Murine Heparin Cofactor II: Purification, cDNA Sequence, Expression, and Gene Structure†

Guang Sen Zhang,‡ Julie H. Mehringer,‡,§ Vivianna M. D. Van Deerlin,‡ Christine A. Kozak,‖ and Douglas M. Tollefsen*,‡

Division of Hematology–Oncology, Washington University School of Medicine, St. Louis, Missouri 63110, and National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892

Received August 26, 1993; Revised Manuscript Received December 13, 1993®

ABSTRACT: Heparin cofactor II (HCII) is a glycoprotein in human plasma that inhibits thrombin rapidly in the presence of dermatan sulfate or heparin. Unexpectedly, we found that HCII activity in murine plasma is present in two proteins of 68 and 72 kDa. The two proteins have the same N-terminal amino acid sequence, and both react with an antibody raised against the C-terminal nine amino acid residues of murine HCII predicted from the cDNA sequence. Treatment of the two proteins with peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase yields a single 54-kDa band. Thus, murine plasma contains two forms of HCII that appear to have identical amino acid sequences but differ in the composition of their N-linked oligosaccharides. HCII cDNA clones isolated from a murine liver library include a 1434 bp open reading frame following the first Met codon, a TAA stop codon, and 580 bp of 3′-untranslated sequence terminating in a poly(A) tail. The amino acid sequence deduced from the cDNA contains the N-terminal sequence of purified murine plasma HCII preceded by a 23-residue hydrophobic sequence presumed to be the signal peptide. The amino acid sequence of murine HCII is 87% identical to that of human HCII, the greatest variability occurring in the N-terminal portion of the protein. Northern blot analysis reveals a 2.3-kb HCII mRNA in murine and human liver, but no HCII mRNA is detectable in heart, brain, spleen, lung, skeletal muscle, kidney, testis, placenta, pancreas, or intestine. Southern blot analysis of restriction fragment length polymorphisms in progeny of interspecific and intersubspecific crosses indicates that mice have a single HCII gene (designated *Hcf2*), which maps to chromosome 16 between *Prm-1* and *Igl*. The murine HCII gene is ~7.1 kb in size and consists of at least four exons and three introns. The intron/exon origanization is identical to that of the human HCII gene except at the 5′ end, where the murine gene may lack a large intron in the 5′-untranslated region. Our results indicate that HCII is more highly conserved than the human and murine homologues of other serpins such as α_1-antitrypsin and α_1-antichymotrypsin.

Heparin cofactor II (HCII)[1] is a 66-kDa glycoprotein in human plasma that inhibits thrombin and chymotrypsin [for reviews, see Pratt et al. (1989), Tollefsen (1992), and Van Deerlin and Tollefsen (1992)]. The rate of inhibition of thrombin by HCII increases >1000-fold when heparin or dermatan sulfate is present (Tollefsen et al., 1983). HCII is a member of the serpin family, a group of >40 homologous proteins most of which have serine protease inhibitor activity (Huber & Carrell, 1989). The human serpin genes that have been characterized to date are each present as a single copy per haploid genome. By contrast, the genes for the murine serpins α_1-protease inhibitor and contrapsin have been duplicated to form clusters of at least 5 and 10 copies, termed the *Spi-1* and *Spi-2* loci, respectively (Borriello & Krauter, 1991; Inglis & Hill, 1991). The genes in each cluster have diverged from one another primarily in their reactive site coding sequences. This process appears to have produced inactivation or altered protease specificity of the duplicated gene products.

The mouse would be an attractive model in which to investigate the physiologic function of HCII, because of the possibility of gene deletion in embryonic stem cells. Before these studies can be considered, it is necessary to determine whether HCII is present in the mouse and whether the HCII gene exists in multiple copies. Pratt et al. (1988) have detected a dermatan sulfate-dependent thrombin inhibitor in murine plasma, but it is unclear whether the activity resides in a protein closely related to human HCII or, alternatively, in proteins that have arisen by duplication of other serpin genes. In this report, we show that murine plasma contains two proteins with different molecular weights that form complexes with thrombin in the presence of dermatan sulfate. The two proteins appear to have identical N-terminal and C-terminal polypeptide sequences but differ in the composition of their N-linked oligosaccharides. Both proteins are encoded by a murine liver cDNA that is very similar in sequence to that of human HCII. In addition, we show that a single copy of the HCII gene is present in the murine genome, that the intron/exon boundaries and chromosome localization are conserved between the two species, and that both murine and human HCII are expressed predominantly in the liver.

† Supported by NIH Grant HL-14147 (Specialized Center of Research in Thrombosis) and by a grant from the Monsanto Co. The first two authors of this paper contributed equally to the project.

* Author to whom correspondence should be addressed at the Division of Hematology–Oncology, Box 8125, Washington University School of Medicine, 660 South Euclid Ave., St. Louis, MO 63110.

‡ Washington University School of Medicine.

§ Fellow of the American Heart Association, Missouri Affiliate.

‖ National Institute of Allergy and Infectious Diseases.

® Abstract published in *Advance ACS Abstracts,* March 1, 1994.

[1] Abbreviations: HCII, heparin cofactor II; *Hcf2*, murine HCII gene; *HCF2*, human HCII gene; *Spi-1*, murine locus containing the α_1-protease inhibitor gene; *Spi-2*, murine locus containing the contrapsin gene; serpin, serine proteinase inhibitor; PCR, polymerase chain reaction; RFLP, restriction fragment length polymorphism; PEG, poly(ethylene glycol); TS, Tris–saline buffer; TCE, Tris–saline/EDTA buffer; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; PVDF, poly(vinylidene difluoride); PNGase F, peptide-N^4-(*N*-acetyl-β-glucosaminyl)-asparagine amidase.

0006-2960/94/0433-3632$04.50/0

MATERIALS AND METHODS

Materials. Human HCII and human thrombin were prepared as previously described (Blinder et al., 1989). Thrombin was labeled with ^{125}I by the chloramine T method to a specific activity of $\sim 1 \times 10^6$ cpm/pmol (Tollefsen & Blank, 1981). Bovine lung heparin and porcine skin dermatan sulfate were purchased from Sigma. The dermatan sulfate was treated with nitrous acid to remove traces of contaminating heparin (Teien et al., 1976). Poly(ethylene glycol) 8000 (PEG) was purchased from Sigma. Restriction enzymes were purchased from Promega and New England Biolabs. Oligonucleotides were synthesized in the Protein Chemistry Facility of Washington University.

Chromogenic Assay of HCII Activity in Murine Plasma. Murine plasma (0–4 μL) was incubated with human thrombin (7 nM) and dermatan sulfate (200 μg/mL) or heparin (67 μg/mL) at room temperature in 100 μL of 0.15 M NaCl, 0.02 M Tris-HCl, and 1 mg/mL PEG, pH 7.4 (TS/PEG buffer). Thrombin was added last to initiate the reaction. After 60 s, 500 μL of 100 μM tosyl-Gly-Pro-Arg-*p*-nitroanilide (Chromozym TH, Boehringer Mannheim) in TS/PEG buffer was added, and the absorbance at 405 nm was determined continuously for 100 s. The rate of change of absorbance was proportional to the concentration of active thrombin that remained in the incubation.

Formation of ^{125}I-Thrombin–HCII Complexes. Murine plasma (2.5 μL) was incubated with human ^{125}I-thrombin (5 nM) and dermatan sulfate (0–500 μg/mL) for 5 min at room temperature in 50 μL of TS/PEG buffer. The reaction was stopped by addition of 50 μL of electrophoresis sample buffer [40 mg/mL SDS, 10% (v/v) β-mercaptoethanol, 20% (v/v) glycerol, and 0.125 M Tris-HCl, pH 6.8], followed by heating at 100 °C for 5 min. A 20-μL portion of the sample was subjected to SDS–PAGE on a 7.5% polyacrylamide gel in the presence of β-mercaptoethanol according to Laemmli (1970). Autoradiography was performed as described (Tollefsen & Blank, 1981).

Isolation of Murine HCII. Frozen murine plasma anticoagulated with citrate was obtained from Pel-Freez Biologicals. The plasma was thawed at 37 °C and mixed with soybean trypsin inhibitor (20 mg/L), benzamidine (5 mM), and PEG (35 g/L) for 30 min at 4 °C. The resulting precipitate was removed by centrifugation. One-tenth volume of 1 M $BaCl_2$ was added slowly to the supernatant, the solution was stirred for 30 min at 4 °C, and the precipitate was removed by centrifugation. PEG (100 g/L) was mixed with the supernatant solution for 60 min at 22 °C, and the resulting precipitate was removed by centrifugation. PEG (150 g/L) was again mixed with the supernatant solution for 40 min at 22 °C. The precipitate was recovered by centrifugation, dissolved in 150 mL of 50 mM Tris, 20 mM sodium citrate, and 5 mM EDTA, pH 7.4 (TCE buffer), and applied to a 5 × 26 cm column of heparin–Sepharose (Pharmacia) equilibrated in TCE buffer. The column was washed and then step-eluted with TCE buffer containing 0.2 M NaCl. The eluate was pooled, dialyzed against 20 mM Tris-HCl, pH 7.4, and concentrated to ~3.7 mL using an Amicon YM-30 ultrafiltration membrane. The concentrated sample was applied to a Mono Q HR 5/5 column (Pharmacia) equilibrated with 20 mM Tris-HCl, pH 7.4. The column was washed and then eluted with a 50-mL linear gradient of 0–0.7 M NaCl in 20 mM Tris-HCl, pH 7.4, at a flow rate of 1 mL/min. HCII activity, which eluted from the Mono Q column in ~0.32 M NaCl, was determined with the chromogenic assay described above. The concentration of the final product was estimated by the absorbance at 280 nm, using the extinction coefficient (32 430 $M^{-1} \cdot cm^{-1}$) calculated according to Edelhoch (1967) from the amino acid composition of murine HCII deduced from the cDNA sequence.

N-Terminal Sequence Analysis. Approximately 3.6 μg of purified murine HCII was subjected to SDS–PAGE under reducing conditions and electrophoretically transferred onto a poly(vinylidene difluoride) (PVDF) membrane (Millipore). The membrane was stained with Coomassie Brilliant Blue R-250 as described by Matsudaira (1987). N-Terminal sequence analysis was performed on protein bands cut from the PVDF membrane using an Applied Biosystems Model 470A gas-phase sequencer.

Immunoblot Analysis of Purified Murine HCII. The peptide CGKVTNPAKS, which contains the C-terminal nine amino acid sequence of murine HCII, was synthesized in the Protein Chemistry Facility of Washington University. Rabbits were immunized with the peptide coupled to keyhole limpet hemocyanin via the sulfydryl group (Harlow & Lane, 1988), and the resulting immunoglobulins were purified on protein A-Sepharose (Sigma). Purified murine HCII was subjected to SDS–PAGE under reducing conditions, transferred to a nitrocellulose membrane, and probed with the rabbit anti-murine C-terminal peptide antibody followed by ^{125}I-labeled goat anti-rabbit IgG antibody (ICN Radiochemicals) as described previously (Blinder et al., 1988).

