A Mesangium-predominant Gene, Megsin, Is a New Serpin Upregulated in IgA Nephropathy

Toshio Miyata,* Masaomi Nangaku,* Daisuke Suzuki,* Reiko Inagi,* Ken-ichi Uragami,* Hideto Sakai,* Kousaku Okubo,‡ and Kiyoshi Kurokawa*

*Molecular and Cellular Nephrology, Institute of Medical Sciences and Department of Medicine, Tokai University School of Medicine, Isehara, Kanagawa 259-1193, Japan; and ‡Institute for Molecular and Cellular Biology, Osaka University, Suita, Osaka 565, Japan

Abstract

Mesangial cells play an important role in maintaining a structure and function of the glomerulus and in the pathogenesis of glomerular diseases. To identify a specific gene expressed in human mesangial cells, we used a rapid large-scale DNA sequencing and computerized data processing to compare the transcripts in cultured human mesangial cells with various different cells and organs. Using this novel approach, we discovered a new mesangium-predominant gene termed "megsin." We obtained a full-length cDNA clone of megsin, which coded for a novel 380–amino acid protein. Amino acid homology search revealed that megsin belonged to the serpin (serine protease inhibitor) superfamily. The amino acid sequences in the reactive loop site of megsin showed characteristic features of functional serpins. Northern blot and reverse-transcribed PCR analyses of various tissues and cells demonstrated that megsin was predominantly expressed in human mesangial cells. In situ hybridization studies showed the megsin expression in the mesangium of normal glomeruli, while it increased in the expanded mesangium of glomeruli from patients with IgA nephropathy with the degree of mesangial proliferation. Here we report a new human mesangium-predominant gene that may function as an inhibitory serpin in normal and abnormal biological processes of glomerulus. (*J. Clin. Invest.* 1998. 102:828–836.) Key words: 3′-directed cDNA library • cell-specific protein • serpin • mesangial proliferation • glomerulonephritis

Introduction

Sixty trillion (6×10^{13}) cells in the human body share essentially identical genomic DNA. Nonetheless, for normal physiological function distinct for each cell type, expression of genes is tightly regulated by the cell lineage. It is, therefore, critical to uncover which genes are specifically expressed by individual cell types for further understanding of cell biology and body function.

Mesangial cells play a central role in maintaining a structure and function of the glomerulus. In the pathophysiology of glomerular diseases, the proliferation of mesangial cells and the accumulation of extracellular mesangial matrix are primary events leading to the progression to glomerulosclerosis in patients with a variety of glomerular disorders, such as chronic glomerulonephritis and diabetic nephropathy, two major causes of end-stage renal failure.

To determine and characterize mesangium-specific genes is, therefore, essential to understand the pathogenesis of glomerular diseases. We used a novel strategy using a rapid large-scale DNA sequencing and computerized data processing for this purpose. Here we report cloning and characterization of a new human mesangium-predominant gene, megsin, which was a new member of the serpin superfamily and was upregulated in IgA nephropathy, a form of mesangial proliferative glomerulonephritis.

Methods

Construction of the 3′-directed mesangium cDNA library. Details of construction of the cultured human mesangial cell library have been described elsewhere (1). Briefly, human glomeruli were isolated by differential sievings from a normal human kidney obtained from surgical specimens obtained during nephrectomy in a 58-yr-old male with kidney tumor. Isolated glomeruli were resuspended in medium 199 (GIBCO BRL, Gaithersburg, MD) containing 25 mM Hepes, 10% Nu-serum (Collaborative Biomedical Products, Bedford, MA), and antibiotics (10 mg/ml of penicillin, streptomycin, and fungizone), and were incubated in a 5% CO_2 incubator. Mesangial cells at passage 3 were characterized by a series of criteria (1, 2): typical morphologic features, tolerance to trypsin, puromycin and D-valine, positive immunostaining for α-smooth muscle cell actin, and the absence of staining for von Willebrand factor, a marker for endothelial cells, and cytokeratine, a marker for parietal epithelial cells. Total RNA was isolated from human mesangial cells at passage 6 using the standard guanidine isothiocyanate method. PolyA$^+$ RNA was isolated by oligo dT-cellulose column (Pharmacia, Tokyo, Japan) according to the manufacturer's protocol.

A vector primer is dam-methylated at the MboI sites (GATC) and has a 3′ protrusion of homo dT stretch at one end. The double-stranded cDNA moiety was cleaved with MboI. The vector that was attached with the 3′-cDNA (from polyA to the nearest MboI site) was then circularized by ligation and was transformed into *Escherichia coli*. The resulting transformants were selected randomly and were lysed individually by brief boiling. The lysate was used as a template for PCR amplification. The resulting short double-stranded DNA was used for the cycle-sequencing reaction and was analyzed using an autosequencer.

Cloning of megsin. To identify the 5′ end of megsin cDNA, the 5′-Full RACE Core Set (Takara, Tokyo, Japan) was used. The reverse transcriptase (RT)[1] reaction was performed with 1.0 μg of hu-

1. *Abbreviations used in this paper:* IgA-N, IgA nephropathy; PAI, plasminogen activator inhibitor; RSL, reactive site loop; RT, reverse transcriptase; SCCA, squamous cell carcinoma antigen.

man mesangial cell polyA$^+$ RNA with a gene-specific oligonucleotide, 5′-pTCAGAGAGGTCATTC. After degradation of the template mRNA with RNase H, the first strand cDNA was circulated by T4 RNA ligase and an aliquot was used as a template for the first PCR amplification with gene-specific primers, P1 (5′-AGATTCTTGAGCTCAGAT) and P2 (5′-TCATTGATGGGTCCTCAA). The resulting PCR product was then used as a template for the second PCR amplification with nested gene specific primers, P3 (5′-AATGGTGGCATAAACATG) and P4 (5′-ACAGACAAATTGAACTTC). The PCR products were cloned into the pCRII vector (Clontech, Palo Alto, CA) and DNA sequences were determined by dideoxy chain termination reaction using a DNA autosequencer.

Expression of megsin fusion protein. Expression of megsin fusion protein was performed as follows. The entire coding sequence of megsin cDNA was obtained by RT-PCR. Single-strand cDNA was prepared from human mesangial cell polyA$^+$ RNA using oligo (dT) primer and RT. The resulting first-strand cDNA was amplified using specific primers from the 5′ region and the 3′ region of the coding sequence, P5 (5′-GTCGACTTATCAAGGGCAAGAAACTTTGCC) and P6 (5′-GAATTCCATGGCCTCCCTTGCTGCAGCAAA) containing EcoRI and SalI recognition sequences at their 5′ end each. Amplified megsin DNA was cloned originally into pCRII vector (Clontech) for nucleotide sequencing, and then ligated into the maltose-binding fusion protein vector pMALc2 (New England Biolabs, Beverly, MA) using EcoRI and SalI site. Fusion protein was expressed in *E. coli* (DH5α), affinity purified according to the manufacturer's instructions (New England Biolabs), and dissolved in 50 mM Tris-HCl, pH 9.0.

