INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Cloning and characterization of the cDNA and genomic DNA coding for human class I heparin-binding growth factor

Wang, Wen-Pin, Ph.D.

The Ohio State University, 1991
CLONING AND CHARACTERIZATION
OF THE CDNA AND GENOMIC DNA
CODING FOR HUMAN CLASS I HEPARIN-BINDING GROWTH FACTOR

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Wen-Pin Wang, B.S.

The Ohio State University

1991

Dissertation Committee:
Dr. Ing-Ming Chiu
Dr. Clark L. Anderson
Dr. Paul A. Fuerst
Dr. Lee F. Johnson

Approved by:

Advisor

The Molecular, Cellular and Developmental Biology Program
To My Parents, My wife and My Sons
ACKNOWLEDGEMENTS

I thank Dr. Ing-Ming Chiu for academic guidance and support throughout the research. I thank Dr. Clark L. Anderson for helpful discussion, and even running the sequencing gel for me. I would like to thank Dr. Lee F. Johnson for his valuable advice and suggestions. Thank you, Dr. Paul A. Fuerst for consistent comments and help, especially during the early stages of my graduate study. Special thanks to members of Dr. Chiu's laboratory: Dr. René Myers, Kirsten Lehtoma, Matt Poulin, Karen Waddell, Bob Payson, Halit Canatan, Kevin Patrie, Laurine Connors and Dr. Dawei Li for much needed friendship. I would like to express appreciation to Dr. Lai Chu Wu, Dr. Kevin Hackshaw, Dr. Sissy M. Jhiang, Dr. Long-Sheng Chang, Dr. Donald E. Thornton, Dr. Dan Caruso and Linda Ernst. I also thank Julia Adams Wilson and Barbara Mazzotta for their secretarial help on the manuscripts and my undergraduate classmate, Ju Chou, for helping with some of the computer-generated art work. Finally, sincere thanks are extended to my parents, Chang-Fang and Lee-chyan, and family. I hope you will become better. To my wife, Bih-Fen, and my sons, Alan and Gene, I offer my deepest thanks and love.
VITA

Jan. 23, 1959
Born in Taiwan, Republic of China

1977-1981
B.S., Department of Biology, Tunghai University, Taiwan

1981-1983
Lieutenant (ROTC), Taiwan

1983-1984
Research Associate, Department of Microbiology, National Yung-Ming Medical College, Taiwan

1984-1985
Graduate Student, Institute of Neurosciences, National Yung-Ming Medical College, Taiwan

1985-1986
Laboratory Technician, Department of Pediatric Cardiology, Veterans General Hospital, Taiwan

1986-1987
Graduate Teaching Assistant, Molecular Genetics
The Ohio State University, Columbus, Ohio

1987-present
Graduate Research Assistant, Molecular Genetics
The Ohio State University, Columbus, Ohio

PUBLICATIONS


FIELDS OF STUDY

Major Field: Molecular, Cellular and Developmental Biology

Studies in Molecular Biology of Growth Factors
# TABLE OF CONTENTS

**DEDICATIONS** ............................................. ii  
**ACKNOWLEDGEMENTS** ....................................... iii  
**VITA** ................................................. iv  
**LIST OF TABLES** ......................................... xi  
**LIST OF FIGURES** ......................................... xii  
**ABBREVIATIONS** ........................................... xv  

**INTRODUCTION** ........................................... 1  
  Polypeptide Growth Factor ............................. 1  
  Nomenclature, Purification and Characterization of HBGF-1 2  
  Primary Structure of HBGF-1 .......................... 4  
  Biological Properties of HBGF-1 ...................... 5  
  The HBGF-1 Gene and Chromosome Abnormalities ....... 12  
  Structure - Function Relationships of HBGF-1 ....... 13  
  Oncogenic Members of the HBGF Family .............. 15  
    A. **int-2** ......................................... 15  
    B. **hst/KS3** ....................................... 21  
    C. **FGF-5** ......................................... 22  
  Other HBGF-Related Genes ............................. 23  
    A. **FGF-6** ......................................... 23  
    B. **KGF** ........................................... 24  
  Transforming Potential of HBGF-1 and HBGF-2 ....... 28
The HBGF Receptors ................................... 31

A. cek1/fg/NE2 ................................... 33
B. hek/K-sam/cek3 .................................. 34
C. cef ........................................... 35

The HBGF Secretion Mechanisms .......................... 36

A. Intracrine Interaction ................................ 37
B. Cell Leakage ................................... 37
C. Matrix Interaction .............................. 38
D. Alternative Splicing .............................. 39
E. Other Mechanisms ................................ 39

Aims of Studies ..................................... 40

MATERIALS AND METHODS .............................. 42

Materials, Tissues and Cell Lines .................... 42
RNA Isolation and Oligo(dT)-Cellulose Chromatography ..... 43
Northern (RNA) Blot Analysis ......................... 45
Ribonuclease (RNase) Protection Assay ................. 46
cDNA Library Screening ................................ 48
Preparation of Competent Cells ....................... 49
Transformation ....................................... 49
Colony Hybridization/Microwave Colony Screening ....... 50
Isolation of Plasmid, Phage and High Molecular Weight Genomic DNA .................. 51
Isolation of DNA Fragments ............................ 53
Restriction Enzyme Mapping and Gel Electrophoresis .... 55
Cloning of DNA Fragments Into Vectors/Direct Cloning of cDNA Inserts ................ 55

vii
Preparation of Radiolabeled DNA Probes ........................................ 56
Southern (DNA) Blotting and Hybridization/Genomic Blots and Zoo Blots ........................................ 57
Genomic Library Screening ...................................................... 58
Purification, Desalting and Radiolabeling of Oligonucleotides .............. 59
Isolation of Single-Stranded (ss) DNA by Superinfection ..................... 61
Manual DNA Sequencing and Sequence Comparison .................................. 61
Unidirectional Deletion Cloning .................................................. 62
Automatic DNA Sequencer and Cycle Sequencing .................................. 63
Genomic DNA Polymerase Chain Reaction (PCR) .................................. 64
Reverse Transcription and PCR (RT-PCR) ......................................... 65
Anchored PCR or Rapid Amplification of cDNA Ends (RACE) ...................... 66
Direct PCR Screening of Colonies and Plaques ..................................... 66
Cloning PCR Products ............................................................... 67

RESULTS .......................................................................................... 68

SECTION I. MOLECULAR CLONING AND CHARACTERIZATION OF HUMAN HBGF-1 cDNA ................................................. 68
  1.1. Screening Strategy .................................................................. 68
  1.2. cDNA Clones and Restriction Enzyme Mapping ......................... 68
  1.3. cDNA Sequence Analysis ....................................................... 70
  1.4. Alternative Splicing Generates Two Forms of HBGF-1 pHBGF-1.4 and pHBGF-1.5 ......................... 74
  1.5. HBGF-1 mRNA Expression in Human Brain, Kidney and Heart ...........
  1.7. Cloning of cDNA Clones Containing Upstream Exon Sequences .......... 91

viii
### SECTION II. ISOLATION AND CHARACTERIZATION OF HUMAN HBGF-1 GENE

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. Screening Strategy</td>
<td>99</td>
</tr>
<tr>
<td>2.2. Genomic DNA Clones and Exon-Intron Organization</td>
<td>99</td>
</tr>
<tr>
<td>2.3. Genomic DNA Sequence Analysis</td>
<td>107</td>
</tr>
<tr>
<td>2.4. Mapping the 5'-end of First Protein Coding Exon</td>
<td>112</td>
</tr>
<tr>
<td>2.5. Mapping the 3'-end of the Gene</td>
<td>114</td>
</tr>
<tr>
<td>2.6. Sequence Conservation Among Different Species</td>
<td>122</td>
</tr>
<tr>
<td>2.7. Conservation of Intervening Sequences Among Different Species</td>
<td>125</td>
</tr>
<tr>
<td>2.8. Isolating the Genomic Clones Containing Upstream Untranslated Exons</td>
<td>125</td>
</tr>
</tbody>
</table>

### SECTION III. INVESTIGATION OF HUMAN HBGF-1 LOCUS IN ACUTE NON-LYMPHOCYTIC LEUKEMIA PATIENTS

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1. HBGF-1 and 5q- Chromosome</td>
<td>130</td>
</tr>
<tr>
<td>3.2. Genomic Southern Blots Using Different HBGF-1 Probes</td>
<td>130</td>
</tr>
<tr>
<td>3.3. RT-PCR of HBGF-1 Expression in Leukemia Patients</td>
<td>137</td>
</tr>
</tbody>
</table>

### DISCUSSION

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOLECULAR CLONING AND CHARACTERIZATION OF HUMAN HBGF-1 cDNA</td>
<td>140</td>
</tr>
<tr>
<td>Alternative Upstream Untranslated Exons</td>
<td>140</td>
</tr>
<tr>
<td>Ectopic Expression of HBGF-1 cDNA</td>
<td>143</td>
</tr>
<tr>
<td>HBGF-1 mRNA Expression in Various Tissues</td>
<td>143</td>
</tr>
<tr>
<td>ISOLATION AND CHARACTERIZATION OF HUMAN HBGF-1 GENE</td>
<td>149</td>
</tr>
<tr>
<td>Structure Organization of HBGF-1 Gene</td>
<td>149</td>
</tr>
<tr>
<td>Mapping of the First Protein-Coding Exon</td>
<td>149</td>
</tr>
<tr>
<td>Alternative Polyadenylation sites of HBGF-1 Gene</td>
<td>151</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1. Sequence similarity among the members of HBGF family . 18

Table 2. Nucleotide and amino acid sequence similarity of HBGF-1 among different species ............... 27

Table 3. Karyotype analysis of nine leukemia patients .......... 132
LIST OF FIGURES

Figure 1. Mitogenic effect of bovine HBGF-1 on mouse Swiss/3T3 cells ........................................ 6

Figure 2. Dosage effect of bovine HBGF-1 on Swiss/3T3 cells .................................................... 7

Figure 3. Homology of protein sequence among members of human HBGF family ............................. 16

Figure 4. Homology of HBGF-1 protein sequence among different species ................................. 26

Figure 5. Restriction enzyme map of human HBGF-1 cDNA clones .......................................... 69

Figure 6. Sequencing strategies for the determination of entire sequence of cDNA clones .............. 71

Figure 7. Nucleotide and deduced amino acid sequences of human HBGF-1 cDNA clones ............... 72

Figure 8. Northern blotting and hybridization analysis of human fetal RNA with HBGF-1 and β-actin cDNA probes ................................................................. 77

Figure 9. Amplification and nucleotide sequence of upstream un-translated exon of pHBGFl.J of human HBGF-1, l.JA\text{PstI} ................................................. 79

Figure 10. Applications of PCR cDNA clone l.JA\text{PstI} ..................................................... 81

Figure 11. Reverse transcription-polymerase chain reaction of l.J\text{Pvu} ........................................... 85

Figure 12. Amplification and nucleotide sequence of upstream untranslated exon of pHBGFl.H of human HBGF-1, l.H\Delta ......................................................... 87

Figure 13. RT-PCR reaction of l.H\text{Pvu} .................................................................................. 90

Figure 14. RT-PCR reaction of various upstream exons and first protein-coding exon of human HBGF-1 in different tissues ......................................................... 92
Figure 34. Leukemia patient DNA probed with the 2.1-kbp HindIII fragment and the 1.4-kbp fragment of the HBGF-1 gene ........................................ 135

Figure 35. Leukemia patient genomic DNA probed with the 2.1-kbp HindIII fragment of the HBGF-1 gene and the β-globin gene ........................................ 136

Figure 36. RT-PCR of RNAs isolated from different tissues and cell lines ........................................ 138

Figure 37. Summary of the splicing patterns and tissue distributions of the three upstream exons of HBGF-1 ........................................ 144

Figure 38. Summary of the genomic organization of HBGF-1, including the three upstream untranslated exons ........................................ 157
LIST OF ABBREVIATIONS

bek  bacterially expressed kinase
bp   base pair
cDNA complementary DNA
cek  chick embryo kinase
Cl   Curie
CIA  chloroform:isoamyl alcohol (24:1)
cpm  counts per minute
dATP 2'-deoxyadenosine 5'-triphosphate
dNTP deoxynucleoside triphosphate
ds DNA double-stranded DNA
EDTA ethylene diamine tetraacetic acid
FGF  fibroblast growth factor
fgf  fms-like gene
GuTC guanidinium thiocyanate
HBGF heparin-binding growth factor
K-FGF Koposi sarcoma-derived FGF
kbp  kilobase pair
KGF  keratinocyte growth factor
LB   Luria-Bertani medium
mRNA messenger RNA
nt. nucleotide
O.D. optical density
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcription and PCR</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SSC</td>
<td>salt-sodium citrate buffer</td>
</tr>
<tr>
<td>ss DNA</td>
<td>single-stranded DNA</td>
</tr>
<tr>
<td>SSPE</td>
<td>salt-sodium phosphate EDTA buffer</td>
</tr>
<tr>
<td>TAE</td>
<td>tris-acetate electrophoresis buffer</td>
</tr>
<tr>
<td>TBE</td>
<td>tris-borate electrophoresis buffer</td>
</tr>
<tr>
<td>Tris</td>
<td>tris(hydroxymethyl)aminomethane</td>
</tr>
</tbody>
</table>
INTRODUCTION

Polypeptide Growth Factors

Polypeptide growth factors are hormone-like modulator of a number of cellular functions, including cell proliferation and differentiation in vitro and in vivo. However, unlike hormones, they are produced diffusely from a wide variety of cell and tissue types. A convenient classification is based on the method of transport to its target. Endocrine substances are released from a specialized cell in discrete glands and are transported to a distant set of target cells by means of the blood stream. Paracrine agents affect only the group of target cells bearing receptors adjacent to the cell or origin of the trophic agents. A third class, autocrine interaction implies cells respond to substances that they themselves produce.

The first growth factor discovered after continual searching for the biological response modifier for mammalian cell growth is nerve growth factor (NGF) (Levi-Montalini, 1987). Growth of the neurons of the sympathetic nervous system requires NGF. The discovery of epidermal growth factor (EGF) by Stanley Cohen closely followed, which is a protein required for the growth of epithelial cells and 3T3 cells in culture (Gill et al., 1987). Platelet-derived growth factor (PDGF) and insulin-like growth factor-I and -II (IGF-I and IGF-II) have also been described. PDGF is a mitogen for cultured cells of mesenchymal origin. In vivo, PDGF may
promote wound healing (Ross et al., 1986; Deuel, 1987). IGF-I (also named as Somatomedin C) mediates the action of the pituitary growth hormone (also named as Somatotropin; GH). The role of IGF-II in vivo is uncertain (Foresch et al., 1985). Transforming growth factors (TGF) permit 3T3 cells to grow without attachment. TGF-α is structurally related to EGF and binds to the same receptors. TGF-β is a distinct growth factor which can enhance the growth of fibroblasts in culture, but inhibits the growth of some epithelial cells in culture (Keski-Oja and Moses, 1987; Massague, 1987; Roberts and Sporn, 1988). Hematopoietic growth factors include erythropoietin (Epo), colony stimulating factors (CSF: included G-CSF, M-CSF, GM-CSF and Multi-CSF), interleukin 2 (IL2) (Webb et al., 1985). Other factors include γ-interferon which causes virally infected cells to stop cycling and makes them refractory to mitogen stimuli, tumor necrosis factor (TNF; also named as cachectin) which is produced by stimulated macrophages, and also lymphotoxin which is produced by activated lymphocytes and has antitumor activity (Pennica et al., 1984; Beutler et al., 1985).

Nomenclature. Purification and Characterization of HBGF-1

Heparin-binding growth factor (HBGF) is one of the polypeptide growth factors originally identified as an activity in extracts of pituitary and brain that stimulated the growth of 3T3 cells (Armelin, 1973; Gospodarowicz, 1974). Prior to the purification and characterization of HBGF, a number of laboratories had identified polypeptide mitogens from different sources and assayed for their responsive target cells. These mitogens included tumor angiogenesis factor (TAF), fibroblast growth
factor (FGF), endothelial cell growth factor (ECGF), astroglial growth factor (AGF), eye-derived growth factor (EDGF), retina-derived growth factor (RDGF), hypothalamus-derived growth factor (HGF), brain-derived growth factor (BDGF), cartilage-derived growth factor (CDGF), chondrosarcoma-derived growth factor and hepatoma-derived growth factor (HDGF) (review in Baird et al., 1986a; Gospodarowicz et al., 1987a; Burgess and Maciag, 1989; Chiu et al., 1990b).

Most of the above mitogens appear to have very similar biochemical properties. Further comparison turned out to be virtually impossible without further purification and characterization. In 1983, Thorton and his colleagues demonstrated that sulfated glycosaminoglycan, heparin, could potentiate the biological activity of crude preparations of acidic FGF (Thorton et al., 1983). In 1984, Shing et al. were the first to describe the interaction between heparin and chondrosarcoma-derived growth factor (Shing et al., 1984) and this property greatly simplified the purification of these mitogens (Gospodarowicz et al., 1984; Maciag et al., 1984; Shing et al., 1984). In a short period of time, most of the endothelial cell mitogen were purified to homogeneity by using heparin-sepharose affinity chromatography. According to the different binding affinity and other criteria, Lobb and his co-worker suggested these mitogens could be separated into two classes: The class 1 heparin binding growth factor (HBGF-1) describing the acidic polypeptide and the class 2 heparin binding growth factor (HBGF-2) describing the basic polypeptide (Lobb et al., 1986a; Lobb et al., 1986c). However, in recent literature more people used acidic fibroblast growth factor (aFGF) for HBGF-1 and basic fibroblast growth factor (bFGF) for HBGF-2.
HBGF-1 is a single chain polypeptide with a molecular weight of 16,000-18,000 Dalton (Da) (Burgess et al., 1985; Burgess et al., 1986), a pI of 5.0, whereas HBGF-2 has similar molecular weight with a pI of 9.6. HBGF-1 has been isolated from neural tissues such as brain, hypothalamus and retina (Lobb et al., 1986a; Lobb et al., 1986c) while HBGF-2 have been isolated from a variety of cell lines and tissues (Gospodarowicz, 1987). Only recently, HBGF-1 was also found to be present in kidney (Gautschi-Sova et al., 1987), smooth muscle cells (Winkles et al., 1987), a mesoblastic nephroma (Witte et al., 1989) several glioblastoma (Libermann et al., 1987) and tumor cell lines (Lobb et al., 1986b).

Primary Structure of HBGF-1

HBGF-1 has been identified in three forms: (1) 154 a.a. form has a deletion of the NH₂-terminal methionine of the precursor protein (Burgess et al., 1986; Crabb et al., 1986), which gives the 18 kilodalton (kDa) form; (2) 140 a.a. form results from proteolytic cleavage of intact HBGF-1 between lysine-14: phenylalanine 15; and (3) 134 a.a. form results from cleavage between glycine-20: asparagine-21. The latter two truncated forms give the low molecular weight mixtures of 16 kDa (Thomas et al., 1985). Whether three different forms of HBGF-1 occur in vivo or are created artificially during the purification processes is not known. In the case of HBGF-2, the protein was cleaved at different sites by tissue-specific proteases found in brain or tumor cells (Klagsbrun et al., 1987). High molecular weight HBGF-2 was shown to result from translation initiated at idiosyncratic CUG codons (Florkiewicz and Sommer, 1989; Prats et al., 1989). Interestingly, subcellular fate of different HBGF-2
proteins were determined by the choice of different initiation. The lowest molecular weight form of HBGF-2, which is initiated at AUG start codon, is localized in cytoplasm whereas the other two higher molecular weight forms, which are initiated at upstream CUG codons, were found primarily in nucleus (Renko et al., 1990).

Analysis of the amino acid sequences of HBGF-1 and HBGF-2 demonstrate a 55% identity (Thomas, 1987). This sequence similarity together with similar biological activities indicate that the genes may have a common ancestor. A weak but definitive similarity between HBGF and interleukin 1 (IL-1) was also detected (Thomas and Gimenez-Gallego, 1986). This is further supported by the crystallographic studies of the three-dimensional structures of HBGF-1 and HBGF-2 which showed that they resemble the topological folding pattern of IL-1α and IL-1β (Zhu et al., 1991).

Biological Properties of HBGF-1

I. Induction of DNA Synthesis and Cell Division in vitro

HBGF-1 and HBGF-2 are potent inducers of DNA synthesis in a number of normal mammalian cell types of mesodermal or neuroectodermal origin (Gospodarowicz et al., 1986; Thomas and Gimenez-Gallego, 1986). The mitogenic effect was usually measured by [³H]thymidine incorporation into DNA. Fig. 1 shows that mouse Swiss/3T3 fibroblast cells responded to the stimulation of bovine HBGF-1 at 5 ng/ml and maximum effect was detected 22-24 hours after adding the mitogen. The mitogenic activity of HBGF-1 can be potentiated by heparin while that of HBGF-2 is not augmented by
Figure 1. Mitogenic effect of bovine HBGF-1 on mouse Swiss/3T3 cells. Swiss 3T3 cells were grown to confluency and chances to Dulbecco’s modified Eagle medium containing 0.5% calf serum to starve for two days. HBGF-1 (5 ng/ml) was then added and each at specific time thereafter, [\(^3\)H]-thymidine was added, cells were incubated in the presence of radioisotope for two hr. The amount of radioactivity incorporated into DNA as TCA-precipitable materials was measured as the mitogenic activity (Chiu et al., 1990b).
Figure 2. Dosage effect of bovine HBGF-1 on Swiss/3T3 cells. Swiss/3T3 cells were starved in Dulbecco's modified Eagle medium containing 0.5% calf serum for two days. The cells were pulsed with $[^3]$H-thymidine from 18-24 hr after the addition of the specified amount of bovine HBGF-1 and DNA synthesis rate was determined. Alternatively, 5 units/ml of heparin in addition to the HBGF-1 were added to the media (Chiu et al., 1990b).
heparin. At a concentration of 1 ng/ml of HBGF-1, heparin can increase the mitogenic activity of HBGF-1 by more than ten fold (Fig. 2).

The potentiation of HBGF-1 biological activity by heparin is multifaceted and may have at least three interrelated mechanisms (review in Maciag, 1990). First, HBGF-1: heparin complex has altered the tertiary structure which presents a more astute ligand for receptor binding. Second, heparin can prevent covalent modification and inactivation of HBGF-1 by cellular proteolytic enzymes. Lastly, heparin protects HBGF-1 against thermal and chemical denaturation.

II. Regulation of Protein Synthesis and Secretion

Analytical two-dimensional polyacrylamide gel analysis of BALB/c 3T3 cells demonstrates that HBGF-1 and HBGF-2 mostly induce identical changes in intracellular polypeptide expression (Rybak et al., 1988). In cardiac muscle, HBGF-1 and HBGF-2 regulate at least five genes in common, including α and β myosin heavy chains, atrial natriuretic factor and the sarcoplasmic reticulum calcium ATPase; however, they differentially control expression of striated α-actins (Parker et al., 1990). Removal of HBGF-1 from the culture media results in an approximately five-fold increase in plasminogen activator inhibitor (PAI-1) mRNA and protein secretion by human umbilical vein endothelial cell (HUVEC) (Konkle et al., 1990). A marked diminution in the production of prostacyclin (PGI2) by HUVEC in the presence of HBGF-1 and heparin has been reported (Weksler, 1990). The decrease in PGI2 production correlates with reduction in the cellular levels of prostaglandin H synthase and protacyclin synthase. Addition of HBGF-1 to quiescent HUVEC increases the level of PDGF A-chain
mRNA, but not PDGF B-chain mRNA. Nuclear run-on experiments indicate that the transcription rate of PDGF A-chain gene was transiently increased after HBGF-1 addition. HBGF-1 stimulated the synthesis and secretion of PDGF (Gay and Winkles, 1990). HBGFs also stimulate thyrotrophin and prolactin secretion in rat pituitary cells (Baird et al., 1985).

III. Induction of Cell Mobility and Migration in vitro

HBGF-1 has been shown to be chemotactic for endothelial cells, fibroblasts, and astroglial cells (Terranova et al., 1985; Senior et al., 1986). High concentrations of HBGF-1 inhibit thymidine incorporation into 3T3 cells (see Fig. 2; Ke et al., 1990). However, it stimulates chemokinesis and chemotaxis of these cells. These indicate that the chemotactic and mitogenic activities of HBGF-1 may be distinct. It is not known whether the chemostatic activities are mediated through the same receptor systems involved in the mitogenic activities.

IV. Modulation of Differentiation and Development

HBGF-1 stimulate PC12 cells, a rat pheochromocytoma-derived cell line, to undergo neuronal differentiation including the promotion of neurite extension (Togari et al., 1985; Neufeld et al., 1987). Induction of neurite outgrowth in PC12 cells is augmented by heparin (Wagner and D'Amore, 1986). HBGFs also promote the survival and differentiation of a variety of cells that derived from the neural crest, including hippocampal neurons and some neurons from frontal cortex, parietal cortex, occipital cortex, (Morrison et al., 1986; Neufeld et al., 1987). The processes outgrowth of rat retinal ganglion cells was found to be promoted to a far
greater extent by HBGF-1 than by HBGF-2 (Lipton et al., 1988). Thus, next to nerve growth factor, HBGFs have become the best-characterized neurotrophic polypeptide growth factors for neuronal differentiation (Walicke and Baird, 1988).

Biological actions of HBGFs also include repression of skeletal muscle differentiation. In BC3H-1, a mouse myogenic cell line, HBGFs inhibit myogenin gene transcription, which is an important regulator of skeletal muscle cell differentiation (Clegg et al., 1987; Brunetti and Goldfine, 1990).

One of the important activities of HBGFs seems to concern inducer during early embryonic development (Reviewed in Whitman and Melton, 1989). Members of the HBGF family, including HBGF-1 and HBGF-2, are potent inducers of mesoderm formation in vitro (Kimelman and Kirschner, 1987; Slack et al., 1987). Godsave et al. (1988) have argued that mesoderm inducing factors fall into two distinct classes: HBGF and factors similar to Xenopus tissue culture mesoderm-inducing factor (XTC-MIF), which is a TGF-β-related protein also known as activin (Smith et al., 1990). HBGF mRNA and protein are present in both oocytes and embryos of Xenopus (Kimelman et al., 1988; Slack and Isaacs, 1989). It has been speculated that different members of the HBGF family may have different functions during early embryonic development.

V. Angiogenic Activity In Vivo

HBGFs are the most potent angiogenic factors discovered so far and they have potential for enhancement of tissue repair and wound healing (reviewed in Maciag, 1989). Angiogenesis, i.e., new blood vessel
formation, is a rather complex process involving the degradation of basement membrane, the orderly migration, proliferation and differentiation of vascular endothelial cells and tube formation (Folkman, 1986). The turnover rate of normal capillary endothelial cells is measured in years (Denekamp, 1984) and angiogenesis is an infrequent event except in female reproductive system and embryogenesis. Pathological angiogenic events are also associated with atherosclerosis, chronic inflammation (Polverini et al., 1977), diabetic retinopathy (Lutty et al., 1986) immune reaction (Auerbach, 1981), solid tumor growth (Folkman and Klagsbrun, 1987) and other disorders.

