Biochemistry

 Volume 34, Number 39 October 3, 1995

Accelerated Publications

Species-Specific Microhelix Aminoacylation by a Eukaryotic Pathogen tRNA Synthetase Dependent on a Single Base Pair

Cheryl L. Quinn,*,‡ Nianjun Tao,‡ and Paul Schimmel§

Cubist Pharmaceuticals, Inc., 24 Emily Street, Cambridge, Massachusetts 02139, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received July 14, 1995®

ABSTRACT: We report here that tyrosyl-tRNA synthetase from the eukaryotic pathogen *Pneumocystis carinii* is a 370 amino acid polypeptide with characteristic elements of a class I aminoacyl-tRNA synthetase and aligns with the prokaryotic tyrosyl-tRNA synthetases in the class-defining active site region, including the tRNA acceptor helix-binding region. The expressed enzyme is a dimer that aminoacylates yeast tRNA but not *Escherichia coli* $tRNA^{Tyr}$. Like most tRNAs, prokaryotic tyrosine tRNAs have a G1·C72 base pair at the ends of their respective acceptor helices. However, the eukaryote cytoplasmic tyrosine tRNAs have an uncommon C1·G72 base pair. We show that *P. carinii* tyrosyl-tRNA synthetase charges a seven base pair hairpin microhelix ($microhelix^{Tyr}$) whose sequence is derived from the acceptor stem of yeast cytoplasmic $tRNA^{Tyr}$. In contrast, the enzyme does not charge *E. coli* $microhelix^{Tyr}$. Changing the C1·-G72 of yeast $microhelix^{Tyr}$ to G1·C72 abolishes charging by the *P. carinii* tyrosyl-tRNA synthetase. Conversely, we found that *E. coli* tyrosyl-tRNA synthetase can charge an *E. coli* $microhelix^{Tyr}$ and that charging is sensitive to having a G1·C72 rather than a C1·G72 base pair. The results demonstrate that the common structural framework of homologous tRNA synthetases has the capacity to coadapt to a transversion in a critical acceptor helix base pair and that this coadaptation can account for species-selective microhelix aminoacylation. We propose that species-selective acceptor helix recognition can be used as a conceptual basis for species-specific inhibitors of tRNA synthetases.

Pneumocystis carinii is an opportunistic pathogen that invades lung cells and causes an often fatal pneumonia in immunocompromised patients (Haverkos, 1984). Once thought to be a protozoan, *P. carinii* is now classified as a fungus on the basis of the sequence of its ribosomal RNA (Edman et al., 1988). The lack of knowledge about the biology and essential proteins of this unicellular eukaryotic microorganism has been a major obstacle to developing new treatments for *P. carinii* infection (Laughon et al., 1991). With this in mind, we have attempted to identify potential intracellular drug targets in this pathogen, focusing on aminoacyl-tRNA synthetases as a group of essential enzymes. These proteins have been widely studied in other organisms (Meinnel et al., 1995) and are the targets of some antibiotics (Hughes & Mellows, 1980). While polypeptide chains for some tRNA synthetases are over 900 amino acids, the tyrosyl-tRNA synthetases are among the smallest of the class I synthetases, typically consisting of only 400−500 amino acids. The reduced size has advantages for structure−function investigations and drug design. For these reasons, we concentrated on obtaining tyrosyl-tRNA synthetase

* Corresponding author. Tel.: (617) 576-1999. FAX: (617) 576-0232.
‡ Cubist Pharmaceuticals, Inc.
§ Massachusetts Institute of Technology.
® Abstract published in *Advance ACS Abstracts,* September 15, 1995.

0006-2960/95/0434-12489$09.00/0

(TyrRS)[1] from *P. carinii* and investigating whether species-specific differences in tRNA acceptor helix recognition by TyrRS might be used as a basis for drug design.

We noted that the acceptor helices for tyrosine tRNAs of prokaryotes and eukrayotic organelles (mitochondria and chloroplasts) had the conventional G1·C72 base pair found in most tRNAs. However, the tyrosine tRNAs of eukaryotic cytoplasma and archaebacteria have an uncommon C1·G72 base pair (Steinberg et al., 1993). Interestingly, even the 20 known tRNA-like structures of plant viral RNAs which are aminoacylated with tyrosine in the host cytoplasma have a C·G base pair at the corresponding position (Florentz & Giegé, 1995). Because the G1·C72 base pair commonly occurs in the acceptor stems of tRNAs for many different amino acids, it is not thought to be important for synthetase discrimination of one G1·C72-containing tRNA from another. However, a C1·G72 base pair is less common. The significance of the C1·G72 *versus* G1·C72 distinction between eukaryotic cytoplasmic and prokaryotic tyrosine-tRNA was indicated when a C1·G72 base pair was shown to be required for aminoacylation of an *Escherichia coli* amber suppressor with tyrosine by yeast extracts (Lee & RajBhandary, 1991). In addition, class II-specific discrimination of the 1·72 base pair has been noted (Liu et al., 1995). Therefore, we thought the C1·G72 base pair of eukaryotic tyrosine tRNA might be required for aminoacylation by the eukaryotic TyrRS. This would also explain, at least in part, the observation by Chow and RajBhandary (1993) that yeast cytoplasmic TyrRS can not charge *E. coli* tRNA.

The importance of acceptor helix interactions in determining aminoacylation specificity has been repeatedly manifested by the specific aminoacylation of seven base pair microhelix substrates that are based on tRNA acceptor stems [see Frugier et al. (1994), Hamann and Hou (1995), and Martinis and Schimmel (1995)]. These substrates enable investigation of acceptor helix interactions in isolation from the rest of the tRNA molecule. Although there have been no previous reports of aminoacylation of microhelices with tyrosine, it was possible that microhelix substrates might be used to test the significance of the 1·72 base pair for aminoacylation by a eukaryotic TyrRS such as the *P. carinii* enzyme.