PNGase F Treatment. Recombinant peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase (PNGase F) was obtained from Genzyme. Approximately 7 μg of purified murine HCII was denatured by heating at 100 °C for 5 min and then incubated at 37 °C for 24 h with 0.75 unit of PNGase F in 25 μL of 1.25% (v/v) NP-40, 1.7 mg/mL SDS, and 0.2 M Tris-HCl, pH 8.6. The reaction products were analyzed by SDS–PAGE and immunoblotting as described above.

Amplification of a Murine HCII Gene Fragment. A portion of the HCII gene was amplified by the polymerase chain reaction (PCR) from murine genomic DNA with reagents obtained from Perkin-Elmer Cetus. The primers used were 20 nt in length and spanned nt 938–1178 of the human HCII cDNA sequence (Blinder et al., 1988). Approximately 1 μg of murine genomic DNA (Clontech) was incubated with 1.0 μM of each primer, 200 μM of each deoxynucleotide triphosphate, and 2.5 units of *Taq* DNA polymerase in a total volume of 100 μL. PCR was performed for 30 cycles under the following conditions: denaturation for 2 min at 94 °C; annealing for 2 min at 40 °C; and extension for 3 min at 72 °C. The PCR product (~240 bp in length) was ligated directly into the plasmid supplied in the TA Cloning Kit (Invitrogen) and sequenced.

Isolation of Murine HCII cDNA Clones. The PCR product was excised from the TA cloning plasmid with *Eco*RI and *Hind*III, purified by electrophoresis in low melting point agarose, and biotin-labeled by nick translation according to instructions supplied with the PhotoGene Labeling and Detectin System (Gibco BRL). The labeled PCR product (designated probe A) was used to screen a murine liver cDNA library in λZAP II (Stratagene) constructed by Dr. Rick Wetsel, Department of Pediatrics, Washington University. Approximately 1.2×10^5 λZAP II bacteriophage were plated on *Escherichia coli* XL1-Blue cells, and plaque DNA was transferred to Hybond-N nylon membranes (Amersham). Hybridization and detection with streptavidin–alkaline phosphatase conjugate were carried out according to the instructions supplied with the PhotoGene kit. Two positive phage were identified, and their DNA inserts (designated mHCII1.5

and mHCII1.8) were recovered in the pBluescript SK phagemid by coinfection with helper phage R408 (Stratagene).

Because neither of the original clones contained the complete 5′-coding sequence, a murine liver cDNA library in λgt11 enriched in 5′ sequences (5′-stretch, Clontech) was screened with the *Bam*HI/*Bgl*I fragment isolated from the 5′ end of clone mHCII1.8 (probe B). Probe B labeled by random-primed incorporation of digoxigenin-11–dUTP (Genius 1 Kit, Boehringer Mannheim) was used to screen 6×10^5 phage plaques. Hybridization and detection with alkaline phosphatase-conjugated antibody to digoxigenin were performed according to the manufacturer's directions. Six positive phage were plaque-purified. Phage DNA was isolated by a plate lysis method (Sambrook et al., 1989). The inserts released by *Eco*RI digestion were subcloned into pBluescript KS (Stratagene) and partially sequenced. The insert extending furthest in the 5′ direction (mHCII2.11) was sequenced in its entirety.

Northern Blot Analysis. Probes used were a human *Bsu*36I/*Xho*I fragment (nt 526–1194) (Blinder et al., 1988) and a murine *Bam*HI/*Sna*BI fragment (nt 392–1072) from the corresponding cDNAs. All probes were purified following restriction digestion by agarose gel electrophoresis and Geneclean (Bio 101). A control human β-actin probe was supplied by Clontech for use with both human and murine blots. Probes were labeled with [α-^{32}P]dCTP (DuPont–New England Nuclear) using the Mixed Primer Labeling System II (Clontech) according to kit instructions. Human and murine multiple tissue Northern blots were purchased from Clontech. Hybridization was performed as recommended by the manufacturer.

Isolation of Murine HCII Genomic Clones. 5′ murine HCII (5′ HCII, nt 392–769) and 3′ murine HCII (3′ HCII, nt 1219–1942) probes were generated by *Bam*HI/*Bgl*I digestion and by *Hin*dIII digestion, respectively, of murine cDNA. Probes were labeled with digoxigenin-11–dUTP as described above. A murine liver genomic library in EMBL-3/SP6/7 λ phage (Clontech) was propagated in *E. coli* LE392 and plated at a density of 50 000 plaques per 150-mm plate. Duplicate lifts on Hybond-N nylon filters (Amersham) were hybridized to digoxigenin-labeled probes. Hybridization and detection were performed according to the Genius System kit except that hybridization was carried out at 65 °C for 14 h, and detection was with Lumi-Phos 530 (Boehringer Mannheim). Positive phage were plaque-purified, and phage DNA was purified by the liquid-lysis method of Garber et al. (1983). The λ clones were digested with *Xho*I/*Sac*I or *Sac*I alone and the insert fragments were ligated into pBluescript KS phagemid (Stratagene) without further purification.

PCR Amplification of Introns. DNA was amplified by PCR under conditions described previously (Blinder et al., 1989). When extended regions were amplified (>500 bp), the extension cycles were lengthened to 5 min. Synthetic oligonucleotide primers were 19–21 nt in length.

Chromosome Localization. Two genetic crosses were used for genetic mapping: (NFS/N or C58/J × *Mus musculus musculus*) × *M. m. musculus* (Kozak et al., 1990) and (NFS/N × *Mus spretus*) × *M. spretus* or C58/J (Adamson et al., 1991). Progeny of these crosses have been typed for over 425 markers distributed on the 19 autosomes and the X chromosome. Southern blot analysis was used to type DNAs from these progeny for *Hcf2* using the 5′ HCII probe, and for the following chromosome 16 markers: *Prm-1* (protamine-1) using as probe a 500 bp *Sal*I fragment excised from clone GZPrm-1 (Kleene et al., 1983) obtained from N. Hecht (Tufts University, Medford, MA); *Igl* (immunoglobulin λ light chain) using as probe p 1 (Scott et al., 1982) obtained from K. Huppi (National Cancer Institute, National Institutes of Health); *Smst* (somatostatin) using as probe a 2.7-kb *Eco*RI/*Hin*dIII fragment of the clone pgHS7-2.7 obtained from the American Type Culture Collection (ATCC, Rockville, MD); *Mtv-37* (mammary tumor virus integration-37) using as probe a 1.4-kb *Pst*I fragment of the C3H MMTV (Majors & Varmus, 1981) obtained from R. Callahan (National Cancer Institute, National Institutes of Health); and *Pit-1* (pituitary specific transcription factor 1) using as probe the *Bam*HI/*Hin*dIII insert from the clone pmPit1 obtained from the ATCC.

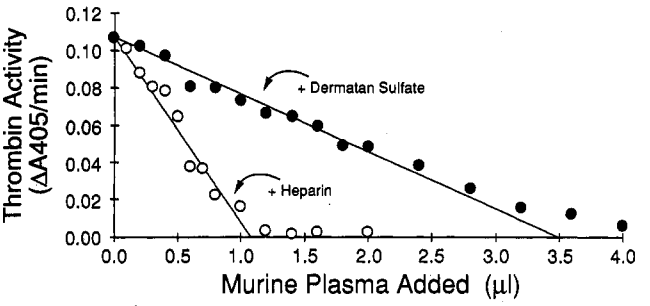

FIGURE 1: Glycosaminoglycan-dependent thrombin inhibitors in murine plasma. Murine plasma (0–4 μL) was incubated with human thrombin (7 nM) in 100 μL of TS/PEG buffer in the presence of dermatan sulfate (200 μg/mL) (●) or heparin (67 μg/mL) (○). After 60 s, the remaining thrombin activity was determined by hydrolysis of the chromogenic substrate tosyl-Gly-Pro-Arg-*p*-nitroanilide (ΔA_{405}/min). No inhibition of thrombin occurred during a 5-min incubation with 4 μL of plasma in the absence of dermatan sulfate (not shown).

Prm-1 was typed in the *musculus* cross following digestion with *Sca*I and in the *spretus* cross using *Hin*dIII; *Igl* was typed in these crosses using, respectively, *Apa*I and *Pvu*II. Progeny of the *spretus* cross were typed for an *Eco*RI polymorphism of *Smst* and a *Bam*HI polymorphism of *Pit-1*. *Mtv-37* (Siracusa et al., 1991) is an *M. spretus* specific retroviral integration which was identified in the C58/J progeny of the *spretus* cross as a 19.4-kb *Pvu*II fragment. Recombination between locus pairs and standard errors were calculated according to Green (1981).

DNA Sequencing. Sequencing of plasmid DNA was performed by a modification (Hattori & Sakaki, 1986) of the dideoxy chain termination technique (Sanger et al., 1977) using modified T7 DNA polymerase (Sequenase 2, United States Biochemical Corp.), [α-^{35}S]ATP (DuPont–New England Nuclear), and 6% denaturing polyacrylamide gels. Sequences were determined in both directions using nested synthetic oligonucleotide primers. DNA sequences were analyzed on a VAX/VMS computer using software obtained from the University of Wisconsin Genetics Computer Group (Devereux et al., 1984).

Molecular Modeling. Surface and buried amino acid residues were modeled by superimposing the polypeptide chain of HCII upon the tertiary structure of ovalbumin (Stein et al., 1990) using the program Sybyl (Tripos).

