

Isolation and characterization of cDNAs encoding the heavy chain of human inter- α -trypsin inhibitor (I α TI): Unambiguous evidence for multipolypeptide chain structure of I α TI

(α_1 -microglobulin/calcium-binding domain/thiol-proteinase inhibitor)

J. P. SALIER*[†], M. DIARRA-MEHRPOUR*, R. SESBOUE*, J. BOURGUIGNON*, R. BENAROUS[‡], I. OHKUBO[§], S. KURACHI^{||}, K. KURACHI[§], AND J. P. MARTIN*

*Institut National de la Santé et de la Recherche Médicale U-295, U.E.R. de Médecine-Pharmacie, BP 97, F-76800 Saint-Etienne-du-Rouvray, France; [†]Institut National de la Santé et de la Recherche Médicale, CHU Cochin-Port-Royal, F-75014 Paris, France; [‡]Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109; and [§]Howard Hughes Medical Center, University of Michigan, Ann Arbor, MI 48109

Communicated by James V. Neel, August 5, 1987

ABSTRACT Human inter- α -trypsin inhibitor (I α TI) is a plasma glycoprotein of M_r 180,000, which has been described as a single polypeptide chain. Recently, however, we proposed that I α TI might be composed of a heavy (H) chain (M_r = 95,000) and a light (L) chain (M_r = 40,000) synthesized by two separate mRNAs. In the present study we have characterized cDNAs for the H chain of I α TI. These cDNAs collectively covered two sequences (977 and 1450 base pairs in length) with single open reading frames. The deduced amino acid sequences were highly homologous to each other and well matched with partial amino acid sequences obtained from purified serum I α TI. RNA blot analyses of liver RNAs with H- or L-chain cDNAs as probes clearly identified two distinct mRNAs of 3.3 and 1.3 kilobases, which corresponded to H or L chain, respectively. Poly(A)⁺ RNAs hybrid-selected with H-chain cDNAs coded for polypeptide chains of M_r 90,000-95,000. These results unambiguously establish that I α TI is made of multipolypeptides, possibly including one H and two L chains. The H chain contains potential calcium-binding sites and also regions homologous to the proposed reactive site for thiol-proteinase inhibitors. These data indicate that I α TI is a complex, multifunctional protein. mRNAs for both the H and L chains were found only in liver.

Inter- α -trypsin inhibitor (I α TI) is a serine protease inhibitor of M_r 180,000 present in human plasma. Polypeptides of lower M_r that immunologically cross-react with I α TI also have inhibitory activity and are present in normal plasma, urine, and bronchial mucus. These are called I α TI derivatives (1). Since I α TI shows a M_r of 180,000 on NaDodSO₄/PAGE, even in the presence of reducing agents, it has long been thought to be a single polypeptide chain. Recently, however, we have reported data to suggest that I α TI is synthesized by two separate mRNAs that code for the heavy (H) chain and the light (L) chain, respectively (2). The L chain gives rise to the I α TI derivatives (2). More recently, cDNAs for the L chain were isolated, and the L chain was found to contain amino acid sequences for two proteins: α_1 -microglobulin (α_1 m), an acute phase reactant, and a M_r 30,000 derivative of human I α TI (HI-30) (3, 4). HI-30 contains two homologous domains with an inhibitory site for elastase as well as an inhibitory site for trypsin and chymotrypsin (5, 6). The amino acid sequence of the H chain has not been determined, and no biological function has been reported for it.

In this paper, we report the isolation and nucleotide sequencing of partial cDNAs for the H chain.^{||} The H chain contains potential calcium-binding sites as well as sequences highly

homologous to those at the reactive site of thiol-proteinase inhibitors. These data clearly indicate that I α TI is a multifunctional protein comprised of multipolypeptide chains that are synthesized in liver by two distinct mRNA species.

MATERIALS AND METHODS

Reagents. Various rabbit anti-human I α TI antisera were prepared as described (7). These included anti-I α TI antiserum containing both anti-H and anti-L chain antibodies, anti-L chain antiserum and anti-H chain antiserum (2). Biotinylated anti-rabbit IgG antibodies and avidin:biotinylated horseradish peroxidase complex were from Vector Laboratories (Burlingame, CA). Various baboon tissues were kindly provided to K. Kurachi by the Regional Primate Center (University of Washington, Seattle). T4 DNA polymerase and its Klenow fragment, dNTP, and ddNTP were from Boehringer Mannheim. Restriction enzymes and M13mp18 and M13mp19 phage DNA were from Bethesda Research Laboratories. *Sma* I-cut, dephosphorylated M13mp8 DNA, M13 sequencing primer, deoxyadenosine 5'-[α -³⁵S]thiotriphosphate (400 Ci/mmol; 1 Ci = 37 GBq), and [³²P]dCTP (3000 Ci/mmol) were from Amersham. GeneScreenPlus nylon membranes were from New England Nuclear. All other chemicals were of analytical grade.