The gene encoding TyrRS was cloned using a recently described approach based on an alignment-guided cross-species polymerase chain reaction (Shiba et al., 1994). The sequence of the *P. carinii* TyrRS gene aligns with those of *Saccharomyces cerevisiae* cytoplasmic TyrRS and four noneukaryotic tyrosyl-tRNA synthetases. The alignment is the strongest in the N-terminal half which contains the class-defining active site domain with the residues needed for docking of the tRNA acceptor helix. Despite the similarities between the sequences of the *P. carinii* and the bacterial enzymes in the N-terminal domain, we detected aminoacylation of yeast but not *E. coli* tRNA by the *P. carinii* enzyme. This observation instigated our use of RNA microhelix substrates to test the significance of the eukaryote-specific C1·G72 base pair for aminoacylation by *P. carinii* TyrRS and encouraged us to consider whether, in principle, species-selective acceptor helix interactions could be exploited for drug development.

[1] Abbreviations: TyrRS, tyrosyl-tRNA synthetase; CP1, connective polypeptide 1; IPTG, isopropyl β-D-thiogalactopyranoside; SDS, sodium dodecyl sulfate.

MATERIALS AND METHODS

Reagents. Rat-derived *P. carinii* genomic DNA was provided by C. Dykstra, University of North Carolina, and prepared as described by Fletcher et al. (1993). The *P. carinii* cDNA library, cloned into the λZAP expression vector, was obtained from the NIH AIDS Reagent Program. The pGEX-4T-1 expression vector (Pharmacia, Piscataway, NJ) was used to construct and express the GST-fusion protein. Oligonucleotides were synthesized on a Millipore Expedite nucleic acids synthesizer. The RNA microhelices were chemically synthesized on this machine using phosphoramidites from ChemGene (Waltham, MA) and deprotected and purified as described (Musier-Forsythe et al., 1991). Purified *E. coli* $tRNA^{Tyr}$ was purchased from Sigma (St. Louis, MO), and crude Brewer's yeast tRNA was purchased from Boehringer Mannheim Biochemicals (Indianapolis, IN). [^{3}H]Tyrosine for enzyme assays, [α-^{32}P]-dATP for labeling DNA probes, and [α-^{35}S]dATP for DNA sequencing were purchased from DuPont/NEN (Boston, MA).

DNA Manipulations. PCR amplifications were completed in a 50 μL reaction volume containing 2 μM of each degenerate primer or 50 ng of each specific primer, 100 μM of each of the 4 dNTP's, 1.25 units of Taq DNA polymerase (Boehringer Mannheim Biochemicals), 60 mM Tris (pH 8.5), 2 nM $MgCl_2$, and 15 mM NH_4SO. PCR reactions with 40 ng of genomic DNA were cycled 40 times at 94 (30 s), 45 (60 s), and 72 °C (60 s). PCR amplifications from plasmids had 10 ng of plasmid template DNA and were cycled 35 times at 94 (30 s), 50 (60 s), and 72 °C (60 s). PCR fragments were gel purified using the Geneclean II kit (Bio101; Vista, CA). Plasmid DNA (purified with the Promega Wizard Minipreps kit, Madison WI) was sequenced using the Sequenase kit from USB (Cleveland, OH).

Protein Manipulations. The *P. carinii* GST-TyrRS fusion protein was expressed in *E. coli* DH5α cells (GIBCO BRL, Gaithersburg, MD) grown to mid-log phase at 37 °C and then transferred to 18 °C for induction with 0.1 mM IPTG for 3 days. Cells were lysed in phosphate-buffered saline (0.15 M NaCl, 0.15 M $NaPO_4$; pH 7.2) using a French press, and the protein was purified from the soluble fraction on a glutathione−agarose column following the protocols of Pharmacia. The GST fragment was removed by digestion with thrombin (10 units of thrombin/mg of protein at 16 °C for 20 h followed by 2−3 h at room temperature). *E. coli* TyrRS (a generous gift of T. Avruch, Cubist Pharmaceuticals, Inc.) was overexpressed in *E. coli* strain BL21 (DE3) (Strategene, La Jolla, CA) with a C-terminal tag of six histidine residues. Using the procedures described by Qiagen (Chatsworth, CA), the protein was purified on a nickel affinity column (Ni-NTA, Qiagen) from the supernatant of lysed cells.

Denaturing protein gels of 10% polyacrylamide/SDS were prepared and run as described (Laemmli, 1970). Protein concentrations were assayed using the Coomassie Plus protein reagent from Pierce (Rockford, IL) (Bradford, 1976), using a bovine serum albumin standard. N-Terminal sequencing was done by Analytical Biotechnology Services (Boston, MA). Quaternary structures of the GST-TyrRS and the thrombin-treated TyrRS were determined by FPLC gel filtration analysis of the purified proteins using a Superose 6 HR 10/30 column (Pharmacia), equilibrated in 0.05 M

$NaPO_4$ with 0.15 M NaCl (pH 7) at a flow rate of 0.4 mL/min. Molecular weight protein standards were thyroglobulin (670 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa), and vitamin B-12 (1.3 kDa).

