Biochem. J. (1992) 281, 645–650 (Printed in Great Britain)

The gene for a novel protein, a member of the protein disulphide isomerase/Form I phosphoinositide-specific phospholipase C family, is amplified in hydroxyurea-resistant cells

Murari M. CHAUDHURI,* Patricia N. TONIN,† William H. LEWIS†‡ and P. R. SRINIVASAN*§

*Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, New York, NY 10032, U.S.A., and †Department of Microbiology, Fitzgerald Building, University of Toronto, Toronto, Canada M5S 1A8

Cell lines selected in multiple steps for increasing resistance to hydroxyurea have been shown to have corresponding increases in ribonucleotide reductase activity. We have isolated a number of cDNA clones from a cDNA library constructed from a highly hydroxyurea-resistant hamster cell line, 600H, in which the activity of ribonucleotide reductase is elevated more than 80-fold. These clones correspond to genomic DNA sequences amplified in the 600H cell line compared with the V79 parental line. One of these cDNA clones, termed P5, codes for a 50 kDa protein detected by *in vitro* translation of poly(A)$^+$ RNA isolated by hybridization/selection. The cDNA sequence contains a single open reading frame of 1317 nucleotides which encodes a polypeptide of 439 amino acids. The amino acid sequence deduced from the cDNA insert contains two copies of the 11-amino-acid sequence Val-Glu-Phe-Tyr-Ala-Pro-Trp-Cys-Gly-His-Cys. Duplicate copies of this sequence also occur in the active site of rat and human protein disulphide isomerase (also known as the β-subunit of human prolyl 4-hydroxylase, tri-iodothyronine-binding protein) and in Form I phosphoinositide-specific phospholipase C, indicating that P5 falls into this newly defined superfamily of proteins. Genomic sequences similar to the cDNA clone are amplified 10–20-fold in hamster cells selected for resistance to increasing concentrations of hydroxyurea, a phenomenon observed earlier with cDNA clones for the M2 subunit of ribonucleotide reductase and ornithine decarboxylase. RNA blots probed with P5 cDNA show two poly(A)$^+$ RNA species which are elevated in hydroxyurea-resistant cells.

INTRODUCTION

The first unique step in DNA replication is the conversion of the four ribonucleoside diphosphates into deoxyribonucleoside diphosphates, catalysed by the enzyme ribonucleotide reductase [1]. The mammalian enzyme is composed of two non-identical subunits, M1 and M2. The M1 subunit is responsible for the complex allosteric regulation of the enzyme and contains nucleotide-binding sites for the substrates and for nucleotide modulators [2]. The M2 subunit contains a non-haem-bound iron and a tyrosyl free radical essential for activity [3]. This free radical can be destroyed by the drug hydroxyurea, a potent inhibitor of ribonucleotide reductase and of cell division [4]. Mammalian cell lines selected for resistance to hydroxyurea have been shown to have an increased level of ribonucleotide reductase activity [5,6]. The elevation in enzyme level has been correlated with increases in free radical content [3,6], and this increase has been shown to be due to amplification of the gene for the M2 subunit [7–9].

We have constructed a cDNA library from the highly hydroxyurea-resistant hamster cell line 600H in order to clone the genes associated with elevated ribonucleotide reductase activity. Several cDNA clones from this library hybridize to genomic sequences amplified in the 600H cell line when compared with the parental hamster cell line. Earlier we described the isolation of a cDNA clone (S7) which contained the coding sequence for the cell-cycle-regulated enzyme, ornithine decarboxylase [9]. We now report on the sequence of yet another cDNA clone, which codes for a novel 50000 Da protein termed P5. The amino acid sequence deduced from the cDNA sequence contains two copies of an 11-amino-acid sequence: Val-Glu-Phe-Tyr-Ala-Pro-Trp-Cys-Gly-His-Cys. This same 11-amino-acid sequence is also repeated twice in protein disulphide isomerase [10], in cellular tri-iodothyronine (T_3)-binding protein [11], in the β-subunit of human prolyl 4-hydroxylase [12] and in Form I phosphoinositide-specific phospholipase C [13]. *In situ* hybridizations indicate that sequences similar to cDNAs for ornithine decarboxylase, the M2 subunit of ribonucleotide reductase and the P5 protein are located on the same band of human chromosome 2p [14]. In addition, using a panel of mouse–hamster somatic cell hybrids, we have mapped sequences similar to M2 subunit, ornithine decarboxylase and P5 cDNAs to hamster chromosome 7 [15]. Based on these various findings, we propose that the gene for P5 protein has been co-amplified in the selection of the hydroxyurea-resistant phenotype along with the genes for the M2 subunit of ribonucleotide reductase and for ornithine decarboxylase.

MATERIALS AND METHODS

Cell culture

All cell lines were routinely grown in α-minimal essential medium (Flow Laboratories) supplemented with 10% fetal bovine serum. V790^R-1 is a derivative of the Chinese hamster lung cell line V79/V6 selected for resistance to 3 mM-ouabain. The hydroxyurea-resistant cell lines were selected for their ability to survive increasing concentrations of hydroxyurea. The cell lines 35H, 100H, 230H and 600H were selected for growth in 35,

Abbreviation used: T_3, tri-iodothyronine.

‡ Died September 1987.

§ To whom correspondence should be addressed.

The nucleotide sequence data reported here will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X62678.

100, 230 and 600 μg of hydroxyurea/ml respectively, and have been described in detail [6].