HBGFs, at concentration of 10 to 100 ng, are angiogenic in vivo using rabbit cornea, chick chorioallantoic membrane or hamster cheek pouch bioassay. HBGFs, at very low concentrations, can promote and coordinate the site-specific formation of neovessel in situ (Thompson et al., 1989). Other angiogenic factors include angiogenin, TGFα, TGFβ, TNF, platelet-derived endothelial cell growth factor and newly identified vacular endothelial growth factor (Folkman and Klagsbrun, 1987; Ishikawa et al., 1989; Leung et al., 1989).

In summary, HBGF possess diverse functions with regard to their biological activities. HBGF-1 and HBGF-2 display overlapping, but not identical, spectra of activities. Nonetheless, HBGF-1 has several documented activities which HBGF-2 has not: (1) The ability to induce neurite outgrowth from rat retinal ganglion cells (Lipton et al., 1988); (2) The ability to induce cells derived from a rat bladder carcinoma to lose their epithelial characteristic and to acquire some properties of
typical mesenchymal cells (Valles et al., 1990); and (3) The ability to regulate hepatocyte regeneration (Kan et al., 1989).

**The HBGF-1 Gene and Chromosome Abnormalities**

Complementary DNA clones encoding human HBGF-1 have been isolated and sequenced (Jaye et al., 1986). Chromosomal localization analysis places the gene for HBGF-1 on human chromosome 5 between bands q31.3 and q33.2 (Jaye et al., 1986; Huebner et al., 1990). For comparison, cDNA clones of human (Abraham et al., 1986b; Kurokawa et al., 1987) and bovine (Abraham et al., 1986a) HBGF-2 have also been cloned and the human gene is located on chromosome 4 (Mergia, 1986). The relative positions of intron and exon boundaries of HBGF-1 and HBGF-2 are similar (Abraham et al., 1986b; Gospodarowicz et al., 1987b; Wang et al., 1989).

Nucleotide sequence analysis of human HBGF-1 cDNA demonstrates that an open reading frame, which corresponds to the protein sequence of HBGF-1 (154 a.a.; Burgess et al., 1986; Crabb et al., 1986) is flanked by a upstream in frame TGA translation termination codon (Jaye et al., 1986), a feature not found in the HBGF-2 cDNA (Abraham et al., 1986b). The sequence flanking the initiation codon (5'-ACCAUGG-3') conforms well to the nucleotide sequence that is optimal for protein synthesis in eukaryotes (Kozak, 1986).

The critical region on the long arm of chromosome 5, at which human HBGF-1 is located, happens to be the location for a variety of growth factors and growth factor receptor genes. For example, other genes coding for hematopoietic growth factors (including GM-CSF, IL-3, IL-4, IL-5, CSF-1), growth factor receptors (for CSF-1, PDGF and HBGF), receptors for
adrenergic drugs, glucocorticoid and dopamine, a myeloid cell differentia-
tion marker (CD14) and a putative transcriptional regulator (for reviews,
see Huebner et al., 1990; Sunahara et al., 1990) are all located in this
region. Interstitial deletions involving this region of chromosome 5
have been observed in bone marrow cells from patients with acute
nonlymphocytic leukemia (ANLL, including therapy-related ANLL, myelodys-
plasia and de novo ANLL) and a myelodysplastic syndrome characterized as
5q- refractory anemia (5q- RA) (reviewed in Nimer and Golde, 1987; Huebner
et al., 1990). Patients with abnormalities of chromosome 5 are known to
have poor responses to standard induction chemotherapy and their overall
survival are poor (Bloomfield, 1986; Le Beau et al., 1986). Particular
disorders with deletions on the long arm of chromosome 5 were shown to
involve the gene coding for c-fms, GM-CSF and IL-3,4 and 5 (Nimer and
Golde, 1987; Le Beau et al., 1989).

Structure - Function Relationships of HBGF-1

As described earlier, HBGF-1 and HBGF-2 share 55% amino acid
sequence identity. They are highly conserved among different species
during evolution. Baird et al. (1986b) first reported the synthesis of 25
peptides, which together encompass and overlap the entire sequence of
HBGF-2. They identified two functional domains, (24-68)-NH₂ and
(106-115)-NH₂ (using numbering system of 1-143 for HBGF-2) based on the
ability of the synthetic peptides to interact with HBGF receptor, bind
radiolabeled heparin and inhibit HBGF-2 stimulation of thymidine
incorporation. Similarly, Schubert et al. (1987) demonstrated that
(1-24)-NH₂, (24-68)-NH₂ and (93-120)-NH₂ synthetic peptides are able to
stimulate cell-substratum adhesion of PC12 cells. So far, only one report
described a synthetic peptide corresponding to (49-72)-NH₂ of HBGF-1 (using
numbering system of 1-154 for full length HBGF-1) which is able to compete
with HBGF-1 for heparin binding (Mehlman and Burgess, 1990). This region
corresponds to (24-68)-NH₂ of HBGF-2 as a heparin-binding domain.

The finding of two cysteines which are conserved in all species of
HBGF-1 and HBGF-2 has led investigators to assume that these residues are
involved in the formation of intracellular disulfide bonds. Harper et al.
(1986) presented indirect evidence for a disulfide bond between the two
cysteines residues. McKeehan and Crabb (1987) also showed that the
multiple fractions obtained during reverse phase HPLC elution of HBGF-1
may be due to scrambled disulfide bonds. It has been reported that
quantitative alkylation of recombinant human HBGF-1 abolished all
detectable receptor binding activity (Jaye et al., 1987). In contrast,
Crabb et al. (1986) reported that alkylation of the four cysteines of
human HBGF-1 without prior reduction has no effect on its mitogenic
activity. It turns out that the two conservative cystines are facing
opposite directions and are not sufficiently close to each other to allow
the formation of intramolecular disulfide bonds based on x-ray crystallo-
graphy (Zhu et al., 1991).

Harper and Lobb (1988) were able to show that limited reductive
methylation of bovine HBGF-1 (15-154 form) resulted in significant
methylation only of lysine-132. The modified protein has a reduced
heparin-binding, receptor-binding and mitogenic activities. Burgess et
al. (1990b) used site-directed mutagenesis to change lysine-132 to
glutamic acid. The mutant protein could still stimulate tyrosine kinase
activity and induce proto-oncogene expression. The results are consistent with previous results documenting reduced heparin-binding and mitogenic activities. Imamura et al. (1990) identified a putative nuclear translocation sequence at amino-terminus of HBGF-1. Mutant HBGF-1 without this sequence fails to induce DNA synthesis and cell proliferation.

So far, functional human HBGF-1 proteins (15-154, and 21-154 forms) have been expressed in E. coli (Jaye et al., 1987). Full-length human HBGF-1 (1-154 form) has been expressed in yeast Saccharomyces cerevisiae (Barr et al., 1988). Most recently, high level expression of at least 50 mg/l culture has been achieved (Ke et al., 1990). It can be anticipated that many further reports of the HBGF-1 cDNA expression in prokaryotic and eukaryotic cells and site-directed mutagenesis will be forthcoming and these reports will lead us to understand the structure-function relationships of HBGF-1 protein.

Oncogenic Members of the HBGF Family: int-2, hst/KS3 and FGF-5

To date, five other members of the HBGF family have been identified and characterized (Fig. 3 and Table 1). Three of them are oncogenes, since they were originally isolated from transformed cells using the focus formation assay.

A. int-2

Oncogene int-2 was identified as a transforming gene by studying the consensus integration sites of mouse mammary tumor virus (MMTV) in virally-induced tumors (Dickson et al., 1984; for a review, see Peters, 1990). int stands for INTEGRation site and at least one other such
Figure 3. Homology of protein sequence among members of human HBGF family. Sequences shared by four or more proteins are underlined. The 5'-end of human FGF-6 coding region has not been determined. The length of human FGF-6 protein is estimated to be the same as the murine homologue (de Lapeyriere et al., 1990). The arrowheads indicated the junctions of exons. The middle exons of the genes coding for each of the first six proteins all have 104 nucleotides with exactly the same boundary, except the gene structure of KGF has not been determined.
<table>
<thead>
<tr>
<th>Protein</th>
<th>HST/KS3</th>
<th>FGFR-5</th>
<th>FGFR-6</th>
<th>KGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBGF-2</td>
<td>MSQGTR</td>
<td>FSHLILSAR</td>
<td>TLLPQVLa</td>
<td>FSHLILSAR</td>
</tr>
<tr>
<td>INT-2</td>
<td>KMRLPL</td>
<td>LLSFLLLFL</td>
<td>PGHFLP</td>
<td>FSHLILSAR</td>
</tr>
<tr>
<td>HST/KS3</td>
<td>VPRQPK</td>
<td>RGPGRRLRR</td>
<td>VFRHFL</td>
<td>FSHLILSAR</td>
</tr>
<tr>
<td>FGFR-5</td>
<td>VFPQAR</td>
<td>VFRHFL</td>
<td>LLLLPQV</td>
<td>FSHLILSAR</td>
</tr>
<tr>
<td>FGFR-6</td>
<td>VLQLS</td>
<td>VFRHFL</td>
<td>LLLLPQV</td>
<td>FSHLILSAR</td>
</tr>
<tr>
<td>KGF</td>
<td>EQAEPQ</td>
<td>VFRHFL</td>
<td>LLLLPQV</td>
<td>FSHLILSAR</td>
</tr>
</tbody>
</table>

**Figure 3**
Table 1. Sequence similarity among the members of HBGF family.

<table>
<thead>
<tr>
<th></th>
<th>HBGF-1</th>
<th>HBGF-2</th>
<th>int-2</th>
<th>hst/KS3</th>
<th>FGF-5</th>
<th>FGF-6</th>
<th>KGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBGF-1</td>
<td>100.0% (155)</td>
<td>54.1% (157)</td>
<td>33.6% (143)</td>
<td>37.8% (127)</td>
<td>38.3% (128)</td>
<td>33.1% (127)</td>
<td>37.0% (120)</td>
</tr>
<tr>
<td>HBGF-2</td>
<td>100.0% (155)</td>
<td>40.0% (135)</td>
<td>43.0% (135)</td>
<td>41.4% (128)</td>
<td>43.2% (125)</td>
<td>39.0% (120)</td>
<td></td>
</tr>
<tr>
<td>int-2</td>
<td>100.0% (239)</td>
<td>32.4% (145)</td>
<td>43.0% (142)</td>
<td>37.1% (143)</td>
<td>40.0% (120)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hst/KS3</td>
<td>100.0% (206)</td>
<td>48.4% (128)</td>
<td>70.1% (127)</td>
<td>33.0% (120)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGF-5</td>
<td>100.0% (233)</td>
<td>47.7% (132)</td>
<td>41.0% (120)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGF-6</td>
<td>100.0% (128)</td>
<td>35.0% (120)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KGF</td>
<td>100.0% (194)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The numbers in parentheses are the numbers of amino acids in each protein or the numbers of amino acids for comparison between each pair of proteins. Gaps are introduced to each amino acid sequence to achieve maximum similarity using the Needleman-Wunsch method.
chromosome site int-1 exists. Disruption of the int-2 locus by a provirus had activated the transcription of the gene (Moore et al., 1986). The proviruses are positioned so that their transcription orientation is opposite from the direction of transcription of int-2 gene, therefore, it is the prototype of activation by enhancer insertion (Peters et al., 1989a). The human homolog of mouse int-2 was then cloned and analyzed (Casey et al., 1986; Brookes et al., 1989a). The gene comprises typical HBGF gene organization, namely three protein-coding exon, which spread over about 11-kbp on chromosome 11q13. The human int-2 protein contains an open reading frame coding for 239 amino acid residues and is 89% homologous to its murine counterpart. Comparison of the predicted amino acid sequence of human int-2 with human HBGF-1 and HBGF-2 showed significant homologies (Dickson and Peters, 1987; see Fig. 3 and Table 1). The protein contains a relatively short amino-terminal extension and a longer carboxyl-terminal extension.

The int-2 gene expresses only in early gestation stages of development and in teratocarcinoma (Jakobovits et al., 1986; Wilkinson et al., 1988). No transcripts have yet been identified in normal adult tissues. During subsequent fetal development, the int-2 transcripts are detected in Purkinje cells of the cerebellum, the developing retina, teeth and inner ear (Wilkinson et al., 1989).

In embryo and mouse embryonal carcinoma cell lines, several classes of int-2 transcripts can be detected, which arise through alternative promoter usage and alternative polyadenylation (Mansour and Martin, 1988; Smith et al., 1988). Murine int-2 protein can be synthesized as amino-terminally extended forms initiating at an in-frame CUG codon. The
leucine- and methionine-initiated proteins are found to be in the nucleus and secretory pathway, respectively (Acland et al., 1990). This phenomenon resembles human HBGF-2 protein, as mentioned earlier.

The \textit{int-2} protein appears to be a secreted growth factor because its amino-terminal domain contains a short hydrophobic sequence, which could function as an atypical signal peptide. However, neither tissue culture derived, nor cell-free synthesized \textit{int-2} protein gives a detectable mitogenicity. Expression of secreted forms of \textit{int-2} cause a morphological transformation in subset of transfected 3T3 cells, and anchorage-independent growth of SW13 adrenal cortical tumor cells (Merlo et al., 1990). \textit{int-2} gene product can act as an epithelial growth factor in transgenic mice (Muller et al., 1990).

\textit{int-2} gene was found to be amplified in 15-20\% of breast carcinoma (Lidereau et al., 1988; Zhou et al., 1988; Adanane et al., 1989; Ali et al., 1989; Fantel et al., 1990; Meyers et al., 1990). Another oncogenic member of the HBGF family, \textit{hst/KS3}, is tightly linked to \textit{int-2} on 11q13 and is co-amplified (Yoshida et al., 1988b). Mouse homolog of \textit{hst/KS3} gene is also adjacent to \textit{int-2} and is activated by proviral insertion in some virally induced mammary tumors (Brookes et al., 1989b; Peters et al., 1989b). However, DNA amplification does not usually result in detectable expression of either the \textit{int-2} or \textit{hst/KS3} gene (Liscia et al., 1989; Theilet et al., 1989). \textit{int-2} and \textit{hst/KS3} are also co-amplified in esophageal (50\%; Tsutsumi et al., 1988), bladder (Tsutsumi et al., 1988), head and neck carcinoma (Somers et al., 1990), melanoma (Adelaide et al., 1988), and hepatocarcinoma (Hatada et al., 1988).
B. hst/KS3

hst/KS3 was isolated as a transforming gene in an NIH/3T3 transfection assay from a human stomach tumor (hst; Yoshida et al., 1987) and kaposis sarcoma (KS3; Delli-Bovi et al., 1987) by two independent investigators (reviewed in Basilico et al., 1989). It was subsequently found that the hst/KS3 genes isolated from a noncancerous portion of stomach mucosa (Sakamoto et al., 1986), normal human leukocyte (Sakamoto et al., 1988) and a variety of other tumors (hepatocarcinoma, colo-carcinoma and melanoma), all have transforming activity which is not due to mutation or rearrangement in the coding region. It was reasoned that the expression of this gene is normally suppressed by a cis-acting element, and that rearrangement in the flanking region of the gene during gene transfer process might activate the transcription of the gene. Thus, hst/KS3 oncogene was the most frequently detected transforming gene after the RAS oncogene, and is the only (with another member of FGF gene family, FGF-6) known proto-oncogene with potent transforming activity.

The structure of hst/KS3 is similar to that of other HBGF genes and consists of three exons (see Fig. 3 and Table 1). The hst/KS3 polypeptide contains a long amino-terminal extension predicted to encode a signal peptide. It is secreted efficiently into the culture medium as a glycosylated protein (Delli-Bovi et al., 1988).

hst/KS3 is a potent mitogen for NIH/3T3 cells (Delli-Bovi et al., 1987) and for endothelial cells derived from capillaries or umbilical cord vein (Delli-Bovi et al., 1988). The processed form of the hst/KS3 protein synthesized in silkworm cells and in E.coli show potent angiogenic
activity in vivo. **hst/KS3 and HBGF-2** stimulate the same pattern of protein phosphorylation and may bind to the same receptors.

**hst/KS3**, like **int-2** oncogene, is transcribed in embryos and embryonic cells. **hst/KS3** transcripts are expressed in undifferentiated F9 cells whereas the homologous **int-2** messengers are present at low level (Velcich et al., 1987; Yoshida et al., 1988a). The DNA regulatory elements which are responsible for the specific activation of the **hst/KS3** gene in undifferentiated embryonal carcinoma (EC) cells reside within a DNA fragment which is part of the non-coding 3' region of the third exon. These elements are able to stimulate transcription only in undifferentiated cells, thus behaving as developmental specific enhancers (Curatola and Basilico, 1990).

C. **FGF-5**

**FGF-5** oncogene was originally isolated from a human bladder carcinoma cell line by DNA transfection assay and then selecting for transformed cells under growth factor-deficient conditions (Zhan et al., 1987). The rearrangement which activated the **FGF-5** gene had juxtaposed a retrovirus transcription enhancer just upstream from the native promoter of the gene (Zhan et al., 1988). The normal **FGF-5** has later been cloned from a 1-day-old human brain stem cDNA library (Zhan et al., 1988). **FGF-5** gene bears the three exon structures typical for HBGF members. The gene is mapped to chromosome 4q21 (Nguyen et al., 1988; Dionne et al., 1990b), the same chromosome to which **HBGF-2** is located.

The **FGF-5** protein consists of 267 amino acid residues which include extended amino- and carboxyl-terminal domain relative to HBGF-1 or HBGF-2
A signal peptide predicted from the cDNA sequence is present at the amino-terminal domain of FGF-5. It has been shown that FGF-5 transformed 3T3 cells secreted a mitogen which eluted from heparin-Sepharose columns at 1.0 M NaCl and has properties described for HBGFs.

Expression of murine FGF-5 RNA is highly restricted in developmental time and space (Haub et al., 1990).

Other HBGF-Related Genes: FGF-6 and KGF

A. FGF-6

FGF-6 gene was cloned by screening a mouse cosmid library with a human hst/KS3 probe under conditions of reduced stringency (Marics et al., 1989). The human FGF-6 gene was subsequently isolated and sequenced. This gene shows strong sequence homology (70%) with hst/KS3 over the c-terminal two-third portion of the putative proteins (see Fig. 3 and Table 1). The cloned normal human FGF-6 is able to transform NIH/3T3 cells using both focus and tumorigenicity assays (Marics et al., 1989). However, it has never been identified from human tumor DNA as a transforming gene in NIH/3T3 assays. The gene is composed of three coding exons and is localized on chromosome 12p13 as determined by in situ hybridization.

Mouse FGF-6 mRNA levels are developmentally regulated with a peak expression in the developing fetus at day 15.5 of gestation and moderate levels during late gestation and in the neonate (de Lapeyriere et al., 1990). In the adult, it can be detected in testes, heart and skeletal muscle. The biochemical and biological properties of the FGF-6 gene product remain to be characterized. It has been shown that FGF-6 is able
to transform NIH/3T3 cultured cells. Interestingly, FGF-6 uses alternative AUG exons to generate two major proteins processing signal peptides.

B. KGF

Recently, Rubin et al. (1989) have identified a growth factor specific for epithelial cells in the conditioned medium of a human embryonic fibroblast cell line. This keratinocyte growth factor (KGF) lacks mitogenic activity on either NIH/3T3 fibroblast or human saphenous vein endothelial cells. KGF is both acid and heat labile and consists of a single polypeptide chain of ca. 28 kDa. In a chemically defined medium, purified KGF was able to complement IGF-1/insulin growth requirement of keratinocytes and was shown to act through a signal-transduction pathway shared with EGF, TGF-α and HBGFs. KGF acts as a potent mitogen for human keratinocytes in culture, equivalent to or more active on a molar basis than EGF (Marchese et al., 1990). KGF binds heparin with less affinity and can be eluted from the heparin-Sepharose column at 0.6M NaCl. Sequence analysis of the cDNA coding for KGF reveals that it is a distinctive molecule with significant structural homology to HBGFs (see Fig. 3 and Table 1). The gene is expressed in the stroma underlying epithelial cells. This location suggests a role in interactions between mesenchymal and epithelial tissues. Thus, it appears to be a major paracrine effector of epithelial cell proliferation (Finch et al., 1989).

In summary, several features are shared by all or most of the members of HBGF family of polypeptide. Comparison of the structure of six HBGF-related genes (except for KGF whose gene structure is not available
as yet for comparison) reveals that they all consist of a similarly positioned three protein-coding exon format. Most prominently, the middle coding exons of the gene coding for six HBGFs all have 104 nucleotides with exactly the same 5'- and 3'-boundaries. The intervening sequences are of different lengths. The distance between the first and the second coding exons of HBGF-1 is 13.6-kbp and the distance between the second and the third exons is 5.3-kbp (Wang et al., 1989; Chiu et al., 1990b). The sizes of the two introns of HBGF-2 are at least 16-kbp and 16-kbp, respectively (Florkiewicz et al., 1991). The sizes of the two introns of human hst/KS3 are 617 and 638 bp, respectively (Yoshida et al., 1987), whereas those of human int-2 are 2,289- and 5,648-bp, respectively (Brookes et al., 1989a). The intron sizes of FGF-5 gene has not been determined precisely. The sizes of the two introns for FGF-6 are approximately 1.0- and 8.2-kbp, respectively (Moore et al., 1986). Thus, these genes are likely to have a common evolutionary origin, and the introduction of intervening sequences must have preceded gene duplication.

The established or derived primary structure of HBGFs are 32.4% to 70.1% identical (see Table 1) and the relative positions of several amino acid residues are conserved (see Fig. 3). Within a "core" of 120 a.a. residues, 19% of residues (included two cysteine residues) are invariant among all seven HBGFs (Basilico et al., 1989, Finch et al., 1989; Goldfarb, 1990). Each HBGF is strikingly conserved through mammalian evolution. The homology of HBGF-1 protein sequence among six different species is shown in Fig. 4. The nucleotide and amino acid sequence similarities are 76.6-95.0% and 88-100%, respectively (Table 2).
<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence 1</th>
<th>Sequence 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
<tr>
<td>Bovine</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
<tr>
<td>Rat</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
<tr>
<td>Hamster</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
<tr>
<td>Mouse</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
<tr>
<td>Chicken</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence 1</th>
<th>Sequence 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
<tr>
<td>Bovine</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
<tr>
<td>Rat</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
<tr>
<td>Hamster</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
<tr>
<td>Mouse</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
<tr>
<td>Chicken</td>
<td>++T++T+++K+H++L+++++++++++Y++++++T+++K++++++</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Homology of HBGF-1 protein sequence among different species. The identical amino acids in all species are represented by "++". The amino acids which are different from the major types are bolded. The arrowheads indicate the junction of exons (middle column). The human sequences is from brain stem cells (Jaye et al., 1986), the bovine from bovine retinal cells (Hailey et al., 1989), the rat from prostate transplantable Dunning tumor 3327 AT-3 cells (Goodrich et al., 1989), the hamster from DDT-1 cells derived from smooth muscle of the ductus deferens (Hall et al., 1990), the mouse from mouse embryo (Hebert et al., 1990), and the check from embryonic brain (Schnurch and Risau, in press).
Table 2. Nucleotide and amino acid sequence similarity of HBGF-1 among different species.

<table>
<thead>
<tr>
<th></th>
<th>human</th>
<th>bovine</th>
<th>rat</th>
<th>hamster</th>
<th>mouse</th>
<th>chick</th>
</tr>
</thead>
<tbody>
<tr>
<td>human</td>
<td>100.0%(100%)</td>
<td>91.4%(90%)</td>
<td>88.8%(95%)</td>
<td>91.6%(97%)</td>
<td>90.5%(95%)</td>
<td>79.6%(90%)</td>
</tr>
<tr>
<td>bovine</td>
<td>100.0%(100%)</td>
<td>84.9%(90%)</td>
<td>86.2%(90%)</td>
<td>86.4%(90%)</td>
<td>76.6%(88%)</td>
<td></td>
</tr>
<tr>
<td>rat</td>
<td>100.0%(100%)</td>
<td>100.0%(100%)</td>
<td>91.8%(98%)</td>
<td>95.0%(100%)</td>
<td>77.6%(90%)</td>
<td></td>
</tr>
<tr>
<td>hamster</td>
<td></td>
<td>100.0%(100%)</td>
<td>92.4%(99%)</td>
<td>78.5%(89%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mouse</td>
<td></td>
<td></td>
<td>100.0%(100%)</td>
<td>78.5%(90%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chick</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100.0%(100%)</td>
</tr>
</tbody>
</table>

The percentages in parentheses are amino acid similarity.
All HBGF mRNA transcripts contain relatively long 3'-untranslated regions. hst/KS3 and FGF-5 contain a second open reading frame in this region (Taira et al., 1987; Zhan et al., 1988). The 3'-untranslated region also contains several sequences of the ATTT(A) type, which have been described as characteristic of unstable, short-lived mRNAs (Shaw and Kamen, 1986). For example, ATTTA motif is repeated six times within the 3'-untranslated sequence of HBGF-1 (Crumley et al., 1989; Wang et al., 1991) and is repeated five times in the corresponding region of hst/KS3 (Delli-Bovi et al., 1987).

**Transforming Potential of HBGF-1 and HBGF-2**

The notion of the relationship between growth factors and oncogenes was first vindicated by several groups that showed the *sis* oncogene of simian sarcoma virus derived from the cellular gene for PDGF and the human *c-sis* proto-oncogene as the structure gene for the B-chain of PDGF (Chiu et al., 1984; for a review, see Chiu, 1989). Subsequently, it was shown that the *fms* oncogene is a transmembrane protein derived from the cellular gene for the receptor of colony stimulating factor-1 (CSF-1; also known as M-CSF; Sherr et al., 1985); proto-oncogene *erbB* codes for the receptor for EGF (Downward et al., 1984). Two other oncogenes, *neu* (also known as HER-2 or *c-erbB*-2) and *c-kit* are also transmembrane proteins with tyrosine kinase domain. They may encode growth factor receptors. In fact, recently a 30 kDa protein (gp30) is identified to be a ligand for p185*neu* receptor (Lupu et al., 1990). Mast cell growth factor (MGF; also known as stem cell factor SCF, or hematopoietic growth factor KL) is a ligand for the *c-kit* receptor (Huang et al., 1990; Williams et al., 1990; Zsebo et
al., 1990). Together they provide a compelling reason for testing the transforming potential of HBGF-1 and HBGF-2.

