Nucleic Acids Research, Vol. 20, No. 16 4173–4178

Structure of the phenylalanyl-tRNA synthetase genes from *Thermus thermophilus* HB8 and their expression in *Escherichia coli*

Roland Kreutzer, Volker Kruft[1], Ekaterina V.Bobkova[2], Olga I.Lavrik[2] and Mathias Sprinzl*
Laboratorium für Biochemie, Universität Bayreuth, Universitätstraße 30, 8580 Bayreuth,
[1]Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, Ihnestraße 73, 1000 Berlin 33, Germany and [2]Institute of Bioorganic Chemistry, Siberian Division of the Russian Academy of Sciences, Lavrentiev Prospect 8, 630090 Novosibirsk, Russia

Received June 5, 1992; Revised and Accepted July 24, 1992 EMBL accession no. X65609

ABSTRACT

A 4459 bp long *Bam*HI restriction fragment containing the two genes for the *Thermus thermophilus* HB8 phenylalanyl-tRNA synthetase was cloned in *Escherichia coli* and its nucleotide sequence was determined. The genes *pheS* and *pheT* encode the α- and β-subunits with a molecular weight of 39 and 87 kD, respectively. Three conserved sequence motifs typical for class II tRNA synthetases occur in the α-subunit. Secondary structure predictions indicate that an arm composed of two anti-parallel α-helices similar to that reported for the *E.coli* seryl-tRNA synthetase may be present in its N-terminal portion. In the β-subunit clusters of hydrophilic amino acids and a leucine zipper motif were identified, and several pronounced α-helical regions were predicted. The particular arginine and lysine residues in the N-terminal portion of the β-subunit, which were found to participate in tRNA binding in the yeast and *E.coli* PheRSs, have their counterparts in the *T.thermophilus* protein. The 5'-portion of an open reading frame downstream of *pheT* was found and codes for a yet unidentified, extremely hydrophobic peptide. The *pheST* genes are presumably cotranscribed and translationally coupled. A novel type of a putative transcriptional terminator in *Thermus* species was identified immediately downstream of *pheT* and other *Thermus* genes. The genes *pheS* and *pheST* were expressed in *E.coli*.

INTRODUCTION

Although all aminoacyl-tRNA synthetases have a common function in protein biosynthesis, namely to catalyse the aminoacylation of tRNA with one specific cognate amino acid, they are widely diverse in structure. Therefore, the investigation of each individual tRNA synthetase gives novel insights into, how structurally different protein molecules accomplish similar reactions, as well as into the evolution of the genetic code.

However, three dimensional structures of only five tRNA synthetases are available, i.e. the tyrosyl-tRNA synthetase of *Bacillus stearothermophilus* (1), glutaminyl- (2), methionyl- (3), seryl-tRNA synthetase of *Escherichia coli* (4), and the aspartyl-tRNA synthetase of yeast (5). Based mainly on these structural data and on computer evaluations of known amino acid sequences of tRNA synthetases these enzymes have recently been classified into two groups (6,7). Class I synthetases show two common consensus sequences ('HIGH' and 'MSKS'), whereas class II synthetases share three different sequence motifs which may be manifested more in their tertiary structure than in their primary structure. The PheRS was assigned to class II because the small subunit of this enzyme displays the characteristic elements of this group (6). However, among the class II synthetases the PheRS appears to be peculiar with respect to its oligomeric structure, its primary attachment site of the activated amino acid, and the fact that its substrate is not one of the primordial types. In the first case, the PheRS has an oligomeric state of $\alpha_2\beta_2$ (8) with the exception of the mitochondrial *Saccharomyces cerevisiae* PheRS, which is active as a monomer composed of sequences corresponding to the prokaryotic α-subunit and to the C-terminal portion of the β-subunit (9). The small subunit of the *Thermus thermophilus* PheRS was designated as α-, the large as β-subunit (10). Apart from PheRS this heterotetrameric structure is found only in the glycyl-tRNA synthetase which possesses the class II-specific sequence motif 3 but lacks the other two motifs (6). Secondly, the PheRS is an exception from the rule that class II synthetases aminoacylate tRNA on the 3'-OH of the ribose (6). Instead, it aminoacylates at the 2'-OH (11,12), a property shared with most of the class I synthetases. Thirdly, the presumably more ancestral class II synthetases are believed to use primordial amino acids as substrates, which are spontaneously formed under prebiotic conditions (13). Phenylalanine, however, probably does not come into existence under these conditions. All these unusual properties point towards an exceptional position of PheRS in evolution.

The α- and β-subunits of the PheRS are encoded by the genes *pheS* and *pheT* which are adjacent to each other in both *E.coli*

* To whom correspondence should be addressed

and *Bacillus subtilis* (14,15). In *E.coli* their expression is regulated by the classical form of attenuation involving two alternative stem-loop structures and a phenylalanine-rich leader peptide (16). In *B.subtilis*, on the other hand, the expression of the *pheST* genes is believed to be regulated by an anti-termination mechanism similar to that described for lamboid phages (15). Since attenuation may also occur in *T.thermophilus*, as discussed for the methionyl-tRNA synthetase gene (*metS*) (17), it was of special interest for us to examine the upstream and downstream regions of the *pheST* genes for possible regulatory elements.

