

## HEPARIN-BINDING EGF-LIKE GROWTH FACTOR: CHARACTERIZATION OF RAT AND MOUSE cDNA CLONES, PROTEIN DOMAIN CONSERVATION ACROSS SPECIES, AND TRANSCRIPT EXPRESSION IN TISSUES

Judith A. Abraham,<sup>1</sup> Deborah Damm,<sup>1</sup> Adrianna Bajardi,<sup>2,3</sup> Judy Miller,<sup>1</sup>  
Michael Klagsbrun,<sup>2</sup> and R. Alan B. Ezekowitz<sup>3</sup>

<sup>1</sup>Scios Nova Inc., Mountain View, CA 94043

<sup>2</sup>Departments of Surgical Research and Biological Chemistry and Molecular Pharmacology,  
Children's Hospital and Harvard Medical School, Boston, MA 02115

<sup>3</sup>Division of Hematology/Oncology, Children's Hospital and Dana-Farber  
Cancer Institute, Boston, MA 02115

Received November 24, 1992

---

**Summary:** Clones were obtained that encode the rat and mouse forms of heparin-binding EGF-like growth factor (HB-EGF), a potent mitogen for smooth muscle cells, fibroblasts and keratinocytes that is proposed to be derived from a transmembrane precursor. Within the HB-EGF precursor sequences predicted from these cDNAs, the region corresponding to the secreted ("mature") factor was found to represent one of the least well conserved areas when compared to human or monkey HB-EGF (73 - 76% sequence identity). Regions of high sequence conservation included the proposed juxtamembrane and transmembrane domains, as well as a proposed heparin-binding region within the "mature" factor. Northern blotting experiments using the HB-EGF clones as probes revealed HB-EGF transcript expression in multiple tissues, particularly lung, skeletal muscle, brain, and heart. © 1993 Academic Press, Inc.

---

The epidermal growth factor (EGF) protein family encompasses a number of mitogens that appear to share not only amino acid sequence similarity, but also the property of being derived from transmembrane precursors (1 - 5). One such mitogen is heparin-binding EGF-like growth factor (HB-EGF), which was first purified from the conditioned medium of the human monocyte/macrophage-like cell line U-937 (3). Structural characterization studies have demonstrated that the HB-EGF secreted from U-937 cells can contain at least 86 amino acids, with the carboxyl-terminal half of the protein sharing approximately 40% sequence identity with human EGF (3,6). The amino-terminal portion of the secreted factor has no counterpart in EGF (3) and contains a highly hydrophilic stretch of amino acid residues that has been suggested as

---

**Abbreviations:** EGF, epidermal growth factor; HB-EGF, heparin-binding EGF-like growth factor; bp, base pair; kb, kilobase.

a potential heparin-binding region (6). Analysis of human cDNA clones has indicated that the secreted HB-EGF is derived from the middle portion of a 208-residue precursor (3); the precursor itself has been proposed to be an integral membrane protein since, as with other mitogens in the EGF family (1,2,4,5), the carboxyl-terminal portion of the precursor contains a strongly hydrophobic domain that is closely followed by charged residues [a putative "stop transfer" signal (7)].

Recently, Naglich *et al.* (8) reported that a cell-surface associated, presumably transmembrane form of monkey HB-EGF can serve as a receptor mediating the binding and uptake of diphtheria toxin into cells. This result has suggested that, as has been speculated in the case of other EGF-like mitogens (1,4,9), membrane-anchored forms of the factor could be playing key biological roles, and precursor sequences not included in the "mature" secreted mitogen might be critical for functions beyond the biosynthesis and release of the soluble factor. We have examined this possibility further by isolating and characterizing cDNA clones encoding the rat and mouse forms of the HB-EGF precursor, since critical functional domains in the precursor might be expected to be highly conserved across species. In addition, comparison of the predicted sequences of "mature" rat and mouse HB-EGF to the sequence of human (U-937-derived) HB-EGF has suggested regions of the secreted factor important for biological function. We have also made use of the rat and mouse cDNAs (along with a previously-isolated human cDNA) to examine the tissue distribution of HB-EGF gene expression.