TyrRS Enzyme Assays. The aminoacylation assay conditions were modified from Chow and RajBhandary (1993). In general, a 60 μL reaction (at 25 °C) contained 150 mM Tris·HCl (pH 7.5), 150 mM KCl, 10 mM $MgCl_2$, 20 mM β-mercaptoethanol, 4 mM ATP, 20 μM [^{3}H]tyrosine (specific activity 2000 cpm/pmol for assays with tRNA and 6000 cpm/pmol for assays with microhelices), and tRNA or RNA microhelix. The reactions were initiated by adding 10 μL of enzyme sample to a final concentration of 6 nM for assays with full tRNA and 4−8 μM for assays with microhelices.

RESULTS

Cloning of Gene for P. carinii Tyrosyl-tRNA Synthetase. A fragment of the TyrRS gene was amplified by PCR from *P. carinii* genomic DNA. The amplified fragment was then used to screen the *P. carinii* cDNA library for a full-length clone-encoding the enzyme. Because the pathogen must be cultured in rat lung cells, it was not practical to purify the protein for sequence information to design PCR primers. Therefore, we designed degenerate primers not only on the basis of the known coding sequence of the *S. cerevisaie* cytoplasmic TyrRS gene (Chow & RajBhandary, 1993) but also reflecting the codon bias of *P. carinii* (Fletcher et al., 1993). The primers were designed to amplify a short conserved region that aligned with the Rossmann nucleotide-binding domain of the *B. stearothermophilus* enzyme between the second and the fifth β-strands (Brick et al., 1989).

The resulting 29-nucleotide forward primer was AAA YTW TAT TGG GGW ACW GCW CCW ACW GG, where Y indicates a pyrimidine and W indicates A or T and corresponds to codons 40−49 of the yeast cytoplasmic enzyme. The 27-nucletide reverse primer was TTG ATC WAC WCC WCC AAA TTG ACA ATC and corresponds to codons 184−192. The TyrRS fragment amplified from *P. carinii* genomic DNA was ligated into the pT7Blue T-vector (Novagen, Madison, WI) to yield plasmid pQY11. Sequencing of the complete PCR fragment on both strands showed that the *P. carinii* TyrRS PCR fragment was 590 bp in length and had an AT content of 70%. On the basis of the alignment of open reading frames with the yeast gene (Chow & RajBhandary, 1993) and consensus splice sites for *P. carinii* (Edlind et al., 1992), three introns were identified in the first 250 bp of the fragment. These three introns are similar to other introns reported in *P. carinii* genomic DNA: they are AT rich (84%, 78%, and 75% AT); all three contain a GTA 5′ splice site; all three contain an AG 3′ splice site; and they are all shorter than 100 nucleotides in length.

The *P. carinii* cDNA library was screened with a ^{32}P-labeled probe of the plasmid pQY11 insert. There were six positive plaques identified out of approximately 50 000 plaques of the *P. carinii* cDNA library which were screened; three of these hybridized to the probe when purified to homogeneity. One isolate, plasmid pNTY4, contained a 1.35 kbp *Eco*RI−*Xho*I insert which was completely sequenced on both strands. Of the other isolates, one had a shorter insert (about 1 kbp) and the third isolate was missing the *Xho*I site.

Sequence of the P. carinii Tyrosyl-tRNA Synthetase. The plasmid pNTY4 insert encoded the full-length cDNA for the *P. carinii* TyrRS gene. The complete gene sequence from the cDNA clone included 100 bases of untranslated DNA before the ATG start codon. Extending for 1113 nt, the open reading frame ends with a TAA stop codon about 100 nt before the polyadenylation site and encodes a protein of 370 amino acids with a predicted $M_r = 41\,000$. As expected for *P. carinii* cDNAs (Fletcher et al., 1993), this one is 69% AT and the codon usage is biased toward an A or T at the wobble nucleotide.

Figure 1 presents the translated amino acid sequence of the *P. carinii* TyrRS gene aligned with the known sequences of TyrRS from *S. cerevisiae* (cytoplasmic), *Bacillus caldotenax*, *Bacillus subtilis*, *Bacillus stearothermophilus*, and *E. coli*. Also shown are the locations of secondary structure elements for the *B. stearothermophilus* enzyme (Brick et al., 1989). The nucleotide-binding fold of alternating α-helices and β-strands is comprised of strands S1−S6 and helices H2−H4 and H9 and extends to just beyond the "KMSKS" element (Figure 1). The nucleotide-binding fold is divided by the connective polypeptide insertion known as CP1. The six sequences align well throughout their entire lengths, including CP1. The *P. carinii* amino acid sequence is 60% identical to the yeast cytoplasmic TyrRS, while both *S. cerevisiae* and *P. carinii* TyrRS sequences are less than 25% identical to any of the published mitochondrial or bacterial TyrRS sequences. Most of the overall homology is due to similarities in the N-terminal region; the C-terminal regions have few conserved residues between any species.

The 370 amino acid *P. carinii* TyrRS is the shortest of the six sequences (Figure 1). Class I synthetases in *E. coli* (an organism for which all the aaRS genes have been sequenced) vary in length from 334 amino acids (tryptophanyl-tRNA synthetase) to 951 amino acids (valyl-tRNA synthetase). Thus, the *P. carinii* TyrRS at 370 amino acids is one of the shortest known class I tRNA synthetases. The six TyrRS enzymes vary in length from 370 to 423 amino acids, with the *P. carinii* enzyme missing C-terminal residues present in the other enzymes after helix 14 (H14).

