Biochemical and Biophysical Research Communications **261,** 183–187 (1999)
Article ID bbrc.1999.0962, available online at http://www.idealibrary.com on IDEAL®

Cloning and Functional Expression of a Human Heparanase Gene

Paul H. Kussie,[1] Jeffrey D. Hulmes, Dale L. Ludwig, Sheetal Patel, Elizabeth C. Navarro, Andrew P. Seddon,[2] Nicholas A. Giorgio, and Peter Bohlen
ImClone Systems, Inc., 180 Varick Street, New York, New York 10014

Received June 4, 1999

We have cloned a gene (*HSE1*) from a human placental cDNA library that encodes a novel protein exhibiting heparanase activity. The cDNA was identified through peptide sequences derived from purified heparanase isolated from human SK-HEP-1 hepatoma cells. *HSE1* contains an open reading frame encoding a predicted polypeptide of 543 amino acids and possesses a putative signal sequence at its amino terminus. Northern blot analysis suggested strong expression of *HSE1* in placenta and spleen. Transient transfection of *HSE1* in COS7 cells resulted in the expression of a protein with an apparent molecular mass of 67-72 kDa. HSE1 protein was detectable in conditioned media but was also associated with the membrane fraction following cell lysis. The HSE1 gene product was shown to exhibit heparanase activity by specifically cleaving a labeled heparan sulfate substrate in a similar manner as purified native protein.

The proteoglycan form of heparan sulfate (HS) synthesized by cells becomes incorporated into the cell membrane and the extracellular matrix (ECM). The endoglycosidases that cleave the HS proteoglycans into smaller fragments are called heparanases; these enzymes may assist in the remodeling of the ECM and aid migration of cells such as macrophages (1) and tumor cells (2, 3). Heparanase activity has been shown to be related to metastatic potential (4, 5) and may be involved in inflammation (6) and angiogenesis (7).

Heparanase activity has been reported in platelets, placenta, CHO cells, melanoma, hepatoma, and prostate carcinoma (8–13). The 9 kDa connective tissue activating peptide-III (CTAP-III) has been reported to have heparanase activity when isolated from human platelets (14) or expressed recombinantly (15). A separate heparanase activity has been described and partially purified from human platelets, placenta, and CHO cells with an apparent MW ranging from 30–134 kDa (8, 10, 16–20). At the present time it is not clear whether these size discrepancies are due to multiple proteins or alternate forms of the same protein. In this study, we describe the peptide sequence analysis of a protein derived from a hepatoma cell line and the cloning of a novel cDNA that encodes a protein with heparanase activity.

MATERIALS AND METHODS

Peptide isolation and amino-terminal sequence determination. A protein extract from SK-HEP-1 cells containing heparanase activity was a gift from Dr. Israel Vlodavsky (Hadassah University Hospital, Jerusalem). Heparanase was purified by methods similar to those previously described, using consecutive chromatography steps on CM-Sepharose, heparin, Concanavalin A and Mono S columns (12, 19). Fractions containing heparanase activity were precipitated with trichloroacetic acid and digested with trypsin as described (21). The precipitated pellet was dissolved in 50 μL of 0.4 M NH_4HCO_3 with 8 M urea, reduced with 4 mM dithiothreitol for 15 min at 37°C and carboxamidomethylated by addition of iodoacetamide to 8 mM at room temperature. The mixture was diluted 4-fold with water before 0.2 μg of trypsin was added and the sample was incubated for 16 hours at 37°C. Tween-20 (to 0.025%) was added to improve recovery and the tryptic peptides were separated by reverse phase HPLC using a C8 column (Vydac). Peak fractions (monitored at 210 nm) were collected and subjected to automated Edman sequencing on a ABI model 477A pulsed-liquid sequencer equipped with an on-line PTH analyzer (PE Applied Biosystems Inc.).

Database search. Peptide sequences were used as queries to search the EST database with the tblastn sequence comparison program (22) installed at the National Center for Biotechnology Information server (http://www.ncbi.nlm.nih.gov). I.M.A.G.E. Consortium expressed-sequence tag (EST) cDNA #442803 (GenBank accession number N41349) was obtained from Genome Systems Inc. (St. Louis, MO).