RESULTS

Quantification of Dermatan Sulfate Cofactor in Murine Plasma. To quantify the glycosaminoglycan-dependent thrombin inhibitors in murine plasma, we incubated a constant amount of human thrombin with 0–4 μL of pooled murine plasma for 60 s in the presence of dermatan sulfate or heparin and then determined the residual thombin activity with a chromogenic substrate (Figure 1). The *x*-intercepts of the

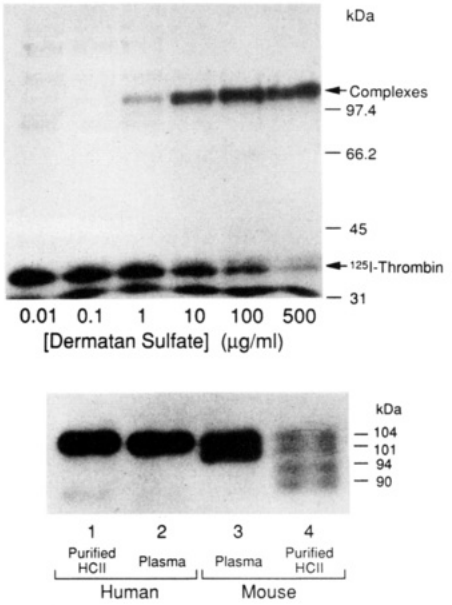

FIGURE 2: Formation of ^{125}I-thrombin–inhibitor complexes. Top panel, A: Murine plasma (2.5 μL) was incubated with human ^{125}I-thrombin (5 nM) and dermatan sulfate (0.01–500 μg/mL) for 5 min in 50 μL of TS/PEG buffer. A portion of each incubation was subjected to SDS–PAGE (200 V, 2 h) under reducing conditions. An autoradiograph of the gel is shown. Botton panel, B: Unfractionated plasma or purified HCII was incubated with human ^{125}I-thrombin (5 nM) and dermatan sulfate (100 μg/mL) for 5 min in 50 μL of TS/PEG buffer and then subjected to SDS–PAGE (200 V, 3 h) and autoradiography. The region of the autoradiograph containing ^{125}I-thrombin–inhibitor complexes is shown. *Lane 1*, purified human HCII (0.6 μg). *Lane 2*, human plasma (1 μL). *Lane 3*, murine plasma (2.5 μL). *Lane 4*, purified murine HCII (3.4 μg) after storage for 20 days at 4 °C. In each lane, the total amount of ^{125}I-thrombin complexed to HCII was approximately the same.

titration curves indicate that ~0.2 μmol of thrombin is inhibited per liter of plasma in the presence of dermatan sulfate and ~0.6 μmol per liter in the presence of heparin.

Evidence for Two Dermatan Sulfate Cofactors in Murine Plasma. We incubated murine plasma with ^{125}I-thrombin for 5 min in the presence of dermatan sulfate and then performed SDS–PAGE with autoradiography to detect stable thrombin–inhibitor complexes. Maximal complex formation occurred in the presence of 10–100 μg/mL dermatan sulfate (Figure 2A, top panel). Prolonged electrophoresis resolved the complexes formed in the presence of 100 μg/mL dermatan sulfate into two bands of equal intensity having apparent masses of 101 and 104 kDa (Figure 2B, bottom panel, lane 3). To rule out the possibility that the 101-kDa band was derived from the 104-kDa band by limited proteolysis after complex formation, we incubated plasma with ^{125}I-labeled thrombin in the presence of 100 μg/mL dermatan sulfate for various times (0.5–30 min) prior to SDS–PAGE. Equal amounts of the two complexes were observed at each time point (data not shown), which indicates that the 104-kDa complex was not progressively converted to the 101-kDa complex. Similar experiments with human plasma and purified human HCII each yielded a single complex of 104 kDa (Figure 2B, lanes 1 and 2). These results suggest that murine plasma contains approximately equal amounts of two dermatan sulfate-dependent thrombin inhibitors that differ in mass by ~3 kDa.

Purification of Murine HCII. The dermatan sulfate cofactor activity was purified from murine plasma by a procedure that included poly(ethylene glycol) precipitation followed by chromatography on heparin–Sepharose and Mono Q (see Materials and Methods for details). The final product consisted of two bands of equal intensity having apparent masses of 72 and 68 kDa as determined by SDS–PAGE under reducing conditions (Figure 3, lane 2). The 72-kDa band comigrated with purified human HCII (lane 3). Both the 72-kDa band and the 68-kDa band formed covalent complexes with ^{125}I-thrombin in the presence of dermatan sulfate (see below). The yield of protein from 150 mL of starting plasma was ~12 nmol based on the A_{280} of the purified product (see Materials and Methods), which equals ~40% of the dermatan sulfate confactor estimated to be present in the starting material.

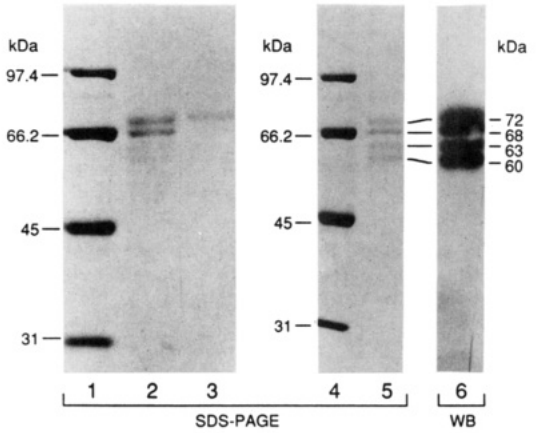

FIGURE 3: SDS–PAGE of purified murine HCII. HCII purified from murine or human plasma as described under Materials and Methods was subjected to SDS–PAGE under reducing conditions. *Lanes 1–5 (SDS–PAGE)* were stained with Coomassie Brilliant Blue R-250. A Western blot of *lane 6 (WB)* was probed with rabbit antibody against the murine HCII C-terminal peptide followed by ^{125}I-labeled goat anti-rabbit IgG antibody. An autoradiograph of the Western blot is shown. No bands were detected on a duplicate Western blot probed with preimmune IgG. *Lanes 1 and 4*, molecular mass standards. *Lane 2*, purified murine HCII (3.4 μg). *Lane 3*, purified human HCII (0.6 μg). *Lanes 5 and 6*, purified murine HCII (3.4 μg) after storage at 4 °C.

During storage of the purified preparation at 4 °C, we observed the appearance of two degradation products having apparent masses of 63 and 60 kDa (Figure 3, lane 5). Incubation of the partially degraded preparation with ^{125}I-thrombin in the presence of dermatan sulfate yielded four complexes ranging from 90 to 104 kDa (Figure 2B, lane 4), which indicates that all four bands possess dermatan sulfate cofactor activity. The upper two complexes, derived from the 68- and 72-kDa bands, comigrated with the two ^{125}I-thrombin–inhibitor complexes formed in unfractionated murine plasma (Figure 2B, lane 3). Therefore, the 68- and 72-kDa proteins appear to comprise the bulk of the dermatan sulfate cofactor in murine plasma.

N-Terminal Sequence Analysis. The protein shown in Figure 3 (lane 5) was transferred to a PVDF membrane following SDS–PAGE, and each of the four bands was subjected to amino acid sequence analysis. The N-terminal amino acid sequences of the 72- and 68-kDa bands were identical to one another (EQLTNEXLTTSFLPANFHKEN; X indicates a position at which no amino acid could be identified) and were 65% identical to the sequence between positions 24 and 44 of human HCII (Blinder et al., 1988). Therefore, the 72- and 68-kDa bands represent two forms of murine HCII. The 63- and 60-kDa bands shared a different N-terminal sequence (XFXKENTVTN), which suggests that they arose from cleavage after Ala^{15} of the 72- and 68-kDa forms, respectively. As shown in Figure 3 (lane 6), all four bands reacted with a rabbit antibody raised against a synthetic

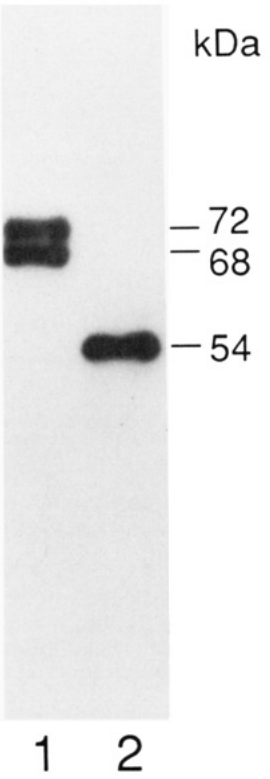

FIGURE 4: PNGase F treatment of purified murine HCII. Purified murine HCII (7 μg) was treated with PNGase F as described under Materials and Methods. A Western blot of the reaction mixture probed with rabbit antibody against the murine HCII C-terminal peptide is shown. *Lane 1*, untreated sample. *Lane 2*, sample treated with PNGase F.

peptide corresponding to the C-terminal nine amino acid residues of murine HCII deduced from the cDNA sequence (see below). Treatment of the 72- and 68-kDa forms with PNGase F yielded a single band of 54 kDa (Figure 4). Therefore, the 72- and 68-kDa forms of murine HCII appear to have identical N-terminal and C-terminal amino acid sequences but differ in the composition of their N-linked oligosaccharides.

Isolation and Characterization of Murine HCII cDNAs. We isolated cDNA clones that allowed us to deduce the complete amino acid sequence of murine HCII. Initially, we amplified a fragment of murine genomic DNA by PCR using a forward primer that corresponded to nt 938–957 and a reverse primer complementary to nt 1159–1178 of the human HCII cDNA. Both of these sequences are present in exon 3 of the human HCII gene (Herzog et al., 1991). The PCR product contained 201 nt exclusive of the primer sequences and was 85% identical to the human HCII cDNA sequence. The cloned PCR product (probe A, Figure 5) was used to screen a murine liver cDNA library in λZAP II. Approximately 120 000 phage were screened, yielding 2 positive clones (mHCII1.5 and mHCII1.8) with inserts of 1.5 and 1.8 kb, respectively. The 5′ end of the mHCII1.8 insert corresponded to the human HCII cDNA sequence beginning with nt 297 (codon 71), and the 3′ end contained a poly(A) sequence. The sequence of the mHCII1.5 insert was identical to the 3′ region of the mHCII1.8 insert. To obtain additional 5′-coding sequence, we used the *Bam*HI/*Bgl*I fragment of the mHCII1.8 insert (probe B, Figure 5) to screen a murine liver cDNA library in λgt11 enriched in 5′ sequences. Screening of approximately 600 000 phage yielded 6 positive clones with inserts ranging in size from 1.1 to 2.5 kb. One of the clones (mHC1II2.11) contained a 1.2-kb insert that extended 301 bp 5′ from the end of clone

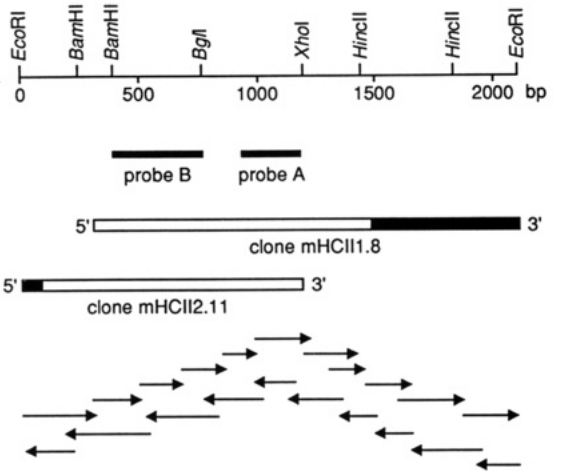

FIGURE 5: Sequencing strategy and partial restriction map of murine HCII cDNA clones. Clone mHCII1.8 was isolated from a murine liver cDNA library in λZAP II probed with a PCR product derived from murine genomic DNA (*probe A*). Clone mHCII2.11 was isolated from a murine liver "5′-stretch" cDNA library in λgt11 probed with the *Bam*HI/*Bgl*I fragment of mHCII1.8 (*probe B*). The cDNA inserts were sequenced on both strands using nested synthetic oligonucleotide primers as shown by the *arrows*. The *open areas* of the two inserts indicate the open reading frame. A partial restriction map derived from the sequence of the composite cDNA is indicated.

mHCII1.8. The other clones did not provide any additional 5′–sequence information. Clones mHCII1.8 and mHCII2.11 were sequenced on both strands as shown in Figure 5. The complete sequence of the murine HCII cDNA obtained from the two overlapping clones and the deduced amino acid sequence beginning with the first Met codon are shown in Figure 6.