As a class I synthetase, TyrRS is expected to have the 11 amino acid signature sequence which ends in the HIGH tetrapeptide and a region containing a conserved KMSKS sequence (Burbaum & Schimmel, 1991). The TyrRS structure and the closely related tryptophanyl-tRNA synthetase structure show that these conserved sequences form parts of the nucleotide-binding pocket (Doublié et al., 1995). However, the HIGH sequence in *P. carinii* TyrRS and in cytoplasmic *S. cerevisiae* TyrRS (Chow & RajBhandary, 1993) is actually HCGY and the KMSKS sequence is actually KMSAS. The H → Y change of the second histidine in the HIGH sequence is notable in light of the proposed role for the second histidine in stabilizing the binding of ATP by hydrogen bonds in the *B. stearothermophilus* TyrRS (Lowe et al., 1985). Furthermore, the second K of the KMSKS is proposed to be involved in stabilizing the tyrosyl adenylate intermediate in *B. stearothermophilus* TyrRS (First & Fersht, 1995), a role unlikely to be filled by the A in this position of the eukaryotic enzymes. Thus, the details of how a conserved sequence element interacts with ATP may vary from enzyme to enzyme.

Expression and Characterization of P. carinii TyrRS. The primer PcY5XE was designed to allow PCR amplification of the TyrRS gene, without untranslated DNA sequences, for cloning as an in-frame gene fusion to the glutathione-

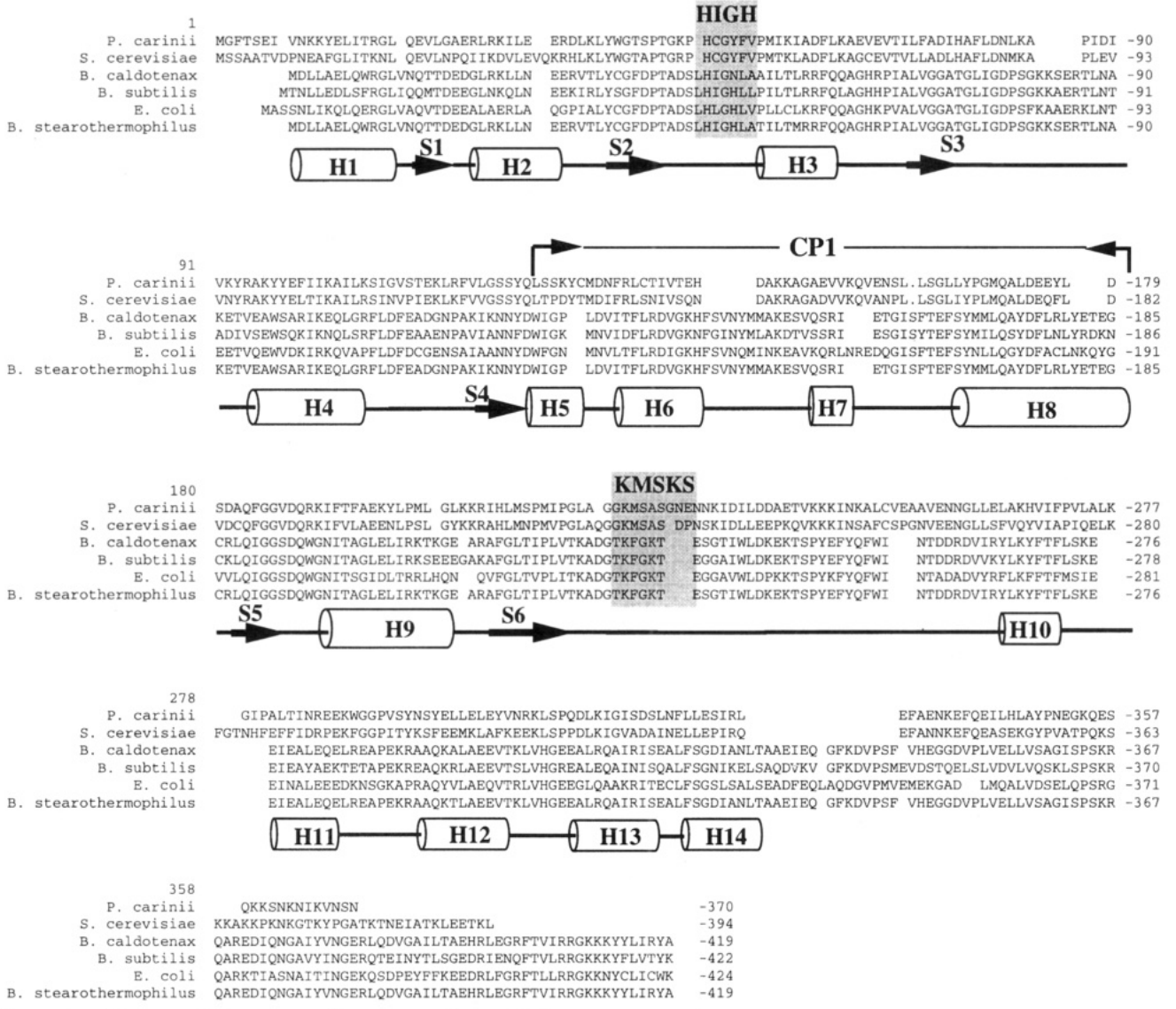

FIGURE 1: Alignment of the amino acid sequences for TyrRS from various species. The sequences shown are for TyrRS of *P. carinii* cytoplasmic (this work), *S. cerevisiae* cytoplasmic (Chow & RajBhandary, 1993), *B. caldotenax* (Jones et al., 1986), *B. subtilis* (Henkin et al., 1992), *E. coli* (Barker et al., 1982), and *B. stearothermophilus* (Winter et al., 1983). Numbers above the species names on the left correspond to positions of the *P. carinii* residues, while numbers on the right correspond to positions specific for each species. The "HIGH" and "KMSKS" regions (shaded) start at positions 51 and 223, respectively, of the *P. carinii* sequence, and Glu152 of the *B. stearothermophilus* sequence aligns with Lys149 of *P. carinii* TyrRS. Elements of the secondary structure based on the three-dimensional structure of *B. stearothermophilus TyrRS* (Brick et al., 1989) are shown below the sequences. Secondary structures for the C-terminal end were not resolved in the crystal structure. Cylinders depict helices, numbered from H1 to H14, and arrows depict β-strands, numbered from S1 to S6. The CP1 polypeptide which splits the nucleotide-binding fold is indicated above the sequences.

