Cloning and sequence of a cDNA coding for the human β -migrating endothelial-cell-type plasminogen activator inhibitor

(vascular fibrinolysis/serine protease inhibitor/placental cDNA expression library/DNA sequence analysis)

T. Ny^{*†}, M. Sawdey^{*}, D. Lawrence^{*}, J. L. Millan[‡], and D. J. Loskutoff^{*§}

*Department of Immunology, Scripps Clinic and Research Foundation, La Jolla, CA 92037; †Department of Applied Cell and Molecular Biology, University of Umea, S-90187 Umea, Sweden; and ‡La Jolla Cancer Research Foundation, La Jolla, CA 92037

Communicated by Ernest Beutler, June 5, 1986

ABSTRACT A λgt11 expression library containing cDNA inserts prepared from human placental mRNA was screened immunologically using an antibody probe developed against the β -migrating plasminogen activator inhibitor (β -PAI) purified from cultured bovine aortic endothelial cells. Thirty-four positive clones were isolated after screening 7×10^5 phages. Three clones ($\lambda 1.2$, $\lambda 3$, and $\lambda 9.2$) were randomly picked and further characterized. These contained inserts 1.9, 3.0, and 1.9 kilobases (kb) long, respectively. Escherichia coli lysogenic for λ 9.2, but not for λ gt11, produced a fusion protein of 180 kDa that was recognized by affinity-purified antibodies against the bovine aortic endothelial cell β -PAI and had β -PAI activity when analyzed by reverse fibrin autography. The largest cDNA insert was sequenced and shown to be 2944 base pairs (bp) long. It has a large 3' untranslated region [1788 bp, excluding the poly(A) tail] and contains the entire coding region of the mature protein but lacks the initiation codon and part of the signal peptide coding region at the 5' terminus. The two clones carrying the 1.9-kb cDNA inserts were partially sequenced and shown to be identical to the 3.0-kb cDNA except that they were truncated, lacking much of the 3' untranslated region. Blot hybridization analysis of electrophoretically fractionated RNA from the human fibrosarcoma cell line HT-1080 was performed using the 3.0-kb cDNA as hybridization probe. Two distinct transcripts, 2.2 and 3.0 kb, were detected, suggesting that the 1.9-kb cDNA may have been copied from the shorter RNA transcript. The amino acid sequence deduced from the cDNA was aligned with the NH₂-terminal sequence of the human β -PAI. Based on this alignment, the mature human β -PAI is 379 amino acids long and contains an NH₂-terminal valine. The deduced amino acid sequence has extensive (30%) homology with α_1 -antitrypsin and antithrombin III, indicating that the β -PAI is a member of the serine proteinase inhibitor (serpin) superfamily.

The generation of plasmin from plasminogen provides an important source of proteolytic activity in cells, tissues, and biological fluids (1, 2). Precise regulation of plasminogen activator (PA) activity may thus constitute a critical feature of many biological systems (3). Such control may be at the level of the formation and resolution of fibrin itself (4), at the level of the interaction of PAs with cells (5, 6), or by specific PA inhibitors (PAIs; ref. 7).

Available evidence indicates that there are at least three immunologically distinct PAIs, including the placental PAI (8), protease nexin (9), and the endothelial cell-derived PAI (10-12). The PAI synthesized by cultured bovine aortic endothelial cells (BAEs) has been purified and partially characterized (13). It differs from the placental PAI and protease nexin in that it exhibits β -mobility when analyzed by agarose zone electrophoresis (14). Moreover, it inhibits tissue-type PA (tPA) as well as urokinase-type PA (uPA), whereas protease nexin and the placental PAI are primarily uPA inhibitors (15). Antiserum to the β -migrating PAI from BAEs has been developed and employed to show that the β -PAIs from human endothelial cells, plasma, serum, and platelets are immunologically related (16).

The level of β -PAI mRNA produced by cultured BAEs varies depending on culture conditions (17), suggesting that the β -PAI gene is regulated by external factors. The production of β -PAI by human endothelial cells (18) and by rat HTC cells (19) is stimulated several fold by interleukin 1 and dexamethasone, respectively, whereas gonadotropins decrease the β -PAI activity of granulosa cells (20). β -PAI is also subject to regulation at the protein level, since both activated protein C (21) and oxidants (22) directly neutralize its activity.

To facilitate the precise biochemical characterization of β -PAI, and to eventually understand the nature of factors regulating β -PAI gene expression, we have undertaken the molecular cloning of the gene. Here we describe the isolation of β -PAI cDNA from a human placental expression library and demonstrate that the β -PAI is a member of the serine protease inhibitor (serpin) superfamily (23, 24).

MATERIALS AND METHODS

Materials. Restriction enzymes, alkaline phosphatase, bacteriophage T4 DNA ligase, *Escherichia coli* DNA polymerase I, Klenow fragment of DNA polymerase I, and T4 DNA polymerase were purchased from Boehringer Mannheim. $[\alpha^{-32}P]dGTP$ (3000 Ci/mmol) and 5'- $[\alpha^{-35}S]$ thio]dATP (600 Ci/mmol; 1 Ci = 37 GBq) were purchased from Amersham. Human α -thrombin was a generous gift of J. Fenton (Albany, NY), and fibrinogen was purchased from Calbiochem-Behring. The purification of human plasminogen and urokinase (25) and the purification of BAE β -PAI and the development of antibodies to it (13) were as described. Antiserum to the placental PAI was purchased from American Diagnostica (Greenwich, CT).

Preparation and Analysis of Crude Placental Extract. Frozen placenta (3.5 g) was washed with phosphate-buffered saline (PBS: 0.14 M NaCl/0.01 M sodium phosphate, pH 7.2) and then extracted with 15 ml of PBS containing 0.5% Triton X-100 at 4°C. The tissue was homogenized using a Dounce homogenizer, and cellular debris was removed by centrifugation at 10,000 \times g for 10 min. The extracts were analyzed for inhibitor activity by reverse fibrin autography (25). Monospecific antisera against human placental-type PAI and

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: PA, plasminogen activator; tPA, tissue-type PA; uPA, urokinase-type PA; PAI, PA inhibitor; BAE, bovine aortic endothelial cell; bp, base pair(s); kb, kilobase(s).

⁸To whom correspondence should be addressed at: Department of Immunology, Scripps Clinic and Research Foundation, 10666 N. Torrey Pines Road, La Jolla, CA 92037.

bovine β -PAI were coupled to protein A-Sepharose (Pharmacia) and used as described (13, 17) to immunoprecipitate the PAIs present in the extract. Antiserum against bovine β -PAI recognizes human β -PAI (16).

Immunological Screening of a Human λ gt11 cDNA Library. The library, derived from a premature (34 weeks of gestation) human placenta and consisting of 10⁶ independent recombinant phages (26), was screened immunologically (27–29) for β -PAI, using the affinity-purified IgG fraction (30) of antibodies to the purified BAE β -PAI as antibody probe (13). To visualize antibody binding, ¹²⁵I-labeled protein A (55 mCi/mg) was employed. Autoradiography was performed by exposing the filters to Kodak XAR5 film with an intensifying screen at -80° C.

Immunoblot Analysis of E. coli Lysates. $\lambda gt11$ and recombinant lysogens were induced and crude extracts were prepared as described (29). For immunoblot analysis, 50 μ l of crude extract was fractionated by NaDodSO₄/PAGE (31). The proteins were electrophoretically transferred to nitrocellulose paper and immunoblotted as described (32, 33), using the immunoglobulin fraction of antiserum purified on either β -PAI affinity columns (above) or on the fusion protein. For the affinity purification of antisera on the fusion protein, 900 μ l of crude extract from induced E. coli lysogens (29) was fractionated by NaDodSO₄/PAGE and transferred to nitrocellulose paper. Strips containing proteins of 150-200 kDa were excised from the nitrocellulose sheets and used for the affinity purification of antisera. Blocking of the nitrocellulose filter strips, binding of specific antibodies, and washings were performed as described for the screening of $\lambda gt11$ libraries with antibody probes (29). To elute bound antibody, the filter strips were incubated twice (3 min each) with 200 μ l of 0.1 M glycine/HCl buffer (pH 2.5) containing 0.02% fetal bovine serum. The eluted material was neutralized by the addition of 140 µl of 0.5 M Tris/HCl (pH 8.0), dialyzed overnight, and used as the primary antibody in immunoblot analysis.

Nucleic Acid Methods. Phage particles were prepared by the plate-lysate method and phage DNA was purified by CsCl equilibrium centrifugation (34). Plasmid DNA was isolated by the method of Birnboim and Doly (35), followed by two consecutive ethidium bromide/CsCl equilibrium centrifugations. Enzyme reactions were carried out according to the conditions suggested by the suppliers. Total RNA was prepared (36) from cultured HT-1080 cells (American Type Culture Collection CCL 121), fractionated by agarose gel electrophoresis in the presence of formaldehyde (37), and subjected to blot hybridization analysis (38).

DNA from λ gt11 clones was digested with *Eco*RI endonuclease, and the excised cDNA insert was subcloned in bacteriophage M13 cloning vector mp9 (39). M13 clones containing the cDNA insert in both orientations were isolated, and deletion libraries of both strands were constructed using the single-stranded M13 method of Dale *et al.* (40). Before sequencing, the size of the M13 templates was determined by electrophoresis in 0.7% agarose gels, and selected templates were sequenced by the dideoxy chaintermination method (41). Both DNA strands were sequenced and more than 80% of each strand was sequenced two or more times. DNA sequence data was processed using the Staden program (42). Homology searches were done by the Pearson fast protein homology program (43).

RESULTS AND DISCUSSION

Identification of Endothelial Cell-Type β -PAI Activity in Placenta. When 20 μ l of the crude placental extract was analyzed for PAI activity by NaDodSO₄/PAGE (30) and reverse fibrin autography (25), two inhibitor zones of 50–55 kDa were revealed (data not shown). Immunoprecipitation

experiments demonstrated that the two inhibitor zones resulted from the presence of both the placental-type PAI (8) and the endothelial cell-type PAI (10–13). Quantitation by radioimmunoassay (44) indicated that the extract contained 270 ng of β -PAI per ml. Since the placental tissue had been extensively washed before extraction, this β -PAI was most likely synthesized by cells contained in placenta and not a serum contaminant. Placenta was therefore employed as a source for the isolation of a cDNA for β -PAI.

Isolation of Human β -PAI cDNA. Approximately 7×10^5 recombinant phages from a \gt11 expression library containing cDNA inserts prepared from human placental mRNA (26) were screened immunologically to obtain cDNAs for the β -PAI. Thirty-four positive clones were obtained, half of which continued to be positive through a second screening. Three positive clones were randomly picked and plaquepurified, and phage DNA was prepared. The phage DNA from the three clones ($\lambda 1.2$, $\lambda 3$, and $\lambda 9.2$) was digested with EcoRI and the cDNA inserts were determined to be 1.9, 3.0, and 1.9 kilobases (kb) long, respectively. The 3.0-kb cDNA insert from $\lambda 3$ was subcloned into a plasmid vector (45), excised with EcoRI, and purified from an agarose gel. The cDNA insert was nick-translated and shown to hybridize with $\lambda 1.2$ and $\lambda 9.2$ DNA under conditions of high stringency, indicating that the DNA inserts in the three clones were related.

Three lines of evidence support the conclusion that the isolated clones code for the human β -PAI. (i) Induction of an *E. coli* lysogenic strain prepared by infecting a high-frequency-of-lysogeny strain (Y1089) with λ 9.2 resulted in the expression of a recombinant fusion protein (180 kDa) that was recognized by an affinity-purified IgG from antiserum raised against the purified BAE β -PAI (Fig. 1, lane B). An *E. coli* strain lysogenic for λ gt11 and thus lacking the cDNA insert did not produce such an immunoreactive protein (Fig. 1, lane C). (*ii*) The 180-kDa recombinant fusion protein and the BAE β -PAI share antigenic epitopes, since affinity purification of the antiserum on the recombinant fusion protein yielded antibodies that recognized the purified BAE β -PAI on immunoblots (Fig. 1, lane D). (*iii*) Analysis of *E. coli*



FIG. 1. Immunoblot analysis of *E. coli* crude extracts. Extracts prepared from induced lysogenic *E. coli* strains were fractionated by NaDodSO₄/PAGE and analyzed by immunoblotting as described in *Materials and Methods*. Lanes A, D, and F: 300 ng of purified BAE β -PAI. Lanes B and E: 50 μ l of extract from strain Y1089 lysogenized with λ 9.2. Lane C: 50 μ l of extract from strain Y1089 lysogenized with λ gt11. For the immunoblotting experiments shown in lanes A-C, affinity-purified IgG against the BAE β -PAI was used as primary antibody. For lanes D-F, the antiserum used was affinitypurified on various proteins bound to nitrocellulose paper: for lane D, antiserum affinity-purified on the BAE β -PAI; and for lane F, antiserum affinity-purified on proteins of 150-200 kDa from the λ gt11 lysogen. The autoradiograms were exposed for 16 hr. Positions of standards (molecular mass in kDa) run in parallel are at left.