## MATERIALS AND METHODS

***Isolation of rat HB-EGF cDNA clones.*** The human HB-EGF cDNA fragment JMU7-1, which spans 461 base pairs (bp) of protein coding sequence and 622 bp of 3' untranslated sequence, was labeled with  $^{32}\text{P}$  by nick-translation and used to probe a Lambda ZAP II cDNA library (#936510; Stratagene) made from the RNA of SV40-transformed rat peritoneal macrophages. Filter hybridizations were carried out at 42°C in 40% formamide, 50 mM sodium phosphate (pH 6.5), 5 X Denhardt's solution (1 X Denhardt's is 0.02% Ficoll, 0.02% polyvinylpyrrolidone, and 0.02% bovine serum albumin), 50 µg/ml denatured DNA, and 5 X SSC (1 X SSC is 150 mM NaCl, 15 mM sodium citrate, pH 7.0). Filter washes were at 47°C in 1 X SSC, 0.1% SDS. Two of the positively - hybridizing phage (λRat4-1 and λRat7-6) were subjected to sequence analysis. Additional clones were obtained by screening a λgt10 rat macrophage cDNA library (RL1028a; Clontech), using the 1.5 kilobase (kb) cDNA insert from λRat7-6 as the probe. Hybridizations in this screen were at 42°C in 50% formamide, 5 X Denhardt's solution, 20 µg/ml denatured DNA, and 5 X SSC; washes were at 50°C in 0.1 X SSC, 0.1% SDS. One of the positively - hybridizing phage detected (λRat1-2) was purified for sequence analysis. DNA sequencing was carried out using the dideoxy method (10) after subcloning of fragments into M13 vectors.

***Isolation of mouse HB-EGF cDNA clones.*** The 0.6 kb cDNA insert in λRat4-1 was labeled with  $^{32}\text{P}$  and used to screen a mouse peritoneal macrophage cDNA library (11; gift of Dr. P. Crocker, Molecular Medicine Institute, University of Oxford) prepared in the Lambda ZAP II vector (Stratagene). Hybridizations were at 50°C in 50mM sodium phosphate (pH 6.5), 5 X Denhardt's solution, 50 µg/ml denatured DNA, and 5 X SSC; washes were at 50°C in 2 X SSC. The sequence of the cDNA insert in one of the hybridizing phage, λMouse1, was partially

determined by double-stranded sequencing using the Sequenase protocol (Version 2.0 DNA sequencing kit; United States Biochemical).

***Northern analyses of tissue RNAs.*** Human, mouse, and rat Multiple Tissue Northern (MTN) blots were obtained from Clontech. Each of these blots represented poly(A)<sup>+</sup> RNA samples that had been isolated from a variety of different tissues, fractionated on a denaturing formaldehyde agarose gel (2 µg RNA/lane), and transferred to a nylon membrane. The blots were probed according to the manufacturer's instructions with either the human cDNA fragment JMU7-1 (for the human MTN blot), the 1.5 kb cDNA insert in λRat7-6 (for the rat MTN blot), or a 1.6 kb KpnI- EcoRI restriction fragment derived from the 3' end of the cDNA insert in λMouse1 (for the mouse MTN blot).

