
The core region of human glutaminyl-tRNA synthetase homologies with the *Escherichia coli* and yeast enzymes

Pia Thömmes, Radegunde Fett, Beate Schray, Norbert Kunze and Rolf Knippers

Fakultät für Biologie, Universität Konstanz, D-7750 Konstanz, FRG

Received March 29, 1988; Revised and Accepted May 20, 1988

Accession no.X07466

ABSTRACT

We have isolated from a Lambda-gt 11 library a human cDNA clone with one open reading frame of about 2400 bases. A stretch of about 350 amino acids in the deduced amino acid sequence is up to 40 percent identical with parts of the known amino acid sequences of *E.coli* and yeast glutaminyl (Gln)-tRNA synthetase. The isolated cDNA sequence corresponds to an internal section of a 5500 bases long mRNA that codes for a 170 kDa polypeptide associated with Gln-tRNA synthetase. Thus, the human enzyme is about three times larger than the *E.coli* and two times larger than the yeast Gln-tRNA synthetase. The three enzymes share an evolutionarily conserved core but differ in amino acid sequences linked to the N-terminal and C-terminal side of the core.

INTRODUCTION

Aminoacyl-tRNA synthetases play a central role in protein biosynthesis by catalyzing the aminoacylation of tRNAs with cognate amino acids. In most cases this reaction involves the synthesis of an aminoacyl adenylate followed by the transfer of the aminoacyl group to the 3'-end of tRNA. In spite of the generality of this reaction scheme aminoacyl-tRNA synthetases are widely diverse in their molecular masses, their subunit composition and amino acid sequences (recently reviewed by Schimmel, ref.nr.1). Exceptions are short regions of similar amino acid sequences which are found at comparable locations in several synthetases with different amino acid specificities. These similarities include short regions around a version of the HIGH (His-Ile-Gly-His) motif, the "signature" sequence (1,2), found quite frequently in an N-terminal section of several bacterial and yeast synthetases. These sequences are thought to be required for nucleotide binding, an elementary

reaction of the aminoacylation pathway (3).

But except for these short regions of homology the individual synthetases are so different that their evolution from a common ancestor would be rather unlikely. In contrast, at least some synthetases with identical amino acid specificities appear to be evolutionarily conserved between bacteria and yeast (1). This may indicate that a synthetase with a given amino acid specificity developed early in the history of life and remained relatively constant throughout evolution. To further support this concept it would be interesting to obtain the amino acid sequences of synthetases from higher eukaryotes.

We have isolated from a human cDNA library a clone, termed PZ cDNA, about 2.4 kilobases (kb) in length. The deduced amino acid sequence includes regions which are up to 50 percent identical with the N-terminal two thirds of the bacterial and with a large internal section of the yeast Glutamyl-(Gln)-tRNA synthetase (4,5). These regions contain the HIGH motif, a possible nucleotide binding fold of several E.coli and yeast synthetases.

However, Northern blottings showed that the PZ sequence is located within a 5.5 kb long mRNA. Antibodies, raised against the bacterially expressed products of the PZ cDNA, react with a 170 kDa protein, which was found to be associated with Gln-tRNA synthetase activity. Our findings indicate that a human Gln-tRNA synthetase resides on a polypeptide of about 1500 amino acids and is thus much larger than the bacterial and yeast enzymes which are composed of 551 and 809 amino acids, respectively (4,5).

MATERIAL AND METHODS

Cells: HeLa S3 cells were grown in suspension culture in minimal essential medium (Gibco) containing 5% fetal calf serum (Gibco). Human embryonic lung fibroblasts (HEL 299, Flow) were grown on plastic dishes in Dulbecco's modified Eagle's medium (Gibco) containing 10% fetal calf serum.

Selection of clones: Total RNA of HeLa cells was prepared by polysome precipitation (6). The mRNA was purified by two cycles of chromatography on oligo-dT-cellulose (Sigma) (7). cDNA

was prepared using M-MLV reverse transcriptase (BRL), RNase H and E.coli DNA polymerase I (Stehelin) in a modified version of the procedure of Gubler & Hoffman (8). The cDNA was methylated with Eco RI methylase (New England Biolabs), ligated to Eco RI linker and cloned into the Eco RI site of lambda gt 11 (Promega) (9).

Antibodies against a crude preparation of DNA polymerase α (10) were raised in rabbits. Purified IgG of these sera were used to screen the expression cDNA library by a modification of the procedure of Huynh et al. (11). For detection of immunoreactive clones we used a biotin-streptavidin system with peroxidase-conjugated streptavidin for the colour reaction (Amersham).

Sequencing and computer analysis: DNA restriction fragments were subcloned in M13mpl8 and M13mpl9. Sequencing of these subclones was by the dideoxymethod of Sanger et al. (12).

Computer analysis of the DNA sequence was carried out using the programs FASTP and ALIGN of the BSA program library and the main frame IBM 4381 computer at the German Cancer Research Centre, Heidelberg.