FIG. 2. Analysis of inhibitor activity in extracts from induced E. coli lysogens. Extracts from lysogenic E. coli strains were fractionated by NaDodSO₄/PAGE and then analyzed by reverse fibrin autography (lanes A–C) and immunoblotting (lanes D–F). Lanes A and D: 300 ng of purified bovine β -PAI. Lanes B and E: 50 μ l of extract from the strain lysogenic for λ 9.2. Lanes C and F: 50 μ l of extract from the strain lysogenic for λ 911. The autoradiograms for lanes E and F were exposed for 2 weeks.

extracts by NaDodSO₄/PAGE followed by reverse fibrin autography revealed that the λ 9.2 lysogen containing the 1.9-kb insert, but not the λ gt11 lysogen, expresses PAI activity (Fig. 2, lane B). Surprisingly, two PAIs of 180 kDa and 40 kDa, respectively, were present in these extracts. To

1 .1

٨

investigate the relationship of these PAIs, proteins from the lysogens were fractionated by NaDodSO₄/PAGE and again analyzed by immunoblotting, but this time the autoradiograms were developed after a longer exposure (Fig. 2, lanes E and F). Two peptides were detected, and these comigrated with the inhibitor activities (compare lanes B and E). The great majority of the β -PAI antigen produced by the λ 9.2 lysogen was detected at \approx 180 kDa; however, a small amount of antigen was also detected at 40 kDa (Fig. 2, lane E). Since the two PAIs share immunologic and biologic properties, the smaller is most likely derived from the larger by proteolytic processing of the β -galactosidase-PAI fusion protein. The specific activity of the released 40-kDa protein may be higher than that of the larger fusion protein because it is no longer sterically hindered.

Nucleotide Sequence of DNA Coding for Human β -PAI and Assignment of Protein Sequence. The DNA sequence from both strands of the 3.0-kb cDNA insert of clone $\lambda 3$ was established by sequencing deletion subclones constructed by the method of Dale *et al.* (40). Fig. 3 shows the DNA sequence along with the inferred amino acid sequence. This is not a full-length cDNA, since the untranslated region, the initiation codon, and most of the signal peptide are missing from the 5' end of the molecule. To identify the codon coding for the NH₂ terminus of the mature protein, we aligned the deduced amino acid sequence with that of the human and bovine β -PAIs (unpublished data). Based on that alignment, the valine designated number 1 in Fig. 3 is the NH₂-terminal residue of the human β -PAI. The β -PAI is secreted and is therefore likely to contain a signal peptide (46). The signal