Apart from the investigation of the gene expression, the elucidation of the structure-function relationships related to the peculiarities of the PheRS has been a predominant challenge for several working groups in recent years. Consequently, several biochemical data have been collected especially concerning the smaller α-subunit of the PheRS, which may be similar in structure to the seryl-tRNA synthetase (4). Studies on the *E.coli* and yeast PheRSs with altered amino acids or deletions in consensus motif 3 (18,19) support the assumption that this motif is, as a part of the active site, involved in phenylalanine binding. Little is known about structure and function of the β-subunit, which shows no homology to other tRNA synthetases. It also contributes to the active site and is involved in tRNA binding as shown by affinity labelling and crosslinking experiments (20–22).

In addition to these data, X-ray structural analyses would contribute to an important progress in our understanding of the catalytic mechanism of the PheRS. Due to their high stability, proteins from thermophilic bacteria have proven particularly useful for crystallography. Recently, the PheRS of *Thermus thermophilus*, an extremely thermophilic bacterial species, has been purified, crystallized (10,23,24) and previously studied (25). Both, the determination of the primary structure and the construction of strains overexpressing the PheRS have now become a prerequisite for further studies. In our laboratory we have recently cloned the glutamyl- and aspartyl-tRNA synthetases of *T.thermophilus* HB8. In this work we report the cloning and sequencing of the genes encoding the *T.thermophilus* HB8 PheRS and their expression in *E.coli*, which may be an essential contribution in order to advance the structural and functional investigations of this enzyme.

MATERIALS AND METHODS

Bacterial strains and plasmids

Thermus thermophilus HB8 (ATCC 27634) was grown as reported previously (26). *Escherichia coli* strains JM109 (27), DH5α (28) and SG13009 (29) were grown in Luria Bertani nutrient medium (30). Plasmid pREP4 (Diagen GmbH, Düsseldorf, Germany) encodes the Lac repressor and the neomycin phosphotransferase. Plasmids pUC18 (27) and pKK223-3 (Pharmacia LKB, Uppsala, Sweden) were used for cloning and sequencing, and for expression, respectively.

Peptide sequencing

For the amino acid sequence determination of PheRS fragments of the individual α- and β-subunits, which were obtained as outlined in (24), 2 nmoles of protein were digested over night with endoproteinase Glu-C (Boehringer Mannheim, Mannheim, Germany) at an enzyme: substrate ratio of 1: 20 in 100 mM KH_2PO_4 at pH 7.5. Peptides were separated by reversed phase HPLC on a laboratory-packed C_{18} column (250×4mm, Vydac C_{18}, 5 μm, 300 Å) using a linear gradient of acetonitrile in water, both containing 0.1% trifluoroacetic acid. Protein and peptide sequencing were done on a pulsed-liquid phase sequencer equipped with a PTH-analyzer (Applied Biosystems, Foster City, USA, models 477A and 120A, respectively) (31).

Recombinant DNA-techniques

Genomic DNA of *T.thermophilus* HB8 was isolated in accordance to the protocol of (32). Restriction enzymes and other DNA modifying enzymes were used as recommended by the manufacturers (Boehringer Mannheim GmBH, Mannheim, Germany; Angewandte Gentechnologie GmBH, Heidelberg, Germany).

Oligodeoxyribonucleotides to be used as primers for PCR were either designed based on the amino acid sequence of one of the PheRS β-subunit peptide fragments taking into account the preferential codon usage of known *T.thermophilus* genes or derived from the determined nucleotide sequence. The oligodeoxyribonucleotides were synthesized using an automatic synthesizer (Applied Biosystems, Foster City, USA, model 381A) by phosphoramidite chemistry on solid phase.

Colony and Southern hybridizations (33,34) were performed according to the protocols of (35) using the digoxigenin DNA labelling and detection kit (Boehringer Mannheim GmBH, Mannheim, Germany) in six times concentrated standard saline citrate solution (0.15 M NaCl, 0.015 M trisodium citrate, pH 7.0) at 70°C.

PCR was performed using the AmpliTaq DNA amplification kit (Perkin Elmer, Norwalk, USA) with 100 to 500 pmoles of oligodeoxyribonucleotide primers and 500 ng of genomic *T.thermophilus* HB8 DNA partially digested with *Bam*HI or 50 ng plasmid DNA of pPheRS2 linearized with *Eco*RI, respectively, as a template. Products from PCR were analyzed on 15% denaturing polyacrylamide or on 0.8% agarose gels. For preparative purposes gel pieces containing the desired bands were cut out, and the DNA was electroeluted. Blunt ended amplified DNA fragments were phosphorylated using polynucleotide kinase in imidazole buffer according to the protocol of (35) prior to their ligation with pUC18. For the amplification and ligation of the *pheS* gene with the expression vector pKK223-3 a strategy analogous to the previously described procedure (36) was applied.