The fusion protein was then reacted with Endo Lys-C (Acromobacter Protease I, Wako Pure Chemicals, Osaka, Japan) (enzyme/substrate ratio = 1:50, W/W) at 37°C for 8 h. The digested peptides were separated by high performance liquid chromatography (model LC-10A; Shimadzu, Kyoto, Japan) on a C18 reverse-phase column (4.6 × 250 nm; Waters, Tokyo, Japan), detected by their absorbance at 214 nm, collected, and subjected to peptide sequencing on an automatic protein sequencer (Model 491 gas-phase sequencer; Perkin-Elmer, Chiba, Japan) or to SDS-PAGE using a 4–20% gradient acrylamide gel and Coomassie brilliant blue staining.

Analysis of the amino acid sequence. Motif search was performed using a Motif Finder program written by A. Ogiwara (Institute for Chemical Research, Kyoto University, Kyoto, Japan). This program searches PROSITE (3, 4) and MotifDic (5). To find a signal sequence, we applied PSORT (K. Nakai, National Institute for Basic Biology, Japan), a program to predict protein localization sites. Sequence homology search was performed using the FASTA program (6). Sequence alignment and pustell matrix comparison were performed according to the protein-scoring matrix pam250 with MacVector (International Biotechnologies, Inc., New Haven, CT). The Kyte-Doolittle scale (7) was used to investigate a hydrophilicity profile with MacVector. Secondary structure prediction was performed by PHD (Profile fed neural network system from Heidelberg) sec, according to the Rost and Sander secondary structure prediction method (8).

Northern blot analysis. RNA of cultured human mesangial cells and fibroblasts was separated on a 1% agarose gel containing 2.2 M formamide and transferred to a nitrocellulose filter. The filter was hybridized in a Rapid Hyb solution (Amersham Corp., Arlington Heights, IL). Blot was washed to a final stringency of 0.1×SSPE/0.1% SDS at 60°C.

A human multiple tissue Northern blot and a human cancer cell line Northern blot were purchased from Clontech. A human multiple tissue northern blot contains 2 μg of polyA$^+$ RNA from heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas. A human cancer cell line Northern blot has 2 μg of polyA$^+$ RNA from promyelocytic leukemia (HL-60), HeLa cell (S3), chronic myelogenous leukemia (K-562), lymphoblastic leukemia (MOLT-4), Burkitts' lymphoma (Raji), colorectal adenocarcinoma (SW480), lung carcinoma (A549), and melanoma (G361). Hybridization and washing was performed as described above.

RT-PCR. Total RNA was isolated from monocytes, polymorphonuclear leukocytes, cultured human mesangial cells, and other human primary cultured cells (vascular smooth muscle cells, fibroblasts, endothelial cells, renal epithelial cells, and keratinocytes) purchased from Takara (Tokyo, Japan) by guanidinium thiocynate-phenol-chloroform extraction. The RNA was reverse transcribed using a T-Primed-First-Strand kit (Pharmacia) according to the manufacturer's instructions.

PCR amplification was performed in 25 μl reaction (with 1 ml of the reverse-transcribed products) containing 0.625 U of ExTaq (Takara), 200 mM each of the four deoxynucleoside triphosphates, 20 pmol of sense primer (5′ end), and 20 pmol of antisense primer (3′ end). The oligonucleotide primers for megsin sequences were: sense, 5′-ATGATCTCAGCATTGTGAATG-3′ and antisense, 5′-ACTGAGGGAGTTGCTTTTCTAC-3′. The expected size of the amplified fragment was 773 bp. The β-actin served as an internal RNA control to allow comparison of RNA levels among different specimens. Specimens were amplified in a DNA Thermal Cycler (Perkin Elmer Cetus). Each cycle consisted of incubation for denaturation at 94°C for 1 min, annealing at 60°C for 2 min, and extension at 72°C for 2 min. The PCR products were resolved by electrophoresis in a 1% agarose gel and visualized by ethidium bromide staining.

In situ hybridization. In situ hybridization was performed as described previously (9). Nucleotide sequences from 391 to 428 of human megsin cDNA (5′-AGCGTGGACTGAGGGAGTTGCTTTTCTACAATATTACT-3′) were used as a probe. To evaluate the specificity of the signals, RNase pretreatment or a competitive study with 100-fold excess amount of homologous or unrelated, unlabeled oligonucleotides was performed. To quantitate the megsin mRNA expression, all nuclei and nuclei with surrounding positive cytoplasm in at least 10 randomly selected glomeruli (cross-section through their vascular poles) were blindly enumerated, and results were expressed as percentage of positive cells of total nuclei. Mann-Whitney U test was used for the statistical comparisons.

Results

Identification of a new human mesangium-specific gene. To identify specific gene(s) expressed in mesangial cells, we used a rapid large-scale DNA sequencing and computerized data processing to compare the transcripts in various different cells and organs (1, 10, 11). A large scale sequencing of a 3′-directed cDNA library of cultured human mesangial cells was performed to determine partial sequences of 1836 randomly picked clones. The sequence similarities of the clones were compared with each other and with the DNA databank GenBank using the FASTA program (6). We selected six clones that had not been recorded in any other databases and were thought to be predominantly expressed in mesangial cells. Dot blot analysis of mRNAs from various organs and cells supported mesangium-predominant expression of these clones (data not shown). One clone was obtained, which was most abundant among those detected only in our human mesangial cell cDNA library. This clone consisted of 0.3% of all the mRNA population.