| Glu | Gly | Ser | Ala | Val | His | His | Pro | Pro | Ser | Tyr | Val | Ala | His | Leu | Ala | Ser | Asp | Phe | Gly | Val | Arg | Val | Phe | Gln | Gln | Val | Ala | Gln | Ala | 26 |
|--|---|---|--|---|---|---|---|---|---|---|---|--|--|---|---|---|---|---|---|--|---|---|---|---|--|---|--|---|--|---|
| GAA | GGG | TCT | GCT | GTG | CAC | Cat | CCC | CCA | TCC | TAC | GTG | GCC | CAC | CTG | GCC | TCA | GAC | TTC | GGG | GTG | AGG | GTG | TTT | CAG | CAG | GTG | GCG | CAG | GCC | |
| Ser TCC | Lys AAG | Asp GAC | Arg CGC | Asn AAC | Val GTG | Val GTT | Phe TTC | Ser TCA | Pro | Tyr Tat | Gly GGG | Val GTG | Ala GCC | Ser TCG | Val GTG | Leu TTG | Ala GCC | Met ATG | Leu CTC | Gln CAG | Leu CTG | Thr ACA | Thr ACA | Gly G GA | Gly GGA | Glu GAA | Thr ACC | Gln CAG | Gln CAG | 56 |
| Gln | Ile | Gln | Ala | Ala | Met | Gly | Phe | Lys | Ile | Asp | Asp | Lys | Gly | Met | Ala | Pro | Ala | Leu | Arg | His | Leu | Tyr | Lys | Glu | Leu | Met | Gly | Pro | Trp | 86 |
| CAG | ATT | CAA | GCA | GCT | ATG | GGA | TTC | AAG | ATT | GAT | GAC | AAG | GGC | ATG | GCC | CCC | GCC | CTC | CGG | CAT | CTG | TAC | AAG | GAG | CTC | ATG | GGG | CCA | TGG | |
| Asn | Lys | Asp | Glu | Ile | Ser | Thr | Thr | Asp | Ala | Ile | Phe | Val | Gln | Arg | Asp | Leu | Lys | Leu | Val | Gln | Gly | Phe | Met | Pro | His | Phe | Phe | Arg | Leu | 116 |
| AAC | AAG | GAT | GAG | ATC | AGC | ACC | ACA | GAC | GCG | ATC | TTC | GTC | CAG | CGG | Gat | CTG | AAG | CTG | GTC | CAG | GGC | TTC | ATG | CCC | CAC | TTC | TTC | AGG | CTG | |
| Phe | Arg | Ser | Thr | Val | Lys | Gln | Val | Asp | Phe | Ser | Glu | Val | Glu | Arg | Ala | Arg | Phe | Ile | Ile | Asn | Asp | Trp | Val | Lys | Thr | His | Thr | Lys | Gly | 146 |
| TTC | CGG | AGC | ACG | GTC | AAG | C AA | GTG | GAC | TTT | TCA | GAG | GTG | GAG | AGA | GCC | AGA | TTC | ATC | ATC | AAT | GAC | TGG | GTG | AAG | ACA | CAC | ACA | AAA | GGT | |
| Met | Ile | Ser | Asn | Leu | Leu | Gly | Lys | Gly | Ala | Val | Asp | Gln | Leu | Thr | Arg | Leu | Val | Leu | Val | Asn | Ala | Leu | Tyr | Phe | Asn | Gly | Gln | Trp | Lys | 176 |
| ATG | ATC | AGC | AAC | TTG | CTT | GGG | AAA | GGA | GCC | GTG | GAC | CAG | CTG | ACA | CGG | CTG | GTG | CTG | GTG | AAT | GCC | CTC | TAC | TTC | AAC | GGC | CAG | TGG | AAG | |
| Thr | Pro | Phe | Pro | Asp | Ser | Ser | Thr | His | Arg | Arg | Leu | Phe | His | Lys | Ser | Asp | Gly | Ser | Thr | Val | Ser | Val | Pro | Met | Met | Ala | Gln | Thr | Asn | 206 |
| ACT | CCC | TTC | CCC | GAC | TCC | AGC | ACC | CAC | CGC | CGC | CTC | TTC | CAC | AAA | TCA | GAC | GGC | AGC | ACT | GTC | TCT | GTG | CCC | ATG | ATG | GCT | CAG | ACC | AAC | |
| Lys | Phe | Asn | Tyr | Thr | Glu | Phe | Thr | Thr | Pro | Asp | Gly | His | Tyr | Tyr | Asp | Ile | Leu | Glu | Leu | Pro | Tyr | His | Gly | Asp | Thr | Leu | Ser | Met | Phe | 236 |
| AAG | TTC | AAC | TAT | ACT | GAG | TTC | ACC | ACG | CCC | GAT | GGC | Cat | TAC | TAC | GAC | ATC | CTG | GAA | CTG | CCC | TAC | CAC | GGG | GAC | ACC | CTC | AGC | ATG | TTC | |
| Ile | Ala | Ala | Pro | Tyr | Glu | Lys | Glu | Val | Pro | Leu | Ser | Ala | Leu | Thr | Asn | Ile | Leu | Ser | Ala | Gln | Leu | Ile | Ser | His | Trp | Lys | Gly | Asn | Met | 266 |
| ATT | GCT | GCC | CCT | TAT | GAA | AAA | GAG | GTG | CCT | CTC | TCT | GCC | CTC | ACC | AAC | ATT | CTG | AGT | GCC | CAG | CTC | ATC | AGC | CAC | TGG | AAA | GGC | AAC | ATG | |
| Thr | Arg | Leu | Pro | Arg | Leu | Leu | Val | Leu | Pro | Lys | Phe | Ser | Leu | Glu | Thr | Glu | Val | Asp | Leu | Arg | Lys | Pro | Leu | Glu | Asn | Leu | Gly | Met | Thr | 296 |
| ACC | AGG | CTG | CCC | CGC | CTC | CTG | GTT | CTG | CCC | AAG | TTC | TCC | CTG | GAG | ACT | GAA | GTC | GAC | CTC | AGG | AAG | CCC | CTA | GAG | AAC | CTG | GGA | ATG | ACC | |
| Asp | Met | Phe | Arg | Gln | Phe | Gln | Ala | Asp | Phe | Thr | Ser | Leu | Ser | Asp | Gln | Glu | Pro | Leu | His | Val | Ala | Gln | Ala | Leu | Gln | Lys | Val | Lys | Ile | 326 |
| GAC | ATG | TTC | AGA | CAG | TTT | CAG | GCT | GAC | TTC | ACG | AGT | CTT | TCA | GAC | CAA | GAG | CCT | CTC | CAC | GTC | GCG | CAG | GCG | CTG | CAG | AAA | GTG | AAG | ATC | |
| Glu | Val | Asn | Glu | Ser | Gly | Thr | Val | Ala | Ser | Ser | Ser | Thr | Ala | Val | Ile | Val | Ser | Ala | Arg | Met | Ala | Pro | Glu | Glu | Ile | Ile | Met | Asp | Arg | 356 |
| GAG | GTG | AAC | GAG | AGT | GGC | ACG | GTG | GCC | TCC | TCA | TCC | ACA | GCT | GTC | ATA | GTC | TCA | GCC | CGC | ATG | GCC | CCC | GAG | GAG | ATC | ATC | ATG | GAC | AGA | |
| Pro CCC | Phe TTC | Leu CTC | Phe TTT | Val GTG | Val GTC | Arg CGG | His CAC | Asn AAC | Pro CCC | Thr ACA | Gly GGA | Thr ACA | Val GTC | Leu CTT | Phe TTC | Met ATG | Gly GGC | Gln CAA | Val GTG | Met ATG | Glu GAA | Pro CCC | TGA | ccc | TGG | GGA | AAG | ACG | сст | |
| 171 TCATCTGGGACAAAACTGGAGATGCATCGGGAAAGAAAGA | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Glug GAA Serr C Glag C Asn AAC F T C Asn AAC F T C Asn AAC Asn AAC C Lys G AST T C C Asn AC C C Asn AC C C Asn AC C C C C C C C C C C C C C C C C C C | Glu Gly GAA GGG Ser Lys TCC AAG Gln Ile CAG ATT Asn Lys AAC AAG Phe Arg TTC CGG Met Ile ATG ATC CTC CGG ACT CCC Lys Phe AAG TTC CIL SAG TTC ILE AAG TTC CIL AAG TTC CIL AAG TTC CAC ACT CCC ILE AAG TTC CIL CAC ATG GLU VAL GAC CATG ACC ACG CTCTTCAT CCCACACGT CTTTCAT CCCCACACGT CTCTTCAT CTCACCAGT CTCTTCAT CTCTTCAT CTCCCCACACT | Glu Gly Ser GAA GGG TCT GAA GGG TCT Ser Lys Asp TCC AAG GAC Gln Ile Gln CAG ATT CAA Asn Lys Asp AAC AAG GAT Phe Arg Ser TTC CGG AGC Met Ile Ser ATG ATC AGG Thr Pro Phe ACT CCC TTC Lys Phe Asn AAG TTC AAC Ile Ala Ala ATT GCT GCC Thr Arg Leu ACC AGG CTG Asp Met Phe GAC ATG TTC Glu Val Asn GAG GTG AAC Pro Phe Leu CCC TTC CTC TCATCTGGGAC/ AGGAGCATTG' CTCATCAGCAGCAC GCTCTTCACCCC CTACCAGCAC CTACAGCAGG | Glu Gly Ser Ala GAA GGG TCT GCT Ser Lys Asp Arg TCC AAG GAC CGC Gln Ile Gln Ala CAG ATT CAA GCA Asn Lys Asp Glu Ac AAG GAT GAG Phe Arg Ser Thr TTC CGG AGC ACG Met Ile Ser Asn ATG ATC AGC AAC Thr Pro Phe Pro ACT CCC TTC CCC Lys Phe Asn Tyr AAG TTC AAC TAT Ile Ala Ala Pro ACT GCT GCC CCT Thr Arg Leu Pro ACT AGG CTG CCC Asp Met Phe Arg GAC ATG TTC AGA Glu Val Asn Glu GAG GTG AAC GAG Pro Phe Leu Phe CCC TTC CTC TTT TCATCTGGGACACAAA ACGATCTTCATCTCAGG GCTCTTCACCTCCC CAGACGGCAGAGGAG | Glu Gly Ser Ala Val GAA GGG TCT GCT GTG Ser Lys Asp Arg Asn TCC AAG GAC CGC AAC Gln Ile Gln Ala Ala CAG ATT CAA GCA GCT Asn Lys Asp Glu Ile AAC AAG GAT GAG ATC Phe Arg Ser Thr Val TTC CGG AGC ACG GTC Met Ile Ser Asn Leu ATG ATC AGC AAC TTG Thr Pro Phe Pro Asp ACT CCC TTC CCC GAC Lys Phe Asn Tyr Thr ATG GTC GCC CCT TAT Thr Arg Leu Pro Arg ACC AGG CTG CCC CGC Asp Met Phe Arg Gln GAC ATG TTC AGA CAG Glu Val Asn Glu Ser GAG GTG AAC GAG AGT Pro Phe Leu Phe Val CCC TTC CTC TTT GTG TCATCTGGGACAAAACTGG ACAGACTTTGTTGTGCAGCAG GTCTTTAATGAGAGCATGGTATGTGCATG GCTCTTCACCCCCCAATCT TGACTCACCCCCCCCAATCT TGACTCACCCCCCCCCC | Glu Gly Ser Ala Val His GAA GGG TCT GCT GTG CAC Ser Lys Asp Arg Asn Val TCC AAG GAC CGC AAC GTG Gln Ile Gln Ala Ala Met CAG ATT CAA GCA GCT ATG Asn Lys Asp Glu Ile Fer AAC AAG GAT GAG GAT CAGC Phe Arg Ser Thr Val Lys TTC CGG AGC ACG GTC AAG Met Ile Ser Asn Leu Leu ATG ATC AGC AAC TTG CTT Thr Pro Phe Pro Asp Ser ACT CCC TTC CCC GAC TCC Lys Phe Asn Tyr Thr Glu AAG TTC AAC TAT ACT GAG Ile Ala Ala Pro Tyr Glu ATT GCT GCC CCT TAT GAA Thr Arg Leu Pro Arg Leu ACC AGG CTG CCC CGC CTC Asp Met Phe Arg Gln Phe GAC ATG TTC AGA CAG TTT Glu Val Asn Glu Ser Gly GAG GTG AAC GAG AGT GGC Pro Phe Leu Phe Val Val CCC TTC CTC TTC TGTGGTCCCCGGAGAC CTCATCTGGGACAAAACTGGAGAT ACCAAGTTTCATCTAAGGCTGCAGGGTCCA CAGGACCTCTCATCAGGGTAGGGCA TTCAACTCAATCAATTTATTATA AACTCTAATAGAAGGCTAACAGC | Glu Gly Ser Ala Val His His GAA GGG TCT GCT GTG CAC CAT Ser Lys Asp Arg Asn Val Val TCC AAG GAC CGC AAC GTG GTT Gln Ile Gln Ala Ala Met Gly CAG ATT CAA GCA GCA ATG GGA Asn Lys Asp Glu Ile Fer Thr AAC AAG GAT GAG ATC AGC ACC Phe Arg Ser Thr Val Lys Gln TTC CGG AGC ACG GTC AAG CAA Met Ile Ser Asn Leu Leu Gly ATG ATC AGC AAC TTG CTT GGG Thr Pro Phe Pro Asp Ser Ser ACT CCC TTC CCC GAC TCC AGC Lys Phe Asn Tyr Thr Glu Phe AAG TTC GCC CCT TAT GAA AAA Thr Arg Leu Pro Arg Leu Leu ACC AGG CTG CCC CGC CTC CTG Asp Met Phe Arg Gln Phe Gln GAC ATG TTC AGA AGT GGT CAG Pro Phe Leu Phe Val Val Arg CCC TTC CTC GGGAGCACG Pro Phe Leu Phe Val Val Arg CCC TTC CTC GGGAGCACG TCATCTGGGACAAAACTGGAGATGCAT ACCAGATCTTGTGTGCCAGGGTAGTAT GCTGTCCCCCCAAGCATCTGGGGTAGTAT ACCAGAACATTGTGTGCCTGGGGAGAGCATT ACCAGATCTTGTGTGCCCGGGTAGTATT GCTCTCACCCCCCAAGCATGGGCACAGCATCAAGAGCCTTGGGCCCCCGCGCCCCCGCGCCCCCAGAGGTTCCAAGACGAGAGAGCAGATC TTAATCTAAGAGCCTAGGGCACGATT CCAGCAGCATCAGTGCCATGGGCACAAGCCTTCAGGGCCACAGGGTAGGCCCTCGGGGACCACTT AGAGGCTACCAGAGGGTAGCATT AGCAGGCCCCCCCCCC | Glu Gly Ser Ala Val His His Pro GAA GGG TCT GCT GTG CAC CAT CCC Ser Lys Asp Arg Asn Val Val Phe TCC AAG GAC CGC AAC GTG GTT TTC Gln Ile Gln Ala Ala Met Gly Phe CAG ATT CAA GCA GCT ATG GGA TTC Asn Lys Asp Glu Ile Fer Thr Thr AAC AAG GAT GAG ATC AGC ACC ACA Phe Arg Ser Thr Val Lys Gln Val TTC CGG AGC ACG GTC AAG CAA GTG Met Ile Ser Asn Leu Leu Gly Lys ATG ATC AGC AAC TTG CTT GGG AAA Thr Pro Phe Pro Asp Ser Ser Thr ACT CCC TTC CCC GAC TCC AGC ACC Lys Phe Asn Tyr Thr Glu Phe Thr AAG TT GCT GCC CCT TAT GAA AAA GAG Thr Arg Leu Pro Arg Leu Leu Val ACC AGG CTG CCC CGC CTC CTG GTT Asp Met Phe Arg Gln Phe Gln Ala GAC ATG TTC AGA GAG GAT CAG CTT Glu Val Asn Glu Ser Gly Thr Val GAG GTG AAC GAG AGT GGC ACG GTG Pro Phe Leu Phe Val Val Arg His CCC TTC CTC CTC GAGAGATGCATCGGGG ACCAGATCTGTCTCAAGACCTTGGGCATCTGGGC ACCAGACCTTCTCAAGACCTTGGGCACC CCAGACGTTTCTCAAGACCTGGGCACCTGGGC ACCAGACCTTCATCAAGATGTTTTTGGAGAGCCTTGGCCACC CTCAACAGGGTGGCCCCCGCGCCCCCTCTT CAAGGACCAGAGTGGCACGGGGAACCTTGGGC ACCAGACCTTTGTGTGCCCGGAGACCTCGGGG ACCAGACCTTTGTGTGCCCGCGACACTCGGGGA ACCAGAGCTTTGTGTGCCCGCGACACTCGGGGA CCCAGACCTTCGCCCCGCATCTGGGCGACCTGGGC ACCAGACCTTCGAGGCCACTGGGCGACCTGGGC ACCAGACCTTCGAGGCCACGGTG ACCAGACCTTCGAGGCCCCCCCCGCGCCCCCCGC ACCAGACCTTCGAGGCCCCCCCGCGCCCCCCGC ACCAGACCTTCGAGGCCCCCCCCGCGCACATCTGGGC ACCAGACCTTCGAGGCCCCCCCCGCGCCCCCCGC ACCAGACCTTCGAGGCCCCCCCCGCGCCCCCCGC ACCAGACCTTCGGCCCCCCCCGCGCCCCCCGC ACCAGACCTTCGAGGCCCCCCCCGCGCCCCCCCCCC | Glu Gly Ser Ala Val His His Pro Pro GAA GGG TCT GCT GTG GTG CAC CAT CCC CCA Ser Lys Asp Arg Asn Val Val Phe Ser TCC AAG GAC CGC AAC GTG GTT TTC TCA Gln Ile Gln Ala Ala Met Gly Phe Lys CAG ATT CAA GCA GCA ATG GGA TTC AAG Asn Lys Asp Glu Ile Fer Thr Thr Asp AAC AAG GAT GAG ATC AGC ACC ACA GAC Phe Arg Ser Thr Val Lys Gln Val Asp TTC CGG AGC ACG GTC AAG CAA GTG GAC Met Ile Ser Asn Leu Leu Gly Lys Gly ATG ATC AGC AAC TTG CTT GGG AAA GGA Thr Pro Phe Pro Asp Ser Ser Thr His ACT CCC TTC CCC GAC TCC AGC ACC CAC Ile Ala Ala Pro Tyr Glu Phe Thr Thr AAG TTC AGC CCT TAT GAG ATC ACG AGG Thr Arg Leu Pro Arg Leu Leu Val Leu ACT GCT GCC CCC CGC CTC CTG GTT CTG Asp Met Phe Arg Gln Phe Gln Ala Asp GAC ATG TTC AGA AGT GGC ACG GTC GAC Glu Val Asn Glu Ser Gly Thr Val Ala GAG GTG AAC GAG AGT GGC ACG GTG GCC Pro Phe Leu Phe Val Val Arg His Asn CCC TTC CCC CAGTCTGTGGGAAGCACTGTGGCACC AGGGCTGTCCCCCGCTCTGTGGGGAAGGAGCACTGTGCCCCCGCGCCCCGCGCCCCCCCGCCCCCCCGCAAAC TCATCTGGGGACAAAACTGGAGATGCATCGGGGAAGGAGGGT TTCATCTCAGGACAGATGCATGGGCACCAGGCCCCCCGCCCCGCCCCCCCGCCCCCCCC | Glu Gly Ser Ala Val His His Pro Pro Ser GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC Ser Lys Asp Arg Asn Val Val Phe Ser Pro TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC Gln Ile Gln Ala Ala Met Gly Phe Lys Ile CAG ATT CAA GCA GCT ATG GGA TTC AAG ATT Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala AAC AAG GAT GAG ATC AGC ACC ACA GAC GCC Phe Arg Ser Thr Val Lys Gln Val Asp Phe TTC CGG AGC ACG GTC AAG CAA GTG GAC TTT Met Ile Ser Asn Leu Leu Gly Lys Gly Ala ATG ATC AGC AAC TTG CTT GGG AAA GGA GCC Thr Pro Phe Pro Asp Ser Ser Thr His Arg ACT CCC TTC CCC GAC TCC AGC ACC CAC CGC Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro AAG TTC AGC CAT TA GAG ATT CAG GTT CGC CCC Thr Arg Leu Pro Arg Leu Leu Val Leu Pro ACT GCT GCC CCT TAT GAA AAA GAG GTG CCT Thr Arg Leu Pro Arg Leu Leu Val Leu Pro ACC AGG CTG CCC CGC CTC CTG GTT CTG CCC Asp Met Phe Arg Gln Phe Gln Ala Asp Phe GAC ATG TTC AGA GAG GTG GCC TCC Pro Phe Leu Phe Val Val Arg His Asn Pro CCC TTC CTC TTT GTG GTC CGC CAAC CGC CCC TCATCTGGGACAAAACTGGAGATGCATCGGGAAAGAAA ACCAGATCTGTCTCAAGACCTTGGCCTCCTTC GLU Val Asn Glu Ser Gly Thr Val Ala Ser GAG GTG AAC GAG AGT GGC ACC GTG CTCC TCATCTGGGACAAAACTGGAGATGCATCGGGAAAGAAAAACGATGTGCTCTCAAGACCTTGGCCTGCTCTCTCAAGACCTTGGCCTGCCT | Glu Gly Ser Ala Val His His Pro Pro Ser Tyr GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC TAC Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC TAT Gln Ile Gln Ala Ala Met Gly Phe Lys Ile Asp CAG ATT CAA GCA GCT ATG GGA TTC AAG ATT GAT Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile AAC AAG GAT GAG ATC AGG ACC ACA GAC GCG ATC Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser TTC CGG AGC ACG GTC AAG CAA GTG GAC TTT TCA Met Ile Ser Asn Leu Leu Gly Lys Gly Ala Val ATG ATC AGC AAC TTG CTT GGG AAA GGA GCC GTG Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg ACT CCC TTC CCC GAC TCC AGC ACC CAC GCC GCC Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp AAG TTC ACC CAT ACT GAG TTC CAC AGC CCC GAT Thr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys ACC AGG CTG CCC CGC CTC CTG GTT CTG CCC AAG Asp Met Phe Arg Gln Phe Gln Ala Asp Phe Thr GAC ATG TTC AGA GAG TTT CAG GCT GCC CAC Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser GAG GTG AAC GAA GTG GCC CTC CTG Trr Arg Leu Pro Arg Leu Leu Val Lau Pro Lys ACC AGG CTG CCC CGC CTC CTG GTG CCC CAC Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser GAG GTG AAC GAA GAG GGC ACG GTG GCC TCC TCA TCATCTGGGACAAAAACTGGAGATGCATCGGGAAAGAAAAACT ACCAGATCTGTCCCAAGACCTGGGCACTGGGCACCTGCTCCTG GTC TAAC TAT ACTGAGG CACG GTG GCC TCC TA GAG GTG AAC GAG AGT GGC ACG GTG GCC TCC TA Asp Met Phe Arg Gln Phe Gln Ala Asp Phe Thr GAC AGG TA CAG AGT GGC ACG GTG GCC TCC TA Pro Phe Leu Phe Val Val Arg His Asn Pro Thr CCC TCC TTT TG GG TC CGG CAC AAC CCC ACA TCATCTGGGACAAAACTGGGGAAGCATCTGGGGAACGAAGAAGAAACT ACCAGACTGTCTCCAAGAGCTGTGGCACTGGGCACCTGCTTCCAA AGGACCTTTGTGTGCCCGGGCACCAGGGTTACTTGAGGCACTCCAGGCCCTCCTCTT AAGGAGCCTTTGTGTGCCCGGCACCCGCGCCCCCCCGGCCCCTCCTCT AGGAGCCTTTGTGTGCCAGGGCACAAGAGAGAAGAAACT ACCAGGCCGCACAAACTGGGCACCAGGGTTACTTAGGTAGG | Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC TAC GTG Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC TAT GGG Gln Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp CAG ATT CAA GCA GCT ATG GGA TTC AAG ATT GAT GAC Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile Phe AAC AAG GAT GAG ATC AGC ACC ACA GAC GCG ATC TTC Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu TTC CGG AGC ACG GTC AAG CAA GTG GAC TTT TCA GAG Met Ile Ser Asn Leu Leu Gly Lys Gly Ala Val Asp ATG ATC AGC AAC TTG CTT GGG AAA GGA GCC GTG GAC Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Leu ACT CCC TTC CCC GAC TCC AGC ACC CAC GCC CGC CTC Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp Gly AAG TTC AGC CAT TA CT GAG TTC ACC ACG CCC CAG GCC Ile Ala Ala Pro Tyr Glu Lys Glu Val Pro Leu Ser TTr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe ACC AGG CTG CCC CGC CTC CTG GTT CTG CCC AAG TTC Asp Met Phe Arg Gln Phe Gln Ala Asp Phe Thr Ser GAC ATG TTC AGA CAG TTC CAG CAC CAC CCC CAG