S-transferase gene (GST) in the expression vector pGEX-4T-1. Primer PcY5XE has the following sequence: cgc *ggatcc* ATG GGA TTT ACA AGT GAA ATA G, where the italicized nucleotides are a *Bam*HI restriction site and the capitalized nucleotides are the codons for the N-terminal amino acids of the *P. carinii* TyrRS open reading frame. The primer was paired with the commercially available universal primer (which anneals downstream from the *Xho*I clonging site) to amplify *P. carinii* TyrRS from plasmid pNTY4. The resulting PCR fragment was cloned into the *Bam*HI and *Xho*I sites of plasmid pGEX-4T-1 to form plasmid pGEXYG18. This plasmid was sequenced to verify the insertion of the gene with the correct sequence at the expected cloning sites.

E. coli DH5α cells transformed with pGEXYG18 were used to express the *P. carinii* TyrRS with an N-terminal fusion to GST. Most of the expressed protein appeared to form insoluble inclusion bodies when the cells were grown at 37 °C; therefore, the protein was expressed in *E. coli* DH5α cells grown at 18 °C and induced with 0.1 mM IPTG for 3 days. The GST-TyrRS fusion protein, with an M_r of 69 000, was found in both the soluble and the insoluble fractions of the lysed cells. After affinity purification of the GST fusion protein from the soluble fraction, the GST fragment was removed by digestion with thrombin. Approximately 10 mg of purified GST-TyrRS protein was obtained from 1 L of culture after 3 days of induction and growth at 18 °C. After thrombin cleavage for over 20 h, about 80% of pure TyrRS protein was recovered. N-Terminal sequencing of this thrombin-treated protein showed that cleavage occurred at the engineered fusion site, leaving the dipeptide Gly-Ser attached to the initiator methionine of TyrRS.

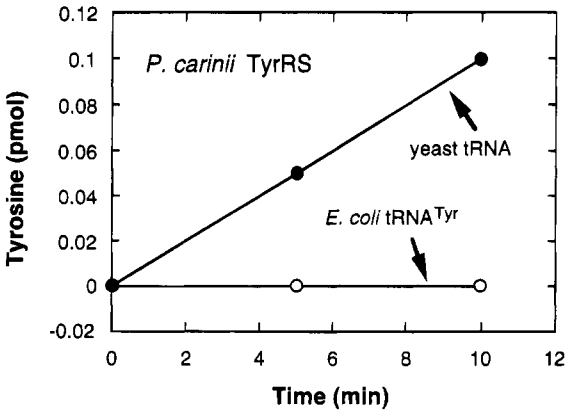

FIGURE 2: Aminoacylation of tRNA by *P. carinii* GST-TyrRS. These assays were performed as described in the Materials and Methods using 6 nM enzyme and 45 μM crude yeast tRNA for 12 nM enzyme with 2 μM *E. coli* tRNA^{Tyr} (type II).

Because both subunits of the bacterial TyrRS homodimer are required for binding tyrosine-tRNA (Labouze & Bedouelle, 1989), we investigated the quaternary structure of the *P. carinii* enzyme. Size analysis of the purified proteins by gel filtration showed that both the GST-TyrRS fusion and free TyrRS form functional dimers. The GST-TyrRS protein eluted with the major peak corresponding to 130 kDa, while the thrombin-treated protein gave a major peak corresponding to 97 kDa, with an error rate of about 10% in estimation of elution times for protein peaks (data not shown). Therefore, these protein peaks correspond to dimers which would have predicted sizes of 134 and 82 kDa, respectively.

P. carinii TyrRS Aminoacylation of tRNA. Function of the *P. carinii* GST-TyrRS and free TyrRS was tested using a crude preparation of yeast tRNA. Both *P. carinii* GST-TyrRS and free TyrRS charged yeast tRNA but not *E. coli* tRNA^{Tyr} (Figure 2). The apparent K_M for tyrosine for both forms of the enzyme was 10 μM for tyrosine concentrations in the range 2.5−40 μM. [In comparison, the *B. stearothermophilus* TyrRS has a K_M for tyrosine of 2 μM for tyrosine concentrations below 40 μM (Wilkinson et al., 1983)]. Because of the similar K_M values and the apparent dimerization of both forms of the TyrRS, it was assumed that the GST fusion did not interfere with the enzyme activity of this protein. Subsequent characterizations were done with purified GST-TyrRS.

P. carinii Tyrosyl-tRNA Synthetase Aminoacylates RNA Microhelices with a Species-Specificity That Recapitulates That Seen with tRNAs: Role of the Terminal Base Pair. We noted that sequences of six bacterial, three mitochondrial, and one chloroplast tyrosine tRNAs have a G1·C72 base pair at the beginning of the acceptor stem (Steinberg et al., 1993). (The exception is the *N. crassa* mitochondrial tRNA^{Tyr} which has an A1·U72 base pair.) In contrast, all 13 published cytoplasmic tRNA^{Tyr} sequences have a C1·G72 base pair. Because it is an uncommon base pair, this C1·G72 could be important in distinguishing tRNA^{Tyr} from other tRNAs in the eukaryotic cytoplasm. We speculated that the species-specific aminoacylation of yeast *versus E. coli* tRNA^{Tyr} by the *P. carinii* enzyme could be explained by the C·G *versus* G·C difference at the end of the acceptor stem.