Selection of overlapping clones: Overlapping cDNA clones were isolated according to Maniatis (13) using (^{32}P)labeled nick translated DNA as probes.

Purification of a protein A-fusion protein and preparation of antibodies: Competent E.coli cells were transformed with pRIT2T (Pharmacia) containing the PZ-cDNA ligated into the Eco RI cloning site (14). Bacteria were grown at 30°C in LB medium (13) containing 100ug/ml ampicillin and 12.5ug/ml tetracyclin. Expression of the fusion protein was induced by shifting the temperature to 42°C for 1.5 h. Bacteria were harvested by centrifugation at 6000 g and resuspended in a small volume of phosphate buffered saline. After shock freezing the bacteria were opened by sonification. Bacterial debris was removed by centrifugation at 10,000 xg for 15 min. The fusion protein was purified by chromatography on IgG-Sepharose according to the manufacturer's protocol (Pharmacia).

Rabbits were injected five times with 500ug purified fusion protein. After four weeks sera were tested for the pre-

sence of antibodies by immuno blotting. The IgG antibodies were purified by chromatography on protein A-Sepharose (Pharmacia).

Preparation of cell extracts: Protein extracts were prepared essentially as described before (15) using a hypotonic Hepes buffer supplemented with 10ug/ml each of pepstatin, leupeptin and aprotinin as protease inhibitors. The crude extracts were clarified by centrifugation at 10,000 xg for 15 min.

Proteins were radioactively labeled by growing HEL cells for 2 h in DME medium supplemented with 20uCi/ml (³⁵S)methionine (800 Ci/mmol). The cells were then washed three times with phosphate buffered saline and used for protein extraction as above.

Immunoblotting: Immunoblotting was performed essentially according to Towbin (16). Proteins were separated by SDS polyacrylamide gel electrophoresis according to Laemmli (17) and transferred to PVDF-Membranes (Millipore). Transfer was performed at 100mA for 16 h in a buffer composed in 2.5 mM Tris-glycine, pH 8.3; 0,03% SDS; 20% methanol. Marker proteins were stained with Ponceau S (Serva). Marker proteins were myosin (205 kDa), β -galactosidase (116 kDa), phosphorylase B (97 kDa), bovine serum albumin (BSA; 67 kDa), ovalbumin (45 kD) (Sigma). Membranes were saturated with 3% instant non fat milk powder in TBST (10 mM Tris; pH 7.5; 150 mM NaCl; 0.05% Tween 20), and incubated with 50ug/ml antibody for at least one hour in TBST containing 3% BSA. As a secondary antibody we used a goat-anti-rabbit antibody conjugated to alkaline phosphatase (Sigma). Colour reaction was performed with 5-bromo-4-chloro-3-indolylphosphat and Nitro- blue-tetrazolium (Sigma).

Immunoprecipitation: Crude cell extracts were incubated with purified IgG for at least one hour on ice. The immunocomplexes were then precipitated by formalin-fixed Staphylococcus aureus cells. The immunopellets were washed three times with NET buffer (150mM NaCl; 50mM Tris-HCl; pH 7.5; 5mM EDTA and 0.05% NP40) containing 0.5 M LiCl. The pellets were resuspended in sample buffer (0.125 M Tris-HCl; pH 6.8; 4% SDS; 20% glycerol; 10% β -mercaptoethanol), boiled and centrifuged to remove insoluble materials. The supernatant proteins were separated by polyacrylamide gel electrophoresis (17).

For activity analysis protein A sepharose (Pharmacia) was used to precipitate the immunocomplex. The pellet was washed twice with TMN-buffer (50mM Tris-HCl; pH 7.5; 5mM MgCl₂ and 50mM NaCl) and resuspended in the reaction mix for tRNA-synthetase activity.

Glutaminyl-tRNA synthetase activity: Synthetase activity was determined by charging of tRNA with radioactive amino acid. The reaction mixture contained in a final volume of 50ul 50mM Tris-HCl; pH 7.6; 25mM KCl; 5mM MgCl₂; 2.5mM ATP; 5 mg/ml calf liver-tRNA (Boehringer); 0.5uM (³H)glutamine (39 Ci/mmol) or 3.5uM(¹⁴C)-glutamine (285 mCi/mmol). Incubation was for 15 min at 37°C. The reaction was stopped by spotting aliquots onto 3MM paper (Whatman) and precipitation in trichloroacetic acid.

Northern blotting: RNA of HEL cells was prepared according to Favalaro et al. (18). Total RNA or poly A⁺-RNA, prepared by chromatography on oligo-dT-cellulose, was separated on formaldehyde-agarose-gels according to Thomas (19). 18S and 28S rRNA were used as internal size markers. RNA was transferred to Hybond-membrane (Amersham) and fixed by uv-crosslinking. PZ cDNA, subcloned in pUC9, was labelled with α -(³²P)dATP by multiprime-labelling resulting in a specific radioactivity of 10⁹ cpm/ug DNA. Hybridization was performed in 50% formamide, 0.75 M NaCl and 75mM Na-citrate for 36 h at 42°C. The membranes were washed in 0.3 M NaCl, 30mM Na-citrate at 50°C.

