Primary Structure of Human Placental Ribonuclease Inhibitor†

Frank S. Lee, Edward A. Fox, Hai-Meng Zhou,‡ Daniel J. Strydom, and Bert L. Vallee*

Center for Biochemical and Biophysical Sciences and Medicine, Department of Pathology, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Received June 14, 1988; Revised Manuscript Received July 14, 1988

ABSTRACT: The primary structure of the human placental ribonuclease inhibitor (PRI), a tight-binding inhibitor of angiogenin and pancreatic ribonucleases, has been determined from the cDNA. The sequence of the mature protein is composed of 460 amino acids, yielding a molecular mass of 49 847 g/mol. Peptides comprising 92% of the predicted sequence were isolated from a tryptic digest of PRI, and direct sequence information obtained for 65% of the molecule agreed at all positions with the sequence predicted from the cDNA. The amino acid sequence of PRI contains seven direct internal repeat units, each 57 amino acids in length. These repeat units comprise 87% of the molecule. The average degree of identity between any two is 39%. A region within each repeat unit displays similarity to tandem, leucine-rich repeats found in six other proteins. Modification of PRI with iodoacetic acid, *p*-(hydroxymercuri)benzoate, and 5,5′-dithiobis(2-nitrobenzoic acid) reveals that at least 30 of the 32 cysteine residues of PRI are in the reduced form.

Human placental ribonuclease inhibitor (PRI)[1] is a member of a family of proteinaceous cytoplasmic RNase inhibitors that occur in the tissues of many mammalian species and bind to both intracellular and extracellular RNases (Roth, 1967; Blackburn et al., 1977; Blackburn & Moore, 1982). It is widely utilized experimentally to protect RNA in transcription and translation reactions from degradation by adventitious RNases (Robbi & Lazarow, 1978; Scheele & Blackburn, 1979; de Martynoff et al., 1980). Physiologically, these inhibitors may be involved in at least two distinct biological systems. The first is the control of intracellular RNases and, consequently, cytoplasmic RNA levels (Imrie & Hutchison, 1965; Kraft & Shortman, 1970). The second, suggested more recently, is the regulation of angiogenin (Shapiro & Vallee, 1987).

Angiogenin was originally purified from human tumor cell conditioned medium and later from human plasma on the basis of its capacity to induce neovascularization (Fett et al., 1985; Shapiro et al., 1987). Primary structure analysis of the protein revealed that it exhibits approximately 35% identity with pancreatic RNases (Strydom et al., 1985). Angiogenin was subsequently shown to catalyze the cleavage of RNA, although its activity differs markedly from that of pancreatic RNase (Shapiro et al., 1986; St. Clair et al., 1987). On the basis of angiogenin's structural and catalytic similarity with pancreatic RNases, the interaction of angiogenin with PRI was investigated. It was shown that PRI inhibits both the ribonucleolytic and angiogenic activities of angiogenin (Shapiro & Vallee, 1987).

In order to understand the interaction between angiogenin and PRI, both a kinetic examination of the interaction and a structural analysis of PRI have been undertaken. Angiogenin and PRI form an exceedingly tight 1:1 complex (Shapiro & Vallee, 1987; Lee et al., 1988a,b). The apparent association rate constant, 1.8×10^8 M^{-1} s^{-1}, is close to the diffusion-controlled limit. The dissociation rate constant of the complex is 1.3×10^{-7} s^{-1}, yielding a half-life of about 60 days. Therefore, the K_i value calculated from the association and dissociation rate constants is extremely low, 7.1×10^{-16} M, approximately 60 times lower than that for bovine pancreatic RNase A. Results from this laboratory suggest that PRI is a competitive inhibitor of angiogenin (Lee et al., 1988b).

Blackburn et al. (1977) have reported that PRI is an acidic protein with a molecular mass of approximately 51 000 g/mol and is rich in leucine (~18% on a molar basis) and cysteine (~6%). It is irreversibly inactivated by sulfhydryl reagents. Furthermore, exposure of the RNase A–PRI or angiogenin–PRI complexes to pHMB results in their dissociation (Blackburn et al., 1977; Shapiro & Vallee, 1987). The determination of the primary structure of PRI, including the oxidation state of its cysteines, is a first step in the delineation of the structural basis of its functional properties.

Experimental Procedures

Materials. PRI was purified from human placenta by the method of Blackburn (1979). Cyanogen bromide activated Sepharose 4B was obtained from Pharmacia. Goat anti-rabbit IgG conjugated to alkaline phosphatase was from Kierkegaard and Perry. The human placental cDNA library was obtained from Clontech, Inc. Nitrocellulose filters were purchased from Schleicher & Schuell. Restriction endonucleases were from Bethesda Research Laboratories, New England Biolabs, or International Biotechnologies, Inc. T4 DNA ligase was from International Biotechnologies, Inc. Exonuclease III, mung bean nuclease, the Klenow fragment of *Escherichia coli* DNA polymerase I, pUC18, M13mp18, M13mp19, the M13 pentadecamer universal primer, and RNA molecular weight

†This work was supported by funds from Hoechst, A.G., under an agreement with Harvard University. F.S.L. was supported by U.S. Public Health Service Grant NRSA 5T32GM07753-09 from the National Institute of General Medical Sciences. E.A.F. was supported by U.S. Public Health Service Grant NRSA HL-07582-01 from the National Heart, Lung, and Blood Institute.

*Address correspondence to this author at the Center for Biochemical and Biophysical Sciences and Medicine, Harvard Medical School, 250 Longwood Ave., Room 106, Boston, MA 02115.

‡Permanent address: Department of Biological Sciences and Biotechnology, Tsinghua University, Beijing, People's Republic of China.

[1] Abbreviations: PRI, placental ribonuclease inhibitor; RNase, ribonuclease; RNase A, bovine pancreatic ribonuclease A; pHMB, *p*-(hydroxymercuri)benzoate; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); SDS, sodium dodecyl sulfate; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; Hepes, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; PBS, phosphate-buffered saline; HPLC, high-performance liquid chromatography; TFA, trifluoroacetic acid.

0006-2960/88/0427-8545$01.50/0

standards were from Bethesda Research Laboratories. Calf intestinal alkaline phosphatase was from Boehringer Mannheim. A dideoxy sequencing kit employing modified T7 DNA polymerase was obtained from United States Biochemical. GeneScreen nylon membranes, [α-^{32}P]dCTP, and [α-^{35}S]-dATP were from Du Pont-New England Nuclear. A DNA labeling kit was from Amersham. Iodoacetic acid, pHMB, and DTNB were from Sigma. Guanidine hydrochloride was from Pierce Chemical Co.

Antibody Preparation. New Zealand white rabbits were immunized by subcutaneous injection of 300 μg of PRI in Ribi's adjuvant (Ribi ImmunoChem Research, Inc.) followed by sequential biweekly booster injections of 100 μg (subcutaneous), 50 μg (subcutaneous), 25 μg (intramuscular), and 50 μg (intravenous). Rabbits were bled monthly. The PRI–Sepharose affinity resin was prepared by coupling 0.82 mg of PRI to 0.5 g of cyanogen bromide activated Sepharose 4B in 0.1 M $NaHCO_3$, pH 9, for 11 h at 4 °C. Antiserum, diluted with an equal volume of PBS, was applied to a column packed with affinity resin and previously equilibrated with PBS. Affinity-purified antibody was eluted with 3.5 M $MgCl_2$ containing 10% dioxane (v/v) (Johnstone & Thorpe, 1982), immediately dialyzed against PBS, and concentrated on a Centricon 30 microconcentrator (Amicon).

cDNA Library Screening. A human placental cDNA library in λgt11 was screened with affinity-purified rabbit anti-human PRI antibodies (Young & Davis, 1983). Antibodies were used at a concentration sufficient to detect 1 ng of purified PRI spotted onto nitrocellulose. Plaques that bound anti-PRI antibodies were detected with goat anti-rabbit IgG conjugated to alkaline phosphatase (Blake et al., 1984). Positive plaques were purified, and DNA was isolated from them by the plate lysate method (Maniatis et al., 1982). The sizes of the cDNA inserts were estimated by agarose gel electrophoresis after cleavage of the DNA with *Eco*RI. The inserts were then cloned into the *Eco*RI site of pUC18. Restriction endonuclease maps were obtained for each clone.