The new gene, megsin, is a new member of the serpin superfamily. Using PCR techniques, a full length cDNA of this gene was obtained (Fig. 1). The nucleotide sequencing analysis revealed an open reading frame of 1140 nucleotides coding for a predicted protein of 380 amino acids with four potential *N*-glycosylation sites (the deduced molecular weight, 43 kD). The sequence around the initiation codon (GCAATGG) was consistent with Kozak's consensus sequence (A/GCCATGG). Stop codons were in all three reading frames at the upstream of the initiation codon.

```
   1  GTACAGGGAAGCTCTCCTTCATCACCTTCCTAAGTGCATGGGGGAAAATACCTAGGGCTCAACAGTCTTGAGAAGTGTGGAAACATTTTC   80
  61  TTTGTGAGTGAGAACAGATCACCTAGAGAAAGGAAACCAGATTCCCATCACTGCTTCTGGGTATCAGATGCTAGCGCTGCACTCCATTTT  180
             M   A   S   L   A   A   A   N   A   E   F   C   F   N   L   F   R   E   M
 181  GCA ATG GCC TCC CTT GCT GCA GCA AAT GCA GAG TTT TGC TTC AAC CTG TTC AGA GAG ATG            240
      D   D   N   Q   G   N   G   N   V   F   F   S   S   L   S   L   F   A   A   L
 241  GAT GAC AAT CAA GGA AAT GGA AAT GTG TTC TTT TCC TCT CTG AGC CTC TTC GCT GCC CTG            300
      A   L   V   R   L   G   A   Q   D   D   S   L   S   Q   I   D   K   L   L   H
 301  GCC CTG GTC CGC TTG GGC GCT CAA GAT GAC TCC CTC TCT CAG ATT GAT AAG TTG CTT CAT            360
      V   N   T   A   S   G   Y   G   N*  S   S   N   S   Q   S   G   L   Q   S   Q
 361  GTT AAC ACT GCC TCA GGA TAT GGA AAC TCT TCT AAT AGT CAG TCA GGG CTC CAG TCT CAA            420
      L   K   R   V   F   S   D   I   N*  A   S   H   K   D   Y   D   L   S   I   V
 421  CTG AAA AGA GTT TTT TCT GAT ATA AAT GCA TCC CAC AAG GAT TAT GAT CTC AGC ATT GTG            480
      N   G   L   F   A   E   K   V   Y   G   F   H   K   D   Y   I   E   C   A   E
 481  AAT GGG CTT TTT GCT GAA AAA GTG TAT GGC TTT CAT AAG GAC TAC ATT GAG TGT GCC GAA            540
      K   L   Y   D   A   K   V   E   R   V   D   F   T   N   H   L   E   D   T   R
 541  AAA TTA TAC GAT GCC AAA GTG GAG CGA GTT GAC TTT ACG AAT CAT TTA GAA GAC ACT AGA            600
      R   N   I   N   K   W   V   E   N*  E   T   H   G   K   I   K   N   V   I   G
 601  CGT AAT ATT AAT AAG TGG GTT GAA AAT GAA ACA CAT GGC AAA ATC AAG AAC GTG ATT GGT            660
      E   G   G   I   S   S   S   A   V   M   V   L   V   N   A   V   Y   F   K   G
 661  GAA GGT GGC ATA AGC TCA TCT GCT GTA ATG GTG CTG GTG AAT GCT GTG TAC TTC AAA GGC            720
      K   W   Q   S   A   F   T   K   S   E   T   I   N   C   H   F   K   S   P   K
 721  AAG TGG CAA TCA GCC TTC ACC AAG AGC GAA ACC ATA AAT TGC CAT TTC AAA TCT CCC AAG            780
      C   S   G   K   A   V   A   M   M   H   Q   E   R   K   F   N*  L   S   V   I
 781  TGC TCT GGG AAG GCA GTC GCC ATG ATG CAT CAG GAA CGG AAG TTC AAT TTG TCT GTT ATT            840
      E   D   P   S   M   K   I   L   E   L   R   Y   N   G   G   I   N   M   Y   V
 841  GAG GAC CCA TCA ATG AAG ATT CTT GAG CTC AGA TAC AAT GGT GGC ATA AAC ATG TAC GTT            900
      L   L   P   E   N   D   L   S   E   I   E   N   K   L   T   F   Q   N   L   M
 901  CTG CTG CCT GAG AAT GAC CTC TCT GAA ATT GAA AAC AAA CTG ACC TTT CAG AAT CTA ATG            960
      E   W   T   N   P   R   R   M   T   S   K   Y   V   E   V   F   P   Q   F
 961  GAA TGG ACC AAT CCA AGG CGA ATG ACC TCT AAG TAT GTT GAG GTA TTT TTT CCT CAG TTC           1020
      K   I   E   K   N   Y   E   M   K   Q   Y   L   R   A   L   G   L   K   D   I
1021  AAG ATA GAG AAG AAT TAT GAA ATG AAA CAA TAT TTG AGA GCC CTA GGG CTG AAA GAT ATC           1080
      R   M   M   H   K   S   Y   I   E   V   T  |E   E   G   T   E   A   T   A   A|
1081  TTT GAT GAA TCC AAA GCA GAT CTC TCT GGG ATT GCT TCG GGG GGT CGT CTG TAT ATA TCA           1140
     |T   G   S   N   I   V   E   K| Q   L   P   Q   S| T   L   F   R   A   D   H
1141  AGG ATG ATG CAC AAA TCT TAC ATA GAG GTC ACT GAG GAG GGC ACC GAG GCT ACT GCT GCC           1200
      P   F   L   F   V   I   R   K   D   D   I   I   L   F   S   G   K   V   S   C
1201  ACA GGA AGT AAT ATT GTA GAA AAG CAA CTC CCT CAG TCC ACG CTG TTT AGA GCT GAC CAC           1260
      P   Stop
1261  CCA TTC CTA TTT GTT ATC AGG AAG GAT GAC ATC ATC TTA TTC AGT GGC AAA GTT TCT TGC           1320
1325  CCT TGA AAATCCAATTGGTTTCTGTTATAGCAGTCCCCACAACATCAAAGRACCACCACAAGTCAATAGATYTGRGTTTAATTGGA  1376
1385  AAAATGTGGTGTTTCCTTTGAGTTTATTTCTTCCTAACATTGGTCAGCAGATGACACTGGTGACTTGACCCTTCCTAGACACCTGGTTGA  1496
1505  TTGTCCTGATCCCTGCTCTTAGCATTCTACCACCATGTGTCTCACCCATTTCTAATTTCATTGTCTTTCTTCCCACGCTCATTTCTATCA  1586
1595  TTCTCCCCCATGACCCGTCTGGAAATTATGGAGRGTGCTCAACTGGTAAGGAGAACGTAGAAGTAGCCCTAGGGATCCTTTTTGAAACTC  1676
1685  TACAGTTATCGCAGATATTCTAGCTTCATTGTAAGCAATCTAGGAAATAAGCCCTGCTGCTTTCTAGAAATAAGTGTGAAGGATAAATTT  1766
1775  TCTTTGTTGACCTATGAAGATTTTAGAGTTTACCTTCATATGTTTGATTTTAAATCAGTGTATAATCTAGATGGTAAAAAATGTGAAATT  1856
1865  GGGATTAGGGACCTACCAAAATATTTCATTAATGCTTTCAATTGACAAATTTTGGCCTTTCTTTGATAAGACAATATGTACATGTTTTTT  1946
1955  CAAATATAAAGATCTTTTAACTGTTGGCAGTTGTTTATCTACAGAATCATATTTCATATGCTGTGTAGTTTATAAGTTTTTCCTCTATTT  2036
2045  ATCAGAATAAAGAAATACAACATACCTGTAAA   2068
```

