Nucleic Acids Research, 1993, Vol. 21, No. 3 615–620

Molecular cloning, sequence, structural analysis and expression of the histidyl-tRNA synthetase gene from *Streptococcus equisimilis*

Corazon A.Menguito, Michael J.Keherly, Chuan-ye Tang, John Papaconstantinou and Paul H.Weigel*

Department of Human Biological Chemistry & Genetics, The University of Texas Medical Branch, Galveston, TX 77555-0647, USA

Received September 30, 1992; Revised and Accepted December 18, 1992 EMBL accession no. Z17214

ABSTRACT

The histidyl-tRNA synthetase gene (*hisS*) from *Streptococcus equisimilis* was cloned and sequenced. The gene for this aminoacyl-tRNA synthetase has an open reading frame of 1278 nucleotides. The deduced amino acid sequence encodes a protein of 426 amino acids with MW = 47,932. The protein is predicted to be soluble with a pI = 5.27. The protein sequence has extensive overall identity/similarity with the *Escherichia coli* and the yeast histidyl-tRNA synthetases (~58% and ~20%, respectively). A putative promoter for gene transcription lies within two hundred nucleotides of the polypeptide start codon. The enzyme was overexpressed, to a level of about 18% of total cellular protein, as a fusion protein (containing an additional 15 amino acids) in *E.coli* using the pT7 expression system containing the T7 RNA polymerase/promoter (Tabor and Richardson, *Proc. Natl. Acad. Sci. U.S.A. 82:1074–1078, 1985). The predicted MW for the hisS* gene product is in good agreement with the size of the fusion protein determined by SDS-PAGE (M_r = 53,700). Amino acid sequencing of the intact fusion protein and proteolytic fragments confirmed the deduced sequence of the synthetase at many positions throughout the protein. The expressed protein catalyzed the specific aminoacylation of $tRNA^{His}$ *in vitro*.

INTRODUCTION

Aminoacyl-tRNA synthetases are a family of structurally diverse enzymes that catalyze the covalent addition of amino acids to the 3′ terminus of their cognate tRNA molecules. These enzymes are responsible for ensuring the fidelity of the mRNA translation process. All twenty *Escherichia coli* tRNA synthetases have now been cloned and sequenced (1,2). Despite catalyzing the same reaction of attaching amino acids to tRNA molecules, tRNA synthetases exhibit a wide diversity in size and subunit compositions. The sizes of the polypeptides range from 303 amino acids for the α subunit of *E.coli* GlyRS (3) to 1440 amino acids for the human GlnRS (4), the largest monomeric tRNA synthetase. The quaternary structures of the synthetases are also diverse, ranging from monomers to homotetramers (α, α_2, α_4) or heterotetramers ($\alpha_2\beta_2$). For the twenty *E.coli* enzymes, the native molecular masses range from 52,000 for CysRS (5) to 380,000 for AlaRS (6,7).

Based on the presence of conserved signature sequences and certain sequence motifs (1,2,8), the twenty synthetases can be divided into two types, Class I and Class II, consisting of ten synthetases each. Class I synthetases contain the signature sequence HIGH and the KMSKS motif. On the other hand, the Class II synthetases share three regions of degenerate similarity known as motifs 1, 2, and 3. These motifs are conserved sequences of 10–12 amino acids. Two motifs contain a highly conserved and characteristic tetrapeptide sequence. Motif 1 is found only in synthetases that form α_2 dimers and is part of the interface region between dimers. Motif 2 is characterized by the consensus tetrapeptide FRNE. Motif 3, which is within the ATP-binding domain, contains the consensus sequence GLER. The active site of the Class II synthetases appears to be formed in part from residues making up motif 2 and/or motif 3 (1). To date, the 3-dimensional structures of five aminoacyl-tRNA synthetases have been determined: *Bacillus stearothermophilus* TyrRS (9), *E.coli* GlnRS (10), MetRS (11), and SerRS (12), and *Saccharomyces cerevisiae* AspRS (13). The first three are Class I, while the latter two are Class II synthetases.

Here we report to the best of our knowledge the first sequence of an aminoacyl- tRNA synthetase gene from the Gram-positive *Streptococcus* species. The *hisS* gene encodes the histidyl-tRNA synthetase enzyme (HisRS), which belongs to the family of Class II synthetases.

EXPERIMENTAL PROCEDURES

Materials

Restriction endonucleases and T4 DNA ligase were from International Biotechnologies, Inc. (IBI) and Promega. Calf intestinal alkaline phosphatase and nick translation kit were from Bethesda Research Lab. The Cyclone kit and the TAQuence

* To whom correspondence should be addressed

sequencing kit were from IBI and U.S. Biochemical Corp. (USB), respectively. X-gal and isopropylthiogalactoside (IPTG) were from Research Organics and USB. [α-^{32}P]dCTP and [α-^{35}S]dATP were from DuPont New England Nuclear. Tris−HCl, EDTA, trypsin (TPCK-treated), V-8 protease, antibiotics, bulk *E.coli* tRNA and other reagents were from Sigma Chemical Co. L-[2,5-^{3}H]histidine was from the Amersham Corp. Scintillation fluids Bio-Safe II and 3a70B were from Research Products International Corp.