DNA Sequence Analysis. For the largest PRI clone, λPRI-A, and the clone λPRI-F, selected restriction fragments were purified by electrophoresis in low-melting agarose and cloned into M13mp18 or M13mp19. For the second largest PRI clone, λPRI-I (54 bases shorter than λPRI-A), the cDNA was first cloned into the *Eco*RI site of M13mp18. Two clones, each containing the cDNA in a different orientation, were isolated and utilized for unidirectional sequence analysis. A set of overlapping deletion clones was generated by the method of Henikoff (1984) as modified by Hoheisel and Pohl (1986), with the additional substitution of mung bean nuclease (buffer was 10 mM sodium acetate, pH 5.2, containing 50 mM NaCl, 0.1 mM $ZnSO_4$, 5% glycerol, and 0.8 mM β-mercaptoethanol) for S1 nuclease. Sequencing was conducted by the dideoxynucleotide chain termination method (Sanger et al., 1977) using a modified T7 DNA polymerase (Tabor & Richardson, 1987) and either the pentadecamer or heptadecamer universal primer. DNA/protein sequences were analyzed with the Microgenie (Beckman Instruments, Inc.) and MSEQ (University of Michigan) programs.

Northern Blot Analysis. Poly(A+) RNA from full-term human placenta was prepared by oligo(dT) chromatography (Aviv & Leder, 1972) of RNA isolated by the method of Chirgwin et al. (1979) as modified by Chomczynski and Sacchi (1987). Five micrograms of this RNA was electrophoresed in a 1.2% denaturing agarose gel containing 6% formaldehyde (Maniatis et al., 1982) and transferred to a nylon membrane. The λPRI-A cDNA insert was radiolabeled by the random hexamer method (Feinberg & Vogelstein, 1983) and used as a probe. Following hybridization and three washes, the final of which was in 1.5 mM sodium citrate, pH 7, containing 15 mM NaCl and 0.1% SDS at 65 °C for 60 min, the filter was exposed to X-ray film for 2 days.

Tryptic Digests of PRI. Attempts to isolate nanomolar quantities of PRI that had been alkylated at sulfhydryl groups were unsuccessful; therefore, the studies were done on native protein preparations. PRI, 16 nmol, was desalted by gel filtration on a Sephadex G-25 column (0.8 × 8 cm) in 50 mM Hepes, pH 7.0. In a typical digest, 12 nmol of desalted PRI was adjusted to pH 8.0 and a final concentration of 50 mM ammonium bicarbonate. Trypsin [HPLC purified according to Titani et al. (1982)] was added in a ratio of 3.3% (w/w) to PRI (~4 mg/mL) and the reaction mixture held at 37 °C under N_2 for 18 h.

The mixture was fractionated by reversed-phase chromatography on a Synchropak RP-P octadecylsilane column (Synchrom, Inc.). The eluents were (A) 0.1% TFA in water and (B) 75% acetonitrile–0.09% TFA. A 180-min linear gradient from 5 to 60% solvent B was employed at a flow rate of 1 mL/min. The frontally eluted fractions were partially evaporated under N_2, diluted with eluent A, and rechromatographed on an Altex IP-octadecylsilane column (Beckman Instruments, Inc.). The same solvents and gradient were employed.

Amino Acid and Protein/Peptide Sequence Analysis. Amino acid analyses were performed by the phenyl isothiocyanate precolumn derivatization methodology (Bidlingmeyer et al., 1984) as described (Strydom et al., 1986). A Beckman 890 sequencer, updated to microsequencing status, was employed for sequence studies on the protein and peptides (Strydom et al., 1986).

Cysteine Analysis. The number of free sulfhydryl groups in PRI was quantitated by three methods. For each method, DTT in the storage buffer was removed either by gel filtration or by repeated concentration on a Centricon 30 microconcentrator. In the first method, 20 μg of PRI was carboxymethylated with 20 mM iodoacetic acid in 0.1 M Hepes, pH 8, containing 1 mM EDTA for 2 h in the dark at 25 °C. After dialysis in the dark against 0.1 M acetic acid and then water, the (carboxymethyl)cysteine content was determined by amino acid analysis. In the second method, the increase in absorbance at 250 nm was monitored following the addition of aliquots of pHMB to 1.15 μM PRI in 50 mM Hepes, pH 7, at 27 °C. The concentration of the pHMB stock solution was determined spectrophotometrically with a molar absorptivity of 16 900 M^{-1} cm^{-1} at 233 nm (Boyer, 1954). In the third method, DTNB was added at a final concentration of 50 μM to a solution of 1.15 μM PRI in 0.1 M Hepes, pH 8, containing 1 mM EDTA at 27 °C. Spectrophotometric quantitation of the thionitrobenzoate ion released was based on a molar absorptivity of 13 600 M^{-1} cm^{-1} at 412 nm (Ellman, 1959).

Results

Isolation and Nucleotide Sequencing of PRI cDNA Clones. A cDNA library in expression vector λgt11, constructed with mRNA from human placenta, was screened with affinity-purified polyclonal anti-PRI antibodies. Of 450 000 plaques that were screened, 7 expressed antigen that bound the antibodies. The seven putative positives were plaque purified, and the cDNA inserts were subcloned into pUC18. Restriction mapping revealed that six of the seven clones were related and differed only in length, primarily at one end. The restriction map of the seventh was dissimilar, and partial sequence analysis confirmed that it was distinct from the other six.

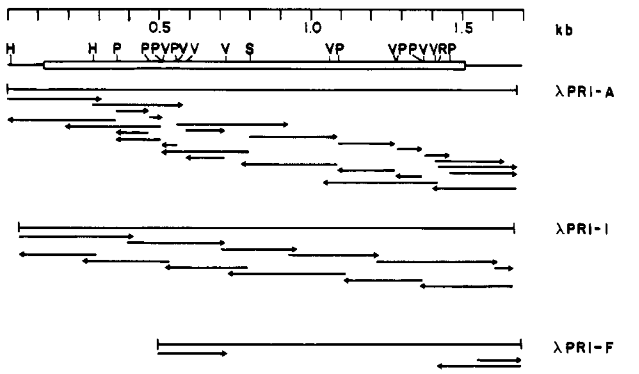

FIGURE 1: Partial restriction map and sequencing strategy of PRI cDNA clones. In the restriction map, the open bar is the coding region for PRI, and abbreviations for restriction sites are as follows: H, *Hinc*II; P, *Pst*I; V, *Pvu*II; S, *Stu*I; R, *Rsa*I. Below the map, the lines with vertical bars indicate the lengths of PRI cDNA clones, and the arrows show the direction and lengths of sequences obtained.