Northern blot analysis. An 890-bp DNA fragment encoding a portion of the EST 442803 ORF was amplified by PCR using the forward primer 5′-GATATTTTCATCAATGGGTCGC-3′ and reverse primer 5′-TATATGAGAAAGCTGGCAAGCC-3′. The gel purified fragment was labeled with [α-^{32}P]-dCTP (Prime-It II Random Primer Labeling Kit, Stratagene) and hybridized to a Human Multiple

[1] To whom correspondence should be addressed. Fax: 212-645-2054.
[2] Present address: Pfizer Inc., Central Research Division, Eastern Point Road, Groton, Connecticut 06340.

0006-291X/99 $30.00

TABLE 1

Sequence Summary of Tryptic Heparanase Peptides from SK-HEP-1

Fraction number	Approximate yield (pmol)	Peptide sequence
10	30	L-L-R
20	20	S-F-L-K
24	30	(L)-Y-G-P-D-V-G-Q-P-R
27	50	G-L-S-P-A-Y-L-R
35	25	(S)-V-Q-L-(N)-(G)-(L)-(T)-L-
67	5	(Q)-V-F-F-E-A-(G)-N-Y-(H)-L-V-D-E-N-F-

Note. Residues in parentheses are ambiguous calls.

Choice Northern Blot (Origene). Hybridization was performed in ExpressHyb solution (Clontech) and washed according to manufacturer's instructions. The same blot was stripped and rehybridized under the same conditions with a human β-actin cDNA probe to check quantities of RNA loaded onto each lane.

cDNA library screen. A human placental cDNA library was screened by PCR (Origene Technologies, Inc.) to obtain the full-length gene. The screen utilized an arrayed oligo-dT primed library cloned into pCMV6-XL3. Three primers (forward 5′-TCCTGATGTTGGTCAGCCTCG-3′, reverse 5′-GTAACCGCAAGTACTTGGTGACG-3′, reverse 5′-TTGAACCTTGCACGCTTGCC-3′) corresponding to regions within the open reading frame of EST 442803 were used for PCR amplification screening. Positive clones were rescreened using a forward vector-specific primer and reverse gene-specific primer to estimate the clone length. DNA sequence analysis was performed by dideoxy sequencing (University of Georgia Molecular Genetics Facility, Athens, GA). Computer assisted analysis of the translated *HSE1* gene was performed using PSORT (23) at http://psort.nibb.ac.jp:8800/.

Transient expression in COS cells. To generate a fusion protein with a carboxyl-terminal V5-epitope, the full open reading frame of the *HSE1* gene, including 12 bases upstream of the start codon, was cloned into the plasmid pcDNA3.1/V5-His (Invitrogen) by PCR using Vent polymerase (New England Biolabs). The fragment was cloned using the EcoRI and XbaI sites of the plasmid so that the *HSE1* gene was fused at the carboxy-terminal to a V5 epitope and poly-histidine encoding sequence, contributing 29 additional amino acids to the predicted size. Control transfections used vector alone. Transient transfection of COS7 cells was performed using Lipofectamine (Life Technologies, Inc.) according to manufacturer's instructions and cells were cultured in serum-free media (OptiMEM, Life Technologies, Inc.). Three days after transfection, conditioned media and COS7 cells were harvested. The COS7 cells were processed following published protocols for the isolation of heparanase from platelets (8). The cells were subjected to repeated freeze/thawing and centrifuged to separate soluble and insoluble fractions. Subsequently, the insoluble fraction was treated with mild detergent exposure (1% CHAPS for 30 minutes at 25°C, followed by centrifugation) in an attempt to solubilize membrane-associated protein as described (17). Resulting soluble and insoluble fractions were run on an SDS-PAGE gel and the recombinant fusion-protein was detected by Western blotting using anti-V5 monoclonal antibodies (Invitrogen).

