Diss. ETH Nr. 9468

Investigations on *Escherichia coli* Phenylalanyl-tRNA Synthetase at the Molecular Level:

Identification and Genetic Engineering of a Phenylalanine Specificity Determinant and Possible Application of a Relaxed Specificity Mutant

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZÜRICH
for the degree of
DOCTOR OF NATURAL SCIENCES

presented by
PETER KAST
Dipl. Natw. ETH
born on June 12, 1961
citizen of Wetzikon (ZH) and Fischingen (TG)

accepted on the recommendation of
Prof. Dr. H. Hennecke, examiner
Prof. Dr. T. Leisinger, co-examiner

Zürich 1991

Dank

An erster Stelle danke ich Herrn Prof. Hauke Hennecke für seine ausserordentlich grosszügige und kompetente Unterstützung meiner wissenschaftlichen Tätigkeit und ganz besonders für die mir gebotene Möglichkeit, meine Arbeit in Übereinstimmung mit meinen Vorstellungen verwirklichen zu können. Er hat mir mit dem "pheS-Projekt" ein Thema anvertraut, das mich noch heute fasziniert.

Herrn Prof. Thomas Leisinger danke ich für die Übernahme des Korreferats und für seine wertvollen Anregungen zu weiterführenden Experimenten.

Claudia Wehrli, mit der ich zwei Jahre lang zusammenarbeiten durfte, gebührt ein besonderer Dank. Ohne ihren grossen Einsatz und ihr technisches Können wäre es nicht möglich gewesen, die grosse Zahl von PheRS Mutanten in Kapitel III zu konstruieren und zu analysieren.

Barbara ("Babs") Keller hat während ihrer Semester- und Diplomarbeit die in Kapitel V beschriebenen Projekte bearbeitet. Ich danke ihr herzlich für die sehr angenehme und ebenso erfolgreiche Zusammenarbeit.

Sergio Cantoreggi schulde ich Dank für die während seiner Diplomarbeit synthetisierten Phenylalanin-Derivate. Herrn Prof. D. Seebach und Herrn Dr. B. Martinoni danke ich für die anregenden Diskussionen, und Klaus Girgenrath und Lucas Dürselen bin ich für die Scintillationsmessungen dankbar.

Den früheren und aktuellen Mitgliedern des "Oligoteams" (Christoph Kündig, Beat Thöny, Stefan Irniger, Armand Savioz, Tobias Schmidheini, Dietmar Stax, Sergio Schmid) danke ich für die gute und kameradschaftliche Zusammenarbeit.

Ich danke allen Mitgliedern der "Rhizo"-Gruppe, und im besonderen meinen LaborkollegInnen Sabine Ebeling, Christoph Kündig und Brigitte Herzog-Pfluger, für das ausgezeichnete Arbeitsklima und die Tatsache, dass sie mich trotz meines exotischen Projekts recht gut akzeptiert haben.

Allen Angehörigen der Familien Kast und Gamper bin ich für ihre Unterstützung während meiner Studenten- und Doktoranden-Zeit Dank schuldig. Und ***auch*** *Marianne Gamper. Sie hat einen wesentlichen Anteil am Gelingen dieser Arbeit.*

TABLE OF CONTENTS

CHAPTER I

Synopsis of the Work

1. Introduction: the importance of phenylalanyl-tRNA synthetase in cellular metabolism

Every living cell uses a similar basic equipment for the accurate transmission of the genetic information. The first step involves the synthesis of mRNA, a transcript of the protein-encoding gene(s), which is then translated by the ribosomes in a second step. Both processes need to proceed with high precision to ensure the accuracy of protein biosynthesis. The correct function of another set of reactions occurring before the ribosomal step, however, is at least as crucial. The 'true translation' of the genetic code, *i. e.* the assignment of the proteinogenic amino acids to nucleic acid triplets, is already made through the aminoacylation reactions (Schimmel, 1991). These reactions are carried out by the aminoacyl-tRNA synthetases which couple the amino acids to a tRNA molecule that possesses the corresponding amino acid-specific anticodon (Hershey, 1987; Schimmel, 1987). Typically, a (bacterial) cell has 20 aminoacyl-tRNA synthetases, one for each triplet-encoded amino acid. Recently, however, examples for the existence of two enzymes per amino acid were discovered (Lévêque *et al.*, 1990; Clark & Neidhardt, 1990; Putzer *et al.*, 1990). The peculiarity of the aminoacyl-tRNA synthetase-catalysed reactions is their extraordinarily high specificity. These enzymes must be able to discriminate strictly between cognate and non-cognate substrates (of both amino acids and tRNAs) which may have rather similar structures. This high precision is accomplished by sophisticated substrate recognition and product proof-reading mechanisms (Freist, 1989; Söll, 1990).

The focus in this work is on phenylalanyl-tRNA synthetase (PheRS) from

Abbreviations

ATP, adenosine triphosphate; DNA, deoxyribonucleic acid; K_M, Michaelis Menten constant, indicative for substrate affinity; mRNA, messenger ribonucleic acid; *p*-F-Phe, *para*-fluoro-phenylalanine; Phe, phenylalanine; PheRS, phenylalanyl-tRNA synthetase; $tRNA^{(Phe)}$, transfer ribonucleic acid (specific for Phe).

Escherichia coli. The enzyme catalyses the attachment of phenylalanine (Phe) to its cognate $tRNA^{Phe}$; the reaction is driven by the hydrolysis of ATP. PheRS is one of the largest enzymes in the very heterogeneous class of aminoacyl-tRNA synthetases and possesses an $\alpha_2\beta_2$ quarternary structure. The primary structure of *E. coli* PheRS is known from the sequence analysis of *pheS* and *pheT*, the genes encoding the α (small) and β (large) subunits, respectively. The presence of three short consensus sequence motifs identifies PheRS as a member of class II aminoacyl-tRNA synthetases (Eriani *et al.*, 1990). Up to now, there is no information available on the tertiary structure of any PheRS. (For details and references see Introduction to Chapter II).