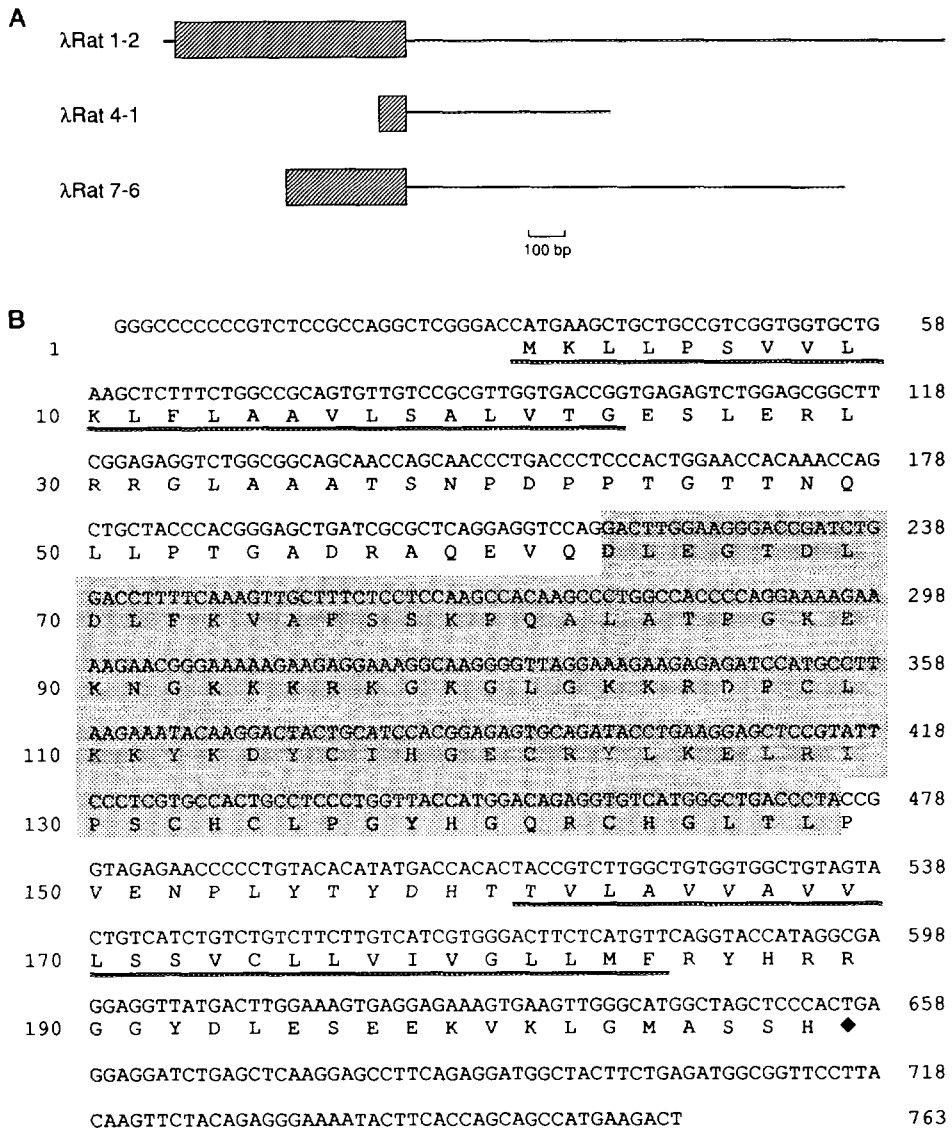
## RESULTS AND DISCUSSION

***Cloning and analysis of rat and mouse HB-EGF cDNAs.*** Given that HB-EGF was originally purified from the conditioned medium of the human monocyte/macrophage-like U-937 cell line (3), and HB-EGF-like activity has been reported in conditioned media samples collected from human peripheral blood monocytes and macrophages (12), we chose cDNA libraries derived from macrophages to screen for rat and mouse HB-EGF clones. A fragment of human HB-EGF cDNA (3) was first used to screen a cDNA library made from SV40-transformed rat peritoneal macrophages. None of the positively-hybridizing clones detected in this screen contained the entire coding region for rat HB-EGF; however, two of the clones (λRat4-1 and λRat7-6; Fig. 1A) contained portions of the coding region. The cDNA insert in λRat7-6 was then isolated and used to screen a second rat macrophage cDNA library. Restriction mapping and DNA sequencing revealed that one of the clones detected in this screen, λRat1-2, contained an approximately 2.1 kb cDNA insert (Fig. 1A) that spanned the full 208-codon HB-EGF precursor coding region (Fig. 1B).

To obtain a clone encoding mouse HB-EGF, the cDNA insert from clone λRat4-1 was isolated and used as a probe to screen a cDNA library made from mouse peritoneal macrophages. One of the positively-hybridizing clones detected, λMouse1, was found to contain a cDNA insert that included the entire HB-EGF coding region (Fig. 2).