Synthetic RNA hairpin microhelices are sequence-specific substrates for aminoacylation with at least nine different aminoacyl tRNA synthetases [see Frugier et al. (1994), Hamann and Hou (1995), and Martinis and Schimmel (1995)]. The sequences of these substrates are based on those of the tRNA acceptor stems that end in the N^{73}CCA single-stranded tetranucleotide, where N^{73} is the so-called discriminator base. Although aminoacylation efficiency of the microhelices can be orders of magnitude lower than with the full tRNA, the reactions are highly specific with specificity determined by the N^{73} nucleotide and one or two base pairs within the first four of the acceptor stem. Thus, these microhelix substrates provide a means to dissect functional synthetase-acceptor stem contacts in isolation from interactions with the rest of the tRNA molecule. For these reasons, we investigated the ability of *P. carinii* TyrRS to specifically aminoacylate a microhelix based on the sequence of the acceptor stem of yeast tRNA^{Tyr} and tested the significance of the C1·G72 base pair. Because there is no sequence available for any *P. carinii* tRNA, we used microhelix sequences based on the GψA isoacceptor of yeast tRNA^{Tyr} (Figure 3).

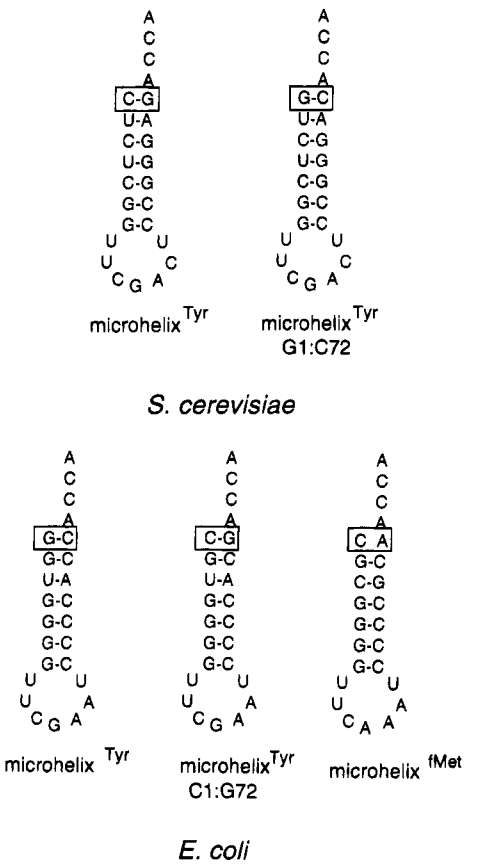

FIGURE 3: RNA microhelices. The 1·72 base pair of each microhelix is boxed. The synthetic yeast and *E. coli* microhelix structures, derived from the acceptor stems and TψC loops of an *S. cerevisiae* cytoplasmic tyrosine tRNA and of *E. coli* tyrosine tRNA.

A synthetic microhelix based on the yeast tRNA^{Tyr} was charged with tyrosine (Figure 4, top). This aminoacylation was specific and depended on more than just the A73 discriminator base, because a microhelix based on the acceptor stem of *E. coli* tRNA^{fMet} (Martinis & Schimmel, 1992), which has A73, was not acylated. Reflecting the species-specificity seen with the full tRNAs (Figure 2), the *E. coli* microhelix^{Tyr} was not aminoacylated with the *P. carinii* TyrRS.

To test the role of the first base pair in recognition of these microhelices, we changed the C1·G72 base pair of yeast

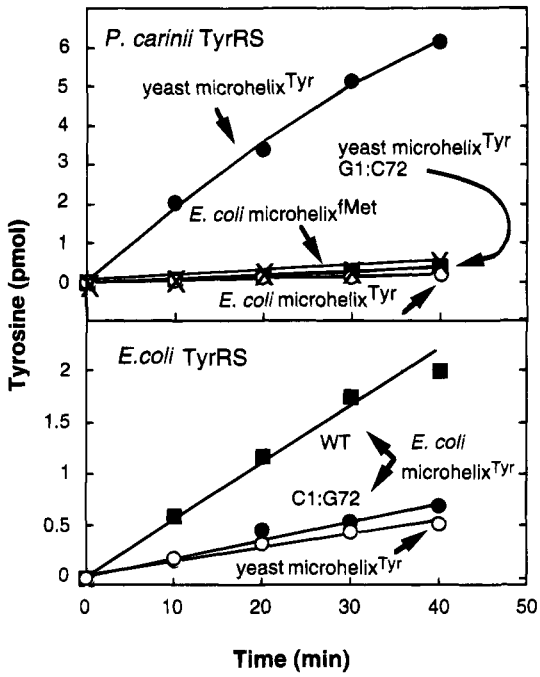

FIGURE 4: Aminoacylation of RNA microhelices. Top, aminoacylation of microhelices by *P. carinii* GST-TyrRS was assayed with 4 μM enzyme and 200 μM yeast microhelix^{Tyr}, 4 μM enzyme and 200 μM *E. coli* microhelix^{Tyr}, 8 μM enzyme and 200 μM *E. coli* microhelix^{fMet}, or 8 μM enzyme and 400 mM yeast G1·C72 microhelix^{Tyr}. Bottom, aminoacylation of RNA microhelices by *E. coli* TyrRS was assayed with 6 μM *E. coli* TyrRS and 200 μM microhelices.

microhelix^{Tyr} to G·C. This single change abolished aminoacylation by the *P. carinii* enzyme (Figure 4, top). Therefore, the C1·G72 base pair is essential for aminoacylation of microhelices with *P. carinii* TyrRS.