RESULTS

The PZ cDNA clone. We have used antibodies, raised in rabbits against DNA polymerase α , isolated by conventional column chromatography (10), to screen a human cDNA library in the Lambda-gt 11 expression vector. Several cDNA clones were isolated. One of these clones, termed PZ cDNA, was sequenced and found to contain one open reading frame of 2376 nucleotides.

However, the deduced amino acid sequence had no similarity with the amino acid sequences of known DNA polymerase, including that of the large subunit of DNA polymerase α (20). Instead a computer assisted alignment revealed a 40% similarity of the PZ encoded amino acids 30-390 with known sequences of the bacterial and yeast Gln-tRNA synthetases (Fig.1). In fact,

```

Pz      1   DUSTTKRANVFEKKQDUGKFLVPGRAENGKUTAFPPPSGSLYLIHGAKR
Gln4   222  GFLGDLKHUGENPQAVPELMEHIEUTGKUNTAFFPPPEPNGLYLIHGAKR
Ecgl nS
          SEARRAFTNFIRQIIDEEDASGKHTTUNTAFFPPPEPNGLYLIHGAKS

51  AILNIMQUNFKGLIRAFDQTNPEKEKDFEKULLEDAHMLIK-PTQTVISQVLETI
271  IMUNFGYAKYHNGITMLAFDQTNPEKERPEVFESIKRMSLGLFK-PNKITVSSDYVDEL
47  IQLNFGIRADYKGLNLRAFDQTNPKEDIEVYESIKDQEWLGFHWSGNURVSSDYVDEL

110  MKYARELLIQEGKAYVDDTPREQKREPL-----EQRIESKHAKNPIKLNQLMIEEK
330  YRIRAEULIKNGKAVUCHCTAREIKRAGIKEDGTPGGEPYACKHADSIEENLQEFROMR
107  HAVRIELINKEIAYUDELTPEQIREVAGILT-----QPGKNSPTADPSIEENLLEFMR

161  KGSDFGHSCLRAKIDTSSNNGCTADPTLVKCKIQPHPTGKQNYVNYPTVDFACPIVDSI
390  DQYKPGAEALAKQNLNSPSPQNLIAVAVLNAPHPTGKQNYVNYPTVDFTHLQVDSI
162  AGGFEEGKCLRAKIDTSSPFIUTADPLVYIKVHEHQTKGKLCIVPTVDFTHCISDAL

221  EGUTHLHTEVYHDEDFVYIEALGLIR-KYVILEYSALNHNITULSKRKLITWFMDEL
450  ENITHSLCTTEFVLSRESVLELDDYHUHF-APPAVEYCALNITGTULSKRKLIAQLUDEKF
222  EGITHSLCTTEFVLSRESVLELDDYHUHF-APPAVEYCALNITGTULSKRKLITWFMDEL

280  MGGDDPALPTVREQLRAGNIVEGLKQFIARQSSASVUNPEQDKIWAFFNKQIIDUPFA
509  MAKHDDPALPTLEIRIARAGUPPGAIFLFIINTLGVTTSTHIIQVAFESARAKMLIEDTTPR
282  MGGDDPAFTIISLARAGVTRASIREFCKRIGUTKQDNTIEMASLESCLIEQLENHNPRA

340  YVRLKK-EVLPUNPEAQEETKEVAKHPKNPEGLKPLVYSPKUFIEGADRETFSQEM
569  LFIILDPAEUMADLSDDYEELRMIIVYRGTPEFGERTUPFTNKVYIEASL-----
342  RIRVLDPA-UKLUIENYQGEETIINCHANKPEGSAQUPFSDEIWIIDPAFREEANKQV

399  UTFIIGRSTLQK-IT---TKNMEKSYLHNSL-ITKTKI-TRKPL--RSL--GLQR LH
620  ---FSEVDDKE-F---FALTPNQUGIiku-----S-HTUS--KSL--EKDEAG
401  KRLULGKELRANAVIKAEPVEKDRIGHITTIIFCTYDADLISKPADGKUKGVIHWS

448  NLLPPIPIICUTVEQITKPLKDFDFKQYVKNKSCHEELNLGDPDLKDKKQIIDLQ
659  KIIIRIINVDNKHVEEGSKPKPKTYIQWPIISSKYNSPLATETRVVNLFLKSENPSSHP
461  AAHALPUEIRLMDALFSDNPFADDFLSULNPESSVIKQGFREPSLKDRVAKKAFDFER

508  AGFFIICDVPVSPVMSCKEAPCULIYIPDGHTKNPTSGSKKTKVEATKNETSAPFKE
719  EGFLEKDIIMPESEVUMKESUMEHNFQDUUKNSPHUUDSUKNSFVVEEDKDSKEUCFDR
521  EGFLEKDSVSTAEKPUFNATUGLADTGK
    
```