The two largest of the six related clones, λPRI-A and λPRI-I, were sequenced completely on both strands by the dideoxynucleotide chain termination method (Sanger et al., 1977). A third clone, λPRI-F, was 19 bp longer than λPRI-A at the 3′ end and was sequenced at that end. The sequencing strategy is shown in Figure 1. The composite nucleotide sequence of the PRI cDNA, spanning 1698 bp, is shown in Figure 2. It consists of a 5′ untranslated region of 121 nucleotides, an open reading frame of 1383 nucleotides, a termination codon, a 3′ untranslated region of 175 nucleotides, and a poly(A) tail of 16 nucleotides. The initiator codon at nucleotide 122, the first ATG codon in the cDNA, occurs in a sequence, CCACCATGA, that differs at only one position from a consensus sequence for initiation by eukaryotic ribosomes, CCACCATGG (Kozak, 1984). A single potential polyadenylation signal, ATTAAA, is present 19 nucleotides preceding the poly(A) tail. This polyadenylation signal is found in 12% of eukaryotic mRNAs (Wickens & Stephenson, 1984). The nucleotide sequence, excluding the poly(A) tail, contains 64% G+C.

The sequences of the three PRI clones agreed at all nucleotide positions, except one. At nucleotide 1536, in the 3′ untranslated end, λPRI-A and λPRI-I encode a C whereas λPRI-F encodes a T. This difference may either be an artifact introduced by the cloning procedures or be due to the presence of different PRI alleles.

```
                10        20        30        40        50        60        70        80        90       100       110       120
        CGGGGTTGTCGACACGTTCAACCCGTTCTGCTGGCTCGAGAACGAAGTAGGCCGTCTCGCTCTGGGTCTCCAGGCCCGCGACCGTCCGCCAGTCGTCCCGAGGCCACTCTTCACCTCCACC
               130       140       150       160       170       180       190       200       210       220       230       240
        ATGAGCCTGGACATCCAGAGCCTGGACATCCAGTGTGAGGAGCTGAGCGACGCTAGATGGGCCGAGCTCCTCCCTCTGCTCCAGCAGTGCCAAGTGGTCAGGCTGGACGACTGTGGCCTC
     0  (M) S  L  D  I  Q  S  L  D  I  Q  C  E  E  L  S  D  A  R  W  A  E  L  L  P  L  L  Q  Q  C  Q  V  V  R  L  D  D  C  G  L
                                                                 *  *  *  *  *  *  *  *  *                   *  *  *  *  *  *
               250       260       270       280       290       300     · 310       320       330       340       350       360
        ACGGAAGCACGGTGCAAGGACATCAGCTCTGCACTTCGAGTCAACCCTGCACTGGCAGAGCTCAACCTGCGCAGCAACGAGCTGGGCGATGTCGGCGTGCATTGCGTGCTCCAGGGCCTG
    40   T  E  A  R  C  K  D  I  S  S  A  L  R  V  N  P  A  L  A  E  L  N  L  R  S  N  E  L  G  D  V  G  V  H  C  V  L  Q  G  L
         *  *  *  *        *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *        *  *  *  *  *  *  *  *
               370       380       390       400       410       420       430       440       450       460       470       480
        CAGACCCCCTCCTGCAAGATCCAGAAGCTGAGCCTCCAGAACTGCTGCCTGACGGGGGCCGGCTGCGGGGTCCTGTCCAGCACACTACGCACCCTGCCCACCCTGCAGGAGCTGCACCTC
    80   Q  T  P  S  C  K  I  Q  K  L  S  L  Q  N  C  C  L  T  G  A  G  C  G  V  L  S  S  T  L  R  T  L  P  T  L  Q  E  L  H  L
                           *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *
               490       500       510       520       530       540       550       560       570       580       590       600
        AGCGACAACCTCTTGGGGGATGCGGGCCTGCAGCTGCTCTGCGAAGGACTCCTGGACCCCCAGTGCCGCCTGGAAAAGCTGCAGCTGGAGTATTGCAGCCTCTCGGCTGCCAGCTGCGAG
   120   S  D  N  L  L  G  D  A  G  L  Q  L  L  C  E  G  L  L  D  P  Q  C  R  L  E  K  L  Q  L  E  Y  C  S  L  S  A  A  S  C  E
         *  *  *  *  *  *  *  *  *  *                                                  *  *  *  *  *  *  *  *  *  *  *  *  *  *
               610       620       630       640       650       660       670       680       690       700       710       720
        CCCCTGGCCTCCGTGCTCAGGGCCAAGCCGGACTTCAAGGAGCTCACGGTTAGCAACAACGACATCAATGAGGCTGGCGTCCGTGTGCTGTGCCAGGGCCTGAAGGACTCCCCCTGCCAG
   160   P  L  A  S  V  L  R  A  K  P  D  F  K  E  L  T  V  S  N  N  D  I  N  E  A  G  V  R  V  L  C  Q  G  L  K  D  S  P  C  Q
         *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *                       *  *  *  *  *  *  *  *  *  *  *  *
               730       740       750       760       770       780       790       800       810       820       830       840
        CTGGAGGCGCTCAAGCTGGAGAGCTGCGGTGTGACATCAGACAACTGCCGGGACCTGTGCGGCATTGTGGCCTCCAAGGCCTCGCTGCGGGAGCTGGCCCTGGGCAGCAACAAGCTGGGT
   200   L  E  A  L  K  L  E  S  C  G  V  T  S  D  N  C  R  D  L  C  G  I  V  A  S  K  A  S  L  R  E  L  A  L  G  S  N  K  L  G
                        *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *              *  *  *  *  *     •     *  *
               850       860       870       880       890       900       910       920       930       940       950       960
        GATGTGGGCATGGCGGAGCTGTGCCCAGGGCTGCTCCACCCCAGCTCCAGGCTCAGGACCCTGTGGATCTGGGAGTGTGGCATCACTGCCAAGGGCTGCGGGGATCTGTGCCGTGTCCTC
   240   D  V  G  M  A  E  L  C  P  G  L  L  H  P  S  S  R  L  R  T  L  W  I  W  E  C  G  I  T  A  K  G  C  G  D  L  C  R  V  L
         *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *           *  *  *  *  *
               970       980       990      1000      1010      1020      1030      1040      1050      1060      1070      1080
        GGGCCAAGGAGAGCCTGAAGGAGCTCAGCCTGGCCGGCAACGAGCTGGGGGATGAGGGTGCCCGACTGCTGTGTGAGACCCTGCTGGAACCTGGCTGCCAGCTGGAGTCGCTGTGGGTG
   280   R  A  K  E  S  L  K  E  L  S  L  A  G  N  E  L  G  D  E  G  A  R  L  L  C  E  T  L  L  E  P  G  C  Q  L  E  S  L  W  V
                              *  *  *  *  *  *  *  *     *  *  *           *  *  *  *  *  *  *  *  *  *  *  *  *  *
              1090      1100      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
        AAGTCCTGCAGCTTCACAGCCGCCTGCTGCTCCCACTTCAGCTCAGTGCTGGCCCAGAACAGGTTTCTCCTGGAGCTACAGATAAGCAACAACAGGCTGGAGGATGCGGGCGTGCGGGAG
   320   K  S  C  S  F  T  A  A  C  C  S  H  F  S  S  V  L  A  Q  N  R  F  L  L  E  L  Q  I  S  N  N  R  L  E  D  A  G  V  R  E
                                                                                                         *  *  *  *  *  *  *  *
              1210      1220      1230      1240      1250      1260      1270      1280      1290      1300      1310      1320
        CTGTGCCAGGGCCTGGGCCAGCCTGGCTCTGTGCTGCGGGTGCTCTGGTTGGCCGACTGCGATGTGAGTGACAGCAGCTGCAGCAGCCTCGCCGCAACCCTGTTGGCCAACCACAGCCTG
   360   L  C  Q  G  L  G  Q  P  G  S  V  L  R  V  L  W  L  A  D  C  D  V  S  D  S  S  C  S  S  L  A  A  T  L  L  A  N  H  S  L
         *  *  *  *                             *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *
              1330      1340      1350      1360      1370      1380      1390      1400      1410      1420      1430      1440
        CGTGAGCTGGACCTCAGCAACAACTGCCTGGGGGACGCCGGCATCCTGCAGCTGGTGGAGAGCGTCCGGCAGCCGGGCTGCCTCCTGGAGCAGCTGGTCCTGTACGACATTTACTGGTCT
   400   R  E  L  D  L  S  N  N  C  L  G  D  A  G  I  L  Q  L  V  E  S  V  R  Q  P  G  C  L  L  E  Q  L  V  L  Y  D  I  Y  W  S
            *  *  *  *  *  *  *  *  *  *  *                                   *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *
              1450      1460      1470      1480      1490      1500      1510      1520      1530      1540      1550      1560
        GAGGAGATGGAGGACCGGCTGCAGGCCCTGGAGAAGGACAAGCCATCCCTGAGGGTCATCTCCTGAGGCTCTTCCTGCTGCTGCTCTCCCTGGACGACCGGCCTCGAGGCAACCCTGGGG
   440   E  E  M  E  D  R  L  Q  A  L  E  K  D  K  P  S  L  R  V  I  S  •
         *  *  *  *  *  *  *  *  *           *  *  *  *  *  *  *  *  *
              1570      1580      1590      1600      1610      1620      1630      1640      1650      1660      1670      1680
        CCCACCAGCCCCTGCCATGCTCTCACCCTGCATATCCTAGGTTTGAAGAGAAACGCTCAGATCCGCTTATTTCTGCCAGTATATTTTGGACACTTTATAATCATTAAAGCACTTTCTTGG
              1690
        CAAAAAAAAAAAAAAAA
```