Rogelj et al. (1988) were the first to demonstrate that NIH/3T3 cells transfected with bovine HBGF-2 cDNA and expressing the native 18 kDa protein did not appear to trigger autocrine transformation. They could produce small slow-growing, spontaneously regressing tumors at low frequency in nude mice (Rogelj et al., 1989). In contrast, addition of the mouse immunoglobin signal peptide 5' to the HBGF-2 coding sequence (designated as Ig-bFGF) had a profound effect on cellular phenotype and tumorigenic potential. Similar results were obtained by Blam et al. (1988) who used a HBGF-2 cDNA fused to the signal sequence of growth hormone. However, contradiction remains in that the latter reported efficient secretion of the signal peptide-fused HBGF-2 by the transformants while the former did not detect any Ig-bFGF secretion. Nonetheless, secreted HBGF-2 exhibited little mitogenic activity, suggesting that interaction of HBGF-2 with its cognate receptors likely occurs while the fusion protein is being processed along the secretory pathway, as described for the PDGF/sis oncogene (Leal et al., 1985; Robbins et al., 1985; Betsholtz et al., 1984), IL3 (Wong et al., 1987; Browder et al., 1989; Dunbar et al., 1990) and GM-CSF (Lang et al., 1985). Browder et al. (1989) called these instances as "private" autocrine loop to differentiate from traditional view of "public" autocrine loop.

In contrast, two other reports showed either baby hamster kidney-derived fibroblast (BHK-21) or mouse BALB/c 3T3 cells transformed with wild-type HBGF-2 cDNA grown autonomously on soft agar and have detectable protein levels of HBGF-2 in conditioned media (Neufeld et al., 1988;
Sasada et al., 1988). In at least one case (Sasada et al., 1988), the phenotypic alteration in the transformed cells could be reversed by the addition of neutralizing antibodies to the medium. The discrepancies among the four reports could be due to differences in the expressing levels using different cell lines, different expression vector systems or different secretory signals.

Quarto et al. (1989) observed that the transformed phenotype by transfecting NIH/3T3 cells with HBGF-2 cDNA required high level of expression. They also reported the transformation caused down-regulation of the HBGF-2 receptor (Moscatelli and Quarto, 1989). Later, Yayon and Klagsbrun (1990) also showed that surface HBGF receptors in chimeric signal peptide-HBGF-2 transfected cells were unavailable and down-regulated and, therefore, the cells were transformed by an internal autocrine loop. Suramine up-regulated surface HBGF receptor and inhibited autocrine transformation (Yayon and Klagsbrun, 1990).

Thus far, there are only two reports on the expression of HBGF-1 in the mammalian system, one for truncated form (Jaye et al., 1988) and one for full-length form (Bunnag et al., 1991). None of the reports describes the signal peptide-fused HBGF-1. NR6 cells, a Swiss/3T3 mutant lacking EGF-R (Huang et al., 1986), expressing the truncated form of HBGF-1 (deletion of 1-21 amino acid residues) are unable to support autocrine growth on soft agar and produce only small nonprogressive tumors in nude mice (Jaye et al., 1988). However, NIH/3T3 transfectants expressing full-length HBGF-1 are capable of growing on soft agar even in the absence of exogenously-provided HBGF-1, and induce tumor formation when injected into nude mice (Bunnag et al., 1991). Thus, different forms of HBGF-1 may
posses subtle yet distinct biological properties. It is also reported that cells transfected with cDNAs that code only for high molecular weight forms of HBGF-2 have impaired growth (Renko et al., 1990). This is suggestive of specific functional roles for different forms of HBGF-2.

Rifkin and Moscatelli (1990) speculated that it is unlikely that enhanced HBGF-1 or HBGF-2 expression is a primary event in the etiology of natural tumors since sufficient levels of expression would be difficult to achieve. However, it will also be interesting to determine if nature would employ signal peptide translocation as an alternative route for carcinogenesis in human malignancies.

The HBGF Receptors

Radioreceptor binding assays established for HBGF-1 have identified the receptors on a variety of cell types (Friesel et al., 1986; Huang et al., 1986; Neufeld and Gospodarowicz, 1986; Olwin and Hauschka, 1986; DiSorbo et al., 1988; Lee et al., 1989). Radiolabeled HBGF-1 has been cross-linked to high-affinity receptors with molecular weights of 130-165 kDa (Gospodarowicz et al., 1985; Friesel et al., 1986; Olwin and Hauschka, 1986; Libermann et al., 1987; Winkles et al., 1987; Coughlin et al., 1988; Kan et al., 1988; Friesel et al., 1989). The high-affinity HBGF-1 receptor has a Km of 50-500 pM with 0.5 - 5.0 x 10^4 molecules per cell. In addition, HBGF-2 will compete for the binding of HBGF-1 to its receptor and vice versa (Neufeld and Gospodarowicz, 1986; Olwin and Hauschka, 1986; Burrs and Olwin, 1989). These results suggest that they share a common cellular receptor. Recent work also indicates that hst/KS3 also share receptors with HBGF-2 (Mansukhani et al., 1990).
Further work indicated that HBGF receptors are associated with protein tyrosine kinase activity and it is autophosphorylated at the tyrosine residue following binding of ligand, a property shared by many other growth factor receptors (Huang and Huang, 1986; Coughlen et al., 1988; Friesel et al., 1989; Lee et al., 1989; for review see Ullrich and Schlessinger, 1990). There is not a consensus regarding the role for protein kinase-C activation, phosphoinositide hydrolysis, or calcium mobilization in the mechanism of HBGF-1 (Tsuda et al., 1985; Kaibuchi et al., 1986; Magnaldo et al., 1986; Yamamoto et al., 1988). HBGF-1 induces tyrosine phosphorylation of three distinct polypeptides of molecular weights 150, 130 and 90 kDa in NIH/3T3 (Friesel et al., 1989). The identification of MW 150 kDa protein as phospholipase C-γ (PLC-γ) has been established, first by cDNA cloning and subsequently by immunological criteria (Burgess et al., 1990a). The breakdown of phosphatidylinositol biphosphate by PLC-γ can lead to the generation of inositol 1,4,5-triphosphate (IP3), which releases calcium from intracellular storages, and diacylglycerol (DG), which activates protein kinase C, respectively.

It has been demonstrated that there is an approximate inverse relationship between the number of receptors per cell and the content of endogenous HBGF (Moscatelli, 1987), suggesting that receptors are down-regulated by endogenous growth factor. Binding and cross-linking experiments show HBGF-1 can induce rapid down-regulation of its receptors with subsequent internalization by receptor-mediated endocytosis (Schreiber et al., 1985; Huang et al., 1986; Friesel et al., 1986; Clegg et al., 1987; Friesel and Maciag, 1988). Internalized ligand is remarkably resistant to degradation (Friesel and Maciag, 1988), a
phenomenon which is different from several other polypeptide growth factors. Degradation of HBGF-1 occurs in the lysosomal compartment with the generation of relative large fragments of the polypeptide.

Recently, candidate receptor proteins have been purified by HBGF affinity chromatography, and the corresponding cDNAs have been cloned. They can be grouped into three different types.

A. **cek1/flg/NE2**

Lee et al. (1989) first cloned a HBGF-2 receptor cDNA from a chick embryo cDNA library using an oligonucleotide probe (Lee et al., 1989). This chick HBGF-2 receptor is a transmembrane protein that contains three extracellular immunoglobulin (Ig)-like domains, an unusual acidic region, and an intracellular tyrosine kinase domain. The Ig-like motifs together with an insertion within the consensus kinase domain are features common to receptors for PDGF and CSF-1, although the latter bear five Ig-like domains instead of three (for a review, see Ullrich and Schlessinger, 1990).

The chicken HBGF-2 receptor is coincidentally identical to an independently isolated cDNA termed **cek1** (chicken embryo kinase 1; Pasquale and Singer, 1989), which is isolated from a 10-day chicken embryo cDNA expression library using anti-phosphotyrosine antibody. The chicken HBGF-2 receptor cDNA also has ~95% sequence identity to a previously identified partial cDNA clone, **flg** (fms-like gene; Ruta et al., 1988), which is isolated from a human endothelial cell cDNA library by hybridizing at relax stringency using the **v-fms** oncogene (derived from the cellular gene for CSF-1 receptor, a tyrosine-specific protein kinase).
The human fig protein turned out to be the receptor for both HBGF-1 and HBGF-2 (Ruta et al., 1989).

NE2, the mouse HBGF-2 receptor, was identified by Reid et al. (1990) by using the polymerase chain reaction to amplify mRNA sequences from 10-day-old mouse embryo neuroepithelium. These amplified sequences lie between the primers designed from two highly conserved amino acid motifs from the catalytic domain of protein-tyrosine kinases. More importantly, an alternatively spliced transcript encodes a shorter protein lacking the first of the three Ig-like domains (Reid et al., 1990). Mansukhani et al. (1990) also cloned the shorter form of murine HBGF-2 receptor and showed the expressed receptor is activated by HBGF-2 and hst/KS3. Safran et al. (1990) used oligonucleotide-derived from the kinase insert domain of the human fig sequence to screen a mouse brain cDNA library and isolated mouse fig with three Ig-like domains.

Later, the shorter form of human HBGF-2 receptor was isolated from placenta cDNA library (Itoh et al., 1990). Surprisingly, secreted forms of human HBGF receptor was subsequently isolated from human placenta and human umbilical vein endothelial cell cDNA library (Johnson et al., 1990). The secreted forms of the receptors lack transmembrane sequence and have a different carboxyl-terminal half of the third Ig-like domain.

B. bek/K-sam/cek3

The second type of HBGF receptor is called bek (bacterially expressed kinase; Kornbluth et al., 1988). The amino acid sequence predicted from the bek partial cDNA sequence was found to be 85% homologous to the cek1/fig/NE2. Dionne et al. (1990) reported the
complete cDNA cloning of bek and showed the similarity and distinction with cek/flg/NE2. The intracellular tyrosine kinase domain of these receptors are 90% similar at the amino acid level. The kinase domain has been found to be interrupted by a 14 amino acid insert. Two loops form of bek was also first identified as an amplified gene in stomach cancer and named K-sam (stomach cancer-derived cell line, KATO-III cell-derived stomach cancer amplified; Hattori et al., 1990). The cek3 isolated by Pasquale (1990) turned out to be chicken bek with three loops form.

C. cef

The third type of HBGF receptor has been purified from chick embryo extracts and has properties distinct from the protein discussed above (Burrus and Olwin, 1989). The cDNA has been cloned and named cef. The predicted amino acid sequence of cef contains motif characteristics of cell surface receptors. However, cef has no homology to formerly identified tyrosine kinase HBGF receptors.

HBGF receptor is a portal of cellular entry for herpes simplex virus type 1 (HSV-1; Kaner et al., 1990). Transmembrane forms (with two Ig or three Ig domain) confer infectivity by HSV-1. Inhibitor of HBGF-2 binding to it receptor and competitive polypeptide antagonists of HBGF-2 prevent HSV-1 uptake.

Recently, a newly developed expression cloning system has been used to clone the receptor for KGF (Miki et al., 1991). The rationale is that introduction of the cDNA into NIH/3T3 cells which secrete KGF but do not express KGF receptors might create an autocrine transforming loop. The
successfully cloned KGF receptor has affinity for HBGF-1 as well as KGF and is an alternatively-spliced form of bek (Miki et al., 1991).

In summary, multiple alternatively spliced mRNAs have been found for at least three HBGF receptor genes. The high affinity binding of more than one HBGF to each of the different receptors represents a unique double redundancy without precedence among polypeptide growth factor-receptor interactions (Dionne et al., 1990).

The HBGF Secretion Mechanisms

Most of the secretory pre-proteins are extended at amino-terminals by a "leader" or "signal" sequence of 13-30 hydrophobic amino acid residues. This allows the co-translational, vectorial transport of the proteins across the membrane of the rough endoplasmic reticulum (Walter et al., 1984). Secretory proteins are further transported to the Golgi apparatus where they undergo glycosylation, and finally reach the secretory vesicles which, upon fusing with the plasma membrane, release their contents out of the cell.

The lack of a signal peptide in HBGF-1 and HBGF-2, the polypeptides for which several high affinity cell surface receptors have been described, invoked the obvious question of how it is released from the site of biosynthesis. A significant amount of information is available at the molecular level concerning the interaction of HBGF with its membrane receptor. However, the mechanism of secretion is still not known.
A. Intracrine Interaction

The lack of a signal peptide or an internal hydrophobic domain of HBGF-1 (Burgess et al., 1986; Crabb et al., 1986; Jaye et al., 1986) are shared by a number of cytoplasmic growth factors, including the precursor of interleukin-1α and -1β (IL-1α and -1β; Auron et al., 1984; March, 1985.), ciliary neurotrophic factor (CNTF; Stockli et al., 1989), ADF (a thioredoxin-like molecule released by leukemia T-cell; Tagaya et al., 1989), blood coagulation factor XIIIa (Grundmann et al., 1986), yeast mating pheromone-a factor (Kuchler et al., 1989) and the newly discovered platelet-derived endothelial cell growth factor (Ishikawa et al., 1989). A new category of mode of action for growth factors has been termed "intracrine interaction" (for a review, see Logan, 1990). Intracrine growth factors acting in this way need not be secreted. Rather, they remain within the producer cell and act directly as intracellular messengers to regulate cellular function.

As discussed earlier, other members of HBGF family all have signal peptide. The secreted members can act on the identified receptors localized at plasma membrane. Therefore, the receptors are not specific for HBGF-1 and/or HBGF-2. The cytoplasmic localization and nucleus translocation of HBGF-1 (Sano et al., 1990; Imamura et al., 1990) and HBGF-2 (Bouche et al., 1987; Baldin et al., 1990; Renko et al., 1990; Tessler and Neufeld, 1990) support this explanation.

B. Cell Leakage

Human IL-1 has approximately 30% sequence identity with HBGF-1 (Gimenez-Gallego et al., 1985; Thomas et al., 1985). Topologically
equivalent folds have been observed for IL-1 and HBGF-1 (Zhu et al., 1991). It has been postulated that IL-1 is released from stimulated monocytes or macrophages by leakage, perhaps due to cell damage (Auron et al., 1983). On the other hand, Kostura (1989) described an unconventional mechanism for IL-1 release after cleavage by a specific enzyme (convertase). Rubartelli et al. (1990) demonstrated that IL-1β is actively secreted via a pathway of secretion different from the classical endoplasmic reticulum-Golgi route. However, no defined pathway for release of HBGF-1 has been described other than cell death. A few reports have described the occurrence of soluble extracellular HBGF-2 (Neufeld et al., 1988; Schweigerer et al., 1987), the concentration has been so low that release by cell death cannot be excluded.

C. Matrix Interaction

HBGF may not be released in a soluble form. It is possible that HBGF is released from cells in association with extracellular matrix (ECM) components and become an integral part of the insoluble ECM (Baird and Ling, 1987; Vlodavsky et al., 1987; Folkman et al., 1988). The affinity of HBGF for heparin sulfate proteoglycans (HSPGs) and glycosaminoglycans (GAGs) has suggested that these cell-associated complex carbohydrates and the ECM can function as a storage depot and serve as a local source of growth factor. Release and activation of the factor from here may occur when specific enzymes, like proteases and/or heparinases secreted by target cells or invading tumor cells, degrade this structure (Vlodavsky et al., 1987; for a review, see Baird and Walicke, 1989). To date, there is
only one report that shows HBGF-1 deposited in the extracellular matrix of the neonatal rat cardiac myocyte (Weiner and Swain, 1989).

D. Alternative Splicing

It is possible that alternative splicing could occur in the cell secreting HBGF-1 such that a signal peptide could be synthesized. In the literature, different forms of mRNA coding for cellular and secreted protein, which are encoded by the same gene, have been reported (Gough, 1987; Grower et al., 1988; de Sauvage and Octave, 1989). This hypothesis can be addressed by the studies of HBGF-1 cDNA and genomic DNA.

E. Other Mechanisms

As mentioned previously, secreted forms of HBGF receptors have been identified (Johnson et al., 1990). The secreted forms may bind to HBGF in an intracellular compartment and subsequently serve as a means for secreting the factor.

Very recently, one type of transporter proteins named histocompatibility antigen modifier (HAM1 and HAM2) have been found in mice, rats, and humans (Monaco et al., 1990). Their function is to help shuttle peptides across intracellular membrane, therefore, viral protein fragments lacking a key molecular signal can cross the membranes, and displayed on the infected cells surface by a class I major histocompatibility complex (MHC) protein, a phenomenon known as "antigen presentation". HBGF proteins may have similar transport mechanisms.

Secretion of α-factor, a yeast mating pheromone, requires the STE6 gene product, a protein homologous to the mammalian multidrug resistance
(MDR) glycoprotein (Kuchler et al., 1989; McGrath and Varshavsky, 1989). HBGF protein secretion may be related to the MDR function in mammals.

**Aims of Studies**

I have described a number of the features of HBGF-1 and also other members of HBGF family. HBGF-1 and other members have several unusual properties not found in other polypeptide growth-factors. Foremost among these is its lack of a signal sequence and intracellular localization. One hypothesis is mRNA coding for HBGF-1 with signal peptide may be present in some tissues where HBGF-1 is secreted. A definitive answer to this hypothesis needs systematic analysis of cDNA clones and genomic DNA clones coding for HBGF-1. One objective of this study is to address this issue by cDNA cloning. My major accomplishments include isolation of four HBGF-1 cDNA clones, determination of their sequences, and mRNA expression in fetal tissues. I have showed that the human fetal heart expresses high levels of HBGF-1 mRNA, which indicates that HBGF-1 may be involved in mediating some processes such as embryonic development and vascular growth in the heart. I have found that alternative splicing generates at least four different froms of HBGF-1 mRNA with the same coding region, which may result from the usage of alternative promoters in different tissues.

These results suggest that the gene structure and regulatory mechanisms for gene expression of HBGF-1 are complex. In an effort to understand the regulation of expression of the HBGF-1 gene, the next objective is to undertake the cloning and sequencing of the gene coding
for HBGF-1. My major accomplishments include isolation of nine contiguous genomic DNA clones, determination of the complete DNA sequences of all three coding exons and the exon-intron organization, mapping of the 5'-end of the first coding-exon and the 3'-ends of the gene. I have identified multiple polyadenylation sites, and have compared the nucleotide sequences between the human and bovine HBGF-1 and showed the high degree of sequence similarity. I have also detected HBGF-1 transcripts from glioblastoma cell lines, fetal kidney and brain.

Cloning of the HBGF-1 gene allowed me to characterize this locus in acute nonlymphocytic leukemia (ANLL) patients. I have screened a prospective cohort of eight ANLL patients using three probes derived from the HBGF-1 locus. I concluded that there is no allelic loss or gross rearrangement in ANLL patients with or without 5q- deletion. There is also no difference in the expression of this gene in the mononuclear cells of ANLL patients and normal control individuals. These results suggest that HBGF-1 has no apparent role in the etiology of ANLL.
MATERIALS AND METHODS

Tissues, Cell Lines and Materials

T4 DNA ligase, T4 polynucleotide kinase and most of the restriction endonucleases were purchased from New England Biolabs Inc. (NEB, Boston, MA) and Bethesda Research Laboratories (BRL, Gaithersburg, MD). Some restriction endonucleases were purchased from Boehringer Mannheim Biochemicals (BMB, Indianapolis, IN). Calf intestine alkaline phosphatase (CIAP) and large fragment of E. coli DNA polymerase (Klenow fragment) were also from BMB. Radioactive isotopes were from Amersham Corp. (Arlington Heights, IL) and Sequenase kit for DNA sequencing was from United States Biochemical Corp. (USB, Cleveland, OH). Oligonucleotides were synthesized by a Model 380A DNA synthesizer from Applied Biosystems Inc. (ABI, Foster city, CA) at Davis Medical Research Center or Biochemical Instrumentation Center, The Ohio State University. At later stage, oligonucleotides were purchased from Oigos Etc (Guilford, CT). The fluorescent primers for an automated DNA sequencer (Model 370A, ABI) were purchased from ABI. Cloning vectors pBluescript (+)/(-) and phage T7/T3 RNA polymerases were obtained from Stratagene Cloning System (La Jolla, CA). DNA and amino acid analysis programs were purchased from DNA Star, Inc. and were operated on an IBM AT with a Compact Disk Reader.

E. coli HB101, JM101 and DH5α cells containing no plasmid were grown in Luria-Bertani (LB) medium [bacto-tryptone 10 grams, bacto-yeast extract
5 grams, and NaCl 10 grams in 1 liter of double-distilled H$_2$O (ddH$_2$O), pH 7.0] or plated on LB plates (with additional 1.1% bacto-agar for plates). Cells containing pBluescript vectors were grown in medium or plates containing 50 $\mu$g/ml of ampicillin (amp)

Human tissues, e.g. brain, kidney, placenta, heart, etc, were provided by the Tissue Procurement Service of the Ohio State University Comprehensive Cancer Center. Mononuclear cells were isolated from heparinized peripheral blood by density gradient centrifugation on Ficoll-Hypaque (Sambrrook et al., 1982). T cells were separated from non-T cells using rosetting procedure as described (Kanas et al., 1985). Human glioblastoma cell line A2781 was provided by S. Aaronson, National Cancer Institute, Bethesda, MD (Gazit et al., 1984); D-65MG was provided by G.Y. Gillespie, University of Alabama at Birmingham, Birmingham, AL; U1424MG was provided by C. Betsholz, University of Uppsala, Sweden. Human prostate carcinoma cell line LNCap clone FGC was purchased from ATCC (No. CRL 1740). U937 is a human promonocytic cell line and DMS10 is a subclone derived from U937. DMS10 was incubated in the presence of $\gamma$-interferon as described (van de Winkle et al., 1991) before RNA was isolated.

RNA Isolation and Oligo(dT')-Cellulose Chromatography

Total cellular RNA was extracted from tissues and cell lines by the guanidinium thiocyanate (GuTC) method and purified by ultracentrifugation through 5.7 M CsCl (Chirgwin et al., 1979). Usually six confluent T-150 flasks of cultured cells were used. The cells were washed with phosphate buffered saline (PBS), scraped in 7.5 ml GuTC cocktail (15.5 ml 4 M GuTC, 0.11 ml 14.3 M $\beta$-mercaptoethanol, 1-2 drops of antiform A emulsion) with
a rubber policeman. The cells were transferred to Dounce homogenizer and homogenized slowly for at least twenty times. For frozen tissues, usually one gram was pulverized in liquid $N_2$ with a mortar and pestle and then mixed with GuTC cocktail at room temperature. The tissue/GuTC mix was transferred to tissumizer and homogenized at 200 rpm for 2 min. More antiform A emulsion could be added if the homogenate became too foamy. After that, the samples from culture cells or whole tissues were layered over 3.5 ml cushion of 5.7 M CsCl/0.1 M EDTA, pH 7.0. Total cellular RNA was pelleted by ultracentrifugation in a Beckman SW41 rotor at 35,000 rpm for 18 hr in 15°. For small quantity of RNA, tabletop ultracentrifuge TLS-55 was used. Samples were suspended in 1.5 ml GuTC cocktail after homogenization. Solid CsCl could be added to a final concentration of 0.5 gram/ml. Then the mixtures were layered on top of 0.4 ml 5.7 M CsCl/25 mM EDTA cushion and spun at 55,000 rpm for 3 hr. The RNA was recovered by resuspending in TE buffer [10 mM tris(hydroxymethyl)aminomethane (Tris-HCl), 1 mM EDTA, pH 7.4] and ethanol precipitation. RNA was stored in ethanol at -80°. Before usage, 2-5 µg of RNA (according to the spectrophotometric reading at 260 nm) was run on a small formaldehyde/agarose gel and stained with ethidium bromide. From the integrities of the 28S ribosomal RNA (rRNA) and 18S rRNA (4.5-kb and 1.9-kb, respectively), the quality of the RNA was evaluated. The formaldehyde in the formaldehyde/agarose gel may be eliminated without affecting the electrophoretic separation of RNA, which also minimized the adverse effect of formaldehyde on gel staining (Liu and Chou, 1990).

Other methods have been used to eliminate the ultracentrifugation step of GuTC/CsCl method. One involved a single step method by acid
guanidinium thiocyanate-phenol-chloroform extraction (AGPC), as described by Chomczynski and Sacchi (1987). The commercialized product, called RNAzol B (Cinna/Biotecx, Friendswood, TX), was used according to the manufacturer's protocol.

Poly(A) RNA was enriched using an oligo(dT)-cellulose column (Aviv and Leder, 1972). Briefly, RNA was denatured in water at 62-65° for 1-2 min, and cooled rapidly in ice. Then it was adjusted to the final concentration of binding buffer [10 mM Tris-HCl, pH 7.5, 0.5 M NaCl, and 0.1% N-laurylsarcosine (sarkosyl)] and applied to an oligo(dT)-cellulose column (type III, Collaborative Research Inc., Bedford, MA). Flow-through buffer was collected and re-applied to the column. The column was washed with 6-7 ml of binding buffer for poly(A)^+ fractions, then with 3-4 ml of elution buffer [10 mM Tris-HCl, pH 7.5, 0.1% sodium dodecylsulfate (SDS)] for poly(A) fractions. Both poly(A) RNA and the poly(A)^+ RNA were ethanol precipitated and used in later studies. Usually between 2% and 5% of the total RNA applied to the column was recovered as poly (A)^+ RNA.

**Northern (RNA) Blot Analysis**

Twenty µg of total cellular RNA and/or 5 µg of poly(A) RNA for each sample was denatured at 65°C for 5 min in sample buffer (50% deionized formamide, 2.2 M formaldehyde) and loaded onto a 1.1% agarose gel containing formaldehyde (Dobner et al., 1981) and electrophoresed in 1X running buffer (20 mM 3-(N-Morpholino)propanesulfonic acid (MOPS), pH 7.0, 5 mM sodium acetate, and 1 mM EDTA). RNA was blotted onto a nitrocellulose filter from Schleicher & Schuell (S & S, Keene, NH) in 20 X salt-sodium citrate (SSC; 175.3 g NaCl and 88.2 g sodium citrate, pH 7.0, in 1
liter of ddH₂O for 20 X SSC) solution overnight using wick transfer as described by Thomas (1980). The filter was baked for 2 hr at 80° vacuum oven and prehybridized for a minimum of 2 hr to a maximum of overnight at 42°. Prehybridization and hybridization buffer were as follows: 50% deionized formamide with Dowex XG8 resin, 1 X Denhardt’s solution [0.02% ficoll, polyvinylpyrrolidone (PVP) and bovine serum albumin (BSA)], 5 X SSC, 50 mM sodium phosphate, pH 6.5, 0.25 mg/ml yeast tRNA, except hybridization buffer also contained 2 x 10⁶ cpm/ml of ³²P-labeled probe. After hybridization overnight, the filter was then washed with 2 X SSC, 0.1% SDS at room temperature for 2 x 5 min, then with 0.1 X SSC, 0.1% SDS at 42° for 2 x 15 min. The filter was exposed overnight to X-ray film at -70°C with intensifying screen. To ensure equal amounts of RNA were used, the filter was then re-probed with the 1.7-kbp EcoRV-ApaLI DNA fragment derived from a human β actin cDNA clone (Ponte et al., 1984) for normalization.