Expression of HCII mRNA in Tissues. We probed Northern blots to determine the tissue specificity of murine and human HCII mRNA expression. Murine tissues studied were heart, brain, spleen, lung, liver, skeletal muscle, kidney, and testis (Figure 7), and intestine (not shown). Human tissues were the same except for the absence of spleen and testis and the addition of placenta and pancreas. In both species, using probes of about 700 bp derived from cDNAs for murine or human HCII, expression was detected only in the liver (Figure 7). In order to ensure that mRNA was present in all the lanes, the same blots were reprobed with a 2-kb human β-actin cDNA which cross-hybridizes strongly with murine mRNA. The β-actin control probe detected an mRNA of approximately 2 kb in each lane. As previously reported, a smaller form of β-actin mRNA was also present in heart and skeletal muscle (Lamballe et al., 1991; Pari et al., 1991). The murine and human HCII transcripts were both approximately 2.2–2.3 kb in length (Figure 7). An additional blot (not shown), in which murine and human liver RNAs were present in adjacent lanes of the same gel, confirmed that the murine and human HCII transcripts were identical in size.

Isolation of Murine HCII Genomic Clones. Two overlapping clones containing HCII sequences were isolated from an adult murine DBA/2J liver library in EMBL-3/Sp6/7 λ phage. Approximately 500 000 plaques were screened separately with the 5′ HCII and 3′ HCII probes (Figure 8). Both probes were derived from the murine cDNA as described under Materials and Methods. One plaque was positive with the 5′ HCII probe, and two were positive with the 3′ HCII probe. *Sac*I digestion of the clone hybridizing to the 5′ HCII probe (G-1.3) yielded 8- and 3-kb fragments in addition to the λ arms. G-1.3 was digested first with *Xho*I, which cuts in the λ polylinker but not in the insert, and then with *Sac*I, and

```
                                1 CTGTCTGCACATTCCAGTGGGGGTGGTGTCAGCCAGGCTCCCTGCCATTCCTTGTTGTCTGTTTTTCACAGCAGCTTTACGTCTGCAGAA   90
                                                                  ∇                                CHO
-23 Met Lys His Pro Leu Cys Thr Leu Leu Ser Leu Ile Thr Phe Met Cys Ile Gly Ser Lys Gly Leu Ala Glu Gln Leu Thr Asn Glu Asn   7
 91 ATG AAA CAT CCA CTC TGC ACT CTT CTT TCT CTC ATC ACA TTT ATG TGT ATT GGG AGC AAA GGC CTT GCT GAG CAG CTA ACT AAT GAA AAC  180
                                                                                                 SO4
  8 Leu Thr Thr Ser Phe Leu Pro Ala Asn Phe His Lys Glu Asn Thr Val Thr Asn Asp Trp Ile Pro Glu Gly Glu Glu Asp Glu Asp Tyr  37
181 CTG ACC ACG TCC TTC CTG CCT GCT AAC TTC CAC AAA GAA AAC ACA GTC ACC AAT GAC TGG ATC CCA GAG GGG GAA GAG GAT GAG GAC TAC  270
                                    SO4
 38 Leu Asp Leu Glu Lys Leu Leu Gly Glu Asp Asp Asp Tyr Ile Tyr Ile Ile Asp Ala Val Ser Pro Thr Asp Ser Glu Ser Ser Ala Gly  67
271 CTG GAC CTG GAG AAG CTG CTT GGT GAA GAT GAC GAC TAC ATT TAT ATC ATC GAT GCC GTT TCC CCG ACA GAC TCA GAA TCA AGT GCT GGC  360

 68 Asn Ile Leu Gln Leu Phe Gln Gly Lys Ser Arg Ile Gln Arg Leu Asn Ile Leu Asn Ala Lys Phe Ala Phe Asn Leu Tyr Arg Val Leu  97
361 AAC ATC CTG CAG CTT TTC CAA GGC AAG AGC CGG ATC CAG CGT CTT AAT ATC CTC AAT GCA AAG TTT GCC TTC AAC CTT TAC CGA GTC CTG  450

 98 Lys Asp Gln Ala Thr Thr Ser Asp Asn Leu Phe Ile Ala Pro Val Gly Ile Ser Thr Ala Met Gly Met Ile Ser Leu Gly Leu Arg Gly  127
451 AAG GAC CAG GCT ACC ACA TCT GAT AAC CTC TTC ATA GCA CCT GTT GGC ATT TCT ACT GCC ATG GGG ATG ATC TCC TTA GGT CTG AGG GGA  540
                                                  CHO
128 Glu Thr His Glu Glu Val His Ser Val Leu His Phe Arg Asp Phe Val Asn Ala Ser Ser Lys Tyr Glu Val Thr Thr Ile His Asn Leu  157
541 GAG ACC CAT GAA GAA GTA CAC TCA GTT CTA CAC TTC AGA GAC TTT GTC AAT GCT AGC AGC AAG TAC GAG GTT ACC ACC ATT CAC AAT CTC  630

158 Phe Arg Lys Leu Thr His Arg Leu Phe Arg Arg Asn Phe Gly Tyr Thr Leu Arg Ser Val Asn Gly Leu Tyr Ile Gln Lys Gln Phe Pro  187
631 TTT CGA AAG CTG ACC CAT CGT CTA TTC AGG AGG AAT TTT GGG TAC ACA CTT CGG TCA GTT AAT GGC CTT TAT ATT CAG AAG CAG TTT CCC  720

188 Ile Arg Glu Asp Phe Lys Ala Ala Met Arg Glu Phe Tyr Phe Ala Glu Ala Gln Glu Ala Asn Phe Pro Asp Pro Ala Phe Ile Ser Lys  217
721 ATA CGG GAG GAC TTC AAA GCT GCT ATG AGA GAG TTT TAC TTT GCT GAG GCC CAG GAG GCT AAC TTC CCA GAT CCT GCC TTC ATA TCA AAG  810

218 Ala Asn Asn His Ile Leu Lys Leu Thr Lys Gly Leu Ile Lys Glu Ala Leu Glu Asn Ile Asp Pro Ala Thr Gln Met Leu Ile Leu Asn  247
811 GCA AAC AAC CAC ATT CTG AAA CTC ACC AAG GGC CTT ATA AAA GAA GCT CTG GAG AAT ATA GAT CCT GCT ACT CAG ATG CTG ATT CTG AAC  900

248 Cys Ile Tyr Phe Lys Gly Thr Trp Val Asn Lys Phe Pro Val Glu Met Thr His Asn His Asn Phe Arg Leu Asn Glu Arg Glu Val Val  277
901 TGC ATC TAC TTC AAA GGA ACT TGG GTG AAT AAA TTC CCA GTA GAA ATG ACG CAC AAC CAC AAC TTC CGG CTG AAT GAG AGA GAA GTA GTC  990

278 Lys Val Ser Met Met Gln Thr Lys Gly Asn Phe Leu Ala Ala Asn Asp Gln Glu Leu Asp Cys Asp Ile Leu Gln Leu Glu Tyr Val Gly  307
991 AAA GTC TCC ATG ATG CAA ACT AAA GGG AAT TTC CTT GCG GCA AAT GAC CAG GAG CTA GAC TGT GAC ATT CTC CAG CTG GAG TAC GTA GGG 1080

308 Gly Ile Ser Met Leu Ile Val Val Pro Arg Lys Leu Ser Gly Met Lys Thr Leu Glu Ala Gln Leu Thr Pro Gln Val Val Glu Arg Trp  337
1081 GGC ATC AGC ATG CTT ATT GTA GTT CCA CGG AAG CTA TCG GGA ATG AAG ACT CTT GAA GCA CAG CTT ACA CCC CAG GTG GTG GAG AGA TGG 1170
          CHO
338 Gln Lys Ser Met Thr Asn Arg Thr Arg Glu Val Leu Leu Pro Lys Phe Lys Leu Glu Lys Asn Tyr Asn Leu Val Glu Val Leu Lys Ser  367
1171 CAA AAA AGC ATG ACA AAC AGA ACT CGA GAG GTA CTT CTG CCA AAG TTT AAG CTT GAG AAG AAT TAC AAC CTG GTG GAG GTC CTT AAG TCC 1260
                CHO
368 Met Gly Ile Thr Lys Leu Phe Asn Lys Asn Gly Asn Met Ser Gly Ile Ser Asp Gln Arg Ile Ala Ile Asp Leu Phe Lys His Gln Ser  397
1261 ATG GGA ATC ACA AAG CTC TTC AAC AAG AAT GGC AAT ATG TCA GGC ATC TCA GAC CAG AGG ATC GCC ATT GAC CTG TTC AAG CAC CAA AGT 1350
                                                                  ◊
398 Thr Ile Thr Val Asn Glu Glu Gly Thr Gln Ala Ala Ala Val Thr Thr Val Gly Phe Met Pro Leu Ser Thr Gln Val Arg Phe Thr Val  427
1351 ACC ATC ACA GTG AAT GAA GAG GGT ACC CAA GCT GCA GCT GTA ACT ACA GTG GGG TTC ATG CCG CTG TCT ACC CAA GTC CGA TTC ACT GTC 1440

428 Asp Arg Pro Phe Leu Phe Leu Val Tyr Glu His Arg Thr Ser Cys Leu Leu Phe Met Gly Lys Val Thr Asn Pro Ala Lys Ser  *      455
1441 GAC CGT CCC TTC CTG TTC CTA GTG TAT GAG CAC CGA ACC AGC TGT CTG CTC TTC ATG GGA AAA GTG ACC AAC CCT GCC AAG TCT TAA TGG 1530

1531 CAGTGGTCTGGGCACCTGATGTAGTTTGGGGACACTCTGTATTCCTGTTTTCATTCTAACAAAGAGGACAGAGCTATTTTGCTATGATAATCTCTGTGGTTTATGATAATAATCTGAAC 1649

1650 AGCAAACAGTGCCATTAAGTGCACAGATGTTAATGCAACTAATGGTTGTGCCAGCTATTCCCTAGAGGTCCTAGACACACAGACTGCCTTAGCCCCCACACCTCTGTCCCTAGGAGAAT 1768

1769 GAATGGGCAGAGTGGAGCTGACACTCTCTCCTCAGGCAGCCTCCTCAGCCTTTATTCGTCAACGGCCTCTGCTCAGCCAGGCCCTTCCAGACAAAAGGCACTGAGGAACTTGCACCCCT 1887

1888 TCCAGGGAGGGTGTCCAAGTAGCACCGTTCCTGCCCAGTGGGAAAGCAATCCCAACATGTCTGAGATGTTCGGACTTTTAAATTCATGTAAATATCTCCTCTTCTGCTGTGTGTACATA 2006

2007 GGACTACATAGTCTATCCTAGAGATAAATAAACAACCATATTTACCATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2104
```

FIGURE 6: Nucleotide and deduced amino acid sequences of murine HCII cDNA. The proposed signal peptidase cleavage site (∇), reactive site (◊), TAA stop codon (*), asparagine-linked glycosylation sites (CHO), and tyrosine sulfation sites (SO4) are shown. The polyadenylation signal (AATAAA) is underlined. The *dots* mark every tenth nucleotide or amino acid residue.