Construction and screening of cDNA libraries

Total RNA was prepared from 600H cells grown in the presence of hydroxyurea using guanidine thiocyanate extraction [16]. Polyadenylated RNA was isolated by chromatography on oligo(dT)–cellulose [17]. The construction of the cDNA library and its screening has been described [12]. Briefly, the hamster cDNA library was plated on *Escherichia coli* LE392, and duplicate nitrocellulose filter plaque lifts were hybridized with ^{32}P-labelled cDNA probes synthesized from polyadenylated RNAs purified from either wild-type V790^R-1 cells or hydroxyurea-resistant 600H cells. Of approx. 10^5 plaques screened, ten were found to be finally positive with ^{32}P-labelled cDNA prepared from poly(A)$^+$ RNAs of 600H cells, and negative with labelled cDNA from V79 cells. One of these, termed S7, was found to have a partial coding sequence for ornithine decarboxylase [9]. Another isolate which did not cross-hybridize with S7 was designated P5. P5 λ DNA on cleavage with *Eco*RI gave three fragments in addition to λ arms. They were subcloned into the pUC-9 plasmid and were named p5-5, p5-8 and p5-6 respectively. The insert sizes were 0.51 kb, 0.92 kb and 0.71 kb respectively.

Hybridization selection and *in vitro* translation

The procedures outlined in Riciardi *et al.* [18] and Mather *et al.* [19] were used with some modifications, and have been described [9]. Immunoprecipitation of the translation products was carried out by the procedure of Anderson & Blobel [20] using rabbit polyclonal antisera specific for purified P5 protein. Small amounts of P5 protein were purified to homogeneity, based on its size and its position in two-dimensional gels (J. M. Cocking, W. H. Lewis & P. R. Srinivasan, unpublished work). The purified protein was used for the isolation of antibodies and for determining the *N*-terminal amino acid sequence.

DNA sequencing

The various *Eco*RI fragments and other restriction fragments generated from them were cloned into appropriate sites of M13mp19 or M13mp18. Sequencing was performed using the -40 M13 primer or -20 M13 primer (New England Biolabs) and dideoxynucleotide chain-termination methods with [^{35}S]dATP [21]. Overlapping deletion clones were also constructed from the M13mp19 or M13mp18 clones containing p5-8 and p5-6 inserts in both orientations.

DNA purification, Southern blotting and hybridization

Genomic DNA was purified from cell lines, digested with restriction endonuclease, electrophoresed in agarose, blotted on to Gene Screen Plus (Du Pont–New England Nuclear) filters and hybridized as described by the manufacturer. Molecular probes for hybridization were prepared from purified p5-5, p5-8 and p5-6 fragments to a specific radioactivity of approx. 1×10^9 c.p.m./μg by the random primer technique [22] or by nick-translation [23]. Fragment sizes were calculated using λ DNA cleaved with the restriction endonuclease *Hin*dIII as standard.

Northern blotting and RNA hybridization

Total RNA or purified poly(A)$^+$ RNA was electrophoresed through a 1% agarose/6% formaldehyde gel and blotted on to nitrocellulose as described [24].

RESULTS

Isolation and nucleotide sequence of P5 cDNA clones

A cDNA library was constructed from poly(A)$^+$ RNA isolated from the highly hydroxyurea-resistant hamster cell line 600H. The library was screened by differential hybridization to identify those clones representing RNA sequences present in elevated quantities in the drug-resistant line. P5, a cDNA clone isolated in this fashion, contained a 2.2 kb insert and was chosen for further study. On cleavage with *Eco*RI it gave three fragments. To investigate whether the 600H cell line contained amplified DNA sequences homologous to P5, genomic DNA was digested with *Eco*RI, blotted on to a Gene Screen Plus membrane and hybridized to radiolabelled p5-5. After stripping the radioactive probe, the blot was then reprobed with radiolabelled p5-8 and subsequently with radiolabelled p5-6. Fig. 1 (lane 1) shows that p5-5 bound to an 18 kb fragment (Fig. 1*a*), p5-8 to two fragments of 2.8 and 1.4 kb (Fig. 1*b*) and p5-6 to a 2.8 kb fragment (Fig. 1*c*) in the genomic DNA of the parental cell line V790^R-1. To determine whether any of these P5-related sequences were amplified in hydroxyurea-resistant lines, DNA was purified from Hyd^R-4, the first-step hydroxyurea-resistant mutant, and from the multiple-step mutant lines 100H and 600H. DNA from all four cell lines exhibited a similar hybridization pattern (Fig. 1, lanes 2–4) with respect to each fragment. However, the intensity of the hybridizing bands was greater in the hydroxyurea-resistant lines. The intensities of hybridization of the 18 kb fragment (Fig. 1*a*) and the 2.8 kb fragment (Fig. 1*b*) at 1 μg of Hyd^R DNA were almost equivalent to the intensity of 10 μg of V790^R-1 DNA, and thus we estimate that the single-step hydroxyurea-resistant Hyd^R-4 cell line already contains an approx. 10-fold amplification of genomic sequences similar to that of P5.

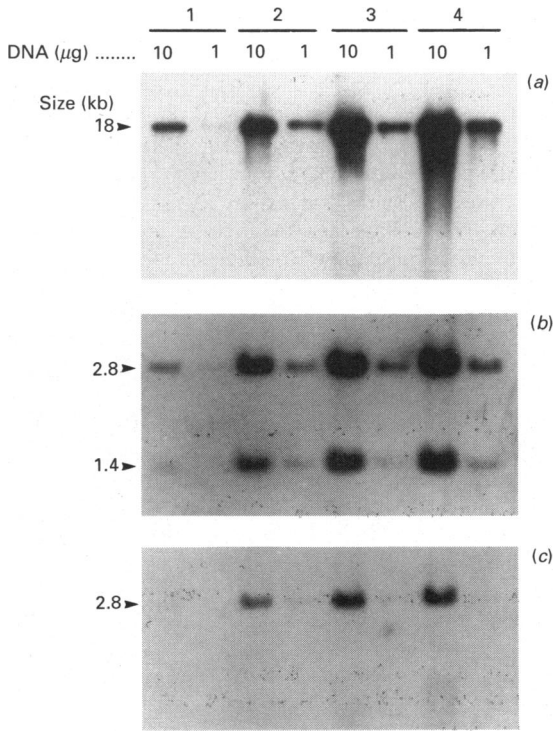

Fig. 1. Southern blot analysis of *Eco*RI-digested genomic DNA hybridized to radiolabelled probes p5-5 (*a*), p5-8 (*b*) and p5-6 (*c*)

Cell lines: lane 1, V790^R-1; lane 2, Hyd^R-4, lane 3, 100H; lane 4, 600H. Each lane contains either 10 or 1 μg of genomic DNA. Arrows indicate fragment size in kilobases.