For hybridization of oligonucleotide probes to Northern blots, a slightly different hybridization buffer was used. The filter was hybridized in 2 X Denhardt’s, 6 X SSC, 50 µg/ml denatured herring sperm (HS) DNA, 50 µg/ml tRNA and 0.1% SDS at 37°C overnight. The filter was first wash with 2 X SSC, 0.05% SDS for one hr at room temperature, and then for 20 min with 6 X SSC, 0.1% SDS at 20°C below Td [Td= 4(G+C) + 2(A+T)].

Ribonuclease (RNase) Protection Assay

The DNA fragments were cloned into appropriate restriction enzyme sites of pBluescript vectors in the proper orientation for the production
of antisense RNA from the phage T7 promoter. The plasmid was then
digested with appropriate restriction enzymes. The digested DNA was
treated with proteinase K (50 μg/ml) at 37°C for 30 min, followed by
phenol/chloroform:isoamylalcohol (CIA, 24:1) at 1:1 extraction and ethanol
precipitated. The DNA was then used as the template for in vitro
transcription to produce uniformly labeled riboprobe with [α-32P] UTP as
described elsewhere (Melton et al., 1984) using T7 RNA polymerase instead
of SP6 RNA polymerase, simply because pBluescript vectors did not have the
SP6 promoter. The riboprobe was treated with DNase to remove the DNA
template, extracted with phenol/CIA 1:1, and recovered from ethanol
precipitation using tRNA as a carrier. By repeating ethanol precipitation
three times, the unincorporated [32P]UTP retained the supernatant.

Twenty μg of total cellular RNA from various tissues, cell lines and
the riboprobe (10⁶ cpm) were dissolved in 30 μl of hybridization buffer
[80% formamide, 40 mM piperazine-N,N'-bis(2-ethanesulfonic acid (PIPES),
pH 6.4, 400 mM NaCl and 1 mM EDTA, as described previously (Zinn et al.,
1983). After being heated to 85°C for 10 min, the RNA-RNA duplex was
allowed to anneal at 45°C overnight. Next day, the RNA mixtures was
proceeded to ribonuclease digestion by dilution with 350 μl of RNase
digestion buffer (10 mM Tris-HCl, pH 7.5, 5 mM EDTA, 300 mM NaCl)
containing 40 μg/ml of RNase A and 2 μg/ml of RNase T1. Incubation was at
30°C for 30 min. Ten μl of 20% SDS and 2.5 μl of 20 mg/ml proteinase K
were added to the protected fragment and incubated at 37°C for 15 min.
After extraction with phenol/CIA 1:1, the protected RNA-RNA duplex were
ethanol precipitated with 1 μl of 10 mg/ml tRNA as a carrier. The
protected fragments were separated on a 6% polyacrylamide/7 M urea
denaturing gel in Tris-borate-EDTA (TBE, 90 mM Tris-borate and 2 mM EDTA) buffer (Maxam and Gilbert, 1980). The size of the RNase resistant fragments were determined by using pBR322 plasmid DNA digested with Hinfl as a marker. The marker was labeled by the fill-in reaction of Klenow fragment in the presence of [α-32P]dATP (Maxam and Gilbert, 1980).

cDNA Library Screening

A λgt11 cDNA library was constructed by a modification of the method of Huynh et al. (1985) using poly(A) RNA prepared from a one-day-old human brain stem (de Ferra et al., 1985). The cDNA library was generously provided to us by Carmie Puckett (California Institute of Technology). The library was plated out in ten 150 mm petri dishes, each containing 100,000 phages, using E. coli Y1088 as the host cells. After overnight incubation at 37°C, the phage plaques were lifted onto the master and replica nitrocellulose filters (S & S) or Hybond N nylon membranes (Amersham) and screened by the method of Benton and Davis (1977). The twenty filters were denatured in 1.5 M NaCl/0.2 N NaOH, and neutralized in 0.5 M Tris HCl, pH 7.2/3 M NaCl. After air drying for several hours to overnight, the filters were baked in a vacuum oven for 2 hr at 80°C. The filters were prehybridized for 2 hr and then hybridized with a 33-base synthetic oligonucleotide, 5'-CTTGTCCGATTCTGGGGATCCGAGGATC-3', based on the published sequence (Jaye et al., 1986) under relaxed conditions. These conditions included hybridization in 6X SSC, 1X Denhardt's solution, 0.5% SDS, 100 μg/ml denatured herring sperm DNA at 37°C overnight with 1 x 10^6 cpm/ml of probe. The filters were first washed in 6x SSC at room temperature for 3 x 15 min, followed by 55°C for just 1 min. When a
cloned DNA fragment was used as a probe, the hybridization temperature was at 55° and the final wash was 0.1X SSC/0.1% SDS for 2 x 15 min at the same temperature. Autoradiography was performed on the filters. Positive signals were aligned between master filters and replica filters. If they were duplicated, the corresponding clones were removed from the master plates and subjected to three or four cycles of plaque purification.

**Preparation of Competent Cells**

Homemade competent cells were prepared by innoculating 10 ml of LB medium with a loop of bacterial glycerol or stab culture. The next day, 0.5 ml of the overnight cultures were added into 50 ml of fresh medium, and shaken for 2-4 hr until OD$_{590}$ was 0.5-0.7. The cells were spun and resuspended in 25 ml cold 50 mM CaCl$_2$. After incubating on ice for 10 min, the cells were again spun and resuspended in 50 mM CaCl$_2$ to a OD$_{590}$ = 10. For preserving the cells, glycerol was added to a final concentration of 15%(v/v) and 100 µl aliquots of the cells were stored at -80°C. When the frozen cells were retrieved, they were placed at 37°C first and then kept on ice. Commercial competent cells were also used for some purposes. Commercial competent HB101 and DH5α cells were from BRL, and JM101 cells were from Stratagene.

**Transformation**

Homemade competent cells (100 µl) were added to 10 µl of DNA ligation mixture. The mixture was incubated on ice for 10 min and heat-shocked at 42°C for 2.5 min. Finally 10 µl and 100 µl of the mixture were plated onto LB plates containing 50 µg/ml ampicillin. If blue/white
selection of DH5α or JM101 cells was desired, 20 μl of 100 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 60 μl of 2 % 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) were added to the competent cell mixture before spreading on the plates.

Commercial competent cells were used in a small scale. One μl of diluted DNA (1-10 ng) was added to 20 μl of cells. After incubation on ice for 30 min and heat shock at 42°C for 40 sec, 80 μl of SOC (containing 2% bacto-tryptone, 0.5% bacto-yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, and 20 mM glucose) was added to the tube and shaken for one hr at 37°C. Later, the entire mixture was applied onto one plate with IPTG and X-gal for blue/white color selection.

**Colony Hybridization/Microwave Colony Screening**

Bacterial colonies were toothpicked to the filters in a grid pattern, or they were "lifted" from a plate of colonies, and re-grown overnight. The filters were then processed by placing them sequentially upon Whatman paper soaked in the following solutions: 10% SDS for 5 min, 0.5 M NaOH for 5 min, 1 M Tris-HCl, pH 7.4, for 5 min, 0.5 M Tris-HCl/1.5 M NaCl. The bacterial debris were removed by scrubbing with a gloved hand using 2X SSC. Prehybridization and hybridization were the same as for the Southern blot.

Recently, a rapid procedure of microwave colony screening was used according to Buluwela et al. (1989). The filters were laid on Whatman paper soaked in 2X SSC/0.5% SDS for 2 min., and then treated for 2.5 min in a 650W microwave oven. The colony lysis, DNA denaturation and fixation occurred in this single step. The treated filters could be subjected to
standard prehybridization and hybridization without UV fixation.

**Isolation of Plasmid, Phages and High Molecular Weight Genomic DNA**

The mini alkaline extraction procedure was used for screening clones containing plasmid DNA. To prepare double-stranded plasmid DNA template for sequencing, more phenol/CIA extractions before and after ribonuclease digestion were performed. This purified mini-lysate plasmid DNA was sufficient for manual DNA sequencing and cycle sequencing on a automatic DNA sequencer (see later section). For rapid isolation of mini-prep plasmid DNA suitable for DNA sequencing, a method described by Jones and Schofield (1990) was used. It produced DNA within 30 min from pelleting the overnight bacterial culture.

A boiling mini-prep procedure was used for DH5α cells, as described by Holmes and Quigley (1981). RNase treatment of the DNA was performed at the same time as restriction enzyme digestion. Additional phenol/CIA treatments could generate template for DNA sequencing.

Large scale supercoiled (RFI) plasmid DNA extraction was performed using a CsCl gradient, as described in detail by Maniatis et al. (1982). For general usage, pZ523 columns (5 Prime → 3 Prime, Inc., West Chester, PA) were used as an alternative to separate plasmid DNA from chromosomal DNA.

Phage DNA was isolated as follows: 100 μl high-titer phages in phage storage buffer (PSB, 0.1 M NaCl, 10 mM Tris-HCl, pH 7.4, 10 mM MgCl₂ and 0.05% gelatin) were mixed with 100 μl CaMg (10 mM CaCl₂ and 10 mM MgCl₂) and 100 μl of an overnight bacterial culture. After incubation at 37°C for 30 min, the mixtures were added to 50 ml fresh NZY medium (5 g NaCl,
2 g MgCl$_2$.6H$_2$O, 10 g NZ amine and 5 g bacto-yeast extract, pH 7.0, in 1 liter of ddH$_2$O). The next day, after adding 2.5 ml chloroform to help lyse the bacteria, the phage lysate was separated from the bacterial debris by centrifugation at 3,000 rpm for 15 min at 4°C. Twenty-five ml of the supernatant was stored while the other 25 ml was used for DNA isolation. The supernatant was treated with 10 µg/ml DNase and 25 µg/ml RNase at 37°C for 1 hr. After 2 phenol extractions and one CIA extraction, the phage DNA was ethanol precipitated.

A modified procedure was also used for making phage DNA from small-scale liquid lysates, as described by Ausubel et al. (1989). The major change from the protocol above was a centrifugation at 27,000 rpm for 1.5 hr in SW28 rotor after DNase and RNase digestion. Isolated phage DNA could be recovered by passing through a Sepharose CL-4B (Pharmacia, Uppsala, Sweden) column to remove the contaminating RNA. Large-scale phage DNA was isolated from 5 liters of culture and concentrated by a Minitan unit (Millipore Corp., Bedford, MA).

High molecular weight genomic DNA was isolated from tissues and cell lines. Frozen tissues were crushed into fine powder in liquid nitrogen with a mortar and pestle. Phenol and extraction buffer (20 mM Tris-HCl, pH 7.5, 0.5% SDS, 1 M NaCl and 1 mM EDTA) at 1:1 were added and the mixtures was shaken for 2-3 hr. After spinning at 3,000 rpm for 15 min in a table top centrifuge RT6000B, the aqueous phase was recovered. One tenth vol of 3 M NaOAc, pH 5.2 and 2X vol of ethanol were added to the aqueous phase. The genomic DNA was spooled with a glass rod. The spooled DNA was dissolved in TE buffer and subsequently dialyzed in 20 mM Tris-HCl, pH 7.5, 10 mM NaCl and 1 mM EDTA overnight with two changes of
buffer.

High molecular weight genomic DNA from cultured cells and white blood cells was prepared by using slightly different protocols (Bell et al., 1981). Confluent cell cultures were washed three times with Tris/saline (1 liter: 8.0 g NaCl, 0.38 g KCl, 1.0 g Na₂HPO₄, 1.0 g glucose, 3.0 g Tris base, 15 mg phenol red, adjust pH to 7.0 and autoclave) and scraped with a rubber policeman. Harvested cells were lysed with 12 ml of ice cold lysis buffer (0.32 M Sucrose, 10 mM Tris-HCl, pH 7.5, 5 mM MgCl₂ and 1% Triton-X100) by gentle vortexing. The nuclei were recovered by centrifuging at 3,000 rpm for 10 min. Pelleted nuclei were resuspended in 75 mM NaCl/25 mM EDTA. SDS and proteinase K were added to final concentration of 0.5% and 250 μg/ml, respectively. The reactions were inoculated overnight at 37 °C. The next day, the reactions were extracted with phenol/CIA 1:1 once and CIA once. The genomic DNA was spooled out of the solution with a glass rod in the presence of 1/10 vol of 3 M NaOAc and 3 vol EtOH.

Isolation of DNA Fragments

In general, DNA fragments were isolated by a conventional electrophoresis method. The desired DNA fragments were cut from the agarose gel after electrophoresis. The gel slices were put into dialysis bags (Spectra/Por; Spectrum Medical Industries, Inc., Los Angeles, CA) and run at 50-100 volts in 0.5X Tris-acetate buffer (TAE; 1 X TAE = 40 mM Tris-HCl, 1 mM EDTA, pH 8.0) buffer. The EtBr-stained DNA fragments moved out of the gel slices in 0.5-2 hr. Reverse current of 100 volts was applied for 5 sec to prevent DNA fragments sticking to the dialysis
membrane. After phenol and CIA extraction, the DNA fragments were ethanol precipitated. The DNA fragments were recovered by centrifuge either at 8,000 rpm for 30 min (for regular fragments) or 25,000 rpm for 30 min (for diluted DNA fragments).

The Geneclean kit from Bio101 (Bio101, La Jolla, CA) was also used for quick recovery of the DNA fragments. This method was done according to the protocol supplied by the company. Essentially, gel slices were weighed first, then 2.5-3 vol of NaI stock solution was added. After about 5 min at 45-55°C, the gel slices were completely dissolved. Five μl of glassmilk for 5 μg or less of DNA (additional 0.5 μl for each 0.5 μg) was added to the tubes and the tubes were placed on ice for 5 min. To pellet the silica matrix with the bound DNA, the tubes were spun in a microcentrifuge for 5 sec. The pellet was washed three times with the supplied wash solution. Later, 10-20 μl of TE buffer was added to the tubes. After incubation at 45-55°C for 2-3 min, the DNA fragments were eluted from the Glassmilk by spinning for 30 sec in a microcentrifuge. Regular agarose and/or low-melting-point (LMP) agarose were used for conventional electroelution and/or Geneclean. In the case of LMP agarose, the DNA fragments could be recovered by melting the agarose at 65°C for 5 min, extracting with phenol and CIA, and precipitating with ethanol. The isolated DNA fragments were used for various purposes, e.g., making a probe, subcloning, etc.
Restriction Enzyme Mapping and Gel Electrophoresis

Plasmid, phage and high molecular weight genomic DNA were digested by restriction enzymes purchased from various companies. Generally, 1-10 units of enzyme was used for each 1 µg of DNA, and the vol of the enzyme used did not exceed one tenth of the final vol. After incubation at 37°C for 1-2 hr, the digested DNA fragments were separated by agarose gel electrophoresis in TAE buffer containing 1 µg/ml EtBr. The percentage of the gel used was determined by the sizes of the DNA fragments, and the gel size and running condition were decided based on purposes. After electrophoresis, the gel was photographed under UV light. The restriction enzyme maps were established by analysis of various enzymes digestions.

Cloning of DNA Fragments Into Vectors/Direct Cloning of cDNA Inserts

DNA fragments (inserts) were prepared by restriction enzyme digestion. After heat inactivation of the enzymes at 65°C for 5 min (for most of the enzymes), either the DNA fragments were used directly for "shotgun" cloning or a particular DNA fragment was band isolated and used for subcloning. Vectors were prepared similarly, but in some situations the cloning sites needed to be modified before ligation. To prevent re-circularization of the single-enzyme digested vectors, a dephosphorylation reaction was carried out using 2 µl (1 units/µl) calf intestine alkaline phosphatase (CIAP) in 100 mM Tris-HCl, pH 8.0. Several different molar ratios of vector to insert, including 1:1, were used for ligation. In general, the ligation reaction was performed using T4 DNA ligase in 20 µl total volume at 14°C for 16 hr. For blunt-end ligation, polyethylene glycol (PEG) was added to a final concentration of 15% and
the ligation reaction was carried out at 20°C instead of 14°C. Following transformation, the colonies were picked and the plasmid DNA was isolated from overnight culture and analyzed.

cDNA inserts in λgt11 vectors were subcloned directly into the endogenous pBR322 plasmid vector without adding exogenous plasmid vectors. The detailed procedures were described by Chiu and Lehtoma (1990). The phage DNA was isolated from infected Y1088 host cells and digested with EcoRI. The endogenous plasmid pMC9, which contained the 1.7-kbp lacZ and lacI genes inserted into EcoRI site of pBR322, was extruded from the bacteria during lysis and was co-purified, and co-digested with the phage DNA. The ligation was performed as usual except without adding any exogenous plasmid vectors. After transformation, colonies contained the cDNA insert was analyzed by gel electrophoresis.

**Preparation of Radiolabeled DNA Probes**

Double-stranded (ds) DNA probes were prepared using the nick translation method (Rigby et al., 1977) in the early stage of the work. One half μg of DNA was mixed with 5X buffer [250 mM Tris-HCl, pH 7.5, 50 mM MgSO4, 0.5 mM dithiothreitol(DTT) and 250 μg/ml BSA], 10 μl of DNA polymerase I (5 units)/ DNase I (100 pg), and 100 μCi of [α-32P-dATP] in 100 μl final vol. (Amersham). The mixture was incubated for 1-2 hr at 14-15°C. The reaction was stopped by adding 10 μl of 0.5 M EDTA, 10 μl of 20% SDS, and 880 μl of low salt buffer (LSB; 0.2 M NaCl, 20 mM Tris-HCl, pH 7.5, 1 mM EDTA). The mixture was passed through an Elutip column (S & S) to remove unincorporated nucleotides. After washing with 4 ml of LSB, the radiolabeled DNA probe was eluted by 0.6 ml high salt buffer
(HSB; 1.0 M NaCl, 20 mM Tris-HCl, pH 7.5, 1 mM EDTA). The probe was boiled for 5 min to denature the ds DNA and quick chilled on ice before use.

Later, random-primer labeling techniques were used (Feinberg and Vogelstein, 1984). Commercial kits were available from Amersham (multi-prime DNA labeling system) and Stratagene (Prime-it). Essentially 25-50 ng of DNA was denatured by heating to 95-100°C for 2-5 min, then chilled on ice. Buffer, unlabelled dNTPs (with dATP), random hexamer (6-mer), [α-32P] dATP and Klenow fragment were added. The tube was incubated for 30 min at 37°C or at least 3 hr at room temperature. Stratagene's Prime-it kit utilized random nonamer (9-mer) and T7 DNA polymerase, therefore the reaction could be finished in 2-10 min at 37-40°C. The reaction was stopped and desalted as for nick translation.

Southern (DNA) Blotting and Hybridization/Genomic Blots and Zoo Blots

Digested DNA was electrophoresed on agarose gels and analyzed by the method of Southern (1975). The gel was denatured in 0.5M NaOH/1M NaCl for 30 min by slow shaking. Following two washes with distilled water, the gel was neutralized in 1M Tris-HCl/0.6M NaCl for 45 min or longer. Then, the gel was shaken in 6X SSC for 15 min before blotting. The treated gel was laid onto the top of Whatman papers. A piece of nitrocellulose filter (S & S) and/or nylon membrane (Hybond-N or Hybond-N⁺ from Amersham) was placed on top of the gel. About 40 sheets of Blot Block (S & S) or paper towel were placed on top of the gel, then weighted with a 250-500 g object. The transfer solution was 6X SSC for nitrocellulose filter or 2X SSC for nylon filter. After blotting overnight, the nitrocellulose filter
was baked for 2 hr at 80°C in a vacuum oven and the Hybond-N nylon filter was UV-crosslinked for 2 min. No treatment was needed for the Hybond-N* nylon filter. The hybridization conditions employed included 6X SSC, 1X Denhardt's solution, 0.5% SDS, 100 µg/ml herring sperm DNA and 1 x 10^6 cpm/ml of the probe at 62°C. The filter was first washed with 2X SSC/0.1% SDS for 5 min twice at room temperature, followed by 0.1 X SSC/0.1% SDS for 15-30 min twice at 62°C.

Genomic DNA blots were treated as above, except different stringencies of conditions were used. Stringent condition included hybridization in 50% formamide, 5X SSC, 1X PE (50 mM Tris-HCl, pH 7.5, 5 mM EDTA, 0.1% sodium pyrophosphate, 1% SDS, 0.2% PVP, and 0.2% ficoll), 150 µg/ml herring sperm DNA and washing in 0.1 X SSC/0.1% SDS for 30 min twice at 65°C. Non-stringent condition included hybridization in 40% formamide, 6X salt-sodium phosphate-EDTA buffer (SSPE; 175.3 g NaCl, 27.6 g NaH₂PO₄·H₂O and 7.4 g EDTA, pH 7.4, in 1 liter of ddH₂O for 20X SSPE), 3% SDS, 0.5% carnation milk, 10% dextran sulfate and washing in 2X SSC for 30 min once at 65°C. Non-stringent conditions were used for "zoo" blots containing genomic DNA from several different species. The hybridized signal was scanned and analyzed using an LKB densitometer.

Genomic Library Screening

Four different human genomic DNA libraries were used for screening. Two human genomic libraries were constructed in Charon 4A using partial AluI-HaeIII digest and partial EcoRI digested DNA, respectively (Lawn, et al., 1978). A genomic library made from partial Sau 3A digested DNA and cloned into the SalI site of the λFix phage vector was generously provided
by Dr. Y.K. Fung, University of Southern California. A chromosome 5 specific library was obtained from American Type Culture Collection (ATCC), which was constructed by cloning the chromosome 5 DNA after complete HindIII digestion into Charon 21A (Van Dilla and Deaven, 1990). One million phage (represents 5 times the size of the human genome) from the library was plated onto ten 150 mm petri dishes (1 x 10^5 phages/dish) using LE392 as host cells. The plaques were lifted onto nitrocellulose filters, denatured, neutralized, and baked. Prehybridization and hybridization conditions were as described in cDNA library screening. When a cloned DNA probe was used, washing conditions followed the regular Southern blotting and hybridization procedures, except at 55°C.

**Purification, Desalting and Radiolabeling of Oligonucleotides**

After completion of the oligonucleotide synthesis, the chemically protected groups were removed by heating at 55°C for 6-12 hr. The oligonucleotide was then dried down in a Speedvac (Savant Instrument Inc., Farmingdale, NY). Gel exclusion chromatography was used to desalt the detritylated oligonucleotide. Briefly, the lyophilized sample was redissolved in 0.2 ml 10 mM triethylammonium bicarbonate, pH7.0 (TEAB), and then loaded onto Sephadex G50 (fine) column. The larger oligonucleotide came out of the column first as monitored by absorbance at 260 nm. The size of the oligonucleotide could be checked by running 1 μg of oligonucleotide on a 12% polyacrylamide gel and staining with 0.02% methylene blue. Some oligonucleotides were purified by electrophoresis on a polyacrylamide gel and eluted overnight from the excised, crushed gel. The OPC cartridge (ABI) was used to purify tritylated oligonucleotides.
The manufacturer's procedures were followed.

Oligonucleotide probes were labelled at their 5'-ends with T4 polynucleotide kinase. Half µg of oligonucleotide was mixed with 10 X exchange buffer (0.5 M Tris-HCl, pH 7.6, 0.1 M DTT, and 0.5 M MgCl₂), 5 µl (11 units/µl) T4 polynucleotide kinase, and 500 uCi [γ-³²P]dATP. The reaction was incubated for 1 hr at 37°C. The reaction was stopped by 10 µl of 0.2 M EDTA and loaded onto DE52 column. The DE52 column was first equilibrated with 10 ml T₁₀E₁₀N₁₀₀ (10 mM Tris-HCl, pH 8.0, 10 mM EDTA, and 100 mM NaCl), and rinsed with 4 ml of T₁₀E₁₀N₂₀₀ (10 mM Tris-HCl, pH 8.0, 10 mM EDTA, and 200 mM NaCl). The oligonucleotide probe was collected with 3 x 1 ml T₁₀E₁₀N₆₀₀ (10 mM Tris-HCl, pH 8.0, 10 mM EDTA, and 600 mM NaCl). To monitor the efficiency of the reaction further, DE81 filters were used as described in Sambrook et al. (1989). DE81 filters were positively charged and they absorbed oligonucleotides strongly that were too small to be precipitated efficiently with trichloroacetic acid (TCA).

To make the 63-base pair (bp), second-exon-specific oligonucleotide probe, two oligonucleotides with seven complementary bases at their respective 3'-ends were annealed in the fill-in reaction and labeled to high specific activity (approximately 5 x 10⁸ CPM/µg) by the random-primer labeling technique (Feinberg and Vogelstein, 1983). Briefly, the two oligonucleotides were mixed with 10X blunting buffer [80 mM MgCl₂, 100 mM β-mercaptoethanol, 2 mM each deoxynucleotide triphosphate (dNTP), and 10 mM adenosine 5'-triphosphate (ATP)], 5 µl (5 units/µl) Klenow fragment and 450 uCi[α-³²P] dATP. After incubation at 37°C for 1 hr, the reaction was stopped by 10 µl of 0.2 M EDTA. The reaction product was diluted with T₁₀E₁₀N₁₀₀ before loading onto a DE52 column.
Isolation of Single-Stranded (ss) DNA by Superinfection

ss DNA was rescued from bacteria containing pBluescript phagemid by superinfection with helper phage VCS M13 or R408. The procedures were followed according to the Stratagene manual with some modifications. Twenty ml of 2 X YT (16 grams bacto-trypton, 10 grams bacto-yeast extract, and 5 grams NaCl, pH 7.0, in 1 liter of ddH2O) was inoculated with 0.8 ml of an overnight bacterial culture containing pBluescript phagemid. The culture was shaken at 37°C in a 250 ml flask until OD₆₀₀ = 0.3 to 2.5 x 10⁸ bacteria/ml (about 30 min). Then 80 /µl of helper phage (1 x 10¹¹ pfu/ml) was added to the flask. The culture was continuously shaken for 10-16 hr. After centrifugation, the supernatant was treated with 1/4 vol 3.5 M ammonium acetate, pH 7.5/20% PEG and placed on ice for 20 min. ss DNA was pelleted by centrifuge at 10,000 rpm for 10 min. The pellet was resuspend in 200 /µl TE buffer and was extracted with equal vol of phenol/CIA five times, then CIA alone once. The ss DNA was precipitated by adding 150 /µl 7.5 M ammonium acetate, pH 7.5, 600 /µl of ethanol and centrifuge. The amount of the ssDNA was verified by electrophoresis on an agarose gel containing ethidium bromide.

Manual DNA Sequencing and Sequence Comparison

The DNA sequence was determined by the chain termination method of Sanger et al. (1977) using the Sequenase kit from USB (Cleveland, OH). During early stages of the project, ss DNA was used as template for sequence analysis by using T7 primer, T3 primer (Stratagene), or other designed synthetic oligonucleotide primers. During later stages, double-stranded (ds) DNA from minilysate or CsCl gradient preparation was used
predominantly. The ds DNA template was prepared for sequencing as follows: 5 µg of ds DNA was denatured in 0.2 M NaCl/0.2 mM EDTA at 37°C for 30 min. Denatured DNA was precipitated by 0.1 vol NaOAc and 2.4 vol ethanol. Then, 5 ng of primer was hybridized with 1 µg of ss DNA or 3-5 µg of ds DNA and the sequencing reaction was carried out following the manufacture's protocol.