Nested deletions of DNA fragments cloned in pUC18 were obtained by the exonuclease III/ S1 nuclease technique (37). Sequencing of pUC18 derivatives was performed by the chain termination method (38) using the T7 sequenase kit (Pharmacia LKB, Uppsala, Sweden) and α-[^{35}S]ATP (1500 Ci/mmol; Du Pont, Bad Homburg, Germany) with either universal sequencing and reverse primers or synthetic oligodeoxyribonucleotide primers corresponding to *T.thermophilus* DNA sequences.

Expression and determination of the enzyme activity

For the expression of the PheRS genes, *E.coli* SG13009 harbouring the appropriate plasmids was grown in 20 ml Luria Bertani broth with 200 μg/ml ampicillin, and, in the case of the pREP4 containing strains, with 25 μg/ml kanamycin. The *tac*-promoter of pKK223-3 was induced by 1mM IPTG at OD_{600}=0.7. Cells were harvested three hours after induction. Crude cellular lysates were prepared as described in (39) and centrifuged at 12000 g for 15 minutes. The supernatant was analyzed by SDS PAGE (40).

For the determination of the enzyme activity, the supernatant was subjected to thermal denaturation for 15 minutes at 60°C in order to remove thermolabile *E.coli* proteins. After centifugation the activity of the *T.thermophilus* PheRS was assayed at 60°C in 50 mM N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid),

pH 7.6, 30 mM KCl, 6.25 mM $MgCl_2$ 2.5 mM β-mercaptoethanol, 10 mM ATP and 20 μM [^{14}C]-phenylalanine with 50 A_{260}/ml (2 mg/ml) $tRNA^{bulk}$ from *T.thermophilus*.

RESULTS AND DISCUSSION

Cloning of the *pheST* operon and determination of the nucleotide sequence

The N-terminal amino acid sequences of the purified PheRS subunits α and β (24) and of five peptide fragments obtained by enzymatic cleavage of the β-subunit with endoprotease Glu-C were determined (Fig. 3, boldface letters). Different oligodeoxyribonucleotides (17 to 33 nucleotides long) were designed and used as labeled probes in Southern hybridization experiments with cleaved genomic DNA of *T.thermophilus* but failed to deliver clear and specific signals even under varying hybridization conditions. In order to circumvent this problem we used two oligodeoxyribonucleotides (2721 and 2817) corresponding to the N- and C-terminal amino acids, respectively, of one of the sequenced peptide fragments (Fig. 1) in PCR with genomic DNA of *T.thermophilus* HB8 as template. An expected DNA fragment of 80 bp, which was detected among several other amplification products, was eluted from the gel and used as a template in a second PCR assay. This delivered sufficient amounts of this particular DNA fragment, which was cloned in pUC18 and, by sequencing, confirmed to correspond to the amino acid sequence determined for the respective peptide. This DNA fragment proved to be a suitable probe resulting in highly specific signals in hybridization. When genomic DNA hydrolyzed with *Bam*HI was used in Southern hybridization only one single signal was obtained corresponding to a restriction fragment of 4.5 kbp in size (data not shown). *Bam*HI fragments of approximately this size were cloned in *E.coli* and, by colony hybridization, one clone (pPheRS1) was identified to carry a 4.5 kbp *Bam*HI fragment in pUC18 hybridizing to the probe. As later confirmed by DNA sequencing, this *Bam*HI fragment contained the entire genetic information of the PheRS (see below).

For the determination of the nucleotide sequence deletions of the plasmids pPheRS1 and pPheRS2 were generated, the latter carrying the same 4.5 kbp *Bam*HI fragment as pPheRS1 but in reversed orientation. Using 31 selected deletion clones most of the sequence of this 4459 bp long *Bam*HI fragment was determined on both strands. Remaining gaps were filled by priming the sequencing reactions with synthetic oligodeoxyribonucleotides specific for internal sequences. The complete nucleotide sequence is deposited in the European Molecular Biology Laboratory database (accession number X65609).

All amino acid sequences determined for the above mentioned peptides of PheRS were found to be parts of the amino acid sequences deduced from the two open reading frames of 1053 bp and 2358 bp in length. This confirmed the identification of the *pheST* genes. Their codon usage with an extreme bias for G and C in the third position resembles that of the most other *Thermus* genes reflecting their adaption to high growth temperatures. The amino acid composition as deduced from the sequences (data not shown) is in good agreement with that determined by hydrolysis of the protein (25). The 5'-portion of an open reading frame starting at position 4135 downstream of *pheT* was identified (Fig. 2). This is (as evidenced by the remarkable avoidance of rare codons and the high GC-content in the third position which is comparable to the *pheST* genes) probably part of a coding region for an as yet unidentified protein. However, no pronounced promoter- or SD-sequences were found in the region between *pheT* and the assumed coding region. Interestingly, the 108 N-terminal amino acids deduced from the nucleotide sequence constitute an extremely hydrophobic, leucine-rich peptide, which could be folded into several β-sheets (data not shown). Sequencing of the rest of this putative gene may reveal more details.

