.1 as illustrated on the left in Fig. 6 and summarized in Fig. 7. The relationship of the map location of the WNT-5A gene to other regions of interest on 3p is shown on the right in Figs. 6 and 7. The results thus independently confirm the previous mapping to chromosome 3 (25) and further refine the assignment of WNT-5A gene to 3p14.2-3p21.1, between two translocation break points.

**WNT-5A Expression in Early Human Embryos: Implications for Tissue Differentiation**—The expression pattern of WNT-5A mRNAs was studied by in situ hybridization in five human embryos at ages of 28–42 days. WNT-5A mRNAs showed a complex pattern of expression in the developing heart, facial processes, limb, body axis, gonad, and brain (Fig. 8). In all cases, WNT-5A expression in specific sites preceded overt tissue differentiation. In a 28-day human embryo, a low level of WNT-5A transcripts was found in part of the atrium and the ventricle of the heart (Fig. 8A). Expression in the pericardium was transient, positive hybridization being found only in a 33-day embryo (not shown). WNT-5A mRNAs were found to be highly expressed in 28- and 33-day embryos in the frontal nasal mesenchyme (Fig. 8B, C and D) and in the branchial arches (Fig. 8G). Particularly high levels of expression were found in cells that were localized toward the tips of the facial prominence (Fig. 8C).

Expression of WNT-5A transcripts was also found in the mesenchyme surrounding the Wolffian and Müllerian ducts of the 42-day embryo (Fig. 8D). In the developing kidney, WNT-5A transcripts were detected in the mesenchyme sur-
rounding the nephric tubules (Fig. 8, E and H) in the mesonephros of a 38-day embryo but not in a 42-day one (Fig. 8 I).

WNT-5A expression was striking in embryos between 28 and 42 days in the developing central nervous system with a high level of transcripts localized to the tegmentum (floor) of the midbrain (Fig. 8 F). Weaker hybridization signals were in the basal region of the caudal hindbrain and rostral spinal cord (not shown). No WNT-5A transcripts were detected in the forebrain and rostral hindbrain, but strong expression was seen in the hypophysis (Rathke's pouch, Fig. 8 F).

High levels of WNT-5A expression were found in cells populating the peripheral and lateral but not the core of the branchial arches (Fig. 8 G). Particularly high expression was also present in the mesenchyme bordering the pharyngeal clefts and pouches (Fig. 8 G). Along parts of the mid-body axis, expression was also detected within the lateral mesoderm (not shown). Expression of WNT-5A was also seen in the developing gonad. In the indifferent gonad of 38- and 42-day embryos, strong WNT-5A expression was seen in the celomic epithelium and in the precursors of supporting cells destined for follicles or seminiferous tubules (Fig. 8, H–J). Primordial germ cells did not express WNT-5A (Fig. 8 J).

WNT-5A mRNAs also showed preferential localization in the developing limb. In a 33-day embryo, there was a gradient of WNT-5A expression along the proximal-distal axis of the hind limb bud with highest hybridization signals in the progress zone toward the distal tip of the hind limb bud (Fig. 8 K). Expression level of WNT-5A in the developing hind limb increased at 38 and 42 days because most of WNT-5A transcripts were detected in the distal mesenchyme (Fig. 8 L) surrounding the condensing prechondrogenic cells and also at the tips of limb digits (Fig. 8 M). No expression was seen in differentiating chondrocytes (Fig. 8 M) or in the muscle masses. Strong expression was also found in the preaxial and postaxial ectoderm at the rostral and caudal borders of the developing hind limb at the level of joint formation in the 38-day embryo (not shown).