Fig. 1 Amino acid sequences (one letter code) of the core regions of Gln-tRNA synthetases. The PZ cDNA consists of 2376 bp forming one open reading frame with a potential coding capacity for 792 amino acids. We show the first 567 amino acids of the deduced amino acid sequence. The N-terminal first residue of the PZ encoded sequence is given as amino acid nr.1 in the upper line (PZ). The following residues in the PZ-derived sequence are aligned for maximal similarity with the E.coli enzyme sequence (Ecgl nS) and the core sequence of the yeast Gln-tRNA synthetase (Gln4). The entire bacterial sequence of 551 amino acids (4) is presented. The yeast sequence, shown above, starts at residue nr. 222 and ends at residue nr.778 of the published sequence containing a total of 809 amino acids (5).

The entire PZ cDNA sequence appears in the EMBL/Gen Bank Nucleotide Sequence Databases under the accession number XO 7466.

the amino acids 30-135 in the deduced PZ sequence are to almost 50 percent at identical places in corresponding regions of the N-terminal halves of the E.coli and yeast enzymes. This region

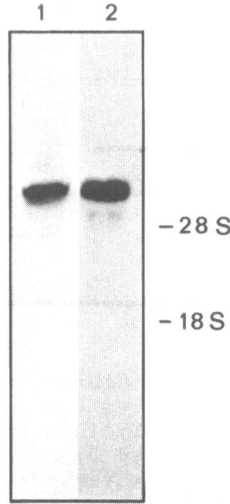


Fig. 2 Northern blots. Poly (A⁺) RNA was prepared from exponentially growing HeLa cells (1) and primary embryonic lung fibroblasts (2) and separated by gel electrophoresis. The RNA was transferred to nitrocellulose filters and hybridized to radioactively labeled PZ cDNA (spec. radioactivity: ca. 1×10^9 cpm/ μ g DNA). We show the autoradiograms after 2 days exposure.

of the PZ clone includes the "signature sequence" (1,2) His--Ile-Gly-His (HIGH) in an environment of a 15-out-of-17 amino acid identity. The part of the PZ sequence between amino acids 141 and 388 was identical to about 40% with comparable sections of the bacterial and yeast synthetase. The three sequences diverge beyond this point and show little or no similarities in their C-terminal sections.

The degree of identity between the three enzymes, shown in Fig.1, is too high to be fortuitous. Thus, it was quite likely that the PZ cDNA contained indeed the coding sequence of the human Gln-tRNA synthetase. To investigate this possibility we first tried to identify the mRNA from which the PZ-cDNA was derived. We found that the PZ-cDNA hybridized to a single, 5.5 kb long mRNA from HeLa and other human cell lines (Fig.2). Using radioactively labeled PZ-cDNA as a probe we have screened several human and mouse cDNA libraries for hybridizing sequences. We have isolated a number of cDNA clones which overlap with the PZ clone and with each other. An arrangement of these clones suggests that the PZ sequence begins approximately 500

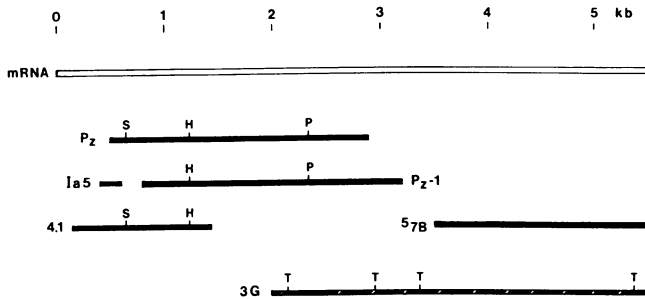


Fig. 3 PZ related cDNA clones. Clones PZ-1, 4.1 and Ia5 were isolated from human cDNA libraries (see: Material and Methods) by hybridization to radioactively labeled PZ-cDNA. Their relationship to PZ-DNA was determined by nucleotide sequencing. Clone 3G was isolated from an Okayama-Berg mouse cDNA library (obtained via N.Brown, Worcester, Mass.) by hybridization to PZ-cDNA. The 5'-terminal Pst I fragment of clone 3G hybridized to clones PZ and PZ-1 but not to clone 4.1. Clone 3G was used to isolate the 3' end clone 5-7 B from a human cDNA library.

Restriction enzyme sites: H, Hind III; P, Pvu II; S, Sph I; T, Pst I.

nucleotides downstream of the putative cap site and ends at approximately 2500 nucleotides upstream of the 3' end of the 5.5 kb mRNA (Fig.3). Thus, the protein sequence of Fig.1 is coded for by a centrally located portion of the mRNA.

