Volume 14 Number 20 1986 Nucleic Acids Research

The mRNA for a proteinase inhibitor related to the HI-30 domain of inter-α-trypsin inhibitor also encodes α-1-microglobulin (protein HC)

John F.Kaumeyer, Joseph O.Polazzi and Michael P.Kotick

Recombinant DNA-Chemistry, Microbiology Research, Biotechnology Group, PO Box 932, Miles Laboratories, Inc., Elkhart, IN 46515, USA

Received 1 August 1986; Revised and Accepted 15 September 1986

ABSTRACT

Inter-α-trypsin inhibitor (ITI) is a 180 kd serine proteinase inhibitor found in human serum. Treatment of 180 kd ITI with trypsin releases a 30 kd fragment (HI-30) which contains the anti-proteolytic activity of the high molecular weight form. We have isolated a cDNA clone from a human liver library which codes for HI-30, and have determined its DNA sequence. The mRNA not only codes for HI-30 but also another serum protein, α-1-microglobulin, which has not been previously associated with ITI or HI-30. The α-1-microglobulin sequence is found in the amino-terminus of the protein and is preceded by a signal sequence. HI-30 is found at the carboxy-terminus. The two protein sequences are separated by two arginine residues.

INTRODUCTION

Inter-alpha-trypsin inhibitor (ITI) is a serine proteinase inhibitor found in human plasma and serum (1). It is unique among the serum proteinase inhibitors in that it occurs as more than one species. It is found as a high molecular weight form of 180 kd and as a lower molecular weight species of 30 kd, designated HI-30 (2,3). This lower molecular weight species is also found in urine (4). HI-30, which is released from 180 kd ITI by treatment with trypsin in vitro (3), contains all of the anti-proteolytic activity of the 180 kd form (5). Although the full structure of ITI is unknown, HI-30 has been purified from urine and the complete amino acid sequence has been determined (6). HI-30 is composed of 50% carbohydrate and contains two tandemly linked anti-proteolytic domains (7). Both domains are structurally related to the Kunitz family of low molecular weight serine proteinase inhibitors (6).

Although 180 kd ITI has been described as a single chain glycoprotein (8-10), recent experiments indicate there are multiple mRNAs in baboon liver coding for polypeptides related to ITI (11). Specifically, HI-30 is synthesized in the liver as part of a 40 kd protein. A cDNA clone of this mRNA has been isolated from human liver mRNA and a partial sequence reported (12). The sequence data indicate that HI-30 is found in the carboxy-terminus of the protein.

We have also isolated a cDNA clone for this polypeptide and have determined the entire sequence. Our data confirm the findings of Bourguignon and co-workers (12) that HI-30 is found at the carboxy-terminus of the encoded protein. In addition we find that the amino-terminus of this polypeptide contains the sequence of the serum protein α-1-microglobulin (protein HC) (13,14), which has not been previously associated with ITI or HI-30. These results shed new light on the source of α-1-microglobulin and HI-30 found in serum.

MATERIALS & METHODS

Materials

Reagents and monomers used for DNA synthesis were purchased from Applied Biosystems. Radioisotopes were purchased from New England Nuclear. The Bluescribe RNA expression vector and reagents used for *in vitro* transcription were purchased from Vector Cloning Systems (San Diego, CA). Guanylyltransferase was purchased from Bethesda Research Laboratories. Rabbit antibodies (IgG fraction) to α-1-microglobulin and inter-α-trypsin inhibitor were purchased from Accurate Chemical and Scientific Corp. (Westbury, NY). IgGsorb was purchased from The Enzyme Center.

Oligonucleotide Probes

DNA oligonucleotides were synthesized on an Applied Biosystems Model 380A synthesizer, using the phosporamidite technology (15). Oligonucleotides were purified as recommended by the manufacturer, including purification on Sephadex G-50 and polyacrylamide electrophoresis in 7M urea. Radioactive probes were synthesized by annealing complementary oligonucleotides and filling in with Klenow fragment from *E. coli* DNA polymerase I. Prior to synthesis, oligomers were annealed at a concentration of 200 μg/ml in 0.1M NaCl. The mixture was heated at 65°C for 10 minutes, 37°C for 20 minutes, 23°C for 20 minutes, 15°C for 20 minutes and 4°C for 20 minutes. The synthesis reaction consisted of 0.2 μg annealed oligomers, 400 μM dATP, 400 μM dTTP, 100 μCi α-^{32}P-dCTP (3000 Ci/mmol), 100 μCi α-^{32}P-dGTP (3000 Ci/mmol), 50 mM tris (pH 7.2), 10mM $MgCl_2$, 0.1 mM dithiothreitol, 50 μg/ml bovine serum albumin, and 10 units Klenow fragment, in a volume of 30 μl. The reaction was incubated at 23°C for 60 minutes and then was made 400 μM with unlabeled dCTP and dGTP. After 15 minutes at 23°C the reaction was stopped with 1 μl of 10% SDS and the labeled oligomers purified by chromatography on Sephadex G-50. Using this method specific activities of 10^9 cpm/μg were routinely reached.