Screening and Characterization of Human Liver cDNAs. About 1.6×10^5 recombinant phage plaques of a normal human liver cDNA library constructed in λ gt11 (provided by R. A. Lazzarini, National Institutes of Health) were screened with an anti-I α TI antiserum by a modification of the Young and Davis procedure (8). Positive phage clones were plaque-purified, and liquid phage stocks were prepared according to standard protocols (9). Then positive clones were immunoscreened for the expression of either the L or the H chain of I α TI. These clones were also screened with L-chain cDNA, designated pHuLIT11 (3), as a probe. Radiolabeled cDNA probes were prepared by nick-translation with [³²P]dCTP.

Hybrid Selection of RNAs and *in Vitro* Translation. *In vitro* translation of hybrid-selected poly(A)⁺ RNAs of human liver and immunoprecipitation of the translation products were performed as described (2, 3).

Nucleotide Sequencing. cDNA inserts were subcloned into M13mp18 or M13mp19. Alternatively, the entire DNA of a recombinant λ gt11 clone was sonicated, and fragments of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: α_1 m, α_1 -microglobulin; I α TI, inter- α -trypsin inhibitor; HI-30, a M_r 30,000 derivative of human I α TI; H, heavy; L, light.
[†]To whom requests should be addressed.

^{||}The sequences of the two domains reported in this paper are being deposited in the EMBL/GenBank data base (Bolt, Beranek, and Newman Laboratories, Cambridge, MA, and Eur. Mol. Biol. Lab., Heidelberg) (accession nos. J03012 and J03013).

500–900 base pairs in length were ligated to M13mp8 at the *Sma*I site according to a standard “shotgun” technique (10). DNA sequencing was performed by the dideoxy chain termination method with deoxyadenosine 5′-[α -³⁵S]thiotriphosphate and analyzed by polyacrylamide gel electrophoresis using a buffer gradient gel (11).

RNA Blot Analysis. RNA blot analysis of human liver poly(A)⁺ RNAs with an I α TI L-chain cDNA probe has been described (3). Total RNAs were also prepared from various tissues of an adult male baboon by the guanidinium isothiocyanate procedure followed by centrifugation through a 5.7 M CsCl cushion (12). RNA preparations were electrophoresed in 1.5% agarose gels containing 6.7% formaldehyde in 20 mM phosphate buffer (pH 7.0). The agarose gels were then blotted onto a GeneScreenPlus membrane. Prehybridization, hybridization with radiolabeled cDNA probe(s), and washing were carried out as recommended for GeneScreenPlus by the manufacturer.

Amino Acid Sequences of I α TI and Its Peptides. Highly purified serum I α TI (13) was subjected to amino acid sequencing to determine its N-terminal end(s) at the protein sequencing facility of the University of Michigan. Recoveries for the first seven residues varied between 19% and 47%. In K. Hochstrasser’s laboratory (Munich, F.R.G.), a set of I α TI peptides was also obtained by digestion of reduced and carboxymethylated I α TI with trypsin or *Staphylococcus aureus* V8 proteinase. Amino acid sequences were determined manually or by Edman degradation in a solid-phase sequencer. A library of peptide sequences for I α TI was kindly provided to us by K. Hochstrasser.

RESULTS

Human cDNAs for I α TI. In the initial screening of a normal human liver cDNA library for I α TI cDNA, 27 positive phage clones were identified with anti-I α TI antiserum and were plaque purified. Ten positive clones were detected with the anti-L chain antiserum, and 7 positive clones were identified with anti-H chain antiserum. None of these clones were detected with both anti-L chain and anti-H chain antisera. The radiolabeled L-chain cDNA probe (pHuLITI1) also hybridized to all clones that were detected with the anti-L chain antiserum (results not shown). Among the clones detected with the anti-H chain antiserum, several (designated λ HuHITI-9, -19, -33) were analyzed in further detail.

Nucleotide Sequence of cDNAs. Sequencing strategies for clones 9, 19, and 33 are shown in Fig. 1. The nucleotide

sequence of clone 19 and the combined sequence of clones 9 and 33 are 977 and 1450 bases long, respectively, and contain a single open reading frame (Fig. 2). The amino acid sequences derived from the nucleotide sequences are in excellent agreement with the amino acid sequences obtained from various peptides derived from I α TI. So far the latter have confirmed about 60% of the amino acid sequence derived from the cDNAs (Fig. 2). No homology has been found between the amino acid sequences of the L (4) and H chains. A computer search of a protein data bank with the GenePro program (Riverside Scientific, Seattle) did not provide any evidence of significant homology to any other reported proteins.