2. Approaches to elucidate the phenylalanine binding site of *E. coli* phenylalanyl-tRNA synthetase

PheRS must discriminate between the cognate substrate Phe and non-cognate amino acids in the cell, such as the rather similar tyrosine. Therefore, the Phe binding site of PheRS was believed to contain a prototype of a binding pocket specific for a non-substituted aromatic ring. In this work it was attempted to localize and characterize this binding site by molecular-genetic means. The investigations that were carried out can be subdivided into three levels, as follows.

Level 1: Localization of a presumptive Phe binding site in *E. coli* PheRS (Chapter II)

Although PheRS does not significantly aminoacylate naturally occurring amino acids other than the cognate substrate Phe, it was shown that different xenobiotic ring-substituted Phe analogues were able to interact with wild-type PheRS; and some of them (*e. g. p*-fluoro-phenylalanine; *p*-F-Phe) were even attached to $tRNA^{Phe}$ (Santi & Danenberg, 1971; Gabius *et al.*, 1983). Hennecke & Böck (1975) described two *E. coli* PheRS mutant strains which were resistant to *p*-F-Phe due to the production of altered PheRS α subunits as the cause for an exclusion of *p*-F-Phe from the enzymatic reaction. These strains seemed extremely well suited for the localization of the Phe binding site in PheRS. Another *E. coli* strain with a PheRS α subunit defect leading to an increased K_M for Phe (*i. e.* a lowered Phe affinity; Grüll *et al.*, 1979) was analyzed in parallel (Chapter II).

The mutant *pheS* genes of these strains were cloned by complementation to

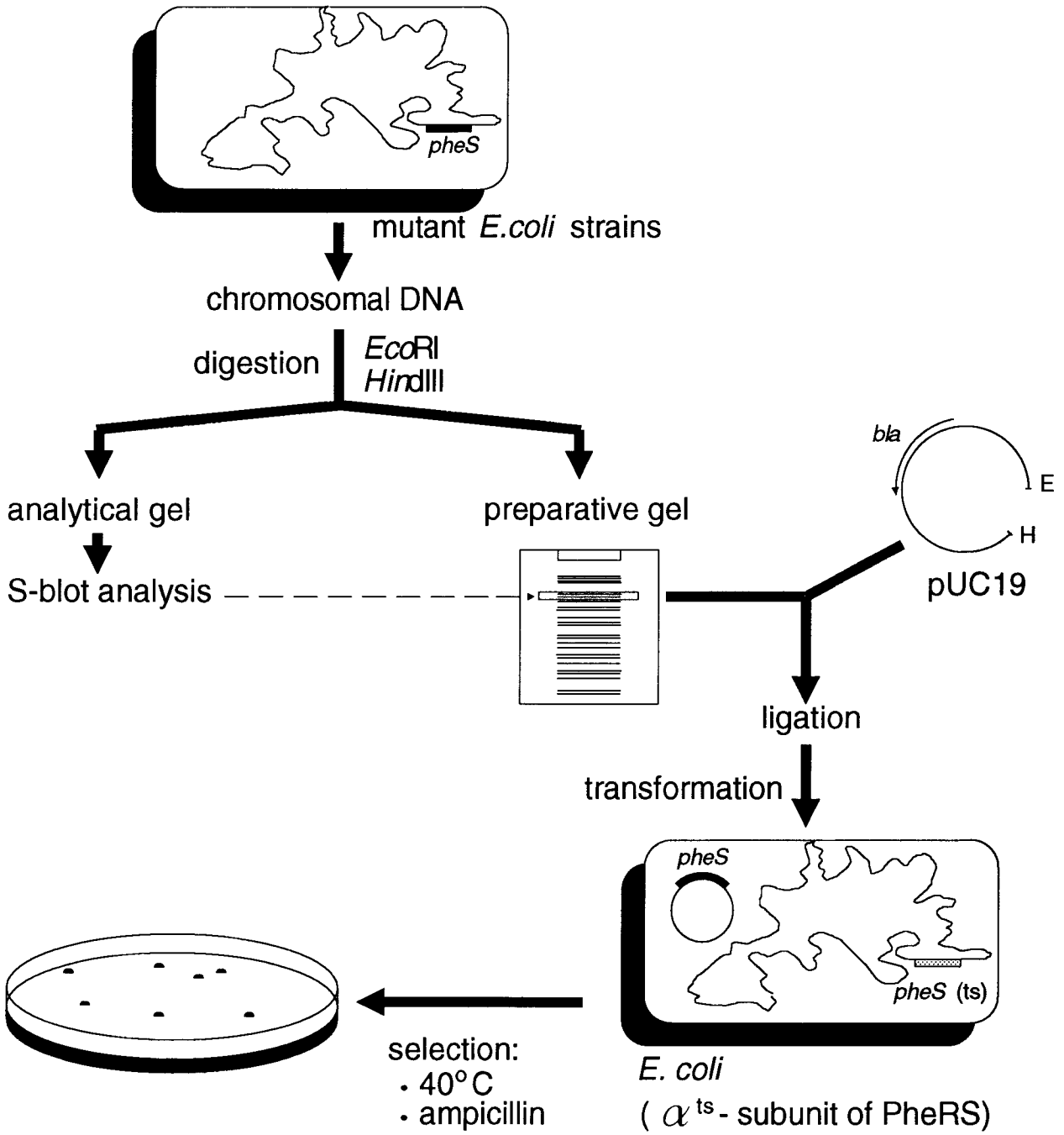


Figure 1. Cloning of mutant *pheS* genes by complementation of a thermosensitive *E. coli* strain. The chromosomal DNA from three *E. coli* mutant strains was isolated and digested with the indicated restriction enzymes. The presence of the fragment carrying the *pheST* genes (described in Chapter II) was verified by Southern blot analysis. DNA fragments of the appropriate size were cut out of a preparative agarose gel and ligated to a plasmid vector (pUC19), previously digested with the same enzymes. The ligation mixture was introduced into a thermosensitive (ts) *E. coli* strain (the mutation in *pheS* leading to thermosensitivity is characterized in Chapter II). Selection at elevated temperature on ampicillin-containing agar plates allowed only growth of cells containing the *pheS* gene on the plasmid. E and H denote sites for *Eco*RI and *Hin*dIII, respectively; *bla*, gene for β-lactamase conferring ampicillin resistance.