**DISCUSSION**

During the past decade there has been a clear realization that proto-oncogenes, the evolutionarily conserved progenitors of mutant genes that contribute to neoplastic growth, are directly implicated in the control of normal development (40). This view has gained support from a number of experimental observations, including the fact that some proto-oncogenes code...
Fig. 8. Expression of the WNT-5A gene in early human embryos from 28 to 42 days of gestation. In situ hybridizations were performed using sense and antisense riboprobes specific for WNT-5A (see “Experimental Procedures”). No specific hybridization was found using sense riboprobes. A–G, I, and K–M are dark field pictures taken under dark field illumination where silver grains appear white; the background color is due to the counterstain. H and J are double exposures with a combination of bright field with a blue filter and dark field illumination under a red filter to produce superimposed images where silver grains appear pinkish red. A, weak expression is found in the differentiating trabeculae of the ventricle (vt) and parts of the epi-mycardium of the atrium (open arrow) of a 28-day embryo but absent in the pericardium (arrow). B, strong expression in the frontonasal mesenchyme (fm) and weaker expression in the heart (h) are detected in a 33-day embryo. C, hybridization was strong in the frontonasal mesenchyme especially in the region of the tip of the facial prominence of a 38-day embryo (arrows). D, expression is found in the mesenchyme (me) surrounding the Muellerian (mt) and Wolffian (wt) ducts in a 42-day embryo. No hybridization signal was detected in the Muellerian and Wolffian ducts. E, strong expression is found in the mesenchyme (arrow) surrounding the nephric tubules (nt) of the developing kidney of the 38-day embryo. F, restricted expression is localized to the tegmentum of the midbrain (mb) but is absent in the forebrain (fb) and rhombomeres (arrows) of the 28-day embryo. Expression is also seen in the hypophysis (hy). G, transcripts within the branchial arch are concentrated within the peripheral and lateral regions but are absent in the core (co) of a 33-day embryo. There is strong hybridization in the subectodermal mesenchyme in the cleft between the arches (small arrows). The crown-rump axis of the embryo is indicated by the large arrow with the arrowhead pointing toward the crown. H, in the urogenital complex of a 38-day embryo, expression is found in the mesenchyme (m) but is absent in the tubules (t) and glomerulus (g) of the developing kidney. In the indifferent gonad (gn), strong signal is detected in the celiac epithelium (ce), interstitial cells, and somatic cells (arrows) of the indifferent gonad of a 42-day embryo. Note the lack of expression in the developing kidney (k) at this stage. J, higher magnification of I. Expression is seen in the interstitial cells (i) and the somatic cells in the forming sex cords of the indifferent gonad. Primordial germ cells (arrowheads) do not express WNT-5A. The dashed line marks grouping of somatic cells surrounding the primordial germ cells. K, preferential expression is found in the progress zone of the hind limb bud of a 33-day embryo (pz, marked by the dashed lines). L, transcripts are preferentially localized in the distal mesenchyme (dm) of the hind limb bud of a 38-day embryo. No hybridization is seen in the condensing myogenic and skeletogenic mesenchyme. M, strong expression is found in mesenchymal cells of the nail bed (nb) and tendinous connection (te) but absent in the chondrocytes (ca) of the hind limb digit of the 42-day embryo (original magnification: A–C, G–I, and K–M, ×65; D, ×160; E and F, ×140; J, ×260).
for growth factors, that mutant forms can arrest, distort, or promote differentiation, and that they are temporally and spatially sequestered during embryogenesis (9). In this study, we investigated the genomic organization, promoter activity, and chromosomal localization of the WNT-5A gene, which codes for a cysteine-rich growth factor involved in cell-cell signaling during embryonic development and perhaps tumor growth. The genomic organization of the human WNT-5A gene was notable in several aspects. The gene was found to consist of 5 exons spanning ~25 kb of continuous DNA with the last exon extending for ~7 kb and harboring multiple polyadenylation sites. A genomic 5' flanking region of 4.8 kb was localized that overlapped with the corresponding 5' terminal sequence of the cDNA. We discovered a new leader exon interrupted by a 411 bp intron that was retained in our original cDNA cloning (25). Specifically, the sequence of this first exon was found to be markedly different when compared with the corresponding region of the mouse cDNA (13). This divergence was surprising inasmuch as the remaining coding sequence of the mouse cDNA was >90% identical to the human cDNA. In addition, our unpublished data on the genomic organization of the murine WNT-5A gene have indicated that the exonic organization of the translated region is also closely conserved. These findings suggest that the 5' sequences including the control regions of the murine and human WNT-5A genes have diverged during evolution. However, it is interesting that the initial 39 bp of exon 2, which is also untranslated, were highly conserved. It should be pointed out that because exon 1 comprises most of the untranslated region of the WNT-5A mRNA, the overall characteristics of the WNT-5A protein sequence would not be affected by these differences in DNA sequences.