Assays for heparanase activity. Heparanase activity was defined as the ability to degrade high molecular weight (40–100 kDa) radiolabeled HS substrate into low molecular weight (5–15 kDa) HS fragments (2, 10, 12, 24) that can be differentiated by gel filtration chromatography (19, 25). Radiolabeled HS substrate was prepared by metabolically [^{35}S]sulfate-labeling the extracellular matrix deposited by bovine corneal endothelial cells as described (19, 25). Soluble substrate was made by releasing the ^{35}S-labeled HS proteoglycans from the culture plate with trypsin. Samples to be tested were incubated at 37°C for 1.5–16 h in 10 mM sodium phosphate-citrate buffer, pH 6.0, with 20–25,000 cpm of ^{35}S-labeled HS substrate. Specificity of heparanase activity was verified by inhibition of HS cleavage by including heparin (10 μg/mL) in the reaction tube. Characterization of ^{35}S-sulfate-labeled HS product by size was done using Superose 6 gel filtration chromatography (24, 25). Blue dextran and phenol red were added to the sample to mark the excluded (Vo) and included (Vt) volumes, respectively.

RESULTS AND DISCUSSION

Heparanase-active protein isolated from human hepatoma cells was digested with trypsin and fragments were separated by reverse-phase HPLC. Eluted fractions containing peptides (5–50 pmol) were subjected to N-terminal sequence analysis. Several of the peptide sequences matched the protease-inhibitor nexin-I sequence (26). To rule out nexin-I as a heparanase, purified nexin-I was tested in a HS degradation assay, but no heparanase activity was detected. Other peptide sequences (Table 1) were obtained which did not match nexin-I or any characterized protein sequence in our database search, indicating that these peptides originated from a novel protein.

Database search. Using the novel peptide sequences we obtained (Table 1), translated sequences within the EST database at the National Center for

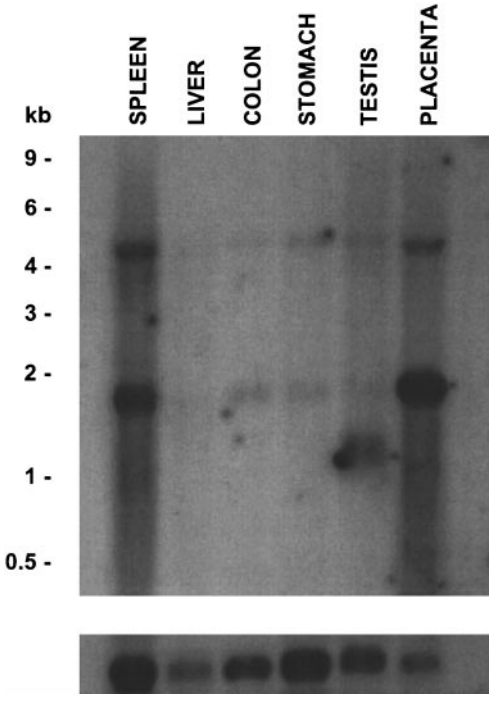

FIG. 1. Tissue specificity of heparanase expression in human tissues. A human multiple choice Northern blot (Origene), in which each lane contained 2 μg of poly(A) RNA from the tissues indicated, was hybridized with a radiolabeled EST442803 DNA fragment. Radiolabeled β-actin cDNA was rehybridized to the same filter.