Library Screening

A λgt11 library of human liver mRNA was constructed by Dr. P. J. Simpson of Miles Laboratories. The library was constructed as described by Huynh et al.

(16) except that the double-stranded cDNA was not treated with EcoRI methylase before being digested with EcoRI. The library consisted of 10^6 independent clones of which 71% of the packaged phage contained inserts. The library was screened as described by Benton and Davis (17). A total of 10^6 phage were plated on Y1088 at a density of 375 phage/cm^2. The plaque lifts were prehybridized for at least 4 hours at 37°C in 20% formamide, 2X Denhardt's solution, 5X SSPE, 0.1% SDS, 100 µg/ml sheared, single-stranded salmon sperm DNA. The filters were then hybridized overnight in the same conditions with 10^6 cpm/ml of radiolabeled oligomers. The hybridization mixture was boiled for 10 minutes before adding to the filters. The filters were washed in 0.2X SSPE, 0.1% SDS at 37°C and autoradiographed with Kodak XAR-2 film using Dupont Hi-Plus intensifying screens at -70°C.

DNA Sequencing

The DNA fragment was gel purified and used to generate a random library of fragments in the M13 vector mp-18 using the sonication technique of Deininger (18). Individual M13 clones were sequenced using the dideoxy chain termination method of Sanger (19). Twenty random clones were first used to determine the DNA sequence. Four primers were synthesized and used to confirm remaining regions of ambiguous sequence. Each nucleotide was confirmed by sequencing at least three times and 95% of the nucleotides were sequenced four times.

Computer Analysis

Data from shotgun DNA sequencing was aligned using the GEL program of Intelligenetics. The National Biomedical Research Foundation Protein Sequence database was searched with the IFIND program. Matching sequences were aligned using the ALIGN program. Open reading frame searching and hydropathy analysis were done with the DNA/Protein Sequence Analysis Software purchased from International Biotechnologies, Inc.

In Vitro RNA Transcription

The HI-30 clone was ligated into the EcoRI site of the Bluescribe RNA expression vector. This construct was cut with HindIII and then transcribed *in vitro* using T7 RNA polymerase to produce an mRNA copy of the insert. The conditions used for synthesis of RNA were the same as those of Melton et al. (20). Analysis of the transcription products by gel electrophoresis showed that the RNA was full length. The RNA transcript was phenol extracted and ethanol precipitated after synthesis.

Capping and In Vitro Translation

RNA transcripts were capped with vaccinia virus guanylyltransferase. The reaction contained 50mM tris (pH 7.9), 1.25 mM $MgCl_2$, 6 mM KCl, 2.5 mM dithiothreitol, 100 µg/ml bovine serum albumin, 100 µM S-adenosyl methionine, 330

μM GTP. One unit of enzyme was used per μg of RNA. The reaction was carried out at 37°C for 45 minutes. It was terminated by addition of SDS to 0.5% and then phenol extracted.

Capped mRNAs were translated in the micrococcal nuclease-treated rabbit reticulocyte lysate translation system purchased from Promega Biotec. The translations were performed as recommended by the manufacturer. Immunoprecipitations were performed as described by Dobberstein et al. (21), with the exception that IgGsorb was substituted for protein A-Sepharose. SDS-polyacrylamide gel electrophoresis was performed as described by Laemmli (22). Samples were boiled for 10 minutes in 4% SDS and 50mM dithiothreitol before electrophoresis.

RESULTS

Our approach to isolating the cDNA clone of HI-30 was to employ the technique of long oligonucleotide probes (23-25). The probes were designed by selecting regions of HI-30 (6) with low codon degeneracy and converting this amino acid sequence into DNA sequence using most preferred human codons (26). The two protein sequences utilized in the design and the oligonucleotides synthesized are given in Fig. 1. One probe represented the sequence Met-49 to Phe-66, which is located in the first anti-proteolytic domain of HI-30. Another was taken from amino acids Cys-115 to Cys-132 found in the second domain. Each probe was constructed from two 33 nucleotide long oligomers which overlapped by 12 nucleotides. Radioactive probes were made by annealing the oligomers and filling in the remaining single stranded regions with Klenow fragment and radioactive nucleotides.