FIGURE 2: Nucleotide sequence of PRI cDNA and inferred amino acid sequence. Nucleotides are numbered beginning with +1 at the first nucleotide of the cDNA. Amino acids are numbered beginning with 0 at the initiator methionine. Underlining shows tryptic peptides identified by amino acid composition. Asterisks indicate residues confirmed by Edman degradation. Closed circle indicates the stop codon. Double underlining shows a potential polyadenylation signal.

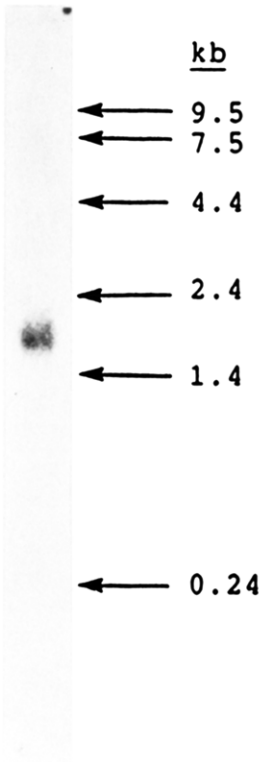

FIGURE 3: Northern blot analysis. Five micrograms of poly(A^+) RNA from full-term human placenta was electrophoresed, transferred, and probed with radiolabeled λPRI-A cDNA insert. The sizes of RNA molecular weight standards are shown to the right of the autoradiogram.

Size of Human PRI mRNA. A Northern blot analysis of human placental poly(A+) RNA revealed the presence of a single band at 1.9 kb when the λPRI-A cDNA insert was used as a probe (Figure 3). The size of the mRNA is slightly larger than that of the cDNA, ~1.7 kb. Since the latter lacks most of the poly(A) tail, which is generally 150–200 bp in length (Brawerman, 1976), this suggests that the cDNA is full length or nearly full length.

Primary Structure of PRI. Thirty-four PRI tryptic peptides were isolated and identified by amino acid compositions as coming from regions covering 92% of the protein (Table I). Thirty of these peptides were sequenced, resulting in sequences comprising 300 residues encoded in areas widely distributed in the largest cDNA open reading frame, thereby establishing it as the correct one.

Evidence that the first methionine in this open reading frame is the initiator codon comes from both amino acid composition and sequence data. Six cycles of Edman degradation on undigested PRI suggested that the N-terminal amino acid is blocked. The amino acid compositions of all the purified tryptic peptides were examined. Save for the initiator methionine, the composition of one peptide agreed closely with that of the N-terminal tryptic peptide predicted from the cDNA sequence (T1, Table I). This peptide was refractory to Edman degradation, consistent with its identity as the N-terminal peptide. Since a seryl residue follows the initiator methionine, these results suggest that the mature protein begins with a modified seryl residue.

Amino acid composition and sequence data are also consistent with the C-terminal amino acid assignment inferred from the cDNA. By amino acid analysis, only one tryptic peptide lacked either lysine or arginine, a characteristic typical of C-terminal peptides in tryptic digests. The amino acid composition of this peptide was consistent with that expected of the C-terminal tryptic peptide (T39, Table I), and the experimentally determined sequence was Val-Ile-Ser, in complete agreement with that predicted from the cDNA.

The following provides further evidence that the open reading frame described encodes PRI: (1) The calculated molecular mass of the mature protein encoded in the open reading frame is 49 847 g/mol. This is in agreement with the molecular mass determined experimentally by SDS–polyacrylamide gel electrophoresis, 50 000 g/mol, and gel filtration chromatography, 52 000 g/mol (Blackburn et al., 1977). (2) The predicted amino acid composition of mature protein closely matches that of hydrolysates of purified PRI (Blackburn et al., 1977).

Cysteine Content of PRI. The number of free sulfhydryl groups in PRI was determined by use of three different reagents—iodoacetic acid, pHMB, and DTNB (Table II). Under nondenaturing conditions, only 4.0 cysteines were carboxymethylated with iodoacetic acid whereas 33.0 and 35.0 cysteines were titrated with pHMB and DTNB, respectively. Under denaturing conditions (4–4.8 M guanidine hydrochloride), all three reagents gave similar results. Carboxymethylation indicated the presence of 30.1 cysteines, close to the values of 33.1 and 35.1 obtained by pHMB and DTNB titration, respectively. Reduction of PRI with DTT under denaturing conditions prior to reaction with iodoacetate resulted in a (carboxymethyl)cysteine content of 31.7 residues, similar to that obtained under denaturing but nonreducing conditions, 30.1.

Discussion

PRI inhibits the enzymatic activity of both neutral cytoplasmic and extracellular RNases; the latter include angiogenin and RNase A. Cloning of the PRI cDNA provides a basis from which to understand both the interaction of PRI with ribonucleases at the protein level and the regulation of PRI at a genetic level.

Anti-PRI antibodies were used to screen a cDNA library derived from human placenta, the source of the protein. Sequences of tryptic peptides then served to confirm the identity of the PRI cDNA clones and to establish the correct reading frame. A combination of amino acid composition and sequence data for two tryptic peptides was employed to determine the N- and C-termini of mature PRI. Finally, the nucleotide sequence of the cDNA established the complete amino acid sequence of the protein. The protein and DNA sequence information confirmed one another. Determination of the oxidation state of the cysteines in the native protein completed the primary structure of PRI.

The serine immediately following the initiator methionine likely is the N-terminal residue of the mature protein. Thus, a peptide with an amino acid composition consistent with that of the N-terminal tryptic peptide (excluding the initiator methionine) was detected in the tryptic digest of PRI (T1, Table I). This peptide as well as the intact protein were refractory to Edman degradation, suggesting that after synthesis the nascent N-terminal methionine is cleaved from the protein and serine, the new N-terminal residue, is modified at its α-amino group. The modifying group probably is acid labile since the full amount of unmodified serine predicted from the cDNA is present in the acid hydrolysate of this peptide. Since acetylation is the most common modification of N-terminal seryl residues in blocked eukaryotic proteins (Wold, 1981; Persson et al., 1985), this acid-labile group may well be an acetyl group.