Bacterial strains and growth media

Streptococcus equisimilis strain D181 (group C) was from the Rockefeller University Collection. *Escherichia coli* strain HB101 was used as the host in the cloning and the overexpression of the *hisS* gene. *E.coli* strain JM101 was used in the isolation of single-stranded template DNA for sequencing. *E.coli* strains were routinely grown at 37°C in LB medium. D181 was grown in brain heart infusion broth (Difco). When appropriate, antibiotics were added at the following concentrations (in μg/ml): ampicillin, 50; tetracycline, 12.5; kanamycin, 40−50; chloramphenicol, 30.

Molecular cloning of the *S. equisimilis hisS* gene

The shuttle vector pSA3 (a gift from Dr. J.J. Ferretti, Univ. of Oklahoma Health Sciences Center; 14) was used in the cloning of the *hisS* gene. Standard methods for molecular cloning, transformation of *E.coli* HB101, Southern blotting, and colony screening were according to Sambrook *et al.* (15). Chromosomal DNA from *S. equisimilis* was isolated by the method of Marmur (16). DNA sequencing was performed by the dideoxy chain termination method of Sanger *et al.* (17) with the TAQuence sequencing kit using M13mp19 (18) as vector. Overlapping deletion subclones of each DNA strand were constructed using the Cyclone kit according to the manufacturer's protocols. Analyses of nucleotide and amino acid sequences were performed using PC/GENE software (IntelliGenetics, Inc.).

Expression of the *hisS* gene as a fusion protein in *E.coli* and cell extract preparation

Since the *hisS* gene product was not expressed well in *E.coli*, a fusion protein capable of overexpression was constructed. The subcloning of the coding region of the *hisS* gene into pT7-7 and overexpression in *E.coli* were based on the method of Tabor and Richardson (19). *E.coli* HB101 carrying pGP1-2 and pT7-7*rec*III (37) was grown overnight at 30°C in enriched medium (2% tryptone, 1% yeast extract, 0.5% NaCl, 0.2% glycerol, 50 mM potassium phosphate, pH 7.2) with 50 μg/ml each of ampicillin and kanamycin. The culture was induced at 42°C for 30 min, rifampicin was added to a final concentration of 200 μg/ml, and the cells were incubated for an additional 2 h at 37°C. The cells were harvested by centrifugation at 3,200×g for 10 min, washed with M9 medium and resuspended in sonication buffer (10 mM Tris−HCl, pH 8, 1 mM EDTA, 2 mM N-ethylmaleimide, 10 μM phenylmethyl-sulfonyl fluoride). The cell suspension was sonicated, centrifuged at 13,000×g for 10 min to pellet cell debris, and then centrifuged at 100,000×g for 60 min to obtain the soluble fraction (S100). The S100 fraction was used to measure charging activity of the expressed protein. Details of the overexpression system, the purification of HisRS and characterization of the enzyme will be presented elsewhere (37).

Protein sequencing

To confirm the N-terminal deduced amino acid sequence of the HisRS, cell pellets from induced HB101 (containing pGP1-2 and pT7-7*rec*III) were resuspended in sample buffer containing SDS and β-mercaptoethanol, boiled for 5 min, and analysed by SDS-PAGE on a 10% gel as described by Laemmli (20). The proteins were transferred by electroblotting onto an Immobilon-P membrane and stained with Coomassie blue (1 mg/mL in 50% methanol). The stained HisRS fusion protein band was excised and sequenced using an Applied Biosystems 475A Protein Sequencer (Protein Core Facility, UTMB). Peptides were prepared by digesting 5 nmol of purified fusion protein with 2% (w/v) Trypsin or V-8 protease at 37°C overnight. The digestion mixture was fractionated over a Synchropak C-8 reverse phase HPLC column (6.5 μ, 100×4.1 mm from Alltech Associates Inc.) and purified peptides were sequenced as above. Yields typically ranged from 30 down to 0.8 pmol per cycle.

tRNA charging assays

The aminoacylation assay was based on the method of Lin *et al.* (21). A reaction mixture of 90 μL consisted of 50 mM Tris−HCl, pH 7.4, 80 mM KCl, 20 mM $MgCl_2$, 10 mM ATP, 0.1 mM EDTA, 0.5 mM dithiothreitol, 2.2 mg/mL total *E.coli* tRNA, and 0.76 μM L-[2,5-^{3}H]histidine (44 C_i/mmol). The reaction was started by the addition of the S100 fraction at 37°C. At 20,40 and 60 sec, a 25 μL aliquot of the reaction mix was quenched by spotting onto 24 mm discs of Whatman 3MM paper. The discs were washed 3 times in cold 5% trichloroacetic acid, once in 95% ethanol and then air dried according to a modified method of Mans and Novelli (22). The radioactivity was then determined in Bio-Safe II scintillation fluid using a Beckman model LS7500 liquid scintillation spectrometer.