A 5.5 kb mRNA has the capacity to code for a protein that is about 3 times larger than the E.coli and 2 times larger than the yeast Gln-tRNA synthetase (4,5). Moreover, the molecular mass of mammalian Gln-tRNA synthetase has been reported to be around 95 kDa (21). Thus, the PZ containing 5.5 kb mRNA appeared to code for an unexpectedly large protein. It remained, of course, possible that the 5.5 kb mRNA was composed of large 5'- or 3'-noncoding sequences. But this possibility was excluded by the experiments described below.

Antibodies against the PZ encoded protein. We have cloned the open reading frame of the PZ cDNA into the Eco RI restriction site of the bacterial expression vector pRIT2T (ref.nr.14). In this construct, the PZ cDNA was linked to a sequence coding for the IgG binding domain of protein A. The fusion protein, composed of a small N-terminal section of protein A and a larger

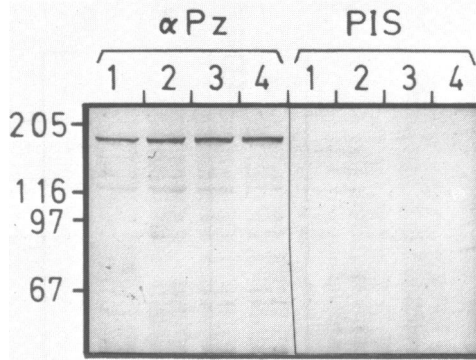


Fig. 4 Immunoblots. Protein extracts from primary lung fibroblasts (1), adenovirus-transformed human 293 cells (2), HeLa cells (3) and proteins from HeLa cells after 100,000 xg centrifugation (4) were separated by polyacrylamide gel electrophoresis and transferred to nitrocellulose filters. The filters were incubated with PZ antibodies (PZ), or with rabbit preimmune serum (PIS) and stained as described in Methods.

C-terminal part, containing the PZ encoded polypeptide, was isolated from induced bacteria and purified to homogeneity by affinity chromatography on IgG-columns (14).

Antibodies, raised against the purified fusion protein in rabbits (PZ antibodies), were used to identify PZ encoded proteins in cytosolic extracts from several human cell lines.

This was done by immunostaining of electrophoretically separated proteins transferred to nitrocellulose filters ("Western blotting"). Our results clearly showed a specific staining of a 170 kDa polypeptide which was present in each one of the three cell lines investigated (Fig.4).

Using the PZ antibodies we performed immunoprecipitation experiments with cytosolic extracts from (^{35}S)methionine labeled primary human lung fibroblasts. The immunoprecipitated proteins were separated by SDS-polyacrylamide gel electrophoresis and visualized by autoradiography. As shown in Fig.5A eight polypeptides were specifically precipitated by the PZ antibodies but not by control antibodies. The most prominent bands of 170, 150 and 135 kDa can also be seen in Coomassie stained gels whereas the faster migrating bands can only be detected by autoradiography of (^{35}S)labeled polypeptides.

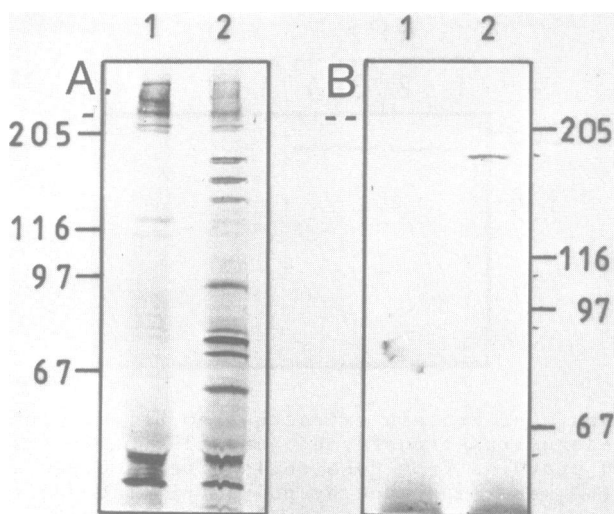


Fig. 5 Immunoprecipitation. (A): Exponentially growing primary cells were radioactively labeled and used for protein preparation. The proteins were incubated with rabbit preimmune sera (1) or with PZ-antibodies (2). The immunoprecipitated proteins were separated by polyacrylamide gel electrophoresis and visualized by autoradiography. The molecular weights of the precipitated proteins have been estimated to be: 170,150, 135,94,78,74,55 and 48 kDa. (B): The immunoprecipitation was performed as in the experiment shown on the left except that the electrophoresed proteins were transferred to a nitrocellulose filter and incubated with PZ antibodies for immunostaining.