The Asn-Xxx-Ser/Thr sequence at residues 396–398 is another site of possible posttranslational modification. While

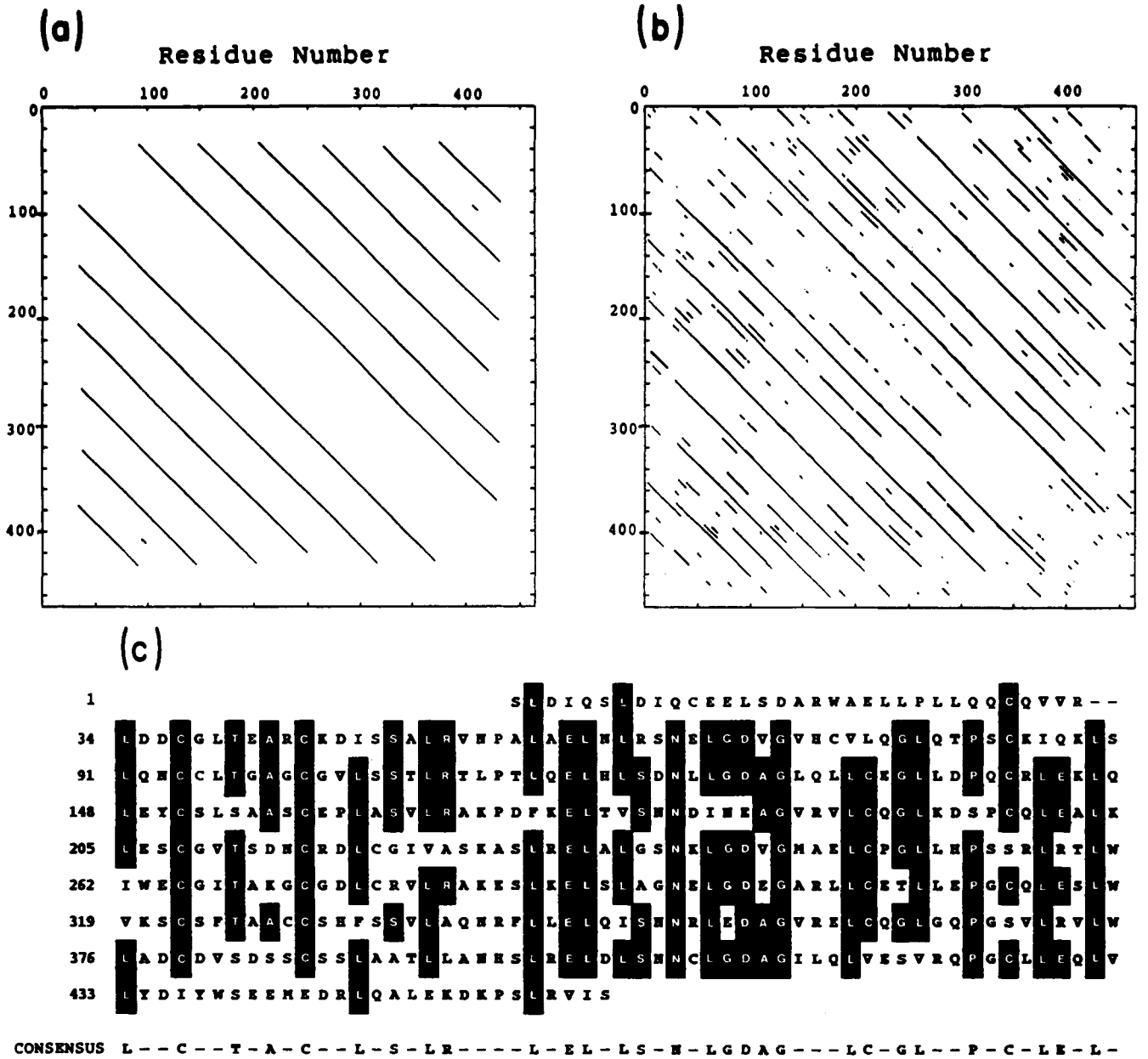

FIGURE 4: Internal repeat structure of PRI. (a) Boswell and McLachlan (1984) self-comparison matrix for PRI. Parameters are as follows: gap penalty = 300, gap extension penalty = 200, damping factor = 0.9, threshold for display in matrix = 4.5 standard deviations. Central diagonal, corresponding to direct alignment, is suppressed. (b) Same as (a), except that threshold for display in matrix = 2.5 standard deviations. (c) Alignment of the seven direct repeats within PRI. The first residue in each line is numbered at the left. Black background is present for identical residues that occur in at least four of the seven repeats. Consensus sequence based on the same criterion is also shown. Two deletions, indicated by dashes, have been introduced into the first line for optimal alignment.

this is a potential site of N-glycosylation, no amino sugars were detected in hydrolysates of the peptide containing this sequence. This is consistent with the fact that neither amino nor reducing sugars were found in the native protein (Blackburn et al., 1977).[2]

The primary structure of PRI is highly repetitive (Figure 4) according to the methodology of Boswell and McLachlan (1984), which yields a self-comparison matrix (Figure 4a) indicative of seven direct internal repeat units comprising 87% of the molecule. Figure 4c shows the alignment of these repeats, extending without deletion from Leu-34 through Val-432, within the primary structure. The repeat units share conserved residues (present in at least four of the seven repeats) at 29 of the 57 aligned positions and contain invariant amino acids at seven of these positions. The average identity between any two is 39%.

Further application of the Boswell and McLachlan algorithm employing less stringent conditions (Figure 4b) suggests that each of the seven strong repeating units of PRI contains an internal duplication, as evidenced by the fainter lines halfway between the six main ones in each half of the matrix. In addition, the seven strong repeat units are flanked at the

[2] We are currently conducting studies to determine whether other posttranslational modifications are present in PRI, and our preliminary findings are as follows: (i) Some of the seryl (105, 385) and cysteinyl (151, 158) residues which were found during sequencer degradations showed either phenylthiohydantoin derivatives of unknown origin in addition to the usual derivatives and their breakdown products or markedly lower yields of the usual products. These sites are under further investigation. (ii) Two peptides, each containing residues 238–256 (T21, Table I), were isolated in equal yields, and sequencing revealed them to be identical at all positions except for position 243; methionine was present in the later eluting peptide, its sulfoxide form in the earlier eluting one.

Table I: Amino Acid Compositions of Tryptic Peptides of Placental Ribonuclease Inhibitor[a]

	sequence position and peptide							
+	1–18 T1	19–33 T2	34–43 T3	46–52 T5	53–63 T6	64–85 T7	86–88 T8	257–258 T22
Asp	2.53 (3)	0.47	1.95 (2)	1.09 (1)	1.96 (2)	1.52 (2)		
Glu	3.52 (4)	3.62 (4)	1.04 (1)	0.19	1.18 (1)	2.64 (3)	1.09 (1)	
Ser	3.21 (3)	0.94	0.14	1.96 (2)	0.15	2.02 (2)		
Gly	1.95	0.69	1.16 (1)	0.14	0.17	2.66 (3)		
His	0.13					0.45 (1)		
Arg	0.79 (1)	0.81 (1)	0.75 (1)	0.95 (1)	1.00 (1)	0.49	0.63	(1)
Thr	0.36	0.33	0.90 (1)			0.96 (1)		
Ala	1.57 (1)	1.14 (1)	1.05 (1)	1.09 (1)	2.07 (2)	0.48		
Pro	0.58	0.87 (1)			1.07 (1)	1.04 (1)		
Tyr	0.30					0.24		
Val	0.53	1.46 (2)		0.14	1.05 (1)	1.82 (3)		
Met	0.20	0.14				0.16		
Ile	1.14 (2)	0.35		1.03 (1)		0.25	1.05 (1)	
Leu	3.21 (3)	3.65 (4)	1.83 (2)	1.11 (1)	2.99 (3)	2.81 (3)	0.67	(1)
Phe	0.27	0.16			0.10	0.35		
Trp (nd)[b]		(1)						
Lys	0.33					0.81 (1)	1.12 (1)	
Cys (nd)	(1)	(1)	(1)	(1)		(2)		
total	18	15	10	7	11	22	3	2