RESULTS

Molecular cloning of the *S. equisimilis hisS* gene

The cloning of the *hisS* gene utilized the shuttle vector pSA3 (14), which can replicate in both *E.coli* and *Streptococcus*. An EcoRI library was constructed from *S. equisimilis* D181 genomic DNA. Recipient HB101 transformant colonies were screened for the presence of a polysaccharide capsule. This screening strategy was devised for the isolation of a different *Streptococcus* gene, a polysaccharide synthase. Based on this screening, we initially isolated a plasmid carrying a truncated gene; the EcoRI fragment contained an ORF which could encode the first 278 amino acids. This partial gene and its deduced protein sequence were highly homologous (23) to the *E.coli* histidyl- tRNA synthetase gene (*hisS*) and its protein (HisRS). Southern blot analysis of SphI-digested genomic DNA showed a 1.8 kb fragment that hybridized to a 3′ Nco I/EcoRI fragment from the truncated gene (used as a probe) but not to the 5′ Xba I/Nco I fragment (not shown). This 1.8 kb fragment was cloned and sequenced. It was found to code for the remaining 148 amino acids of the initial ORF.

Nucleotide and deduced amino acid sequences of the *hisS* genes

Figure 1 shows the nucleotide sequence and deduced amino acid sequence of the *hisS* gene isolated from *S. equisimilis* strain D181. The gene has an open reading frame of 1278 nucleotides, which can encode a protein of 426 amino acids with a molecular weight of 47,932. A possible in-frame alternate initiation codon (TTG) is upstream at codon −5. An in-frame termination codon (TGA) is upstream at position −7. Analysis of the 5′ flanking sequence revealed a putative RNA polymerase promoter '−10' region (24; overlined in Fig. 1) located at nucleotide −157 to −146, which is a 12-nucleotide palindromic sequence. A

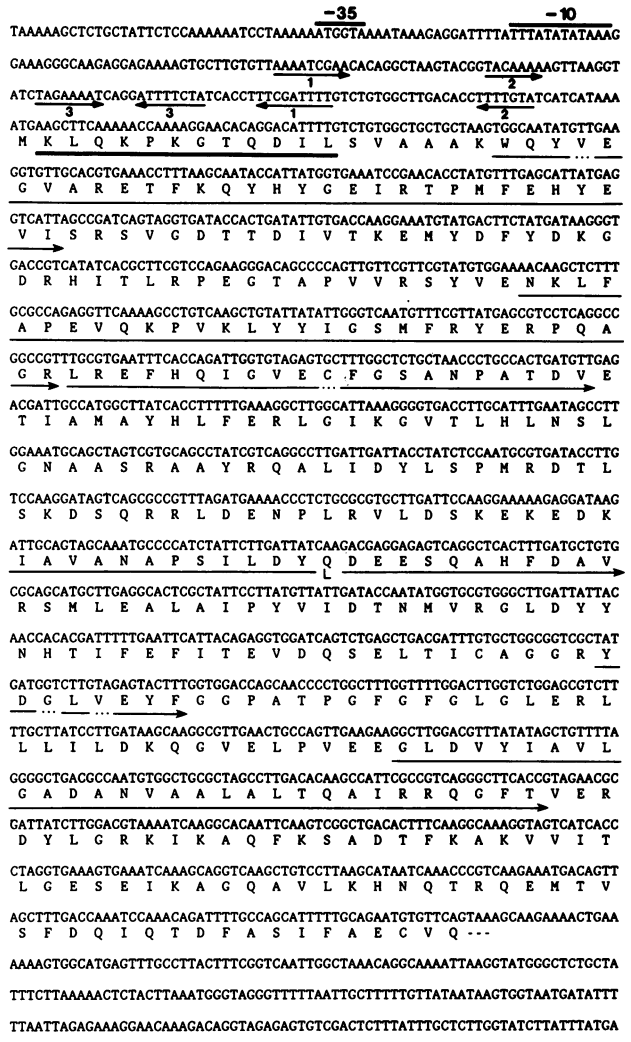

Figure 1. Nucleotide sequence, flanking sequences, and deduced amino acid sequence of the *hisS* gene from *S. equisimilis*. Putative '−10' and '−35' consensus promoter sequences (see Text) are overlined. Paired, numbered arrows represent inverted repeats that potentially can form stem-loop structures. The deduced amino acid sequence that was confirmed by Edman sequencing of the intact fusion protein is indicated with a thick underline. Amino acid sequences confirmed by sequencing HPLC-purified peptides are underlined with an arrow. This sequence has been deposited with GenBank under accession No. Z17214.

corresponding '−35' promoter region was also found at nucleotide −180 to −175. In addition, the 5′ region between nucleotides −146 and +1 exhibits potential secondary structure including several stem-loops (Fig. 2A), which may be important for transcriptional or translational control of gene expression. A Shine-Dalgarno (SD) consensus sequence for a strong ribosome binding site (25) was not found associated with the translational start site. However, possible SD sequences are at nucleotide −136 to −128. This region is upstream of the potential stem-loop structures and could, therefore, be closer to the initiating codon (Fig. 2A). Several possible stem-loop structures were also present in the 3′ flanking sequence of the gene (Fig. 2B).