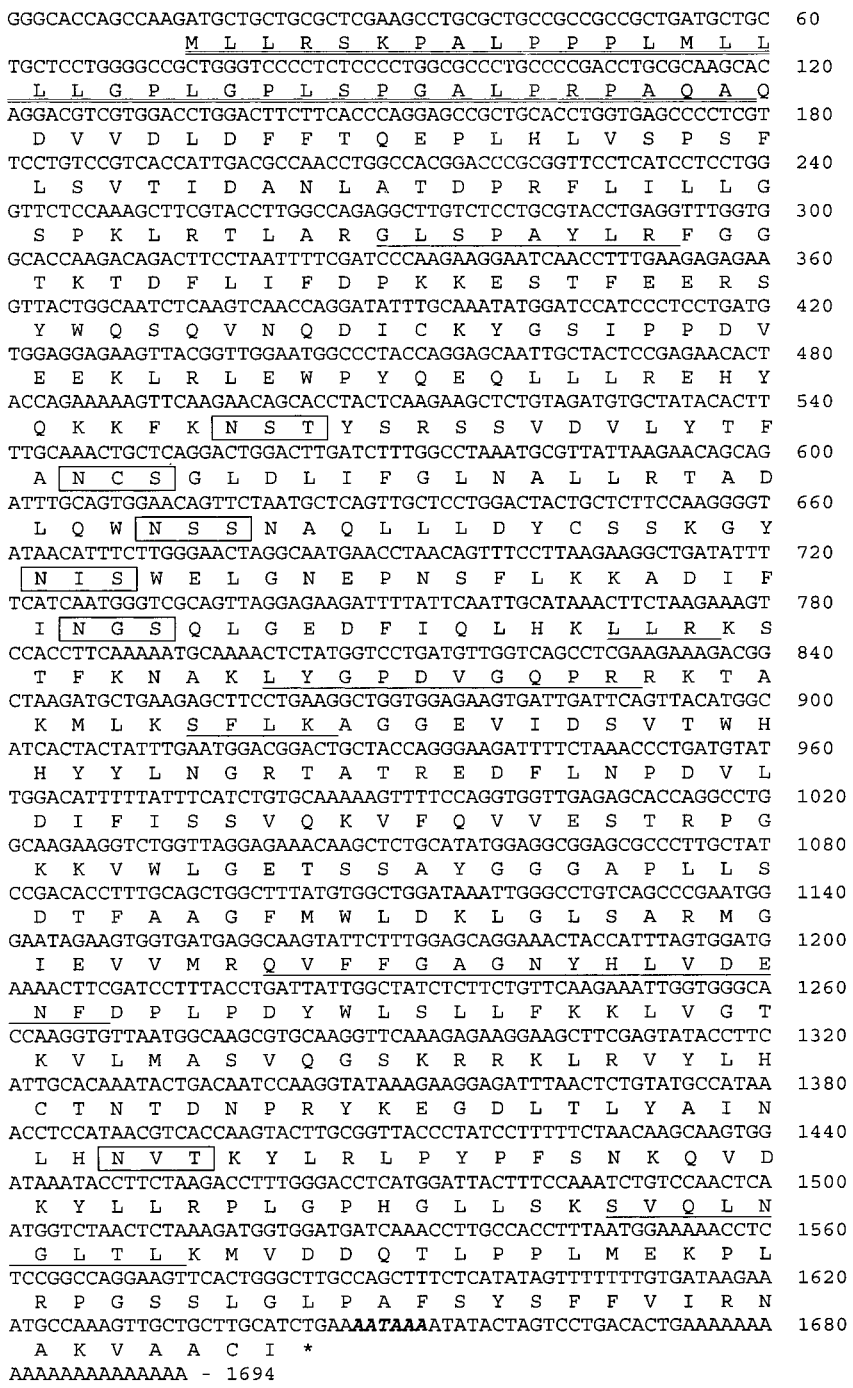

FIG. 2. cDNA sequence of *HSE1* and deduced amino acid sequence of heparanase gene. Potential sites for N-linked glycosylation are boxed while the predicted signal peptide is double underlined (═). Peptide sequences identified by Edman analysis are underlined and the polyadenylation signal is italicized and bold. The nucleotide sequence of heparanase has been submitted to the GenBank Database under Accession No. AF152376.

Biotechnology Information were searched using the tblastn sequence comparison program (22). One of these peptides, (L)-Y-G-P-D-V-G-Q-P-R, was found to match a sequence deposited in the GenBank Human EST database. EST clone 442803, from a human placental cDNA library, was identified as possessing an open reading frame that matched this and three other putative heparanase peptide sequences from Table 1. The complete sequence of EST 442803 insert was then determined and additional regions of translated sequence were found that matched the novel peptide sequences (Table 1 and see below). The complete EST

442803 sequence contained an uninterrupted open reading frame (ORF) of 975 bp, a stop codon, and a poly(A) tail, but lacked a translation initiation codon with a Kozak sequence.

Northern blot analysis. Expression of the putative heparanase gene was studied by Northern blot analysis of poly(A) RNA prepared from six different human tissues using a ^{32}P-labeled cDNA probe prepared from EST 442803. Specific mRNA was found to be highly expressed in the placenta and spleen but low levels were detectable in all the tissues tested (Fig. 1). Two mRNA transcripts were seen in all tissues, a predominant 1.8 kb transcript and a second 5 kb transcript. A third mRNA transcript of 1.4 kb was seen in testis tissue only. Results of a Southern blot analysis suggested that the different sized mRNA transcripts were due to a single gene (data not shown).