thermoresistance of an *E. coli* strain normally unable to grow at elevated temperatures because it possesses a thermosensitive PheRS α subunit (Fig. 1). The nucleotide sequence analysis revealed that the mutation responsible for the K_M increase resulted in a glycine-to-aspartic acid exchange at position 191 of the α subunit. Interestingly, this mutation mapped to motif 2, one the conserved class II aminoacyl-tRNA synthetase sequences. The mutation responsible for *p*-F-Phe resistance in both analyzed *pheS* alleles caused an alanine-to-serine exchange at position 294 of the α subunit, precisely within motif 3, another typical class II tRNA synthetase sequence. It seemed reasonable to suggest, therefore, that the mutations identified the conserved motifs 2 and 3 as regions that participate in the formation of the Phe binding site of PheRS and, more generally, in the amino acid binding site of class II aminoacyl-tRNA synthetases (Chapter II).

Level 2: Comparative analysis of PheRSs from different origins (Chapters II and V)

By comparing the PheRS α subunit sequence of *E. coli* with the homologous sequences of *Bacillus subtilis* and yeast cytoplasmic and mitochondrial PheRSs, a particularly good conservation of the region around position 294 of the *E. coli* protein became apparent (Chapter II). It is generally accepted that conservation of primary structure motifs is an indication for a functional or structural importance of the corresponding residues. The alignment of PheRS sequences around position 294 revealed a much more pronounced conservation in this region than in the immediately adjacent, overlapping motif 3 region. This suggested that the mutation leading to *p*-F-Phe resistance has affected a site that is important and specific for the function of PheRSs.

To further test this hypothesis, it was of interest to examine more PheRS sequences from different origins for the degree of sequence conservation within the motif 3 region. In Chapter V, the cloning and partial sequencing of the *pheS* gene from *Salmonella typhimurium*, a close relative of *E. coli*, is described. The cloning was achieved by complementation of a thermosensitive *E. coli* PheRS mutant in an analogous way as shown in Fig. 1. The second organism chosen for cloning of the *pheS*-homologous gene was *Thermus thermophilus* which is only very distantly related to *E. coli* (Chapter V). Suitable DNA fragments carrying the *T. thermophilus pheS* gene were identified by Southern blot hybridizations. Two types of hybridization probes were used: the heterologous *E. coli pheS* gene and five oligonucleotides that were designed on the basis of conserved regions in the known PheRS sequences and were adapted to the extreme codon usage of *T. thermophilus*. The

cloned *T. thermophilus pheS* gene was finally identified by colony hybridization and then sequenced partially (Chapter V).

Figure 2 shows an alignment of PheRS small subunit sequences including the corresponding regions from *S. typhimurium* and *T. thermophilus*. The newly determined sequences confirmed the previous observation that the region around position 294 of the *E. coli* PheRS α subunit is well conserved among PheRSs from different origins. Such a PheRS-specific site might well be (part of) the Phe binding site on the enzyme.

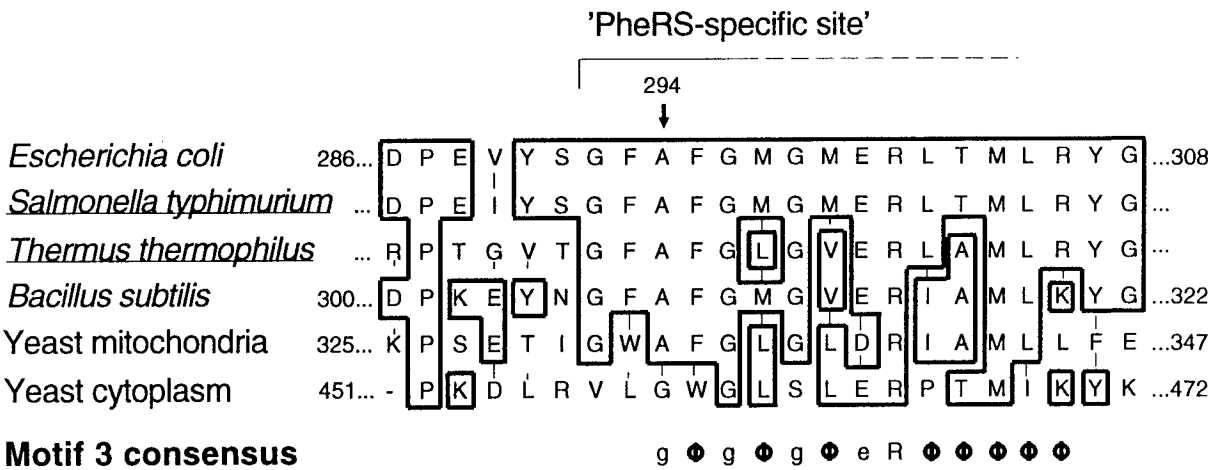


Figure 2. Alignment of PheRS small subunit sequences in a region affected by the *p*-F-Phe resistance mutation. The Figure contains PheRS sequences corresponding to the region around the *E. coli* α subunit residue 294 (marked by an arrow). The PheRS sequences from *S. typhimurium* and *T. thermophilus* (species names underlined) were determined in Chapter V. Identical amino acids (in one-letter-code) are boxed; related residues are connected by thin lines; Φ in the motif 3 consensus (adapted from Eriani *et al.*, 1990) stands for hydrophobic amino acids. More details and references to the sequences are given in Chapter II.