Recently, sequencing analysis was performed on an automatic DNA sequencer (Model 370A, ABI) as described later. Sequencing projects were managed with the SEQMAN program of DNA Star, Inc. (Madison, WI). The coding regions in human genomic sequences were identified by alignment with cDNA sequences using ALIGN program (Wilbur and Lipman, 1983). The DOTPLOT program was used to compare DNA sequences from different species.

**Unidirectional Deletion Cloning**

Majority of the DNA sequence was determined from clones generated by the unidirectional Exonuclease III/ mung bean nuclease deletion cloning (Henikoff, 1984). The deletional cloning kit was from Stratagene. Essentially, CsCl banded pBluescript containing insert was digested to completion with a unique 3'-overhang restriction site and a unique 5'- or blunt restriction site that lies between the insert and the chosen 3'-site. The linearized plasmid was then digested with 20 units of Exonuclease III at 37°C for various time intervals. Later, mung bean nuclease was added to generate blunt-ended molecules. The molecules were ligated, transformed and mini-prep plasmid DNA was analyzed to verify the degree of deletion. Clones containing appropriately spacing deletions were picked, and used for sequencing analysis by universal primer to
reduce the cost and time spent on synthesizing oligonucleotide primers.

**Automatic DNA Sequencer and Cycle Sequencing**

Some of the sequencing analysis was performed on an automated DNA sequencer (Model 370A, Applied Biosystems) and ds DNA was used as a template with Taq polymerase. The template denaturation was similar to manual sequencing. Then the template was divided into 4 different tubes and annealed with dye-labeled FAM, JOE, TAMRA, and ROX primers in a primer-to-template molar ratio of 3.5 : 1. The extension, enzyme "chase", ethanol precipitation, preparation for loading procedures, and electrophoresis were done according to the manufacture's protocol.

Recently, this fluorescence-based sequencing with Taq polymerase was linked to the usage of thermal cycler, therefore the entire procedures were automatically performed. This so called "cycle sequencing" procedures were essentially as follows: First of all, the template was mixed with dye primers and Taq polymerase in 4 different tubes. Then, the mixture was put into a thermal cycler and put through twenty-seven cycles of 98°C for 2 sec, 55°C for 15 sec and 65°C for 2 min. After electrophoresis on an automatic sequencer, routinely more than 250-bp sequence could be determined.
Genomic DNA Polymerase Chain Reaction (PCR)

A pair of PCR oligonucleotide primers, forward primer and reverse primer, were used. The PCR reaction was carried out in a final volume of 100 µl containing 12.5 ng to 1 µg of genomic DNA (isolated from tissues, cell lines, or blood cells), 25 pmol of each oligonucleotide, 200 µM dNTPs, and 10 µl of 10× reaction buffer [100 mM Tris-HCl, pH 8.3 at 25°C, 500 mM KCl, 15 mM MgCl₂, 0.01% (w/v) gelatin]. The PCR reaction was performed using a Perkin-Elmer DNA thermal cycler (Perkin Elmer Cetus, Norwalk, CT). Prior to adding 2.5 units of Taq polymerase, the reaction mixture was heated at 95°C for 5 min and then allowed to cool to 72°C. After adding Taq polymerase, the denaturation step was carried out at 94°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 2 min for a total of 40 cycles. The reaction mixture was extended at 72°C for 15 min and then stored at 4°C for further characterization. Amplified DNA was subjected to agarose gel electrophoresis. After electrophoresis, PCR products were directly visualized under UV after ethidium bromide staining. Later, restriction enzyme digests were used to confirm the amplification of the expected product. The product was subcloned (see later section) and used for further analysis.

Several improvement procedures to increase the specificity of PCR were used according to the actual situation. 1-10% of denaturant dimethyl sulfoxide (DMSO), 5-15% of glycerol, 1 x 10⁻⁵ M to 1 x 10⁻⁴ M of tetramethylammonium chloride (TMAC), and/or 1.25%-10% of formamide were tested when needed.
Reverse Transcription and PCR (RT-PCR)

The reverse transcription reaction was as follows: 2 µg of RNA was mixed with 2 µl of oligo (dT)12-18 (500µg/ml). The mixture was heated to 70°C for 10 min and quick chilled on ice for 2 min. The contents of the tube was then collected by brief centrifugation for 10 to 20 sec. After that, 4 µl of 5X RT buffer (250 mM Tris-HCl, pH 8.3 at room temperature, 375 mM KCl,15 mM MgCl₂), 2 µl of 0.1 M DTT, and 4 µl of 2.5mM dNTP were added to the tube. The tube was placed at 37°C for 2 min to equilibrate the temperature. Finally, 2 µl of Moloney murine leukemia virus RNase H\(^{-}\) reverse transcriptase [M-MLV H- RT (Superscript) from BRL; 200 units/µl] was added and the tube was incubated at 37°C (up to 40°C) for 1 hr. Amplification was similar to that of the genomic DNA PCR, except that the template used was the reverse transcribed first-strand of cDNA from RNA. The cycling program was 94°C for 30 sec, 50° or 55°C for 45 sec, and 72°C for 2 min for a total of 50 cycles.

The GeneAmp RNA PCR kit from Perkin-Elmer Cetus was also used. Supplied protocols for reverse transcription of RNA and PCR amplification of cDNA in a single reaction tube were precisely followed. To eliminate possible presence of genomic DNA in the RNA sample, the reverse-transcribed cDNA products was digested with a restriction enzyme which is positioned between the two primers before adding Taq polymerase for the PCR. The cDNA was amplified using a thermal cycler (Perkin-Elmer Cetus) for 35 cycles: 94°C, 1 min; 60°C, 1 min. The PCR products were analyzed on a 3% Nusieve and 1% Seakem agarose gel and visualized by staining with ethidium bromide.
Anchored PCR or Rapid Amplification of cDNA Ends (RACE)

Anchored PCR, also known as Rapid Amplification of cDNA Ends (RACE), was used to generate cDNAs by PCR amplifying copies of the region between a specific point in the transcript and the 5'-end (or 3'-end). The procedures were followed essentially as described by Frohman et al. (1988) with some modifications. In reverse transcription reactions, 5 μg of total RNA was mixed with other reagents including 1 pmol of a gene-specific primer. A homopolymer of dA was then appended to the first-strand reaction products using terminal transferase. Finally, amplification was accomplished using 2 pmol of adaptor-(dT)$_{17}$ hybrid primer, 25 pmol of adaptor primer and 25 pmol of a second, internal gene-specific primer. Sometimes, adding a second round of amplification using nested primers added an additional level of specificity and amplification.

Direct PCR Screening of Colonies and Plaques

Plasmid screening from colonies, circumventing DNA preparation, could be carried out by first lightly touching the edge of the colonies with the tip of the toothpick. Colonies were swirled into 0.5 ml ddH$_2$O and boiled for 5 min. After centrifugation for 2 min, 5 μl of the supernatant was used for 30 rounds of temperature cycling: 94°C for 1 min, 55°C for 1 min and 72°C for 2 min. (Gussow and Clackson, 1989).

Phage were transferred by gently stabbing the center of a well-isolated plaque with a toothpick, and swirling the toothpick briefly in a 20 μl PCR reaction mix. PCR was performed directly.
Cloning PCR Products

For directional cloning, Taq polymerase was inactivated with phenol/CIA extraction and ethanol precipitation. Then, the PCR amplified products were digested with appropriate restriction enzymes, run on a gel, and purified by electroelution or Geneclean. The subcloning procedure was the same as regular subcloning.

For blunt-end cloning, the PCR buffer and excess primers were removed by ethanol precipitation, three rounds of Centricon 100 filtration (Amicon, Danvers, MA) or gel purification. Then, the products were blunt-ended using T4 polymerase. The blunt-ended products were ligated in the presence of PEG to the vector previously digested with SmaI or other restriction enzymes which generate blunt-end. The recombinants were identified with alkaline mini-lysate, boiling mini-prep (for DH5α cells), or colony hybridization/microwave colony screen screening.

The TA cloning kit from Invitrogen Corp. (San Diego, CA) was also used for direct insertion of PCR products into a plasmid vector. This method took advantage of the non-template dependent activity of Taq polymerase that adds (dA)s to the 3'-end of PCR duplex products. The vector used provides (dT) overhangs.
RESULTS

SECTION I. MOLECULAR CLONING AND CHARACTERIZATION OF HUMAN HBGF-1 cDNA

1.1. Screening Strategy

The DNA probe used was a 33-base synthetic oligonucleotide based on the published sequence coding for human HBGF-1 (Jaye et al., 1986). Specific oligonucleotide probe is better than degenerate oligonucleotide probe because the latter contains mixed sequences with different GC content and therefore different melting temperature (Tm). For this reason, a specific probe was chosen to screen a brain stem cDNA library. Brain tissues are known to contain large amounts of HBGF-1. Poly (A+) RNA isolated from a 1-day-old human neonatal brain was used to construct a cDNA library in λgt11 (de Ferra et al., 1985). The cDNA library was generously provided by Carmie Puckett (California Institute of Technology, CA).

1.2. cDNA Clones and Restriction Enzyme Mapping

One million phage clones from the brain stem cDNA library plated out on E. coli Y1088 were screened using a kinase-labeled 33-base oligonucleotide probe under non-stringent conditions. Two positive plaques were identified. These positive phage clones were further purified through three or four cycles of rescreening. The cDNA inserts were subcloned and designated pHBGF-1.1 and pHBGF-1.2, respectively (Fig. 5). Two additional
Figure 5. Restriction enzyme map of human HBGF-1 cDNA clones. $1 \times 10^6$ phages from a human brain stem cDNA library constructed in λgt 11 were plated on a lawn of *E. coli* Y1088 and screened with a synthetic 33-mer, 5'CTTGT CCCAT CCACT GTGCC ATCCG GAAGG ATC3'. The open box represents the open reading frame. The ends of the each clone correspond to synthetic EcoRI linkers used in the construction of the cDNA library (de Ferra et al., 1985). The 3'-end of the pHBGFl.1 cDNA insert contains a AATAAA polyadenylation sequence and poly (A) tail.
clones, pHBGF-1.4 and pHBGF-1.5, were subsequently isolated and subcloned from the same cDNA library using the cDNA insert from pHBGF-1.1 as a probe, (Fig. 5). The DNA inserts from these four phage clones were isolated and subcloned into plasmid vectors by standard methods (Maniatis et al., 1982) or directly cloned into pBR322 without providing exogenous plasmid DNA as described (Chiu and Lehtoma, 1990a; Chiu et al., 1991). The plasmid DNA was subjected to restriction enzyme mapping analysis by digestion with various enzymes. The restriction enzyme maps of each of the cDNA inserts were aligned based on the common restriction enzyme sites. Clones pHBGF-1.1 and 1.2 were similar to those previously reported (Jaye et al., 1986). pHBGF-1.4 and 1.5 appear to extend further toward the 5'-end of the cDNA. The total length compiled from these cDNA inserts was about 2.35-kbp (Fig. 5).

Restriction enzyme analysis also showed that there is a PstI site near the 5'-end of pHBGF-1.4 but not in pHBGF-1.5. These data implied that these two clones had different upstream sequences (see below).

1.3. cDNA Sequence Analysis

The cDNA sequences were determined by the enzymatic chain termination method (Sanger et al., 1977). The sequencing strategy was shown in Fig. 6 and the complete cDNA sequences was shown in Fig. 7. It consists of 2342-bp which includes 465-bp of HBGF-1 coding sequence. Differences between this sequence and the previously published, partial sequence (Jaye et al., 1986) are an G residue insertion at nucleotide 518 and a deletion of a C residue between nucleotide 566 and 567 in Fig. 7. For these
Figure 6. Sequencing strategies for the determination of entire sequence of cDNA clones. The open box represents the coding sequence for human HBGF-1. The short horizontal arrows indicate the sequencing strategies used to sequence both strands of the 2.35-kb region. They also indicate the length and direction of the sequences that were determined. Close circles indicate sequencing by M13, reverse M13, T7, or T3 primers. Open squares indicate sequencing by designed oligonucleotide primers.
Figure 7. Nucleotide and deduced amino acid sequences of human HBGF-1 cDNA clones. The plasmid containing the HBGF-1 cDNA inserts was sequenced by the chain termination method. The first nucleotide of the sequence corresponds to the 5'-end of the first protein coding exon (Wang et al., 1989; see later). The 5'-end of pHBGF 1.1 and pHBGF 1.2 started at nucleotide 43 and 9, respectively, while the 5'-end of pHBGF 1.4 and pHBGF 1.5 extended 59-bp and 83-bp further upstream from the 5'-end of the first protein-coding exon. The upstream sequences of pHBGF 1.4 and 1.5 are indicated with a solid line and a dashed line, respectively. Three triangles indicate the exon junctions. The polyadenylation signal sequence, AATAAA, is indicated with a wavy line. This DNA sequence has been submitted to Genbank; the accession number is X51943.
discrepancies, this cDNA sequence is identical to the corresponding genomic DNA sequence as shown later (Wang et al., 1989; Fig. 22). The 3' end of pHBF-1.1 cDNA insert contained a polyadenylation sequence, AATAAA, as indicated with a wavy line, and a poly (A) tail of seventeen A residues.

1.4. Alternative Splicing Generates Two Forms of HBGF-1 mRNA—pHBGF-1.4 and pHBGF-1.5

Sequence analysis of pHBF-1.4 and 1.5 revealed that sequence downstream from nucleotide 1 in Fig. 7 were identical to each other and also to the sequences determined for pHBF-1.1 and 1.2. Nucleotide 1 was established by ribonuclease protection analysis to be the 5'-end of the first protein-coding exon, as shown later. However, pHBF-1.4 and 1.5 differed from each other upstream from nucleotide -10.

When the cDNA sequence was compared with genomic DNA sequence, the difference was observed immediately upstream from the nucleotide 1 (the 5'-end of the first protein-coding exon). The upstream sequences of these two clones probably represent two different upstream untranslated exon which were spliced alternatively to the first protein coding exon.

A postdoctorate in the lab, Dr. R. Myers, used a plasmid derived from pHBF-1.5, designated as pHBF-1.5ABam, as a template to generate a riboprobe by in vitro transcription for ribonuclease protection assays. Fully protected fragments were identified in the mRNA isolated from adult human brain and several glioblastoma cell lines (Myers, R., Wang, W.-P. and Chiu, I.-M., unpublished results). Human kidney and some glioblastoma cell lines contained lesser amounts of pHBF-1.5 mRNA. A large portion of
the riboprobe was protected up to the junction of the upstream untranslated exon and first protein-coding exon. These results indicate that different upstream untranslated exons are preferentially spliced to the coding region in different tissues and alternative splicing generated at least two different forms of HBGF-1 mRNA.

However, when Dr. Myers used pHBF-1.4ΔBam, a pHBF-1.4 derived plasmid, as a template for similar RNase protection assays, fully protected fragment were not detected. Tissues expressing pHBF-1.5 upstream untranslated exon showed protected fragments up to ten nucleotides beyond the exon junction since the pHBF-1.5 sequence was identical to that of the pHBF-1.4 up to nucleotide -10. Therefore, if the pHBF-1.4 transcript was present, it probably exists at an extremely low level.

Recently, several genomic clones containing the pHBF-1.5 upstream untranslated exons were isolated and characterized. An oligonucleotide derived form pHBF-1.4, HBGF 1003 (nt. -50 - -18 in Fig. 7), was used as a probe and hybridized to the upstream genomic clones. The hybridization region was very close to the pHBF-1.5 upstream region. Later, it was found that the identical sequence of pHBF-1.4 was present in the opposite strand and was 22 nucleotides upstream from the most 5'-end of pHBF-1.5 cDNA clone. All of these information indicates that the pHBF-1.4 most likely represents a cloning artifact.

### 1.5. HBGF-1 mRNA Expression in Human Brain, Kidney and Heart

HBGF-1 protein was isolated from bovine kidney (Gautschi-Sova et al., 1987), human brain (Gautschi-Sova et al., 1986; Gimenez-Gallego et al., 1986; Harper et al., 1986), smooth muscle cells (Winkles et al.,
1987), and various glioma cell lines (Libermann et al., 1987). Using the 2.2-kbp EcoRI insert isolated from pHGF-1.1, Northern blotting and hybridization was used to determine the presence of HBGF-1 mRNA in various fetal tissues.

Sample RNA was isolated from the lung, heart, intestine, liver, brain and placenta of a 13-week-old fetus. A 4.5-kb HBGF-1 transcript was detected in the RNA isolated from heart and brain but not in that from the lung, intestine, or liver (Fig. 8). The 4.5-kb hybridizing signal was not due to 28S ribosomal RNA nonspecific cross-hybridization because when poly (A)+ RNA was used for Northern analysis, the same 4.5-kb message was detected. Surprisingly, fetal heart expressed a higher level of HBGF-1 mRNA than fetal brain. HBGF-1 expression in the fetal placenta could not be detected even when as much as 5 μg of poly (A)+ RNA was used. RNA from adult kidney was also included in the same filter (Fig. 8). A slight, yet reproducible variation in the mobility of the 4.5 kb HBGF-1 mRNA isolated from brain and kidney was observed. This phenomenon might reflect the difference in the upstream exons, as discussed previously, because ribonuclease protection analysis indicated that they all use the same major polyadenylated sites, as shown later. The same filter was hybridized to a human β-actin cDNA probe (Ponte et al., 1984) to demonstrate that the equal amount of RNA was used for each sample (Fig. 8). The RNA coding for isoforms of actin were detected in the fetal heart tissue, such that β- and γ-actin mRNA (2.4-kb) were present in all tissues while α-actin mRNA (1.6-kb) was present only in fetal heart.
Figure 8. Northern blotting and hybridization analysis of human fetal RNA with HBGF-1 and \( \beta \)-actin cDNA probes. Twenty \( \mu \)g of total cellular RNA were denatured in 2.2 M formaldehyde and 50% formamide and fractionated by gel electrophoresis in a 1.1% agarose gel containing 2.2 M formamide. In the case of fetal placenta, 5 \( \mu \)g of poly(A)\(^+\) RNA were analyzed. RNA was transferred to a nitrocellulose filter and hybridized to human HBGF-1 cDNA probe (top panel). The filter was then washed and exposed overnight to X-ray film with two intensifying screens. The same filter was then hybridized to human \( \beta \)-actin cDNA probe and exposed for 20 min (lower panel). The 28S (4.5-kb) and 18S (1.9-kb) were used as molecular weight makers. The RNA coding for isoforms of actin were also observed in the human fetal heart sample (lane 2), such that \( \beta \)- and \( \gamma \)-actin mRNA are 2.1-kb while \( \alpha \)-actin mRNA is 1.6-kb.
Crumley et al. (1990) published a HBGF-1 cDNA clone isolated from human brain stem cDNA library, designated as pHBGF-1.J, which has a 324-bp different upstream untranslated exon from pHBGF-1.4 and pHBGF-1.5. They also isolated a genomic clone containing the 324-bp cDNA sequence. A canonical splice donor site on the 3' side of this upstream exon and a splice acceptor site on the 5' side of the first coding exon were detected. In order to verify these clones and to generate probes from this region, novel methods using polymerase chain reaction (PCR) were adopted.

First, I carried out PCR analysis on genomic DNA using two designed oligonucleotides, HBGF 1201/1301 (for sequences see Fig. 9, Panel B), according to the published sequences. The expected 324-bp HBGF-1 cDNA fragments of pHBGF 1.J were observed when different amounts of genomic DNA templates were used (Fig. 9, panel A). This PCR product was digested with PstI to confirm the amplification of the correct product. Then, the PstI-digested fragment (229-bp) was further subcloned into PstI and SmaI-digested pBluescript KS(+) vector and designated as 1.JAPstl. White colonies were picked and DNA was isolated using boiling mini-prep procedure. The miniprep DNA was used as a template for sequencing. The sequence was identical to the published sequence (Fig. 9, panel B; Crumley et al., 1990).

The 1.JAPstl PCR product of 229-bp in length was used as a probe for genomic Southern blot analysis (Fig. 10). The 1.JAPstl PCR probe hybridized to 2.4-kbp EcoRI, 5.0-kbp BamHI, 9.6-kbp BglII and 1.4-kbp
Figure 9. Amplification and nucleotide sequence of upstream untranslated exon of pHBGF1.J of human HBGF-1, 1.JA/PstI.

Panel A: Amplification of upstream untranslated exon of pHBGF1.J, 1.JA, using two designed oligonucleotide primers, HBGF 1201/HBGF 1301. Samples of 100 µl contain standard buffer, 200 µM each dNTP, and 25 pmol each primers. Human WBC genomic DNA 1 µg (lane 2), 500 ng (lane 3), 100 ng (lane 4), 50 ng (lane 5), 25 ng (lane 6), 12.5 ng (lane 7), and 2.5 units of Taq polymerase were subjected to 40 cycles of amplification. Each sample (10 µl) was resolved on a 3% NuSieve GTG agarose gel plus 1% Seakem GTG agarose gel and visualized by ethidium bromide fluorescence. Lane M: Markers are HindIII-digested phage λ (500 ng) and HaeIII-digested φX 174-RF (250 ng). Lane 1: Control template and control primers supplied by Perkin-Elmer Cetus. Arrow indicates the amplification products of 324-bp.

Panel B: The nucleotide sequence of upstream untranslated exon of PCR cDNA clone 1.JA/PstI. For directional cloning, Taq polymerase was inactivated with phenol/CIA extraction and ethanol precipitation. The 324-bp fragment was digested with PstI and subjected to GeneClean purification. The digested fragment was subcloned into Smal-PstI-digested pBluescript vector and sequenced by the chain termination method of Sanger et al. (1977). PCR primers, HBGF 1201 and HBGF 1301, are underlined. Two PstI sites are also indicated. Triangle indicates the junction of upstream untranslated exon and first protein-coding exon. Numbers given are relative to the downstream initiator ATG.
A. bp

<table>
<thead>
<tr>
<th>M</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1078</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>872</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>603</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>310</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>281/271</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>234</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>194</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>324 bp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B.

-324 ATCCCAA GGCTAGGAGG ACACACTCT AACCAGGTGGG TGCGTAGGT GTGCGTTC
→HBGF1201

-267

-207

-147

PstI

-87

PstI

-27

-1

Figure 9
Figure 10. Applications of PCR cDNA clone 1.4PstI.

Panel A: Genomic Southern blotting analysis of human placenta DNA with the 229-bp 1.4PstI PCR probe. The genomic DNA were digested with EcoRI (lane 1), BamHI (lane 2), BglII (lane 3), HindIII (lane 4), and then hybridized to the 229-bp 1.4PstI PCR probe. The image was processed with a Betagen beta scanner (Waltham, MA) for 4 hours. The arrowheads indicate the hybridizing bands.

Panel B: Conservation of upstream untranslated exon of pHGFl.4 among primates. Genomic DNA was prepared from fetal rhesus monkey lung cells (lanes 1, 2 and 5), human placenta (lane 3), chimpanzee skin cells (lane 4), and baboon cells (lane 6). Genomic DNA was digested with EcoRI except DNA in lanes 1 and 2, which was digested with PvuII and PstI, respectively. The digested DNA was blotted and hybridized to the 229-bp 1.4PstI PCR probe. The arrowheads indicate hybridizing bands.

Panel C: Screening of EcoRI restriction polymorphisms with the 229-bp 1.4PstI PCR probe. Genomic DNA was isolated from eight individuals and digested with EcoRI. The digested DNA was hybridized to the 229-bp 1.4PstI PCR probe. The arrowheads indicate three alleles with bands at 3.2-kbp, 2.4-kbp and 1.0-kbp, respectively.
Figure 10
HindIII fragments (Fig. 10, Panel A). This upstream untranslated exon probe also cross-hybridize to genomic DNA from other primates, chimpanzee, rhesus monkey and baboon (Fig. 10, panel B, lanes 4-6). According to a previous report by Heitz et al. (1990), a similar probe detected 3.2-kbp (11.6%), 2.35-kbp (82.9%) and 1.0-kbp (5.5%) alleles. Genomic DNA from eight individuals was screened using this upstream exon probe. The similar EcoRI polymorphism was observed: 3.2-kbp, 2.35-kbp and 1.0-kbp (Fig. 10, panel C). Human HBGF-1 gene has been located to chromosome 5q 31.3-33.2 (Jaye et al., 1986; Huebner et al., 1990). I have combined this information in a sequence-tagged site (STS), designated FGFA.7/5q31-33 for inclusion in the human genome map (Olson et al., 1989). The above results demonstrate the direct utility of the PCR product as a probe for retrieving clone containing this STS from libraries or total genomic DNA (Wang, W.-P. and Chiu, I.-M. manuscript submitted). R. Payson used this 1.JAPstI PCR probe to hybridize to recently identified phage and cosmid clones. He was able to localize the hybridized region which is about 10.5-kbp upstream from the 1.5 exon (see summary in Fig.38). Sequencing of this region further confirmed this assignment and complete sequencing of this region is in progress. Ribonuclease protection assay and primer extension using this specific upstream exon are also currently under investigation.

Reverse transcription-PCR (RT-PCR; Verus et al., 1987) was performed to clone the cDNA fragment containing the upstream untranslated region and the 5'-end of the downstream protein-coding region using HBGF 1201 and HBGF 151 (across the junction of first and second protein-coding exons at nt. 188-210 in Fig. 7). A 540-bp fragment designated as 1.JAPVU was
amplified from kidney RNA (Fig. 11, panels A and B). A shorter product of 307-bp was also amplified because the upstream primer non-specifically primed to an internal site, which was confirmed by sequence analysis (Fig. 11, panels A and B). These PCR products were first digested with BamHI to confirm amplification of the correct products. Once confirmed, the PCR products were subcloned into pCR1000 vector using TA cloning kit. Microwave colony hybridization was used to screen recombinants containing inserts (Fig. 11, panel C). Sequence analysis confirmed the identity of the entire 540-bp. Dr. Myers used these clones as templates to synthesize riboprobes for protection assays. She showed that adult ribonuclease kidney predominantly expresses 1.J upstream exon, whereas adult human brain predominantly expresses 1.5 upstream exon.

Another novel HBGF-1 cDNA, designated as pHBF-1.H, was isolated from human prostate carcinoma cell line, LNCaP, which contained another upstream exon sequence (S. Harris, W. Alton Jones Cell Science Center, NY, personal communication). Using two designed oligonucleotides, HBGF 1501/1601 (sequences see Fig. 12, Panel D), this 90-bp upstream untranslated exon probe was generated by genomic DNA PCR (Fig. 12, panel A). An amplified product of primer dimer was observed when lesser amounts of genomic DNA template was used (Fig. 12, panel A, lanes 4-7).