Figure 1. Nucleotide sequence of the megsin cDNA and its deduced amino acid sequence. The sequencing of cDNA revealed an open reading frame of 1140 nucleotides coding for a 380–amino acid polypeptide displayed over the nucleotide sequence in one-letter code. A double overline indicates a serpin signature. Pn-numbered portion enclosed by a box indicates the RSL and the arrow indicates the putative reactive site. Two putative hydrophobic regions are indicated by single underlines. Four potential *N*-glycosylation sites are indicated by asterisks. Confirmed peptide sequences of the fusion protein produced by *E. coli* expression system are indicated by single overlines. Initial nucleotide sequence obtained by a rapid large-scale DNA sequencing of a 3′-directed cDNA library of human cultured mesangial cells is indicated by a double underline. The accession number for the nucleotide sequence of megsin is AF027866 (DNA Data Bank of Japan, European Molecular Biology Laboratory, and GenBank).

The size of the fusion protein produced by *E. coli* expression system was ∼ 85 kD (Fig. 2), which was equivalent to the calculated molecular weight of the deduced amino acid sequence of the fusion protein with the maltose-binding protein (42 kD). The amino acid sequencing analysis of two Endo Lys-C-digested peptides of the fusion protein confirmed that this open reading frame was correct (single overlines in Fig. 1).

Amino acid homology search by the FASTA program throughout the SwissProt database revealed that this gene was highly homologous to members of the serpin superfamily (12–16), which is a group of structurally related proteins and generally serves as extracellular, irreversible serine protease inhibitors. Squamous cell carcinoma antigen 1 (SCCA1) showed the highest homology to megsin (41.2% identity, see also Fig. 3*A*),
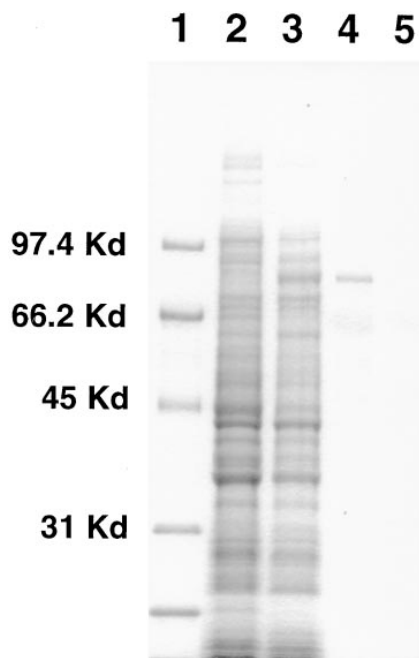
Figure 2. SDS-PAGE analysis of the fusion protein (maltose-binding protein megsin) expressed in *E. coli*. After SDS-PAGE on a 4–20% gradient acrylamide gel, the gel was stained with Coomassie brilliant blue. Lane *1*, molecular weight markers; lane *2*, *E. coli* lysate with pMalc2 (5 µg); lane *3*, *E. coli* lysate with pMalc2 containing megsin cDNA (5 µg); lane *4*, purified fusion protein (maltose-binding protein megsin) (1 µg); lane *1*, Endo Lys-C–digested peptides of the fusion protein (5 µg).

followed by other members of the serpin superfamily, including SCCA2 (40.6%), leukocyte elastase inhibitor (37.5%), and plasminogen activator inhibitor-2 (PAI-2) (35.2%). Therefore, we termed this gene megsin after mesangial cell-predominant gene with a homology to serpin.

Motif search of the amino acid sequence of megsin confirmed the serpin signature at the COOH terminus. Comparison with the other members of the serpin superfamily suggested that amino acids 331–352 corresponded to the reactive site loop (RSL) (P17 to P5′) (17–20) (a box in Fig. 1 and Fig. 3 *B*).

Northern blot analysis and RT-PCR confirmed that megsin was predominantly expressed in human mesangial cells. By Northern blot analysis with a megsin cDNA probe, a single transcript of 3.0 kb was detected in cultured mesangial cells, but not in any other organs or cancer cell lines tested (Fig. 4 *A*). Although megsin transcript was detected in primary cultured human fibroblasts, the amount of the transcript was apparently much less than in mesangial cells (Fig. 4 *B*). The megsin transcripts were detected only in cultured human mesangial cells by RT-PCR at 20 cycles (Fig. 5 *A*). When the samples were amplified up to 40 cycles, faint bands corresponding to megsin could be observed in kidneys from two different subjects, some other primary cultured cells, and circulating cells (Fig. 5 *B*). Without reverse transcription, no bands could be detected (data not shown). These results demonstrated that megsin was predominantly expressed in human mesangial cells.

In situ hybridization studies demonstrated that expression of megsin was upregulated in the proliferating mesangium. The meg-

sin mRNA expression was evaluated on human renal tissues from 18 patients with IgA nephropathy (IgA-N) and three normal kidneys by in situ hybridization. Patients with IgA-N were divided into two subgroups: those with mesangial cell proliferation with minimal glomerular sclerosis (proliferative stage, *n* = 9) and those with glomerular sclerosis in > 30% of glomeruli (sclerotic stage, *n* = 9). Megsin mRNAs were demonstrated exclusively in the glomeruli of both normal subjects (Fig. 6 *A*) and IgA-N (Fig. 6 *B*). Within the glomeruli, megsin transcripts were detected in location consistent with mesangial cells (Fig. 6 *C*). Pretreatment of tissues with RNase before hybridization eliminated most of the signal detected with the megsin probe, and the megsin signals disappeared with a large amount of unlabeled homologous oligonucleotide competitors but not with nonhomologous oligonucleotides (data not shown). The number of positive cells for megsin in IgA-N of the proliferative stage was significantly greater than those in normal kidneys (Table I).
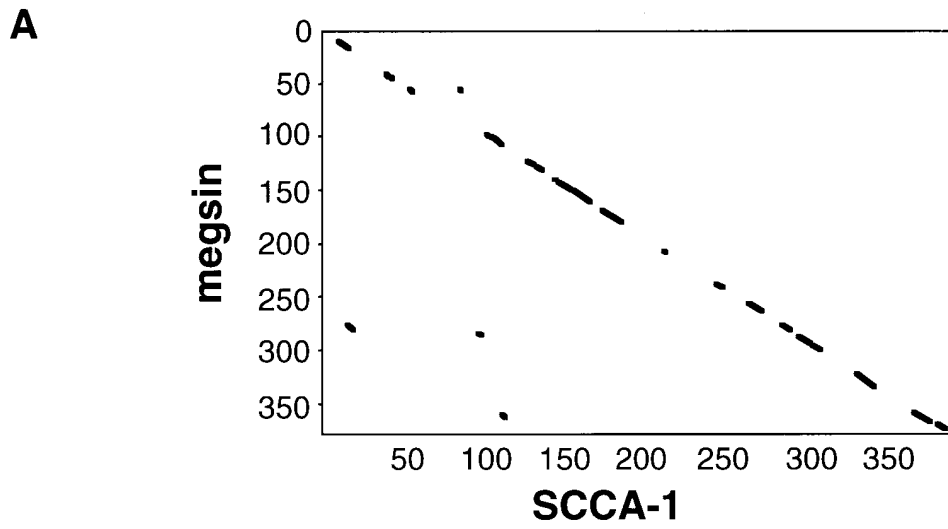
Discussion