Level 3: Mutational analysis of a potential Phe binding site in *E. coli* PheRS (Chapter III)

As stated above (level 1), the sequencing of *pheS* alleles revealed a point mutation leading to an alanine-to-serine exchange at position 294 of the PheRS α subunit as the cause for *p*-F-Phe resistance. From a mechanistic point of view it might be hypothesized that the additional hydroxyl group of serine prevents the entering of the slightly bigger substrate analogue *p*-F-Phe into the binding pocket of PheRS (Fig. 3). Taking into account the conservation of amino acids around position 294 (see level 2), a working model for the presumptive aromatic binding site

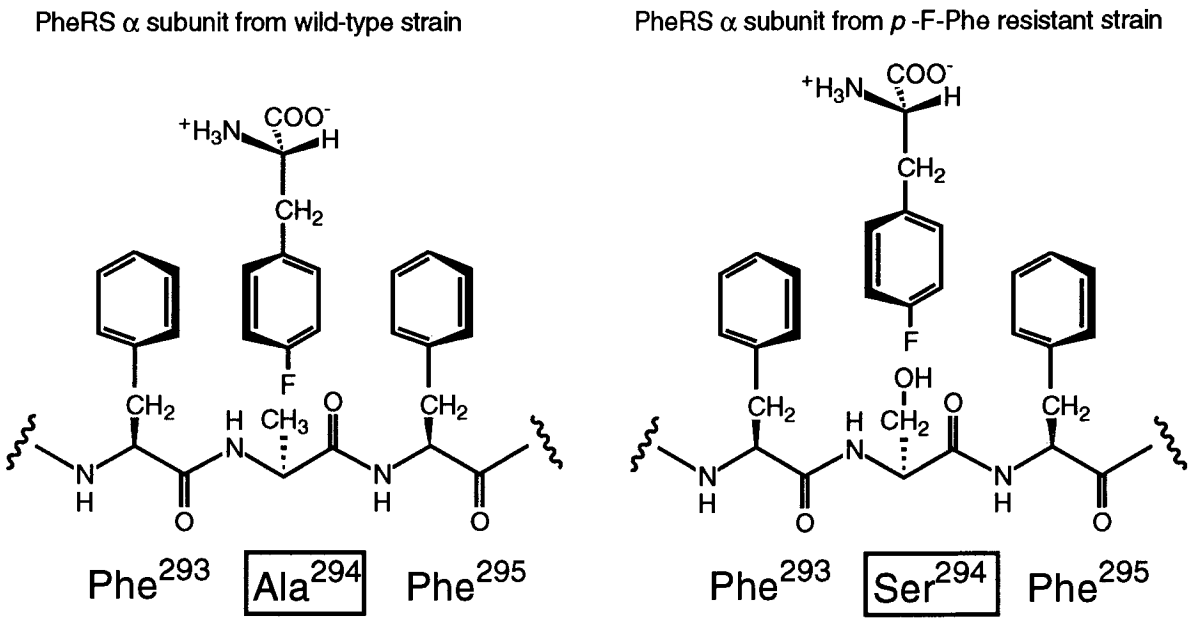


Figure 3. Working hypothesis for the Phe binding site in *E. coli* PheRS and the mechanism of *p*-F-Phe resistance. In this preliminary working model the binding site for the substrate's aromatic ring is formed by residue 294 (contacting the *para*-position) and the two phenylalanines 293 and 295. The discrimination against *p*-F-Phe in the mutant enzyme Ser^{294} (on the right side) is accomplished by steric interference of the additional hydroxyl group with the *para*-substituent of this substrate analogue. This may prevent productive binding of the analogue in the Phe binding pocket.

in PheRS was proposed (Chapter III). This working hypothesis anticipated interactions of the residue at position 294 with the *para*-position of the amino acid substrate's aromatic ring, whereas the flanking aromatic residues 293 and 295 would interact with the ring system of the substrate (Fig. 3).

This model was now tested at a functional level by site-directed mutagenesis of the alanine at position 294 as well as of the flanking Phe residues. The effects of the introduced point mutations were analyzed *in vitro* by measuring PheRS activity in strains carrying the mutated *pheS* genes or by a complementation assay *in vivo* (Chapter III). In contrast to the predictions made from the working model (Fig. 3), it turned out that the replacement of Phe^{293} and Phe^{295} by selected other amino acids did not directly affect Phe binding. However, PheRS stability seemed to be dependent on the nature of those residues. As drawn in Fig. 4 and discussed in Chapter III, the aromatic residues at positions 293 and 295 might rather play an important role in stabilizing subunit interactions in PheRS.

The effects from mutagenesis of the alanine at position 294 clearly showed that the binding of Phe was influenced by the individual replacements. This became obvious from the altered affinities of the mutant enzymes for phenylalanine