To find out which of the immunoprecipitable polypeptides corresponded to the PZ encoded protein we performed an immunoprecipitation experiment, exactly like that shown in Fig. 5A, but transferred the electrophoresed polypeptides to a nitrocellulose filter for immunostaining. We found that only the 170 kDa polypeptide reacted with the PZ antibodies (Fig.5B). This result indicated that the smaller polypeptides, seen in the autoradiogram of the immunoprecipitates (Fig.5A), were not degradation products of the 170 kDa polypeptide. In addition, the 150 kDa polypeptide could be stained in a parallel experiment with an antibody raised against the yeast Ile-tRNA synthetase, and the 135 kDa polypeptide was stained by an antibody against the yeast Leu-tRNA synthetase (data not shown). (Both antibodies were gifts of U. Englisch, Göttingen).

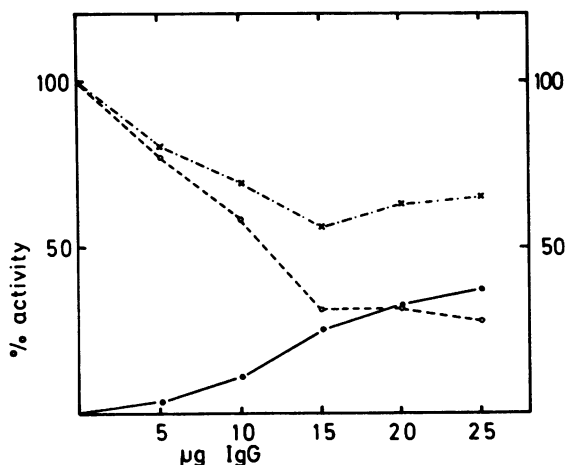


Fig. 6 Immunoprecipitable Gln-tRNA synthetase activity. Proteins, extracted from HeLa cells, were incubated on ice with PZ antibodies in the concentrations indicated above. After 60 min, the immunocomplexes were precipitated with protein A-Sepharose. The entire sepharose pellet was resuspended in the reaction mixture described under Materials and Methods. The aminoacylation of tRNA was assayed in the immunopellet (●) and in the supernatant remaining after immunoprecipitation (○) using radioactively labeled glutamine. Note that the sum of the enzymatic activities (x) in the immunopellet and in the supernatant did not add up to 100 percent. A possible explanation is that the polyclonal PZ antibodies include activities directed at epitopes in the active center of the enzyme. Antibodies with this specificity would inhibit the aminoacylation reaction whereas antibodies directed against other parts of the protein are innocuous.

The 100 percent value in our assay system corresponds to 11 pmole glutamine transferred to trichloroacetic acid precipitable tRNA after 15 min at 37°C using 90 µg of unfractionated protein extract under the conditions described under Methods.

The nature of the other immunoprecipitable polypeptides is not yet known.

The data of Fig.5 suggest that human Gln-tRNA synthetase forms a tight complex with other tRNA synthetases, and this complex remains stable during the immunoprecipitation procedure.

Enzymatic activities associated with the immunoprecipitable proteins. The bacterially expressed fusion protein has no detectable enzymatic activity. We have therefore determined

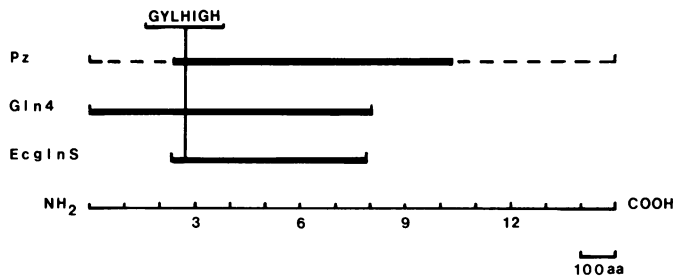


Fig. 7 Localization of the "signature sequence" in Gln-tRNA synthetases. The amino acid sequence Gly-Tyr-Leu-His-Ile-Gly-His (GYLHIGH) is a motif common for the Gln-tRNA synthetase from human cells (PZ), yeast (Gln 4) and E.coli (Ecgl nS). This motif marks the N-terminal region of the enzyme core. Known amino acid sequences are indicated by the thick horizontal lines. The broken line corresponds to unsequenced sections of the human Gln-tRNA synthetase. The localization of the PZ derived region is based on the data of Fig.3.

the activity of the enzyme in specific immunoprecipitates using calf liver tRNA as substrates in an aminoacylation assay and found activities transferring labeled glutamine (Fig.6) as well as labeled isoleucine (not shown) to tRNA. The precipitable enzymatic activity increased with increasing antibody concentrations whereas the supernatants were depleted of comparable enzymatic activity (Fig.6). These findings support our conclusion that the PZ sequence codes for a section of the human cell Gln-tRNA synthetase and, furthermore, that this enzyme is closely associated with Ile-tRNA synthetase.