Given that the *P. carinii* and *E. coli* enzymes are homologous proteins derived from the same ancestor (Figure 1), we predicted that the *E. coli* TyrRS might also charge a microhelix, but one based on the *E. coli* tRNA^{Tyr} isoacceptor. In contrast to the *P. carinii* TyrRS, we found that *E. coli* TyrRS aminocylated *E. coli* microhelix^{Tyr} but had significantly reduced activity with yeast microhelix^{Tyr} (Figure 4, bottom). This aminoacylation was also sensitive to the 1·72 base pair, because a simple G·C → C·G transversion significantly reduced the efficiency of aminoacylation (Figure 4, bottom). It should be noted that early work on *E. coli* tRNA^{Tyr} amber suppressors showed that 1·72 substitutions resulted in the loss of specificity of charging (Celis, 1979).

DISCUSSION

The work reported here represents the first demonstration of the aminoacylation of a microhelix with tyrosine. Other tRNA synthetases that show activity on microhelices include those for alanine, aspartate, cysteine, glycine, histidine, isoleucine, methionine, serine, and valine (Frugier et al., 1994; Hamann & Hou, 1995; Martinis & Schimmel, 1995). The efficiency of aminoacylation varies greatly in these systems and appears to reflect the degree to which interactions outside of the acceptor stem contribute to the efficiency of aminoacylation of the full tRNA. In general, the highest efficiency has been seen with the class II enzymes such as alanyl-, aspartyl-, and histidyl-tRNA synthetases. As is the case for other class I enzymes studied, the efficiency of microhelix aminoacylation of *P. carinii* and *E. coli* tyrosyl-tRNA synthetases is reduced by several orders of magnitude compared to that observed with full tRNA. Nevertheless, in all cases, including the examples investigated here, aminoacylation is highly specific as demonstrated by sensitivity to particular nucleotides in the microhelix. The sequence-specific aminoacylation of microhelices has been proposed to represent an operational RNA code for amino acids which was possibly one of the earliest relationships between RNA sequence/structure and specific amino acids (Schimmel et al., 1993). The existence of an operational code is consistent with proposals that the microhelix-containing domain of the tRNA molecule arose earlier than and had an origin independent from the anticodon-containing domain (Maizels & Weiner, 1993; Noller, 1993).

The present results are a demonstration of how tRNA synthetases have coadapted in evolution to changes in critical acceptor stem nucleotides. The results differ in some ways, however, from those recently reported for the class II glycine system. Human and *E. coli* glycyl-tRNA synthetases do not cross-acylate their respective tRNAs (Shiba et al., 1995). This species-specificity is paralleled by species-specific aminoacylation of minihelix substrates based on the acceptor-TψC stems of the corresponding glycine tRNAs (Hipps et al., 1995). In contrast to the tyrosyl enzymes, the sequences of the *E. coli* and human glycyl enzymes give no evidence that they arose from a common ancestor. In spite of their sequence dissimilarities, aminoacylation by both enzymes is sensitive to the N73 nucleotide (U73 for *E. coli* minihelix^{Gly} and A73 for human minihelix^{Gly}) and this nucleotide difference accounts for the species-specific aminoacylation (Hipps et al., 1995). Thus, the location of critical nucleotides within the RNA helix, if not the nucleotides themselves, is conserved for enzymes whether or not they are clearly related in evolution. This conclusion emphasizes that the operational RNA code should be thought of as an RNA–protein complex that coadapts at particular places in the RNA structure to selective pressures on both components of the complex.

Although the crystal structure of the *B. stearothermophilus* TyrRS has been determined (Brick et al., 1989), only the active site domain has been resolved and no cocrystals with tRNA^{Tyr} have been reported. Therefore, we cannot make a reasonable guess as to the locations of the residues in the TyrRS sequences which explain the G1·C72 *versus* C1·G72 selectivities of the *E. coli* and *P. carinii* enzymes. However, in the cocrystal of the class I *E. coli* glutaminyl-tRNA synthetase with tRNA^{Gln}, interactions near the end of the acceptor helix are mediated by residues in the CP1 polypeptide (Rould et al., 1989). In *B. stearothermophilus* TyrRS, Glu152 within CP1 has been mapped as proximal to A73 and C74 of the bound tRNA^{Tyr} acceptor stem (Vidal-Cross & Bedouelle, 1992) and shown by mutational analysis to be important for "negative discrimination" of noncognate tRNAs (Bedouelle et al., 1993). The role of Glu152 is consistent with the idea that acceptor stem interactions are mediated by residues within CP1 and that differences within CP1 may account for G1·C72 *versus* C1·G72 discrimination.

The clear distinction between eukaryotic and bacterial TyrRS acceptor helix recognition offers an opportunity to find or design drugs which selectively bind to the position 1·72 acceptor helix recognition site on the enzyme. Because the fine structure of this site must differ between prokaryotes and eukaryotes, it should be possible, in principle, to find

drugs that are able to interact with either the bacterial or the eukaryotic TyrRS but not with both. Of further interest are the human enzyme and the question of whether enough differences occur in acceptor helix interactions at places other than the 1·72 position to find a drug selective for the *P. carinii* TyrRS that does not interfere with the human protein.