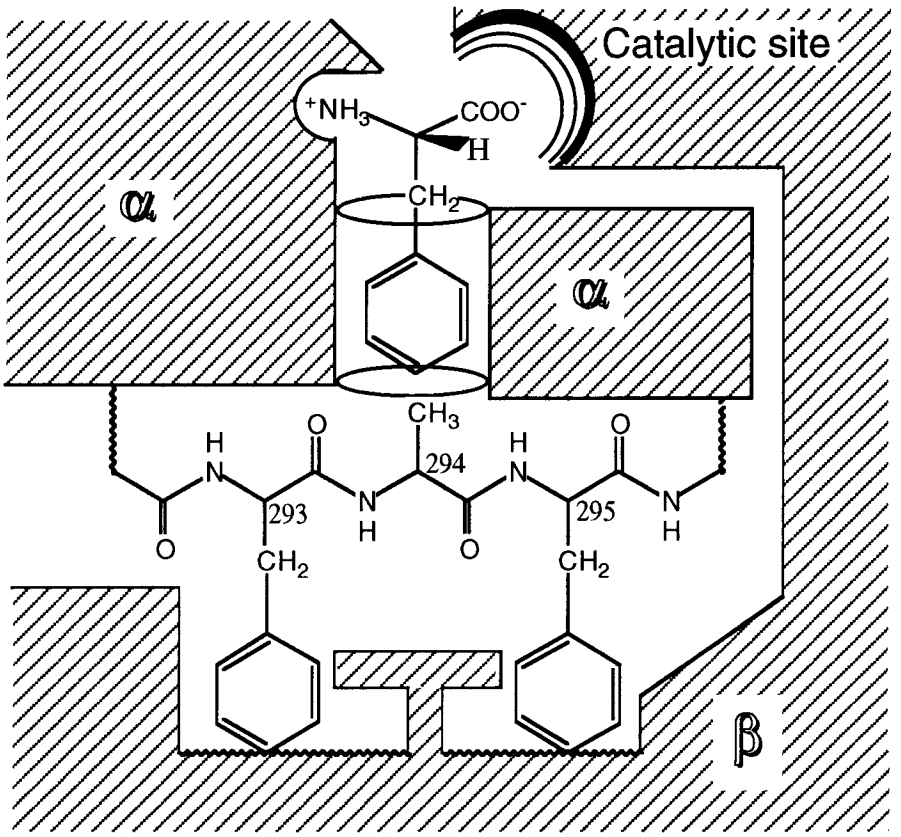


Figure 4. Hypothetical scheme of the Phe binding site in PheRS. The model summarizes the interpretation of results from amino acid exchanges at positions 293, 294 and 295 of the α subunit (cf. Chapter III). Residue 294 is assumed to be in contact with the *para*-position of the substrate phenylalanine. The adjacent residues (293 and 295), however, do not directly influence substrate binding, but might contribute to protein stability by interacting with another region of PheRS, possibly the β subunit. As discussed in Chapter III, the catalytic site of PheRS may be at the α/β subunit interface.

and by the pronounced shifts in the spectrum of substrate analogues able to interact with mutant PheRSs, as compared to the wild-type enzyme. The most interesting mutation was the exchange of alanine by the smaller glycine. The resulting mutant PheRS Gly^{294} showed a much better aminoacylation of the Phe analogue *p*-F-Phe than the wild-type enzyme. Furthermore, as one might have predicted from the working model in Fig. 3, the Gly^{294} variant allowed interactions even with the larger *p*-chloro and *p*-bromo derivatives of Phe; moreover, all three *para*-halogenated Phe analogues became highly toxic *in vivo* in *E. coli* strains producing the Gly^{294} PheRS. In contrast, growth of the wild-type strain was less affected by *p*-F-Phe; and *p*-chloro and *p*-bromo analogues of Phe showed no effect at all. Obviously, the methyl group of alanine in the wild-type binding pocket prevented the

accommodation of the larger *para*-chloro and -bromo substituents.

The results from replacements at position 294 confirmed the hypothesis that direct contacts occur between amino acid 294 and the *para*-position of the substrate's aromatic ring (Chapter III). As displayed in Fig. 4, it thus seems very likely that at least part of the Phe binding site in PheRS is formed by this motif 3 residue in the enzyme of *E. coli* and perhaps also of other organisms.

3. Further perspectives

Exploitation of a PheRS mutant with relaxed substrate specificity (Chapter IV)

One major result of the mutagenesis experiments described in Chapter III is the finding that position 294 of the PheRS α subunit determines specificity for *para*-substituted Phe analogues. The example with PheRS variant Gly^{294} showed that even a highly precise aminoacyl-tRNA synthetase can be manipulated by single amino acid exchanges to use a broader substrate spectrum. Furthermore, these findings opened up a perspective for the application of the Gly^{294} PheRS in the production of artificial proteins: this PheRS mutant with relaxed substrate specificity may be exploited in an *in vivo* system for the deliberate incorporation of various Phe analogues into proteins.

In Chapter IV the possible set-up of such systems and the potential applications to production of analogue-substituted, pharmaceutically interesting target proteins are discussed.

Towards the determination of the PheRS tertiary structure (Chapter V)

As mentioned above, the *pheS* gene from *T. thermophilus* was cloned to obtain a sequence for comparisons among PheRS α subunits from different origins. Another important reason for chosing this organism was the availability of *T. thermophilus* PheRS crystals suitable for X-ray diffraction analysis (Ankilova *et al.*, 1988). The complete sequencing of the *pheS* and *pheT* genes from *T. thermophilus* may finally permit the elucidation of the first high resolution structure of an aminoacyl-tRNA synthetase of the $\alpha_2\beta_2$ type (Chapter V). This would subsequently allow the definitive determination of the topology of the Phe binding site on the enzyme.

4. References

Ankilova, V. N., Reshetnikova, L. S., Chernaya, M. M. & Lavrik, O. I. (1988). Phenylalanyl-tRNA Synthetase from *Thermus thermophilus* HB8. Purification and Properties of the Crystallizing Enzyme. *FEBS Lett.* **227**, 9-13.

Clark, R. L. & Neidhardt, F. C. (1990). Roles of the Two Lysyl-tRNA Synthetases of *Escherichia coli*: Analysis of Nucleotide Sequences and Mutant Behavior. *J. Bacteriol.* **172**, 3237-3243.

Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990). Partition of tRNA Synthetases into Two Classes Based on Mutually Exclusive Sets of Sequence Motifs. *Nature (London)*, **347**, 203-206.

Freist, W. (1989). Mechanisms of Aminoacyl-tRNA Synthetases: A Critical Consideration of Recent Results. *Biochemistry*, **28**, 6787-6795.

Gabius, H.-J., von der Haar, F. & Cramer, F. (1983). Evolutionary Aspects of Accuracy of Phenylalanyl-tRNA Synthetase. A Comparative Study with Enzymes from *Escherichia coli*, *Saccharomyces cerevisiae*, *Neurospora crassa*, and Turkey Liver Using Phenylalanine Analogues. *Biochemistry*, **22**, 2331-2339.

Grüll, J. M., Hennecke, H., Fröhler, J., Thomale, J., Nass, G. & Böck, A. (1979). *Escherichia coli* Mutants Overproducing Phenylalanyl- and Threonyl-tRNA Synthetase. *J. Bacteriol.* **137**, 480-489.

Hennecke, H. & Böck, A. (1975). Altered α Subunits in Phenylalanyl-tRNA Synthetases from *p*-Fluorophenylalanine-Resistant Strains of *Escherichia coli*. *Eur. J. Biochem.* **55**, 431-437.

Hershey, J. W. B. (1987). Protein Synthesis. In *Escherichia coli and Salmonella typhimurium*: *Cellular and Molecular Biology* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M. & Umbarger, H. E., eds), vol. 1, pp. 613-647, Am. Soc. Microbiol., Washington, DC.