Two highly homologous domains (domains 1 and 2) were found in the partial amino acid sequences of the H chain (Fig. 2). Their relative positions in the H chain are not known at the present time. The sequence Val-Val-Ala-Gly-Lys in domain 2 of the H chain (amino acids 145–149) was found to be highly homologous to the reactive site of human α -2-thiol-proteinase inhibitor (14). A similar sequence is also found in domain 1. A preliminary experiment to determine thiol-proteinase inhibitor activity of I α TI showed that I α TI apparently does not inhibit papain, whereas human calpain I was weakly inhibited by I α TI. As shown in Fig. 3, sequences that are highly homologous to a well-defined calcium-binding site sequence (15) were also found in both domains of the H chain.

Analysis of Poly(A)⁺ RNAs Hybrid Selected with cDNAs of λ HuHITI-9 or -19. Immunoprecipitation of the *in vitro* translation products of human liver poly(A)⁺ RNAs with anti-I α TI antiserum have yielded two different polypeptide chains with *M_r* values of 90,000–95,000 and 40,000, which correspond to the H and L chains of I α TI, respectively (2). In the present study, hybrid-selection experiments also showed that the DNA of clone 19 or 9 specifically retained mRNA fractions that encoded polypeptide chains of *M_r* 90,000–95,000. These polypeptides were immunoprecipitated by an anti-I α TI antiserum, whereas no L chain was concurrently detected (results not shown).

RNA Blot Analysis of Human Liver Poly(A)⁺ RNAs. The radiolabeled cDNA insert of λ HuHITI-9 or -19 was used as a hybridization probe to analyze liver poly(A)⁺ RNAs. With either probe, a single population of mRNAs of about 3.3 kilobases (kb) was found (Fig. 4), whereas a probe containing the 3′ half of the L-chain cDNA detected mRNA species of 1.3 kb, which is in good agreement with previous data (3). The latter result was also obtained with a probe derived from the

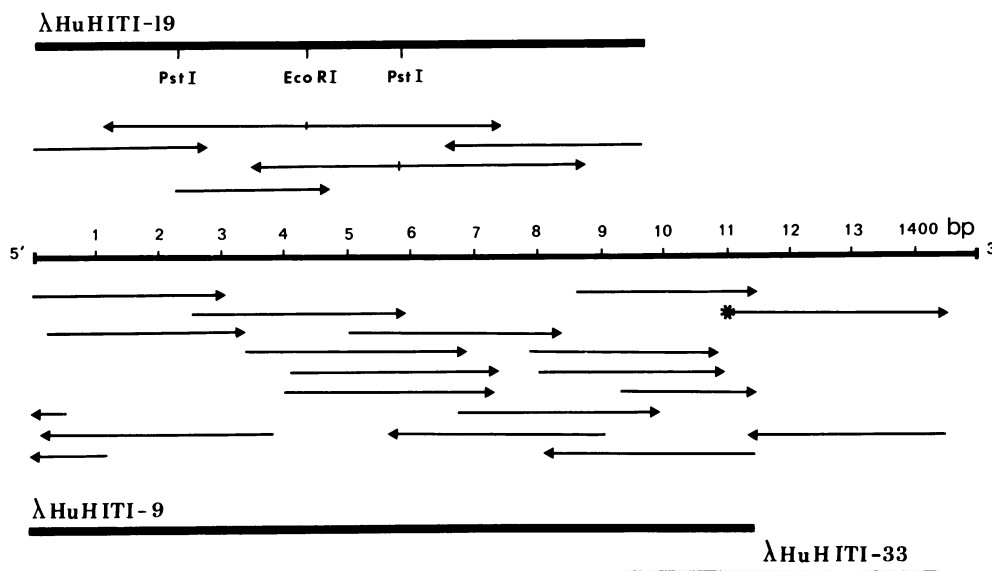


FIG. 1. Strategy for nucleotide sequencing of cDNAs λ HuHITI-19 (Upper) and λ HuHITI-9 and -33 (Lower). The arrows indicate the direction and extent of sequences determined for each cDNA fragment. The sequence with an asterisk at its 5′ end was primed with a specific, synthetic oligonucleotide. Most M13 subclones of clone λ HuHITI-9 were obtained by a shotgun technique.