Using a novel approach entailing a rapid large-scale DNA sequencing and computerized data processing, we discovered a new human mesangium-predominant gene "megsin." Megsin was composed of 380 amino acids and was a new member of the serpin superfamily. Although some serpins have functions distinct from protease inhibition, three evidences suggest that megsin is an inhibitory serpin. First, the Peven residues of the megsin RSL were uncharged, small, and nonpolar, which is a characteristic feature of the inhibitory serpins. Second, inhibitory serpins have the sequence of Ala-Ala/Thr-Ala-Ala at the NH$_2$-terminal region of the RSL (P12 to P9), called the hinge region. P12 to P9 of the RSL of megsin were ATAA. The P17-P8 sequence of megsin (EEGTEATAAT) was indeed consistent with the consensus sequence among inhibitory serpins (EEGTEAAAAT) (Fig. 3). Third, it is known that the stretch of β-sheet immediately precedes the NH$_2$-terminal side of the RSL, which is a prerequisite for proteinase inhibition, and limits the sizes and charges of amino acids in the hinge region to achieve appropriate conformational change. Megsin preserved this β-sheet stretch.

Residues flanking the putative scissile bonds within the RSL (P1 and P1′), which are thought to be important in determining the substrate specificity (21–23), were Lys-Gln. No other inhibitory serpins have related sequences at this site. Serpins such as the Kunitz type bovine basic protease inhibitor

Table I. Quantitative Analysis of Megsin mRNA from Normal and IgA-N Kidneys by in situ Hybridization

	Percentage of positive cells
IgA-N/Proliferative stage (*n* = 9)	25.21±3.91*
IgA-N/Sclerotic stage (*n* = 9)	18.39±4.91
Normal kidney (*n* = 3)	14.47±1.96

To quantitate the megsin mRNA expression by in situ hybridization, all nuclei and nuclei-surrounding positive cytoplasm in at least 10 randomly selected glomeruli were blindly enumerated, and the results were expressed as percentage of positive cells of total nuclei. Data are mean±SD. *$P < 0.01$ versus normal kidney or IgA-N/Sclerotic by Mann-Whitney U test.

A

B

	P15				P10					P5				P1 P1'				P5'
megsin	E	E	G	T	E	**A**	T	**A**	**A**	T	G	S	N	**I**	V	E	K	Q	**L**	**P**	Q	S

(Inhibitory serpins)

SCCA1	E	E	G	A	E	**A**	**A**	**A**	**A**	T	**A**	**V**	**V**	G	**F**	G	S	S	**P**	T	S	T
ILEU	E	E	G	T	E	**A**	**A**	**A**	**A**	T	**A**	G	**I**	**A**	T	**F**	C	**M**	**L**	**M**	**P**	–
PAI-2	E	E	G	T	E	**A**	**A**	**A**	G	T	G	G	**V**	**M**	T	G	R	T	G	H	G	G

Non-inhibitory serpins

OVA	E	**A**	G	R	E	**V**	**V**	G	S	**A**	E	**A**	G	**V**	D	**A**	**A**	S	**V**	S	–	–
Gene Y	E	E	G	T	E	**A**	T	G	S	T	G	**A**	**I**	G	N	**I**	K	H	S	**L**	E	**L**

Figure 3. Comparison of the amino acid sequences of megsin with other members of the serpin superfamily. (*A*) A comparison of the protein sequences between megsin and SCCA1. Matching regions are shown as diagonal lines. These sequences were aligned according to the protein-scoring matrix pam 250 with window size 8, minimum percent score 60, and hash value 2. (*B*) Comparison of serpin RSLs. Inhibitory serpin RSLs from P17–P5′ (Schechter and Berger numbering [37]) were aligned. Nonpolar residues are shown in bold. P17–P8, indicated by a box, corresponds to the consensus sequence for inhibitory serpins (*EEGTEAAAAT*) (38). *ILEU*, leukocyte elastase inhibitor; *OVA*, ovalbumin.

are known to have Lys at P1 and inhibit trypsin strongly. The target serine protease of megsin might be, therefore, a lysine-cleaving protease.

Hydrophilicity profile showed two hydrophobic regions in α helix A (amino acids 1–16) and B (amino acids 22–44) of megsin (single underlines in Fig. 1), whereas we could find no apparent NH$_2$-terminal signal peptide sequence. Several serpins appear to exist as dualistic molecules that can be either secreted or cytoplasmic, due to these noncleavable internal signal sequence in a helix A and B (24–28). It is possible that these α helix regions play important roles for translocation of megsin as dualistic molecules.

Northern blot analysis revealed the predominant expression of megsin in cultured human mesangial cells. Megsin transcript was not detected in poly A$^+$ RNA from other cancer cells or organs tested by Northern blotting. Localization of megsin-producing cells in our in situ hybridization studies of human renal tissues from normal subjects and patients with IgA-N were consistent with mesangial cells. Furthermore, we cloned a megsin homologue from cultured rat mesangial cells, which were stained positive for Thy-1 antigen, a marker of rat mesangial cells (29) (our manuscript in preparation). These findings strongly suggest the predominant expression of megsin in mesangial cells.