Lévêque, F., Plateau, P., Dessen, P. & Blanquet, S. (1990). Homology of *lysS* and *lysU*, the Two *Escherichia coli* Genes Encoding Distinct Lysyl-tRNA Synthetase Species. *Nucl. Acids Res.* 18, 305-312.

Putzer, H., Brakhage, A. A. & Grunberg-Manago, M. (1990). Independent Genes for Two Threonyl-tRNA Synthetases in *Bacillus subtilis*. *J. Bacteriol.* **172**,

4593-4602.

Santi, D. V. & Danenberg, P. V. (1971). Phenylalanyl Transfer Ribonucleic Acid Synthetase from *Escherichia coli*. Analysis of the Phenylalanine Binding Site. *Biochemistry*, **10**, 4813-4820.

Schimmel, P. (1987). Aminoacyl tRNA Synthetases: General Scheme of Structure-Function Relationships in the Polypeptides and Recognition of Transfer RNAs. *Ann. Rev. Biochem.* **56**, 125-158.

Schimmel, P. (1991). Classes of Aminoacyl-tRNA Synthetases and the Establishment of the Genetic Code. *Trends Biochem. Sci.* **16**, 1-3.

Söll, D. (1990). The Accuracy of Aminoacylation - Ensuring the Fidelity of the Genetic Code. *Experientia*, **46**, 1089-1096.

CHAPTER II

Phenylalanyl-tRNA Synthetase Mutants of *Escherichia coli* Analyzed at the Molecular Level

(Submitted to: *Journal of Molecular Biology*)

Peter Kast and Hauke Hennecke

Mikrobiologisches Institut, Eidgenössische Technische Hochschule,
Schmelzbergstrasse 7,
CH-8092 Zürich, Switzerland

Running title: Phenylalanyl-tRNA Synthetase Mutants

Keywords: *p*-fluoro-phenylalanine resistance / phenylalanine binding site / phenylalanyl-tRNA synthetase / *pheS* gene / polymerase chain reaction

Non-standard abbreviations

bp: basepair(s); IPTG: isopropyl-β-D-thiogalactopyranoside; kb: kilobasepair(s); PCR: polymerase chain reaction; *p*-F-Phe: *p*-fluoro-phenylalanine; R: resistant; S: sensitive; ts: thermosensitive; XxxRS: aminoacyl-tRNA synthetase for amino acid Xxx in 3-letter-code.

1. Abstract

Neither the tertiary structure nor the location of active sites are known for phenylalanyl-tRNA synthetase (PheRS), a member of class II tRNA synthetases. In an attempt to detect the phenylalanine binding site, three *Escherichia coli* mutant strains producing a PheRS with altered substrate binding properties were analyzed genetically. The mutations were previously shown to map in *pheS*, the gene encoding the (small) α subunit of the tetrameric PheRS ($\alpha_2\beta_2$ structure). The mutant *pheS* genes were cloned by complementation of the thermosensitive strain NP37 ($PheRS^{ts}$), and the deviations from the wild-type gene were determined by sequence analysis. The point mutation leading to a decreased affinity for phenylalanine in PheRS from strain G1 resulted in a Gly^{191} to Asp^{191} exchange within motif 2, one of three conserved sequence motifs in class II aminoacyl-tRNA synthetases. The mutation causing resistance to *p*-fluoro-phenylalanine led to an Ala^{294} to Ser^{294} exchange in the α subunits from two independent analogue resistant strains (AB1360-12 and K10-F6). This alteration (S294) mapped in the well conserved C-terminal part of the α subunit, precisely within motif 3, another typical class II sequence. We thus propose that motifs 2 and 3 participate in the phenylalanine binding site of PheRS. This hypothesis is substantiated in the light of the recently resolved crystal structure of SerRS, another class II tRNA synthetase. Mutation S294 was also the key for proposing a mechanism by which the substrate analogue *p*-fluoro-phenylalanine is excluded from the enzymatic reaction. This may be achieved by steric interactions between the *para*-position of the aromatic ring and the amino acid residue at position 294. Accidentally, the mutation responsible for the thermosensitive PheRS of strain NP37 was also cloned and then verified by direct sequencing of amplified NP37 DNA generated by asymmetric PCR. It is suggested here that the resulting amino acid exchange (Gly^{98} to Asp^{98}) might cause subunit disaggregation due to electrostatic repulsion.

2. Introduction

By coupling amino acids to their cognate tRNAs, aminoacyl-tRNA synthetases catalyze key reactions in translating the genetic code. Aminoacyl-tRNA synthetases responsible for the same amino acid are generally well conserved across species. However, in spite of their common reaction mechanism, the individual synthetases for each of the 20 amino acids in a given organism are widely diverse in sequence, subunit size and quarternary structure (for reviews see Schimmel & Söll, 1979; Schimmel, 1987). To know the reason for this diversity might help unravel the origin of the genetic code. The dissimilarity could partly be due to independent evolution (Weiner & Maizels, 1987) and/or to the fact that the enzymes catalyze, besides the aminoacylation, a variety of unrelated reactions (Rapaport *et al.*, 1985; Grunberg-Manago; 1987), *e. g.* an intron-splicing by TyrRS in *Neurospora crassa* (Cherniack *et al.*, 1990) and LeuRS in yeast mitochondria (Herbert *et al.*, 1988; Labouesse, 1990). In addition, individual domains for precise substrate recognition (Rould *et al.*, 1989) and proofreading mechanisms (Freist, 1989) could have been acquired in later stages of evolution.