The precursor sequences predicted from the cDNA clones of rat and mouse HB-EGF are compared to the predicted precursors of human (3) and monkey (8) HB-EGF in Fig. 3. All four predicted precursors are 208 residues in length with no insertions or deletions between species, and all contain the two hydrophobic stretches of amino acids that have been previously proposed to represent a secretion signal peptide at the amino terminus and an internal transmembrane domain (3,6,8,13). Overall, the human and monkey HB-EGF precursors share 97% sequence identity (8), while the rat and mouse precursors are 92% identical (Fig. 4); between the primate precursors and the rodent precursors, however, there is only about 80% sequence identity.

A comparison of the degrees of conservation in various regions of the HB-EGF precursor revealed that two of the least conserved areas correspond to the proposed propeptide and



**Figure 1.** Characterization of rat HB-EGF cDNA clones. (A) Schematic diagram comparing the extents of the cDNA inserts in rat HB-EGF clones  $\lambda$ Rat1-2,  $\lambda$ Rat4-1, and  $\lambda$ Rat7-6. Hatched boxes indicate the portion of each clone that represents protein coding region sequence. (B) Partial nucleotide sequence of the cDNA insert in clone  $\lambda$ Rat1-2, and predicted primary translation product for rat HB-EGF. Nucleotides are numbered at the end of each line on the right side of the figure; predicted amino acids encoded by the nucleotide sequence are given in one-letter code and are numbered on the left side of the figure. Bold underlines indicate two strongly hydrophobic regions in the predicted precursor, representing a putative secretion signal peptide (amino acids 1 - 23) and a proposed transmembrane domain (amino acids 161 - 184). The shaded area indicates a stretch of 86 amino acids that, by analogy to human U-937-derived HB-EGF (6), is predicted to be contained in secreted rat HB-EGF. The 5' ends of the cDNA inserts in clones  $\lambda$ Rat4-1 and  $\lambda$ Rat7-6 lie at nucleotides 580 and 331, respectively, in the sequence shown. The nucleotide sequence derived for rat HB-EGF has been submitted to GenBank (accession number L05489).