Northern blot analysis did not detect megsin transcripts in human kidney samples, although RT-PCR analysis of human kidneys and some other primary cultured or circulating cells showed corresponding bands after the samples were amplified up to 40 cycles. One possible explanation for this discrepancy is due to the small number of mesangial cells in the kidney, which is also composed of endothelial cells, epithelial cells, and other various cells. Alternatively, cultured mesangial cells are known to acquire new phenotypes, and become more prone to proliferate or become activated (30, 31). Our in situ hybridization studies demonstrated that the expression of megsin was upregulated in IgA-N of the proliferative stage. Therefore, it is more likely that expression of megsin is upregulated only when mesangial cells are activated.
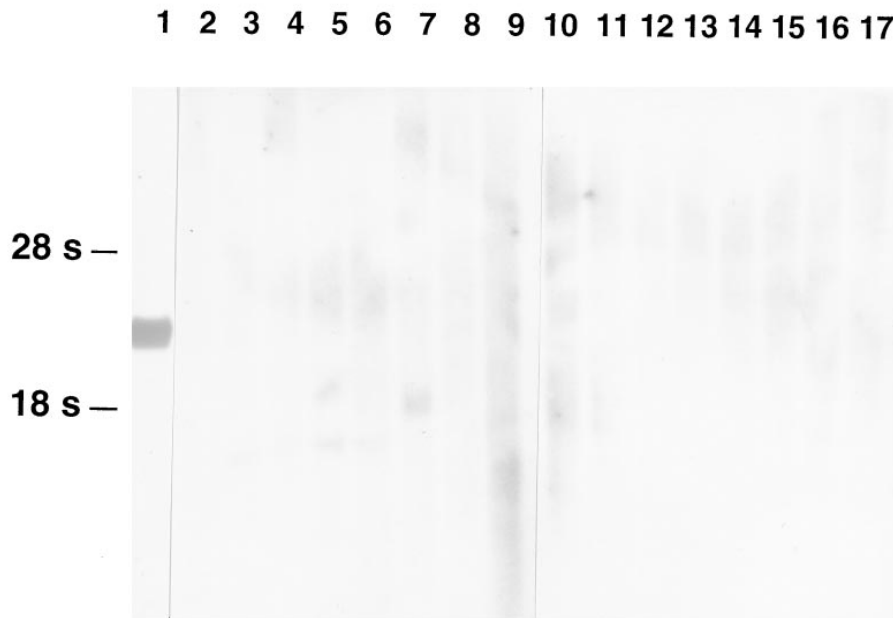
A

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

28 s —

18 s —

B

1 2 3 4

28 s –

18 s –

Figure 4. Detection of megsin transcripts. (*A*) Expression of megsin in cultured mesangial cells, various human organs, and cancer cell lines were examined by Northern blot. Each lane contained 2 μg of poly A+ RNA. The expression of megsin was observed in mesangial cells. Lane *1*, cultured mesangial cells; lane *2*, pancreas; lane *3*, kidney; lane *4*, skeletal muscle; lane *5*, liver; lane *6*, lung; lane *7*, placenta; lane *8*, brain; lane *9*, heart; lane *10*, melanoma cells (G361); lane *11*, lung carcinoma cells (A549); lane *12*, colorectal adenocarcinoma cells (SW480); lane *13*, Burkitts' lymphoma cells (Raji); lane *14*, lymphoblastic leukemia cells (MOLT-4); lane *15*, chronic myelogenous leukemia cells (K-562); lane *16*, HeLa cells (S3); lane *17*, promyelocytic leukemia cells (HL-60). (*B*) Northern blot analysis of primary cultured mesangial cells and fibroblasts. The expression levels of megsin in mesangial cells were much higher than in fibroblasts. Lane *1*, 20 μg of total RNA from mesangial cells; lane *2*, 20 μg of total RNA from fibroblasts; lane *3*, 4 μg of total RNA from mesangial cells; lane *4*, 4 μg of total RNA from fibroblasts. Horizontal bars on the right correspond to the sizes of 28S and 18S ribosomal RNA (5.1 kb and 1.9 kb, respectively).
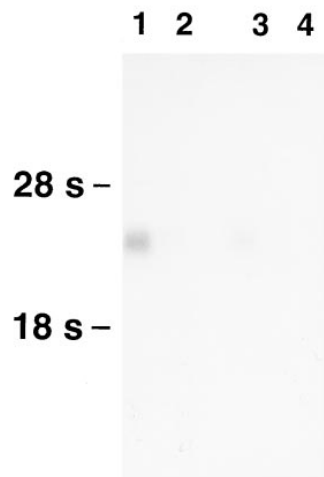
The mechanism of mesangium-predominant expression of megsin is another important subject to be pursued. Further studies of the promoter and transcriptional factor of megsin will give an important insight to a mechanism of cell type–dependent gene expression. This information can also be useful in targeting specific cells when molecular genetics and gene transfer techniques are applied.

Furthermore, mesangium-predominant genes can also serve as biological markers for mesangial cells. Although Thy-1 antigen has been used to identify rat mesangial cells, rat Thy-1 antigen is also expressed by other cell lineages such as thymocytes and, more importantly, human mesangial cells lack this molecule (32). Human mesangium-specific markers provide great tools to investigate glomerular diseases, which renders identification of megsin even more importance.

Sequence homology and similar biological activity of megsin with PAI-2 and other inhibitory serpins gives us a clue for its putative function. Serpins play a central role in the regulation of a wide variety of physiologic and pathological processes including coagulation, fibrinolysis, matrix metabolism,
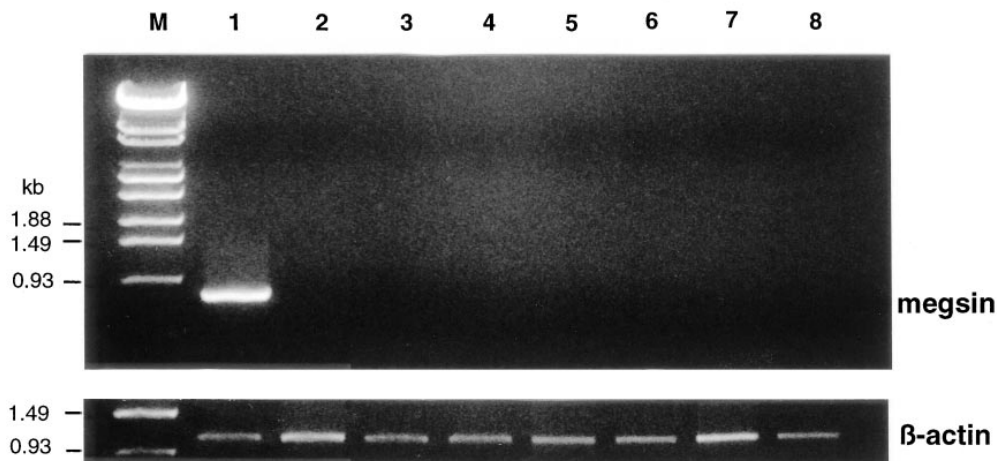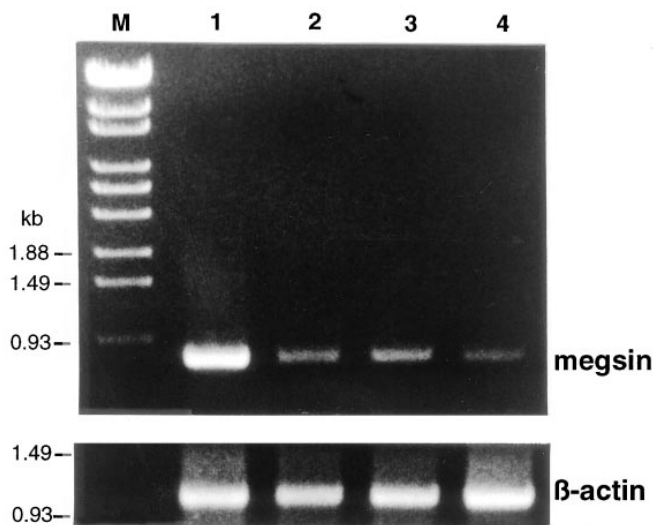
A