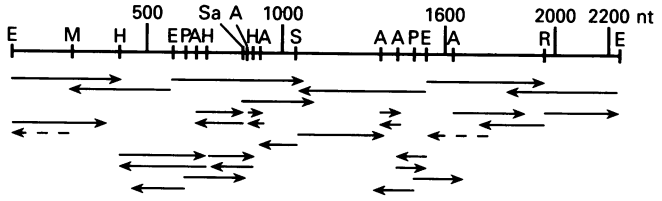

Fig. 2. Sequencing strategy and partial restriction endonuclease map of P5 cDNA clone

Horizontal arrows indicate the strategy for sequencing, representing directions and the lengths of sequencing runs. Enzymes: E, *Eco*RI; H, *Hae*III; A, *Alu*I, R, *Rsa*I; S, *Sac*I; Sa, *Sau*3A; P, *Pst*I; M, *Mbo*II. Two segments indicated by broken arrows were sequenced by using oligodeoxynucleotide primers.

When the highly hydroxyurea-resistant cell line 600H was compared with Hyd^R-4, the intensity of the hybridizing bands appeared to be increased by a further 2-fold (Fig. 1, lane 4). From these experiments, we estimate that the 600H cell line contains a 15–20-fold amplification in P5-related genomic DNA sequences. As a further control, the blot was stripped of the P5 probe and rehybridized with a probe for the catalase gene. The four hamster lines exhibited identical intensities of hybridization, indicating that equal amounts of DNA had been transferred to the filter (results not shown).

In order to identify the amino acid coding potential of the P5 clone, the complete nucleotide sequence of the clone was determined. The sequencing strategy and a partial restriction endonuclease map is given in Fig. 2. The three *Eco*RI fragments p5-5, p5-8 and p5-6 were first individually sequenced and their sequences gave a clue as to the order in which they should be aligned. p5-6 contains the cleavage/polyadenylation signals and a string of adenylate residues, strongly favouring its position at the 3′ end of the cDNA clone. Translation of the sequences of p5-5 and p5-8 indicated that p5-5 should precede p5-8, as p5-8 contains the termination codon UGA at its 3′ end. This order was confirmed by sequencing *Pst*I fragments which overlap the *Eco*RI sites between p5-5 and p5-8, and between p5-8 and p5-6 (Fig. 2). The P5 fragment contains 2234 nucleotides, with an open reading frame of 1317 nucleotides beginning with AUG, followed by 896 nucleotide stretch containing two cleavage/polyadenylation signals AATAAA, (underlined in Fig. 3), one after base 1691 and the other at position 2188 (Fig. 3). The calculated molecular mass of the protein is 48218 Da.