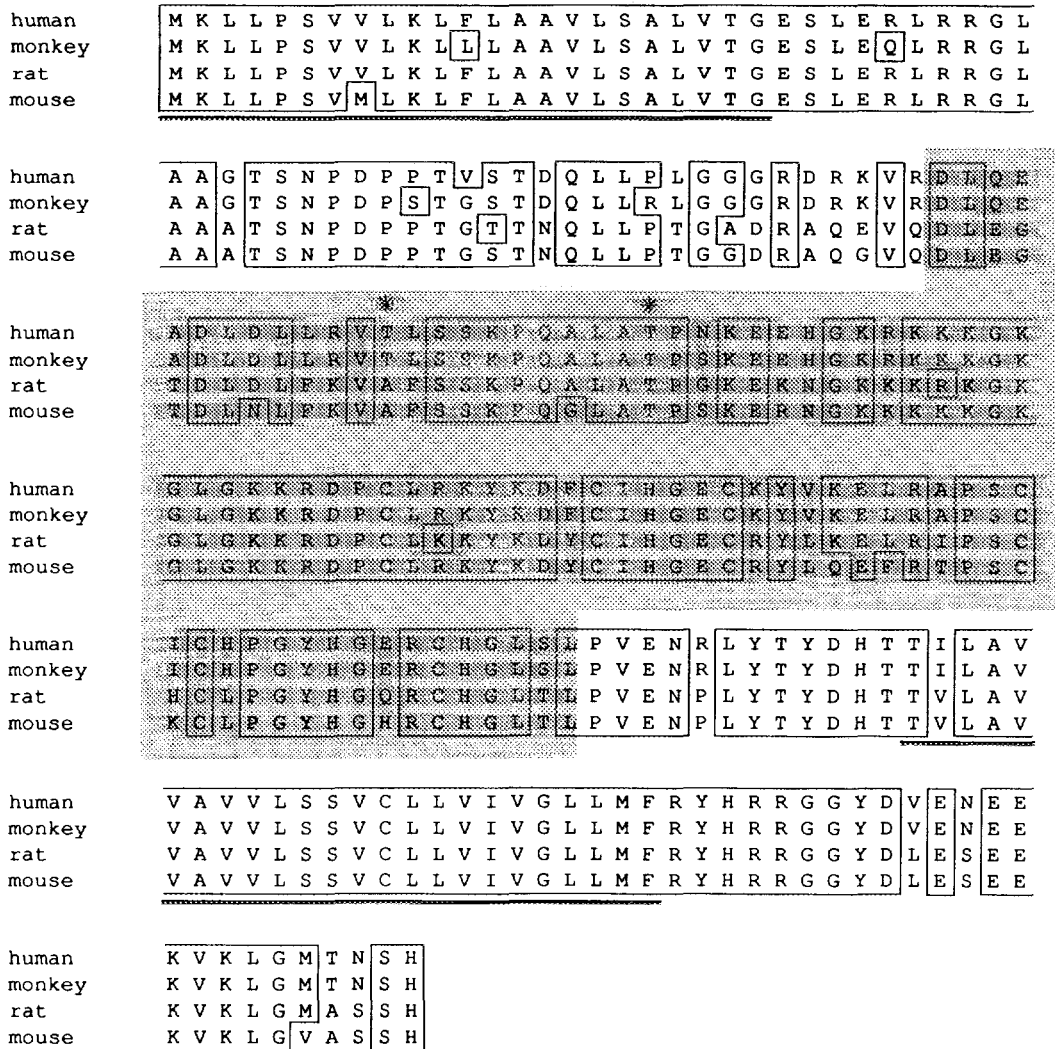
```

CGGCCGCCAGACCTTCAAGGGCTGGAGTGGACGCGCGGACCGACTCTGAACAGACAGACG 60
AACCGGCCGCCAAGGTTCCAGACAGGATCTACCCAGAGGCAGGCAGCGGACAGTGCC 120
TTAGTGGAACCTCGTGTCTCCACCGCCTGGCCCCGGTGCAGGCGTCCAGTGCCCGCCG 180
CATCCAAAGTGATCGTGCCTCCCGTCTCCGCCAGCTCGGGACCATGAAGCTGCTGCCG 240
1          M K L L P
TCGGTGATGCTGAAGCTCTTCTGGCCGAGTGTGTCCGCGTTGGTGACCGGTGAGAGT 300
6 S V M L K L F L A A V L S A L V T G E S
CTGGAGCGGCTTCGGAGAGTCTGGCGGCAGCAACCAGCAACCCCTGACCCCTCCCACTGGA 360
26 L E R L R R G L A A A T S N P D P P T G
TCCACAAACCAGCTGCTACCCACGGGAGGTGATCGTGTCTCAGGGGGTCCAGGACTTGGAA 420
46 S T N Q L L P T G G D R A Q G V Q D L E
GGGACAGACTGAACCTTTTCAAAGTTGCTTCTCCTCCAAGCCACAGGCCCTGGCCACC 480
66 G T D L N L F K V A F S S K P Q G L A T
CCAAGCAAAGAAAGGAATGGAAAAGAAGAAGAAAGGAAAGGGTTAGGGAAGAAGAGA 540
86 P S K E R N G K K K K G K G L G K K R
GACCCATGCCTCAGGAAATACAAGGACTACTGCATCCACGGGAGTGCAGATACTGCAG 600
106 D P C L R K Y K D Y C I H G E C R Y L Q
GAGTTCGGTACTCCCTCTTCAAATGCCTCCCTGGTTACCCAGGACACAGGTGTCATGGG 660
126 E F R T P S C K C L P G Y H G H R C H G
CTGACTCTACCAGTGGAGAATCCCTATACATATGACCACACTACAGTCTTGGCTGTG 720
146 L T L P V E N P L Y T Y D H T T V L A V
GTGGCTGTAGTACTGTCGTCCTGTCTTCTTGTTCATCGTGGGACTTCTCATGTTTAGG 780
166 V A V V L S S V C L L V I V G L L M F R
TACCACAGGAGAGGAGGTTATGACTTGGAAAGTGAAGAGAAAGTGAAGTTGGGCGTGGCT 840
186 Y H R R G G Y D L E S E E K V K L G V A
AGTCCCACTGAGGAGGACCTGAGCTATAGGAACCTTCAGAGGCTACTTCTGAGACAGTG 900
206 S S H ♦
GTTCGTTACAGTCTACATAGAGGAGAATATTTACCAGCAGCCATGAAAACGTCTTC 960
ATCATTTCAGTTGCTACCTGACTGGCCCTCCTGTAAT 1000