Another long open reading frame of 965 nucleotides was found 242 nucleotides downstream of the *hisS* coding region. This second ORF was truncated and the amino terminal sequence of 321 amino acids deduced from the partial nucleotide sequence was found to be homologous to the *E.coli* (26) and yeast (27) aspartyl-tRNA synthetases (not shown).

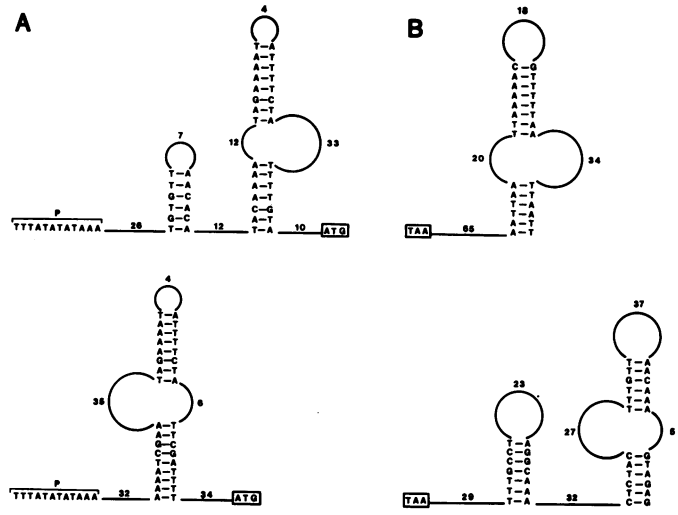

Figure 2. Potential stem-loop structures found in the 5′ and 3′ flanking regions of the *hisS* gene. The 5′ (A) and 3′ (B) flanking nucleotide sequences were analyzed for inverted repeats using the PC/GENE program REPEATS. Two structures with the largest number of base pairs in the stems are shown. The numbers indicate the number of nucleotides in a loop or interval between stems.

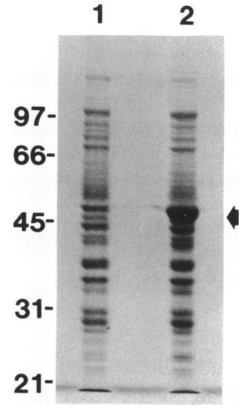

Figure 3. Expression of the *S. equisimilis hisS* gene in *E.coli*. Cell extracts were prepared and analysed by SDS-PAGE using cells containing the T7 RNA polymerase/promoter system (19) as described in Experimental Procedures. Lane 1, Uninduced bacteria. Lane 2, Bacteria induced at 42°C. The arrow at the right indicates the overexpressed HisRS fusion protein. Positions of the molecular mass markers (in kilodaltons) are shown on the left.

Expression of the *S. equisimilis hisS* gene in *E.coli*

To confirm that the product of the *hisS* gene is a protein and to verify that it has aminoacylating or tRNA charging activity, we fused the coding region to the T7 promoter of the T7 RNA polymerase/promoter vector pT7-7 developed by Tabor and Richardson[2] (19). The expected fusion protein contains an additional 15 amino acids at the amino end between Met_1 and Lys_2. The fusion construct was then electroporated into HB101 carrying another plasmid, pGP1-2, which has the T7 RNA polymerase gene under regulation of a temperature-sensitive repressor (*cI857*). By temperature induction at 42°C, the repressor is inactivated allowing the T7 RNA polymerase to be expressed and to transcribe the gene linked to the T7 promoter. By inhibiting the *E.coli* RNA polymerase with rifampicin, gene expression is limited almost exclusively to genes under the control