Structure of the PheRS α- and β-subunits

The α- and β-subunits of the PheRS encoded by *pheS* and *pheT* comprise proteins of 350 and 785 amino acids with a calculated molecular weight of 39 and 87 kD, respectively. No amino acid sequence deviations from the recently published *T.thermophilus* HB8 PheRS sequence (41) were found, apart from a stretch of 14 amino acids in the middle of the β-subunit. This difference is probably due to sequencing mistakes resulting in a frameshift comprising that particular region. The alignment of the PheRS

```
673                                                                    714
LeuArgLeuProLeuProAspLysProLeuAlaPheGlnAspProSerArgHisProAlaAlaPheArgAspLeuAlaVal
CTCCGCCTGCCCCTCCCGGACAAACCCCTGGGCTTCCAAGACCCCTCCCGCCACCCCGCGGCCTTCCGCGACCTGGCCGTG
***** ** ******** ***                                  ***** ***** *****
5'-CTCCGGCTCCCCCTCCCCGAC-3'                            3'-AAGGCCCTGGAGCGGCA-5'
     Oligo 2721                                             Oligo 2817
```

Figure 1. Construction of a hybridization probe for the identification of the *pheT* gene. In the top line the amino acid sequence as determined for a peptide fragment of the PheRS α-subunit is shown; in the bottom line are the nuceotide sequences of two oligodeoxyribonucleotides derived from the amino acid sequence, which were used as primers in PCR in order to amplify the intervening DNA. In between the proper nucleotide sequence is displayed with the asterisks to indicate matches with oligodeoxyribonucleotide 2721 and the complement of oligodeoxyribonucleotide 2817.

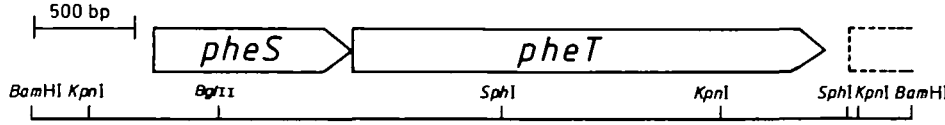

Figure 2. The restriction map of the cloned 4459 bp *Bam*HI fragment is shown in relation to the positions of the *pheST* genes (open arrows) and the 5'-portion of an open reading frame (dashed box) above.