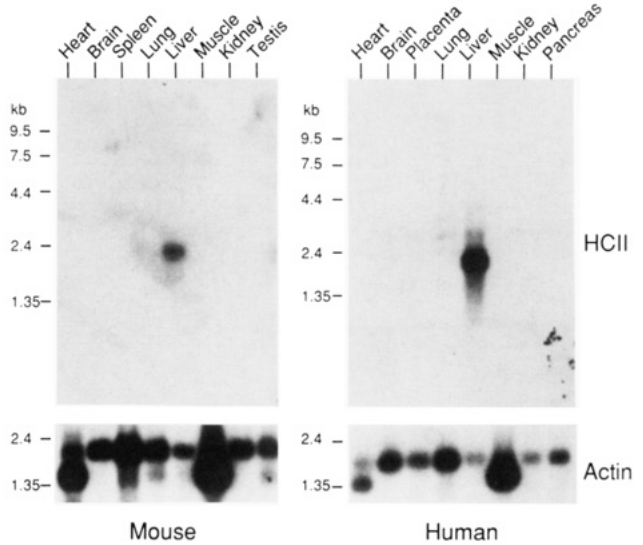

FIGURE 7: Expression of murine HCII mRNA in tissues. Murine (*left*) and human (*right*) multiple-tissue Northern blots (Clontech) were hybridized with HCII cDNA probes (*top*) or with a control human β-actin probe (*bottom*). HCII probes were a murine *Bam*HI/*Sna*BI cDNA fragment (nt 392–1072) and a human *Bsu*36I/*Xho*I (nt 526–1194) cDNA fragment. Equal amounts of poly(A)$^+$ RNA were present in each lane.

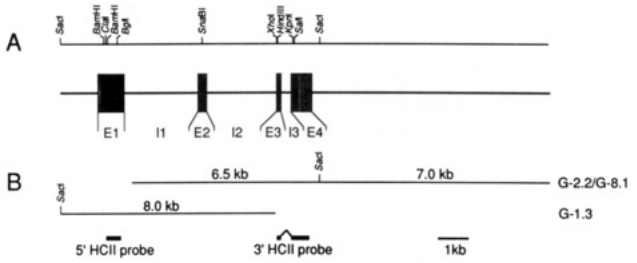

FIGURE 8: Structure of the murine HCII gene and genomic clones. (A) Partial restriction map and structure of the murine HCII gene. The *boxes* represent exons 1–4 (E1–E4). The *shaded areas* represent untranslated regions of exons 1 and 4. The three introns are designated I1–I3. (B) Inserts from EMBL-3 λ clones isolated from a murine genomic library. Shown below the clones are the 5′ HCII and 3′ HCII probes which were derived from the murine cDNA as described. The 5′ HCII probe was used to isolate λ clone G-1.3. The 3′ HCII probe was used to isolate λ clones G-2.2 and G-8.1.

fragments were ligated into pBluescript KS phagemid. Nine subclones from the G-1.3 ligation chosen for further analysis all contained 8-kb inserts. The 3-kb fragment was not recovered.

*Sac*I digestion of both λ clones hybridizing to the 3′ HCII probe (G-2.2 and G-8.1) yielded 6.5- and 7.0-kb fragments which were subcloned into pBluescript KS. Two G-8.1 subclones had 6.5-kb inserts, two G-2.2 subclones had 6.5-kb inserts, and three G-2.2 subclones had 7-kb inserts. Restriction mapping of the λ inserts, combined with PCR and sequence data, indicates that the λ clones overlap as shown in Figure 8. Both of the internal *Sac*I sites in the λ clones were in

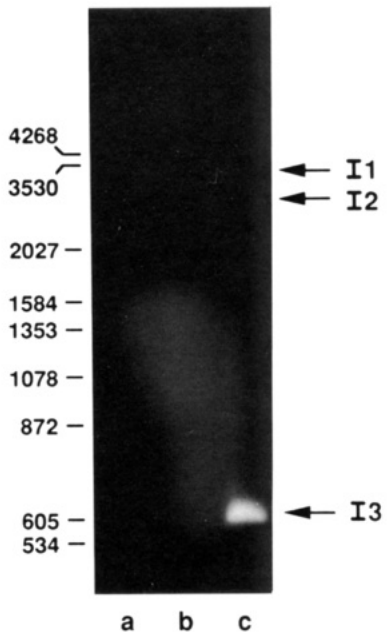

FIGURE 9: PCR analysis of intron size. Intron size was estimated by PCR amplification of genomic clones using primers near the intron/exon boundaries. The amplified product sizes were determined by mobility in agarose gel electrophoresis as compared to standards shown on the left. *Lane a*, intron 1 (I1) amplified from clone G-1.3; *lane b*, intron 2 (I2) amplified from clone G-2.2; *lane c*, intron 3 (I3) amplified from clone G-2.2. Calculated sizes for the PCR products were the following: intron 1, 3300 bp; intron 2, 2900 bp; intron 3, 630 bp. Subtraction of the distance between primers in the cDNA yielded net sizes for the introns as follows: intron 1, 2900 bp; intron 2, 2700 bp; intron 3, 400 bp.

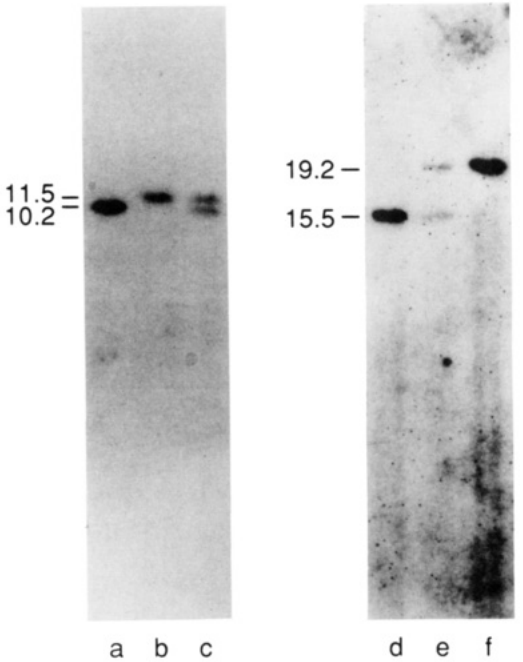

FIGURE 10: Southern blot analysis of HCII RFLPs. Genomic DNAs from parental mice and from progeny of (NFS/N × *M. spretus*) × C58/J (left) or (NFS/N × *M. m. musculus*) × *M. m. musculus* (right) mice were digested with *Hin*dIII or *Sst*I, respectively. Fragments were separated by agarose gel electrophoresis and hybridized to [α-^{32}P]ATP-labeled murine 5′ HCII probe. Blots were washed with 0.2 × SSC at 65 °C. *Lane a*, *M. spretus*; *lane b*, NFS/N; *lane c*, progeny of the *spretus* cross; *lane d*, *M. m. musculus*; *lane e*, progeny of the *musculus* cross; *lane f*, NFS/N.

noncoding regions since none of the exon sequence (see below) contains a *Sac*I site.

Intron/Exon Boundaries. Intron/exon boundaries were sequenced using primers derived from the cDNA sequence except in the case of the exon 3/intron 3 boundary, in which a primer corresponding to the 3′ end of intron 2 was used. Alignment of the genomic and cDNA sequences indicates that the gene consists of at least four exons and three introns (Figure 8). Exon 1 contains nt 1–916 of the murine HCII cDNA,[2] exon 2, nt 917–1190; exon 3, nt 1191–1335; and exon 4, nt 1336–3′-untranslated sequence. To rule out the presence of unsuspected introns, the entire coding sequence of the gene was determined. No polymorphisms were observed between the coding sequence of the gene and the cDNA sequence, even though the genomic clones were derived from a different mouse strain (DBA/2J) than the cDNA clones (B10.D2 and BALB/c).

Estimation of Intron Size. Intron size was estimated by PCR analysis using primers near the intron/exon boundaries. The amplified product sizes were determined by agarose gel electrophoresis (Figure 9). Control reactions in which templates were omitted did not yield PCR products (not shown). Subtraction of the distance between primers in the cDNA yielded net sizes for the introns as follows: intron 1, 2900 bp; intron 2, 2700 bp; and intron 3, 400 bp. The estimates of intron size were confirmed by restriction mapping of the genomic clones using the restriction sites shown in Figure 8 (data not shown). The 500 bp PCR product of intron 1 visible in Figure 9 (lane a) was considered to be an artifact.

Chromosome Localization. On Southern blots, the 5′ HCII probe hybridized with a single band in each of 13 restriction digests used to type NFS/N and *M. m. musculus* and in 9 restriction digests of NFS/N and *M. spretus* (Figure 10 and data not shown). This pattern is consistent with the presence of a single copy of the HCII gene in these strains of mice. Analysis of parental DNAs of the two genetic crosses with the 5′ HCII probe identified *Sst*I fragments of 19.2 kb in NFS/N and C58/J and 15.5 kb in *M. m. musculus*. *Hin*dIII fragments of 10.2 and 11.5 kb were identified in *M. spretus* and in NFS/N and C58/J, respectively (Figure 10). Inheritance of the polymorphic fragments in the two crosses indicated linkage of *Hcf2* to markers on proximal chromosome 16 (Table 1). Combined data from both crosses are consistent with the following gene order and distances in cM: *Prm-1* – 4.9 ± 1.3 – *Hcf2* – 2.0 ± 0.9 – *Igl* – 2.5 ± 2.5 – *Smst* – 25.0 ± 7.7 – *Mtv-37* – 5.4 ± 2.6 – *Pit-1*.