	GT GGA GGC ACA AAC ATC AAC GAA GCA CTC CTA CGG GCA ATC TTC	44
	<u>Gly Gly Thr Asn Ile Asn Glu Ala Leu Arg Ala Ile Phe</u>	14
	ATT TTG AAT GAA GCC AAT AAC TTG GGA CTG TTA GAC CCC AAC TCC	89
	<u>Ile Leu Asn Glu Ala Asn Asn Leu Gly Leu Leu Asp Pro Asn Ser</u>	29
	CTC TCG CTG ATC ATT TTG GTT TCT GAT GGA GAT CCA ACA GTG GGC	134
	<u>Val Ser Leu Ile Ile Leu Val Ser Asp Gly Asp Pro Thr Val Gly</u>	44
	GAA CTA AAA CTG TCA AAA ATT CAG AAA AAC GTT AAG CAG AAC ATC	179
	<u>Glu Leu Lys Leu Ser Lys Ile Gln Lys Asn Val Lys Glu Asn Ile</u>	59
	CAA GAC AAT ATC TCC TTG TTC AGT TTG GGC ATG GGA TTT GAT GTG	224
	<u>Gln Asp Asn Ile Ser Leu Phe Ser Leu Gly Met Gly Phe Asp Val</u>	74
	GAC TAT GAT TTT TTG AAG AGA CTG TCC AAT GAC CAC AAT CAG ATT	269
	<u>Asp Tyr Asp Phe Leu Lys Arg Leu Ser Asn Glu Asn His Gly Ile</u>	89
	GCA CAA AGG ATT TAT GGA AAC CAG GAC ACG TCT TCC CAG CTT AAG	314
	<u>Ala Gln Arg Ile Tyr Gly Asn Gln Asp Thr Ser Ser Gln Leu Lys</u>	104
	AAA TTC TAC AAC CAG GTC TCC ACT CCA TTG CTC CGG AAT GTT CAG	359
	<u>Lys Phe Tyr Asn Gln Val Ser Thr Pro Leu Leu Arg Asn Val Gln</u>	119
	TTG AAC TAT CCC CAT ACA TCA GTC ACG GAC GTC ACT CAA AAC AAT	404
	<u>Phe Asn Tyr Pro His Thr Ser Val Thr Asp Val Thr Gln Asn Asn</u>	134
	TTT CAT AAC TAC TTT TTT GGA GGC TCA GAG ATT GTG GTG CCA GGA	449
	<u>Phe His Asn Tyr Phe Gly Gly Ser Glu Ile Val Val Ala Gly Lys</u>	149
	TTT GAC CCT GCT AAA TTG GAT CAA ATA GAG AGC GTT ATC ACG CGG	494
	<u>Phe Asp Pro Ala Lys Leu Asp Gln Ile Glu Ser Val Ile Thr Ala</u>	164
	ACT TCG GCT AAC ACG CAG TTA GTC TTG GAG ACC CTG GCG CAG ATG	539
	<u>Thr Ser Ala Asn Thr Gln Leu Val Leu Glu Thr Leu Ala Gln Met</u>	179
	GAC GAC TTG CAG GAT TTT CTA TCG AAA GAC AAG CAT GCA GAT CCC	584
	<u>Asp Asp Leu Gln Asp Phe Leu Ser Lys Asp Lys His Ala Asp Pro</u>	194
	ATT TTC ACC AGG AAA CTG TGG GCC TAT CTA ACC ATC AAC CAA CTG	629
	<u>Asp Phe Thr Arg Lys Leu Trp Ala Tyr Leu Thr Ile Asn Gln Leu</u>	209
	CTA GCT GAA CGA AGC CTG GCT CCT ACA GCT GCC GCC AAG AGA AGA	674
	<u>Leu Ala Glu Arg Ser Leu Ala Pro Thr Ala Ala Ala Lys Arg Arg</u>	224
	ATT ACA AGA TCG ATC CTG CAG ATG TCT CTA GAC CAC CAC ATT GTG	719
	<u>Ile Thr Arg Ser Ile Leu Gln Met Ser Leu Asp His His Ile Val</u>	239
	ACT CCG CTG ACC TCG CTG GTG ATC GAG AAC GAG GCT GGG GAT GAG	764
	<u>Thr Pro Leu Thr Ser Leu Val Ile Glu Asn Glu Ala Gly Asp Thr</u>	254
	CGG ATG CTG GCG GAT GCC CCA CCG CAG GAT CCC TCC TGC TCA	809
	<u>Arg Met Leu Ala Asp Ala Pro Pro Gln Asp Pro Ser Cys Cys Ser</u>	269
	GGG GCC CTG TAT TAC GGC AGC AAA GTG GTT CCA GAT TCC ACC CCG	854
	<u>Gly Ala Leu Tyr Tyr Gly Ser Lys Val Val Pro Asp Ser Thr Pro</u>	284
	TCT TGG GCC AAT CCT TCA GCA ACG CCC CTG ATC TCC ATG CTG GCA	899
	<u>Ser Trp Ala Asn Pro Ser Ala Thr Pro Val Ile Ser Met Leu Ala</u>	299
	CAA GGA TCT CAG GTG CTA GAG TCC ACG CCA CCC CCA CAT GTG ATG	944
	<u>Gln Gly Ser Gln Val Leu Glu Ser Thr Pro Pro Pro His Val Met</u>	314
	AGA GTT GAA AAT GAC CCA CAT TCC ATC ATT TAT CTA CAA AAA AGC	989
	<u>Arg Val Glu Asn Asp Pro His Ser Ile Ile Tyr Leu Pro Lys Ser</u>	329
	CAA AAG AAC ATT TGT TTC AAT ATT GAC TCA GAA CCT GGA AAA ATC	1034
	<u>Gln Lys Asn Ile Cys Phe Asn Ile Asp Ser Glu Val Pro Gly Lys Ile</u>	344
	CTC GAC CTG GCT TCT GAC CCA GAA TCA GGA ATT GTA GTG AAC GGT	1079
	<u>Leu Asp Leu Ala Ser Asp Pro Glu Ser Gly Ile Val Val Asn Gly</u>	359
	CAG CTT GTT GGT GCC AAG AAG CCC AAC AAT GGA AAA CTA AGC ACC	1124
	<u>Gln Leu Val Gly Ala Lys Lys Pro Asn Asn Gly Lys Leu Ser Thr</u>	374
	TAT TTT GGA AAA CTG GGA TTT TAT TTC CAA AGT GAA GAC ATA AAA	1169
	<u>Tyr Phe Gly Lys Leu Gly Phe Tyr Phe Gln Ser Glu Asp Ile Lys</u>	389
	ATA GAA ATC AGC ACT GAG ACC ATC ACC CTG AGC CAT GGT TCT AGC	1214
	<u>Ile Glu Ile Ser Thr Glu Thr Ile Thr Leu Ser His Gly Ser Ser</u>	404
	ACA TTC TCC TTG TCC TGG TCC GAC ACG GCT CAA GTC ACG AAT CAG	1259
	<u>Thr Phe Ser Leu Ser Trp Ser Asp Thr Ala Gln Val Thr Asn Gln</u>	419
	AGG GTG CAG ATC TCA GTG AAG AAA GAA AAA GTG GTA ACT ATC ACC	1304
	<u>Arg Val Gln Ile Ser Val Lys Lys Glu Lys Val Val Thr Ile Thr</u>	434
	CTG GAT AAA GAG ATG TCC TTT TCT GTT TTA CTT CAT CGT GTT TGG	1349
	<u>Leu Asp Lys Glu Met Ser Phe Ser Val Leu Leu His Arg Val Trp</u>	449
	AAG AAG CAT CCC GTC AAT GTT GAC TTT CTG GGA ATC TAC ATA CCC	1394
	<u>Lys Lys His Pro Val Asn Val Asp Phe Leu Gly Ile Tyr Ile Pro</u>	464
	CCT ACA AAC AAG TTC TCA CCT AAA GCC CAC GGA CTA ATA GGC CAG	1439
	<u>Pro Thr Asn Lys Phe Ser Pro Lys Ala His Gly Leu Ile Gly Gln</u>	479
	TTT ATG CAG GA	1450
	<u>Phe Met Gln</u>	482