```

**Figure 2.** Partial nucleotide sequence of the cDNA insert in clone  $\lambda$ Mouse1, and predicted primary translation product for mouse HB-EGF. Underlines, shading, and residue numbering are as in Fig. 1. The nucleotide sequence derived from  $\lambda$ Mouse1 was submitted to GenBank and given the accession number L07264.

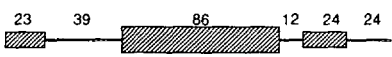
cytoplasmic domains (Fig. 4). Interestingly, the region corresponding to the 86 residues that have been identified in secreted forms of human HB-EGF ["mature" HB-EGF (6)] is also one of the least-well conserved: while there is 99% sequence identity between the human and monkey HB-EGF sequences in this region and 87% identity between the rat and mouse HB-EGF sequences, there is only 73 - 76% identity when the primate sequences in this region are compared with either rat or mouse (Fig. 4). By comparison, mature human and rat EGF share 68% sequence conservation, similar to what is reported here for HB-EGF, but the human and rat forms of another EGF family member -- transforming growth factor- $\alpha$  -- are much more highly conserved,



**Figure 3.** Comparison of predicted HB-EGF precursor sequences. Residues conserved in at least three of the four species are indicated by the outlines. Bold underlines indicate the proposed signal peptide and transmembrane domain; gray shading marks the region corresponding to the amino acids so far identified in forms of human HB-EGF secreted from U-937 cells. Asterisks mark two proposed sites of O-linked glycosylation in human HB-EGF (3,6).

sharing 92% sequence identity (1). It is also interesting to note that only one of the two threonine residues proposed as sites of O-linked glycosylation in human HB-EGF (3,6) is conserved in the rat and mouse forms of the factor (Fig. 3).

As shown in Fig. 4, the highest degree of sequence conservation in the HB-EGF precursor is most consistently found in the signal peptide, juxtamembrane, and transmembrane domains, suggesting strong selective pressure on these regions. While the high degree of conservation could be dictated by the events involved in the biosynthesis of secreted HB-EGF, these results



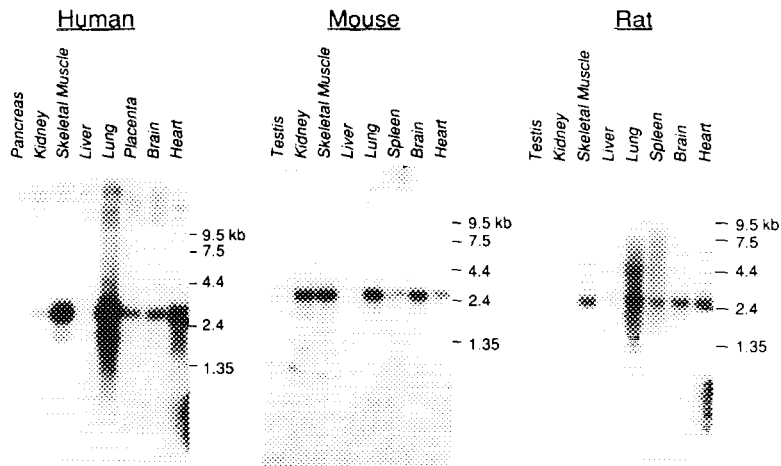
		sig	pro	"mature"	JM	TM	cyto	overall
human	vs. monkey	96	90	99	100	100	100	97
	vs. rat	100	72	76	92	96	83	82
	vs. mouse	96	77	73	92	96	79	81
monkey	vs. rat	96	67	76	92	96	83	80
	vs. mouse	91	72	74	92	96	79	80
rat	vs. mouse	96	92	87	100	100	96	92