Table 1[a]

	Hcf2	*Igl*	*Smst*	*Mtv-37*	*Pit-1*
Prm-1	13/264 4.9 ± 1.3	15/254 5.9 ± 1.5	3/39 7.7 ± 4.3	22/77 28.6 ± 5.1	80/254 31.5 ± 2.9
Hcf2		5/254 2.0 ± 0.9	1/40 2.5 ± 2.5	23/100 23.0 ± 4.2	68/253 26.9 ± 2.8
Igl			1/40 2.5 ± 2.5	21/78 26.9 ± 5.0	68/242 28.1 ± 2.9
Smst				8/32 25.0 ± 7.7	9/37 24.3 ± 7.1
Mtv-37					4/74 5.4 ± 2.6

[a] Genomic DNA from (NFS/N or C58/J × *M. m. musculus*) × *M. m. musculus* and (NFS/N or C58/J × *M. spretus*) × *M. spretus* backcross mice was typed for *Hcf2* and for chromosome 16 markers by Southern blot analysis of RFLPs, as described. The number of crossovers between markers per number of mice typed for each combination is shown on the top line of each row. The recombination frequency ± standard error (bottom line of each row) was calculated according to Green (1981).

[2] The corresponding sequence in the human gene is encoded by exons 1 and 2 (Herzog et al., 1991). Thus, exons 2–4 of the murine gene correspond to exons 3–5 of the human gene.

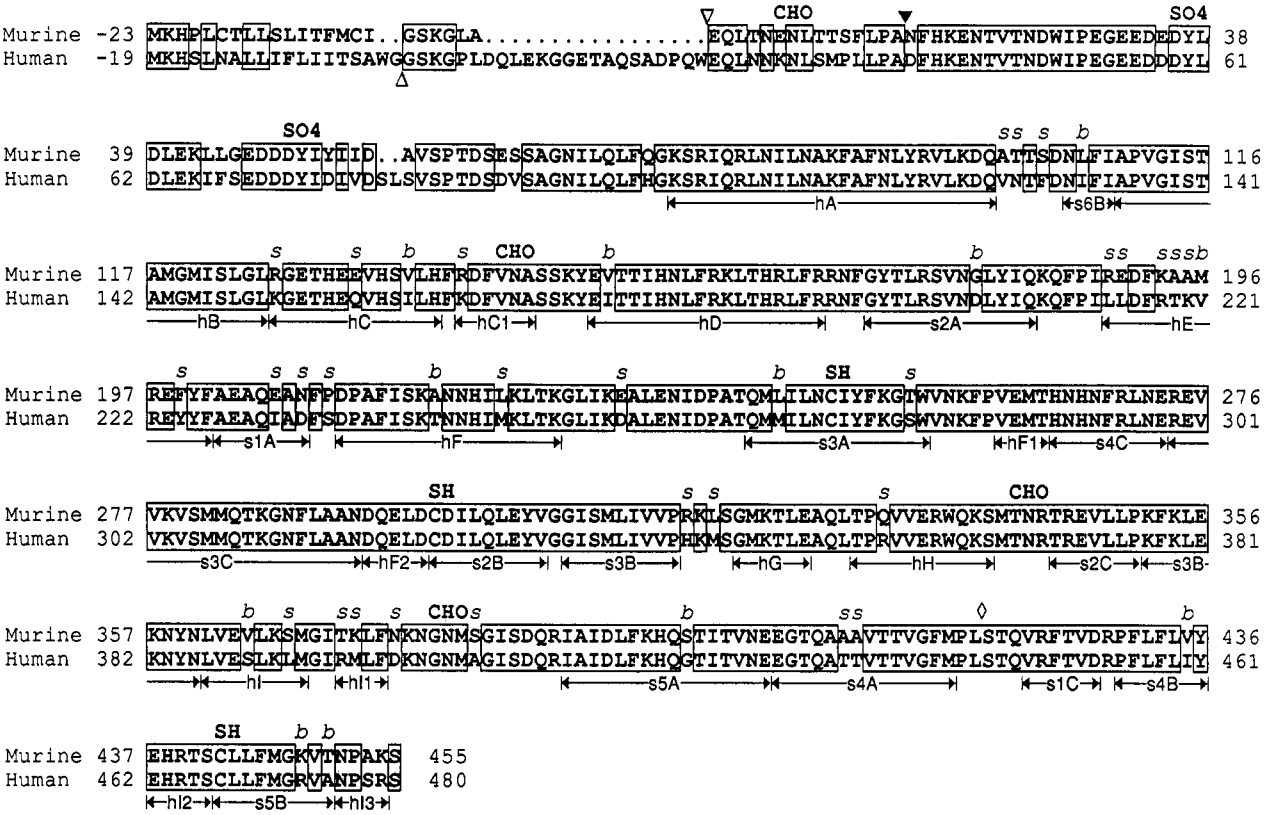

FIGURE 11: Alignment of murine and human HCII amino acid sequences. Identical residues are enclosed in *boxes*. The *dots* indicate gaps introduced to optimize the alignment. The proposed signal peptidase cleavage sites for murine (▽) and human (△) HCII are shown. The reactive site (◊), asparagine-linked glycosylation sites (CHO), tyrosine sulfation sites (SO4), and cysteine residues (SH) of murine HCII are also shown. The cleavage site in the partially degraded form of murine HCII is indicated (▼). Regions corresponding to α-helices (*hA–hI3*) and β-sheets (*s1A–s6B*) in cleaved α_1-antitrypsin (Huber & Carrell, 1989), and nonconserved amino acids corresponding to surface (*s*) or buried (*b*) residues in ovalbumin (Stein et al., 1990), are indicated.

DISCUSSION

HCII accounts for virtually all of the dermatan sulfate-dependent thrombin inhibitor activity (i.e., dermatan sulfate cofactor) in human plasma (Tollefsen et al., 1983). Thus, incubation of human plasma with ^{125}I-thrombin in the presence of dermatan sulfate yields a single protease–inhibitor complex of 104 kDa on SDS–PAGE (Figure 2B). By contrast, both HCII and antithrombin have heparin cofactor activity, and two distinct complexes are formed after incubation of ^{125}I-thrombin with human plasma in the presence of heparin (Tollefsen & Blank, 1981). Murine plasma contains ~0.2 μM dermatan sulfate cofactor activity and ~0.6 μM heparin cofactor activity (Figure 1). Unexpectedly, incubation of ^{125}I-thrombin with murine plasma in the presence of dermatan sulfate yields two complexes of equal intensity (101 and 104 kDa; Figure 2B), and the purified murine dermatan sulfate cofactor consists of two proteins of 68 and 72 kDa (Figure 3). Both proteins bind ^{125}I-thrombin in the presence of dermatan sulfate to yield complexes that comigrate with those that are formed in unfractionated plasma.

The two proteins have identical N-terminal amino acid sequences, and both react with an antibody raised against the C-terminal peptide (nine residues) of murine HCII deduced from the cDNA sequence. Removal of N-linked oligosaccharides from either protein with PNGase F yields a single product with an apparent mass of 54 kDa. Therefore, the 68- and 72-kDa proteins represent two forms of murine HCII that appear to differ in the extent of glycosylation. Two partially degraded forms of murine HCII were observed after storage of the protein at 4 °C (Figure 3). Both degraded forms had the same N-terminal sequence (beginning with Asn^{16} of native murine HCII) and reacted with the C-terminal peptide antibody. Therefore, the partially degraded forms probably arose by proteolytic cleavage of the Ala^{15}–Asn^{16} peptide bonds of the native 68- and 72-kDa proteins. The 8–9-kDa difference in apparent mass between the native and partially degraded forms of murine HCII is probably explained, in part, by glycosylation of Asn^{7} in the deleted N-terminal peptide. Moreover, the inability to identify Asn^{7} by Edman degradation is consistent with glycosylation at this site.

The murine HCII cDNA (2104 bp) appears to be nearly full-length by comparison to the HCII mRNA (2.2–2.3 kb) detected in murine liver by Northern blot analysis. The cDNA includes a 1434 bp open reading frame following the first Met codon (nt 91–93), a TAA stop codon (nt 1525–1527), and 580 bp of 3′-untranslated sequence terminating in a poly(A) tail. A polyadenylation signal (AATAAA) is located 23 nt upstream from the poly(A) sequence. A purine nucleotide is present three residues upstream from the first Met codon, which is compatible with initiation of translation at this site (Kozak, 1986). The cDNA encodes the N-terminal amino acid sequence of purified murine HCII (EQLTNE...) preceded by a 23-residue hydrophobic sequence presumed to be the signal peptide. The proposed signal peptidase cleavage site between Ala^{-1} and Glu^{1} conforms to the "(−3,−1) rule" of von Heijne (1984), in which a small, neutral residue must be present at positions −1 and −3 for efficient cleavage to occur. The mature polypeptide consists of 455 amino acid residues with a calculated mass of 52 003 Da.

An alignment of the deduced amino acid sequences of murine and human HCII is shown in Figure 11. The sequences are collinear except for deletions in murine HCII of 2 residues between Ile^{-7} and Gly^{-6}, 17 residues between Ala^{-1} and Glu^{1}, and 2 residues between Asp^{55} and Ala^{56}. The fact that the

first 23 amino acid residues at the N-terminal end of human HCII are not represented in murine HCII suggests that this portion of the protein is not required for activity. Furthermore, the partially degraded forms of murine HCII lack an additional 15 residues but remain able to bind thrombin in the presence of dermatan sulfate (Figure 2B, lane 4). These observations are consistent with the previous finding that deletion of the N-terminal 52 amino acid residues of recombinant human HCII has no effect on the rate of thrombin inhibition in the presence or absence of a glycosaminoglycan (Van Deerlin & Tollefsen, 1991).

Murine and human HCII are 87% identical in amino acid sequence. Murine HCII contains both of the potential tyrosine sulfation sites (Tyr^{37} and Tyr^{50}) and all three potential asparagine-linked glycosylation sites (Asn^{7}, Asn^{144}, and Asn^{343}) present in human HCII. The murine protein contains an additional glycosylation site at Asn^{379}. The three cysteine residues of human HCII, which do not appear to form intramolecular disulfide bonds (Church et al., 1987), are conserved in murine HCII (Cys^{248}, Cys^{298}, and Cys^{442}). Forty of the amino acids that differ between murine the human HCII can be modeled on the basis of the tertiary structure of ovalbumin (Stein et al., 1990). Twenty-eight of the nonconserved residues appear to occur on the surface of the protein (Figure 11), where amino acid substitutions might affect the interaction of HCII with other macromolecules. Nevertheless, the glycosaminoglycan binding site, the reactive site, and the N-terminal acidic region, which interacts with thrombin, are highly conserved.