FIG. 2. Nucleotide and amino acid sequences of H-chain cDNAs. (Left) λ HuHITI-19 (domain 1). (Right) λ HuHITI-9 and -33 (domain 2). Underlining indicates the amino acid sequences that match with those obtained from $I\alpha$ TI polypeptides (K. Hochstrasser, personal communication). Boxed areas indicate sequences that are highly homologous. ●, Potential reactive sites as thiol-proteinase inhibitors; ▲, Ca^{2+} -binding sites; ◆, potential sites for asparagine-linked carbohydrates.

5' half of L-chain cDNA, which contains the sequence for α_{1m} (see below).

Tissue Distribution of Expression for $I\alpha$ TI Genes. In RNA blot analyses, three cDNA probes were used. One is the insert from λ HuHITI-9, which detected the H chain. The other two probes of 591 and 501 base pairs were prepared from cDNA for L chain by *Ava* I digestion and correspond to the regions containing α_{1m} and HI-30 domains, respectively (J.B., M.D.-M., R.S., J.P.M., and J.P.S., unpublished data). Of 13 different tissues obtained from a baboon, only liver was

found to express H or L chain (Fig. 5). Low level expression was also found in liver of a human fetus at the 18th week of gestation but was not detected in two human hepatoma cell lines, an umbilical cord endothelium cell line, or a B-lymphoblastoid cell line.

DISCUSSION

In the present study, we have isolated and sequenced cDNA clones coding for $I\alpha$ TI. Amino acid sequences derived from cDNA clones λ HuHITI-9, -19, and -33, which were among