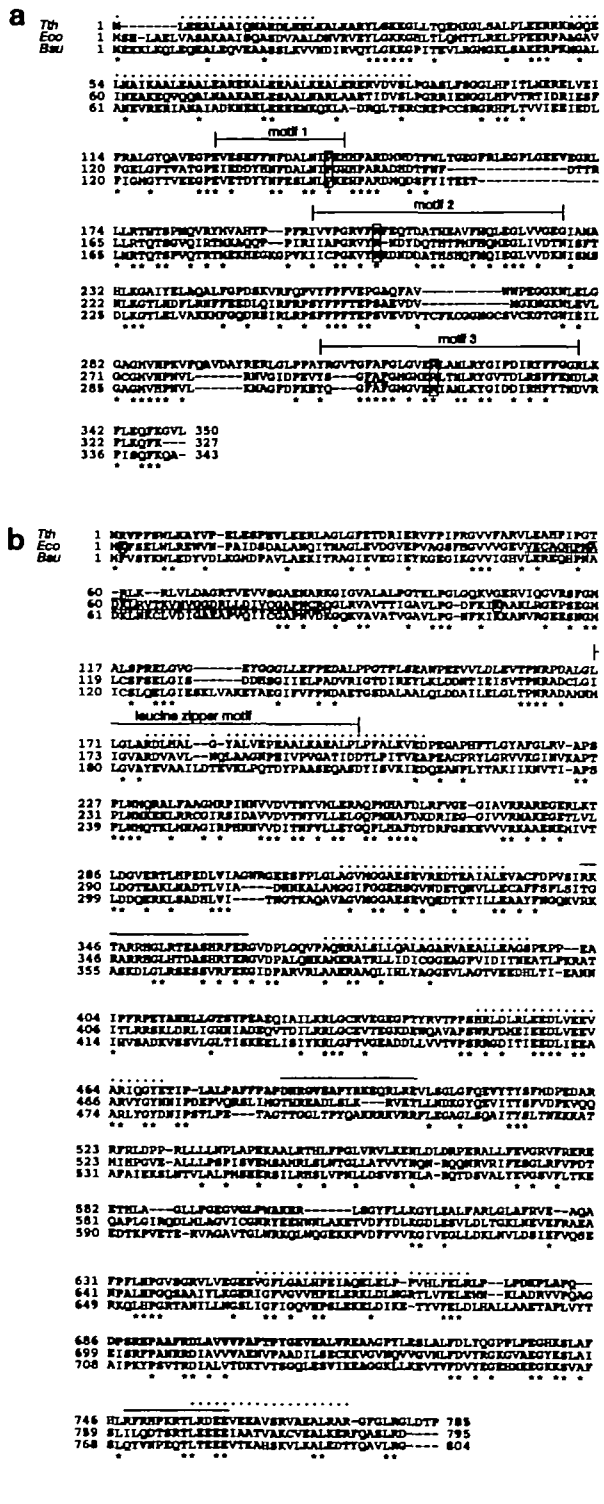

Figure 3. Alignment of the *T.thermophilus* HB8 PheRS α- (a) and β-subunit (b) amino acid sequences (*Tth*) with those of *E.coli* (*Eco*) and *B.subtilis* (*Bsu*) as taken from (14,15,52). Asterisks indicate residues identical in the three known prokaryotic PheRS sequences. Amino acid sequences determined for peptide fragments of the purified PheRS subunits are in boldface letters. (a) The conserved sequence motifs 1, 2 and 3 (6) are marked by a line above with the invariant residues therein in boxes. The 'FAF' motif in the *E.coli* sequence mutagenized by (18) is underlined. Two putative anti-parallel α-helices are labeled by dotted lines above. (b) Regions with an accumulation of hydrophilic residues are overlined, and predicted α-helical regions are marked by dotted lines above. A zink-finger motif in the *E.coli* protein proposed by (17) is underlined and the 'ADKL' motif common to the yeast and *E.coli* PheRSs (22) as well as two lysine residues labeled by oxidized $tRNA^{Phe}$ which are boxed. A leucine zipper motif is indicated.

Table 1. Comparison of *T.thermophilus* HB8 promoters with the *E.coli* −35 and −10 promoter consensus sequences.

Gene/operon	−35 region	bp space	−10 region	reference
E.coli consensus	TTGACA	17	TATAAT	
T.thermophilus HB8				
leuB	TTGACC	18	TACCCT	(54)
23S rRNA	TTGACA	18	TATCTT	(55)
4.5S rRNA	TAGCCT	18	TATACT	(56)
16S rRNA	TTGACA	17	TAGCAT	(57)
thrU(tuf2)	TTGACA	17	TAGCCT	(58)
trpE	TGGACA	17	TATCCT	(59)
pheS	TTCGCG	20	TATCCT	this work

amino acid sequences with those of *E.coli* and *B.subtilis* (Fig. 3) reveals an elevated degree of conservation in several particular regions of the proteins rather than a random distribution of conserved residues. This points towards the functional relevance of these regions. In the α-subunit three conserved sequence motifs occur which are typical for class II tRNA synthetases (6). Motif 1 contains an invariant proline residue (position 140), whereas motifs 2 and 3 contain invariant arginine residues (positions 204 and 321). Moreover, in motif 3 an alanine at position 314 is present corresponding to alanine 294 in the *E.coli* PheRS. This residue in the *E.coli* enzyme presumably interacts with the substrate phenylalanine and contributes to the determination of the substrate specificity (18) and is located within a stretch of five amino acids identical in the three known prokaryotic PheRSs, again emphasizing the relavance of this region. The secondary structure prediction according to Chou and Fasman (42) indicates α-helical regions near the N-terminus from amino acids 1 to 22 and from amino acids 51 to 90. This may form an arm composed of two long anti-parallel helices as seen in the seryl-tRNA synthetase of *E.coli* (4). It has been speculated that this unique structural element interacts with the long variable loop of the $tRNA^{Ser}$. However, if the existence of this arm in PheRS turnes out to be valid, its function as a support for a long variable loop in tRNA binding becomes improbable, since $tRNA^{Phe}$ of *T.thermophilus* lacks this feature (43).

The *T.thermophilus* PheRS β-subunit appears to be highly structured according to its predicted secondary structure. Pronounced α-helical regions occur around positions 175−205, 315−335, 370−395, 450−470, 645−675 and 755−775. Clusters of hydrophilic amino acids around positons 350−365, 480−500 and 740−760, all found in the vicinity of predicted turns-in the secondary structure, are presumably located on the surface of the protein. A zink-finger motif as proposed for the *E.coli* PheRS and several other tRNA synthetases (17) was not found in the N-terminal portion of the *T.thermophilus* PheRS α-subunit. Also, a sequence motif ('ADKL') common to yeast and *E.coli* PheRSs is not present in the *T.thermophilus* PheRS. Because it could be labeled with periodate-oxidized $tRNA^{Phe}$ (21,22) the lysine within this motif was shown to participate in tRNA binding. An arginine rather than lysine occurs at the corresponding position (position 60) in the *T.thermophilus* PheRS β-subunit. Two other lysine residues in the *E.coli* PheRS were labeled with oxidized tRNA, i. e. lysine 2 and lysine 106 (21), which correspond to arginine 2 and lysine 102 in the *T.thermophilus* protein. The arginine residues 2 and 60 in *T.thermophilus* may serve as substitutes in tRNA binding for the corresponding lysines in *E.coli*. Attempts

```
Tth pheT  4013 TAATGCCA CCCCATGCTGGC TTGC GCCAGCATGGGGG CCCC-GGCAAAAGGCC 4064
Taq mdh   1995 TTACACCA CCCCATGCTGGC TTGC GCCAGCATGGGGG CCCC-GGCAAAAGCTC 2046
Tfl mdh   1647 TTATACCA CCCCATGCTGGC GCAA GCCAGCATGGGGG CCCCCGGCAAAAGCTC 1699
Tth metS  2214 TAATACCT CCCCAcGCCGGg GCAA GCCGGCATGGGGG CCCC-GGCAAAGGGTT 2265
               * *  **  ***** ** **       *** ********** **** ****** *
```