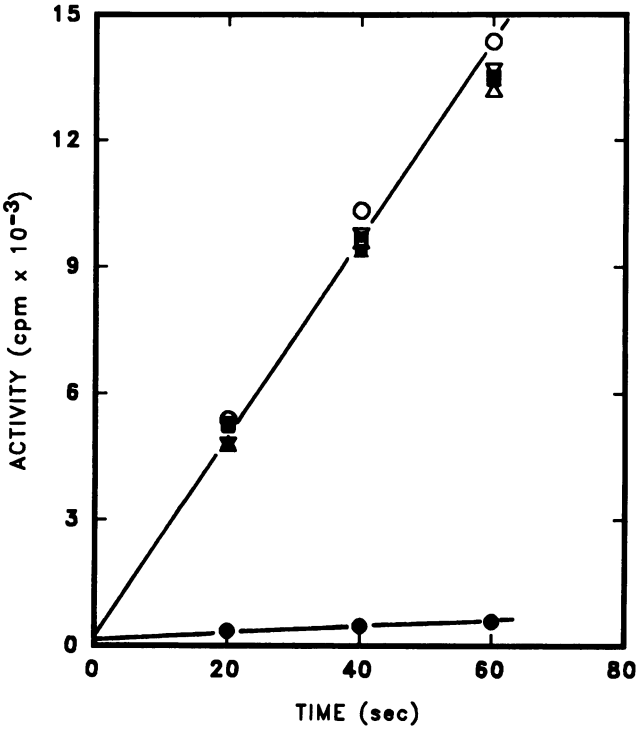

Figure 4. Aminoacylation activity of the HisRS fusion protein. tRNA charging activity in the S100 fraction from induced cells was measured using ^{3}H-His in the absence and presence of a 100-fold molar excess of nonradiolabeled amino acids as described in Experimental Procedures. The aminoacylation activity was determined with ^{3}H-His only (○) or with unlabeled histidine (●) or with other unlabeled amino acids. Other symbols each represent a mixture containing several of the other amino acids: Ala, Val, Leu, Ile, Pro, Met (□); Phe, Trp (■); Ser, Thr, Cys (△); Gln, Asn (▲); Glu, Asp (▽); Lys, Arg (▼). Gly and Tyr, not shown in this figure, also did not compete. Uninduced cells showed < 10% of the activity of the induced cells.

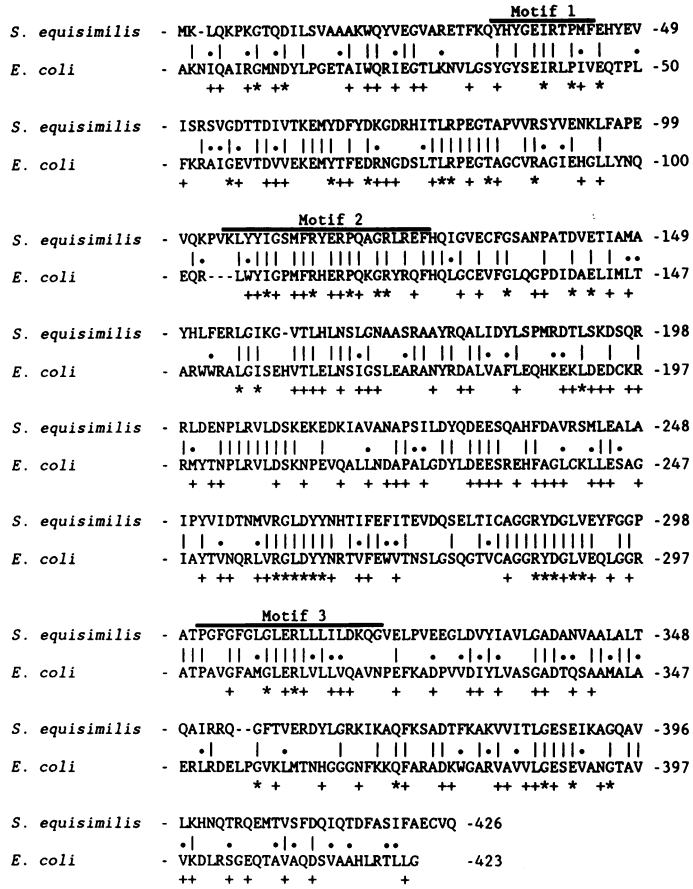

Figure 5. Comparison of the amino acid sequences of histidyl-tRNA synthetases from *E.coli* and *S. equisimilis*. The two protein sequences were aligned using the PC/GENE program PALIGN. Identical amino acids are shown by a vertical line and similar amino acids are shown by a dot. Amino acids that are identical or similar in all known histidyl-tRNA synthetases from bacteria to man are shown by an asterisk or a plus sign, respectively. The sequence motifs characteristic of Class II synthetases and identified as motifs 1, 2, and 3 are overlined. Dashes in the sequence represent gaps inserted to maximize the alignment.

of the T7 promoter. Under these conditions, the expected fusion protein was expressed at a level > 18% of the total cellular protein (Fig. 3). The predicted MW of the fusion protein is 49, 864. The apparent M_r of the fusion protein determined by SDS-PAGE is 53,700. A portion of the deduced amino acid sequence at the amino end (Lys_2 through Leu_{13}) was confirmed by Edman sequencing of the intact protein (Fig. 1). Additionally, six other tryptic or V-8 peptides were purified by HPLC and their amino acid sequences confirmed the deduced sequence at positions throughout ORF1. One discrepancy was found at position 229. Amino acid sequence analysis revealed Leu at this position, whereas the nucleotide sequence (CAA) indicated Gln.