The probes were pooled for screening the λgt11 human liver library. More

```
   DOMAIN I

   Met Ala Cys Glu Thr Phe Gln Tyr Gly Gly Cys Met Gly Asn Gly Asn Asn Phe

5'-ATG GCC TGC GAG ACC TTC CAG TAC GGC GGC TGC-3'
                            3'-ATG CCG CCG ACG TAC CCG TTG CCG TTG TTG AAG-5'

   DOMAIN II

   Cys Gln Gly Asn Gly Asn Lys Phe Tyr Ser Glu Lys Glu Cys Arg Glu Tyr Cys

5'-TGC CAG GGC AAC GGC AAC AAG TTC TAC AGC CAG-3'
                            3'-AAG ATG TCG GTC TTC CTC ACG TCT CTC ATG ACG-5'
```

Figure 1. Oligonucleotide probes used for isolation of the α-1-microglobulin-HI-30 cDNA clone. The protein sequence used for designing the probes is taken from Wachter and Hochstrasser (6). The sequence from domain I is Met-49 to Phe-66, and that from domain II is Cys-115 to Cys-132.

than 100 positive clones were identified in the 426,000 phage which were screened. Twelve positive clones were picked and plaque purified. They were analysed by characterizing their melting temperature with each probe and determining the size of the inserted DNA. All 12 clones hybridized to each probe individually. The probe for domain I remained hybridized to each clone at 60°C in 0.1X SSPE, 0.1% SDS. The probe for domain II washed off between 50°C and 60°C. The sizes of the inserts ranged from ~700 to ~1300 nucleotides as determined by polyacrylamide gel electrophoresis. The clone with the largest insert was chosen for further analysis by DNA sequencing.

The DNA sequence presented in Fig. 2 reveals that the EcoRI fragment of this clone is 1221 nucleotides long. The fragment contains an open reading frame of 352 amino acids which initiates with methionine at nucleotide position 46 and terminates at nucleotide 1101. The clone also contains an eight nucleotide stretch of adenylate residues at the 3' end. The signal for poly(A) addition, AATAAA, is found 19 nucleotides upstream. The complete sequence for HI-30 is found at the carboxyl-end of the open reading frame between nucleotides 661 and 1095. The oligonucleotides used to construct the probes have high homology with the cDNA sequence. The domain I probe has 94% homology with the cDNA and the domain II probe has 91% homology. These high levels of homology are consistent with the melting characteristics of these probes.

The amino acid sequence of HI-30 predicted by the DNA sequence of the cDNA is almost identical to that determined by Wachter and Hochstrasser for HI-30 purified from human urine (6). There are minor differences at two locations. The previously published protein sequence assigns Val-Ile at amino acid positions 86-87 of HI-30 and Glu at position 138. The DNA sequence which we determined for HI-30 assigns an Ile-Val pair and a Gly at these positions, respectively. Our DNA sequence is almost identical to that determined by Bourguignon et al. (12) for a cDNA clone from human liver. Their report presents the last 264 nucleotides of a similar mRNA. The only difference between the two sequences is an extra cytosine in our sequence found at position 1183, in the 3' untranslated region.