	sequence position and peptide									
+	89–109 T9	110–142 T10	146–166 T12	167–172 T13	271–277 T24	173–187 T14	188–194 T15	195–204 T16	205–216 T17	217–225 T18
Asp	1.27 (1)	4.10 (4)	0.45	1.36 (1)	(1)	3.56 (4)	0.36	0.94 (1)	1.93 (2)	1.04 (1)
Glu	1.77 (1)	5.23 (5)	3.12 (3)	0.16		2.29 (2)	1.07 (1)	2.04 (2)	1.32 (1)	0.16
Ser	2.82 (3)	2.17 (1)	3.70 (4)	0.25		1.27 (1)	0.25	0.92 (1)	1.81 (2)	0.95 (1)
Gly	3.07 (3)	3.24 (3)	0.48	0.93	(2)	1.90 (1)	1.18 (1)	0.20	1.16 (1)	1.06 (1)
His	0.15	1.01 (1)								
Arg	0.98 (1)	0.60 (1)	0.90 (1)	0.29	(1)	0.92 (1)			0.94 (1)	
Thr	1.70 (2)	1.96 (2)	0.13	0.18		0.96 (1)	0.12	0.14	0.89 (1)	
Ala	1.32 (1)	1.17 (1)	2.88 (3)	1.18 (1)		1.36 (1)	0.11	1.12 (1)	0.11	1.04 (1)
Pro	0.31	2.15 (2)	1.03 (1)	1.09 (1)		0.14		1.10 (1)	0.18	
Tyr			0.78 (1)			0.17				
Val	1.09 (1)	0.19	1.10 (1)	0.18		1.95 (2)	1.15 (1)	0.12	0.95 (1)	0.75 (1)
Met	0.12	0.14	0.11			0.14				
Ile	0.42	0.11	0.30			1.02 (1)				0.68 (1)
Leu	4.77 (5)	10.67 (11)	4.85 (5)	0.49	(1)	1.51 (1)	2.03 (2)	2.10 (2)	1.07 (1)	1.08 (1)
Phe	0.26		0.12	1.00 (1)		0.31				
Trp (nd)										
Lys	0.38		0.13	1.42 (2)		0.18	0.94 (1)	0.82 (1)		0.90 (1)
Cys (nd)	(3)	(2)	(2)		(2)		(1)	(1)	(2)	(1)
total	21	33	21	6	7	15	7	10	12	9

	sequence position and peptide								
	226–229 T19	230–237 T20	238–256 T21	259–270 T23	278–280 T25	283–286 T27	287–301 T28	302–320 T29	352–358 T32
Asp		0.95 (1)	0.91 (1)	0.50	0.28	0.28	1.70 (2)	0.39	1.26 (1)
Glu	0.22	1.02 (1)	1.23 (1)	1.23 (1)	0.35	1.07 (1)	3.04 (3)	3.65 (4)	1.05 (1)
Ser	0.75 (1)	0.93 (1)	2.04 (2)	0.71	0.60	0.88 (1)	1.31 (1)	1.45 (1)	0.22
Gly	0.11	1.10 (1)	3.38 (3)	1.15 (1)	0.28	0.39	3.52 (3)	1.72 (1)	1.09 (1)
His			0.89 (1)	0.11		0.11		0.10	
Arg	0.86 (1)		0.98 (1)	0.29	0.91 (1)	0.13	0.71 (1)	0.22	0.97 (1)
Thr			0.27	1.20 (2)	0.17	0.16	0.17	0.97 (1)	
Ala	0.86 (1)	0.92 (1)	1.31 (1)	0.87 (1)	0.64	0.36	2.20 (2)	0.34	1.05 (1)
Pro			2.09 (2)	0.23	0.10	0.13	0.23	0.99 (1)	0.20
Tyr			0.10			0.10	0.11	0.10	
Val	0.35		1.19 (1)	0.32	0.72 (1)	0.30	0.16	1.02 (1)	0.95 (1)
Met			0.81 (1)			0.15	0.12	0.12	
Ile			0.12	1.01 (2)		0.11	0.10	0.19	
Leu	0.98 (1)	1.69 (2)	3.86 (4)	1.50 (1)	1.07 (1)	1.07 (1)	3.14 (3)	5.14 (6)	1.16 (1)
Phe			0.37	0.17		0.10		0.13	
Trp (nd)				(2)				(1)	
Lys		0.71 (1)	0.22	0.67 (1)	0.20	0.55 (1)	0.14	0.99 (1)	0.21
Cys (nd)			(1)	(1)				(2)	
total	4	8	19	12	3	4	15	19	7

	sequence position and peptide						
	359–372 T33	373–400 T34	401–422 T35	423–445 T36	446–451 T37	452–457 T38	458–460 T39
Asp	0.15	3.58 (4)	3.97 (4)	2.22 (2)	0.26	1.11 (1)	0.16
Glu	3.01 (3)	0.42	3.23 (3)	6.08 (6)	2.00 (2)		0.17
Ser	1.13 (1)	5.51 (6)	2.24 (2)	1.25 (1)	0.35	1.03 (1)	1.40 (1)

Table I (Continued)

	sequence position and peptide						
	359–372 T33	373–400 T34	401–422 T35	423–445 T36	446–451 T37	452–457 T38	458–460 T39
Gly	3.32 (3)	0.58	2.44 (2)	1.38 (1)	0.32	0.12	0.30
His		0.90 (1)					
Arg	0.98 (1)	1.01 (1)	1.04 (1)	1.01 (1)	0.22	1.06 (1)	
Thr	0.11	0.98 (1)	0.12	0.26	0.21		
Ala	0.19	3.74 (4)	1.25 (1)	0.29	1.03 (1)		
Pro	1.09 (1)	0.14	0.12	1.16 (1)	0.20	1.05 (1)	
Tyr				1.58 (2)	0.19		
Val	1.01 (1)	2.12 (2)	1.98 (2)	1.16 (1)	0.18		0.79 (1)
Met		0.13		0.81 (1)			
Ile		0.18	1.02 (1)	1.06 (1)			0.69 (1)
Leu	2.80 (3)	5.59 (6)	5.25 (5)	4.40 (4)	1.83 (2)	1.00 (1)	0.11
Phe	0.22	0.11					
Trp (nd)		(1)		(1)			
Lys				0.11	0.79 (1)		
Cys (nd)	(1)	(2)	(1)	(1)			
total	14	28	22	23	6	6	3

[a] Relative molar amounts of amino acids are given, on the basis of the numbers of amino acids inferred from the cDNA sequence (in parentheses). [b] nd, not determined.

Table II: Cysteine Content of PRI[a]

	conditions		
reagent	−reducing[b] −denaturing[c]	−reducing[b] +denaturating[c]	+reducing[b] +denaturing[c]
iodoacetic acid	4.0	30.1	31.7
pHMB	33.0	33.1	ND[d]
DTNB	35.0	35.1	ND[d]

[a] PRI concentration was determined by titration against RNase A (Lee et al., 1988a) when pHMB or DTNB was employed. [b] Sample reduced with 4 mM DTT under denaturing conditions in 0.1 M Hepes, pH 8, containing 1 mM EDTA for 1.5 h at 25 °C. [c] 4.8, 4, and 4 M guanidine hydrochloride present when iodoacetic acid, pHMB, and DTNB were used, respectively. [d] Not determined.