Demonstration of histidyl-tRNA synthetase activity

To confirm the function of the *hisS* gene product, S100 extracts from the overexpressing cells were prepared and assayed for their ability to charge $tRNA^{His}$. Nonradiolabeled His competed > 90% for the incorporation of ^{3}H-His in the amino-acylation reaction (Fig. 4). None of the other nineteen amino acids tested, present in 100-fold molar excess over radiolabeled histidine, was able to compete in the tRNA charging assay. Extracts prepared from *E.coli* HB101 (pGP1−2) cells alone showed no charging activity, indicating that the endogenous *E.coli* enzyme was not detected under the same conditions. Likewise S100 extract from the same cells transformed with a fusion construct with the coding sequence in the opposite orientation showed no $tRNA^{his}$ charging activity. These results prove that this gene product is a histidyl-tRNA synthetase and that the enzyme is specific for histidine.

DISCUSSION

To date, the histidyl-tRNA synthetases from man (28,29), hamster (28), yeast (30), *E.coli* (23), and *S. equisimilis* (this report) have been sequenced. These enzymes exhibit a relatively narrow range of polypeptide sizes, from 424 amino acids for the *E.coli* enzyme to 546 amino acids for the yeast mitochondrial enzyme. Homologies among the three eucaryotic HisRSs are high, with the greatest relatedness being 92% identity/similarity between the human and hamster HisRS enzymes. The yeast enzyme is 51% and 58% identical/similar to the human and hamster enzymes, respectively.

Comparison of the amino acid sequences of HisRS from *E.coli* (23) and *S. equisimilis* (Fig. 5) show that the two bacterial enzymes are 58.4% homologous, with 42.6% identical and

```
   Motif 1:      GφXXφXXPφφ
S. e.  28   RETFKQYHYGEIRTPMFEH   46
E. c.  29   KNVLGSYGYSEIRLPIVEQ   47
S. c.  80   SGLFKKHGGVTIDTPVFEL   98
M. a.  81   ICCFKRHGAEVIDTPVFEL   99
H. s.  81   IRCFKRHGAEVIDTPVFEL   99

   Motif 2:   +φφXφXXXFRXE      +φX-F
S. e.  102   KPVKLYYIGSMFRYERPQ--AGRLREFHQIGVECFG   135
E. c.  103   ---RLWYIGPMFRHERPQ--KGRYRQFHQLGCEVFG   133
S. c.  144   QSIKRYHIAKVYRRDQPAMTKGRMREFYQCDFDVAG   179
M. a.  145   TNIKRYHIAKVYRRDNPAMTRGRYLNSITVDFDIAG   180
H. s.  145   TNIKRYHIAKVYRRDNPAMTGGRYPNSITVDFDIAG   180

                Region 1                          Region 2
S. e.  257   MVRGLDYYNHTI   268      286   GRYDGLVEYFG   296
E. c.  256   LVRGLDYYNRTV   267      285   GRYDGLVEQLG   295
S. c.  324   LARGLDYYTGLI   335      375   GRYDNLVNMFS   385
M. a.  324   LARGLDYYTGVI   335      360   RRYDGLVGMF-   369
H. s.  324   LARGLDYYTGVI   335      360   RRYDGLVGMF-   369
             *********                  *******

   Motif 3:   φGφGφGφERφφφφ
S. e.  301   PGFGFGLGLERLLLILDKQGVELPVE   326
E. c.  300   PAVGFAMGLERLVLLVQAVNPEFKAD   325
S. c.  396   PCVGISFGVERIFSLIKQRINS-STT   434
M. a.  377   PCVGLSIGVERIFSIVEQRLEALEEK   402
H. s.  377   PCVGLSIGVERIFSIVEQRLEALEEK   402
```

Figure 6. Comparison of the sequence motifs shared by histidyl-tRNA Class II synthetases from five species. The three sequence motifs shared by the Class II synthetases are given above the protein sequences. Conserved residues in the sequences are indicated by single-letter amino acid code for the invariant residue; ϕ for hydrophobic residues (F, Y, W, I, L, V, M, and A); − for negatively charged residues (D, E, N, and Q); + for positively charged residues (H, R, and K); and X for any residue. For regions 1 and 2, identities are indicated by asterisks, and conservative substitutions are indicated by dots. The five sequences from the five different species compared are S. e., *Streptococcus equisimilis*; E. c., *Escherichia coli* (23); S. c., *Saccharomyces cerevisiae* (30); M. a., *Mesocricetus auratus* (28); and H. s., *Homo sapiens* (28,29).

15.8% conservatively substituted amino acids. The *E.coli* HisRS is 11%, 38%, and 44% identical/similar to the human, hamster, and yeast HisRS enzymes, respectively. The *S. equisimilis* HisRS, however, is only 16%, 18%, and 20% identical/ similar to the human, hamster, and yeast HisRS enzymes, respectively.