Recently, most aminoacyl-tRNA synthetases were classified into two non-related groups, according to the presence or absence of typical short consensus sequences (Burbaum *et al.*, 1990; Eriani *et al.*, 1990*a*). For three class I members, the three-dimensional structures derived from X-ray analyses are available: TyrRS from *Bacillus stearothermophilus* (Brick *et al.*, 1989), GlnRS from *E. coli* (Rould *et al.*, 1989) and an active fragment of *E. coli* MetRS (Brunie *et al.*, 1990). The structures suggest that these aminoacyl-tRNA synthetases are composed of distinct domains (Schimmel, 1987) that are conserved in secondary and tertiary rather than primary structure (Blow *et al.*, 1983; Rould *et al.*, 1989). In addition, the class I-specific consensus sequence elements were shown to carry the active sites. A HIGH-like sequence participates in the N-terminal section of a structurally conserved nucleotide binding fold ('Rossmann fold'; Rossmann *et al.*, 1974) for the binding of ATP (Blow *et al.*, 1983; Leatherbarrow *et al.*, 1985; Webster *et al.*, 1987; Burbaum *et al.*, 1990). Another sequence ('KMSKS') maps near the C-terminal end of the Rossmann fold (Burbaum *et al.*, 1990). It was originally identified by affinity labelling experiments (Hountondji *et al.*, 1985) and is thought to be close to the 3'-end of the bound tRNA substrate (Hountondji *et al.*, 1986; Starzyk *et al.*, 1987; Meinnel *et al.*, 1990) and seems to interact with ATP (Fersht *et al.*, 1988; Rould *et*

al., 1989; Hountondji *et al.*, 1990).

To date, the crystal structure of only one enzyme belonging to the class II aminoacyl-tRNA synthetases has been published. With the *E. coli* SerRS structure Cusack *et al.* (1990) confirmed the partition of tRNA synthetases into two classes also on a conformational basis by demonstrating the absence of the characteristic 'Rossmann fold' of class I enzymes. The SerRS substrates could not be co-crystallized; therefore, their binding sites have not yet been determined and related to the class II consensus motifs.

Besides structural analyses, extensive investigations at the functional level are currently being carried out. Their aim is to elucidate how the individual enzymes accomplish the highly specific recognition of their substrates, and to identify enzyme domains directly involved in aminoacylation. Most of this research concerned the problem of cognate tRNA recognition (Schimmel, 1987; Mellot *et al.*, 1989; Perona *et al.*, 1989; Edwards & Schimmel, 1990; Ghosh *et al.*, 1990; Meinnel *et al.*, 1991) and tRNA identity (Normanly & Abelson, 1989) and led to the notion of a 'second genetic code' (de Duve, 1988; Waldrop, 1989; Moras, 1990).

Our own research focusses on phenylalanyl-tRNA synthetase (PheRS) of *E. coli*. This enzyme is among the largest aminoacyl-tRNA synthetases known and displays an $\alpha_2\beta_2$ quarternary structure (Fayat *et al.*, 1974; Hanke *et al.*, 1974; Schimmel, 1987). Recently, the relative topology of its 4 subunits was investigated by neutron small-angle scattering (Dessen *et al.*, 1990). Both types of subunits are required for enzyme activity (Hennecke & Böck, 1975). The PheRS genes *pheS* and *pheT* (encoding the α and β subunits, respectively) had been isolated on a λ transducing phage (Hennecke *et al.*, 1977) and their nucleotide sequences were determined (Fayat *et al.*, 1983; Mechulam *et al.*, 1985). The gene arrangement and the two-step reaction mechanism of PheRS are shown in Fig. 1.

According to the presence of certain consensus sequence elements, PheRS belongs to the less well characterized class II tRNA synthetases (Eriani *et al.*, 1990*a*). Neither the tertiary structure of PheRS nor the location of the binding sites for Phe or ATP within the primary sequence are known. Affinity labelling experiments implied the large β subunit to contain $tRNA^{Phe}$ binding sites (Khodyreva *et al.*, 1985), and Hountondji *et al.* (1987) identified sequences in the N-terminal region of the β subunit to be near the binding site of the 3'-end of $tRNA^{Phe}$. Some of the analogous experiments carried out with yeast cytoplasmic PheRS confirmed this location (Sanni *et al.*, 1991), whereas other data implicated the small subunit of the yeast enzyme to contain the major tRNA binding sites (Renaud *et al.*, 1982;