The amino terminal end of the open reading frame contains a 205 amino acid sequence which has not been previously identified with either 180 kd ITI or HI-30. To determine if this protein sequence might be related to that of other serum proteins, we used this sequence to search the NBRF Protein Sequence Database. Surprisingly the sequence showed an almost perfect match with the sequences of two proteins α-1-microglobulin (13) and protein HC (14). A comparison of the sequence from the amino-terminal region of the open reading

```
            10          20          30          40          50          60
            *           *           *           *           *           *
TAG ACC GAG CCT GTG GGA TAT ACC AAG GCA GAG GAG CCC ATA GCC ATG AGG AGC CTC GGG
                                                        Met Arg Ser Leu Gly

                                                alpha-1-mic
            70          80          90                      110         120
            *           *           *                       *           *
GCC CTG CTC TTG CTG CTG AGC GCC TGC CTG GCG GTG AGC GCT GGC CCT GTG CCA ACG CCG
Ala Leu Leu Leu Leu Leu Ser Ala Cys Leu Ala Val Ser Ala Gly Pro Val Pro Thr Pro

           130         140         150         160         170         180
            *           *           *           *           *           *
CCC GAC AAC ATC CAA GTG CAG GAA AAC TTC AAT ATC TCT CGG ATC TAT GGG AAG TGG TAC
Pro Asp Asn Ile Gln Val Gln Glu Asn Phe Asn Ile Ser Arg Ile Tyr Gly Lys Trp Tyr

           190         200         210         220         230         240
            *           *           *           *           *           *
AAC CTG GCC ATC GGT TCC ACC TGC CCC TGG CTG AAG AAG ATC ATG GAC AGG ATG ACA GTG
Asn Leu Ala Ile Gly Ser Thr Cys Pro Trp Leu Lys Lys Ile Met Asp Arg Met Thr Val

           250         260         270         280         290         300
            *           *           *           *           *           *
AGC ACG CTG GTG CTG GGA GAG GGC GCT ACA GAG GCG GAG ATC AGC ATG ACC AGC ACT CGT
Ser Thr Leu Val Leu Gly Glu Gly Ala Thr Glu Ala Glu Ile Ser Met Thr Ser Thr Arg

           310         320         330         340         350         360
            *           *           *           *           *           *
TGG CGG AAA GGT GTC TGT GAG GAG ACG TCT GGA GCT TAT GAG AAA ACA GAT ACT GAT GGG
Trp Arg Lys Gly Val Cys Glu Glu Thr Ser Gly Ala Tyr Glu Lys Thr Asp Thr Asp Gly

           370         380         390         400         410         420
            *           *           *           *           *           *
AAG TTT CTC TAT CAC AAA TCC AAA TGG AAC ATA ACC ATG GAG TCC TAT GTG GTC CAC ACC
Lys Phe Leu Tyr His Lys Ser Lys Trp Asn Ile Thr Met Glu Ser Tyr Val Val His Thr

           430         440         450         460         470         480
            *           *           *           *           *           *
AAC TAT GAT GAG TAT GCC ATT TTC CTG ACC AAG AAA TTC AGC CGC CAT CAT GGA CCC ACC
Asn Tyr Asp Glu Tyr Ala Ile Phe Leu Thr Lys Lys Phe Ser Arg His His Gly Pro Thr

           490         500         510         520         530         540
            *           *           *           *           *           *
ATT ACT GCC AAG CTC TAC GGG CGG GCG CCG CAG CTG AGG GAA ACT CTC CTG CAG GAC TTC
Ile Thr Ala Lys Leu Tyr Gly Arg Ala Pro Gln Leu Arg Glu Thr Leu Leu Gln Asp Phe

           550         560         570         580         590         600
            *           *           *           *           *           *
AGA GTG GTT GCC CAG GGT GTG GGC ATC CCT GAG GAC TCC ATC TTC ACC ATG GCT GAC CGA
Arg Val Val Ala Gln Gly Val Gly Ile Pro Glu Asp Ser Ile Phe Thr Met Ala Asp Arg

           610         620         630         640         650         660
            *           *           *           *           *           *
GGT GAA TGT GTC CCT GGG GAG CAG GAA CCA GAG CCC ATC TTA ATC CCG AGA GTC CGG AGG
Gly Glu Cys Val Pro Gly Glu Gln Glu Pro Glu Pro Ile Leu Ile Pro Arg Val Arg Arg

HI-30
           670         680         690         700         710         720
            *           *           *           *           *           *
GCT GTG CTA CCC CAA GAA GAG GAA GGA TCA GGG GGT GGG CAA CTG GTA ACT GAA GTC ACC
Ala Val Leu Pro Gln Glu Glu Glu Gly Ser Gly Gly Gly Gln Leu Val Thr Glu Val Thr

      I    730         740         750         760         770         780
            *           *           *           *           *           *
AAG AAA GAA GAT TCC TGC CAG CTG GGC TAC TCG GCC GGT CCC TGC ATG GGA ATG ACC AGC
Lys Lys Glu Asp Ser Cys Gln Leu Gly Tyr Ser Ala Gly Pro Cys Met Gly Met Thr Ser

           790         800         810         820         830         840
            *           *           *           *           *           *
AGG TAT TTC TAT AAT GGT ACA TCC ATG GCC TGT GAG ACT TTC CAG TAC GGC GGC TGC ATG
Arg Tyr Phe Tyr Asn Gly Thr Ser Met Ala Cys Glu Thr Phe Gln Tyr Gly Gly Cys Met
```

```
          850         860         870         880         890 ||      900
           *           *           *           *           *          *
GGC AAC GGT AAC AAC TTC GTC ACA GAA AAG GAG TGT CTG CAG ACC TGC CGA ACT GTG GCG
Gly Asn Gly Asn Asn Phe Val Thr Glu Lys Glu Cys Leu Gln Thr Cys Arg Thr Val Ala

          910         920         930         940         950         960
           *           *           *           *           *          *
GCC TGC AAT CTC CCC ATA GTC CGG GGC CCC TGC CGA GCC TTC ATC CAG CTC TGG GCA TTT
Ala Cys Asn Leu Pro Ile Val Arg Gly Pro Cys Arg Ala Phe Ile Gln Leu Trp Ala Phe

          970         980         990        1000        1010        1020
           *           *           *           *           *          *
GAT GCT GTC AAG GGG AAG TGC GTC CTC TTC CCC TAC GGG GGC TGC CAG GGC AAC GGG AAC
Asp Ala Val Lys Gly Lys Cys Val Leu Phe Pro Tyr Gly Gly Cys Gln Gly Asn Gly Asn

         1030        1040        1050        1060        1070        1080
           *           *           *           *           *          *
AAG TTC TAC TCA GAG AAG GAG TGC AGA GAG TAC TGC GGT GTC CCT GGT GAT GGT GAT GAG
Lys Phe Tyr Ser Glu Lys Glu Cys Arg Glu Tyr Cys Gly Val Pro Gly Asp Gly Asp Glu

         1090        1100        1110        1120        1130        1140
           *           *           *           *           *          *
GAG CTG CTG CGC TTC TCC AAC TGA CAA CTG GCC GGT CTG CAA GTC AGA GGA TGG CCA GTG
Glu Leu Leu Arg Phe Ser Asn ---

         1150        1160        1170        1180        1190        1200
           *           *           *           *           *          *
TCT GTC CCG GGG TCC TGT GGC AGG CAG CGC CAA GCA ACC TGG GTC CAA ATA AAA ACT AAA

         1210        1220
           *           *
TTG TAA ACT CCT GAA AAA AAA
```