GT Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser GAG GTG AAC GAG AGT GCG ACC GTG GCC TCC TCC Trc TCT CTC TTC GGG CAC GTG GCC TCC TCC TCC TCC TTC CCC GGC CTC CTG GTT CTG CCC AAG TTC ASp Met Phe Arg Gln Phe Gln Ala Asp Phe Thr Ser GAC ATG TTC AGA CAG GTT CAG GCT GAC TTC ACG AGT Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser GAG GTG AAC GAG AGT GGC ACC GTG GCC TCC TCA TCC Pro Phe Leu Phe Val Val Arg His Asn Pro Thr Gly CCC TTC CTC TTTTG TG GTC CCG CAC CTCC CTC | Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val Ala GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC TAC GTG GCC Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC TAT GGG GTG Gln Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp Lys CAG ATT CAA GCA GCT ATG GGA TTC AAG ATT GAT GAC AAG Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile Phe Val AAC AAG GAT GAG ATC AGC ACC ACA GAC GCG ATC TTC GTC Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val TTC CGG AGC ACG GTC AAG CAA GTG GAC TTT TCA GAG GTG Met Ile Ser Asn Leu Leu Gly Lys Gly Ala Val Asp Gln ATG ATC AGC AAC TTG CTT GGG AAA GGA GCC GTG GAC CAG Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Leu Phe ACT CCC TTC CCC GAC TCC AGC ACC CAC CGC CGC CTC TTC Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp Gly His ATT GCT GCC CCT TAT GAA AAA GAG GTG CCT CTC TCC CTT Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe Ser ACC AGG CTG CCC CGC CTC CTG GTT CTG CCC AAG TTC CCC Asp Met Phe Arg Gln Phe Gln Ala Asp Phe Thr Ser Leu GAC ATG TTC AGA CAG TTT CAG GCT GAC TTC CAC GAG CTT Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser Thr GGT GTC AAC GAA CAG TTT CAG GCT GAC TTC ACG AGT CTC CTC TCC CCC GC CTC CTG GTT CTG CCC AAG TTC TCC Asp Met Phe Arg Gln Phe Gln Ala Asp Phe Thr Ser Leu GAC ATG TTC AGA CAG TTT CAG GCT GAC TTC ACG AGT CTT Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser Thr GAG GTG AAC GAG AGT GGC ACG GTG GCC TCC TCA TCC CAA Pro Phe Leu Phe Val Val Arg His Asn Pro Thr Gly Thr CCC TTC CTC TTT GTG GTC CGG CAC CAC CTC CTCA TCCAAGAAACACGACTGTGCCACTGTGTGCACTCTGGGGACCTTCCACCAGACACTTGCCCCCCCC | Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val Ala His GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC TAC GTG GCC CAC Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC TAT GGG GTG GCC Gln Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp Lys Gly CAG ATT CAA GCA GCT ATG GGA TTC AAG ATT GAT GAC AAG GGC Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile Phe Val Gln AaC AAG GAT GAG ATC AGC ACC ACA GAC GCG ATC TTC GTC CAG Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val Glu TTC CGG AGC ACG GTC AAG CAA GTG GAC TTT TCA GAG GTG GAG Met Ile Ser Asn Leu Leu Gly Lys Gly Ala Val Asp Gln Leu ATG ATC AGC AAC TTG CTT GGG AAA GGA GCC GTG GAC CAG CTG Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Leu Phe His ACT CCC TTC CCC GAC TCC AGC ACC CAC GCC CGC CGC CTC TTC CAC Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp Gly His Tyr AAG TTC AAC TAT ACT GAG TTC ACC ACG CCC GAT GGC CAT TAC Ile Ala Ala Pro Tyr Glu Lys Glu Val Pro Leu Ser Ala Leu ATT GCT GCC CCT TAT GAA AAA GAG GTG CCT CTC TCC CCC AGG CTG CCC CGC CTC CTG GTT CTG CCC AAG TTC TCC CTG Thr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe Ser Leu ACC AGG CTG CCC CGC CTC CTG GTT CTG CCC AAG TTT CAC Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser Thr Ala GAG GTG AAC GAG AGT TC CAG GCC CAAC CCC ACA GGC TTT CCA CTC TCTC TCT CTT GTG GGC ACC CTC TCT ACG ACG CTC TTC ACA GIU Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser Thr Ala GAG GTG AAC GAG AGT GGC ACG GTG GCC TCC TCA TCC ACA GCT Pro Phe Leu Phe Val Val Arg His Asn Pro Thr Gly Thr Val CCC TTC CTC TTT TGGG GTC CGG CACC CTCC TTTACAGAAAACTTGGAGAACAGTCTGGCCAGCTGTGTTATTTGGAGGTCACCTGGCCACCTGCCTCCTCCAGGAAACTCCGAAAACAGTGGAGCGTTATTTGGAGGCACCTTGGCCACCTGCCTCCTCCAGGAAACTCCCGGAAACCAGCC CCCAGAAACAGTGTGCCCGGGACCACTGGGCACCTGCCTCCTCCAGGAAACAGCC CTGAGAGCCAAAACTGGGAGAGGCACTGGGCACCTGGCTGCTCCTCCGGGACCGTGTTATTTGGAGGCCAGATTTCCCCCTCCCCCGCGCCCCCCCC | Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val Ala His Leu GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC TAC GTG GCC CAC CTG Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser TCC AAG GAC CGC AAC GTG GTT TTC TCA CAC CTAT GGG GTG GCC TCG Gln Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp Lys Gly Met CAG ATT CAA GCA GCT ATG GGA TTC AAG ATT GAT GAC AAG GGC ATG Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile Phe Val Gln Arg AAC AAG GAT GAG ATC AGC ACC ACA GAC GGG ATC TTC GTC CAG CGG Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val Glu Arg TTC CGG AGC ACG GTC AAG CAA GTG GAC TTT TCA GAG GTG GAG AGA Met Ile Ser Asn Leu Leu Gly Lys Gly Ala Val Asp Gln Leu Thr ATG ATC AGC AAC TTG CTT GGG AAA GGA GCC GTG GAC CAG CTG ACA Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Ieu Phe His Lys ACT CCC TTC CCC GAC TCC AGC ACC CAC CGC CGC GTC TTC CAC AAA Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp Gly His Tyr Tyr AAG TTC AGC CAT TA CT GAG TTC ACC ACG CCC GAT GGC CAT TAC TAC Thr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe Ser Leu Glu ACC AGG CTG CCC CGC CTC CTG GTT CTG CCC AAG GTT GCC CTC ACC Thr Arg Leu Pro Arg Leu Leu Val Lau Pro Lys Phe Ser Leu Ser Asp GAC ATG TTC AGA CAG TTT CAG GCT GAC TTC CAC GAC Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser Thr Ala Val GAG GTG ACC GAG CACG GTG GCC CTC CTC TCT CCAC GAC GTT AGA CAG ACT TT CAG GCT GAC TTC CAC GAG CTT TCA GAC Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser Thr Ala Val GAG GTG AAC GAG AGT GGC ACG GTG GCC TCC TCA TCC ACA GCT TTCATCTGGGACAAAACTGGAGATGCATCGGGGAAGAAAACTCCGAAGAAAACTCCGAAAAACTGGGGACCTTGGGCACCTCCTCTTTAGGTCAAGGGTTACTTAAGAAAACAGGTGGACCTTGGGCACCCCCCCC | Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val Ala His Leu Ala GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC TAC GTG GCC CAC CTG GCC Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser Val TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC TAT GGG GTG GCC TCG GTG Gln Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp Lys Gly Met Ala CAG ATT CAA GCA GCT ATG GGA TTC AAG ATT GAT GAC AAG GGC ATG GCC Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile Phe Val Gln Arg Asp AAC AAG GAT GAG ATC AGC ACC ACA GAC GCG ATC TTC GTC CAG CGG GGA TTC CGG AGG ACG GTC AAG CAA GTG GAC TTT TCA GAG GTG GAG AGA GCC Met Ile Ser Asn Leu Leu Gly Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala TTC CGG AGC ACC GTC GAG CAA GTG GAC TTT TCA GAG GTG GAG AGA GCC Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Leu Phe His Lys Ser ACT CCC TTC CCC GAC TCC AGC ACC CAC CGC CGC CTT TC CAC AAA TCA Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp Gly His Tyr Tyr Asp AAG TTC AAC TAT ACT GAG TTC ACC ACG CCC GAT GGC CAT TAC TAC GAC Ile Ala Ala Pro Tyr Glu Lys Glu Val Pro Leu Ser Ala Leu Thr Asn ATT GCT GCC CCC TTT CAG AAAA GAG GTG CCT CTC TC CC CAC ACC Thr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe Ser Leu Gla Arr ACC AGG TG CCC CGC CTC TG GTT CGG GAC TTC CCC GAG CCA Asp Met Phe Arg Gln Phe Gln Ala Asp Phe Thr Ser Leu Ser Asp Gln GAC ATG TTC AGA CAG TTT CAG GCT GAC TTC ACG AGT CTT TCA GAC CAA Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser Thr Ala Val Ile GAG GTG AAC GAG AGT GGC CCC GCC TCCTG GAG ACAC TTC ACG AGT CTT TCA GAC CAA GTTC AGA CAG AGT GGC CGG GCC CCC CCC CC CAC ACA GTC GTT TCCA GAAC CAA GCCAAGGTGACCCCCCCCCCCGGGCCCCCCCCCCCC | Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val Ala His Leu Ala Ser GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC TAC GTG GCC CAC CTG GCC TCA Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser Val Leu TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC TAT GGG GTG GCC CCG GTG TG Gln Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp Lys Gly Met Ala Pro CAG ATT CAA GCA GCT ATG GGA TTC AAG ATT GAT GAC AAG GGC CAG GCG GAT CTG Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val Gln Arg Asp Leu AAC AAG GAT GAG ATC AGC ACC ACA GAC GCG ATC TTC GTC CAG CGG GAT CTG Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala Arg TTC CGG AGC ACG GTC AAG CAA GTG GAC TTT CA GAG GTG GAG GAG AGA GCC GAG Met Ile Ser Asn Leu Leu Gly Lys Gly Ala Val Asp Gln Leu Thr Arg Leu ATG ATC AGC AAC TTG CTT GGG AAA GGA GCC GTG GAC CAG CTG ACA CGG CTG Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Leu Phe His Lys Ser Asp ACT CCC TTC CCC GAC TCC AGC ACC CAC CGC CGC TTC CAC AAA TCA GAC Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp Gly His Tyr Tyr Asp Ile ATT GATC AGC CAT TA CT GAG ATC ACC ACG CCC GAT GGC CAT TAC TAC GAC ATT Thr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe Ser Leu Glu Thr Asn Ile ATT GCT GCC CCG CCC CGC CTC GG GTG GCC TTC TCC CCA GAG ACT GAC AGG TTG AGC AGG TG CGG CTC TTC CAC CAA GCT Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser Thr Ala Val Ile Val GAG GTA GAC GAG AGT GGC CCTC GTC CTC CAC CAC GGT GAC ACG AGT CCC CGC CCC GG CTC CTG GAT TCA CC AGC GTG GAC ATT TA AGT GAC AGA GTG GCC CCC CAC GG CTT TCC AGA ACT Thr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe Ser Leu Glu Thr Glu ACC AGG CTG CCC GG CTC CTG GAT CAC ACC GCC GAA GTC TTA GAC AAA Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser Thr Ala Val Ile Val GAG GTA AAC GAG AGT GGC CGC CTC CTC ACC ACC GTG CTA TAC GAG GTC TTCAAACTGAGAACTGGGAAGAGAGAGAGAAACTCCGAAGAACACTCCTTAGGGGTATTCT CCATCAGGGACAAACTGGAGATGCATCGGGGAACTACTGCAATACTCACTAGGGGACATTCT GCCACAGACAGGGTAGGCCTACTGGGGGACCTCTCCTTGGGGACCTCTACTTGGGGCCTGCTCCTTGTGTGCCACTGGGGGCCTCTCCTTGGGGGGGG | Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val Ala His Leu Ala Ser Asp GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC TAC GTG GCC CAC CTG GCC TAC GAC Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser Val Leu Ala TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC TAT GGG GTG GCC TCG GTG TTG GCC Gln Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp Lys Gly Met Ala Pro Ala CAG ATT CAA GCA GCT ATG GGA TTC AAG ATT GAT GAC AAG GGC ATG GCC CCC GCC Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile Phe Val Gln Arg Asp Leu Lys Anc AAG GAT GAG ATC AGC ACA GAG GGG ATC TTC GTC GAG CGG GAT CTG AAG Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala Arg Phe TTC CGG AGC CAG CTC AAG GAG GTG GAC TTT TCA GAG GTG GAG AGA GCC AGT GTA ATTC AGC AAC TTG CTT GGG AAA GGA GCC GTG GAC CAG CTG AAA TCA AGC AAC TTG CTT GGG AAA GGA GCC GTG GAC CAG CTG AAA CCA GG CTG GTG Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Leu Phe His Lys Ser Asp Gly ACT CCC TTC CCC GAC TCC AGC ACC CAC CGC GCC GTC TTC CAC AAA TCA GAC GGC Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp Gly His Tyr Tyr Asp Ile Leu AAG TTC AAC TAT ACT GAG TTC ACC ACG GCC CTC TCT CGC CAA CAC ATT CTG Thr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe Ser Leu Glu Thr Glu Yr GAC AGG CCC CCC CTC TTC GG GAC CAG GCC CTC TCT CG GAG ACA GAT CTG ACA GTG CCC CCC GC CTC TAG GAA AAA GAG GTG CCT CTC TCT GCC CTA CAC AGA GTC Asp Met Phe Arg Gln Phe Gln Ala Asp Phe Thr Ser Leu Ser Asp Gln Glu Pro GAC ATG TTC AGA GAT GTC GGG CAG GGC CTC TCT CCC GAG GCC CTA TG GTG GCC CTT CTC TTT GTG GTC CGG GCA CTC CCC CAC GGT CTT TCA GAC CAA GAG CTT GU Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser Thr Ala Val Ile Val Ser AGG GTG AAC GGA GTG GGC CCC CTC TCT CTC GCC CTA GTG GTG CTT GCT TTG GTG GTC CGG GG CAC GGG GCC CCC CTC TCT CTC CCCCATA GGGC TCATCTGGGACAAGATGTGGCTCTGGGCACCGGTGTGTGT | Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val Ala His Leu Ala Ser Asp Phe GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC TAC GGG GCC CAC CTG GGC TCA GAC TTC Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser Val Leu Ala Met TCC AAG GAC CGC AGC GTG TTT TC TCA CCC CC TAT GGG GTG GCC TCG GTG TTG GCC ATG Gln Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp Lys Gly Met Ala Pro Ala Leu CAG ATT CAA GCA GCT ATG GGA TTC TAA CAC GAC GGA TG GCC CCC GCC CTC Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile Phe Val Gln Arg Asp Leu Lys Leu ACA AAG GAT GAG ATC AGC ACC ACA GCA GCG ATC TTC GTC CAG GGG GAT CTG AAG CTG Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala Arg Phe Ile TTC CGG AGC ACG GTC AAG CAA GTG GAC CTT TCA GAG GTG GAG AGA GCC AGA TTC ATC Met Ile Ser Asn Leu Leu Gly Lys Gly Ala Val Asp Gln Leu Thr Arg Leu Val Leu ATG ATC AGC ACT TTG CTG GG AAA GGA GCC GTG GAC CGG CTG GTG GTG GTA CTG CAC CTT CCC GAC TCC AGC ACC CGC CGC CTC TTC CAC AAA TCA GAC GGC GGC GTG Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Leu Phe His Lys Ser Asp Gly Ser ACT CCC TTC CCC GAC TCC AGC ACC CAC CGC CGC CTC TTC CAC AAA TCA GAC GGC AGC Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp Gly His Tyr Tyr Asp Ile Leu Glu AAG TTC AAC TAT ACT GAG TTC CAC AGC CCC GAT GGC CAT TAC TAC GAC ATC CTG GAT Thr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe Ser Leu Glu Thr Asn Ile Leu SA ATT GCT GCC CCT TAT GAA AAA GGT GCT CTC TC CAC GAG CAC CAC CAC CTC CAC CG CC CTC TT GGA ATT CAG GCT GCT CTC ACA GGT CTT TTC AGC GAC AGG CTC CTC Glu Val Asn Glu Ser Gly Thr Val Ala Ser Ser Thr Ala Val Ile Val Ser Ala GAG GTG AAC GAG AGT GGC ACG GTG GCC TCC TTC ACA AGT CTT TTTA AGC CAA GGC CTC CTC CTTC TTT GTG GTC CGG CACA CCC GCA CACA GGC CTC TTTC ATA GTC ATA GTC CACAGACT GGCCACTCTCCCCGGAGCCCTCCCCAGGGGGGCAGGTGGGCAGCATTGGCCTCCCCGGGGCCCACTGCCGCCCCCCCC | Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val Ala His Leu Ala Ser Asp Phe Gly GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC TAC GTG GCC CAC CTG GCC TCA GAC TTC GGG Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser Val Leu Ala Met Leu TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC CCT ATG GGG GGG CCC GCG GTG TGG GCC CCC Gln Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp Lys Gly Met Ala Pro Ala Leu Arg CAG ATT CAA GCA GCT ATG GGA TTC TAA CAC GAC GAT GAC AAG GGC ATG GCC CCC GCC CTC CGG Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile Phe Val Gln Arg Asp Leu Lys Leu Val AAC AAG GAT GAG ATC ACA CAC ACA GAC GGC GT CTC GTC CAG GGG GAT TCG AAG CTG GTC Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala Arg Phe Ile Ile TTC CGG AGC ACG GTC AAG CAA GTG GAC TTT TCA GAG GTG GAG AGA GCC AGA TTC ATC ATC ATC ATC AGC ACT TG CTT GGG AAA GGA GCC GTG GAC CAG CTG ACA CGG CTG GTG GTG Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Leu Phe His Lys Ser Asp Gly Ser Thr ATG ATC AGC AAC TG CTC AGG ACC CAC CGC CGC CTC TTC CAC AAA TCA GAC GGC AGC ACT Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp Gly His Tyr Tyr Asp Ile Leu Glu Leu AAG TTC AAC TTA ACT GAG TTC ACC ACG CCC GAT GGC CAT TAC CAC AAT CTG GAA CTG Thr ATA CT GGC CTT TG AAA AGA GGG TC CTC CTC TC TC CAC AAA TCA GAC ATC CTG GAA CTG CC CC TTT GA GAA AGA GGG TC CC CTC TC TC CAC AAA TCA GAC ATC CTG GAA CTG Ile Ala Ala Pro Tyr Glu Lys Glu Val Pro Lys Phe Ser Leu Glu Thr Glu Val Asp Leu ACC AGG CTG CCC GGC CTC CTG GTT CTG CCC CAC AGC CAT CTC AGC CGC CTTC AAC AGA ATT CAG GGG CAC GTG GCC TCC TTA CTA AGC CAA AGC CTC CACC CCC TTC CTC TTT GTG GTC GGG CAC AGC GTC CTC TTC CAC AGC AGA AGT CTG CTC CACC GGC TTC ATA CAG GAG TGGC CTC CTC TTC CCC AGC ATA GTC TAC AGC CGC CTC CCC TTC TTG GTG GCC AGC GGC GCC CCC CAC AGG CTT TAC TAC GAC AGG CTG CTC CCCC CGG GTG CCC CGG CTC CTG GTT CTG CCC AAA GTC CTT TTC AGA GAG CTG CTC CCCC CCC GGC CTC CTG GTT CTG GCC GGC GCC CCC C | Glu Gly Ser Als Val His His