Figure 4. Alignment of the nucleotide sequence downstream of *pheT* with sequences downstream of the *mdh* genes of *Thermus aquaticus* B (*Taq*) (46) and *Thermus flavus* (*Tfl*) (53) and the *metS* gene of *T.thermophilus* HB8 (*Tth*) (17). The stop triplets of the *pheT* and *metS* genes are underlined, those of the *mdh* genes are 46 bp further upstream. The numbering is taken from the original publications. An inverted repeat is indicated by oppositely oriented arrows above. Complementary nucleotides forming a putative stem are in boldface capital letters and non-complementary nucleotides within in lower case letters. The asterisks indicate nucleotides identical in all four sequences.

to modify the *T.thermophilus* PheRS by periodate-oxidized $tRNA^{Phe}$ from *T.thermophilus* were not successful. A leucine zipper motif occurs at positions 170 to 198. If existing, this element could either exhibit the site of dimerization of the β-subunits, or it points towards a regulatory role of the *T.thermophilus* PheRS in transcription.

Regulation of the *pheST* expression

Although the *pheST* operon in *E.coli* is regulated by attenuation, no attenuator-like sequences were found in the region upstream of *pheS* in *T.thermophilus*. In fact, there is an imperfect inverted repeat at position 215 to 263 which may form a stem-loop structure. However, no open reading frame for a leader peptide overlapping the inverted repeat was found. Thus, the criteria of a proper attenuator are not met. Instead, two motifs occur 19–50 bp upstream of *pheS* which display similarities to the −35 and −10 sequences of *T.thermophilus* promoters (Table 1) suggesting that the expression of the *pheST* operon in *T.thermophilus* is either constitutive or involves other regulatory mechanisms. The SD-sequence (5′-GAGG) 8 bp upstream of the *pheS* start triplet is complementary to four contiguous nucleotides near the 3′-end of the *T.thermophilus* 16S rRNA (44) and may be sufficient for ribosome binding of the mRNA. The start codon of *pheT* overlaps the stop codon of *pheS* (AUGA) suggesting that both genes are translationally coupled. A similar arrangement is observed with the *T.thermophilus* HB27 *trpB* and *trpA* genes (45) and occurs at many other prokaryotic operons whose gene products are required in equimolar amounts. This may also be the case with the products of the *T.thermophilus pheST* genes, as supported by the $\alpha_2\beta_2$ oligomeric structure of PheRS (8). A 44 bp long sequence located downstream of *pheT* was found to be identical to a sequence downstream of the *T.aquaticus* B malate dehydrogenase (*mdh*) gene and strikingly similar to sequences downstream of the *T.flavus mdh* and *T.thermophilus* HB8 methionyl-tRNA synthetase (*metS*) genes (Fig. 4). This conserved sequence comprises an inverted repeat which may form a stem-loop structure followed by a special motif. The sequence of the four bp long loop at the *T.flavus mdh* and the *T.thermophilus metS* genes is reversed and complementary to this loop at the two other genes, which may be the result of an inversion (46). It is possible that the entire conserved sequence exhibits a repetitive element as found in the genome of other bacterial species (47). However, due to the putative formation of a stem-loop structure and the position downstream of the genes we assume that it represents a novel type of a transcriptional terminator in *Thermus* species. It is possible that a Rho-like factor is involved in the termination, because the series of several U residues typical for Rho-independent terminators lacks in the mRNA downstream of the stem-loop structure. Provided that the stem-loop structure is transcribed, it could also stabilize the 3′-portion of the mRNA.

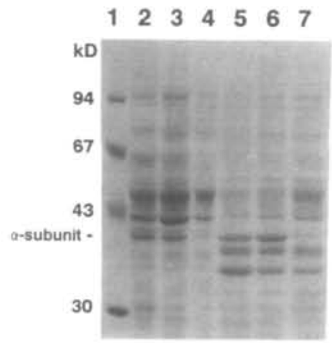

Figure 5. Expression of the PheRS subunits analyzed by SDS PAGE. Lane 1: molecular weight marker; lanes 2–4: crude cellular extracts; lanes 5–7: cell debris dissolved in 10% SDS; *E.coli* SG13009(pREP4+pPheS1) (lanes 2 and 5), *E.coli* SG13009(pREP4+pPheST1) (lanes 3 and 6), and *E.coli* SG13009(pREP4) without expression plasmid (lanes 4 and 7).