Figure 2. Nucleotide and protein sequence of the α-1-microglobulin-HI-30 protein. The sequence of α-1-microglobulin (α-1-mic) begins at nucleotide 103 and extends to 651. HI-30 begins at 661 and extends to 1095. Domain I begins at 724; domain II starts at 891. Amino acid sequences which were used for designing oligonucleotide probes are underlined. Positions which differ from the protein sequence of HI-30 purified from urine are enclosed.

frame with these two proteins is shown in Fig. 3. The sequence of protein HC is more complete. There is a mismatch at position 55 and a two amino acid gap at positions 56 and 57. The sequence stops three amino acids short of the Arg-Arg dipeptide adjacent to the first amino acid of the HI-30 region. Both α-1-microglobulin and protein HC sequences begin at position Gly-20 of the open reading frame. Hydropathy analysis of the open reading frame indicates that the first 19 amino acids are highly hydrophobic and probably represent the signal sequence involved in secretion of this protein from liver cells.

Fig. 4 summarizes the structure of this cDNA clone for α-1-microglobulin and HI-30. It contains a signal sequence followed by the sequence of α-1-microglobulin. At the junction of α-1-microglobulin and HI-30 are two arginine residues. The sequence of HI-30 begins immediately after these residues and terminates with an additional Ser-Asn dipeptide. The protein has a calculated molecular weight of 42 kd which is similar in size to the ITI- related protein

```
        10         20         30         40         50         60
                    GPVPTPPDNIQVQENFNISRIYGKWYNL----------KIMDRMTV     MIC
                    ||||||||||||||||||||||||||||          ||||||||
MRSLGALLLLLSACLAVSAGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLKKIMDRMTV     ITI
                    ||||||||||||||||||||||||||||||||||||  ||||||||
                    GPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPL--KIMDRMTV     HC

    70         80         90        100        110        120        130
STLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHKSKWNITMESYVVHTNYDEY     MIC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
STLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHKSKWNITMESYVVHTNYDEY     ITI
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
STLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHKSKWNITMESYVVHTNYDEY     HC

       140        150        160        170        180        190
AIFLTK-FSRHHGP-ITAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECVPGEQQP     MIC
||||||  |||||| ||||||||||||||||||||||||||||||||||||||||||||||||| |
AIFLTKKFSRHHGPTITAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECVPGEQEP     ITI
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AIFLTKKFSRHHGPTITAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECVPGEQEP     HC

  200
EPI                                                                    MIC
|||
EPILIPRVRR                                                             ITI
|||||||
EPILIPR                                                                HC
```