Site-directed mutagenesis of human HCII has suggested that the binding sites for dermatan sulfate and heparin overlap but are not identical. Thus, mutations of Lys^{173}, Arg^{184}, and Lys^{185} decrease the affinity of HCII for heparin, whereas mutations of Arg^{184}, Lys^{185}, Arg^{189}, Arg^{192}, and Arg^{193} diminish dermatan sulfate binding (Blinder et al., 1989; Blinder & Tollefsen, 1990; Ragg et al., 1990a; Whinna et al., 1991). All of the basic residues identified in the glycosaminoglycan binding site of human HCII are conserved in the murine protein (Figure 11). Murine and human HCII are activated by similar concentrations of dermatan sulfate [cf. Figure 2A and Tollefsen et al. (1983)] and, therefore, appear to have similar affinities for the glycosaminoglycan.

Proteolytic attack at the reactive site Leu^{444}–Ser^{445} (P1–P1′) peptide bond in human HCII results in formation of the stable complex in which thrombin is inactive (Griffith et al., 1985). Single amino acid substitutions of P1 or other residues near the reactive site of a serpin are often associated with loss of function or altered protease specificity (Huber & Carrell, 1989). Human and murine HCII have identical reactive site sequences between residues P9 and P15′ (Figure 11). In preliminary experiments, we have found that purified murine HCII inhibits human thrombin and human chymotrypsin with apparent second-order rate constants of 9×10^4 and 6×10^5 M^{-1} min^{-1}, respectively, in the absence of a glycosaminoglycan (data not shown). The rate constants are similar to those reported for human HCII (Church et al., 1985; Van Deerlin & Tollefsen, 1991) and, together with the sequence data, suggest that human and murine HCII have identical protease specificities.

The N-terminal region of human HCII from Glu^{53} to Asp^{75} contains a unique cluster of acidic residues that may occupy the glycosaminoglycan binding site in the absence of a glycosaminoglycan (Ragg et al., 1990b; Van Deerlin & Tollefsen, 1991). Binding of a glycosaminoglycan appears to displace the acidic domain, allowing it to interact with anion binding exosite I of thrombin and thereby contribute to the increased rate of inhibition (Phillips et al., 1993; Ragg et al., 1990a; Sheehan et al., 1993; Van Deerlin & Tollefsen, 1991). All of the acidic residues in this region are conserved with the exception of Asp^{75}, which is replaced by Tyr in murine HCII. Other functional regions that have been described in human HCII include a chemotactic peptide (Asp^{39}–Ile^{66}) that is released by leukocyte proteases (Church et al., 1991) and a conserved pentappetide near the C-terminal end (Phe^{456}–Ile^{460}) that may be involved in binding thrombin–HCII complexes to the serpin–enzyme complex receptor of hepatocytes (Joslin et al., 1993). Similar sequences are also present in murine HCII.

The simple pattern of HCII bands detected by Southern blot analysis indicates that there is a single copy of the HCII gene in the murine genome. This finding is in contrast to the murine α_1-protease inhibitor and contrapsin genes, which have been duplicated to form clusters of at least 5 and 10 genes, respectively, termed the *Spi-1* and *Spi-2* loci (Borriello & Krauter, 1991; Inglis & Hill, 1991). Thus, genes closely related to the α_1-protease inhibitor and contrapsin genes have been detected as multiple bands on Southern blots. Although the duplicated sequences in each of these multigene families are highly conserved (e.g., <4% overall divergence for genes at the *Spi-1* locus), they differ markedly in their reactive site coding sequences. The 5′ HCII cDNA probe that we used for the Southern blotting experiments is located outside the region of greatest divergence among the murine serpin genes. If multiple HCII-related genes exist, this probe should hybridize to other members of the gene family. In support of the Southern blot analysis, the murine HCII genomic clones contained sequences identical to the cDNA clones in the reactive site coding region. In addition, we detected a single mRNA species for HCII on Northern blot analysis, whereas multiple transcripts have been detected for genes at the *Spi-1* and *Spi-2* loci (Borriello & Krauter, 1991; Inglis & Hill, 1991). Both the murine and human HCII genes are transcribed in the liver but not in any other tissues examined (Figure 7). Our results are consistent with previous reports that human HCII is expressed in the liver (Ragg, 1986) and in hepatoma-derived Hep G2 cells (Ragg & Preibisch, 1988) with a transcript size of about 2.3 kb.

The murine HCII gene (*Hcf2*) is located on chromosome 16 at a position between *Prm-1* and *Igl*. *Prm-1* is homologous to the human protamine P1 gene (*PRM1*), and *Igl* is homologous to the human immunoglobulin λ light chain gene cluster (*IGL*). These genes map to human chromosomal regions 16p13 (Viguie et al., 1990) and 22q11 (Emanuel et al., 1984), respectively. The human HCII gene (*HCF2*) has been mapped to chromosomal region 22q11 (Herzog et al., 1991), which is consistent with the linkage of murine *Hcf2* and *Igl*. The proximal position of *Hcf2* relative to *Igl* extends the region of murine chromosome 16 with linkage homology to human chromosome 22. Several other human and murine serpin genes have been mapped. The genes for murine α_1-protease inhibitor and contrapsin are tightly linked on chromosome 12 (Hill et al., 1985), and the genes for human α_1-antitrypsin and α_1-antichymotrypsin are tightly linked on chromosome 14 (Rabin et al., 1986). Thus, the HCII gene is not linked to either of these genes.

The three introns identified in the murine HCII gene are located at positions identical to introns in the human HCII gene (Figure 12). Introns are found at similar positions in the human genes for α_1-antitrypsin and angiotensinogen (Strandberg et al., 1988) and in the murine gene for α_1-protease

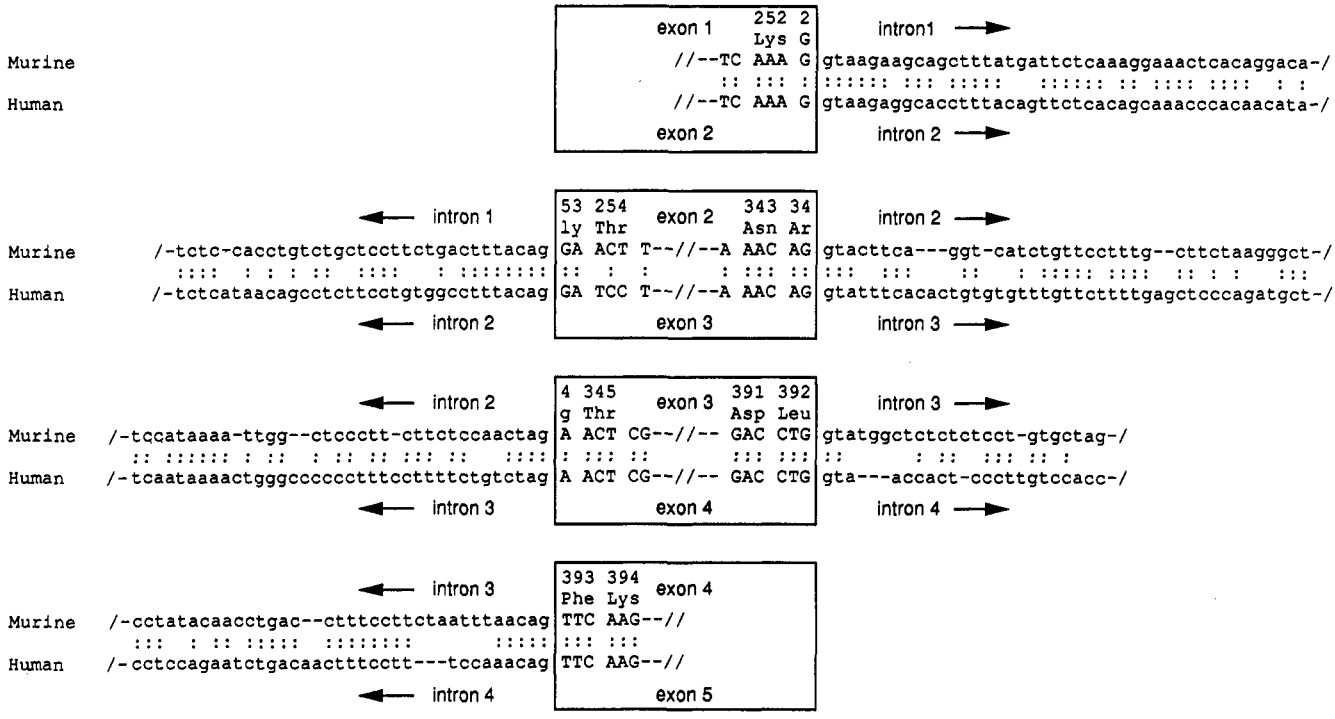

FIGURE 12: Sequences of intron/exon boundaries. The murine intron/exon boundary sequences are compared to the previously published human sequences (Ragg et al., 1990b). Partial intron sequences are shown in *lower case leters*. Partial exon sequences are shown in *upper case letters* and contained within the *center boxes*. The translated murine amino acid sequence is shown above the DNA sequence. Amino acid numbering is according to the cDNA-derived amino acid sequence, where 1 corresponds to the N-terminus of HCII purified from murine plasma. Conserved nucleotides (:), breaks between the beginning and end of the complete intron or exon sequences (–//–), and gaps in the sequence inserted to obtain optimal alignment (–) are indicated.

inhibitor (Krauter et al., 1986). Other homologous members of the serpin family show little similarity to HCII in genomic organization, and the evolutionary relationships among the serpin genes are unclear (Strandberg et al., 1988). All of the intron/exon boundaries in the murine HCII gene obey the GT-AG rule (Breathnach & Chambon, 1981). The human and murine introns have similar sequences near the intron/exon junctions (Figure 12), although the corresponding introns of the two species differ in total length [cf. Herzog et al. (1991)]. The human HCII gene contains a 5151 bp intron beginning 17 nt upstream from the start site for translation. An intron may not be present at the corresponding position of the murine gene, since the murine genomic and cDNA sequences are collinear for 90 nt upstream from the start site for translation (Figure 13). Since the murine transcript detected on Northern analysis (2.2–2.3 kb) is comparable in size to the longest available cDNA (2.1 kb), it is unlikely that the transcribed sequence extends much further in the 5′ direction. The 5′ sequence of the murine gene contains a TATA-like sequence (TATTATTATTTTTAT) at nt –74 to –60 which could serve as an RNA polymerase binding site. Alternatively, the 5′ end of the murine HCII cDNA could represent part of an unspliced intron, since nt 1–74 of the cDNA appear to be homologous to the 3′ end of intron 1 of the human HCII gene (Figure 13).