Cloning of the full length gene. To obtain the full-length complement of the putative heparanase gene, a pooled array placental cDNA library was screened using PCR primers specific for EST 442803. Two distinct clones were recovered containing inserts of 1.7 kb and 4.6 kb. The 1.7 kb insert was sequenced and determined to contain an open reading frame of 1629 base pairs including the entire previously characterized EST 442803 sequence. The ORF was predicted to encode a protein of 543 amino acids with a calculated molecular mass of 61192 daltons (Fig. 2). A consensus polyadenylation signal, AATAAA, was located downstream of the first in-frame stop codon and 27 bp upstream of the poly(A) tail. The deduced amino acid sequence included a predicted signal peptide leader sequence of 35 residues, for direction to the secretory apparatus, and lacked any recognizable membrane anchor sequence. The sequence contained six potential N-linked glycosylation sites (Fig. 2). Within the protein sequence were three peptides sequences reported to be from the human platelet heparitinase (16). A search of protein databases with the full length protein sequence failed to identify significant similarity to any protein yet reported, including the bacterial heparinases, CTAP-III, and other heparin binding proteins (27–29). Sequence analysis of the larger 4.6 kb cDNA clone identified the presence of a longer 3′ untranslated region containing approximately 3.6 kb of additional sequence not present in the 1.7 kb clone. The 5′ end of the transcript was truncated by 441 bp, perhaps due to an incomplete cDNA synthesis of the longer 5 kb mRNA species detected with the EST probe. The two distinct transcripts and the smaller 1.4 kb testis transcript may represent alternative splicing of nascent RNA and are currently being investigated.

Functional expression of heparanase from the full length cDNA. To determine whether the cloned cDNA encoded a protein with heparanase activity, COS7 cells were transfected with the full-length 1.7 kb cDNA cloned into a V5-epitope expression vector. Three days after transfection the conditioned media and COS7 cells were harvested. The COS7 cells were processed following published protocols for the isolation of heparanase from platelets (8). Recombinant V5-epitope-tagged protein was detected in both the soluble and insoluble fractions (Fig. 3, lanes 1, 2 and 3), in a manner consistent with published reports for native heparanase (17). Mild detergent exposure released recombinant protein from the insoluble fraction (Fig. 3, lane 4) in agreement with a previous finding (17). The apparent molecular weight of the recombinant fusion protein (Fig. 3), which includes 29 amino acid residues from the epitope-tag (3.2 kDa) is estimated to be 70–75 kDa by SDS-PAGE analysis. We estimated the expressed ORF alone (without the epitope-tag) to be 67–72 kDa. This would suggest that one or more of the predicted glycosylation sites is used since the calculated molecular weight of the deduced protein is only 61 kDa (including the leader sequence of 35 residues that may get cleaved during sorting). Several smaller molecular weight forms were detected in the insoluble material (Fig. 3, lane 3) and may represent proteolytic processing at the amino-terminal (since the epitope-tag is still present). Soluble fractions that were positive for recombinant protein were then each tested for enzymatic activity. Soluble material from the COS cell transfection (Fig. 3, lane 2) degraded radiolabeled HS substrate into distinct molecular weight fragments and this degradation could be inhibited by heparin (Fig. 4). COS cell conditioned media (Fig. 3, lane 1) and partially purified native heparanase (following method of (17)) displayed similar specific activity while the respective controls were all inactive (not shown). More extensive characterization of the HSE1 protein and its