In vitro translation

Hybrid selection and *in vitro* translation experiments were carried out to elucidate the size of the protein encoded in the P5 cDNA sequence. DNA from pUC-9 plasmid and from pUC-9 plasmid containing p5-8 insert were bound to nitrocellulose filters. The filters were hybridized to poly(A)$^+$ RNA purified from 600H cells. The RNAs binding to the filters were eluted and translated in a rabbit reticulocyte system with [^{35}S]methionine and analysed by polyacrylamide-gel electrophoresis. Fig. 4 shows that p5-8-selected mRNA, when translated *in vitro*, produced a polypeptide of molecular mass of approx. 50000 kDa (Fig. 4, lanes 3–5). In lanes 3 and 4, 20-fold and 5-fold dilutions respectively of p5-8 selected mRNA were added. Additional polypeptides of lower molecular masses were apparent, but only the 50 kDa polypeptide band decreased in intensity on dilution of p5-8-selected mRNA (Fig. 4, lanes 3 and 4). To confirm that the 50 kDa protein band selected by p5-8 cDNA is antigenically related to P5 protein, immunoprecipitations of the translation

```
GAA TTC CGC CGG CTT CTC GGC ATG GCT CGC CTT GGG TTC GGT CTG GTG AGC TGT ACC TTC - 60
                           Met Ala Arg Leu Gly Phe Gly Leu Val Ser Cys Thr Phe

TTT CTG GCT GCC AGC GGT CTC TAT TCC TCT AGT GAC GAT GTC ATT GAA TTA ACG CCA TCA - 120
Phe Leu Ala Ala Ser Gly Leu Tyr Ser Ser Ser Asp Asp Val Ile Glu Leu Thr Pro Ser

AAT TTC AAC AGA GAA GTT ATT CAG AGC AAT AGT CTG TGG CTT GTA GAG TTT TAT GCT CCA - 180
Asn Phe Asn Arg Glu Val Ile Gln Ser Asn Ser Leu Trp Leu Val Glu Phe Tyr Ala Pro

TGG TGT GGT CAT TGC CAG AGG TTA ACA CCA GAA TGG AAG AAA GCA GCC ACT GCA TTG AAA - 240
Trp Cys Gly His Cys Gln Arg Leu Thr Pro Glu Trp Lys Lys Ala Ala Thr Ala Leu Lys

GAT GTT GTC AAA GTC GGT GCA GTT GAT GCA GAT AAG CAT CAG TCC CTG GGA GGT CAG TAT - 300
Asp Val Val Lys Val Gly Ala Val Asp Ala Asp Lys His Gln Ser Leu Gly Gly Gln Tyr

GGT GTC CAG GGA TTC CCT ACC ATC AAG ATA TTC GGA GCT AAC AAA AAC AAA CCA GAA GAT - 360
Gly Val Gln Gly Phe Pro Thr Ile Lys Ile Phe Gly Ala Asn Lys Asn Lys Pro Glu Asp

TAT CAG GGT GGC AGA ACT GGG GAG GCC ATC GTA GAT GCT GCC CTC AGT GCT TTG CGC CAG - 420
Tyr Gln Gly Gly Arg Thr Gly Glu Ala Ile Val Asp Ala Ala Leu Ser Ala Leu Arg Gln

CTC GTG AAG GAT CGC CTT AGT GGG CGG AGT GGT GGG TAC AGT TCT GGA AAA CAG GGC AGA - 480
Leu Val Lys Asp Arg Leu Ser Gly Arg Ser Gly Gly Tyr Ser Ser Gly Lys Gln Gly Arg

GGT GAT AGT TCC AGT AAG AAG GAT GTG ATA GAG CTG ACT GAC GAC ACC TTT GAT AAG AAT - 540
Gly Asp Ser Ser Ser Lys Lys Asp Val Ile Glu Leu Thr Asp Asp Thr Phe Asp Lys Asn

GTC CTG GAT AGT GAC GAT GTT TGG ATG GTT GAA TTC TAT GCT CCA TGG TGT GGA CAC TGC - 600
Val Leu Asp Ser Asp Asp Val Trp Met Val Glu Phe Tyr Ala Pro Trp Cys Gly His Cys

AAA AAC CTG GAG CCA GAA TGG GCC ACT GCA GCC ACA GAG GTA AAG GAG CAA ACG AAG GGG - 660
Lys Asn Leu Glu Pro Glu Trp Ala Thr Ala Ala Thr Glu Val Lys Glu Gln Thr Lys Gly

AAA GTA AAG CTG GCA GCC GTG GAC GCT ACG GTG AAC CAG GTC CTG GCC AAC CGG TAT GGG - 720
Lys Val Lys Leu Ala Ala Val Asp Ala Thr Val Asn Gln Val Leu Ala Asn Arg Tyr Gly

ATT AGA GGA TTC CCT ACA ATC AAG ATA TTT CAG AAA GGC GAG GCT CCT GTG GAC TAT GAT - 780
Ile Arg Gly Phe Pro Thr Ile Lys Ile Phe Gln Lys Gly Glu Ala Pro Val Asp Tyr Asp

GGT GGA CGG ACA AGA TCA GAC ATA GTG TCG AGG GCC CTG GAT CTG TTC TCG GAT AAT GCC - 840
Gly Gly Arg Thr Arg Ser Asp Ile Val Ser Arg Ala Leu Asp Leu Phe Ser Asp Asn Ala

CCT CCT CCA GAG CTG CTT GAG ATA ATC AAC GAA GAC GTG GCC AAG AAG ATG TGT GAG GAG - 900
Pro Pro Pro Glu Leu Leu Glu Ile Ile Asn Glu Asp Val Ala Lys Lys Met Cys Glu Glu

CAC CAG CTC TGT GTC GTG GCC GTG CTG CCC CAC ATC CTG GAC ACT GGG GCC GCC AGG AAC - 960
His Gln Leu Cys Val Val Ala Val Leu Pro His Ile Leu Asp Thr Gly Ala Ala Arg Asn

TCT TAC TTG GAA ATT CTT CTG AAG CTG GCT GAC AAA TAC AAG AAG AAG ATG TGG GGG TGG - 1020
Ser Tyr Leu Glu Ile Leu Leu Lys Leu Ala Asp Lys Tyr Lys Lys Lys Met Trp Gly Trp

CTG TGG ACA GAA GCT GGA GCT CAG TCT GAG CTG GAG AAT GCA CTG GGG ATT GGA GGG TTT - 1080
Leu Trp Thr Glu Ala Gly Ala Gln Ser Glu Leu Glu Asn Ala Leu Gly Ile Gly Gly Phe

GGG TAC CCT GCC ATG GCA CGC ATC AAC GCT CGC AAA ATG AAA TTT GCT CTT CTC AAA GGG - 1140
Gly Tyr Pro Ala Met Ala Arg Ile Asn Ala Arg Lys Met Lys Phe Ala Leu Leu Lys Gly

TCT TTC AGT GAA CAA GGC ATT AAT GAG TTT CTC AGG GAA CTG TCT TTT GGA CGT GCC TCC - 1200
Ser Phe Ser Glu Gln Gly Ile Asn Glu Phe Leu Arg Glu Leu Ser Phe Gly Arg Ala Ser

ACA GCA CCC GTG GGA GGT GGT TCC TTT CCT GCC ATC ACT GCC AGG GAG CCC TGG GAT GGT - 1260
Thr Ala Pro Val Gly Gly Gly Ser Phe Pro Ala Ile Thr Ala Arg Glu Pro Trp Asp Gly

AGA GAT GGT GAG CTT CCT GTG GAG GAT GAC ATT GAC CTC AGC GAT GTG GAA CTT GAT GAC - 1320
Arg Asp Gly Glu Leu Pro Val Glu Asp Asp Ile Asp Leu Ser Asp Val Glu Leu Asp Asp

CTG GAG AAG GAC GAG TTG TGA GGC CAC AGC TCA GCC TTC CAA TGT CTT CTC TTG GGA GCG - 1380
Leu Glu Lys Asp Glu Leu ---

GGC AGT TTT CCG GCA GTG AAG GCT CTG GTC AGC TGT CTG CCC TTG GCC TTT CCG AGC AGC - 1440
ACT TGC TGG CTG GCC GGC CAC TAA AAA CAC TGC AGC AGT GAA CCT CGG CGT CTC AAG AAA - 1500
ACA CTG AAG AAT TCT ATG AAT TGT AGC AGT GAA TTG GAT TGC GTT CTT TGG TAT ATT TTG - 1560
AAG CAA GGT GGG TTG TTG AAA CAT TTT CCC TCC TGA CTG CTG CTT GAA TGT TCT TGG AAG - 1620
CTG TTT CTT ATA TAT AGG GTT TTT AAA ATG TGA TTC CTT TGT TTG AAT ATT AAT GGC TTT - 1680
TTC CAT TAA AGA ATA AAA TGA TAT TTT GGA CAA TGC TGA TAA ATG TAT GAA ATT AGT GGC - 1740
ACA TCA TAA AAT CGG AGT CTG GTG GTT AGC CTG AAA ACT AAC CCT GTG CAC TGA TGA CAC - 1800
CAC ACC ACG GCT TCC TGT AAA AGT GTT TTT AAA TAT GAA GAT AAG ACC AGG CTG GAG AGA - 1860
TGG CTC AGC CAT TAA AGG CTA GAC TCA CAA CCA AAA ATG AAG ACA GAC CTT TTG ACA AGG - 1920
GTG TGG TTC TGT GTG GTA AAC TTG GTA CCG AGG GAG ACC ACA CTC ATG CCC CAA ACA AAG - 1980
TCT TGG TCT TCG GGC TTC ACC AGT GAC AAG GAG GCT CAG GTT TTC AAC ATT CTC ACT GAT - 2040
TTC CCA TTA AAC GTC TCA TGA AGA CCA GAG CCC CAG TGG ATT ATC CAA GCT ATT TTC ACT - 2100
TTG TGT AAA ACG CTC AGG AGA CAG TTT GGA GTT ATG TTA GGA CAC CAG AAA AGC CTC AGA - 2160
CTG AGG CAT TTT ATT GTA ACT TAA AAA AAT AAA TAC ACT CTA ACT AAA AAA AAA AAA AAA - 2220
AAA AAA AAA AAA AA - 2234
```