Figure 3. Alignment of sequence homologies between α-1-microglobulin (MIC), protein HC (HC), and the first 205 amino acids of the α-1-microglobulin-HI-30 protein.

synthesized in baboon liver (11). To demonstrate the structure of this protein we synthesized an mRNA from the cDNA clone and translated the mRNA in a reticulocyte lysate in vitro translation system. The protein product was then challenged with antibodies to α-1-microglobulin and inter-α-trypsin inhibitor. The results of this experiment are shown in Fig. 5.

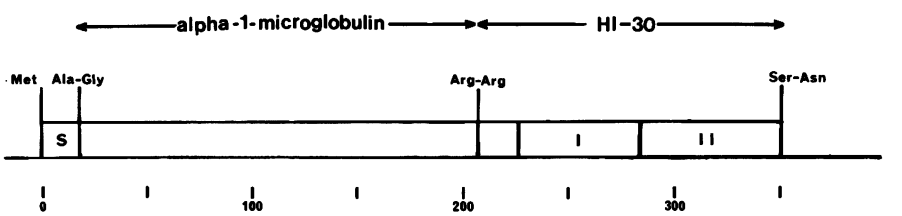

Figure 4. Structure of the α-1-microglobulin-HI-30 protein. Met is the first amino acid of the open reading frame. "S" indicates the signal sequence and Ala-Gly represents the potential signal peptidase cleavage site. The Arg-Arg dipeptide is found at the boundary of α-1-microglobulin and HI-30 regions. The HI-30 sequence consists of 21 amino acids followed by the Kunitz-like domains I and II. The Ser-Asn dipeptide occurs at the end of the open reading frame.

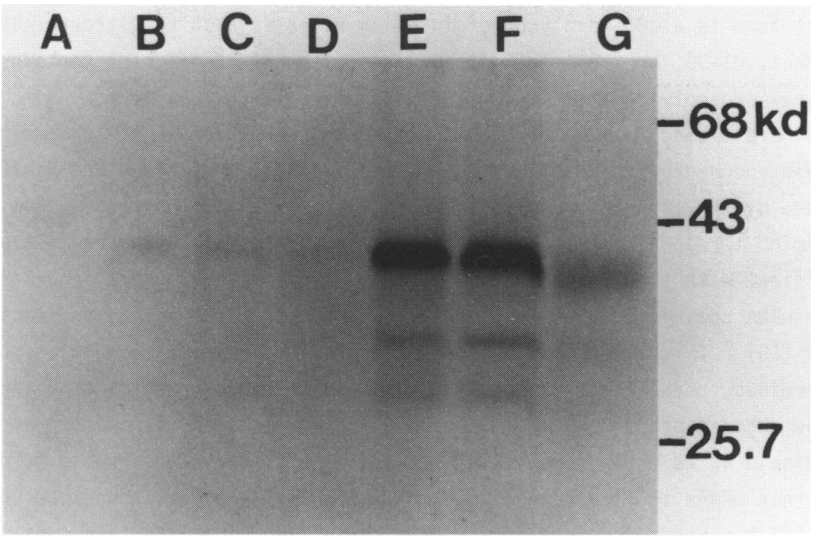

Figure 5. Immunoprecipitation of the translation product of α-1-microglobulin-HI-30 mRNA. RNA was transcribed from the insert cloned into the Bluescribe expression vector. The RNA was capped in vitro and translated with a micrococcal nuclease-treated rabbit reticulocyte lysate incorporating 35-S methionine. Lane A is a translation with no RNA added and lane G is a translation of in vitro synthesized mRNA. Lanes B, C, D are mRNA translations precipitated with non-immune antibody, anti-β-galactosidase, and no antibody. Lanes E and F were precipitated with anti-α-1-microglobulin and anti-inter-α-trypsin inhibitor.