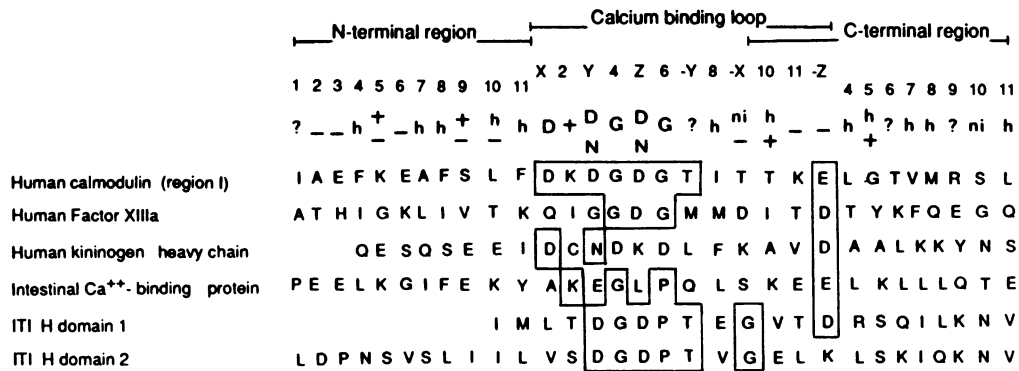


FIG. 3. Comparison of the potential Ca^{2+} -binding site sequences in domains 1 and 2 of the H chain with those found in other proteins. X, Y, Z, -Y, -X, and -Z refer to the vertices of the Ca^{2+} coordinating octahedron. Hydrophobic (h), negative (-), positive (+), nonionic (ni), and weakly conserved (?) residues are indicated. The one-letter code for amino acids is used. The model of Ca^{2+} -binding site is obtained from ref. 15. The amino acid sequences of the first four proteins are from refs. 16-19, respectively. Homologies between proteins or with the model are boxed.

those detected with an anti-H chain antiserum, unambiguously matched the partial amino acid sequences obtained from various peptides of $\text{I}\alpha\text{TI}$. In RNA blot analyses, a mRNA band of 3.3 kb was detected when the cDNA insert of $\lambda\text{HuHITI-9}$ or -19 was used as a probe. The size of this mRNA is in good agreement with the translation product (M_r 90,000-95,000) identified with an anti- $\text{I}\alpha\text{TI}$ antiserum in a cell-free translation assay employing mRNAs isolated in a hybrid-select experiment using an H-chain cDNA obtained from either $\lambda\text{HuHITI-9}$ or -19.

In a previous paper (2), we proposed that $\text{I}\alpha\text{TI}$ may be composed of H- and L-polypeptide chains, which are synthesized by separate mRNAs in liver, based on the data obtained by immunoprecipitation of translation products by anti- $\text{I}\alpha\text{TI}$ antisera. Isolation and characterization of cDNAs for the H chain in the present study, in addition to the data reported for L-chain cDNA (3, 4), provide unambiguous evidence that $\text{I}\alpha\text{TI}$ is composed of H and L chains, which are synthesized separately in liver tissue. This is also in good agreement with the reported observation that $\alpha_1\text{m}$, which is contained in the L chain, is an acute phase protein (20), whereas the H chain is not (1). Amino acid sequence analysis of highly purified serum $\text{I}\alpha\text{TI}$ in the present study gave two major N-terminal sequences. One sequence was Ser-Leu-Pro-Gly-Glu, which may correspond to the N-terminus of the H chain. The other sequence was Ala-Val-Leu-Pro-Gln-Glu, which corresponds to the N-terminal sequence of HI-30 (6). This clearly indicates that L chain of isolated $\text{I}\alpha\text{TI}$ has been

cleaved between $\alpha_1\text{m}$ and HI-30 by an unidentified protease, which results in the loss of $\alpha_1\text{m}$ moieties from $\text{I}\alpha\text{TI}$. This is further supported by the fact that $\text{I}\alpha\text{TI}$ isolated from serum

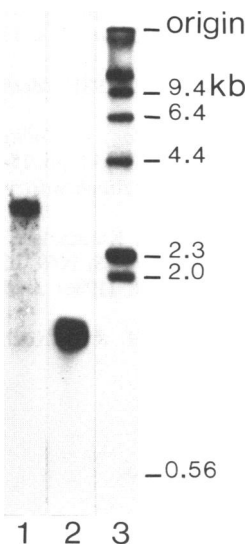


FIG. 4. RNA blot analysis of poly(A)⁺ RNAs from human liver. Each lane was loaded with 5 μg of mRNA. Anode is at the bottom. Lanes: 1, hybridization with H-chain-specific probe (cDNAs $\lambda\text{HuHITI-9}$ or -19); 2, hybridization with L-chain-specific probe (pHuLITI1; ref. 3); 3, size markers (in kb) (*Hind*III-restricted λDNA that was end-labeled by the Klenow filling-in reaction).