PCR buffer containing 5% formamide enhanced the specific product of 90-bp even when a lesser amount of genomic DNA template was used (Fig. 12, panel B, lanes 2-3). This PCR product was subcloned directly into pCR 1000 vector, designated as 1.HA, and direct PCR screening of white colonies was used to screen recombinants containing inserts (Fig. 12, panel C). A modified mini-prep procedure was used to prepare the template
Figure 11. Reverse transcription-polymerase chain reaction of 1.JaPvu.

Panel A: One μg of total RNA isolated from human brain (lane 1) or human kidney (lane 2) was reverse transcribed into cDNA and subsequently amplified using upstream primer derived from upstream exon pHBGF1.J, HBGF 1201, and downstream primer derived from first protein coding exon, HBGF 151. Arrow indicates the expected fragment of 540-bp. Arrowhead indicates one of the internal primed products, 307-bp.

Panel B: Re-amplification of 1 μl of samples from panel A (1:20 dilution).

Panel C: Microwave colony screening of recombinants carrying the 540-bp and 307-bp fragment inserts. The 307-bp major fragment of panel A, lane 2, and 540-bp major fragment of panel B, lane 2, were excised from agarose gel and subcloned into pCR1000 vector using TA cloning kit. The white colonies were picked onto nylon membranes and allowed to re-grow. The colony lysis, DNA denaturation and fixation were carried out in a single step using microwave. The filters were hybridized to a 176-bp EcoRI-BamHI fragment from HBGF-1 cDNA.
Figure 11
Figure 12. Amplification and nucleotide sequence of upstream untranslated exon of pHBGF1.H of human HBGF-1, 1.Ha, using two designed oligonucleotide primers, HBGF 1501/HBGF 1601.

Panel A: Amplification of upstream untranslated exon of pHBGF1.H, 1.Ha. Human WBC genomic DNA 1 µg (lane 2), 500 ng (lane 3), 100 ng (lane 4), 50 ng (lane 5), 25 ng (lane 6), and 12.5 ng (lane 7) were used as templates for amplification. Lane M: Markers are HindIII-digested phage λ (500 ng) and HaeIII-digested φX 174-RF (250 ng). Lane 1: Control template and control primers supplied by Perkin Elmer Cetus. The condition used was described in MATERIAL AND METHODS. Arrow indicates the amplification products of expected 90-bp fragment. Arrowhead indicates the non-specific amplification products of 60-bp.

Panel B: Amplification of upstream untranslated exon of pHBGF1.H, 1.Ha, using PCR enhancer. Human WBC genomic DNA 500 ng (lane 1), 100 ng (lane 2), 25 ng (lane 3) were used as templates for amplification. The condition used was identical to that of panel A, except that the PCR buffer contained 5% formamide.

Panel C: Direct PCR screening of colonies carrying the 90-bp (1.Ha) and 60-bp fragment inserts. The sample of panel B, lane 1, was subcloned into pCR1000 vector using TA cloning kit. The white colonies were swirled into 0.5 ml ddH2O and boiled for 5 min. Five µl of the supernatant was used for PCR using the same pair of primers as in panel A.

Panel D: The nucleotide sequence of PCR cDNA clone 1.Ha. A modified procedure, according to Jones and Schofield (1990), was used to prepare the DNA template for sequencing in 30 min. PCR primers, HBGF 1601 and HBGF 1501, are underlined. Positions of identity between a hamster HBGF-1 cDNA clone (Hall et al., 1990) and human clone are marked with open circles. Triangle indicates the exon junction. Numbers given are relative to the downstream initiator ATG.
I K -103 GAATTGC GGCCGCCA.GA CTCTTGGGGG ATTCCTTAGT GAGTGAGTTC ACTGCTCMA -47

HBGF1601

Figure 12

-103 GAATTCG GGCGCGAGA CTCTTGGGGG ATTCCTTAGT GAGTGAGTTC ACTGCTCACAA -47

HBGF1601

-46 GAGGGCTTT GCCACTTCTG CAGGGAAGCC AGCCACGGGC CAGCAGV -1

Figure 12
for sequencing. The sequence was identical to the pHBGF-1.H sequence (Fig. 12, Panel D). R. Payson used this established PCR condition and the HBGF 1501/1601 primers of 1.HA to amplify DNA isolated from several upstream phages and cosmid clones. Payson localized this upstream untranslated exon to be about 31.4-kbp upstream from the first protein-encoding exon, thus, established the relationship of upstream exons as 5'-1.J-1.5-1.H-3' (See summary in Fig. 38).

RT-PCR was performed to clone the cDNA fragment containing upstream untranslated region of pHBGF-1.H and downstream protein-coding region using HBGF 1601/151 primers. Initial attempts used RNA isolated from the same prostate carcinoma cell line LNCaP and several different prostate carcinoma tissues as template failed. Therefore, the screening was extended to glioblastoma cell lines U1242MG and D65MG, which were shown to be non-1.5, non-1.J HBGF-1 producers by Dr. R. Myers using RNase protection assays. A 300-bp fragment of expected size, designated as 1.HAPvU, were amplified from U1242MG and D65MG using oligo (dT) as primers for reverse transcription (Fig. 13, panel A, lanes 4 and 7). These PCR products were digested with BamHI to confirm the amplification of the correct products (Fig. 13, panel B). These PCR products were then subcloned, designated as 1.HAPvu, and sequenced. This clone was used to generate a riboprobe to examine the abundance and tissue distribution of this upstream untranslated exon. The preliminary results obtained by Dr. R. Myers indicated that 1.H upstream exon was expressed at low level in the tissues screened. Several glioblastoma cells lines which express HBGF-1 mRNA did not express 1.J, 1.5, or 1.H upstream exons. This
Figure 13. Reverse transcription-polymerase chain reaction of l.HaPvu.

Panel A: Comparison of oligo(dT), random hexamers, and downstream primer in the RT-PCR. Samples in lanes 1, 4, and 7 were amplified starting with oligo (dT), lanes 2, 5, and 8 with random hexamers, and lanes 3, 6, and 9 with downstream primer. One μg of total RNA isolated from a prostate carcinoma cell line LNCaP (lanes 1-3), glioblastoma cell lines D65MG (lane 4-6), and U1242MG (lane 7-9) were reverse transcribed into cDNA and subsequently amplified using upstream primer derived from upstream exon of pHBGF1.H, HBGF 1601, and downstream primer from first protein coding exon, HBGF 151. Lane M: 2 μg of 123-bp ladder marker from BRL. Arrow indicates the expected fragment of 300-bp.

Panel B: The amplified fragments of 300-bp from panel A, lane 4 and lane 7 were digested with BamHI (lane 1 and lane 2, respectively). Arrow indicates one of the expected fragment of 240-bp after digestion.
observation indicated the possible presence of other unidentified upstream exon(s), pHBGF-1.X (Myers, R. and Chiu, I.-M., unpublished results). Therefore, the upstream untranslated exons of HBGF-1 and tissue distribution are rather complex.

### 1.7. Cloning of cDNA Clones Containing Upstream Exon Sequences

In order to clone the upstream exons, initial experiments examined the expression of different upstream exons in various tissues using RT-PCR. The 1.5aPvu was detected in brain, kidney and heart (Fig. 14, A, lanes 1-3; and E, lane 1). The 1.JaPvu was expressed in kidney and heart (Fig. 14, B, lanes 2 and 3; and E, lanes 3 and 5) and less amount in brain. The 1.HaPvu was also expressed in kidney and much less in heart (Fig. 14, C, lanes 2 and 3) (see summary in Fig. 37).

Various probes were used to screen cDNA libraries. The probes included: 1.JA Pst PCR product, which contains the 1.J upstream exon, 0.4-kbp SphI genomic fragment, which contains the entire 1.5 upstream exon and 1.HA PCR product containing the 1.H upstream exon. No clone has been isolated so far.

The alternative approach used to clone the upstream exon-containing cDNA was anchored PCR (Frohman et al., 1988). The gene-specific oligonucleotide HBGF 1105, corresponding to the last 35 nucleotides of the upstream untranslated 1.5 exon (sequence see Fig. 15, Panel B), was used as a primer for reverse transcription. The products were labeled with [α-3²P]ATP and electrophoresed on a 6% polyacrylamide gel. An extended product of approximately 517-bp was detected when HBGF 1105 was used as a primer.
Figure 14. RT-PCR reaction of various upstream exons and first protein-coding exon of human HBGF-1 in different tissues.

Panel A. RT-PCR of the coding region using primers HBGF 305, located at first protein-coding exon, and HBGF 202, located at second protein-coding exon, in RNAs isolated from brain (lane 1), kidney (lane 2), heart (lane 3), and H2O as negative control (lane 4). The amplified products were electrophoresed on agarose gel, transferred into Hybon-N membrane, and hybridized to a internal 176-bp EcoRI-BamHI fragment from cDNA clone. Arrow indicates the expected fragment of 265-bp.

Panel B. RT-PCR of the 1.5 upstream exon and first protein-coding exon using primers HBGF 1004 and HBGF 151. Template used and the sample treatments are the same as in panel A. Arrow indicates the expected fragment of 1.5ΔPu.

Panel C. RT-PCR of the 1.1 upstream exon and first protein-coding exon using primers HBGF 1201 and HBGF 151. Templates used and the sample treatments are the same as in panel A. Arrowheads indicate the truncated fragments of 1.1ΔPu.

Panel D. RT-PCR of the 1.1 upstream exon and first protein-coding exon using primers HBGF 1601 and HBGF 151. Templates used and sample treatments are the same as in panel A. Arrow indicates the expected fragment of 300-bp 1.1ΔPu.

Panel E. RT-PCR with non-reverse transcription control. Lane 1 and 2: HBGF 1004 and HBGF 151 were used as primers in heart RNA with or without reverse transcription. Lane 3 and 4: HBGF 1201 and HBGF 151 were used in kidney RNA with or without reverse transcription. Lane 5 and 6: HBGF 1201 and HBGF 151 were used in heart RNA with or without reverse transcription. Arrow indicates the expected fragment of 1.5ΔPu. Arrowhead indicates the truncated fragments of 1.5ΔPu.
Figure 14
Figure 15. Anchored polymerase chain reaction.

Panel A. Primer extension products using primer HBGF 1105 derived from upstream exon of pHBGF1.5. The arrow indicates the fragments of about 517-bp in brain (lane 1) and kidney (lane 2). pBR322 DNA digested with HinfI as a marker (lanes M). The marker was labeled by fill-in reaction of DNA polymerase I (Klenow fragment) in the presence of [α-32P]dATP (Maxam and Gilbert, 1980).

Panel B. The nucleotide sequence of the anchored PCR clone. The primer extension products were tailed with (dA) and subsequently amplified using primers HBGF 1003, (dT)-adaptor, and adaptor. The PCR product was subcloned into pBluescript and sequenced. The primer HBGF 1003 was non-specifically primed at upstream region to generate the anchored PCR clone. Primers are underlined. Vertical arrow indicates the putative exon junction. Positions of identity between a bovine HBGF-1 cDNA clone (Halley et al., 1988) and the human clone are marked with open circles. Triangle indicated the exon junction.
Figure 15
indicating that this was the size of the uncharacterized region (Fig. 15 A). The transcribed first-strand cDNA was tailed with (dA) using terminal deoxynucleotide transferase. The tailed cDNA was amplified by PCR using nested gene-specific oligonucleotide HBGF 1003 and oligo(dT)-adaptor primers. One amplified fragment showed a strong hybridization signal and the expected restriction enzyme sites. This fragment was subcloned and sequenced, revealing that the HBGF 1003 primer had primed nonspecifically at about 350-bp upstream from the authentic site (Fig. 15, B). Dr. R. Myers mapped the putative exon junction to be about 11 nucleotides upstream from HBGF 1003 primer (Fig. 15, B). Therefore, this anchored PCR cDNA clone could be derived from mRNA or contaminated genomic DNA. To further character this upstream exon from 1.5 exon, conventional cDNA library screening and anchored PCR will be pursued. One alternative approach is to sequence the primer extension product or use this product directly as a probe for library screening.

It is possible that these upstream exons could be spliced together before spliced into the downstream protein-coding exon. Using a combination of primers derived from various upstream exons in RT-PCR assays, kidney and heart RNA showed expected 1.1J and 1.5 composite products of 456-bp (Fig. 16, lanes 2-3; summary in Fig. 37). This products are being subcloned and sequenced. The sequence analysis could provide the information of the exact junction site.

When 1.1JPvu was used as a probe in RNase protection assays, the composite products would yield two protected fragments due to the intervening 1.5 exon sequence. The protected fragment in kidney RNA only corresponded to the product in which the 1.1J exon is spliced directly to
Figure 16. RT-PCR of the 1.J and 1.5 upstream exons using primers HBGF 1201, located at the 1.J upstream exon (sequence see Fig. 9, Panel B), and HBGF 1501, located at the 1.5 exon (sequence see Fig. 15, Panel B), in RNA isolated from brain (lane 1), kidney (lane 2), heart (lane 3) and H2O as a negative control (lane 4). The amplified products were electrophoresed on a agarose gel, transferred to Hybond-N membrane, and hybridized to the 229-bp 1.JAPstI probe. Arrow indicates the expected fragment of 456-bp.
to the first protein-coding exon. Therefore, this splicing process is a rare event which can only be detected by sensitive RT-PCR methods.
SECTION 2. ISOLATION AND CHARACTERIZATION OF HUMAN HBGF-1 GENE

2.1. Screening Strategy

With the availability of human HBGF-1 cDNA clones, several genomic clones containing the human HBGF-1 gene were isolated using the 2.2-kbp EcoRl insert of pHBGF-1.1 as a probe. Several different human genomic DNA libraries were screened, including a partial AluI-HaeIII library, a partial EcoRI, a partial Sau3A library, and a complete HindIII chromosome 5 specific library.

2.2. Genomic Clones and Exon-Intron Organization of Human HBGF-1 Gene

Eleven recombinant DNA clones from a partial AluI-HaeIII human genomic DNA library were identified initially to hybridize to HBGF-1 cDNA probe. The clones were plaque purified several cycles. The phage DNA inserts were analyzed by restriction enzyme digestion and subsequently hybridized to the HBGF-1 cDNA probe. Based upon the sizes of the EcoRI fragments and the hybridization pattern to the probe, they were grouped into three different categories: Six clones were ChIK 117, four were ChIK 108a and one was ChIK 105a.

A restriction enzyme map of the first category, ChIK 117 clones was shown in Fig. 17. The map was established by single digestion with EcoRI, BamHI, BglII and HindIII, and also appropriate double digestions of ChIK 117 phage DNA. The 9.5-kbp EcoRI fragment from ChIK 117 hybridizing to the HBGF-1 cDNA probe was further characterized. It was subcloned into
Figure 17. Restriction enzyme map of ChIK117. A human partial AluI-HaeIII-digested genomic DNA library was screened with an HBGF-1 cDNA probe from pHBCF 1.1. The wavy lines indicate the left and the right arms of bacteriophage Charon 4A. The far-left and the far-right EcoRI sites were created during the library constructing procedures; therefore, they do not represent the actual EcoRI sites in the genomic DNA. Numbers within the restriction enzyme sites are the sizes of DNA fragments in kbp. Hatched box indicates the 1.0-kbp HindIII-BglII DNA fragment hybridized with the cDNA probe.
pBluescript KS(-) plasmid vector and subjected to a more detailed mapping. Southern blotting and hybridization analysis established that the smallest hybridizing fragment was a 1.0-kbp HindIII-BglII fragment (Fig. 18, lane 6). This fragment has an internal BamHI site (Fig. 18, lane 5) which is also present near the 5' end of four HBGF-1 cDNA clones. I speculated that this fragment might actually contain the first coding exon. This speculation was confirmed by further nucleotide sequence analysis as shown below.

ChIK 108a is a chimeric clone (Fig. 20). It contains a 5.5-kbp EcoRI fragment from human DNA, also a 7.0-kbp EcoRI fragment from phage vector Charon 4A. The 5.5-kbp EcoRI fragment hybridized to the cDNA probe and contained part of the third exon, as determined by nucleotide sequencing.

ChIK 105a had three EcoRI fragments and one of them, 5.9-kbp in size, hybridized very strongly with the cDNA probe. Restriction enzyme map of ChIK 105a was established by similar restriction enzyme digestion and Southern blotting and hybridization (Fig. 19). The relative position of the BglII, HindIII, and PstI sites in the 5.9-kbp EcoRI was identical to those of cDNA clones.

Because the BglII site is unique in the cDNA clones, and is three nucleotides downstream from the translation termination codon, I reasoned that the 5.9-kbp EcoRI fragment from ChIK 105a might contain the 3'-exon. Nucleotide sequencing analysis confirmed this as shown below. The alignment of common restriction enzyme sites between the genomic DNA and cDNA, and the intensity of the hybridization signal, indicated that the
Figure 18. Southern blotting and hybridization of the 9.5-kbp EcoRI fragment of ChIK117 with the HBGF-1 cDNA probe from pHBGF 1.1. The 9.5-kbp EcoRI fragment was subcloned into the EcoRI site of pBluescript KS(-), and further electroeluted from the resulting plasmid. The purified fragment was digested with BamHI (lane 1), BglII (lane 2), HindIII (lane 3), BamHI and BglII (lane 4), BamHI and HindIII (lane 5), and BglII and HindIII double digestion (lane 6). Southern analysis was performed as described. The digested DNA was electrophoresed on a 1.2% agarose gel, transferred to a nitrocellulose filter, and hybridized to the HBGF-1 cDNA probe. The filter was exposed to the X-ray film at -80°C. Lambda DNA digested with HindIII and pX 174 DNA digested HaeIII were used as molecular weight markers.
3’ exon contained a very long, contiguous, 3’-untranslated sequence. This speculation was confirmed when the 3’-end of the gene was mapped. The gene contained a long 3’-untranslated sequence of 3.1-kbp in length, and 1.9-kbp was contained in the pHBGF-1.1 cDNA clone.

Alignment of the DNA sequence of ChIK 117 and ChIK 105a with the cDNA sequence showed a gap of 104-bp which may represent at least one exon. In order to clone this region, another library constructed by partial EcoRI digestion and cloned into Charon 4A was used. Two oligonucleotides were designed from the 104-bp gap, having seven complementary bases at their respective 3’-ends. Two oligonucleotides were annealed and labeled with [α-32p] dATP by the Klenow fill-in reaction. The clone ChIL 208 was isolated using this 63-bp exon-specific probe. Restriction enzyme mapping of ChIL 208 showed that it overlaps with ChIK 105a (Fig. 19).

Since the ChIK 117 did not overlap with clones ChIK 108a, ChIK 105, or ChIL 208, the same partial AluI-HaeIII library was screened with a unique sequence probe derived from the 3’-end of ChIK 117 in order to isolate clones bridge the gap. A new clone, designated ChIK 114 (Fig. 20), was isolated and shown to extend toward the 3’-end of ChIK 117 clone. Cross hybridization with downstream three clones was not seen, thus this clone did not bridge these two regions.

To expedite chromosomal walking, a different library was used, which is a chromosome 5 specific library with insert DNA completely digested with HindIII (Van Dilla and Deaven, 1990). The probe used was a 0.2-kbp HindIII-EcoRI fragment derived from the 3’-end of ChIK114. A new clone, ChIM 101, containing a very small insert of 0.4-kbp was isolated. I used
Figure 19. Restriction enzyme maps of ChIL 208 and ChIK 105a. The CHIK105a clone was obtained by screening the partial AluI-HaeIII human genomic DNA library with the HBGF-1 cDNA probe from pHBGF-1.1. The CHIL208 clone was obtained by screening a partial EcoRI human genomic DNA library with a second-exon-specific 63-mer probe. The overlap of these two clones was established by aligning the common restriction enzymes sites and by Southern blot analysis. One BglII and one HindIII sites have not been mapped definitively (?). Exact location of four PstI sites in the 4.2-kbp EcoRI fragment of CHIK105a were not determined. Hatched boxes indicate the locations of the second and third exons.
Figure 20. Restriction enzyme map of the human HBGF-1 gene. Nine genomic DNA phage clones were isolated from different human genomic DNA libraries. Each clone was mapped by appropriate restriction enzyme digestion and hybridized to HBGF-1 cDNA probe or probes derived from overlapping genomic DNA clones. The three closed boxes represent the three coding exons. The three hatched boxes represent the three probes used to characterize genomic DNA from leukemia patients (see later). a: 1.4-kbp EcoRI probe (spanning from 7.0 - 8.4-kbp). b: 2.1-kbp HindIII probe (spanning from 18.5 - 20.6-kbp) and 2.6-kbp BgIII probe (spanning from 38.5 - 41.1-kbp). The three hybridizing EcoRI fragments (5.8-kbp, 19.8-kbp and 11.7-kbp) are also indicated. B: BamHI; G: BgIII; H: HindIII; R: EcoRI Restriction enzymes.
Figure 20
this insert fragment as a probe to isolate two additional overlapping clones ChIK 506 and ChIO 107 (Fig. 20) from the partial AluI-HaeIII and the partial Sau3AI λ-FIX library, respectively.

In an attempt to clone the upstream regions which may contain the cis-regulatory elements, a unique 1.4-kbp EcoRI fragment (probe a in Fig. 19) derived from the most 5′-end of ChIK 117 was used as a probe for this purpose. I identified an upstream clone, ChIO 209, from the λ-FIX library. These nine overlapping phage clones comprise a stretch of 54-kbp of genomic DNA including three coding exon of human HBGF-1 gene in two EcoRI fragments, 19.8-kbp and 11.7-kbp (Fig. 20).

2.3. Genomic DNA Sequence Analysis

The DNA fragments from each phage clone hybridizing to the HBGF-1 cDNA probe, namely the 1.0-kbp HindIII-BgIII fragment of ChIK 117, the 2.5-kbp EcoRI-HindIII fragment of ChIL 208, and the 1.5-kbp EcoRI-BglIII and 2.6-kbp BglII-BglII fragments of ChIK 105a, were cloned into pBluescript vectors for further nucleotide sequence analysis. Unidirectional deletion (Henikoff, 1984) was used to prepare some of the templates for sequencing. Some sequence analyses were performed on an automatic DNA sequencer. The sequencing strategy was shown in Fig. 21.

DNA sequencing showed that in comparison with the published amino acid sequence, these clones contained the entire protein-coding sequence (Gautschi-Sova et al., 1986; Gimenez-Gallego et al., 1986; Harper et al., 1986; Fig. 22). The protein-coding regions were included within three
Figure 21. Sequencing strategies for the determination of the three coding exons and flanking regions of human HBGF-1 gene. The three closed boxes represent the three coding exons. The arrows indicate the direction and extent to which the sequences were determined. Close circles indicate sequencing by M13, reverse M13, T7, and T3 primers. Open rectangles indicate sequencing by designed oligonucleotide primers. Triangles indicate sequencing by an automatic sequencer.
Figure 22. Nucleotide sequence of HBGF-1 gene and its predicted amino acid sequence. Hybridizing DNA fragments from genomic DNA phage clones were subcloned into pBluescript vectors. Deletional clones were generated using exonuclease III and mung bean nuclease as described by Henikoff (1984). ss DNA or ds DNA was sequenced by the chain termination method of Sanger et al. (1977) using T7, M13, or other designed synthetic oligonucleotide primers. Appropriate restriction enzyme sites in the DNA sequences are indicated for references. Symbols: and , splice acceptor and donor sites, respectively; ------, intron (the distance between the first and second exon was 13.6-kbp, whereas the distance between the second and third exon was 5.3-kbp); wavy lines represent several polyadenylation signal sequences. The accession numbers are M23017, M23086, and M23087 (for parts of the third coding exon) for the three coding exons, respectively (Wang et al., 1989). The accession number for the complete sequence of the third coding exon is X59065 (Wang et al., 1991).
exons of 169, 104 and 192-bp in length. All splicing donor and acceptor sites were the same as the canonical splicing sequences (Sharp, 1981).

An in-phase termination codon, TGA, located three nucleotides upstream from the translation initiation codon, was also observed (Fig. 22). This indicates the more upstream sequence was most likely untranslated sequences. Parts of the untranslated sequences in genomic DNA were also shown in the cDNA.

RNase protection analysis established that the major polyadenylation sites were approximately 3.1-kbp downstream from termination codon. This entire region was sequenced. The two major polyadenylation sites identified by RNase protection analysis (Chiu et al., 1990c), as well as three minor sites previously identified by cDNA cloning (Crumley et al., 1989; Chiu et al., 1990c) were presented in the DNA sequence of the third exon (Fig. 22, wavy lines).

2.4. Mapping the 5'-end of the First Protein-Coding Exon

Before the 5' regions of cDNA clones pHGF-1.4 and pHGF-1.5 were available to be sequenced, a 1.0-kbp HindIII BglII genomic fragment derived from ChIK 117 was cloned into pBluescript KS (+) vector. The plasmid DNA was linearized at the KpnI or NcoI site, and the antisense RNA riboprobe was made from in vitro transcription using T7 RNA polymerase and \([\alpha-\text{32p}]\) UTP (Fig. 23). The RNA riboprobe was hybridized with 20 µg of RNA from various sources and then digested with RNase A and T1 (Zinn et al., 1983; Melton et al., 1984). The protected fragment was fractionated on a 6% polyacrylamide gel. Using KpnI transcript, a protected fragment,
Figure 23. Strategies for RNase protection assay. A 1.0-kbp HindIII-BglII fragment containing the first coding exon of the HBGF-1 gene was cloned into the HindIII-BamHI sites of pBluescript KS(+) plasmid DNA. The resultant plasmid DNA placed the antisense strand downstream from the phage T7 promoter. The solid box represents the first coding exon, which was bounded by splicing acceptor and splicing donor sequences. Panel A and B schematize the size of RNA in vitro transcripts when plasmid DNA digested with KpnI and NcoI, respectively, was used as template for mapping. Wavy lines represent corresponding protected fragments after RNA-RNA hybridization and RNase digestion.
-200 bp, was detected in kidney and brain RNA using Kpnl transcript (Fig. 24 A, lanes 1 to 3). Glioblastoma cell line A2781 RNA also showed a fainter signal which suggested a lesser degree of expression (Fig. 24A, lane 4). Through the comparison between cDNA and genomic DNA sequence, I localized the 3'-end of the first protein-coding exon. On the basis of the size of the protected fragment (~200 bp) and the canonical splicing acceptor sequence (Sharp, 1981), the 5'-end was predicted to be 203 bp upstream from the 3'-end. This assignment was confirmed by the observation of a 166 bp protected fragment when a shorter NcoI transcript was used for the same assay (Fig. 24 B and C). When the 5'-end sequence of cDNA clones became available, it confirmed these results obtained from RNase protection assays.

Since the RNase protection assay is a more sensitive method than Northern blot analysis; it was used to confirm that HBGF-1 mRNA was present in a 13-week-old fetal brain and kidney, as previously determined using a cDNA insert for Northern hybridization (Fig. 23 C, lanes 5 and 6). However, the fetal liver and term placenta appeared not to express the HBGF-1 mRNA at a detectable level.