Pro Pro Ser Tyr Val Ala His Leu Ala Ser Asp Phe Gly Val GAA GGG TCT GCT GTG CAC CAT CCC CCA TCC TAC GGT GCC CAC CTG GCC TCA GAC TTC GGG GTG Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser Val Leu Ala Net Leu Gln TCC AAG GAC GCC AAC GTG GTT TTC TCA CCC TAT GGG GTG GCC TCG GTTTG GCC ATG CTC CAG Gln Ile Gln Als Ala Met Gly Phe Lys Ile Asp Asp Lys Gly Net Ala Pro Ala Leu Arg His CAG ATT CAA GCA GCT ATG GGA TTC TAA GCA GCA GCA GCG GCA CTG GCC CTC CGC GCC CTC CGG CAT Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile Phe Val Gln Arg Asp Leu Lys Leu Val Gln AC AAG GTA GA GTA GAC ACC ACA GAC GCG GCA TCT TCC GTC CAG GGA GTG GTC CAG Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala Arg Phe Ile Ile Asn TTC CGG AGC ACC GTC CAG GAC GCG GCA CTTC CAC GAC GGA CTG GTC ATC ATC ATC ATT ATC AGC AAC TTO CTT GGG AAA GGA GCC GTG GAC CAG CTG ACA GG CTG GTG GTG GTG ATT ATC AGC AAC TTO CTT GGG AAA GGA GCC GTG GAC CGC GTG ACA GG CTG GTG GTG GTG AT Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Leu Phe His Lys Ser Asp Gly Ser Thr Val ACT CCC TTC CCC GAC TCC AGC ACC CAC CGC GG GG CCT TTC CAC AAA ATCA GAC GGC AGC ACC GTC CCC Ile Ala Ala Pro Tyr Glu Lys Glu Val Pro Leu Ser Ala Leu Thr Asn Ile Leu Ser Ala Gln ATT GTG GCC CCT TAT GAA AAA GAG GTG CCT CTC TCT GCC CTT ACC AAC ATT CTG AGT GCC CAG Thr Arg Leu Pro Arg Leu Leu Val Leu Val Pro Leu Ser Ala Leu Thr Asn Ile Leu Ser Ala Gln ATT GTG GCC CCT TAT GAA AAA GAG GTG CCT CTC TCT GCC CTC ACC AAC ATT CTG AGT GCC CAG GLU Val Asn Glu Ser Gly Thr Val Ala Ser Ser Thr Ala Val Ile Val Asp Leu Arg GAA GTG AAC GAA CAT TT CAG GTG GAG CTCC TCC ACA GGT CTT TCA GAC CAA GAG CTC CTC GG GLU Val Asn Glu Ser Gly Thr Val Ala Ser Ser Ser Thr Ala Val Ile Val Ser Ala Arg Met Fro Phe Leu Phe Val Val Arg His Asn Pro Thr Gly Thr Val Leu Phe Met Gly Gln Val Met CCC TTC CTC TTT GTG GTC CG GG GCC CCC CTC GTG GTG | Glu Gly Ser Ala val His His Pro Pro Ser Tyr Val Ala His Leu Ala Ser Asp Phe Gly Val Arg GAA GGG TCT GCT GTG GAC CAT CCC CCA TCC TAC GTG GCC CCTG GCC GAC TTC GGC GGC TAG GTG GCC GGC AAC GTG GTT TTC TCA CCC TAT GGG GTG GCC TCG GTG TTG GCC ATG GTC CAG GTG GIn 11e Gin Ala Ala Het Gly Phe Lys 11e Asp Asp Lys Gly Net Ala Pro Ala Leu Arg His Leu CAG ATT CAA GCA GTG GAT TCA AG ATT GAT GAC AAG GGC ATG GCC CCC GCC CCC GGG CAT CTG AST CYA AG GAT GAG ATC GAG ATT AAG ATT GAT GAC AAG GGC ATG GTC CAAG GCC CCC CCC CGG CCC CGG CCC CGG GG TT CAAG GAT GAG ATC AGC ACC ACA GAC GCG ATC TTC GTC CAG GGG GAT CTG AAG CTG GTC CAAG GGC Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala Arg Phe 11e 11e Asn Asp TTC CGG AGC CAC GTC AG CAA GGA GCC GTG GAC CTG GAG GAG GCC CAG GTG GTG GTG GTG AAT GAC ATC ACC ACC TTG CTT GGG AAA GGA GCC GTG GAC CAG GTG GAA GCC AGA TTC ATC ATC AAT GAC Met 11e Ser Asn Leu Leu Gly Lys Gly Ala Val Asp Gln Leu Thr Arg Leu Val Leu Val Asn Ala ATG ATC AGC AAC TTG CTT GGG AAA GGA GCC GTG GAC CAG CTG ACA CGG CTG GTG GTG GTG GTG ATT GCC Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Leu Phe His Lys Ser Asp Gly Ser Thr Val Ser ACT CCC TTC CCC GAC CAC CAC CGC CCC GTC GTC TTC CAC AAA TTCG AGC GGC ACT GTC TTC ATC ACC CCC GCA GTC AGA AGG GGG CCC TCT GT GCC CAA CAT CTG GAA CTG CCC TAC Thr Arg Leu Pro Arg Leu Leu Val Lys Glu Val Pro Leu Ser Ala Leu Thr Asn I1e Leu Ser Ala GIn Leu ATT GCT GGC GTC TTC AGA AAA GGG GGC CTC TCT CT GCC CCA GAC ATT CTA GAG GCC AGC CTC CTC AGG CCC TTAT ACT GAG ATC ACC AGG GGC CTC TCT CT CCC CG CC CTA CAC AAT CTG GAC CCC CGC CTG CTG CTG CTG CTG CCC CAAC ATT CTA GAG CGC CTC TCC CC GGC CTC TCC CTG CGC CTC TCC CG GGC CTC TCC CC GGC CTC TCC CTG GGC CCC CTC CCC GGC CTC CT | Glu Gly Ser Als Val His His Pro Pro Ser Tyr Val Ala His Leu Ala Ser Asp Phe Gly Val Arg Val GAA GGG TCT GG CAC CAT CCC CCA TCC TAC GTG GCC CCG GCC GTG CTTG GCC TAG GG GTG AGG GTG Ser Lys Asp Arg Arn Val Val Phe Ser Pro Tyr Gly Val Ala Ser Val Leu Ala Met Leu Gln Leu Thr TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC TAT GGG GTG GCC TCG GTG TTG GCC ATG CTC CAG GTG ACA Gln 11e Gln Ala Ala Met Gly Phe Lys 11e Asp Asp Lys Gly Met Ala Pro Ala Leu Arg His Leu Tyr CAG ATT CAA GCA GTG ATT CAA GA TT GAT GAC ANG GGC ATG GGC CCC GCC CTC CGG CAT CTG TAC Asn Lys Asp Glu 11e Fer Thr Thr Asp Ala 11e Phe Val Gln Arg Asp Leu Lys Leu Val Gln Gly Phe AAC AAG GAT GAG ATC AGC ACC ACA GAC GCG ATC TTC GTC CAG GGG GAT CTG AAG CTG GTC CAG GGC TTC The Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala Arg Phe 11e 11e Asn Asp Trp TCC GA AGC ACC GTC AG CAA GTG GGC CTT TTC AGA GTG GAC AGA GCC AGA TTC ATC ATC ATC ATC AC AGC ACC TTT C GG AGC ACC GTC CAG CAC CCAC GCC GTG GAC CAG CTG GTG CTG GTG CTG GTG AAT GCC CTC Thr Pro Phe Pro Asp Ser Ser Thr His Arg Arg Leu Phe His Lys Ser Asp Gly Ser Thr Val Ser Val ATT CAC CCC GAC TCC AGC ACC CAC GC GC GCC TTC TC CAC AAA TCA GAC GGC ACT CTC TGT GLys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp Gly His Tyr Tyr Asp 11e Leu Glu Leu Pro Tyr His AAT TCA ACC CAG TTC AGG AGG GTC CTC TCT GTG CTA ACA ACT GTG GTG ACT GCC TAC CAC ACT CCC TAC CAG GAC GTC CAC CAC GC CCC TCT CT CTG CAC ACA ATC CTG GAA CTG GCC CTA CC ACC CCC TAC CAG GAC GTC CTC GTC CTC CTC CTC CTC GGC CTA TCC AGC AATC GTG GAC CTG CTC ATC ATT GTG GGC CCC TTA ACT GAG GTG CTC CTC TCT CTC CC GGA GAC GTG GTC ATC CAC CAC CTC ACC CGC GTG CTC GTG CTG CTC CTC TCT CTC GGC CTA ACC CTA GC CTA CC CAC AGC GTG CCC CGC CTC GTG TTC CCC CC CTC CT | Clu Gly Ser Als val His His Pro Pro Ser Tyr Val Ala His Leu Ala Ser Asp Phe Gly Val Arg Val Phe GAA GGG TCT GTG CAC CAT CCC CCA TCC TAC GTG GCC CAC CTG GCC TCA GAC TTC GGG GTG GTG TTT CC GAG GAC CGC AAC GTG GTT TTC TCA CCC TAT GGG GTG GCC TCG GTC TTG GGC ATG CTC CAG CTG ACA ACA GIN Ile GIN Ala Ala Het Gly Phe Lys Ile Asp Asp Lys Gly Met Ala Pro Ala Leu Arg His Leu Tyr Lys CCA GAT CAA GCA GCT ATG GGA TTC AGA ATT GAT GAC AMG GGC ATG GCC CCC GCC CC GC CCT CCG GC ATC ATG AAG GAT GAG ATT AGG GAT TTC AGA ATT GAT GAC AMG GGC ATG GCC CCC GCC CC GC CCC GC CAG CTG TTC CAG GAT CAAG ATC AGG ACC ACA GAC GGC ATT TCA GTC CAG GGG GAT CTG AGG GTG GTC CAG GGG TTC ATG AGG GAT GAG ATC AGC ACC ACA GAC GCG ATC TTC GTC CAG GGG GAT CTG AGG CTG GTG CGA GGC TTC ATG TTC CGG AGC AGC GTC AAGC ACC CAC GAC GCG ATC TTC GTC CAG CGG GAT CTG AGG GTG GTG CAA GGG CTC ATG TTC CGG AGC AGC GTC AGG CAC TTC CAG GGG GAA GAG AGA GCC GTG GTG GTG AAT GAG GTG GTG ATT AGC AAC ATG GTT GGG GAA GGG CTT TTCA GAG GTG GGA AGA GCC GTG GTG GTG GTA ATG GAC GAG TTC AGC AGC ATG GTT GGG AAA GGA GGC GTG GTG GTG AAT CAAC ATC ATC GAC GGC CTT TCC GAC CAG CGG CTG GTG GTG GTG ATA GGC CTC TTC ATG ATC AGC AAC TTG GTT GGG AAA GGA GGC CTT TTC ACA AAA TCA GAC GGC GAC ATC TCT GTG CCC TTC CCC TTC CCC GAC TCC AGC ACC CAC CGC CGC CTC TTC CAC AAA TCA GAC GGC ACT GTC TCT GTG CCC Iys Phe Aan Tyr Th Glu Phe Thr Thr Pro Asp Gly His Tyr Tyr Asp Ile Leu Glu Lue Pro Tyr His Gly AGG TTC AAG CAAC TTC CTA GG TC ACC CAC CGC CCT CTC TCT GG CCC CAC ACA CTT CG GAC CTA CAC GG Thr Arg Leu Pro Arg Leu Luv SI Luv Pro Lys Phe Ser Leu Glu Thr Asn Ile Luu Sar Ala GIn Leu 11e Ser ATT GCT GCC CCC CTT TTG GAA GGG GT TCT CCT TCT GG GC CTA CAC AAT CTG GAG CTC CAC GAG GCC CTT AGC AGG CTT ATG AAA GAG GTG CCT CTC TCT GG CCC CAC CAC AATT CTG GAG CTC CCC CAC CCC ACC TTC ATG CCC CCC CTT TTG GAA AGG GGT CCT TCT CTG GG CCT GAG CTC CAC GAG CCC CTA CAC CCC CTA CGC CTT CTG GG CCC CTC CTC GG GT CGC CCC CC | Clu Gly Ser Als val His His Pro Pro Ser Tyr Val Ala His Leu Ala Ser Asp Phe Gly Val Arg Val Phe Gln GAA GGG TCT GTG CAC CAC CCC CA TCC TAC GTG GGC CAC CTG GGC TTC GGG GTG GGG AGG GTG TTT CAG Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser Val Leu Ala Met Leu Gln Leu ThT Thr Gly TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC TAT GGG GTG GCC TCG GTG TTG GCC ATG CTC CAG CTG ACA ACA GGA Gln Ile Gln Ala Ala Net Gly Phe Lys Ile Asp Asp Lys Gly Met Ala Pro Ala Euu Arg His Leu Tyr Lys Glu CAG ATT CAA GCA GT ATG GGA TTC AGA ATT GAT GAC AMG GGC ATG GCC CCC GCC CTC CGC CAT CTG TAC AAGA GGA Asn Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile Phe Val Gln Arg Asp Leu Lys Leu Val Gln Gly Phe Met Pro ACA GG GT GGG ATC AGC ACC ACA GAC GCC GTC TTC GTC CAC GG GAT CTG AGC GTG CTG CAG GGC TTC ATG GCC Phe Arg Ser Thr Val Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala Arg Phe Ile Ile Asn Asp Trp Val Lys TTC CGG AGC AGC GTG CAAC CAC GGC GTG TTC CAG GGG GAA GG CCG GTG GTG GTG AAT GGC CTC TTC TTC CGG AGC AGC GTG GTA AGG GGC TTT TCC AGG GTG GGA AGA GCC GTG GTG GTG GTG ATA GGC ATC TTC Thr Pro Phe PtO Asp Ser Ser Thr His Arg Arg Leu Phe His Lys Ser Asp Gly Ser Thr Val Ser Val Pro Met ATC CCC TTC CCC GAC TCC AGC CAC CGC CGC CTC TTC CAC AAA TCA GAC GGC GGC GTC TTC GGG CAC CTG AGG TTC AGG GTA GGG GTG CTC TTC GGC GTC GTC GTG GAA CTG CCC AGG CCC ATC GGC CCC TTC CCC GGC CTC TTC AGC GGC CTC TTC CAC AAA TCA GGC GCC CTC GTC CCC GGC CAC GCC CTC CCC GGC CTC TAT GAA AAG GG GTC CTC TC TC GCC CTA CAC CAC CGG GGC CCC TC AGG TTA AGC AAA GTG GTC ACC CAC CGC CGC CTC TTC CAC AAA TCA GAC GGC CCC GTC TTC CAC GGG GCC Thr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe Ser Leu Glu Th Glu Val Asp Leu Arg Lys Pro Leu Glu AGC AGT GTC CTC GG TTA GGA GTG GCC TTC TC GC GAC CAC AAC ATT CTC GGA GCC CAG GGG CTC TAC AGG GTG AAC GG GTG GCC TCC TCT CTC GG GG CCC TC TCC GG GGC GCC GTC GGC GG | The Gly Ser Als Val His His Pro Pro Ser Tyr Val Als His Leu Ala Ser Asp Phe Gly Val Arg Val Phe Gln Gln GAA GGG TCT GGT GTG CAC CAC CC CA TCC TAC GTG GGC CAC CTG GGC TCA GAC TTC GGG GTG TTT CAG CAG Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser Val Leu Ala Net Leu Gln Leu Thr Thr Gly Gly TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC CAT TGG GTG GGC CTG GTG TTG GGC ATG CTC CAG CTA AAA GAG GGA Gln Ile Gln Ala Ala Net Gly Phe Lys Ile Asp Asp Lys Gly Net Ala Pro Ala Leu Arg His Leu Tyr Lys Glu Leu CGA ATT CAA GGA GGT ATT GGGA TTC AGA CAA GGA GGA ATG GCC CCC GCC CTC CGG CAT CTG TAC AAG GAG GGC ASA ATG GAG GAT TATG GGA TTC AGA CAT GGG CAT GTC CAG CGG GGT GTG CAG GGG CTT C ATG CAAG GGA GGA ASA AMG GAT GAG ATC AGC ACA GAG GGG ATC TTC GTG CGC GGG GGAT CTG AGA GGG GGT TC ATG CCC CAC ASA CAAG GAT GAG ATC AGC ACA GAG GGG CAT TTC GTG CGC GGG GGAT CTG AGA GGG GGT GTG CAG GGG GTG AAG AAC AAG GAT GAG ATC AGC ACA GG GG CAT TT TCA GAG GTG GAG GGA GCT AATG ATC ATC ATC ATC GGG AGA ACA Met Ile Ser Asn Leu Leu Gly Lys Gly Ala Val Asp Gln Leu Thr Arg Leu Val Leu Val Asn Ala Leu Tyr Phe Asn ATG ATC AGC AAC TTG CG AAA GGA GC CGC CGC CTC TTC CAA ATCA AGC GGG GAG CAT GTT CTTG TGG CCC ATG ATG ATC AGC CAC TG CC AGC CAC CGC CGC CGC CTC TTC CAA ATCA GAG GGA GCT GTT GTG TCC TAC TTC AAA ATG ATC AGC GAC TTT GGA ATG GGA CGC GCC CGC CTC TTC CAC AAT CAGC GGC GGA CGC TTC TGT GTG CCC ATG ATG ATG ATC AGC CAC TTG CTA AGC AGC GC CGC CGC CTC TTC CAC AAT CAGC GGC GGA CAC CT TG TGT GTG CCC ATG ATG ATG ATG CCC GCC CTT AGG GAT GGG CCT CT TC CAC CACA CAC CTG GAA CTG CTC GTG TC GTG TCC TAC GC CAC CAC AGG GTG AGA CAC GG TTC AGG GAC GGG CCT CT TC CCC CAC AAT CTG GAG CAG CT TTC TG GGC CAC CAG ATT GCC GCC CTT AT ATG GGA GGG GTC CT CT CT CCC CCA CAC CTG GAC CTG GGC CAC CC TG ATT GCG GCC CCC TAT AAT GAG GGG GCC CT CT TC CT CC CAC CAC ATT CTAC GAC CAC CGG GG CCC CT CC CAC CAC CGG GG CCC CCC CAC GGA GGG CT CT CC CC CAC GGG GGC CCC CCC CAC GGG GGG CT CT CT CC CAC CAC CGG CGC CCC CCC CGC GCC CGC CG | Clu gly ser Âla Yal His His Pro Pro Ser Tyr Val Als His Leu Ala Ser Asp Phe Gly Val Arg Val Phe Gin Gin Val GAA GGG TCT GCT GTG CAC CAT CCC TAC GTG GCC CAC CTG GCC TC GGC TTA GGG GTG GTT CAG GTG Ser Lys Asp Arg Aan Val Val Phe Ser Pro Tyr Gly Val Ala Ser Val Leu Ala Het Leu Gin Leu Thr Th Gly Gly Glu TCC AAG GAC CGC AAC GTG GTT TTC TCA CCC TAT GGG GTG GGC CTG GTG TTG GCC ATG CTC CAG CTG ACA ACA GGA GGA GAA Gin Ile Gin Ala Ala Net Gly Phe Lys Ile Asp Asp Lys Gly Net Ala Pro Ala Leu Arg His Leu Tyr Lys Glu Leu Met CAG ATT CAA GGA GCT ATG GGA TTC MA GA ATG GAT GAC AGG GGC ATG GCC CCG GC CTC CGG CTT CTG TAA GAG GAC TCA TA AGG GAT GAG ATC AGG CAC CAC GAC GGG ATC TTC GTC CAG GGG GTG GGA GTG GAAG GCC TC CAG CTG TC TG TC CAG GCG ATT Lys Asp Glu Ile Fer Thr Thr Asp Ala Ile Phe Val Gin Arg Asp Leu Lys Leu Val Gin Gly Phe Met Pro His Phe AAC AAG GAT GAG ATC AGG CAC ACA GAC GGG ATC TTT CA GAG GTG GAG AGA GCC AGG TTC ATC ATC ATC ATC GAC GGG TTC ATG CCC CAC TTC CGG AGC ACG GTC AAG CAC GAC GGC ATT TCA GAG GTG GAG AGA GCC AGG TTC ATC ATC ATC ATT CAT CAT GAC CGG TTT AGA GAG GAG GAC GAC GG GTG GAG AGA GCC CAG GAC GAC GTC GTG GTG GTG GTG GTG GTG GTG AGA ACA CAC Met Ile Ser Asn Leu Leu Gly Lys Gly Ala Val Asp Gln Leu Thr Arg Leu Val Leu Val Asn Ala Leu Tyr Phe Aan Gly ATG ATC AGC AAC TTG GGA ATG GGA CC CGC CGC CTT CTC CAA ATC AAT CAC GAC GTG CTT GTG CCC ATG TTG CTC TAC TTT AGG CAAC GTG CAC GAC CGC GGC CCC TTC CAC AAT TAC ATC GAC GGA CTG GTG GTG GTG ATG ATG GTT CAC TTA AGG TTC AGC TAT ACT GAG TTC ACC AGC GCC GGC CTT TTC CAC AAT TAC TAC GAC GGA CTG GCC CAC CGC GGA CACC CTC Ile Ala Ala Pro Tyr Glu Phe Thr Thr Pro Asp Gly His Tyr Asp Ile Leu Glu Leu Pro Tyr His Gly Asp Thr Leu AGG TTC AGC CTT TAG GAA AAA GAG GT GC CTC TT GT GC CCC ACC ACC AAA TTC GTA GG CCC TAC GAC GGA CACC CTC GAG GTA ACC GTT ATA CT GAG TTC ACC AAG GT CTT CTA GAG GTG GAA GTG GAC CTG CAG GAA ACC CTC GAG GTA ACC GTA TAC GAG GTT CACC AAG GT CTT CTA GGA GTG GAA GTG GAC GCC CTA GGA CACC CTC GAG GTA CCC GCC CCC TTA GGA TTA GC CCC CCC GAA GGC CTT TCA GAA GAG GCC CTA GGA CTT | Clu Gly set Âl val His His Pro Pro Set Tyr Val Ala His Leu Ala Ser Asp Phe Gly Val Arg Val Arg Val Arg Gro Gro TT CGG CGC CAC CTC CGG GC CAC CTC GGG GT AGG GTG TT CGG CGG GT GTG GG GGG TT CGG CGG | Clu gly set lie vil wil His His Pro Pro Set Tyr Val Als His Leu Als Set Asp Phe Gly Val Arg Val Phe Gin Gin Val Als Gin GAA GG TCT GCT GTG CAC CAT CCC CAC TCC TAC GTG GCC CAC CTG GCC TCA GAC TTC GGG GTG GTT TTC TAG CAG GTG GCC GC GCT GCT GTG CAC CAT CCC CAC TCC TAC GTG GCC CAC CTG GCC TCA GAC TTC GGG GTG GTT TTC TAG ALG CAG GTG TTC AAG GAA CGT GAT GGA TTC TAC CCC TAT GGG GTG GCC TCG GTG TTG GCC CAC CTC CAG CACA GGA GGA GAA ACC CAG ACG AAT CAA GCA GCT ATG GGA TTC AAG ATT GAT GAC AAG GGC CTC GTG GTT TG CCC CAG CTC TAC CAG GGC TC TAG GGG GTG GTC ATG GGA ATTC AAG GAA GGG ATT GAT GAC AAG GGC GCC CTC GGG GCT CTG CAG CTC TAC CAG GGC CCC TC CGG GAT TTA AAG CAG CTC TTC AAG GAC GCC TAT GGG ATT CAAG GCA GCT ATTG GGA GTC ATG GGA TTC AAG GCA GCT ATG GGA TTC AAG GGA TTC GAC GGG GTC TTC TC GTC CAG CGG GCC CTC CGG GC TTC TAC GGC GCC CCC TC CC CC ACT CAA GCA ACA CAC ACA CAG GGG ATT TTA GG GGG GGA GTG AAG CTG GTC CAG CC CAC TTC TC AAG PAC AAG GAT CAA CAC ACC ACA GAG GGG ATTT TAC GGC GGA CTG AAG GC GTC ATG CAC AGA GAC CAC ACA ACA ASA CAG GAT AG CAC ACA CAC AG GGG ATT TAC AG GTG GAG AGA GCC AGA TTC ATC AAT GAC TGG GTG AAG CAC ACA ACA AGA GAT TA CAC ACA CAA CTG GGC ATT TC CAG GGG GGA GCA AGA GCC AGA TTC ATC AAT GAC TGG GTG AAG CAC ACA ACA AGA GAT TTA CTT GGT AAA GGA GCC GTG GAC CAG CTG ACA CG GCTG GTG TTC TAC ATT GAC TGC TTC CT CAC GGC CAG TGG TTC PTO PHE PTO ABE SET SET TH HIS AFG AFG JELU PHE HIS LYS SET ABE CLU YL LYU ALL LEU VAL ATT ACT GAG GAC CAC CAC ACA CAC ACA AGA GTT ACT CAT TAC TAG GAG ATC AGG GGG GGC CTC TTC CC CAG ATC CAC GGG ACA CCC CTG GGC CAC CCC CGG CAC CCC CCC GGG CAC CCC CC | Clu gly set fie val wal val me me pro Pro Set Tyr val Ala His Leu Als Ser Amp Phe gly val Arg val Phe gln gln val Ala gln Ala GAA GGG TCT GCT GTG CAT CCT CC TAC GTG GGC CAC CTG GCC TCA GAC TTC GGG GTG AGG GTG GGG GGG GGG GGG GGG GGG |