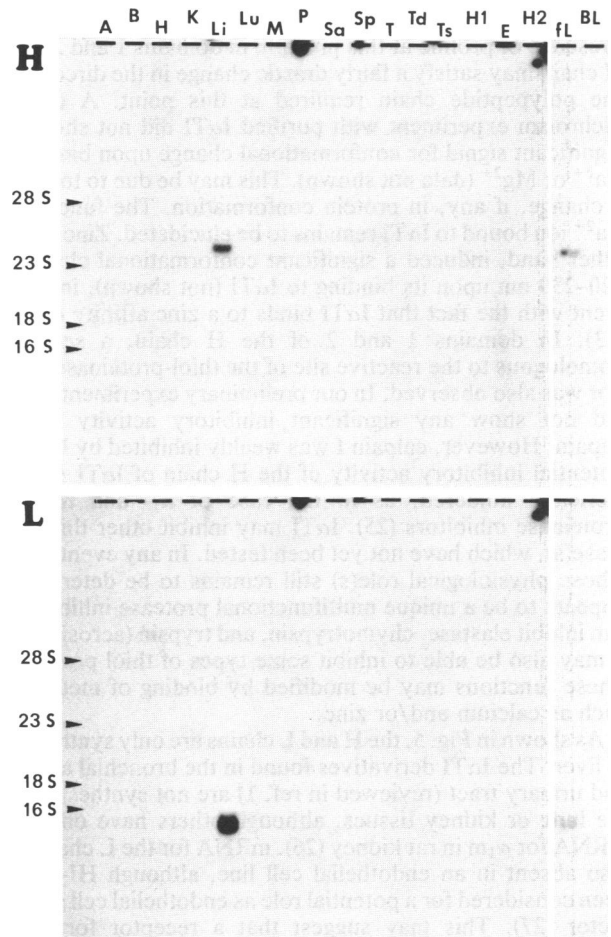


FIG. 5. RNA blot analysis of total RNAs from baboon tissues and human cell lines. Each lane was loaded with 20 μg of total RNA. Anode is at the bottom. (Upper) H-chain probe $\lambda\text{HuHITI-9}$ was used. (Lower) Two L-chain probes containing the $\alpha_1\text{m}$ or HI-30 sequence gave identical results. A, adrenal; B, brain; H, heart; K, kidney; Li, liver; Lu, lung; M, muscle; P, pancreas; Sa, salivary gland; Sp, spleen; T, testis; Td, thyroid; Ts, thymus; H1 and H2, human hepatoma cell lines; E, human umbilical cord endothelium cell line; fl, human fetal liver (18 weeks old); BL, human B-lymphoblastoid cell line. Marker RNAs are indicated on the left.

is not immunoprecipitated by an anti- α_1 m antiserum (J.P.S., unpublished data). Serum I α TI of M_r 180,000, therefore, may be composed of one H chain (M_r 95,000 without carbohydrates) covalently linked to two HI-30 moieties of the L-chain. HI-30 has an apparent M_r of 44,000 in NaDodSO₄/PAGE due to its very high carbohydrate content (21). The bonds linking the H chain and HI-30 are not susceptible to reducing agents, indicating that a bond(s) other than a disulfide bond(s), such as the isopeptide bond(s) found in fibrin or fibronectin (22) or the Lys-Lys bond(s) found in collagen (23), may be involved. The nature of the bonds still remains to be determined.

The H chain of I α TI clearly contains homologous domains as shown in Fig. 2. The relative positions of domains 1 and 2 in the H chain are not yet known. The possibility that the H chain is heterogeneous and each H chain may contain either of these domains cannot be excluded. In domains 1 and 2, we found two similar sequences, which are homologous to the sequence for a well-defined calcium-binding site. In this sequence, an aspartic acid and two asparagine/aspartic acid residues provide chelating ligands for calcium. Position 6 (mostly glycine) permits folding of the polypeptide chain to form a proper orientation of ligands to chelate the metal ion (15). Replacement of aspartic acid in position X with other amino acids has been reported (see Fig. 3). A glycine in position 6 is not an absolute requirement (24) and the presence of proline at this position in domains 1 and 2 of the H chain may satisfy a fairly drastic change in the direction of the polypeptide chain required at this point. A circular dichroism experiment with purified I α TI did not show any significant signal for conformational change upon binding of Ca²⁺ or Mg²⁺ (data not shown). This may be due to too small a change, if any, in protein conformation. The function of Ca²⁺ ion bound to I α TI remains to be elucidated. Zinc, on the other hand, induced a significant conformational change at 220–230 nm upon its binding to I α TI (not shown), in agreement with the fact that I α TI binds to a zinc affinity column (13). In domains 1 and 2 of the H chain, a sequence homologous to the reactive site of the thiol-proteinase inhibitor was also observed. In our preliminary experiments, I α TI did not show any significant inhibitory activity toward papain. However, calpain I was weakly inhibited by I α TI. A potential inhibitory activity of the H chain of I α TI may be sterically hindered, as in the case of α_1 - and α_2 -thiol-proteinase inhibitors (25). I α TI may inhibit other thiol protease(s), which have not yet been tested. In any event, I α TI, whose physiological role(s) still remains to be determined, appears to be a unique multifunctional protease inhibitor. It can inhibit elastase, chymotrypsin, and trypsin (acrosin), and it may also be able to inhibit some types of thiol proteases. These functions may be modified by binding of metal ions such as calcium and/or zinc.

As shown in Fig. 5, the H and L chains are only synthesized in liver. The I α TI derivatives found in the bronchial airways and urinary tract (reviewed in ref. 1) are not synthesized by the lung or kidney tissues, although others have observed mRNA for α_1 m in rat kidney (26). mRNA for the L chain was also absent in an endothelial cell line, although HI-30 has been considered for a potential role as endothelial cell growth factor (27). This may suggest that a receptor for HI-30 molecule is present on the endothelial cell surface.