ACKNOWLEDGMENT

We thank Dr. Jian-Ping Shi for his guidance with the enzyme assays and kinetic analyses, Dr. Susan Martinis for microhelix synthesis as well as for helpful comments on the manuscript, and Dr. Lluís Ribas de Pouplana for constructing Figure 1. The *P. carinii* cDNA Library from BioTechnology General was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH.

REFERENCES

Barker, D. G., Burton, C. J., & Winter, G. (1982) *FEBS Lett. 150*, 419−423.
Bedouelle, H., Guez-Ivanier, V., & Nageotte, R. (1993) *Biochimie 75*, 1099−1108.
Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
Brick, P., Bhat, T. N., & Blow, D. M. (1989) *J. Mol. Biol. 208*, 83−98.
Burbaum, J. J., & Schimmel, P. (1991) *J. Biol. Chem. 266* 16965−16968.
Celis, J. E. (1979) in *Transfer RNA: Structure, Properties, and Recognition* (Schimmel, P. R., Söll, D. & Abelson, J. N., Eds.) pp 539−545, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Chow, C. M., & RajBhandary, U. L. (1993) *J. Biol. Chem. 268*, 12855−12863.
Doublié, S., Bricogne, G., Gilmore, C., & Carter, C. W., Jr. (1995) *Structure 3*, 17-31.
Edlind, T. D., Bartlett, M. S., Weinberg, G. A., Prah, G. N., & Smith, J. W. (1992) *Mol. Microbiol. 6*, 3365−3373.
Edman, J. C., Kovacs, J. A., Masur, H., Santi, D. V., Elwood, H. J., & Sogin, M. L. (1988) *Nature 334*, 519−522.
First, E. A., & Fersht, A. R. (1995) *Biochemistry 34*, 5030−5043.
Fletcher, L. D., Berger, L. C., Peel, S. A., Baric, R. S., Tidwell, R. R., & Dykstra, C. C. (1993) *Gene 129*, 167−174.
Florentz, C., & Giegé, R. (1995) in *tRNA: Structure, Biosynthesis, and Function* (Söll, D., & RajBhandary, U., Eds.) pp 141−163, ASM Press, Washington, DC.
Frugier, M., Florentz, C., & Giegé, R. (1994) *EMBO J. 13*, 2219−2226.
Hamann, C. S., & Hou, Y.-M. (1995) *Biochemistry 34*, 6527−6532.
Haverkos, H. W. (1994) *Am. J. Med. 76*, 501−508.
Henkin, T. M., Glass, B. L., & Grundy, F. J. (1992) *J. Bacteriol. 174*, 1299−1306.
Hipps, D. Shiba, K., Henderson, B., & Schimmel, P. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 5550−5552.
Hughes, J., & Mellows, G. (1980) *Biochem. J. 191*, 209−219.
Jones, M. D., Lowe, D. M., Borgford, T. & Fersht, A. R. (1986) *Biochemistry 25*, 1887−1891.
Labouze, E., & Bedouelle, H. (1989) *J. Mol. Biol. 205*, 729−735.
Laemmli, U. K. (1970) *Nature 227*, 680−685.
Laughon, B. E., Allaudeen, H. S., Becker, J. M., Current, W. L., Feinberg, J., Frenkel, J. K., Hafner, R., Hughes, W. T., Laughlin, C. A., Meyers, J. D., Schrager, L. K., & Young, L. S. (1991) *J. Infect. Dis. 164*, 244−251.
Lee, C. P., & RajBhandary, U. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 11378−11382.
Liu, H., Kessler, J., Peterson, R., & Musier-Forsyth, K. (1995) *Biochemistry 34*, 9795−9800.
Lowe, D., Fersht, A. R., & Wilkinson, A. J. (1985) *Biochemistry 24*, 5106−5109.
Maizels, N., & Weiner, A. M. (1993) in *The RNA World* (Gesteland, R. F., & Atkins, J. F., Eds.) pp 577−602, Cold Spring Harbor Laboratory Press, New York.
Martinis, S. A., & Schimmel, P. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 65−69.
Martinis, S. A., & Schimmel, P. (1995) in *tRNA: Structure, Biosynthesis and Function* (Söll, D., & RajBhandary, U., Eds.) pp 349−370, American Society for Microbiology, Washington, DC.
Meinnel, T., Mechulam, Y., & Blanquet, S. (1995) in *tRNA: Structure, Biosynthesis and Function* (Söll, D., & RajBhandary, U., Eds.) pp 251−292, ASM Press, Washington, DC.
Musier-Forsyth, K., Scaringe, S., Usman, N., & Schimmel, P. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 209−213.
Noller, H. F. (1993) in *The RNA World* (Gesteland, R. F., & Atkins, J. F., Eds.) pp 137−156, Cold Spring Harbor Laboratory Press, New York.
Rould, M. A. Perona, J. J., Söll, D., & Steitz, T. A. (1989) *Science 246*, 1135−1142.
Schimmel, P., Giegé, R., Moras, D., & Yokoyama, S. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8763−8768.
Shiba, K., Schimmel, P., Motegi, H., & Noda, T. (1994) *J. Biol. Chem. 269*, 30049−30065.
Steinberg, S., Misch, A., & Sprinzl, M. (1993) *Nucleic Acids Res. 21*, 3011−3015.
Vidal-Cross, A., & Bedouelle, H. (1992) *J. Mol. Biol. 223*, 801−810.
Wilkinson, A. J., Fersht, A. R., Blow, D. M., & Winter, G. (1983) *Biochemistry 22*, 3581−3586.
Winter, G., Koch, G. L. E., Hartley, B. S., & Barker, D. (1983) *Eur. J. Biochem. 132*, 383−387.

BI951616H