Despite the low level of homology between the *E.coli* and human enzymes and between the *S. equisimilis* and the three eucaryotic enzymes, all five of these histidyl- tRNA synthetases share the three sequence motifs characteristic of Class II tRNA synthetases (Fig. 5). As initially reported (28) the human HisRS did not appear to have motif 3, the GLER tetrapeptide. However, a subsequent correction to this region of the human HisRS (29) indicates that this enzyme does, in fact, contain Motif 3. Since the *S. equisimilis* and *E.coli* HisRSs differ in length by only two amino acids and since the human and hamster enzymes (28) have the same number of amino acids, one would expect to find the three sequence motifs in the same locations. Motifs 1,2 and 3 in the two prokaryotic enzymes are at the same position within two amino acids. Motifs 1 and 2 in the three eucaryotic enzymes are exactly in the same locations (Fig. 6). Although the yeast enzyme (30) is 38 amino acids longer than the human and hamster enzymes, motifs 1 and 2 are at the same site within one amino acid. Motif 3 in the yeast enzyme is 19 amino acids away relative to the position in the hamster enzyme. There are also differences in the composition of the conserved consensus tetrapeptides, FRNE (Motif 2) and GLER (Motif 3). The eucaryotic enzymes have YRRD instead of FRNE, and replace V for L in GLER.

Aside from sharing the three sequence motifs, there are two other regions found between motifs 2 and 3 (designated Regions 1 and 2 in Fig. 5) that are highly conserved among the five proteins. The functions of these two peptide regions are not known at this time, but it is possible that they provide additional interactions to stabilize substrate binding. Region 1 is a six amino acid conserved sequence (RGLDYY) and region 2 is also a six-amino acid sequence (RYDGLV). Based on a hydropathy plot (31), region 1 (amino acid positions 259−264) is hydrophilic in nature and is likely, therefore, to be located on the exterior of the *Streptococcus* HisRS. In contrast, region 2 (positions 287−292) is predicted to be hydrophobic in nature and is presumably located in the interior of the protein.

Compared to the Gram negative *E.coli*, very few genes and their associated promoters have been characterized in the Gram positive *Streptococcus* species. Consequently, much less is known about the variety and sequences of specific gene regulatory regions in *Streptococcus*. Many *Streptococcus* genes have been identified that contain the typical consensus Shine-Dalgarno sequence for ribosome binding and the −10 and −35 consensus sequences for RNA polymerase binding. The *hisS* gene, however, does not have the typical ribosome binding site near the initiating codon. The putative −10 and −35 sequences and Shine-Dalgarno sequence are relatively far upstream. Other *Streptococcus* genes, such as the *dpnM* gene of *Streptococcus pneumoniae*, also do not have the ribosome binding sequence (32). Since these genes, including the *hisS* gene reported here[2], are transcribed, it is likely that other *Streptococcus*-specific regulatory elements are used in the expression of these genes. The possible stem-loop structures spanning the 5′ region from nucleotide −146 to −10 are close to the putative −10 and −35 promoter sites. If present at the DNA level these structures could be involved in the regulation of *hisS* gene transcription.

The 12-nucleotide palindromic sequence at −157 to −146 (Fig. 1) may be a binding site for a regulatory factor that controls expression of the *hisS* gene. This phenomenon is not uncommon for aminoacyl-tRNA synthetase genes. For example, in the *E.coli alaS* gene, Putney and Schimmel (33) found that gene transcription was repressed by its gene product, AlaRS, which binds to a palindromic sequence. The palindrome flanks the transcription initiation site and the TATA sequence. This type of transcriptional autoregulation might also occur with the *hisS* gene. The presence of potential stem-loop structures located upstream of the translation start site (Fig. 2A) may also be involved in regulation at the translational level. The *E.coli* ThrRS autoregulates its own expression by binding to its mRNA at 5′ stem-loop structures and decreasing translation (34−36). The ribosome is prevented from binding to its *rbs* and translation is repressed. This repression is relieved by $tRNA^{Thr}$ which displaces the bound synthetase from its mRNA and allows the synthesis of ThrRS to proceed.

We also note that the cloned DNA contains two Class II synthetase genes found in tandem, only 242 base pairs apart. Putative promoter elements and the Shine-Dalgarno sequence are found upstream of the second ORF. Again, potential stem-loop structures are found at the 5′ untranslated region of this gene (Fig. 2B). It will be interesting to see if other tRNA synthetase genes are located immediately upstream of the *hisS* gene or downstream of the putative *aspS* gene. The *aspS* gene has not been mapped in *E.coli*; thus, it is not known if the same order of these genes seen in *S. equisimilis* is found in *E.coli*.

ACKNOWLEDGMENTS

We gratefully acknowledge Elizabeth Gerhardt, Shirley Chapman and Drs Lillian Chan, Richard Fritz and William Thompson for assistance with the DNA sequence analysis, Janet Oka for help preparing the figures, and Lisa Raney for help preparing the manuscript. This research was supported by grant #004952-11 from the Texas Higher Education Coordinating Board (to PW and JP). Cora Menguito was supported by a predoctoral fellowship from the James W.McLaughlin Fellowship Fund.