Fig. 3. Nucleotide sequence and deduced amino acid sequence of hamster P5 cDNA

The two polyadenylation signals are underlined. The undecapeptide sequence, which is repeated twice, as well as the *N*-terminus of the mature protein are also indicated.

products were performed with antiserum specific for P5 protein. Fig. 4, lane 6 clearly shows that the antiserum immunoprecipitated the 50 kDa protein band selected by the p5-8 cDNA, whereas lane 7 shows that the mRNAs selected with pUC-9 as a control directed the synthesis of no polypeptides capable of being immunoprecipitated by the antiserum.

Size and abundance of P5 mRNAs

Equal amounts of total RNAs from wild-type CHOpro$^-$, V790^R-1 and several of the hydroxyurea-resistant cell lines were fractionated by formaldehyde/agarose electrophoresis, transferred to nitrocellulose and hybridized to ^{32}P-labelled p5-8 cDNA

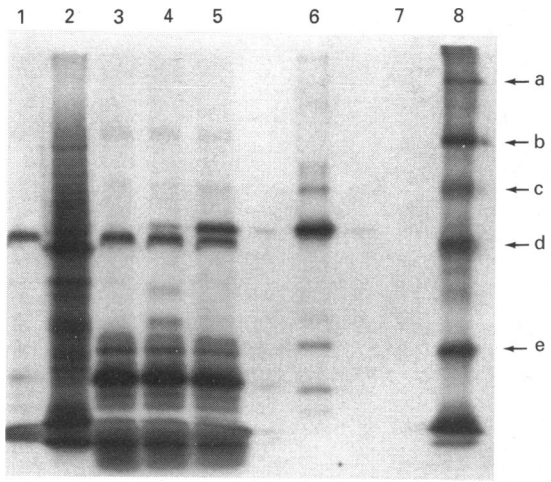

Fig. 4. Electrophoresis of the *in vitro* translation products of hybridization-selected poly(A)$^+$ RNA from 600H cells

Lane 1, no exogenous RNA; lane 2, total poly(A)$^+$ RNA; lanes 3–5, poly(A)$^+$ RNA selected by pUC-9 p5-8 cDNA (lane 3, diluted 1–20; lane 4, diluted 1–5; lane 5, undiluted); lane 6, immunoprecipitation of translation products of poly(A)$^+$ RNA selected by pUC-9 p5-8 cDNA; lane 7, immunoprecipitation of translation products of poly(A)$^+$ RNA selected by pUC-9 control DNA; lane 8, ^{14}C-labelled reference proteins: a, myosin; b, phosphorylase *b*; c, BSA; d, ovalbumin; e, β-lactoglobulin.

as a probe. Fig. 5 shows that RNA extracted from all cell lines contained two hybridizing bands, of approx. 1.9 kb and 2.3 kb. However, the higher-molecular-mass band was barely visible in the CHOpro$^-$, V790^R-1 and Hyd^R-4 cell lines when total RNA was examined. The occurrence of the 2.3 kb band in V790^R-1 cells could be readily observed when poly(A)$^+$ RNA was analysed by Northern blot (Fig. 6). Of the two distinct size classes of mRNAs, the smaller mRNA of 1.9 kb predominated in all cell lines examined. Interestingly, the first-step hydroxyurea-resistant

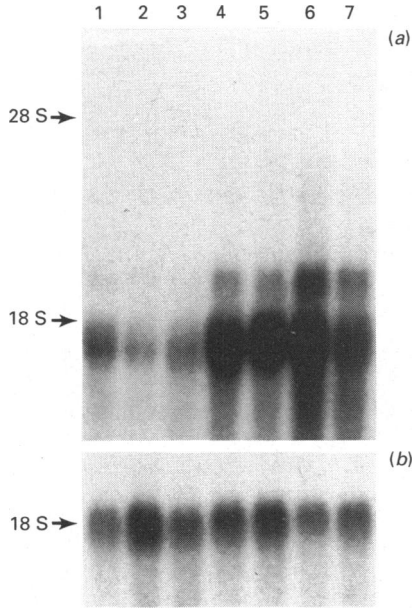

Fig. 5. Northern blot analysis of total RNA hybridized to either radiolabelled p5-8 cDNA (*a*) or radiolabelled actin as control (*b*)

Cells lane 1, CHO; lane 2, V79; lane 3, Hyd^R-4; lane 4, 35H; lane 5, 100H; lane 6, 230H; lane 7, 600H.