**Figure 4.** Sequence conservation in various regions of the predicted HB-EGF precursor. A proposed structure for the HB-EGF precursor is shown schematically at the top of the figure, and is divided into signal peptide ("sig"), propeptide ("pro"), "mature", juxtamembrane ("JM"), transmembrane ("TM"), and cytoplasmic ("cyto") domains. The number of amino acids in each proposed domain is given above the schematic drawing. The column below each proposed domain gives the percent sequence identity in that domain for each of the species comparisons listed to the left. The overall sequence identity in the whole HB-EGF precursor in each species comparison is given in the right-most column.

would also be consistent with a model wherein these regions played other key roles, either separately or as part of a transmembrane form of HB-EGF. Mutagenesis studies may help to elucidate the actual functional significance of each of these regions.

Although overall sequence homology is only 73 - 76% in the "mature" HB-EGF domain, smaller portions of this domain show much higher conservation. One such region is the sequence KRKKKGKGLGKKRDPCLRKYK (residues 93 - 113 of the precursor) which, except for conservative substitutions of lysine for arginine or vice versa, is completely conserved in human, monkey, rat, and mouse HB-EGF (Fig. 3). Coincidentally, *in vitro* mutagenesis and peptide analysis studies have indicated that the amino acids responsible for human HB-EGF binding to heparin lie principally within this sequence (S. Thompson, S. Higashiyama, K. Wood, M. Klagsbrun, and J. Abraham, manuscript in preparation), and peptides derived from this sequence appear to interfere with the ability of HB-EGF to interact with low affinity, heparin-like binding sites on cells (S. Higashiyama and M. Klagsbrun, unpublished observations). Studies on two other heparin-binding mitogens, basic fibroblast growth factor and vascular endothelial growth factor, have indicated that interaction of these factors with high affinity receptors on cells is dependent upon the presence of heparin-like molecules (14 - 16); the high degree of conservation in the proposed heparin-binding region of HB-EGF may indicate that binding to cell-surface heparin-like molecules is also critical to the biological activity of this growth factor.

One interesting aspect to the report (8) that a cell-associated form of HB-EGF can serve as a diphtheria toxin receptor is that, unlike many human and monkey cell types, rat and mouse cells are resistant to this toxin (17 - 19). If this resistance is due to an inability of the toxin to



**Figure 5.** Detection of HB-EGF transcripts in tissues. Northern blots of human, mouse, and rat RNA samples from various tissues were probed as described in Materials and Methods. Autoradiographs of each blot are shown. Migration positions of RNA size standards are indicated to the right of each blot; the size of each standard is given in kilobases (kb).

bind the rat or mouse forms of HB-EGF, then clues as to the site of toxin binding might be obtained by examining the sites of sequence differences in rat and mouse HB-EGF relative to the primate forms (Fig. 3). This line of reasoning would implicate the carboxyl-terminal portion of the "pro" domain and the amino-terminal portion of the "mature" sequences as potential sites for toxin binding, since sequence divergence is greatest in these areas.

***Tissue distribution of HB-EGF gene expression.*** To examine the distribution of HB-EGF gene expression in the adult body, fragments of the rat and mouse cDNA clones were used to probe Northern blots of poly (A)<sup>+</sup> RNAs derived from a variety of tissues. In addition, a human HB-EGF cDNA fragment was used to probe a blot of human tissue RNAs. The results are shown in Fig. 5. Although there appeared to be some variation in the expression pattern between species, the approximately 2.5 kb HB-EGF transcript was consistently detected in a number of tissues, particularly lung, skeletal muscle, brain and heart; relatively low transcript levels were detected in liver in all three species. With regard to a diphtheria toxin receptor role for HB-EGF, the expression of the HB-EGF gene in multiple tissues is consistent with observations that diphtheria toxin does not appear to target a particular organ *in vivo*, but rather produces tissue necrosis and physiological changes in numerous organs (20). Experiments are currently in progress using *in situ* hybridization to localize more precisely the cells responsible for HB-EGF transcription in the different tissues.