In summary, our data indicate that HCII is more highly conserved than some of the other human and murine serpins. Whereas HCII is 87% identical in amino acid sequence in the two species (Figure 11), human α_1-antitrypsin and its murine homologue (α_1-protease inhibitor) are only ~60% identical (Borriello & Krauter, 1991). Similarly, only 1 of the 10 amino acid residues between P5 and P5′ of the reactive site of human α_1-antichymotrypsin is conserved in its most closely related murine counterpart (contrapsin), and the 2 serpins have different protease specificities (Inglis & Hill, 1991). This

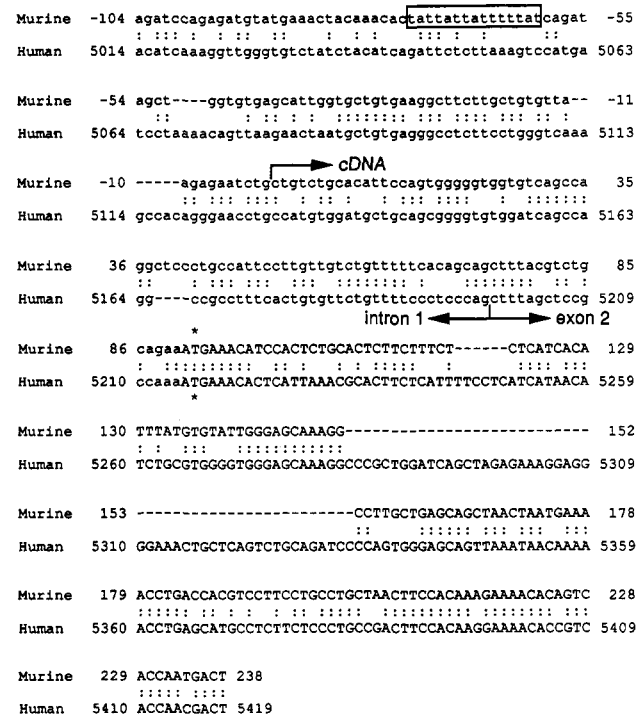

FIGURE 13: Comparison of murine and human 5′-genomic sequences. Murine and human genomic sequences were aligned using Genetics Computer Group software as described. Nucleotide +1 (indicated by an *arrow* in the murine sequence) corresponds, in both cases, to the 5′ end of the available cDNA sequence. Translated sequences are shown in *upper case letters*. The start site for translation is indicated by an *asterisk* (*). Untranslated sequences are shown in *lower case letters*. The intron 1/exon 2 boundary of the human genomic sequence is indicated by *arrows*. Nucleotides conserved between murine and human genes (:) and gaps in the sequence inserted to obtain optimal alignment (–) are indicated. The TATA-like sequence in the murine gene is enclosed by a *box*.

finding implies that α_1-antichymotrypsin and contrapsin have different physiologic functions. Hill and Hastie (1987) proposed that the selective force for duplication and reactive site hypermutation of serpin genes in the mouse may have been exposure to proteases elaborated by various parasites. Our observations that a single gene encodes both molecular weight forms of murine HCII and that murine and human HCII are highly conserved in regions involved in binding to other macromolecules suggest that HCII has the same physiologic function in both species. Experiments *in vitro* indicate that the thrombin–HCII reaction is stimulated by dermatan sulfate proteoglycans synthesized by fibroblasts or vascular smooth muscle cells (McGuire & Tollefsen, 1987). When the vascular endothelium is disrupted, HCII may bind to dermatan sulfate in the subendothelium and regulate activities of thrombin, such as mitogenesis and chemotaxis, that are related to wound healing or inflammation [reviewed in Coughlin et al. (1992)]. The consequences of HCII deficiency, however, may not be known until gene deletion experiments have been performed in the mouse.

ACKNOWLEDGMENT

We thank Dave McCourt, Howard Hughes Medical Institute, Washington University, for performing the amino acid sequence determinations and Jayne Marasa for her help with the recombinant DNA work. We also thank Rick Wetsel, Department of Pediatrics, Washington University, for providing his murine liver cDNA library.

REFERENCES

Adamson, M. C., Silver, J., & Kozak, C. A. (1991) *Virology 183*, 778–781.
Blinder, M. A., & Tollefsen, D. M. (1990) *J. Biol. Chem. 265*, 286–291.
Blinder, M. A., Marasa, J. C., Reynolds, C. H., Deaven, L. L., & Tollefsen, D. M. (1988) *Biochemistry 27*, 752–759.
Blinder, M. A., Andersson, T. R., Abildgaard, U., & Tollefsen, D. M. (1989) *J. Biol. Chem. 264*, 5128–5133.
Borriello, F., & Krauter, K. S. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9417–9421.
Breathnach, R., & Chambon, P. (1981) *Annu. Rev. Biochem. 50*, 349–383.
Church, F. C., Noyes, C. M., & Griffith, M. J. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6431–6434.
Church, F. C., Meade, J. B., & Pratt, C. W. (1987) *Arch. Biochem. Biophys. 259*, 331–340.
Church, F. C., Pratt, C. W., & Hoffman, M. (1991) *J. Biol. Chem. 266*, 704–709.
Coughlin, S. R., Vu, T.-K. H., Hung, D. T., & Wheaton, V. I. (1992) *J. Clin. Invest. 89*, 351–355.
Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.
Edelhoch, H. (1967) *Biochemistry 6*, 1948–1954.
Emanuel, B. S., Cannizzaro, L. A., Tsujimoto, Y., & Croce, C. M. (1984) *Am. J. Hum. Genet. 36*, 202.
Garber, R. L., Kuroiwa, A., & Gehring, W. J. (1983) *EMBO J. 2*, 2027–2036.
Green, E. L. (1981) *Genetics and Probability in Animal Breeding Experiments*, Oxford University Press, New York.
Griffith, M. J., Noyes, C. M., Tyndall, J. A., & Church, F. C. (1985) *Biochemistry 24*, 6777–6782.
Harlow, E., & Lane, D. (1988) *Antibodies: a Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Hattori, M., & Sakaki, Y. (1986) *Anal. Biochem. 152*, 232–238.
Herzog, R., Lutz, S., Blin, N., Marasa, J. C., Blinder, M. A., & Tollefsen, D. M. (1991) *Biochemistry 30*, 1350–1357.
Hill, R. E., & Hastie, N. D. (1987) *Nature 326*, 96–99.
Hill, R. E., Shaw, P. H., Barth, R. K., & Hastie, N. D. (1985) *Mol. Cell. Biol. 5*, 2114–2122.
Huber, R., & Carrell, R. W. (1989) *Biochemistry 28*, 8951–8966.
Inglis, J. D., & Hill, R. E. (1991) *EMBO J. 10*, 255–261.
Joslin, G., Wittwer, A., Adams, S., Tollefsen, D. M., August, A., & Perlmutter, D. M. (1993) *J. Biol. Chem. 268*, 1886–1893.
Kleene, K. C., Distel, R. J., & Hecht, N. B. (1983) *Dev. Biol. 98*, 455–464.
Kozak, M. (1986) *Cell 44*, 283–292.
Kozak, C. A., Peyser, M., Krall, M., Mariano, T. M., Kumar, C. S., Pestka, S., & Mock, B. A. (1990) *Genomics 8*, 519–524.
Krauter, K. S., Citron, B. A., Hsu, M. T., Powell, D., & Darnell, J. E., Jr. (1986) *DNA 5*, 29–36.
Laemmli, U. K. (1970) *Nature 227*, 680–685.
Lamballe, F., Klein, R., & Barbacid, M. (1991) *Cell 66*, 967–979.
Majors, J. E., & Varmus, H. E. (1981) *Nature 289*, 253–258.
Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035–10038.
McGuire, E. A., & Tollefsen, D. M. (1987) *J. Biol. Chem. 262*, 169–175.
Pari, G., Jardine, K., & McBurney, M. W. (1991) *Mol. Cell. Biol. 11*, 4796–4803.
Phillips, J. E., Shirk, R. A., Whinna, H. C., Henriksen, R. A., & Church, F. C. (1993) *J. Biol. Chem. 268*, 3321–3327.
Pratt, C. W., Church, F. C., & Pizzo, S. V. (1988) *Arch. Biochem. Biophys. 262*, 111–117.
Pratt, C. W., Whinna, H. C., Meade, J. B., Treanor, R. E., & Church, F. C. (1989) *Ann. N.Y. Acad. Sci. 556*, 104–115.
Rabin, M., Watson, M., Kidd, V., Woo, S. L., Breg, W. R., & Ruddle, F. H. (1986) *Somatic Cell Mol. Genet. 12*, 209–214.
Ragg, H. (1986) *Nucleic Acids Res. 14*, 1073–1088.
Ragg, H., & Preibisch, G. (1988) *J. Biol. Chem. 263*, 12129–12134.
Ragg, H., Ulshöfer, T., & Gerewitz, J. (1990a) *J. Biol. Chem. 265*, 22386–22391.
Ragg, H., Ulshöfer, T., & Gerewitz, J. (1990b) *J. Biol. Chem. 265*, 5211–5218.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: a Laboratyr Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Scott, C. L., Mushinski, J. F., Huppi, K., Weigert, M., & Potter, M. (1982) *Nature 300*, 757–760.
Sheehan, J. P., Wu, Q., Tollefsen, D. M., & Sadler, J. E. (1993) *J. Biol. Chem. 268*, 3639–3645.
Siracusa, L. D., Jenkins, N. A., & Copeland, N. G. (1991) *Genetics 127*, 169–179.
Stein, P. E., Leslie, A. G. W., Finch, J. T., Turnell, W. G., McLaughlin, P. J., & Carrell, R. W. (1990) *Nature 347*, 99–102.
Strandberg, L., Lawrence, D., & Ny, T. (1988) *Eur. J. Biochem. 176*, 609–616.
Teien, A. N., Abildgaard, U., & Höök, M. (1976) *Thromb. Res. 8*, 859–867.
Tollefsen, D. M. (1992) in *Heparin and Related Polysaccharides* (Lane, D. A., Björk, I., Lindahl, U., Eds.) pp 167–176, Plenum Press, New York.
Tollefsen, D. M., & Blank, M. K. (1981) *J. Clin. Invest. 68*, 589–596.
Tollefsen, D. M., Pestka, C. A., & Monafo, W. J. (1983) *J. Biol. Chem. 258*, 6713–6716.
Van Deerlin, V. M. D., & Tollefsen, D. M. (1991) *J. Biol. Chem. 266*, 20223–20231.
Van Deerlin, V. M. D., & Tollefsen, D. M. (1992) *Semin. Thromb. Hemostasis 18*, 341–346.
Viguie, F., Domenjoud, L., Rousseau-Merck, M.-F., Dadoune, J.-P., & Chevaillier, P. (1990) *Hum. Genet. 85*, 171–174.
von Heijne, G. (1984) *J. Mol. Biol. 173*, 243–251.
Whinna, H. C., Blinder, M. A., Szewczyk, M., Tollefsen, D. M., & Church, F. C. (1991) *J. Biol. Chem. 266*, 8129–8135.