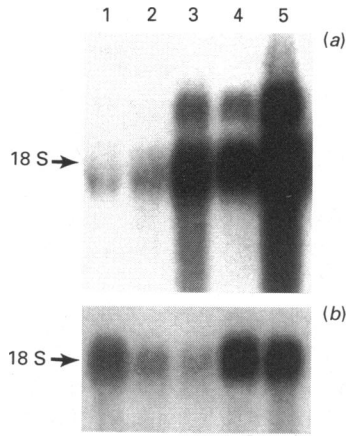

Fig. 6. Northern blot analysis of total RNA and poly(A)$^+$ RNA hybridized to radiolabelled p5-8 cDNA (*a*) or radiolabelled actin as control (*b*)

Lane 1, V79 total RNA, lane 2, Hyd^R-4 total RNA, lane 3, 600H total RNA, lane 4, V79 poly(A)$^+$ RNA; lane 5, 600H poly(A)$^+$ RNA.

mutant Hyd^R-4 showed only a slight elevation of the 1.9 kb species. However, in the other hydroxyurea-resistant cell lines both size classes were elevated. Approx. 5-fold more P5-specific mRNA was present in 35H, 100H and 600H cell lines by dot-blot analysis, but there was only a 50% increase in the first-step hydroxyurea-resistant mutant Hyd^R-4 (results not shown).

DISCUSSION

The technique of differential hybridization has enabled us to isolate a number of cDNA clones whose mRNAs are present in greater copy number in 600H cells compared with parental V79 cells. In an earlier paper [9], we presented evidence that one of these cDNA clones, S7, contained a partial coding sequence for ornithine decarboxylase. The first-step hydroxyurea-resistant mutant, Hyd^R-4, was amplified about 10-fold for genomic sequences similar to S7 cDNA, and the degree of amplification increased by another 2-fold in the highly hydroxyurea-resistant line 600H. Genomic sequences similar to the cDNA of the M2 subunit were also amplified to the same extent in these lines. Now we have analysed another cDNA clone, P5, whose genomic sequences are also amplified to the same degree as M2 and S7 in the first-step hydroxyurea-resistant mutant and in cell lines subsequently isolated with increasing concentrations of hydroxyurea.

The genes for the M2 subunit, for ornithine decarboxylase and for P5 have all been mapped to the same chromosome, chromosome 7 in the hamster [14]. In addition, *in situ* hybridizations indicate that, in human cells, sequences similar to cDNAs for the three genes are located in the same band on the short arm of the chromosome 2 [13]. We have also examined two stable revertant lines with different sensitivities to hydroxyurea and different ribonucleotide reductase activity levels [25]. In these hamster lines, a decrease in hydroxyurea resistance is accompanied by a parallel decrease in gene copies for the M2 subunit, ornithine decarboxylase and the P5 protein. All of these findings, taken together, strongly support the idea that the genes for M2, ornithine decarboxylase and P5 are closely linked and may form part of an amplicon in hamster cells.

Our results indicate that P5 mRNAs are present in two size species of poly(A)$^+$ RNA, of 1.9 kb and 2.3 kb, with both present in greater abundance in 35H, 100H and 600H cells. However, the first-step hydroxyurea-resistant mutant Hyd^R-4 shows only a

Table 1. Comparison of peptide sequences in P5 and other proteins

Proteins	Peptide sequence													References
Hamster P5 protein	47 Leu	Val	Glu	Phe	Tyr	Ala	Pro	Trp	Cys	Gly	His	Cys	59 Gln	Present work
	182 Met	Val	Glu	Phe	Tyr	Ala	Pro	Trp	Cys	Gly	His	Cys	194 Lys	
Rat protein disulphide isomerase	46 Leu	Val	Glu	Phe	Tyr	Ala	Pro	Trp	Cys	Gly	His	Cys	58 Lys	[10]
	390 Phe	Val	Glu	Phe	Tyr	Ala	Pro	Trp	Cys	Gly	His	Cys	402 Lys	
β subunit of human prolyl 4-hydroxylase	Pro	Val	Glu	Phe	His	Ala	Pro	Trp	Cys	Gly	His	Cys	Lys	[12]
	Phe	Val	Glu	Phe	Tyr	Ala	Pro	Trp	Cys	Gly	His	Cys	Lys	
Cellular T_3-binding protein	45 Leu	Val	Glu	Phe	Tyr	Ala	Pro	Trp	Cys	Gly	His	Cys	57 Lys	[11]
	389 Phe	Val	Glu	Phe	Tyr	Ala	Pro	Trp	Cys	Gly	His	Cys	401 Lys	
Form-I phosphoinositide-specific phospholipase C	49 Leu	Val	Glu	Phe	Phe	Ala	Pro	Trp	Cys	Gly	His	Cys	61 Lys	[13]
	397 Leu	Ile	Glu	Phe	Tyr	Ala	Pro	Trp	Cys	Gly	His	Cys	409 Lys	
Thioredoxins	24												36	
E. coli	Leu	Val	Asp	Phe	Trp	Ala	Glu	Trp	Cys	Gly	Pro	Cys	Lys	[36]
Rabbit	23 Val	Val	Asp	Phe	Ser	Ala	Thr	Trp	Cys	Gly	Pro	Cys	35 Lys	[37]
Glutaredoxins	3												15	
E. coli	Thr	Val	Ile	Phe	Gly	Arg	Ser	Gly	Cys	Pro	Tyr	Cys	Val	[38]
Calf thymus	14 Val	Val	Val	Phe	Ile	Lys	Pro	Thr	Cys	Pro	Tyr	Cys	26 Arg	[39]

modest increase of 50% over levels in the parental line, V79. The cDNA sequence (Fig. 3) contains the hexanucleotide sequence AATAAA at nucleotide 1692 and again at nucleotide 2188. The occurrence of this signal sequence for cleavage/polyadenylation at two different sites is in agreement with the presence of two size classes of poly(A)$^+$ RNA.