We have identified a 5' flanking region in the human WNT-5A locus that harbored several features that have been correlated with transcriptional control regions and are typically observed in genes that encode oncoproteins, growth factors, transcription factors, and housekeeping genes (39). This flanking sequence was located in a GpC-rich island and harbored numerous cis-acting elements, including S1, AP1, and AP2 binding motifs, and acted as a functional promoter when tested in transient cell transfection assays. Of note, the strongest promoter activity resided in the proximal promoter, the region contained within the 631 bp upstream of the major transcription start site. The fact that both an internal deletion and the long 3.9 kb promoter showed a 40% decline in functional activity suggests that the distal promoter region is not required for full transcriptional activity. This is also supported by the fact that the highest concentration of elements binding to transcription factors is within the proximal promoter region. Collectively, our findings demonstrate for the first time that the 5' flanking region of the WNT-5A gene has features of a functional promoter that can drive the expression of a reporter gene in HeLa cells.

Using a combination of primer extension and RT-PCR, we detected two major transcription start sites for the human WNT-5A gene. Of note, the murine WNT-1 gene is also transcribed from two major transcription start sites (41). The WNT-1 sequence near the more 3' start site is reminiscent of a classical promoter with a well defined TATA box. In contrast, the sequence around the more 5' start site does not contain a canonical TATA box but is preceded by a GpC-rich stretch of DNA (41) similar to the human WNT-5A. Of note, a WNT-1 inducing factor-1 has been recently shown to represent a novel transcription factor that binds a G/C box motif in WNT-1 and regulates the expression of WNT-1 during neuroectodermal differentiation (42). Interestingly, the human WNT-1 gene has an entirely homologous WNT-1 inducing factor-1 binding site at the same relative position (43). Although the human WNT-5A promoter does not possess an identical sequence, it harbors in the proximal promoter several stretches of DNA that are homologous to the WNT-1 inducing factor-1 binding site.

Unusual results were also found at the 3'-untranslated region of the WNT-5A gene. The terminal 459 bp coding for 152 amino acids plus the stop codon and the following 6.3 kb were found to comprise a single exon. Two polyadenylation signals were found at -4.4 and 7.9 kb, accounting for at least some of the observed transcript sizes. However, the appearance of transcripts detected by Northern hybridization that were larger than 8 kb could not be satisfactorily explained by our model of the gene structure and may include alternatively spliced exons in the 5'-untranslated region.

The human WNT-5A gene falls between two constitutional chromosome breaks in 3p21.1 and 3p14.2, respectively; thus the WNT-5A gene is perhaps closely telomeric to the 3p14.2 constitutional t(3:8) chromosomal translocation associated with familial renal cell carcinoma (44). Because the murine WNT-5A locus maps to the centromeric region of mouse chromosome 14 (9, 45), as do the murine protein tyrosine phosphatase γ Ptprg gene (46, 47) and the calcium channel α1 subunit CCh1α2 gene (48, 49), these loci thus form a small linkage group in mouse and human, whereas most 3p21 markers map to murine chromosome 9. As discussed above (47), abnormalities on murine chromosomal 14 in metastatic murine melanomas (50), suggest that the integrity of the Ptprg locus should be explored in these tumors. Similarly, it would be very interesting to investigate the integrity of the murine WNT-5A locus in such tumors. Interestingly, human chromosome 3 is composed of about 210 million bp, representing about 7% of the human genome (51). Thus, it might be anticipated that the genes responsible for perhaps 7% of human genetic disorders would be located here. To date, however, only a few disease loci have been mapped to the region of chromosome 3 where WNT-5A maps. It is of note that this region of chromosome 3 is often the target of rearrangements and translocation in several malignant neoplasms, most notably in small cell lung carcinomas (52) and, as pointed above, in renal cell carcinomas (44, 53).