Expression of the *pheST* genes in *E.coli*

In order to obtain large amounts of the PheRS subunits for subsequent studies, the genes *pheS* and *pheST* were cloned in *E.coli* using the expression vector pKK223-3, which has successfully been used for the overproduction of other *T.thermophilus* genes (36,48). This was achieved by introducing restriction sites into the DNA upstream (*Eco*RI) and downstream (*Hind*III) of the *pheS* gene by the means of PCR. *E.coli* JM109 was transformed with the ligation mixture of the resulting PCR product hydrolyzed with the respective endonucleases and the linearized vector pKK223-3. For preliminary studies on the expresssion of both genes the *pheT* gene was linked to *pheS* using a *Bgl*II restriction site situated in *pheS* (Fig. 2). The resulting plasmids, designated as pPHES1 and pPHEST1, were used to transform *E.coli* SG13009(pREP4), a strain which is well suited for expression experiments. Crude cellular extracts as well as cell debris dissolved in 10% SDS were analysed by SDS PAGE (Fig. 5). A major band corresponding to the α-subunit of PheRS occurs in the protein patterns of both expression clones. About half of this protein is insoluble, which is probably due to its tendency to aggregate (24). No band corresponding to the β-subunit was found in the case of the pPheST1 containing clone, indicating that the expression of the *pheT* gene is low. However, a specific PheRS activity was detected in the thermally denatured cellular extract of this clone which was about four times higher than the activity determined for the *T.thermophilus* crude cellular extract. This proved the formation of an active enzyme in the host bacteria. The activity cannot originate solely from the α-subunit, since neither of the subunits are catalytically active in aminoacylation (24). It can further be ruled out that an active hybrid between the α-subunit of

T.thermophilus and the β-subunit of *E.coli* is formed, since the enzyme of the mesophilic bacterium is not stable at 60°C (49). The weak expression of *pheT* could be explained by the occurence of an AGG codon for arginine in the second position, which is one of the least used codons in *E.coli* and was shown to drastically reduce expression of the concerned genes (50). This codon was shown to reduce *lacZ* expression five fold when introduced at the second position, although this is the position at which the AGG codon most frequently occurs (51). Changing this codon may improve the expression of *pheT*.

ACKNOWLEDGEMENTS

Dr. E.V.Bobkova was financially supported by the Federation of European Biochemical Societies. Technical assistance of M.Lenk is greatly acknowledged. We thank Dr. V.N.Ankilova for the excellent expertise in enzyme purification and N.Schönbrunner for reading the manuscript.