B

Figure 5. PCR detection of megsin transcripts. (*A*) PCR analysis at 20 cycles confirmed predominant expression of megsin in mesangial cells (*top*). The bottom shows PCR products of β-actin as an internal control. Lane *1*, mesangial cells; lane *2*, smooth muscle cells; lane *3*, fibroblasts; lane *4*, endothelial cells; lane *5*, renal epithelial cells; lane *6*, keratinocytes; lane *7*, monocytes; lane *8*, polymorphonuclear leukocytes. (*B*) PCR demonstrated bands corresponding to megsin in kidneys when the samples were amplified up to 40 cycles. Lane *1*, cultured mesangial cells; lane *2*, kidney (subject A); lane *3*, kidney (subject B); lane *4*, fibroblasts.

development, blood pressure, malignancy, and inflammation (14, 33). Glomerulosclerosis results from accumulation of extracellular matrix proteins as a consequence of their increased synthesis and relative inhibition of degradation by proteases. The enzymatic activity of a protease is regulated by a balance among proteinases and their inhibitors. Recent studies demonstrated that transgenic mice that overexpress the PAI-1 showed more marked lung fibrosis in response to bleomycin than wild-type mice, whereas PAI-1 knockout mice are relatively protected (34), emphasizing the role of serpin in extracellular matrix accumulation. It is also becoming clear that serpins inhibit critical cellular processes such as adhesion, migration, free radical production, and even apoptosis (35, 36).

Detailed functional analysis of megsin is now under investigation.

In conclusion, megsin is a new serpin protein expressed predominantly in mesangial cells. The expression levels of megsin paralleled mesangial proliferation in vivo. Megsin could be the first biological marker for injured mesangial cells, which renders this molecule a high potential from the diagnostic perspective of glomerular diseases. Furthermore, marked upregulation of megsin in the injured mesangium suggests that megsin may play some roles in the pathogenesis of glomerulosclerosis. Studies of megsin might thus open up the way for future therapeutic approaches to treat various glomerular disorders.
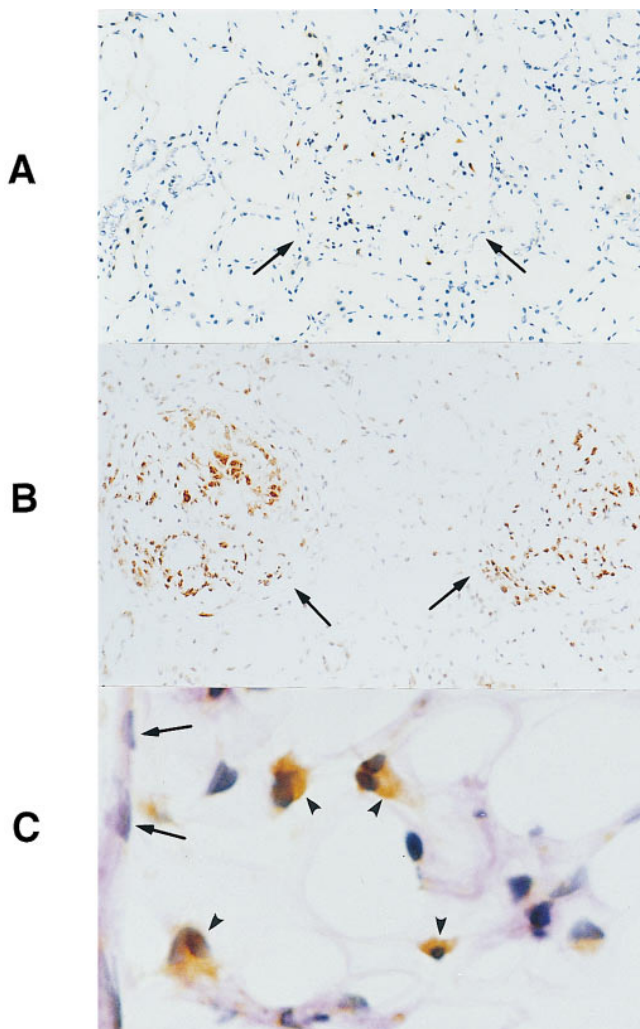
Figure 6. Expression of megsin in the mesangial cells within glomeruli from normal subjects and patients with IgA nephropathy. (*A*) The signals for megsin are observed in a glomerulus (*arrows*) of normal subjects. There are no signals for megsin in tubulointerstitial area. Original magnification, ×40. (*B*) The increased signals for megsin are observed in two glomeruli (*arrows*) of patients with IgA-N. Original magnification ×40. (*C*) A high magnification view shows megsin-producing cells in mesangial area (*arrowheads*), whereas cells of Bowman's capsule (*arrows*) were negative for megsin. Original magnification, ×200.

Acknowledgments

References

1. Yasuda, Y., T. Miyata, M. Nangaku, Y. Iida, K. Maeda, K. Kurokawa, and K. Okubo. 1998. Functional quantitative analysis of the genome in cultured human mesangial cells. *Kidney Int.* 53:154–158.

2. Striker, G.E., and L.J. Striker. 1985. Glomerular cell culture. *Lab. Invest.* 53:122–132.

3. Bairoch, A. 1992. PROSITE: a dictionary of sites and patterns in proteins. *Nucleic Acids Res.* 20:2013–2018.

4. Bairoch, A., P. Bucher, and K. Hofmann. 1996. The PROSITE database: its status in 1995. *Nucleic Acids Res.* 24:189–196.

5. Ogiwara, A., I. Uchiyama, Y. Seto, and M. Kanehisa. 1992. Construction of a dictionary of sequence motifs that characterize groups of related proteins. *Protein Eng.* 5:479–488.

6. Pearson, W.R., and D.J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA.* 85:2444–2448.

7. Kyte, J., and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105–132.

8. Rost, B., and C. Sander. 1993. Prediction of protein structure at better than 70% accuracy. *J. Mol. Biol.* 232:584–599.