FIG. 3. Complete nucleotide sequence of the cDNA insert of λ 3. The nucleotide sequence of the coding strand and the corresponding predicted amino acid sequence are shown. The numbers at left refer to the position of the nucleotides, and those at right, to the position of the amino acids. The value designated as number 1 is the amino terminus, and the amino acid sequence corresponding to the mature protein is numbered 1–379. The amino acid sequence that corresponds to a portion of the signal peptide is represented by the minus numbers in the opposite direction.

Biochemistry: Ny et al.

peptidase normally cleaves to the carboxyl side of residues with small neutral side chains, such as glycine, alanine, and serine (47). Thus, the alanine at the NH_2 -terminal side of the valine designated number 1 may represent the termination of the signal peptide. The reading frame shown in Fig. 3 is the only one without multiple termination codons and codes for 383 residues followed by a TGA stop codon. Removal of the putative signal peptide by cleavage between alanine at position -1 and the value designated as number 1 would result in a mature β -PAI that is 379 residues long and has a calculated molecular mass for the carbohydrate-free molecule of 42,770 Da. This calculation agrees well with the molecular mass of the unglycosylated form of the BAE β -PAI as determined by in vitro translation of its mRNA (17). The β -PAI is glycosylated (13) and the amino acid sequence in Fig. 3 contains three putative glycosylation sites conforming to the canonical Asn-Xaa-Ser/Thr sequence (48) at positions 209-211, 265-267, and 329-331. The 3' untranslated region of the 3.0-kb cDNA is 1788 base pairs (bp) long, excluding the poly(A) tract. The consensus polyadenylylation sequence AATAAA is found 16 bp upstream from the poly(A) attachment site, in agreement with previous reports that this sequence is generally located 15-25 nucleotides upstream from the polyadenylylation site (49).