REFERENCES

1. Brick, P., Bhat, T.N. and Blow, D.M. (1989) *J. Mol. Biol.* **208**, 83–98.
2. Rould, M.A., Perona, J.J., Söll, D. and Steitz, T.A. (1989) *Sci.* **246**, 1135–1142.
3. Brunie, S., Zeiwer, C. and Risler, J.-L. (1990) *J. Mol. Biol.* **216**, 411–424.
4. Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N. and Leberman, R. (1990) *Nature* **347**, 249–255.
5. Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J.C. and Moras, D. (1991) *Sci.* **252**, 1682–1689.
6. Eriani, G., Delarue, M., Poch, O., Gangloff, J. and Moras, D. (1990) *Nature* **347**, 203–206.
7. Burbaum, J.J., Starzyk, R.M. and Schimmel, P. (1990) *Proteins: Struct. Funct. Genet.* **7**, 99–111.
8. Fayat, G., Blanquet, S., Dessen, P., Batelier, G. and Waller, J.P. (1974) *Biochimie* **56**, 35–41.
9. Sanni, A., Walter, P., Boulanger, Y., Ebel, J.-P. and Fasiolo, F. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8387–8391.
10. Ankilova, V.N., Reshetnikova, L.S., Chernaya, M.M. and Lavrik, O.I. (1988) *FEBS Lett.* **227**, 9–13.
11. Sprinzl, M. and Cramer, F. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3049–3053.
12. Fraser, T.H. and Rich, A. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3044–3048.
13. Sprinzl, M. (1991) *Biological Chemistry Hoppe-Seyler* **372**, 759.
14. Mechulam, Y., Fayat, G. and Blanquet, S. (1985) *J. Bacteriol.* **163**, 787–791.
15. Brakhage, A.A., Wozny, M. and Putzer, H. (1990) *Biochimie* **72**, 725–734.
16. Springer, M., Trudel, M., Graffe, M., Plumbridge, J., Fayat, G., Mayaux, J.-F., Sacerdot, C., Blanquet, S. and Grunberg-Manago, M. (1983) *J. Mol. Biol.* **171**, 263–279.
17. Nureki, O., Muramatsu, T., Suzuki, K., Kohda, D., Matsuzawa, H., Ohta, T., Miyazawa, T. and Yokoyama, S. (1991) *J. Biol. Chem.* **2660**, 3268–3277.
18. Kast, P. and Hennecke, H. (1991) *J. Mol. Biol.* **222**, 99–124.
19. Sanni, A., Walter, P., Ebel, J.-P. and Fasiolo, F. (1990) *Nucl. Acids Res.* **18**, 2087–2092.
20. Khodyreva, S.N., Moor, N.A., Ankilova, V.N. and Lavrik, O.I. (1985) *Biochim. Biophys. Acta* **830**, 206–212.
21. Hountondji, C., Schmitter, J.-M., Beauvallet, C. and Blanquet, S. (1987) *Biochem.* **26**, 5433–5439.
22. Sanni, A., Hountondji, C., Blanquet, S., Ebel, J.-P., Boulanger, Y. and Fasiolo, F. (1991) *Biochem.* **30**, 2448–2453.
23. Chernaya, M.M., Korolev, S.V., Reshetnikova, L.S. and Safro, M.G. (1987) *J. Mol. Biol.* **198**, 555–556.
24. Bobkova, E.V., Mashanov-Golikov, A.V., Wolfson, A., Ankilova, V.N. and Lavrik, O.I. (1991) *FEBS Lett.* **290**, 95–98.
25. Bobkova, E.V., Gedrovich, A.V., Ankilova, V.N., Lavrik, O.I., Baratova, L.A. and Shishkov, A.V. (1990) *Biochem. Intern.* **20**, 1001–1009.
26. Castenholz, R.W. (1969) *Bacteriol. Rev.* **33**, 476–504.
27. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene* **33**, 103–119.
28. Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557–580.
29. Gottesman, S., Halpern, E. and Trisler, P. (1981) *J. Bacteriol.* **148**, 265–273.
30. Miller, J. (1972). Experiments in Molecular Genetics, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y..
31. Hewick, R.M., Hunkapiller, M.W., Hood, L.E. and Dreyer, W.J. (1981) *J. Biol. Chem.* **256**, 7990–7997.
32. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Smith, J.A., Seidman, J.G. and Struhl, K. (1987). Current Protocols in Molecular Biology, Greene Publishing Associates, New York.
33. Grunstein, A. and Hogness, D.S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961–3965.
34. Southern, E.M. (1975) *J. Mol. Biol.* **98**, 503–517.
35. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y..
36. Park, H.-J., Kreutzer, R., Reiser, C.O.A. and Sprinzl, M. (1992) *Eur. J. Biochem.* **205**, 875–879.
37. Henikoff, S. (1984) *Gene* **28**, 351–359.
38. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
39. Leberman, R., Antonsson, B., Giovanelli, R., Guariguata, R., Schumann, R. and Wittinghofer, A. (1980) *Anal. Biochem.* **104**, 29–36.
40. Laemmli, U.K. (1970) *Nature* **227**, 680–685.
41. Keller, B., Kast, P. and Hennecke, H. (1992) *FEBS Lett.* **301**, 83–88.
42. Chou, P.Y. and Fasman, G.D. (1978) *Adv. Enzym.* **47**, 45–148.
43. Grawunder, U., Schön, A. and Sprinzl, M. (1992) *Nucl. Acids Res.* **20**, 137.
44. Murzina, N.V., Vorozheykina, D.P. and Matvienko, N.I. (1988) *Nucl. Acids Res.* **16**, 8172–8172.
45. Koyama, Y. and Furukawa, K. (1990) *J. Bacteriol.* **172**, 3490–3495.
46. Nicholls, D.J., Sundaram, T.K., Atkinson, T. and Minton, N.P. (1990) *FEMS Microbiol. Lett.* **70**, 7–14.
47. Hulton, C.S.J., Higgins, C.F. and Sharp, P.M. (1991) *Mol. Microbiol.* **5**, 825–834.
48. Ahmadian, M.R., Kreutzer, R. and Sprinzl, M. (1991) *Biochimie* **73**, 1037–1043.
49. Bobkova, E.V., Stepanov, V.G. and Lavrik, O.I. (1992) *FEBS Lett.* **302**, 54–56.
50. Chen, G.-F.T. and Inouye, M. (1990) *Nucl. Acids Res.* **18**, 1465–1473.
51. Looman, A.C., Bodlaender, J., Comstock, L.J., Eaton, D., Jhurani, P., de Boer, H.A. and van Knippenberg, P.H. (1987) *EMBO J.* **6**, 2489–2492.
52. Fayat, G., Mayaux, J.-F., Sacerdot, C., Fromant, M., Springer, M., Grunberg-Manago, M. and Blanquet, S. (1983) *J. Mol. Biol.* **171**, 239–261.
53. Nishiyama, M., Matsubara, N., Yamamoto, K., Iijima, S., Uozumi, T. and Beppu, T. (1986) *J. Biol. Chem.* **261**, 14178–14183.
54. Croft, J.E., Love, D.R. and Bergquist, P.L. (1987) *Mol. Gen. Genet.* **210**, 490–497.
55. Hartmann, R.K., Ulbrich, N. and Erdmann, V.A. (1987) *Biochimie* **69**, 1097–1104.
56. Struck, J.C.R., Toschka, H.Y. and Erdmann, V.A. (1988) *Nucl. Acids Res.* **18**, 9042–9042.
57. Hartmann, R.K. and Erdmann, V.A. (1989) *J. Bacteriol.* **171**, 2933–2941.
58. Weisshaar, M.P., Ahmadian, R., Sprinzl, M., Satoh, M., Kushiro, A. and Tomita, K. (1990) *Nucl. Acids Res.* **18**, 1902–1902.
59. Sato, S., Nakada, Y., Kanaya, S. and Tanaka, T. (1988) *Biochim. Biophys. Acta* **950**, 303–312.