Earlier [6], we had shown using two-dimensional gels that a protein of approx. 53 kDa is present in greater abundance in cell extracts of the highly hydroxyurea-resistant mutant 600H compared with the parental line V79. We also purified small amounts of this 53 kDa protein to near-homogeneity, and preliminary amino acid sequencing studies with a microsequencer gave the following *N*-terminal amino acid sequence: Leu-Tyr-Xaa-Xaa-Xaa-Asp-Asp-Val-Ile-Glu-Leu-Thr-Pro-Ser-Asn-Phe-Asn-Xaa-Glu-Val-Ile-Gln-Ser-Asn (W. H. Lewis & P. R. Srinivasan, unpublished work). This sequence precisely matches the predicted amino acid sequence found in the cDNA clones (Fig. 3, residues 20–43). Moreover, antiserum to the 53 kDa protein also immunoprecipitated the polypeptides specifically selected by the p5-8 cDNA in hybrid selection experiments. It is of particular interest to point out that the deduced amino acid sequence preceding the *N*-terminal amino acid of the isolated protein (i.e. Met-Ala-Arg-Leu-Gly-Phe-Gly-Leu-Val-Ser-Cys-Thr-Phe-Phe-Leu-Ala-Ala-Ser-Gly) has several characteristics typical of signal sequences. Specifically, it contains a positively charged residue (Arg) at position 3 and an extended hydrophobic sequence followed by Ala-Ser-Gly, which conforms to the consensus sequence for transport into the endoplasmic reticulum and for processing by signal peptidase [26].

The amino acid sequence derived from the cDNA sequence contains an undecapeptide sequence, Val-Glu-Phe-Tyr-Ala-Pro-Trp-Cys-Gly-His-Cys, at positions 48–58 and again at positions 183–193 (Table 1). This same sequence, repeated twice, is seen in protein disulphide isomerase [10] and in the β subunit of prolyl 4-hydroxylase [12]. Other studies indicate that a single polypeptide chain acts both as the β subunit of prolyl 4-hydroxylase and as protein disulphide isomerase [27]. Protein disulphide isomerase, a protein of 57000 Da, catalyses the rearrangement of disulphide bonds in various proteins *in vitro*, and based on this property the suggestion has been put forward that it may function as a catalyst *in vitro* for correct disulphide formation [10,28]. Prolyl hydroxylase, on the other hand, catalyses the formation of 4-hydroxyproline in collagen, and the active enzyme is a tetramer composed of two different polypeptide chains, α and β. The β subunit appears to exist in the free form and also in association with the α-subunits. The β-subunit monomers do not have prolyl 4-hydroxylase activity [27].

The nucleotide sequence of a human cellular T_3-binding protein, which is present in the endoplasmic reticulum and in the nuclear envelope, has been elucidated, and this protein also contains the undecapeptide sequence [11]. Moreover, the coding region of the T_3-binding protein nucleotide sequence has 98% sequence identity with the β subunit of human prolyl 4-hydroxylase and 85% sequence identity with rat protein disulphide isomerase. The suggestion has been made that all three activities are products of the same gene [12].

In contrast, the nucleotide sequence of P5 protein has only 46.5% identity with the nucleotide sequence of protein disulphide isomerase. Moreover, *in situ* hybridization studies indicate that P5 maps to human chromosome 2 (p24 → p25) [14], while the gene coding for both the β-subunit of prolyl 4-hydroxylase and the enzyme protein disulphide isomerase has been assigned to human chromosome region 17 (22-qter) [29]. The undecapeptide contains the interesting sequence Cys-Gly-His-Cys. A similar sequence occurs in the redox proteins thioredoxin and glutaredoxin (Table 1). Thioredoxin was originally considered to be the hydrogen donor for the reduction of ribonucleotides in *E. coli* [30], but this view has changed with the discovery of glutaredoxin in a mutant of *E. coli* lacking thioredoxin [31]. Moreover, in rabbit bone marrow, thioredoxin fails to function

as a hydrogen donor for the related ribonucleotide reductase [32]. A comparison of the amino acid sequence of P5 protein (residues 158–247) with that of *E. coli* thioredoxin (residues 20–104) yields 41 perfect matches, 43 mismatches and seven residues unmatched.

Glutaredoxin has also been isolated from rabbit bone marrow and its structure has been elucidated [33]. When it was examined with rabbit ribonucleotide reductase, it could not serve as a hydrogen donor. Thus even the role of glutaredoxin in mammalian system is not clear and needs further study. In view of these changing concepts, it is important to ask whether P5 protein could have a role as a redox protein in ribonucleotide reduction in mammals. This possibility can be excluded because P5 protein is probably localized in the lumen of the endoplasmic reticulum, as discussed below.

It is interesting that the sequence of yet another protein that is involved in cell signal transduction is now known from molecular cloning to be related to protein disulphide isomerase, namely that of a phosphoinositide-specific phospholipase C [13], an enzyme which hydrolyses phosphatidylinositol 4,5-bisphosphate to two second messengers, 1,2-diacylglycerol and inositol 1,4,5-trisphosphate. The deduced amino acid sequence again contains the undecapeptide sequence, repeated twice (Table 1). Thus the P5 protein belongs to a newly identified group of proteins, the protein disulphide isomerase family [34], a family characterized by the occurrence of an undecapeptide sequence repeated twice, which contains the thiol sequence Cys-Gly-His-Cys and which may have many varied functions.

The P5 protein also contains the *C*-terminal tetrapeptide sequence Lys-Asp-Glu-Leu (KDEL) which is found at the *C*-terminus of the three proteins, protein disulphide isomerase, grp 78 and grp 94, which permanently reside in the lumen of the endoplasmic reticulum [35]. The KDEL sequence appears to be the signal for the retention of a protein in the lumen. Deletion or alterations in this amino acid sequence result in secretion. It is, therefore, likely that the P5 protein is also localized in the lumen of the endoplasmic reticulum. What is surprising is the presence of two proteins in the lumen, protein disulphide isomerase and P5 protein, which contain the identical 11-amino-acid sequence repeated twice and which contain the active site sequence Trp-Cys-Gly-His-Cys. If the P5 protein possesses an enzymic activity similar to that of protein disulphide isomerase, i.e. in forming correct disulphide bonds in proteins, then the possibility must be entertained either that these two proteins could be tissue-specific or the protein substrates for the two enzymes may be different. The relative content of mouse P5 mRNA decreases in the following order: lung > kidney > heart > liver > brain (M. M. Chaudhuri & P. R. Srinivasan, unpublished work). This distribution of P5 mRNA is different from that of protein disulphide isomerase mRNA in the rat [10]. The elucidation of the functional role of P5 protein must await its isolation and purification. This should be feasible with the availability of the cDNA clone for this protein.