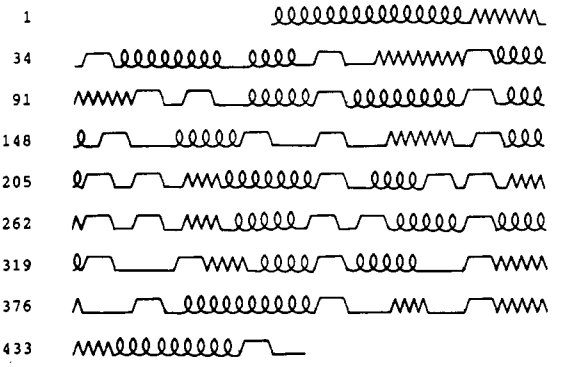

FIGURE 5: Secondary structure predictions for PRI based on algorithm of Chou and Fasman (1978). Sites of predicted α-helix are indicated by loops (number of loops not intended to indicate number of turns), β-sheet by zigzags, and β-turns by humps. Alignment of seven internal repeats is the same as that shown in Figure 4c.

N- and C-termini by a 33-residue (excluding the initiator methionine) and a 28-residue segment, respectively, both of which are similar to the internal repeat units under these conditions.[3] We emphasize that the similarity between half-repeats is significantly less than that between the 57-residue repeat units. Therefore, the dominant periodicity in the primary structure of PRI is that of the 57-residue repeat.

Predictions based on the algorithm of Chou and Fasman (1978) suggest that this 57 amino acid periodicity in primary structure extends to secondary structure (Figure 5). At two positions within each repeat, about one-half and four-fifths of the way as aligned in Figure 5, β-turns are predicted to be fully conserved. At other regions within a repeat, various secondary structures are predicted to be highly conserved between repeats. Together, the periodicity in the primary and predicted secondary structure suggests that PRI contains seven internal domains.

Given previous work examining PRI, the leucyl, cysteinyl, aspartyl, and glutamyl residues are of particular interest with respect to the internal repeat structure. Leucine is, by far, the most abundant component of PRI, comprising ~18% of the residues on a molar basis (Blackburn et al., 1977). The cysteines of PRI are important because PRI is irreversibly inactivated by treatment with sulfhydryl reagents or removal of free thiol from buffers during its purification (Blackburn et al., 1977), indicating that the integrity of the free sulfhydryl groups of PRI is essential for activity. Aspartic acid and glutamic acid may play roles in PRI's mechanism of action as reflected by kinetic data. Thus, increasing sodium chloride concentration decreases the apparent association rate of PRI with angiogenin (Lee et al., 1988a) and increases the K_i for the inhibition of RNase A by PRI (Lee et al., 1988b), suggesting the importance of ionic interactions. These may involve one or more acidic residues of PRI (pI = 4.7; Blackburn et al., 1977) and one or more basic residues of RNase (pI > 9; Richards & Wyckoff, 1971; Fett et al., 1985).

In the primary structure of PRI, 70 of 92 leucines (76%) and 25 of 32 cysteines (78%) are conserved within the repeat units (Figure 4c). The structural features generated by the high degree of conservation of leucine and cysteine residues may contribute to the conformation and function of the native protein. In contrast, only 17 of the 62 acidic residues (29%, Asp + Glu) are conserved (Figure 4c). Eleven of these 17 are glutamic acids and, of these, 7 occur at one invariant position in each repeat. The nonconserved acidic residues in the primary structure are neither clustered nor patterned in any other manner.

The protein was assayed for free sulfhydryl groups to identify the number of cysteine residues whose integrity might be essential for the function of PRI. The present data suggest that at least 30 of the 32 PRI cysteines are in the reduced form. Under nondenaturing conditions, cysteine analysis of PRI is reagent dependent (Table II). In general agreement with previous work (Blackburn et al., 1977), iodoacetic acid alkylated a small fraction of the total. In contrast, both pHMB

[3] It may also be noted that the N-terminal ten residues consist of a direct repeat of five residues.

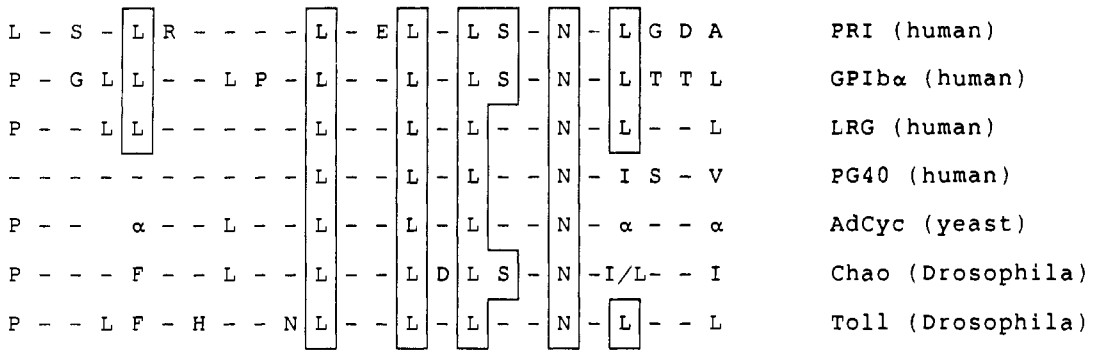

FIGURE 6: Comparison between the repeat consensus sequences of the α subunit of platelet glycoprotein Ib (GPIbα) (Lopez et al., 1987), leucine-rich α_2-glycoprotein (LRG) (Takahashi et al., 1985), fibroblast proteoglycan core protein (PG40) (Krusius & Ruoslahti, 1986; Patthy, 1987), adenylate cyclase (AdCyc) (Kataoka et al., 1985), chaoptin (Chao) (Reinke et al., 1988), the *Toll* protein (Hashimoto et al., 1988), and a portion of that of PRI. Residues identical with those of PRI are boxed. Aliphatic residues (V, L, and I) are indicated by an α.

and DTNB reacted with all 32 cysteines; the experimental values are slightly higher than expected for complete modification and may be due to an uncertainity in PRI concentration determination.

Under denaturing conditions, the results with all three reagents were similar: virtually all cysteines are reactive (Table II). Denaturation and reduction prior to alkylation with iodoacetic acid detected 1.6 more (carboxymethyl)cysteine residues than did denaturation alone. This could reflect the existence of a single disulfide bond or partial oxidation of cysteines during the removal of DTT. The absence of disulfide bonds in PRI would be consistent with its cytoplasmic localization, since the cysteines of cytoplasmic proteins are generally in the reduced form (Torchinskii, 1974).[4]

The internal repeat structure is intriguing in view of the 1:1 stoichiometry of PRI/angiogenin binding. While PRI contains seven repeat units, it binds only one molecule of angiogenin (Shapiro & Vallee, 1987). In proteinase/proteinase inhibitor interactions, the stoichiometry of binding and the number of inhibitor domains deduced from the primary structure usually correspond. Thus, one molecule of adzuki bean inhibitor I binds to two molecules of trypsin (Yoshida & Yoshikawa, 1975); the amino acid sequence of this inhibitor reveals the presence of two homologous repeated domains (Kiyohara et al., 1981). The inhibited proteinases need not be identical. Thus, lima bean trypsin inhibitor simultaneously binds to both trypsin and chymotrypsin (Krahn & Stevens, 1970) in contrast to PRI, which binds either to angiogenin or RNase A but not to both (Shapiro & Vallee, 1987), suggesting that angiogenin and RNase A bind to the same site. It will therefore be of interest to determine the functional constraint that has resulted in the internal repeat structure of PRI.