The two clones carrying cDNA inserts of 1.9 kb were partially sequenced and appear to be identical. These clones are also identical to the 3.0-kb cDNA except that they are truncated and lack much of the 3' untranslated region (i.e., they lack the region 3' from nucleotide 1960). Blot hybridization analysis of electrophoretically fractionated total RNA prepared from the human fibrosarcoma cell line HT-1080, using the 3.0-kb cDNA as probe, indicated the presence of two distinct transcripts, 3.0 and 2.2 kb long (Fig. 4). This observation suggests that the 1.9-kb cDNAs may have been copied from the shorter RNA transcript. A similar size heterogeneity at the 3' termini of mRNAs has been observed in other systems. It may result from expression of more than one gene, from alternative splicing events, or from the use of multiple polyadenylylation signals (50-52). The mechanism in this case is not clear. Although the only polyadenylylation consensus signal (AATAAA) found in this cDNA sequence is at the 3' end of the 3.0-kb cDNA (Fig. 3), a similar but slightly



FIG. 4. Detection of β -PAI mRNA. Lane H Γ : total RNA (10 μ g) from the fibrosarcoma cell line HT-1080 was isolated and subjected to agarose gel electrophoresis in the presence of formaldehyde and, after blotting to nitrocellulose, was hybridized to ³²P-labeled λ 3 cDNA. The autoradiogram is shown. Lane M: markers (*Hind*III-digested λ DNA; lengths in bp at left). The mobilities of the eukaryotic 28S and 18S ribosomal RNAs are indicated at right.