The in vitro translation product is a 42 kd protein (lane G) which is consistent with the size predicted by the open reading frame. The protein band is slightly distorted by co-migrating proteins in the lysate. The protein is precipitated with antibodies to both α-1-microglobulin and 180 kd ITI (lanes E and F). No significant amount of protein product is precipitated by non-immune antibody, anti-β-galactosidase, or no antibody (lanes B, C, and D). Thus the cDNA clone codes for a protein containing the sequences of α-1-microglobulin and HI-30. Both domains are apparently properly folded to present epitopes which are recognized by antibodies made to mature 31 kd α-1-microglobulin and the higher molecular weight form of ITI.

DISCUSSION

Inter-α-trypsin inhibitor is a serum proteinase inhibitor which inhibits trypsin, chymotrypsin, and neutrophil elastase (28-30). However, the specific target of the inhibitor is unknown. This protein is of interest in that it occurs as both high and low molecular weight species. The high molecular

weight form is a 180 kd glycoprotein which releases a 30 kd glycoprotein fragment, HI-30, on treatment with trypsin. The HI-30 fragment contains all of the anti-proteolytic activity of the high molecular weight form.

We have cloned an mRNA from human liver which encodes an HI-30 domain. The protein encoded by this cDNA is only 42 kd in size. Moreover, the HI-30 domain encoded by the cDNA is found at the carboxy-terminus of the protein and not at the amino-terminus, as has been reported for ITI (8, 31). Our results are consistent with those of Bourguignon et al. These workers have demonstrated a liver mRNA coding for a 40 kd protein which reacts with antibodies specific for HI-30 (10). This liver mRNA has been cloned and partially sequenced from the 3' terminus, revealing that the HI-30 domain is found at the carboxy-terminus of the encoded protein (11). Although Bourguignon et al. identified an mRNA encoding a 42 kd protein with HI-30 at its 3' end, they were unable to detect any other mRNAs in liver encoding HI-30 (10). This result is unexpected in that ITI has been described as a single chain, 180 kd polypeptide containing HI-30 at its amino-terminus. However, they did detect mRNAs encoding high molecular weight proteins which reacted with antibodies to ITI but not with antibodies to HI-30. In view of these findings it may be possible that ITI is actually a very stable complex of proteins, containing HI-30.

Our cDNA clone codes for α-1-microglobulin as well as HI-30. α-1-Microglobulin has been characterized as a 31 kd glycoprotein found in serum, urine, and cerebrospinal fluid (32, 33). A similar protein has been isolated by other workers and designated protein HC (34). It contains a yellow-brown chromophoric group and migrates as a heterogeneously charged protein in agarose gels after desialylation with neuraminidase. α-1-Microglobulin is found in serum both as a free form and as a complex with IgA (35, 36). Although no clear function has been demonstrated for α-1-microglobulin, it has been suggested that it may play a role in modulating neutrophil chemotaxis (37). It is not apparent why α-1-microglobulin and HI-30 should be expressed on the same mRNA and then processed into two separately functioning molecules.

The structure of ITI remains unclear. There are no data to indicate that ITI contains α-1-microglobulin. A better understanding of the structure of ITI will require further analysis of the structure of the protein and cloning of the mRNAs which code for ITI.

Note added in revision

The sequence of a similar mRNA coding for α-1-microglobulin has recently been published (38).

ACKNOWLEDGMENTS

We thank Dr. P. J. Simpson for providing the λgt11 human liver cDNA library, David Leland and Janet Boyd for synthesis of oligomers, Mark Stuart for DNA sequencing and Phyllis Vandegrift for preparation of the manuscript. We also thank Drs. E. Truscheit and E. Auerswald, Bayer AG, for their continued interest in this project.