The expression of WNT-5A in early human embryos was remarkably similar to that reported in the developing mouse (13, 54) and chicken (55) with the primary regions of expression of WNT-5A in the midbrain and the developing limb. The murine WNT-5A is the only known member of the WNT gene family that is detected in the developing limb (54) with a pattern of distribution following a proximal-distal gradient along the limb mesenchyme and the ventral half of the limb ectoderm (13). In these examples expression of WNT-5A appears in sites prior to overt tissue differentiation and at times at which important events are occurring in pattern formation. The similarities in expression patterns of the WNT-5A gene in human, mouse, and chicken embryos point to a role for WNT-5A in the patterning and differentiation of the midbrain and limb and reinforces the applicability and importance of using animal models to study the function of this important class of signaling molecules. Of particular interest is the hitherto unreported expression of WNT-5A expression in the developing kidneys and gonads. Commitment of the indifferent gonad to male or female differentiation is likely to be influenced by cells in the supporting cell lineage (56). The SRY gene is involved in initiating that process, but other factors must be involved in the differentiation of germ cells and the gonad. It has also recently been shown that the genital ridge environment is important for providing signals for the initiation of X chromosome reactivation in primordial germ cells (57). Local-
WNT-5A involved in embryonal cell migration (19). For the nervous system and limb development. Its presence in the nogenesis with a distribution that is not limited to the central and pathological states. Together with the mapping data and the expression pattern of WNT-5A in X. laevis with a distribution that is not limited to the central nervous system. The growth factor is involved in the genesis and development of human cancer. This hypothesis is supported by a recent investigation that has shown that both benign and malignant breast tumors exhibit high levels of WNT-5A transcripts (58). In human mammary epithelial cells, the WNT-5A gene is up-regulated by confluence and down-regulated by hepatocyte growth factor (HGF) during the HGF-induced branching in collagen gels (59). These findings suggest that the mammalian WNT-5A gene is involved in mammary epithelial cell motility and are in agreement with the finding in X. laevis where WNT-5A is involved in embryonal cell migration (19).

In summary, our studies have characterized the human WNT-5A gene and shown that the WNT-5A gene has a specific and restricted pattern of expression during early human organogenesis with a distribution that is not limited to the central nervous system and limb development. Its presence in the indifferent gonads and kidneys suggests a novel potential role for the WNT-5A gene in renal and germ cell differentiation. The elucidation of the WNT-5A gene regulatory control regions, together with the mapping data and the expression pattern during development, will make possible a better understanding of the function and regulation of this gene in normal and pathological states.

Acknowledgments—We thank I. Eichstetter for excellent technical assistance, T. Druck, K. W. Gripp, and B. Schlegeler for help with the chromosomal mapping and preparation of the 3,5c13 hybrid, and Patrick Tam and Wai-Sum O for helpful discussion.

REFERENCES
1. McMahon, A. P. (1992) Dev. Biol. 1, 31–60.
2. Nusse, R. (1988) Trends Genet. 4, 291–295.
3. Brown, A. M. C., Wildin, R. S., Prendergast, T. J., and Varmus, H. E. (1986) Mol. Cell. Biol. 10, 1470–1479.
4. Slegfried, E., Wilder, E. L., and Perrimon, N. (1994) Nature 370, 519–523.
5. Brassil, M., Bradley, A., and Hardcastle, T. J. (1993) J. Cell Biol. 123, 247–261.
6. Wary, K., Lou, S., Buchberg, A., Siracusa, L., Druck, T., Iafuoria, S., and Huebner, K. (1993) Cancer Res. 53, 1498–1502.
7. Chin, H., Koizaki, C. A., Kim, H.-L., Mock, B., and McBride, O. W. (1991) Trends Genet. 7, 114–119.
8. Dealy, C. N., Roth, A., Ferrari, D., Brown, A. M. C., and Wainwright, B. (1994) Cancer Res. 54, 219–222.
9. Brauch, H., Zhang, W., and Zakian, V. A. (1992) Proc. Natl. Acad. Sci. U.S.A. 89, 8669–8673.
10. Brauch, H., Zhang, W., and Zakian, V. A. (1992) Proc. Natl. Acad. Sci. U.S.A. 89, 8669–8673.
Characterization of the Complete Genomic Structure of the Human WNT-5A Gene, Functional Analysis of its Promoter, Chromosomal Mapping, and Expression in Early Human Embryogenesis
Keith G. Danielson, Jhuma Pillarisetti, Isabelle R. Cohen, Bijan Sholehvar, Kay Huebner, Ling-Jim Ng, John M. Nicholls, Kathryn S.E. Cheah and Renato V. Iozzo

J. Biol. Chem. 1995, 270:31225-31234.
doi: 10.1074/jbc.270.52.31225

Access the most updated version of this article at http://www.jbc.org/content/270/52/31225

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 61 references, 27 of which can be accessed free at http://www.jbc.org/content/270/52/31225.full.html#ref-list-1