FIG. 5. Comparison of β -PAI, α_1 -antitrypsin (α_1 AT), and antithrombin III (AT III). Sequences around the reactive centers were aligned according to the fast protein homology program (43). The reactive-site peptide bonds are indicated by the vertical line, and the terminology of the P₁-P'₁ reactive-site residues is adopted from Travis and Salvesen (54). The reactive-site methionines are underlined, and homologous residues are boxed. Standard one-letter amino acid abbreviations are used.

modified sequence (AATAAT) was found at positions 1998–2003. If this sequence were used as a signal for poly(A) addition, it could explain the presence of the shorter transcript. In other systems, polyadenylylation has been found to take place in the absence of the AATAAA sequence (53).

Comparison of the deduced amino acid sequence with other proteins, using the fast protein analysis homology program (43), revealed that the β -PAI is 25–30% homologous with antithrombin III, α_1 -antitrypsin, α_1 -antichymotrypsinogen, and ovalbumin (data not shown) and therefore is a member of the serine proteinase inhibitor superfamily of proteins (serpins; refs. 23 and 24). The serpins have diverged from an ancestral molecule over a 500-million-year period (24) and now represent a diverse group of related proteins that control the major proteolytic cascades of the body (e.g., the coagulation, complement, fibrinolytic, and inflammatory cascades; ref. 24).

The inhibitory specificity of the serpins appears to be defined primarily by a single amino acid in the reactive center, the P_1 residue (24). In general, this amino acid reflects the known specificity of the target proteinase. The reactive center of the serpins is located near the COOH terminus and, because it appears to protrude from the rest of the molecule (24), may represent the ideal substrate or "bait" for the proteinase. Inhibition is associated with the formation of 1:1 complexes between inhibitor and enzyme. The amino acid sequences of the reactive centers of β -PAI, α_1 -antitrypsin, and antithrombin III are aligned in Fig. 5. In this alignment, the arginine at position 346 is the P_1 residue of the PAI. Plasminogen activators convert plasminogen into plasmin by cleavage of a single Arg-Val bond (1). Thus, this alignment is consistent with the known arginine-specificity of PAs. The finding that the P_{17} residue is glutamic acid also supports this alignment, since this glutamic acid acts as the "hinge" in serpins and is conserved in all serpins sequenced to date (24). The discussion about the reactive site of β -PAI will remain somewhat speculative until the P_1 and P'_1 residues are actually isolated and identified.

The β -PAI is unusually sensitive to oxidants and rapidly loses its activity in the presence of low concentrations of chloramine-T (22). Since there are no cysteines in the deduced protein sequence (Fig. 3), and since the activity of the oxidatively inactivated β -PAI can be restored by treatment with methionine sulfoxide peptide reductase, the loss of activity may reflect the oxidation of a critical methionine (22). The methionine in the reactive center of β -PAI (i.e., at position 347, the inferred P'₁ position) is a likely candidate, since α_1 -antitrypsin also is sensitive to oxidation and its loss of activity has been related to the oxidation of the P₁ methionine (24). In both cases, the resulting methionine sulfoxide is a bulkier residue and may not readily fit into the pocket of its substrate proteinase.

It has been suggested that the ability to selectively inactivate α_1 -antitrypsin by oxidation of its active-site methionine is an important and unique regulatory feature of this system. Activated neutrophils may neutralize the inhibitor by the secretion of oxygen free radicals (24) at inflammatory sites. This additional level of regulation may provide the means by which essential tissue breakdown can take place, even in the presence of inhibitors that normally inhibit neutrophil elastase. Elevated PA activity has been correlated with tissue destruction, tissue remodeling, and the formation of new organs (for review, see ref. 55). The ability to oxidatively inactivate the β -PAI present in these tissues may also be an important regulatory feature of these systems, enabling PAs to function in the presence of their inhibitor. Thus, the local generation of oxidants may inactivate both α_1 -antitrypsin and β -PAI and in the process unleash a cascade of proteolytic enzymes including elastase, plasmin, and collagenase (56).

We acknowledge unpublished information obtained from Dr. H. Pannekoek and collaborators. We thank Tom Podor for the affinitypurified antibodies to the BAE PAI, Raymond Schleef for determining the β -PAI activity of the placental extract, and Peggy Tayman for typing the manuscript. This work was supported in part by National Institutes of Health grants to D.J.L. and by a National Institutes of Health grant to J.L.M. (CA42595-01). T.N. was supported by a European Molecular Biology Organization fellowship and by NFR B-BU-8473-700, and M.S., by a predoctoral fellowship from the National Science Foundation. This is publication no. IMM4380 from Scripps Clinic and Research Foundation.

- 1. Collen, D. (1980) Thromb. Haemostasis 43, 77-89.
- Erickson, L. A., Schleef, R. R., Ny, T. & Loskutoff, D. J. (1985) in *Clinics in Haematology*, ed. Ruggeri, Z. M. (Saunders, London), Vol. 14, pp. 513-530.
- Astrup, T. (1978) in Progress in Chemical Thrombolysis, eds. Davidson, J. F., Rowan, R. M., Samama, M. M. & Desnoyers, P. C. (Raven, New York), Vol. 3, pp. 1–89.
- Hoylaerts, M., Rijken, D. C., Lijnen, H. R. & Collen, D. (1982) J. Biol. Chem. 257, 2912–2919.
- Hoal, E. G., Wilson, E. L. & Dowdle, E. B. (1983) Cell 34, 273-279.
- Vassalli, J. D., Baccino, D. & Belin, D. (1985) J. Cell Biol. 100, 86-92.
- Loskutoff, D. J., Ny, T., Sawdey, M. & Lawrence, D. (1986) in Proteases in Biological Control and Biotechnology, eds. Cunningham, D. & Long, G. (Liss, New York), in press.
- 8. Astedt, B., Lecander, I., Brodin, T., Lundblad, A. & Low, K. (1985) Thromb. Haemostasis 53, 122-125.
- Scott, R. W., Bergman, B. L., Bajpai, A., Hersh, R. T., Rodriguez, H., Jones, B. N., Barreda, C., Watts, S. & Baker, J. B. (1985) J. Biol. Chem. 260, 7029-7034.
- Loskutoff, D. J., van Mourik, J. A., Erickson, L. A. & Lawrence, D. (1983) Proc. Natl. Acad. Sci. USA 80, 2956–2960.
- 11. Emeis, J. J., van Hinsberg, V. W. M., Verheijen, J. H. & Wijngaards, G. (1983) Biochem. Biophys. Res. Commun. 110, 392-398.
- 12. Thorsen, S. & Philips, M. (1984) Biochim. Biophys. Acta 802, 111-118.
- van Mourik, J. A., Lawrence, D. A. & Loskutoff, D. J. (1984) J. Biol. Chem. 259, 14914–14921.
- 14. Erickson, L. A., Hekman, C. M. & Loskutoff, D. J. (1986) Blood, in press.
- 15. Hekman, C. & Loskutoff, D. J. (1986) Haemostasis, in press.
- Erickson, L. A., Hekman, C. M. & Loskutoff, D. J. (1985) Proc. Natl. Acad. Sci. USA 82, 8710–8714.
- 17. Sawdey, M., Ny, T. & Loskutoff, D. J. (1986) Thromb. Res. 41, 151-160.
- Bevilacqua, M. P., Schleef, R. R., Gimbrone, M. A., Jr., & Loskutoff, D. J. (1986) J. Clin. Invest., in press.
- 19. Loskutoff, D. J., Roegner, K., Erickson, L. A., Schleef,

R. R., Huttenlocher, A., Coleman, P. L. & Gelehrter, T. D. (1986) Thromb. Haemostasis 55, 8-11.

- Ny, T., Bjersing, L., Hsueh, A. J. W. & Loskutoff, D. J. (1985) Endocrinology 116, 1666-1668.
- 21. Sakata, Y., Loskutoff, D. J., Gladstone, C. & Griffin, J. H. (1986) *Blood*, in press.
- 22. Lawrence, D. & Loskutoff, D. J. (1986) Biochemistry, in press.
- 23. Doolittle, R. F. (1983) Science 222, 417-419.
- 24. Carrell, R. & Travis, J. (1985) Trends Biochem. Sci. 10, 20-24. 25. Erickson, L. A. Lawrence, D. A. & Loskutoff, D. I. (1984)
- Erickson, L. A., Lawrence, D. A. & Loskutoff, D. J. (1984) Anal. Biochem. 137, 454-463.
- 26. Millan, J. L. (1986) J. Biol. Chem. 261, 3112-3115.
- Young, R. A. & Davis, R. W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194–1198.
- 28. Young, R. A. & Davis, R. W. (1983) Science 222, 778-782.
- Huynh, T. V., Young, R. A. & Davis, R. W. (1985) in DNA Cloning Techniques: A Practical Approach, ed. Glover, D. (IRL Press, Oxford), Vol. 1, pp. 98-121.
- 30. Cuatrecasas, P. (1969) Biochem. Biophys. Res. Commun. 35, 531-537.
- 31. Laemmli, U. K. (1970) Nature (London) 227, 680-685.
- 32. Lammle, B., Berrettini, M., Schwartz, H. P., Heeb, M. J. & Griffin, J. H. (1986) Thromb. Res. 41, 747-759.
- Johnson, D. A., Gautsch, J. W., Sportsman, J. R. & Elder, J. H. (1984) Gene Anal. Tech. 1, 3-8.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- 35. Birnboim, H. C. & Doly, J. (1979) Nucleic Acids Res. 7, 1513-1522.
- Berger, S. L. & Birkenheimer, C. S. (1979) Biochemistry 18, 5143-5149.
- Fellous, M., Nir, U., Wallach, D., Merlin, G., Rubinstein, M. & Revel, M. (1982) Proc. Natl. Acad. Sci. USA 79, 3082–3086.
- 38. Thomas, P. S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205.
- 39. Messing, J. & Vieira, J. (1982) Gene 19, 269-276.
- Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) Plasmid 13, 31-40.
- 41. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5487.
- 42. Staden, R. (1982) Nucleic Acids Res. 10, 4731-4751.
- 43. Lipman, D. J. & Pearson, W. R. (1985) Science 227, 1435-1441.
- 44. Schleef, R. R., Sinha, M. & Loskutoff, D. J. (1985) J. Lab. Clin. Med. 106, 408-415.
- Bolivar, F., Rodriquez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977) Gene 2, 95-113.
- 46. Blobel, G. & Dobberstein, B. (1975) J. Cell Biol. 67, 852-862.
- 47. von Heijne, G. (1984) J. Mol. Biol. 173, 243-251.
- 48. Marshall, R. D. (1974) Biochem. Soc. Symp. 40, 17-26.
- 49. Proudfoot, N. J. & Brownlee, G. G. (1976) Nature (London) 263, 211-214.
- 50. Crabtree, G. R. & Kant, J. A. (1982) Cell 31, 159-166.
- 51. King, C. R. & Piatigorsky, J. (1983) Cell 32, 707-712.
- Hickok, N. J., Seppanen, P. J., Kontula, K. K., Janne, P. A., Bardin, C. W. & Janne, O. A. (1986) Proc. Natl. Acad. Sci. USA 83, 594-598.
- Ohkubo, H., Kageyama, R., Ujihara, M., Hirose, T., Inayama, S. & Nakanishi, S. (1983) *Proc. Natl. Acad. Sci.* USA 80, 2196-2200.
- 54. Travis, J. & Salvesen, G. S. (1983) Annu. Rev. Biochem. 52, 655-709.
- 55. Dano, K., Andreasen, P. A., Grondahl-Hansen, J., Kristensen, P., Nielsen, L. S. & Skriver, L. (1985) Adv. Cancer Res. 44, 139-266.
- Moscatelli, D., Rifkin, D. B., Isseroft, R. R. & Jaffe, E. A. (1980) in *Proteases and Tumor Invasion*, ed. Struli, P. (Raven, New York), pp. 143-152.