We are indebted to Dr. R. Benarous for helping with cDNA library screening, Drs. J. H. Heaton, R. Zeheb, T. Gelehrter, and D. Kurnit

for providing us with RNAs from human cell lines and RNAs from a human fetus, and Drs. S. Higashiyama and R. Nakamura for performing circular dichroism studies. The secretarial assistance of Ms. E. Vong and Mrs. N. Salier was appreciated. This work was supported by Institut National de la Santé et de la Recherche Médicale, the University of Rouen, and the National Institutes of Health (Grant HL38644 to K.K.). Part of this work was done while J.P.S. was in the laboratory of K.K. at the University of Michigan.

1. Salier, J. P., Sesboue, R., Bourguignon, J., Diarra-Mehrpour, M. & Martin, J. P. (1987) in *Proceedings of the Second International Symposium on Pulmonary Emphysema and Proteolysis*, eds. Mittman, C. & Taylor, J. C. (Academic, New York), pp. 159–166.
2. Bourguignon, J., Vercaigne, D., Sesboue, R., Martin, J. P. & Salier, J. P. (1983) *FEBS Lett.* **162**, 379–383.
3. Bourguignon, J., Diarra-Mehrpour, M., Sesboue, R., Frain, M., Sala-Trepat, J. M., Martin, J. P. & Salier, J. P. (1985) *Biochem. Biophys. Res. Commun.* **131**, 1146–1153.
4. Kaumeyer, J. F., Polazzi, J. O. & Kotick, M. P. (1986) *Nucleic Acids Res.* **14**, 7839–7850.
5. Wachter, E. & Hochstrasser, K. (1979) *Hoppe-Seyler's Z. Physiol. Chem.* **360**, 1305–1311.
6. Reisinger, P., Hochstrasser, K., Albrecht, G. J., Lempart, K. & Salier, J. P. (1985) *Biol. Chem. Hoppe-Seyler* **366**, 479–483.
7. Salier, J. P., Sesboue, R., Vercaigne, D., Bourguignon, J. & Martin, J. P. (1983) *Anal. Biochem.* **133**, 336–343.
8. Young, R. A. & Davis, R. W. (1985) in *Genetic Engineering*, eds. Setlow, J. & Hollaender, A. (Plenum, New York), Vol. 7, pp. 29–41.
9. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
10. Deininger, P. L. (1983) *Anal. Biochem.* **129**, 216–223.
11. Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3963–3965.
12. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
13. Salier, J. P., Martin, J. P., Lambin, P., McPhee, H. & Hochstrasser, K. (1980) *Anal. Biochem.* **109**, 273–283.
14. Ohkubo, I., Kurachi, K., Takasawa, T., Shiokawa, H. & Sasaki, M. (1984) *Biochemistry* **23**, 5691–5697.
15. Garipey, J. & Hodges, R. S. (1983) *FEBS Lett.* **160**, 1–6.
16. Sasagawa, T., Ericsson, L. H., Walsh, K. A., Schreiber, W. E., Fischer, E. H. & Titani, K. (1982) *Biochemistry* **21**, 2565–2569.
17. Takahashi, N., Takahashi, Y. & Putnam, F. W. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8019–8023.
18. Higashiyama, S., Ohkubo, I., Ishiguro, H. & Sasaki, M. (1987) *Biochemistry*, in press.
19. Erickson-Viitanen, S., O'Neil, K. T. & De Grado, W. F. (1987) in *Protein Engineering*, eds. Oxender, D. L. & Fox, C. F. (Liss, New York), pp. 201–211.
20. Mendez, E., Fernandez-Luna, J. L., Grubb, A. & Leyva-Cobian, F. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1472–1475.
21. Hochstrasser, K., Schonberger, O. L., Rossmann, I. & Wachter, E. (1981) *Hoppe-Seyler's Z. Physiol. Chem.* **362**, 1357–1362.
22. Lorand, L., Credo, R. B. & Janus, T. J. (1981) *Methods Enzymol.* **80**, 333–341.
23. Light, N. D. & Bailey, A. J. (1980) in *Biology of Collagen*, eds. Viidik, A. & Vuust, J. (Academic, New York), pp. 15–38.
24. Takagi, T., Konishi, K. & Cox, J. A. (1986) *Biochemistry* **25**, 3585–3592.
25. Higashiyama, S., Ohkubo, I., Ishiguro, H., Kunitatsu, M., Sawaki, K. & Sasaki, M. (1986) *Biochemistry* **25**, 1669–1675.
26. Kastern, W., Bjorck, L. & Akerstrom, B. (1986) *J. Biol. Chem.* **261**, 15070–15074.
27. McKeehan, W. L., Sakagami, Y., Hoshi, H. & McKeehan, K. A. (1986) *J. Biol. Chem.* **261**, 5378–5383.