The investigation was supported by a grant from the American Cancer Society to P.R.S. P.R.S. acknowledges with thanks the advice of Dr. K. Drickamer and Dr. A. Gold in the preparation of this manuscript, and to Ms. Olivia Haub and Dr. M. Goldfarb for providing the total RNAs from mouse tissues.

Received 29 April 1991/12 August 1991; accepted 27 August 1991

REFERENCES

1. Thelander, L. & Reichard, P. (1979) Annu. Rev. Biochem. **48**, 133–158
2. Thelander, L., Eriksson, S. & Akerman, M. (1980) J. Biol. Chem. **255**, 7426–7432
3. Akerblom, L., Ehrneberg, A., Graslund, A., Lankinen, H., Reichard, P. & Thelander, L. (1981) Proc. Natl. Acad. Sci. U.S.A. **78**, 2159–2163
4. Graslund, A., Ehrenberg, A. & Thelander, L. (1982) J. Biol. Chem. **257**, 5711–5715
5. Lewis, W. & Wright, J. (1979) Somatic Cell Genet. **5**, 83–96
6. Lewis, W. & Srinivasan, P. R. (1983) Mol. Cell. Biol. **3**, 1053–1061
7. Thelander, L. & Berg, P. (1986) Mol. Cell. Biol. **6**, 3433–3442
8. Wright, J., Alam, T., McClarty, G., Taggert, A. & Thelander, L. (1987) Somatic Cell Mol. Genet. **13**, 155–165
9. Srinivasan, P. R., Tonin, P. N., Wensing, E. J. & Lewis, W. (1987) J. Biol. Chem. **262**, 12871–12878
10. Edman, J. C., Ellis, L., Blacker, R. W., Roth, R. A. & Rutter, W. J. (1985) Nature (London) **317**, 267–270
11. Cheng, S., Gong, Q., Parkison, C., Robinson, E. A., Appella, E., Merlino, G. T. & Pastan, I. (1987) J. Biol. Chem. **262**, 11221–11227
12. Pihlajaniemi, T., Helaakoski, T., Tasanen, K., Myllyla, R., Huhtala, M., Koivu, J. & Kivirikko, K. I. (1987) EMBO J. **6**, 643–649
13. Frank Bennett, C., Balcarek, J. M., Varichio, A. & Crooke, S. T. (1988) Nature (London) **334**, 268–270
14. Yang-Feng, T., Thelander, L., Lewis, W., Srinivasan, P. R. & Francke, V. (1987) Genomics **1**, 77–86
15. Tonin, P. N., Stallings, R. L., Carman, M. D., Bertino, J. R., Wright, J. A., Srinivasan, P. R. & Lewis, W. H. (1987) Cytogenet. Cell Genet. **45**, 102–108
16. Chirgwin, J., Przbyla, A., MacDonald, R. & Rutter, W. (1979) Biochemistry **18**, 5294–5299
17. Aviv, H. & Leder, P. (1972) Proc. Natl. Acad. Sci. U.S.A. **69**, 1408–1412
18. Riciardi, R., Miller, J. & Roberts, B. (1977) Proc. Natl. Acad. Sci. U.S.A. **76**, 4927–4931
19. Mather, E., Alt, F., Bothwell, A., Baltimore, D. & Koshland, M. (1981) Cell **23**, 369–378
20. Anderson, D. & Blobel, G. (1983) Methods Enzymol. **96**, 111–120
21. Sanger, F., Nicklen, S. & Coulson, A. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
22. Feinberg, A. & Vogelstein, B. (1983) Anal. Biochem. **12**, 6–13
23. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) J. Mol. Biol. **113**, 237–251
24. Maniatis, T., Fritsch, E. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, pp. 202–203, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
25. McClarty, G. A., Tonin, P. N., Srinivasan, P. R. & Wright, J. A. (1988) Biochem. Biophys. Res. Commun. **154**, 975–981
26. Von Heijne, G. (1983) Eur. J. Biochem. **133**, 17–21
27. Koivu, J., Myllyla, R., Helaakosi, T., Pihlajaniemi, T., Tsanen, K. & Kivirikko, K. I. (1987) J. Biol. Chem. **262**, 6447–6449
28. Freedman, R. B. (1984) Trends Biochem. Sci. **9**, 438–441
29. Pajunen, L., Myllyla, R., Helaakoski, T., Pihlajanimi, T., Tasanen, K., Hoyhtya, M., Tryggvasan, K., Solomon, E. & Kivirikko, K. I. (1988) Cytogenet. Cell Genet. **47**, 37–41
30. Laurent, T. C., Moore, E. C. & Reichard, P. (1964) J. Biol. Chem. **239**, 3436–3444
31. Holmgren, A. (1976) Proc. Natl. Acad. Sci. U.S.A. **73**, 2275–2279
32. Hopper, S. & Iurlano, D. (1983) J. Biol. Chem. **258**, 13453–13457
33. Hopper, S., Johnson, R. S., Vath, J. E. & Biemann, K. (1989) J. Biol. Chem. **264**, 20438–20447
34. Freedman, R. B. (1989) Cell **57**, 1069–1072
35. Munro, S. & Pelham, H. R. B. (1987) Cell **48**, 899–907
36. Holmgren, A. (1968) Eur. J. Biochem. **6**, 475–484
37. Johnson, R. S., Matthews, W. R., Biemann, K. & Hopper, S. (1988) J. Biol. Chem. **263**, 9589–9597
38. Hoog, J.-O., Jornvall, H., Holmgren, A., Carlquist, M. & Persson, M. (1983) Eur. J. Biochem. **136**, 223–232
39. Klintrot, I.-M., Hoog, J.-O., Jornvall, H., Holmgren, A. & Luthman, M. (1984) Eur. J. Biochem. **144**, 417–423