2.5. Mapping 3'-ends of the Gene and Alternative Polyadenylation

The most 3'-end of the cDNA clone pHBGF-1.1 contained a poly (A) stretch and the polyadenylation signal (Fig. 7). In order to map the 3'-end(s) of the gene, a series of RNase protection experiment were performed.

First, a 1.4-kbp EcoRI genomic DNA fragment containing the putative polyadenylation signal of pHBGF-1.1 was digested with AlwNI restriction
Figure 24. RNase protection assays and expression of RNA in human fetal tissues, using the KpnI (A) and NcoI (B and C) transcripts. Total cellular RNA was isolated from various tissues and cell lines as described by Chirwin et al. (1979). The KpnI and NcoI transcripts were synthesized in vitro, using phage T7 RNA polymerase and [α-32P]UTP (Melton et al., 1984). The transcript was hybridized to 20 μg of cellular RNA and then digested with RNase A and T1 (Zinn et al., 1983). The protected fragments were separated on a 6% polyacrylamide gel under denaturing conditions. The RNase-resistant fragments are indicated by arrows.

Panel A: lanes 1 and 2, two different human kidneys; 3, human brain; 4, glioblastoma cell line A2781; 5, tRNA.
Panel B: lane 1, NcoI transcript alone without RNase digestion; 2 and 3, two different human kidneys; 4, glioblastoma cell line A2781; 5, tRNA; 6, human brain.
Panel C: lanes 1 and 2, NcoI transcript alone without RNase digestion; 3, human kidney; 4, human brain; 5, human fetal brain; 6, human fetal kidney; 7, human fetal liver; 8, human placenta; 9, glioblastoma cell line A2781; 10, tRNA.
Figure 24
enzyme. The riboprobe generated was subsequently hybridized to human brain and kidney RNA. A major protected fragment was the 0.68-kb full length probe (data not shown). A variety of fainter bands were detected with longer exposure. One of the protected fragments corresponded to the 3'-end of pHBGF-1.1 clone. This result indicates that the major polyadenylation site(s) should reside further downstream.

Therefore, the cloned 2.6-kbp BglII-BglII fragment containing the 1.4-kbp EcoRV was digested with EcoRV and the riboprobe was used for similar assays (Fig. 25, A). A unique 360 bases protected fragment was detected in human brain and kidney (Fig. 26, A, lanes 1 and 2). This fragment was slightly shorter than the original riboprobe, because the probe included an extra 50 bases of sequences derived from the vector. Total RNA isolated from human prostate and tRNA from yeast were used as negative controls (Fig. 26, A, lanes 3 and 4). These results again indicate the entire 2.6-kbp BglII-BglII fragment was part of the exon.

The next cloned downstream fragment, the 1.8-kbp BglII-EcoRI fragment was used (Fig. 25, B). XbaI transcript (12 bp downstream from the BamHI/BglII fusion site in the vector sequence) gave two protected fragments of approximately 520 and 510 bases in both human brain and kidney RNA (Fig. 26, B, lanes 1 and 2). These results were consistent with the results obtained using a different transcript (Ball transcript) as a probe. The two protected fragments were 250 and 240 bases (Fig. 26, C, lanes 1 and 2). Nucleotide sequence analysis also showed two polyadenylation signal sequences, AATAAA and AATAA, in the corresponding regions (Fig. 22).
Figure 25. Strategies for RNase protection assays. DNA fragments derived from the 3'-portion of the third coding exon as well as its downstream regions are displayed. The 5'-BglII site is 3-bp downstream from the translation termination codon.

Panel A: The 2.6-kbp BglII-BglII fragment containing the 3'-end of HBGF-1 cDNA, pHBGF-1.1, was cloned into the BamHI site of pBluescript KS(+) and designated pBg2.6. The DNA fragment was cloned in the proper orientation for the production of antisense RNA from the phage T7 promoter. The size of the in vitro RNA transcript was 410-bases when pBg2.6 was digested with EcoRV and used as the template. The size of the protected fragment after RNase digestion was 360-bases as indicated by the wavy line.

Panel B: The downstream fragment 1.8-kb BglII-EcoRI fragment was cloned into the BamHI-EcoRI sites of pBluescript SK(+) and designated pBE1.8. The size of the antisense RNA transcript was 1.8-kb when the template was digested with XbaI (12-bp downstream from the BamHI/BglII fusion site). The sizes of the protected fragments were 520-bases and 510 bases.

Panel C: The size of the antisense RNA transcript was 1.5-kb when the pBE1.8 template was digested with BalI. The sizes of the protected fragments were 250-bases and 240-bases.

The hatched box represents the 3'-untranslated region of a minor HBGF-1 transcript which was contained in the cDNA clone, pHBGF-1.1. The cross-hatched box represents those of the two major transcripts. The first AATAAA sequence was the polyadenylation signal for one of the minor transcripts (see later HBGF-1 genomic DNA sequence), while the second AATAAA and the AATAA sequences may serve as the polyadenylation signals of the major transcripts.
Figure 25

A. pBg 2.6 digested with EcoRV

B. pBE 1.8 digested with XbaI

C. pBE 1.8 digested with Bal I
Figure 26. RNase protection analysis with DNA probes derived from the 3'-end of the gene. Plasmid DNA was digested with appropriate restriction enzymes and used as template. The transcripts were synthesized in vitro, using phage T7 RNA polymerase and labeled with $[^\alpha-\text{32P}]$UTP. The transcripts were hybridized to 20 μg of cellular RNA and then digested with RNase A and T1. The protected fragments were separated on 6% polyacrylamide gels under denaturing conditions.

Panel (a): plasmid pBG2.6 DNA digested with EcoRV was used as a template. Panel (b): plasmid pBE1.8 digested with XbaI was used as a template. Panel (c): plasmid pBE1.8 digested with Ball was used as a template. Lane M, pBR322 DNA digested with HinfI was used as a marker; lane C, the riboprobe synthesized in vitro without any further treatment was used as a control. RNA isolated from human adult brain (lane 1), human adult kidney (lane 2) or human adult prostate (lane 3), as well as yeast tRNA (lane 4) was used for RNA-RNA hybridization. The RNase-resistant fragments are indicated by arrows and the sizes see figure 24.
Taken together, the two major polyadenylation sites, approximately 10 bp apart, are located 3.1-kbp downstream from the translation stop codon. I also used the 3'-untranslated 2.6-kbp BglII-BglII fragment to detect the HBGF-1 mRNA by Northern blotting and hybridization. A 4.5-kb mRNA species was detected (Fig. 27). Two additional faint bands with sizes of 3.4- and 2.0-kb were also detected in both kidney and brain RNA, which could result from alternative polyadenylation. The size difference of brain and kidney HBGF-1 mRNA as previously indicated is also evident in the Northern blot here (Fig. 27).

2.6. Sequence Conservation Among Different Species

The 3'-untranslated 2.6-kbp BglII-BglII fragment was used to address whether human HBGF-1 is a single copy gene. High molecular-weight genomic DNA was isolated from human placenta and digested with EcoRI, BamHI, BglII, HindIII, PstI, PvuII or SphI. A simple pattern of hybridizing bands, which could be explained by previous restriction enzyme map (Fig. 28) was found. No other additional bands were detected under low-stringency conditions. Other genomic fragments used for genomic Southern blotting and hybridization also supported that the human HBGF-1 gene is a single copy gene. Thus, β-ECGF, α-ECGF, αFGF-1, αFGF-2 and other HBGF-1 preparation known by several other names are encoded by the same gene.

Determination of the sizes of the exons and more importantly the complete nucleotide sequence allowed a detailed sequence comparison with the corresponding region of HBGF-1 cDNA sequences from bovine, rat, hamster, mouse and chick (Halley et al., 1988; Goodrich et al., 1989; Hall
Figure 27. Northern blotting and hybridization of human RNA with the HBGF-1 gene. Total cellular RNA was isolated from two different human kidneys (lane 1 and lane 2), human glioblastoma cell line A2781 (lane 3), and human brain (lane 4) by the guanidinium thiocyanate method (Chirgwin et al., 1979). Total cellular RNA (20 μg) was electrophoresed on a 1.1% agarose gel containing formaldehyde (Dobner et al., 1981). The RNA was transferred to a nitrocellulose filter and probed with the 2.6-kpb BgIII-BglII DNA fragment containing predominantly the third exon. Human glioblastoma cell line A2781 produced significantly less HBGF-1 mRNA than did brain or kidney.
Figure 28. Genomic Southern blotting and hybridization of human genomic DNA with the HBGF-1 gene probe. High molecular weight genomic DNA was isolated from human placenta and digested with EcoRI, BamHI, BglII, HindIII, PstI, PvuII, and SphI, as indicated. The digested DNA was then electrophoresed on a 0.7% agarose gel, blotted into a nitrocellulose filter, and hybridized to the 2.6-kbp BglIII-BglIII DNA probe isolated from CHIK105a. The probe containing predominantly the third coding exon of the HBGF-1 gene. Bacteriophage lambda DNA digested with HindIII was used as a molecular weight marker. No fragments smaller than 2.0-kbp were detected on the original autoradiogram.
et al., 1990; Hebert et al., 1990; Schnurch and Risau, In press). A high degree of similarity (79.6%-91.6%; see Table 2) in the nucleotide sequence of the coding regions was observed in human and each of the other five species. A DOTPLOT analysis with 70% sequence similarity and a window size of 30 nucleotides for the human and bovine cDNA is presented (Fig. 29). The similarity extended 2.4-kbp into the long 3'-untranslated region. If the parameters used for DOTPLOT analysis are increased to 90% similarity, several patches of sequence homology still exist (nt. 1751-1848, nt. 1896-2087, nt. 2471-2652).

2.7. Conservation of Intervening Sequences Among Different Species

Since the 1.4 kb EcoRI-EcoRI fragment (probe a in Fig. 20) did not contain repetitive sequences, I decided to test whether homologous sequences are present in other vertebrates in the zooblot analysis. DNA from human, chimpanzee, rhesus monkey, and baboon was digested with EcoRI and the 1.4-kbp EcoRI-EcoRI was used as a probe. Homologous fragments were identified in the primate DNA with similar size (Fig. 30, A, lanes 1-4). Rhesus monkey DNA also shows discrete hybridizing bands of 1.9-kbp and 3.7-kbp upon digestion with PrulI and PstI (Fig. 30, A, lanes 5,6). The same fragment was used to probe the other non-primate DNA under non-stringent conditions. The results showed homologous sequences in rodent but not in avian DNA (Fig. 30, B). Experiments using this fragment as a probe in Northern blot analysis failed to detect any RNA expression.

2.8. Isolating the Genomic Clones Containing Ustream Untranslated Exns
Figure 29. Homology between the cDNA sequences of human and bovine HBGF-1. The cDNA sequences of human and bovine HBGF-1 were compared using the DOTPLOT program of DNA Star, Inc. The human cDNA sequence was generated by splicing the genomic DNA sequences from three coding exons. The parameters used were 70% sequence similarity with a window size of 30 nucleotides. The hatched boxes represent the protein-coding regions of the two cDNA sequences. Nucleotide number 1 corresponds to the first base of the first coding exon.
Figure 30. Conservation DNA sequence upstream from the HBGF-1 coding region.

Panel A: Hybridization of primate DNA under stringent conditions. Genomic DNA was prepared from human placenta (lane 1), chimpanzee skin cells (lane 2), fetal rhesus monkey lung cells (lanes 3, 5, and 6), and baboon cells (lane 4). Genomic DNA was digested with EcoRI except DNA in lanes 5 and 6, which was digested with PvuII and PstI, respectively. The digested DNA was blotted and hybridized to the 1.4-kbp EcoRI fragment (probe a)

Panel B: Hybridization of non-primate vertebrate DNA under non-stringent conditions. Genomic DNA was prepared from the liver tissues of mouse (lane 1), hamster (lane 2), rat (lane 3), and chick embryo (lane 4). The DNA was digested with EcoRI and hybridized to the 1.4-kbp EcoRI fragment under non-stringent conditions. The arrowheads indicate hybridizing bands in the rodent DNA. Chick DNA did not show any hybridizing bands. Lambda DNA digested with HindIII was used as a marker.
Since the size of the major HBGF-1 transcript is 4.5-kb and the entire coding sequence plus the 3'-untranslated region so far obtained is 3.6-kbp, the 5'-untranslated sequence was deduced to be about 0.9-kbp (Fig. 31). Both genomic DNA walking and oligonucleotide probes designed from upstream sequences of cDNA clones were used to obtain this 5’-untranslated sequence.

Oligonucleotide probes derived from upstream sequences of pHBGF-1.4 and pHBGF-1.5 were used initially to screen various phage and cosmid genomic DNA libraries. When a 23-mer (HBGF 1103, which corresponds to the last 23 nucleotides of the upstream untranslated exon of pHBGF-1.4) was used as a probe, three putative positive clones were isolated from a partial AluI-HaeIII library based on hybridization to the probe. Upon sequencing, each of these clones shows similarity in their sequences but not identity to the probe sequence. K. Lehtoma used a 72-mer from pHBGF-1.5 to isolate a genomic clone, ChIO 101, containing upstream sequence from a peripheral blood leukocyte genomic DNA library constructed in λ-FIX (kindly provided by L. Ernst). R. Payson used end-specific walking probes from ChIO 101 to isolate several overlapping cosmid and phage clones from two chromosome 5 specific genomic DNA libraries constructed in λ-FIX and sCos-1 (obtained from D.L. Deaven). Including upstream untranslated 1.J and 1.5 sequences, these overlapping clones comprise the other greater than 40-kbp genomic DNA, which did not overlap with downstream phage clones. Pulse-field gel electrophoresis is currently being used to examine the size of the gap.
Figure 31. Exon-intron organization of the gene coding for human HBGF-1. The solid line represents the genomic DNA and cDNA regions cloned while the dashed line represents uncharacterized cDNA sequences. Closed boxes in the genomic DNA map represent the three protein-coding exons and the hatched box in the cDNA map represents the protein-coding region. Several human genomic DNA libraries were screened with probes derived from the human HBGF-1 cDNA clones to obtain most of the genomic DNA clones. The remaining phage clones were obtained by chromosome walking using the DNA probes derived from the inserts of existing phage clones. Overlap of the genomic DNA clones was established by aligning the common restriction enzyme sites and by hybridization analysis. Both the location(s) of the 5'-end exon(s) and the transcription initiation site(s) are not mapped.
SECTION III. INVESTIGATION OF HBGF-1 LOCUS IN ACUTE NONLYMPHOCYTIC LEUKEMIA PATIENTS

3.1. HBGF-1 and 5q- Chromosome

HBGF-1 gene has been mapped to the long arm of chromosome 5, bands 5p 31.3-33.2 (Jaye et al., 1986). This location is very close to the distal breakpoint of 5q- deletion observed in some leukemia patients. Three types of 5q- deletion have been identified and all types involve the distal breakpoint clustering in the same region as HBGF-1, while the proximal breakpoints vary (Fig. 32). Rowley (1985) reported that 20% of acute nonlymphocytic leukemia (ANLL) patients have a 5q- deletion. Since various HBGF-1 genomic probes were available, I analyzed the genomic DNA of eight ANLL patients for possible deletion of HBGF-1 gene. Two ANLL patients have 5q- karyotype (Table 3). A patient with acute lymphocytic leukemia (ALL) was included as a control.

3.2. Genomic Southern Blot Using Different HBGF-1 Probes

Patient DNA was digested with EcoRI and hybridized with the 2.6-kbp BglII-BglII fragment containing the majority of the 3'-untranslated region (probe c in Fig. 20, spanning from 38.5-41.1-kbp). The results showed a single hybridizing band of 11.7-kbp in DNA isolated from patients with (Fig. 33, lanes 2 and 10) or without (Fig. 33, lane 1) 5q- deletion. The size of the hybridizing DNA is consistent with that predicted from the restriction enzyme map of HBGF-1 locus. Since the 11.7-kbp EcoRI fragment contains the second and third protein-coding exons, the results suggest that there is no gross gene rearrangement in this region.
Figure 32. Location of HBGF-1 gene and deletions in chromosome 5q'.
The location of HBGF-1 gene was described in Jaye et al. (1986). Three
types of deletions in the long arm of chromosome 5 were frequently
Brackets indicate the limits of uncertainty of breakpoints in chromosome
5q'. The staining pattern was adapted from Yunis and Hoffman (1989).
Table 3. Karyotype analysis of nine leukemia patients.

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Diagnosis</th>
<th>Age</th>
<th>Karyotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>S.N. 101</td>
<td>ANLL</td>
<td>60</td>
<td>XX, 5q-</td>
</tr>
<tr>
<td>S.N. 102</td>
<td>ANLL</td>
<td>63</td>
<td>XY, Normal</td>
</tr>
<tr>
<td>S.N. 103</td>
<td>ANLL</td>
<td>85</td>
<td>XX, Normal</td>
</tr>
<tr>
<td>S.N. 104</td>
<td>ANLL</td>
<td>22</td>
<td>XX, -7</td>
</tr>
<tr>
<td>S.N. 109</td>
<td>ANLL</td>
<td>18</td>
<td>XY, with variable hypodiploidy, but no deletion of chromosome 5</td>
</tr>
<tr>
<td>S.N. 110</td>
<td>ALL</td>
<td>46</td>
<td>XY, 14q+</td>
</tr>
<tr>
<td>S.N. 112</td>
<td>ANLL</td>
<td>64</td>
<td>XY, Normal</td>
</tr>
<tr>
<td>S.N. 116</td>
<td>ANLL</td>
<td>75</td>
<td>XX, Normal, has a point mutation (GGT/Gly → GCT/Ala) in the 12th codon of N-ras gene.</td>
</tr>
<tr>
<td>S.N. 121</td>
<td>ANLL</td>
<td>55</td>
<td>XY, 5q-</td>
</tr>
</tbody>
</table>
Figure 33. Genomic Southern blotting analysis of leukemia patient DNA with the 2.6-kbp BglII DNA probe. Genomic DNA in lane 2 to 10 was isolated from patients 101, 102, 103, 104, 109, 110, 112, 116, and 121, respectively (see table 3). DNA was also isolated from a human placenta (lane 1). The genomic DNA was digested with EcoRI and hybridized to a 2.6-kbp BglII fragment (probe c, spanning from 38.5 - 41.1-kbp in Fig.15) representing a major portion of the 3'-untranslated sequence of HBGF-1 gene. The arrowhead indicates the expected hybridizing band of 11.7-kbp in size. Lambda DNA digested with HindIII was used as a molecular weight marker.
A similar genomic blot was performed and probed with a 2.1-kbp HindIII-HindIII fragment containing the first coding exon (probe b in Fig. 20, spanning from 18.5-20.6-kbp) and a 1.4-kbp EcoRI-EcoRI fragment derived from the 5'-end of the genomic phage clone ChIK 117 (probe a in Fig. 20, spanning 7.0-8.4-kbp). Two hybridizing bands of 19.8-kbp and 5.8-kbp were detected in DNA isolated from ANLL patients (Fig. 33, lanes 1-4), normal control individuals (Fig. 33, lane 5-9) and a promonocytic cell line, U937 (Fig. 33, lane 10). The sizes are consistent with those predicted from the restriction enzyme map (Fig. 20). Again, this result suggests that there is no gross gene rearrangement at the HBGF-1 locus in ANLL patients.

To investigate if there is allelic loss of the HBGF-1 gene in patient chromosomes, a probe derived from the human β-globin gene, a 0.9-kbp EcoRI-BamHI fragment, was used to normalize the signal. Thus, the signal in the previous filter was stripped and the filter was reprobed with 2.1-kbp HindIII-HindIII (probe b) and the β-globin probe. The same filter contained DNA isolated from five normal control individuals and from a promonocytic cell line, U937. I compared the intensities of the 19.8-kbp HBGF-1 hybridizing band and the 5.5-kbp β-globin hybridizing band with a densitometer. No allelic loss of the HBGF-1 gene in ANLL patients with 5q- (Fig. 35, lane 1) or without detectable 5q- (Fig. 35, lanes 2-4) was found.

Combined with the results obtained from the above series of genomic Southern blotting and hybridization, there is no gross rearrangement or allelic loss within 16.5-kbp upstream from the HBGF-1 coding region. Thus, HBGF-1 gene is preserved in ANLL patient genome in the 40-kbp stretch.
Figure 34. Leukemia patient DNA probed with the 2.1-kbp HindIII fragment and the 1.4-kbp fragment of the HBGF-1 gene. Genomic DNA in lane 1 to 4 was isolated from patients 101, 102, 103, and 109, respectively. DNA in lane 5, 6, 8, and 9 was isolated from four normal individuals and DNA in lane 7 was isolated from a human placenta. DNA in lane 10 was isolated from a promonocytic cell line, U937. The genomic DNA was digested with EcoRI and hybridized with both the 2.1-kbp HindIII (probe b, spanning from 18.5 - 20.6-kbp in Fig.15) and the 1.4-kbp EcoRI (probe a, spanning from 7.0 - 8.4-kbp in Fig.15) fragments derived from the HBGF-1 gene. Arrowheads indicate the expected hybridizing bands of 19.8-kbp and 5.80-kbp, respectively. Lambda DNA digested with HindIII was used as a molecular weight marker.
Figure 35. Leukemia patient genomic DNA probed with the 2.1-kbp HindIII fragment of the HBGF-1 gene and the β-globin gene. Genomic DNA in lanes 5, 6, 8, and 9 was isolated from four normal individuals and DNA in lane 7 was isolated from a human placenta. DNA in lane 10 was isolated from a promonocytic cell line, U937. The filter used in Fig.28 was washed with 0.1 X SSC and 0.1% SDS at 100°C for 30 min. The filter was exposed to an X-ray film to ensure complete removal of the probes. The filter was then hybridized to the 2.1-kbp HindIII fragment containing the first coding exon of human HBGF-1 gene (probe b) and 0.9-kbp EcoRI-BamHI fragments derived from β-globin gene. Arrowheads indicate the expected hybridizing bands of 19.8-kbp and 5.5-kbp, respectively.
3.3. RT-PCR of HBGF-1 expression in leukemia patient

It may be possible that rearrangement or deletion further upstream from the HBGF-1 gene alters its transcription. To examine the HBGF-1 mRNA expression in the ANLL patients, the sensitive RT-PCR was used. The HBGF-1 mRNA was detected in brain and kidney (Fig. 36, A, lanes 1 and 2), consistent with previous results obtained using Northern hybridization or RNase protection analysis (Wang et al., 1989; Fig. 8 and Fig. 24). In contrast, mononuclear cells isolated from an ANLL patient with 5q-deletion (Fig. 36, A, lane 3) and a normal individual (Fig. 36, A, lane 4) and two myeloid cell lines (Fig. 36, A, lanes 5 and 6) did not have detectable amounts of HBGF-1 mRNA. The primers used were those derived from the first coding exon, upstream primer nt. 22-45 and downstream primer nt. 210-188 (the numbering system used in that in Fig. 7), therefore the amplified 189-bp product could have represented the PCR product of genomic DNA. However, this is not likely, because I digested the reverse-transcription product with BamHI enzyme which cleaves the template between the primers before proceeding with the polymerase chain reaction. In the parallel control, one nanogram of plasmid DNA containing the HBGF-1 cDNA insert, when digested with BamHI, did not generate any PCR products. We also used another pair of primers which were located at two different exons, upstream primer nt. 22-45 and downstream primer nt. 286-267. These two exons were separated by a 13.6-kbp intron. The expected size is 265-bp. The same results were obtained when this pair of primers were used for PCR confirming the results shown in Figure 35 A. To test the integrity of RNA isolated from both the mononuclear cells and the myeloid cell lines, a pair of primers derived from FcγRI, nt. 391-408 of
Figure 36. RT-PCR of RNAs isolated from different tissues and cell lines. Panel A. RT-PCR of human HBGF-1 mRNA. The arrow indicates the expected amplified product of 189-bp. The template used for RT-PCR are H2O as negative control (lane 1), positive control supplied in the kit (lane 2), or RNA isolated from mononuclear cells of 5q- patient (lane 3), a normal control (lane 4), DMS 10 stimulated with γ-IFN (lane 5), brain (lane 6) and kidney (lane 7). Panel B. RT-PCR of human FcγRI. Templates used are the same as in panel A except in lane 1 where RNA isolated from T-cells was used as a negative control.
exon 1 and nt. 107-90 of exon 3 (van der Winkel et al., 1991), were used for RT-PCR (Fig. 36, B, lanes 3-6). The expected size is 110-bp. It is interesting to note that brain and kidney also express FcγRI mRNA while T-cells did not.
MOLECULAR CLONING AND CHARACTERIZATION OF HUMAN HBGF-1 cDNA

HBGF-1 is a mitogen for a variety of mesoderm- and neuroectoderm-derived cells in vitro as well as an angiogenic factor in vivo. HBGF-1 represents one of the seven members of the HBGF family. The gene for the other six members, HBGF-2 (Abraham et al., 1986), int-2 (Moore et al., 1986), hst/KS3 (Delli Bovi et al., 1987; Yoshida et al., 1987), FGF-5 (Zhan et al., 1988), FGF-6 (Marics et al., 1989; de Lapeyriere et al., 1990) and KGF (Rubin et al., 1989) have been cloned from either murine or human species. Since brain and other neural tissues are major sources of the HBGF-1, the system provide a unique opportunity to study the control of normal growth and oncogenesis in the nervous system.

Alternatively Spliced Upstream Untranslated Exons

Four cDNA clones coding for human HBGF-1 have been isolated from a human brain stem cDNA library. The cDNA sequences for pHBGF-1.1 and 1.2 are identical to those reported for the genomic DNA sequence (Wang et al., 1989) but are slightly different from the previously reported cDNA sequence of HBGF-1 in the 3'-untranslated region (Jaye et al., 1986). Nucleotide sequence analysis revealed that the DNA sequences upstream from the first protein coding exon in two other human HBGF-1 cDNA clones (pHBGF-1.4 and 1.5) are different. The fact that the sequence divergence
begins at the exon-intron boundary strongly argues against the possibility of a cloning artifact. Because the difference occurs in the 5'-untranslated regions, these transcripts may result from the usage of alternative promoters. However, a riboprobe derived from pHBGF-1.4 failed to detect any protected fragment corresponding to this upstream region. After the genomic DNA containing pHBGF-1.5 upstream sequence were cloned, it was found that sequence identical to that of pHBGF-1.4 was present in the opposite strand of the pHBGF-1.5 exon. The cDNA library that I used was constructed by a modification of the method of Huynh et al. (1985), which utilizes self-primed synthesis of the second strand of cDNA and subsequent cleavage of the hairpin structure. It is now recognized that this loopback method results in the loss or rearrangement of sequences corresponding to the 5'-terminus of the mRNA (Land et al., 1981). One example is the cDNA sequence of human complement component C4 (Belt et al., 1985), in which the 5'-end of the cDNA insert covering the initiation codon and part of the leader sequence, was inverted when compared to the sequence of the genomic fragment. It is most likely that the upstream sequence of pHBGF 1.4 is a cloning artifact.