REFERENCES

1. Travis, J. and Salvesen, G. S. (1983) Annu. Rev. Biochem. 52, 655-709.
2. Hochstrasser, K., Feuth, H. and Steiner, O. (1973) Hoppe-Seyler's Z. Physiol. Chem. 354, 927-932.
3. Hochstrasser, K., Bretzel, G., Feuth, H., Hilla, W., and Lempart, K. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 153-162.
4. Hochstrasser, K., Reisinger, P., Albrecht, G. J., Wachter, E. and Schonberger, O. L. (1984) Hoppe-Seyler's Z. Physiol. Chem. 365, 1123-1130.
5. Hochstrasser, K. and Wachter, E. (1979) Hoppe-Seyler's Z. Physiol. Chem. 360, 1285-1296.
6. Wachter, E. and Hochstrasser, K. (1981) Hoppe-Seyler's Z. Physiol. Chem. 362, 1351-1355.
7. Hochstrasser, K., Schonberger, O. L., Rossmanith, I., and Wachter, E. (1981) Hoppe-Seyler's Z. Physiol. Chem. 362, 1357-1362.
8. Morii, M. and Travis, J. (1985) Biol. Chem. Hoppe-Seyler. 366, 19-21.
9. Steinbuch, M. (1976) In Lorand, L. (ed). Methods in Enzymology, Academic Press. New York. Vol. XLV, pp. 760-772.
10. Salier, J. P., Martin, J. P., Lambin, P., McPhee, H., and Hochstrasser, K. (1980) Anal. Biochem. 109, 273-283.
11. Bourguignon, J., Vercaigne, D., Sesboue, R., Martin, J. P. and Salier, J. P. (1983) FEBS Lett. 162, 379-383.
12. Bourguignon, J., Diarra-Mehrpour, M., Sesboue, R., Frain, M., Sala-Trepat, J. M., Martin, J. P. and Salier, J. P. (1985) Biochem. Biophys. Res. Comm. 131, 1146-1153.
13. Takagi, T., Takagi, K., Kawai, T. (1981) Biochem. Biophys. Res. Comm. 98, 997-1001.
14. Lopez, C., Grubb, A., Soriano, F., and Mendez, E. (1981) Biochem. Biophys. Res. Comm. 103, 919-925.
15. Beaucage, S. L., and Caruthers, M. H. (1981) Tetrahedron Letters 22, 1859-1862.
16. Huynh, T. V., Young, R. A., and Davis, R. W. (1985) In Glover, D. M., (ed). DNA Cloning, IRL Press. Maclean VA. Vol. I, pp. 49-78.
17. Benton, W. D. and Davis, R. W. (1977) Science 196, 180-182.
18. Deininger, P. L. (1983) Anal. Biochem. 129, 216-223.
19. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
20. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T. M., Zinn, E., Green, M. R. (1984) Nucl. Acids Res. 12, 7035-7056.
21. Dobberstein, B., Garoff, H., Warren, G., and Robinson, P. (1979) Cell 17, 759-769.
22. Laemmli, U. K. (1970) Nature (London) 227, 680-685.
23. Jaye, M., de la Salle, H., Schamber, F., Balland, A., Kohli, V., Findeli, A., Tolstoshev, P., and Lecocq, J.-P. (1983) Nucl. Acids Res. 8, 2325-2335.
24. Anderson, S. and Kingston, I. B. (1983) Proc. Natl. Acad. Sci. USA 80, 6838-6842.

25. Ullrich, A., Berman, C. H., Dull, T. J., Gray, A., and Lee, J. M. (1984) EMBO J. 3, 361-364.
26. Grantham, R., Gautier, C., Gouy, M., Jacobzone, M., and Mercier, R. (1981) Nucl. Acids Res. 9, r43-r74.
27. Kyte, J. and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105-132.
28. Bromke, B. J. and Kueppers, F. (1982) Biochem. Med. 27, 56-67.
29. Jonsson, B.-M., Loffler, Ch., and Ohlsson, K. (1982) Hoppe-Seyler's Z. Physiol. Chem. 363, 1167-1175.
30. Albrecht, G. J., Hochstrasser, K., and Salier, J.-P. (1983) Hoppe-Seyler's Z. Physiol. Chem. 364, 1703-1708.
31. Reisinger, P., Hochstrasser, K., Albrecht, G. J., Lempart, K., and Salier, J.-P. (1985) Biol. Chemistry Hoppe-Seyler 366, 479-483.
32. Ekstrom, B. and Berggard, I. (1977) J. Biol. Chem. 252, 8048-8057.
33. Ekstrom, B., Peterson, P. A., Berggard, I. (1975) Biochem. Biophys. Res. Commun. 65, 1427-1433.
34. Tejler, L. and Grubb, A. (1976) Biochimica et Biophysica Acta 439, 82-94.
35. Grubb, A. O., Lopez, C., Tejler, L., and Mendez, E. (1983) J. Biol. Chem. 258, 14698-14707.
36. Vincent, C., Bouic, P., Revillard, J.-P., and Bataille, R. (1985) Molec. Immunol. 22, 663-673.
37. Mendez, E., Fernandez-Luna, J. L., Grubb, A., and Leyva-Cobian, F. (1986) Proc. Natl. Acad. Sci. USA 83, 1472-1475.
38. Traboni, C. and Cortese, R. (1986) Nucl. Acids Res. 14, 6340.