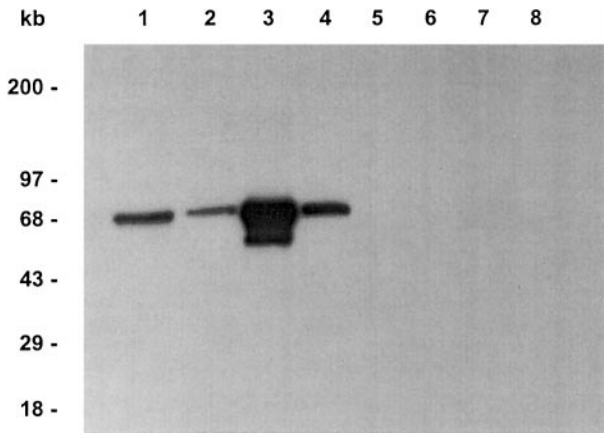

FIG. 3. Expression of carboxyl-terminal V5-tagged recombinant heparanase in COS7 cells as determined by Western blot analysis. Lanes 1–4, from *HSE1*-transfected COS7 cells; lanes 5–8, from control transfected COS7 cells; lanes 1 and 5, conditioned media; lanes 2 and 6, soluble fraction from freeze-thaw cell treatment; lanes 3 and 7, insoluble fraction from freeze-thaw cell treatment; lanes 4 and 8, released from insoluble fraction by detergent exposure.

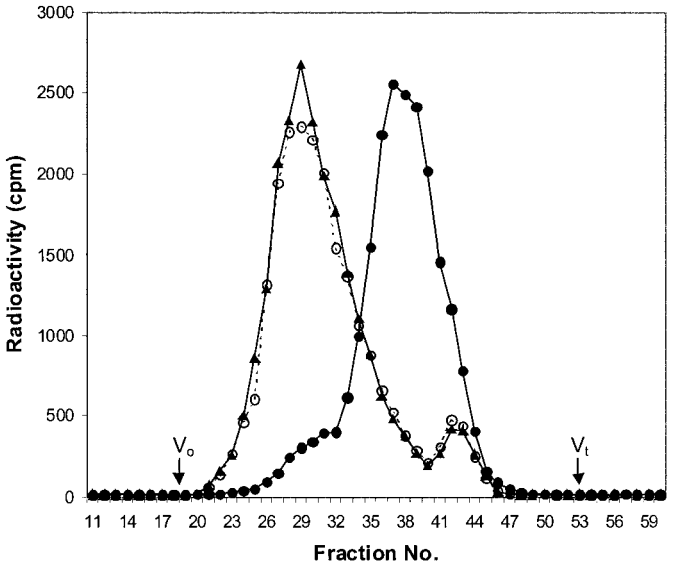

FIG. 4. Degradation of [^{35}S]-labeled HS by *HSE1*-transfected COS7 cells. The soluble fraction from transfected, freeze-thawed COS7 cells was incubated with ^{35}S-labeled HS in the absence (●) or presence (▲) of 10 μg/ml heparin. Control transfected cells were also tested (○). After incubation, the samples were analyzed by gel-filtration chromatography on a 0.7 × 25 cm Superose 6 column run at 0.2 ml/min in PBS (24). One-minute fractions were collected and analyzed for radioactivity.

enzymatic properties is underway. The significant amino acid identity of the cloned gene to peptide sequences from human heparanase, and the demonstration that this cloned gene exhibits heparanase activity, suggests that this gene encodes a novel human heparanase. Therefore, we have designated the gene *HSE1*, **h**eparanase **1.**

The cloning of the *HSE1* gene may help to clarify the actual number of unique heparanase enzymes and the role of heparanase activity from normal and tumor cells in the disassembly of the ECM and in cellular migration. The availability of purified, recombinant heparanase enzyme should aid in the development of screening assays for heparanase inhibitors that may be novel therapeutic agents for the treatment of specific cancers.