DISCUSSION

We have isolated a 2.4 kb cDNA clone (PZ-cDNA) which is very likely a complementary copy of a mRNA coding for a cytoplasmic human Gln-tRNA synthetase. This conclusion is based mainly on two findings: (i) the available nucleotide sequence allows the prediction of a polypeptide chain with striking homologies to parts of the known E.coli and yeast Gln-tRNA synthetases; (ii) antibodies, prepared against the PZ encoded protein, specifically precipitate a Gln-tRNA synthetase activity from the cytosols of human cells. Yet, the human enzyme has a molecular

mass of about 170 kDa and is thus about three times larger than the E.coli and two times larger than the yeast Gln-tRNA synthetase (4,5). It is also larger than the molecular weight of rabbit and sheep Gln-tRNA synthetase which were previously reported to be around 95 kDa (21). The lower molecular weight is probably due to proteolytic degradation during the preparation. In unpublished experiments we have treated cytosolic cell extracts with limited amounts of trypsin and were able to stain additional polypeptides of 120 and 95 kDa in Western blottings with PZ antibodies (not shown).

The PZ reading frame encodes a sequence of almost 800 amino acids which corresponds to only approximately one half of the total human Gln-tRNA synthetase but includes a region homologous to two thirds of the E.coli enzyme. The homologous part of the human synthetase, about 380 amino acids long (Fig. 1), may be considered to be the enzyme core (1,3) which is essential for the charging of tRNA. It has been noted before that a sequence of 300-350 amino acids may be quite sufficient for this elementary function of an aminoacyl-tRNA synthetase (3,23).

Thus, the core of the Gln-tRNA synthetase was conserved throughout evolution, being very similar in organisms as distantly related as bacteria, yeast and humans. However, the three enzymes differ in sequences linked to the core region. Both, the yeast and the human enzymes have extensions of about 200 amino acids on the N-terminal side of the core. The N-terminal extensions may not be necessary for the enzymatic activity as yeast mutants with large deletions in this region have been constructed and found to be viable (24). However, the available sequence data show that the N-terminal regions of the yeast and the human enzyme are not similar (unpublished data). It is therefore not possible to say yet whether the N-terminal extension of the human Gln-tRNA synthetase may similarly be dispensable for enzymatic function. In addition to the N-terminal extension the human enzyme has a sequence of about 500 amino acids linked to the C-terminal end of the core (Fig.7). This is probably a feature of Gln-tRNA synthetases from other mammalian species as a 5.5 kb mRNA, hybridizing to the PZ cDNA, was also found in monkey and

mouse cells (unpublished). The additional, C-terminal sequences must therefore perform a function important for mammalian Gln-tRNA synthetases.

What could this function be? A characteristic feature of some mammalian aminoacyl-tRNA synthetases is their organisation as a high molecular weight complex which may be composed of up to nine or ten different enzymes including Gln-tRNA synthetase (reviewed in ref.nr.25 and 26). These complexes may be anchored to the endoplasmic reticulum, possibly in functional proximity to sites involved in protein synthesis (27,28).

We have shown that the 170 kDa enzyme, reacting with PZ antibodies, coprecipitates with several other polypeptides. Two of the coprecipitated proteins could be identified as tRNA synthetase, namely Ile-tRNA synthetase and Leu-tRNA synthetase. The presence of these proteins in high molecular weight complexes was independently confirmed by sucrose gradient centrifugation showing that the 170 kDa polypeptide sedimented as a broad peak of more than 20 S (not shown) in agreement with similar results of others (25,26).

It has been postulated that a hydrophobic domain, linked to the enzyme core, may be responsible for complex formation (25,26). It appears quite possible that the long C-terminal extension of the human Gln-tRNA synthetase contains this domain. This view is supported by recent work on mammalian His-tRNA synthetases. This enzyme is not part of a high molecular weight complex but appears to occur in an unassociated form in the cytoplasm (26). The mammalian His-tRNA synthetase shares regions of amino acid similarities with the corresponding enzymes from E.coli and yeast. However, it is composed of only 508 amino acids (29), not much different from the E.coli and cytoplasmic yeast His-tRNA synthetases which are composed of 424 and 526 amino acids, respectively (30,31). Thus, mammalian synthetases which are not components of the high molecular weight complex may be similar in size to their prokaryotic counterparts and lack a large C-terminal extension.

Finally, the antibodies, used to isolate the PZ cDNA clone, were prepared against a DNA polymerase α preparation obtained by conventional column chromatography (10). This

preparation as well as three other conventional polymerase preparations from three different laboratories contained the 170 kDa synthetase polypeptide. In fact, the presence of an aminoacyl-tRNA synthetase activity in a polymerase preparation has been described some time ago (32). We believe, however, that the synthetase was a contaminant which fortuitously copurified with DNA polymerase through several column steps as an essentially homogeneous and fully functional DNA polymerase α preparation, purified by an immunoaffinity procedure (33), did not contain detectable amounts of Gln-tRNA synthetase (not shown).