The primary structure of PRI exhibits a significant number of identities with six other proteins of diverse function which all contain leucine-rich repeats: the α subunit of human platelet glycoprotein Ib (Titani et al., 1987; Lopez et al., 1987), human serum leucine-rich α_2-glycoprotein (Takahashi et al., 1985), human fibroblast proteoglycan core protein (Krusius & Ruoslahti, 1986; Patthy, 1987), yeast adenylate cyclase (Kataoka et al., 1985), *Drosophila* chaoptin (Reinke et al., 1988), and the *Drosophila Toll* protein (Hashimoto et al., 1988). Each of these proteins contains at least seven tandem repeat units that are each exactly or close to 24 amino acids in length.[5] The consensus sequences of the repeats in these six proteins are compared to a segment of that of PRI in Figure 6. The high conservation of certain residues, especially leucine, suggests the possibility of common structural motifs in these proteins. It has been hypothesized that the repeats may be important for protein/membrane or protein/protein interactions (Takahashi et al., 1985; Hashimoto et al., 1988; Reinke et al., 1988). The angiogenin/PRI interaction would be an example of the latter. It is not known whether the other proteins inhibit ribonucleases.

The cDNA provides a means for examining the expression of the PRI gene, and knowledge of the primary structure of the protein now provides a basis for exploring its functional and structural interaction with angiogenin and other RNases in greater detail. The sequence of PRI is a foundation for determination of the contact regions between PRI and RNases, for discernment of possible functional domains, and for exploration of isolating active fragments of PRI. Such studies may ultimately contribute to a larger understanding of the role of RNases in physiological processes.

ACKNOWLEDGMENTS

We thank Dr. Robert Shapiro for supplying PRI, both he and Dr. James Riordan for valuable discussions, Drs. Karen Olson and Edward Alderman for providing anti-PRI antiserum, Dr. Barton Holmquist for assistance in preparation of the figures, Albert Shaw for helpful advice, and George Cohen, Sophia Pao, Nazik Sarkissian, and Rebecca Ettling for excellent technical assistance.

Registry No. PRI, 116926-05-7; precursor PRI, 116926-04-6; PRI cDNA, 116926-03-5; nuclease inhibitor, 39369-21-6.

REFERENCES

Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1408–1412.

Bidlingmeyer, B. A., Cohen, S. A., & Tarvin, T. L. (1984) *J. Chromatogr. 336*, 93–104.

Blackburn, P. (1979) *J. Biol. Chem. 254*, 12484–12487.

Blackburn, P., & Moore, S. (1982) *Enzymes (3rd Ed.) 15*, 317–433.

Blackburn, P., Wilson, G., & Moore, S. (1977) *J. Biol. Chem. 252*, 5904–5910.

Blake, M. S., Johnston, K. H., Russell-Jones, G. J., & Gotschlich, E. C. (1984) *Anal. Biochem. 136*, 175–179.

Boswell, D. R., & McLachlan, A. D. (1984) *Nucleic Acids Res. 12*, 457–464.

[4] Also consistent with the cytoplasmic localization of PRI, hydrophobicity calculations (Kyte & Doolittle, 1982) indicate the absence of hydrophobic segments long enough to constitute a transmembrane domain.

[5] The β subunit of human platelet glycoprotein Ib (Lopez et al., 1988) also contains a single 24-residue segment that is similar to the consensus sequences discussed here.

Boyer, P. D. (1954) *J. Am. Chem. Soc. 76*, 4331–4337.
Brawerman, G. (1976) *Prog. Nucleic Acid Res. Mol. Biol. 17*, 117–148.
Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.
Chomczynski, P., & Sacchi, N. (1987) *Anal. Biochem. 162*, 156–159.
Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251–276.
de Martynoff, G., Pays, E., & Vassart, G. (1980) *Biochem. Biophys. Res. Commun. 93*, 645–653.
Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70–77.
Feinberg, A. P., & Vogelstein, B. (1983) *Anal. Biochem. 132*, 6–13.
Fett, J. W., Strydom, D. J., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F., & Vallee, B. L. (1985) *Biochemistry 24*, 5480–5486.
Hashimoto, C., Hudson, K. L., & Anderson, K. V. (1988) *Cell (Cambridge, Mass.) 52*, 269–279.
Henikoff, S. (1984) *Gene 28*, 351–359.
Hoheisel, J., & Pohl, F. M. (1986) *Nucleic Acids Res. 14*, 3605.
Imrie, R. C., & Hutchison, W. C. (1965) *Biochim. Biophys. Acta 108*, 106–113.
Johnstone, A., & Thorpe, R. (1982) *Immunochemistry in Practice*, pp 210–212, Blackwell Scientific, Oxford.
Kataoka, T., Broek, D., & Wigler, M. (1985) *Cell (Cambridge, Mass.) 43*, 493–505.
Kiyohara, T., Yokata, K., Masaki, Y., Matsui, O., Iwasaki, T., & Yoshikawa, M. (1981) *J. Biochem. (Tokyo) 90*, 721–728.
Kozak, M. (1984) *Nucleic Acids Res. 12*, 857–872.
Kraft, N., & Shortman, K. (1970) *Biochim. Biophys. Acta 217*, 164–175.
Krahn, J., & Stevens, F. C (1970) *Biochemistry 9*, 2646–2652.
Krusius, T., & Ruoslahti, E. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7683–7687.
Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.
Lee, F. S., Auld, D. S., & Vallee, B. L. (1988a) *Biochemistry* (in press).
Lee, F. S., Shapiro, R., & Vallee, B. L. (1988b) *Biochemistry* (in press).
Lopez, J. A., Chung, D. W., Fujikawa, K., Hagen, F. S., Papayannopoulou, T., & Roth, G. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5615–5619.
Lopez, J. A., Chung, D. W., Fujikawa, K., Hagen, F. S., Davie, E. W., & Roth, G. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2135–2139.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, pp 371–372, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Patthy, L. (1987) *J. Mol. Biol. 198*, 567–577.
Persson, B., Flinta, C., von Heijne, G., & Jörnvall, H. (1985) *Eur. J. Biochem. 152*, 523–527.
Reinke, R., Krantz, D. E., Yen, D., & Zipursky, S. L. (1988) *Cell (Cambridge, Mass.) 52*, 291–301.
Richards, F. M., & Wyckoff, H. W. (1971) *Enzymes (3rd Ed.) 4*, 647–806.
Robbi, M., & Lazarow, P. B. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4344–4348.
Roth, J. S. (1967) *Methods Cancer Res. 3*, 153–242.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Scheele, G., & Blackburn, P. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4898–4902.
Shapiro, R., & Vallee, B. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2238–2241.
Shapiro, R., Riordan, J. F., & Vallee, B. L. (1986) *Biochemistry 25*, 3527–3532.
Shapiro, R., Strydom, D. J., Olson, K. A., & Vallee, B. L. (1987) *Biochemistry 26*, 5141–5146.
St. Clair, D. K., Rybak, S. M., Riordan, J. F., & Vallee, B. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8330–8334.
Strydom, D. J., Fett, J. W., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F., & Vallee, B. L. (1985) *Biochemistry 24*, 5486–5494.
Strydom, D. J., Harper, J. W., & Lobb, R. R. (1986) *Biochemistry 25*, 946–951.
Tabor, S., & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4767–4771.
Takahashi, N., Takahashi, Y., & Putnam, F. W. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1906–1910.
Titani, K., Sasagawa, T., Resing, K., & Walsh, K. A. (1982) *Anal. Biochem. 123*, 408–412.
Titani, K., Takio, K., Handa, M., & Ruggeri, Z. M. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5610–5614.
Torchinskii, Y. M. (1974) *Sulfhydryl and Disulfide Groups of Proteins* (Dixon, H. B. F., Transl.) pp 212–214, Plenum, New York.
Wickens, M., & Stephenson, P. (1984) *Science (Washington, D.C.) 226*, 1045–1051.
Wold, F. (1981) *Annu. Rev. Biochem. 50*, 783–814.
Yoshida, C., & Yoshikawa, M. (1975) *J. Biochem. (Tokyo) 78*, 935–945.
Young, R. A., & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1194–1198.