REFERENCES

1. Savion, N., Vlodavsky, I., and Fuks, Z. (1984) *J. Cell Physiol.* **118,** 169–178.
2. Vlodavsky, I., Fuks, Z., Bar-Ner, M., Ariav, Y., and Schirrmacher, V. (1983) *Cancer Res.* **43,** 2704–2711.
3. Nakajima, M., Irimura, T., Di Ferrante, D., Di Ferrante, N., and Nicolson, G. L. (1983) *Science* **220,** 611–613.
4. Vlodavsky, I., Michaeli, R. I., Bar-Ner, M., Fridman, R., Horowitz, A. T., Fuks, Z., and Biran, S. (1988) *Isr. J. Med. Sci.* **24,** 464–470.
5. Nakajima, M., Irimura, T., and Nicolson, G. L. (1988) *J. Cell Biochem.* **36,** 157–167.
6. Ihrcke, N. S., Parker, W., Reissner, K. J., and Platt, J. L. (1998) *J. Cell Physiol.* **175,** 255–267.
7. Bashkin, P., Doctrow, S., Klagsbrun, M., Svahn, C. M., Folkman, J., and Vlodavsky, I. (1989) *Biochemistry* **28,** 1737–1743.
8. Oosta, G. M., Favreau, L. V., Beeler, D. L., and Rosenberg, R. D. (1982) *J. Biol. Chem.* **257,** 11249–11255.
9. Klein, U., and Von Figura, K. (1976) *Biochem. Biophys. Res. Commun.* **73,** 569–576.
10. Bame, K. J., Hassall, A., Sanderson, C., Venkatesan, I., and Sun, C. (1998) *Biochem. J.* **336,** 191–200.
11. Nakajima, M., Irimura, T., Di Ferrante, N., and Nicolson, G. L. (1984) *J. Biol. Chem.* **259,** 2283–2290.
12. Pikas, D. S., Li, J. P., Vlodavsky, I., and Lindahl, U. (1998) *J. Biol. Chem.* **273,** 18770–18777.
13. Kosir, M. A., Quinn, C. C., Zukowski, K. L., Grignon, D. J., and Ledbetter, S. (1997) *J. Surg. Res.* **67,** 98–105.
14. Hoogewerf, A. J., Leone, J. W., Reardon, I. M., Howe, W. J., Asa, D., Heinrikson, R. L., and Ledbetter, S. R. (1995) *J. Biol. Chem.* **270,** 3268–3277.
15. Rechter, M., Lider, O., Cahalon, L., Baharav, E., Dekel, M., Seigel, D., Vlodavsky, I., Aingorn, H., Cohen, I. R., and Shoseyov, O. (1999) *Biochem. Biophys. Res. Commun.* **255,** 657–662.
16. Gonzalez-Stawinski, G. V., Parker, W., Holzknecht, Z. E., Huber, N. S., and Platt, J. L. (1999) *Biochim. Biophys. Acta* **1429,** 431–438.
17. Freeman, C., and Parish, C. R. (1998) *Biochem. J.* **330,** 1341–1350.
18. Goshen, R., Hochberg, A. A., Korner, G., Levy, E., Ishai-Michaeli, R., Elkin, M., de Groot, N., and Vlodavsky, I. (1996) *Mol. Hum. Reprod.* **2,** 679–684.
19. Gilat, D., Hershkoviz, R., Goldkorn, I., Cahalon, L., Korner, G., Vlodavsky, I., and Lider, O. (1995) *J. Exp. Med.* **181,** 1929–1934.
20. Graham, L. D., and Underwood, P. A. (1996) *Biochem. Mol. Biol. Int.* **39,** 563–571.
21. Stone, K. L., LoPresti, M. B., Williams, N. D., Crawford, J. M., DeAngelis, R., and Williams, K. R. (1989) *in* Techniques in Protein Chemistry (T. E. Hugli, Ed.), pp. 377–391, Academic Press, San Diego, CA.
22. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410.
23. Nakai, K., and Kanehisa, M. (1992) *Genomics* **14,** 897–911.
24. Freeman, C., and Parish, C. R. (1997) *Biochem. J.* **325,** 229–237.
25. Matzner, Y., Bar-Ner, M., Yahalom, J., Ishai-Michaeli, R., Fuks, Z., and Vlodavsky, I. (1985) *J. Clin. Invest.* **76,** 1306–1313.
26. Gloor, S., Odink, K., Guenther, J., Nick, H., and Monard, D. (1986) *Cell* **47,** 687–693.
27. Godavarti, R., and Sasisekharan, R. (1996) *Biochem. Biophys. Res. Commun.* **229,** 770–777.
28. Wenger, R. H., Wicki, A. N., Walz, A., Kieffer, N., and Clemetson, K. J. (1989) *Blood* **73,** 1498–1503.
29. Wu, D. Q., Kan, M. K., Sato, G. H., Okamoto, T., and Sato, J. D. (1991) *J. Biol. Chem.* **266,** 16778–16785.