ABBREVIATIONS

tRNA, transfer RNA; *hisS*, histidyl-tRNA synthetase gene; ORF, open reading frame; Tris, trishydroxyaminomethane. Throughout this paper, the structural genes encoding the amino acyl-tRNA synthetases are referred to by the standard three-letter code for the amino acids (lower case & italic) followed by S (in italic). The synthetases are denoted by the standard three-letter code followed by RS, *e.g.* HisRS for histidyl-tRNA synthetase.

REFERENCES

1. Burbaum, J.J., and Schimmel, P. (1991) *J. Biol. Chem.* **266,** 16965-16968
2. Moras, D. (1992) *Trends Biochem. Sci.* **17,** 159-164
3. Webster, T.A., Gibson, B.W., Keng, T., Biemann, K., and Schimmel, P. (1983) *J. Biol. Chem.* **258,** 10637-10641
4. Fett, R., and Knippers, R. (1991) *J. Biol. Chem.* **266,** 1448-1455
5. Hou, Y.-M., Shiba, K., Mottes, C., and Schimmel, P. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88,** 976-980
6. Putney, S.D., Royal, N.J., de Vegvar, H.N., Herlihy, W.C., Biemann, K., and Schimmel, P. (1981) *Science* **213,** 1497-1501
7. Putney, S.D., Sauer, R.T., and Schimmel, P.R. (1981) *J. Biol. Chem.* **256,** 198-204
8. Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) *Nature (London)* **347,** 203-206
9. Brick, P., Bhat, T.N., and Blow, D.M. (1988) *J. Mol. Biol.* **208,** 83-98
10. Rould, M.A., Perona, J.J., Söll, D., and Steitz, T.A. (1989) *Science* **246,** 1135-1142
11. Brunie, S., Zelwer, C., and Risler, J.-L. (1990) *J. Mol. Biol.* **216,** 411-424
12. Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N., and Leberman, R. (1990) *Nature (London)* **347,** 249-255
13. Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J.C., and Moras, D. (1991) *Science* **252,** 1682-1689
14. Dao, M.L., and Ferretti, J.J. (1985) *App. Envir. Microbiol.* **49,** 115-119
15. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning A Laboratory Manual.* 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
16. Marmur, J. (1961) *J. Mol. Biol.* **3,** 208-218
17. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74,** 5463-5467
18. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) *Gene* **33,** 103-119
19. Tabor, S., and Richardson, C.C. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82,** 1074-1078 20. Laemmli, U.K. (1970) *Nature* (London) **227,** 680-685
21. Lin, S.X., Shi, J.P., Cheng, X.D., and Wang, Y.L. (1988) *Biochemistry* **27,** 6343-6348
22. Mans, R.J., and Novelli, G.D. (1961) *Arch. Biochem. Biophys.* **94,** 48-53
23. Freedman, R., Gibson, B., Donovan, D., Biemann, K., Eisenbeis, S., Parker, J., and Schimmel, P. (1985) *J. Biol. Chem.* **260,** 10063-10068
24. Rosenberg, M., and Court, D. (1979) *Ann. Rev. Genet.* **13,** 319-353
25. Shine, J., and Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71,** 1342-1346
26. Eriani, G., Dirheimer, G., and Gangloff, J. (1990) *Nucleic Acids Res.* **18,** 7109-7118 27. Sellami, M., Fasiolo, F., Dirheimer, G., Ebel, J.-P., and Gangloff, J. (1986) *Nucleic Acids Res.* **14,** 1657-1666
28. Tsui, F.W.L., and Siminovitch, L. (1987) *Nucleic Acids Res.* **15,** 3349-3367
29. Cusack, S., Härtlein M. and Leberman, R. (1991) *Nucleic Acids Res.* **19,** 3489- 3498.
30. Natsoulis, G., Hilger, F., and Fink, G.R. (1986) *Cell* **46,** 235-243
31. Kyte, J., and Doolittle, R.F. (1982) *J. Mol. Biol.* **157,** 105-132.
32. Mannarelli, B.M., Balganesh, T.S., Greenberg, B., Springhorn, S.S., and Lacks, S.A. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82,** 4468-4472
33. Putney, S.D., and Schimmel, P. (1981) *Nature* (London) **291,** 632-635
34. Springer, M., Graffe, M., Dondon, J., Grunberg-Manago, M., Romby, P., Ehresmann, B., Ehresmann, C., and Ebel, J.-P. (1988) *Biosci. Rep.* **8,** 619-632
35. Moine, H., Ehresmann, B., Romby, P., Ebel, J.-P., Grunberg-Manago, M., Springer, M., and Ehresmann, C. (1990) *Biochim. Biophys. Acta* **1050,** 343- 350
36. Moine, H., Romby, P., Springer, M., Grunberg-Manago, M., Ebel, J.-P., Ehresmann, B., and Ehresmann, C. (1990) *J. Mol. Biol.* **216,** 299-310
37. Menguito, C.A., Papaconstantinou, J. and Weigel, P.H. (manuscript submitted)