ACKNOWLEDGEMENTS

We thank to Neal Brown for purification of the PZ-antigen and especially for his valuable and encouraging discussion. We also thank Marjorie Barnes for subcloning PZ in pRIT2T. Uwe Englisch is thanked for the antibodies against yeast synthetases as well as for help with the synthetase assay and for discussion. We are also grateful to Karl Freese who introduced one of us (R.F.) into computer analysis.

This work was supported by Deutsche Forschungsgemeinschaft (SFB 156).

REFERENCES

1. Schimmel, P. (1987) *Ann.Rev.Biochem.* 56, 125-158
2. Webster, T.A., Tsai, H., Kula, M., Mackie, G.A. and Schimmel, P. (1984) *Science* 226, 1315-1317
3. Jasin, M., Regan, L. and Schimmel, P. (1983) *Nature (London)* 306, 441-447
4. Hoben, P., Royal, N. Chueng, A., Yamao, F., Biemann K. and Söll, D. (1982) *J.Biol.Chem.* 257, 11644-11649
5. Ludmerer, S.W. and Schimmel, P. (1987) *J.Biol.Chem.* 262, 10801-10806
6. Palmiter, R.D. (1974) *Biochemistry* 13, 3606-3614
7. Aviv, H. and Leder, P. (1972) *Proc.Natl.Acad.Sci.USA* 69, 1408-1412
8. Gubler, U. and Hoffmann, B.L. (1983) *Gene* 25, 263-269
9. Young, R.A. and Davis, R.W. (1983) *Proc.Natl.Acad.Sci. USA* 80, 1194-1198
10. Riedel, H.D., König, H. and Knippers, R. (1984) *Biochem.Biophys.Acta* 783, 158-165
11. Huygh, T.V., Young, R.A. and Davis, R.W. (1985) in *DNA cloning*, Glover, D.M., (ed.), Vol. I, pp. 49-78, IRL press, Oxford

12. Sanger, Z., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl.Acad.Sci. USA* 74, 5463-5467
13. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning*. Cold Spring Harbor Laboratory, N.Y.
14. Nilsson, B., Abrahmsen, L. and Uhlen, M. (1985) *The EMBO J.* 4, 1075-1080
15. Thömmes, P., Reiter, T. and Knippers, R. (1986) *Biochemistry* 25, 1308-1314
16. Towbin, H., Staehelin, T. and Gordon, J. (1979) *Proc.Natl. Acad.Sci. USA* 76, 4350-4354
17. Laemmli, U.K. (1970) *Nature (London)* 227, 680-685
18. Favaloro, J. Treisman, R. and Kamen, R. (1960) *Meth. Enzymol.* 65, 718-749
19. Thomas, P.S. (1980) *Proc.Natl.Acad.Sci. USA* 77, 5201-5205
20. Wong, S.W., Wahl, A.F., Yuan, P.M., Prai, N., Pearson, B.E., Arai, K., Korn, D., Hunkapiller, M.W. and Wang, T.S.F. (1988) *The EMBO J.* 7, 37-47
21. Lazard, M., Mirande, M. and Waller, J.P. (1985) *Biochemistry* 24, 5099-5106
22. Englisch, U., Englisch, S. Markmeyer, P., Schischkoff, J., Sternbach, H., Kratzin, H. and Cramer, F. (1987) *Biol. Chem. Hoppe-Seyler* 368, 971-979
23. Hall, C.V., van Cleemput, M., Muench, K.H. and Yanofsky, C. (1982) *J.Biol.Chem.* 257, 6132-6136
24. Ludmerer, S.W. and Schimmel, P. (1987) *J.Biol.Chem.* 262, 10807-10813
25. Deutscher, M.P. (1984) *J.Cell.Biol.* 99, 373-377
26. Dang, C.V. and Dang, C.V. (1986) *Biochem. J.* 239, 249-255
27. Dang, C.V., Yang, D.C.H. and Pollard, T.D. (1983) *J.Cell. Biol.* 96, 1138-1147
28. Mirande, M., Le Corre, D., Louvard, D., Reggio, H., Pailliez, J.P. and Waller, J.P. (1985) *Exp.Cell.Res.* 156, 91-102
29. Tsui, F.W.L. and Siminovitch, L. (1987) *Nucleic Acids Res.* 15, 3349-3367
30. Freedman, R., Gibson, B., Donovan, D., Biemann, K., Eisenbeis, S., Parker, J. and Schimmel, P. (1985) *J.Biol. Chem.* 260, 10063-10068
31. Natsoulis, G., Hilger, F. and Fink, G.R. (1986) *Cell* 46, 235-243
32. Rapaport, E., Zamecnik, P.C. and Baril, E.F. (1981) *Proc. Natl.Acad.Sci. USA* 78, 838-842
33. Nasheuer, H.P. and Grosse, F. (1987) *Biochemistry* 26, 8458-8466