#### ACKNOWLEDGMENTS

This work was supported in part by National Institutes of Health grants GM47397 (to M.K.) and AI23786-05 (to R.A.B.E.). R.A.B.E. is an Established Investigator of the American Heart Association.



## REFERENCES

1. Carpenter, G., and Wahl, M. I. (1990) In *Peptide Growth Factors and Their Receptors I* (M. B. Sporn and A. B. Roberts, Eds.), *Handbook of Experimental Pharmacology* Vol. 95, pp. 69 - 169, Springer-Verlag, Berlin.
2. Plowman, G. D., Green, J. M., McDonald, V. L., Neubauer, M. G., Disteché, C. M., Todaro, G. J., and Shoyab, M. (1990) *Mol. Cell. Biol.* 10, 1969 - 1981.
3. Higashiyama, S., Abraham, J. A., Miller, J., Fiddes, J. C., and Klagsbrun, M. (1991) *Science* 251, 936 - 939.
4. Wen, D., Peles, E., Cupples, R., Suggs, S. V., Bacus, S. S., Luo, Y., Trail, G., Hu, S., Silbiger, S. M., Ben Levy, R., Koski, R. A., Lu, H. S., and Yarden, Y. (1992) *Cell* 69, 559 - 572.
5. Holmes, W. E., Sliwkowski, M. X., Akita, R. W., Henzel, W. J., Lee, J., Park, J. W., Yansura, D., Abadi, N., Raab, H., Lewis, G. D., Shepard, H. M., Kuang, W.-J., Wood, W. I., Goeddel, D. V., and Vandlen, R. L. (1992) *Science* 256, 1205 - 1210.
6. Higashiyama, S., Lau, K., Besner, G. E., Abraham, J. A., and Klagsbrun, M. (1992) *J. Biol. Chem.* 267, 6205 - 6212.
7. Sabatini, D. D., Kreibich, G., Morimoto, T., and Adesnik, M. (1982) *J. Cell Biol.* 92, 1 - 22.
8. Naglich, J. G., Metherall, J. E., Russell, D. W., and Eidels, L. (1992) *Cell* 69, 1051 - 1061.
9. Massagué, J. (1990) *J. Biol. Chem.* 265, 21393 - 21396.
10. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., and Roe, B. A. (1980) *J. Mol. Biol.* 143, 161 - 178.
11. Harris, N., Super, M., Rits, M., Chang, G., and Ezekowitz, R. A. B. (1992) *Blood*, in press.
12. Besner, G., Higashiyama, S., and Klagsbrun, M. (1990) *Cell Regulation* 1, 811 - 819.
13. von Heijne, G. (1986) *Nucleic Acids Res.* 14, 4683 - 4690.
14. Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P., and Ornitz, D. M. (1991) *Cell* 64, 841 - 848.
15. Rapraeger, A. C., Krufka, A., and Olwin, B. B. (1991) *Science* 252, 1705 - 1708.
16. Gitay-Goren, H., Soker, S., Vlodavsky, I., and Neufeld, G. (1992) *J. Biol. Chem.* 267, 6093 - 6098.
17. Middlebrook, J. L., and Dorland, R. B. (1977) *Can. J. Microbiol.* 23, 183 - 189.
18. Middlebrook, J. L., Dorland, R. B., and Leppla, S. H. (1978) *J. Biol. Chem.* 253, 7325 - 7330.
19. Stenmark, H., Olsnes, S., and Sandvig, K. (1988) *J. Biol. Chem.* 263, 13449 - 13455.
20. Collier, R. J. (1975) *Bacteriological Rev.* 39, 54 - 85.