9. Suzuki, D., M. Miyazaki, K. Jinde, T. Koji, M. Yagame, M. Endoh, Y. Nomoto, and H. Sakai. 1997. In situ hybridization studies of matrix metalloproteinase-3, tissue inhibitor of metalloproteinase-1 and type IV collagen in diabetic nephropathy. *Kidney Int.* 52:111–119.

10. Matsubara, K., and K. Okubo. 1993. cDNA analyses in the human genome project. *Gene.* 135:265–274.

11. Okubo, K., N. Hori, R. Matoba, T. Niiyama, A. Fukushima, Y. Kojima, and K. Matsubara. 1992. Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nat. Genet.* 2:173–179.

12. Carrell, R., and J. Travis. 1985. α1-Antitrypsin and the serpins: variation and countervariation. *Trends Biochem. Sci.* 10:20–24.

13. Carrell, R., P.A. Pemberton, and D.R. Boswell. 1987. The serpins: evolution and adaptation in a family of protease inhibitors. *Cold Spring Harbor Symp. Quant. Biol.* 52:527–535.

14. Kruithof, E.K.O., M.S. Baker, and C.L. Bunn. 1995. Biological and clinical aspects of plasminogen activator inhibitor type 2. *Blood.* 86:4007–4024.

15. Potempa, J., E. Korzus, and J. Travis. 1994. The serpin superfamily of proteinase inhibitors: structure, function, and regulation. *J. Biol. Chem.* 269:15957–15960.

16. Remold-O'Donnell, E. 1993. The ovalbumin family of serpin proteins. *FEBS Lett.* 315:105–108.

17. Aertgeerts, K., H.L. De Bondt, C.J. De Ranter, and P.J. Declerck. 1995. Mechanisms contributing to the conformational and functional flexibility of plasminogen activator inhibitor-1. *Nat. Struct. Biol.* 2:891–897.

18. Hopkins, P.C., and J. Whisstock. 1994. Function of mapsin. *Science.* 265:1893–1894.

19. Patston, P.A., and P.G.W. Gettins. 1996. Significance of secondary structure predictions on the reactive center loop region of serpins: a model for the folding of serpins into a metastable state. *FEBS Lett.* 383:87–92.

20. Wright, H.T. 1996. The structural puzzle of how serpin serine protease inhibitors work. *Bioessays.* 18:453–464.

21. Stein, P.E., and R.W. Carrell. 1995. What do dysfunctional serpins tell us about molecular mobility and disease? *Struct. Biol.* 2:96–113.

22. Gettins, P., P.A. Patston, and M. Schapira. 1993. The role of conformational change in serpin structure and function. *Bioessays.* 15:461–467.

23. Creighton, T.E., and I.G. Charles. 1987. Sequences of the genes and polypeptide precursors for two bovine protease inhibitors. *J. Mol. Biol.* 194:11–22.

24. Ye, R.D., T.-C. Wun, and J.E. Sadler. 1988. Mammalian protein secretion without signal peptide removal: biosynthesis of plasminogen activator inhibitor-2 in U-937 cells. *J. Biol. Chem.* 263:4869–4875.

25. Wohlwend, A., D. Belin, and J.D. Vassalli. 1987. Plasminogen activator-specific inhibitors in mouse macrophages: in vivo and in vitro modulation of their synthesis and secretion. *J. Immunol.* 139:1278–1284.

26. Wohlwend, A., D. Belin, and J.D. Vassalli. 1987. Plasminogen activator-specific inhibitors produced by human monocytes/macrophages. *J. Exp. Med.* 165:320–339.

27. Mikus, P., T. Urano, P. Liljestroem, and T. Ny. 1993. Plasminogen-activator inhibitor type 2 (PAI-2) is a spontaneously polymerising SERPIN: biochemical characterization of the recombinant intracellular and extracellular forms. *Eur. J. Biochem.* 218:1071–1082.

28. Genton, C., E.K.O. Kruithof, and W.D. Schleuning. 1987. Phorbol ester induces the biosynthesis of glycosylated and non-glycosylated plasminogen activator 2 in high excess over urokinase-type plasminogen activator in human U-937 lymphoma cells. *J. Cell Biol.* 104:705–712.

29. Miyata, T., K. Isobe, R. Inagi, R. Taguchi, I. Takai, Y. Fujita, T. Iwamoto, T. Hasegawa, K. Maeda, K. Yamada, and I. Nakashima. 1989. Rat mesangial cells actively produce phosphatidylinositol-anchored Thy-1. *Immunology.* 67:531–533.

30. Johnson, R.J., J. Floege, A. Yoshimura, H. Iida, W.G. Couser, and C.E. Alpers. 1992. The activated mesangial cells: a glomerular "myofibroblast?" *J. Am. Soc. Nephrol.* 2:S190–S197.

31. Floege, J., H.R. Radeke, and R.J. Johnson. 1994. Glomerular cells in vitro versus the glomerulus in vivo. *Kidney Int.* 45:360–368.

32. Miyata, T., K. Isobe, R. Dawson, M.A. Ritter, R. Inagi, O. Oda, R. Taguchi, H. Ikezawa, I. Inoue, H. Seo, M. Hasegawa, S. Kobayashi, K. Maeda, K. Yamada, and I. Nakashima. 1990. Determination of the molecular nature and cellular localization of Thy-1 in human renal tissue. *Immunology.* 69:391–395.

33. Bachmann, F. 1995. The enigma PAI-2: Gene expression, evolutionary and functional aspects. *Thromb. Haemostasis.* 74:172–179.

34. Eitzman, D.T., R.D. McCoy, X. Zheng, W.P. Fay, T. Shen, D. Ginsburg, and R.H. Simon. 1996. Bleomycin-induced pulmonary fibrosis in transgenic

mice that either lack or overexpress the murine plasminogen activator inhibitor-1 gene. *J. Clin. Invest.* 97:232–237.

35. Dickinson, J.L., E.J. Bates, A. Ferrante, and T.M. Antalis. 1995. Plasminogen activator inhibitor type 2 inhibits tumor necrosis factor alpha-induced apoptosis. *J. Biol. Chem.* 270:27894–27904.

36. Kumar, S., and C. Baglioni. 1991. Protection from tumor necrosis factor-mediated cytolysis by overexpression of plasminogen activator inhibitor type-2. *J. Biol. Chem.* 266:20960–20964.

37. Schechter, I., and A. Berger. 1967. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* 27:157–162.

38. Schick, C., Y. Kamachi, A.J. Bartuski, S. Cataltepe, N.M. Schechter, P.A. Pemberton, and G.A. Silverman. 1997. Squamous cell carcinoma antigen 2 is a novel serpin that inhibits the chyotrypsin-like proteases cathepsin G and mast cell chymase. *J. Biol. Chem.* 272:1849–1855.