Comparison of the 5'-untranslated sequences of HBGF-1 cDNA clones from 4 different mammalian species (Halley et al., 1988; Goodrich et al., 1989; Hall et al., 1989; Chiu et al., 1990) shows similarities present between the 1.H upstream exon and the 5'-untranslated sequence of the hamster HBGF-1 cDNA clone (Fig. 12); and the 1.5 upstream exon and the 5'-untranslated sequence of the bovine HBGF-1 cDNA clone (Fig. 15). Similarities between the other pairs are not closer than that between two randomly generated DNA sequences. It is not likely that this dissimilarity
is due to evolutionary divergence since the growth factor genes are known to be conservative, but rather due to the cloning of different exons. These data suggest that there is at least one additional exon existing in the human genome, which corresponds to the 5'-untranslated sequence in rat HBGF-1 DNA clones. This assumption is further supported by the observation that some tissues which do not contain any of the 1.J, 1.5 or 1.H upstream exons express HBGF-1 mRNA. We are currently cloning this possible upstream exon using both cDNA cloning and anchored PCR (Frohman et al., 1989; Loh et al., 1989).

A series of RT-PCR and RNase protection assays showed that the human HBGF-1 gene has at least 3 upstream untranslated exons, designated 1.J, 1.5 and 1.H, which are alternatively spliced into the first protein-coding exon. The splicing pattern varies among different tissues and tumor cell lines. The 1.J upstream exon is expressed in kidney, heart and, albeit to a much lesser extent, brain (Fig. 11 and Fig. 15). The 1.5 upstream exon is expressed in brain, kidney and heart (Fig. 15). The 1.H upstream exon is expressed in two glioblastoma cell lines, kidney and heart (Fig. 13 and Fig. 15). RT-PCR also identified a minor portion of mRNA in kidney and heart having a splicing patterns of 1.J into 1.5 (Fig. 16). Preliminary results indicate that upstream region of the 1.J exon may contain a TATA box. It is not known if it corresponds to a functional transcription site presently. If this is the case, HBGF-1 could have alternative promoters in addition to alternative splicing (Summarized in Fig. 37). These two mechanisms could, separately or together, be responsible for the production of more than one mRNA from the HBGF-1 gene.

The different forms of HBGF-1 mRNA can be differentially transcribed
using specific promoters, each responding to specific physiological signals to modulate cell growth and differentiation. Alternative splicing generating different forms of mRNA from a single gene has been documented frequently in the literature (for a review, see Breitbart et al., 1987). For example, tissue-specific usage of alternative promoters is shown in the gene coding for dystrophin (Feener et al., 1989; Nudel et al., 1989). The discovery of tissue-specific isoforms indicates that proteins which interact with dystrophin could differ between tissues in which dystrophin is expressed.

**Ectopic Expression of HBGF-1 cDNA**

A retroviral expression vector carrying the full length HBGF-1 coding sequences using pHBF-1.1 and pHBF-1.2 has been constructed (Bunnag et al., 1991). Transfectants expressing full-length HBGF-1 protein at high level form foci and grow to a higher cell density than the parental NIH/3T3 cells. The transformed cells are capable of growing on soft agar even in the absence of exogenously added HBGF-1 and also induce tumor formation when injected into nude mice. Thus, NIH/3T3 cells acquire a full spectrum of transformed phenotype when full length HBGF-1 was expressed at high levels. This is different from a previous report which showed that the truncated HBGF-1 confers a partial transformed phenotype to the recipient NR6 cells (Jaye et al., 1988). These transfected cell lines present a powerful tool to analyze the mechanism of uncontrolled cell growth. Experiments to characterize the HBGF receptors of the transfected cell lines are also under investigation.
Figure 37. Summary of the splicing patterns and tissue distribution of the three upstream exons of HBGF-1.

Panel A. Four splicing patterns of the 1.J, 1.5, and 1.H upstream exons. TATA boxes located at the most 5'-end of 1.J, 1.5 and 1.H exons have not been identified. ? indicates the possible splicing events occurring further upstream from the 1.J, 1.5 and 1.H exons to unknown regions.

Panel B. Tissue distribution of the four splicing patterns. (+) indicates the low levels of expression. ND stands for "not determined".
Figure 37

<table>
<thead>
<tr>
<th></th>
<th>BRAIN</th>
<th>KIDNEY</th>
<th>HEART</th>
<th>U1242MG</th>
<th>D65MG</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>1.J</td>
<td>Exon 1</td>
<td>(+)</td>
<td>+</td>
<td>ND*</td>
</tr>
<tr>
<td>b.</td>
<td>1.5</td>
<td>Exon 1</td>
<td>+</td>
<td>+</td>
<td>ND*</td>
</tr>
<tr>
<td>c.</td>
<td>1.5</td>
<td>Exon 1</td>
<td>-</td>
<td>+</td>
<td>(+)</td>
</tr>
<tr>
<td>d.</td>
<td>1.J</td>
<td>1.5</td>
<td>Exon 1</td>
<td>(+)</td>
<td>+</td>
</tr>
</tbody>
</table>
HBGF-1 mRNA Expression in Various Tissues

Northern hybridization showed HBGF-1 mRNA was present in fetal brain and kidney, but to a lesser extent than adult brain and kidney (Fig. 8). Later RNase protection assays also confirmed this result (Fig. 24). This suggests that HBGF-1 may play a role in embryonic angiogenesis during fetal development. In contrast, the int-2 gene is expressed in early gestation stages of development (Jakobovits et al., 1986), and hst/KS3 is expressed in the mouse blastocyst and undifferentiated embryonal carcinoma (EC) cells (Hebert et al., 1990).

Despite the fact that HBGF-1 mRNA is not detectable in either the adult or fetal liver (Jaye et al., 1986; Kan et al., 1989; Wang et al., 1989; Fig. 8; Fig. 24), it is expressed in regenerating rat hepatocytes. Expression of HBGF-1 mRNA occurs in both hepatocytes and nonparenchymal cells and persists for seven days after partial hepatectomy. Moreover, two different sizes, 4.8-kb and 1.3-kb, of HBGF-1 mRNA are expressed during different times after hepatectomy (Kan et al., 1989). The temporal expression of different sizes of mRNA further supports the idea of alternative promoters.

The HBGF-1 gene expression in the fetal heart had not been reported previously. Northern hybridization showed the expression of HBGF-1 mRNA in the human fetal heart for the first time (Fig. 8). Recently, Sasaki et al. (1989) purified a bovine heart-derived growth factor (HDGF) and showed that it was identical to HBGF-1 isolated from brain. Weiner and Swain (1989) reported that the primary culture of rat neonatal cardiac myocytes expressed HBGF-1 mRNA and deposited the protein into the extracellular matrix. Later, HBGF-1 was purified from several normal human hearts.
(Casscells et al., 1990). Thus, the embryonic development of the heart and the response of the intact organ to stimuli that may induce vascular growth (Paskyk et al., 1982), myocyte hypertrophy, and nonmyocyte cellular proliferation (Bugaisky and Zak, 1986) may be partly mediated by HBGF-1. HBGF-1 might also be responsible in part for the proliferation of various cell types found in the myocardium (i.e., fibroblasts, smooth muscle cells and endothelial cells) during development or in response to pathological processes, such as atherosclerosis (Klagsburn and Edelman, 1989) and tumor angiogenesis (Folkman and Klagsbrun, 1987). Studies using immunohistochemistry in conjunction with monoclonal antibodies generated against HBGF-2 have demonstrated that HBGF-2 is localized in striated muscle cells and their precursors in 2- through 6-day old chick embryos (Joseph-Silverstein et al., 1989). Immunoreactivity of HBGF-2 has also been reported in the nuclei and intercalated discs of cardiac myocytes (Kardami et al., 1989).

Presence of the angiogenic HBGF-1 in the fetal heart will pave the way for the study of its roles in this organ as well as in the vascular structures. We are currently developing in situ hybridization techniques to identify which cell types of the fetal heart express HBGF-1. It will also be possible to perform immunohistochemical studies to see if the expression of HBGF-1 in the myocardium follows the appearance of recognizable endothelial cells in this tissue as described (Manasek, 1971). Although it is shown that different sizes of HBGF-1 mRNA are expressed during liver regeneration (Kan et al., 1989), we detected only a single 4.5-kb transcript in RNA isolated from fetal heart. Thus, HBGF-1 may function and be regulated differently in the heart and liver.
Since the turnover of endothelial cells is very slow (Denekamp, 1984), it has been suggested that angiogenic factors are sequestered in an active form and released during tissue repair (Vlodavsky et al., 1987). Since HBGF-1 is one of the potent angiogenic factors, one would expect placenta to be a rich source of HBGF-1. However, I could not detect any HBGF-1 mRNA in fetal placenta or term placenta by Northern hybridization (Fig. 8). Later, the more sensitive RNase protection assays also failed to detect any HBGF-1 mRNA expression in placenta (Fig. 24). Unpublished results of in situ hybridization are consistent with these results (Ohlsson, R., and Chiu I.-M.). These results are surprising in view of the vigorous angiogenesis occuring in the placenta tissue. However, a definitive answer will have to await analysis by RT-PCR.

A previous report by Libermann et al. (1987) demonstrated that several human glioma cell lines expressing HBGF-1 mRNA. I also detected a much lower level of the 4.5 kb species of HBGF-1 mRNA in A2781 glioblastoma cells (Fig. 27). Several human glioblastoma cell lines not only express HBGF-1 mRNA, but also contain alternatively spliced untranslated exons (Myers, R., and Chiu, I.-M. unpublished results). Takahashi et al. (1990) reported that HBGF-1 mRNA was significantly expressed in gliomas in vivo. This result suggests that tumor-derived HBGF-1 was involved in the progression of gliomas and also contributes to tumor angiogenesis.
ISOLATION AND CHARACTERIZATION OF HUMAN HBGF-1 GENE

Several different HBGF-1 mRNA species resulting from alternative splicing has been described above. These results suggest that the gene structure and regulatory mechanisms of HBGF-1 are complex. The structural organization of the human HBGF-1 gene had not been characterized before this work.

Structural Organization of HBGF-1 Gene

As a first step toward understanding the HBGF-1 gene structure, I have isolated nine contiguous genomic DNA clones spanning 54-kbp of the HBGF-1 gene by screening genomic DNA libraries with an HBGF-1 cDNA probe and various end-specific walking probes (Fig. 20). The complete DNA sequence of all three coding exons was also determined (Fig. 22). The complete amino acid sequence of human HBGF-1 was deduced from the nucleotide sequence of these genomic DNA clones. The exon-intron structure was determined by Southern hybridization, nucleotide sequence analysis and RNase protection assays (Fig. 29). This gene organization is consistent with that of other genes in the HBGF family; namely, three protein-coding exons, long 3'-untranslated sequences, and intervening sequences in the same positions with respect to the coding sequences.

Mapping of the First Protein-Coding Exon

In order to investigate the possibility of alternative coding sequence for the amino terminus in different tissues which, in turn, may specify the presence or absence of signal peptides for the secretory and
other forms of HBGF-1, RNase protection assays were performed. These assays were also used to map the 5'-end of the first protein-coding exon before the corresponding region of cDNA clones were available for comparison studies.

RNase protection assays of the first protein-coding exon, in which RNA isolated from kidney, brain, fetal brain, fetal kidney, and glioblastoma cell line A2781 were used as templates (Fig. 24), suggested that this exon has a specific acceptor site in all RNA. However, since there are at least three species of HBGF-1 mRNA found in both brain and kidney (Fig. 27), I have not rigorously ruled out the possibility that the observed protected fragments in the assay are due to the minor species (3.4-kb, 2.0-kb, or both). Thus the major bands (4.5-kb) from brain and kidney may have slightly different protein-coding sequences at the amino terminus. In line with this notion is a very weak hybridization signal of the 4.5-kb species on Northern blot when the 1.0-kbp HindIII-BglII fragment containing the first exon was used as a probe. The weak signal was not likely to be due to a high G+C content since the G+C content of this DNA fragment is 50.6%.

Mergia et al. (1989) isolated some regions of the human HBGF-1 gene and used riboprobes generated from different fragments to map the 5'-end of the first coding exon. Their assignment is consistent with my RNase protection results. This 5'-end of the first coding exon does not represent the transcription start site, since it is bounded by a canonical splicing acceptor sequence and no CCAAT or TATA sequences can be detected at the correct distance. In addition, the size of the three exons of HBGF-1 together is 3.6-kbp which is smaller than the size of the major
4.5-kb HBGF-1 mRNA. These data suggest that there are additional sequences further upstream (as described previously).

Since a termination codon (UGA) was identified in phase with and three nucleotides upstream from the putative translation initiation site (Fig. 22), these upstream sequences are most likely untranslated. In a few cases UGA serves as a codon for the unusual twenty-first amino acid, selenocysteine. The examples include formate dehydrogenase H from E. coli, glutathione peroxidase from mouse and human (Zinoni et al., 1986; Chambers et al., 1986; Sukenaga et al., 1987), and type 1 iodothyronine deiodinase (Berry et al., 1991). There is a minor serine tRNA species specific for selenocysteine capable of recognizing this UGA opal nonsense codon (Leinfelder et al., 1988). Non-AUG codons were used for mouse int-2 and human HBGF-2 genes. Kozak (1990) demonstrated that the variability of mammalian ribosomes to initiate at non-AUG codons could be explained by the presence or absence of a stem-loop structure downstream from the codon used. However, this has never been identified in human HBGF-1.

Long 5'-untranslated sequence have been reported in a variety of growth factor genes and oncogenes. This feature also distinguishes HBGF-1 from other members of the HBGF family (Delli Bovi et al., 1987; Dickson and Peters, 1987; Yoshida et al., 1987; Zhan et al., 1988). These sequences may be important for translation regulation and other regulatory roles, as exemplified in the PDGF B-chain mRNA in which the 5'-untranslated sequence is a potent translation inhibitor (Rao et al., 1988).

Alternative Polyadenylation Sites of HBGF-1 Gene

A minor transcription termination site revealed by the RNase
protection assay corresponded to the 3'-end of the pHBGF-1.1 cDNA clone (Fig. 22). A polyadenylation signal sequence, AATAAA spanning nucleotide 2209 to 2214 in Fig. 21, is consistent with this result. There are at least two additional polyadenylation signals in the genomic DNA sequences residing 74-bp and 266-bp further downstream (Fig. 22). The data obtained from the RNase protection assay indicate that these two sites are the 3'-ends of two other minor transcripts. These results are also consistent with the result from cDNA cloning (Crumley et al., 1989)

RNase protection assays suggested that there are two major polyadenylation sites, approximately 10-bp apart, residing 3.1-kbp downstream from the translation termination codon (Fig. 22 and Fig. 26). Taken together, these data support the interpretation that the minor 3.4-kb band corresponds to the transcripts terminating approximately 1.8-2.0-kbp downstream from the translation termination codon, while the major 4.5-kb band ends 1.1-kbp further downstream.

Two partially conserved and mutually exclusive sequence element have been identified downstream from the poly (A) sites in eukaryotic genes. The sequence YGTGTYY, first identified by McLaughlan (1985), and the sequence TT(G/A)NNNTTTTTTT, identified by Renan (1987), are required for efficient formation of 3'-termini. The first three functional HBGF-1 polyadenylation sites have the latter 12-nt consensus sequence downstream. There are five more perfect AATAAA hexamers in the 3'-untranslated region. However, they are likely to be silent as suggested by my RNase protection assays.

The biological significance of the unusually long 3'-untranslated sequences is not known; however, it appears to be a general phenomenon
among the HBGF family. In one case, expression of the hst/KS3 proto-onco-
gene is controlled by regulatory elements present in the 3'-untranslated
region of the gene (Curatola and Basilico, 1990). Three possible octamer
motifs are present in the 3' hst/KS3 DNA sequences. This is the first
known location of enhancer elements in a region of the gene that is
transcribed as the 3'-untranslated portion of the mRNA.

The ATTTA motif shown to destabilize c-fos and GM-CSF mRNAs (Miller
et al., 1984; Shaw and Kamen, 1986) is repeated six times in the
3'-untranslated sequence of HBGF-1. Thus, HBGF-1 mRNA may be transcribed
in response to certain physiological conditions and then degraded soon
after. With the availability of cDNA and genomic DNA clones, the half
life of HBGF-1 mRNA and the destabilizing effect of ATTTA motifs warrant
further investigation.

**Sequence Conservation Among Different Species**

In addition to the coding region, the 3'-untranslated sequence of
HBGF-1 is highly conserved among different mammalian species. The
published bovine cDNA sequence (Halley et al., 1988) provided the longest
stretch of sequence for comparative analysis with our human DNA sequence.
Sequence similarity of 70% in a continuous stretch of 2400-bp in the
3'-untranslated region was observed between the human and bovine sequences
(Fig. 29). Additionally, the divergent region in the bovine cDNA clone is
a cloning artifact (M. Laurent, Paris, France, personal communication).
The 5'-untranslated sequences of the first protein-coding exon and the
entire protein-coding sequences show 88% and 91% sequence similarities,
respectively, between the human and bovine species. The corresponding
regions available for the rat and hamster HBGF-1 genes are also conserved to similar degrees. These data lend further support to our previous assumption that the apparent dissimilarity in the 5'-untranslated sequences among the four different species is not due to species divergence but rather to the selective cloning of different exons which are spliced alternatively to the first protein-coding exon.

RNase protection assays revealed two major polyadenylation signal sequences in the human gene. Sequences from the corresponding regions of other species are not available for comparison. The three minor polyadenylation signal sequences in the human gene were reflected in the bovine cDNA sequence (Halley et al., 1988). In contrast, the ATAAA sequence observed in the rat cDNA, which is presumably used for the synthesis of the rat-specific 1.3-kb HBGF-1 mRNA (Goodrich et al., 1989), is not present in either the human or bovine sequences. This high degree of similarity at the 3'-untranslated region is rather striking and indicates that it may be functionally significant.

Conservation of Upstream Intervening Sequences

Through the characterization of the 19-kbp stretch immediately upstream from the first protein-coding exon, one 1.4-kbp EcoRI fragment was shown to be conserved in both primate and rodent DNA (Fig. 30 A and B). The finding of DNA sequences in other mammals homologous to the 1.4-kbp EcoRI fragment suggests that the sequence in this DNA fragment has been conserved in evolution and may be expressed as a functional gene. This strategy, based on the fact that human DNA sequences that cross-hybridize to the DNA of other mammalian species often correspond to
expressed genes, has been used to clone several human disease genes included Wilm's tumor (Bonetta et al., 1990) and a tumor suppressor gene MCC (mutated in colorectal cancer) involved in colorectal tumors (Kinzler et al., 1991). To date, the nested gene in the mammals has been found only in the factor VIII locus (Levinson et al., 1990) and type 1 neurofibromatosis gene (Wallace et al., 1990).

Preliminary Northern hybridization failed to detect expression from this region in brain, kidney, or placenta. This may be due to a lack of the appropriate tissue sources or a low level expression. Other sensitive expression assays, like the "exon-connection" strategy based on PCR (Fearon et al., 1990), can be used to detect gene expression at low level. In brief, primers are derived from two putative exons of the gene and used in RT-PCR. Further characterization of the 1.4-kbp EcoRI fragment is likely to provide insight into the significance of this high degree conservation.

Cloning the Upstream Untranslated Genomic DNA

By using the upstream exon-specific probe, genomic DNA clones containing the upstream, untranslated exon have been isolated by R. Payson. These upstream clones, which contain the 1.1 and 1.5 upstream exons, do not overlap with the DNA clones shown in Fig. 20. This suggests that the introns are big (at least 59.9-kbp and 49.4-kbp for the 1.1 and the 1.5 exons, respectively), and the splicing patterns are rather complex (Summarized in Fig. 38).

The established PCR conditions using the HBGF 1201/1301 primers for the 1.1 exon was used to isolate six yeast artificial chromosome (YAC)
clones (in collaboration with Drs. C.A. Chinault and T. Caskey, Baylor College of Medicine, TX). As with most current YAC libraries, there is probably at least 9.2% chance of coligation events in any clone. Therefore, chromosome mapping with end probes generated by Alu-vector PCR (Nelson et al., 1989) or other methods (Silverman et al., 1989; Riley et al., 1990) to ensure the contiguity of the cloned region in the genome is under way. Further experiments using infrequently cutting restriction enzymes and pulse field gel electrophoresis will facilitate the mapping and cloning of the unknown gap.

RNase protection assays using genomic DNA fragment, which contained the 1.1 exon as riboprobes, identified that the entire 1.1 exon was 405-bp (Myers, R., and Chiu, I.-M. unpublished results). A putative TATA box started at position -26 relative to the 5'-end of the 1.1 exon was observed. Ongoing primer extension experiments will be used to confirm these results. Moreover, in vitro transcription assays (Farnham and Schimke, 1985) and transient expression of the chloramphenicol acetyl transferase (CAT) gene linked to this 5'-flanking sequence (Gorman et al., 1982) will be used to characterize this putative promoter region of the 1.1 exon.

Similar RNase protection assays indentified that the entire 1.5 exon was 132-bp. However, primer extension product was about 517-bp (Fig. 16). These results indicate that there is an unknown exon(s) residing further upstream from the 1.5 exon. We are currently cloning this exon(s) from cDNA libraries and to examine the regulatory sequences.
Figure 38. Summary of the genomic organization of HBGF-1, including the three upstream untranslated exons. The thick solid lines represent the cloned genomic DNA locus. Dashed lines represent a gap. Closed boxes represent the 1.J, 1.5 and three protein-coding exons. The 1.H exon is represented by an incomplete box to indicate the unknown size.
The intron downstream from the 1.5 exon contains an alternating purine:pyrimidine motif with the potential to form Z-DNA. Other members of the HBGF family, int-2 and hst/KS3, have been shown to have a similar dinucleotide repeat (Polymeropoulos et al., 1990; Tiesman and Rizzino, 1990). Z-DNA formation is involved in the breakpoints of chromosome during translocation at the 5'-flanking region of bcl-2 gene (Adachi and Tsujimoto, 1990). Further experiments using antibodies specific to Z-DNA could address the possibility of Z-DNA formation of HBGF-1 and its relationship with chromosome abnormalities.
Karyotypic analysis reveals that a significant fraction of patients with ANLL or 5q- syndrome have a deletion of a portion of the long arm of chromosome 5, 5q31. Thus, a leukemic suppressor gene may be present at 5q31 and its homozygous loss or inactivation may result in malignant transformation or tumor progression (Sager, 1989). In T-cell acute lymphoblastic leukemias (T-ALL) have lost both alleles of the retinoblastoma gene and have frequent mutations in the p53 tumor suppressor gene (Cheng et al., 1990; Cheng and Haas, 1990). The Philadelphia chromosome in chronic myelogenous leukemia (CML), on the other hand, involves a reciprocal translocation to activate the abl proto-oncogene by bcr (Kurzrock et al., 1989). If deletions of chromosome 5 follow the first mechanism, the cytogenetically normal chromosome 5 in ANLL patients may have small aberrations which can only be detected by molecular approaches. It is suggested that HBGF-1 gene may be such a suppressor gene, because of the following observations; first, the HBGF-1 gene resides in chromosome 5 band 5q 31.3-33.2 (Jaye et al., 1986) and is at or very close to the distal break-point of the 5q- deletion; and second, high concentration of HBGF-1 inhibit DNA synthesis in cultured fibroblasts (Chiu et al., 1990b; Ke et al., 1990). HBGF-2 has a permissive action on human hematopoietic progenitors as reported by Gabbianelli et al. (1990). Thus HBGF may have a role in hematopoiesis and the HBGF-1 gene may be such a suppressor gene.

**HBGF-1 and 5q- Chromosome**

To explore the possibility that the HBGF-1 gene is deleted in 5q-
chromosome, I elected to characterize the HBGF-1 gene locus in the ANLL patients. Using three different probes spanning 40-kbp, I couldn't detect any gross rearrangement or allelic loss with or without 5q- deletion (Fig. 33 to Fig. 35). Even though HBGF-1 has long been considered as a mitogen for cells of mesenchymal and neuroectodermal origin, recent isolation of an HBGF receptor from the K562 cell line (Partanen et al.,1990; Keegan et al., 1991) suggests that HBGF-1 may be involved in the growth and differentiation of the hematopoietic cells. Thus I chose to characterize the expression of HBGF-1 mRNA in mononuclear cells from normal controls as well as ANLL patients. The sensitive RT-PCR method to detect mRNA was employed for this study. I was not able to detect any difference in the expression pattern of this gene in the mononuclear cells of ANLL patients. Thus, the deletion at the 5q- chromosome most likely involves other growth factors, receptors, or gene coding for previously unidentified proteins. Pulse-field gel electrophoresis is currently being used to detect DNA rearrangements in the 100-1000-kbp range of the HBGF-1 locus. In at least one case, one allele of chromosome 5 has a large deletion involving PDGF-R, c-fms and HBGF-1, and the other allele has a small deletion in proximity to but not involving any of these three genes (Thorton, D. E., and Chiu, I.-M. manuscript submitted). This study will likely provide clues to the involvement of HBGF or its neighboring genes in the etiology of leukemia. The loss of such a gene might give rise to profound dysregulation of cellular proliferation and differentiation. Cloning of the DNA sequence flanking the breakpoints of 5q- chromosome may facilitate subtyping or prognosis of ANLL patients.
HBGF-1 research has seen an explosion of interest in the areas regarding gene family characterization, expression patterns, and biological activities. In the past several years, there has been an enormous increase in HBGF-1 research as a result of the development of protein isolation methods, availability of characterized molecular probes, specific antibodies and recombinant protein products. Among the more intriguing current questions are the mechanisms of secretion of HBGF-1 and subsequent deposition in extracellular matrix as latent signal, the interaction of HBGF-receptor system and signal transduction mechanism, tissue expression patterns and regulations of the complicated 5'-untranslated exons.

Further work on HBGF-1's molecular structure using various HBGF-1 cDNA clones will further elucidate the basis of its biological activities. Tissue-specific expression of different 5'-untranslated exons provides a system to study alternative promoter usage, alternative splicing and other post-transcriptional controls. This may lead to the identification of various tissue-specific transcription factors. Moreover, cloning of the longe range locus of HBGF-1 will provide insight into the role of this locus in human disease.
LIST OF REFERENCES


Chambres, I., Frampton, J., Goldfarb, P., Affara, N., McBain, W., Harrison, P.R. (1986). The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the termination codon, TGA. EMBO J. 5:1221-1227.


terminal differentiation occur in $G_1$ phase and is repressed by fibroblast growth factor. J. Cell. Biol. 105:949-956.


Mansour, S., Martin, G. (1988). Four classes of mRNA are expressed from the mouse int-2 gene, a member of the FGF gene family. EMBO J. 7:2035-2041.