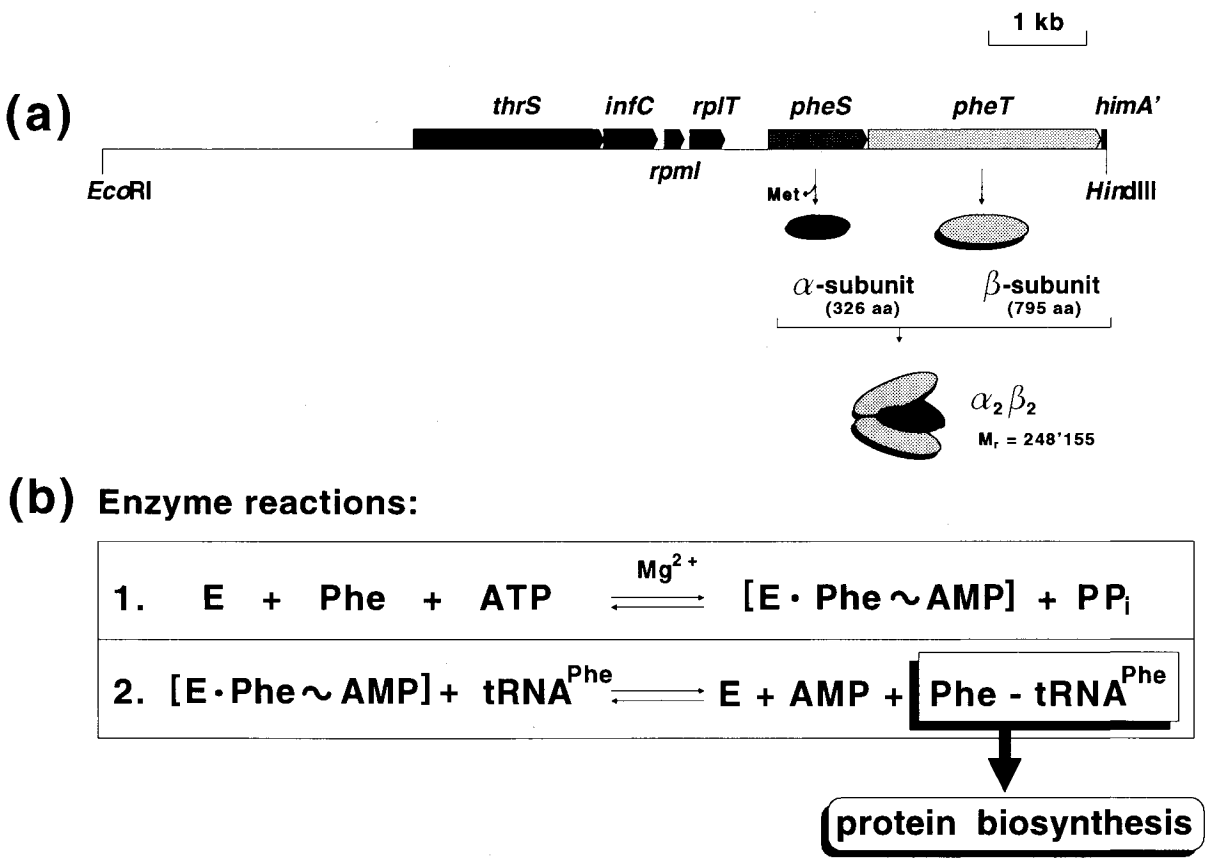


Figure 1. Gene organisation and enzymatic reaction of PheRS. (a) The genes *pheS* and *pheT* (encoding PheRS α and β subunits, respectively) are arranged in an operon at min 37 of the *E. coli* linkage map (Bachmann, 1990). Besides *pheS* and *pheT* the 10.25 kb *Eco*RI/*Hin*dIII fragment contains the genes encoding threonyl-tRNA synthetase (*thrS*), initiation factor IF3 (*infC*), ribosomal proteins A (or L35; encoded by *rpmI*) and L20 (*rplT*), and the N-terminal end of the α subunit of integration host factor (*himA*) (Fayat *et al.*, 1983; Mechulam *et al.*, 1985; Wada & Sako, 1987). PheRS is drawn as suggested by Dessen *et al.* (1990). aa, amino acids. **(b)** The aminoacylation of $tRNA^{Phe}$ proceeds (simplified) via a two-step reaction mechanism (Söll & Schimmel, 1974; Fasiolo & Fersht, 1978; Baltzinger & Holler, 1982). In the first step, the amino acid (Phe) is activated with ATP to form the enzyme-bound aminoacyl adenylate. Then, the amino acid is transferred to its cognate tRNA by releasing AMP. E, PheRS enzyme; PP_i, inorganic pyrophosphate.

Fasiolo *et al.*, 1989). Neither one of these locations, however, corresponds to the sites proposed for other aminoacyl-tRNA synthetases including GlyRS, the only other enzyme having an $\alpha_2\beta_2$ subunit structure (Schimmel, 1987; Toth & Schimmel, 1990*a*). Other substrate crosslinking studies suggested that the active site might be at the αβ interface in PheRS of *E. coli* (Khodyreva *et al.*, 1985) and yeast (Baltzinger *et al.*, 1979). Such a hypothesis was also discussed for GlyRS,

which is immunologically related to PheRS (Nagel *et al.*, 1988), based on an analysis of α and β chain mutants (Toth & Schimmel, 1990*b*).

The aim of this work was to localize the Phe binding site of PheRS by analyzing *E. coli* mutant strains possessing enzymes with altered substrate binding properties. In this approach, two different types of mutants were used. One of the strains is Phe auxotrophic due to an increased K_M of PheRS for this substrate (Grüll *et al.*, 1979). Two other, independently isolated strains are resistant to *p*-fluoro-phenylalanine (*p*-F-Phe). This phenotype is caused by a mutant PheRS that, in contrast to wild-type PheRS, excludes this toxic substrate analogue (Richmond, 1962; Fangman & Neidhardt, 1964) from the enzymatic reaction (Hennecke & Böck, 1975). All three mutations were mapped to *pheS* (Hennecke & Böck, 1975; Grüll *et al.*, 1979); therefore, we cloned and sequenced the mutant *pheS* genes. It was expected that especially the *p*-F-Phe resistance mutation, which led to a discrimination in PheRS between Phe and the slightly bigger substrate analogue, should point directly to the Phe binding site.

3. Materials and Methods

3.1. Bacterial strains, phages and DNA

The bacterial strains, vectors and plasmids used in this work are listed in Table 1. A phage P1kc lysate from strain JC10289/pKY102 (Ihara *et al.*, 1985) for the construction of *recA*⁻ strains was obtained from A. Birkmann (University of Munich, Germany). VCS-M13 helper phage for single-strand DNA isolation was from Stratagene (San Diego CA, USA). Oligonucleotides (listed in Table 2) used for mutagenesis, PCR and conventional (radioactive) sequencing were synthesized on an Applied Biosystems DNA synthesizer model 380B. *pheS*-specific dye-primers for automated DNA sequencing were kindly provided by G. Zon (Applied Biosystems, Foster City CA, USA). Dye-primer -36M13 was purchased from the same company. λ-DNA was from Biofinex (Praroman, Switzerland).

3.2. Microbiological techniques

E. coli strains were usually grown in LB-medium (Miller, 1972) at the permissive temperature. The antibiotics used were ampicillin at 150 μg/ml, tetracyclin at 15 μg/ml or streptomycin at 100 μg/ml. Construction of the *recA*⁻ strain KA2 was accomplished by generalized transduction (Masters, 1985). A P1kc lysate from the *recA*⁻ deletion strain JC10289/pKY102 (Ihara *et al.*, 1985) was used which allowed cotransduction of the *recA*⁻ marker and a tetracyclin resistance gene. The procedure of Miller (1972) was followed, except that incubations were done at 30°C. Multiplicity of infection was 0.01. After preadsorption, the mix was diluted with LB; then sodium citrate was added (to 0.5 M), and samples were spread directly on LB-plates containing tetracyclin and 0.1 M sodium citrate. Purified transductants were checked on plates for sensitivity to UV-light (germicidal lamp, G30T8, 30 W; Sylvania, Switzerland) at a dose of 100 J/m^2. The absence of a lysogenic P1 phage was verified by checking sensitivity to P1 infection (if wild-type *recA* was supplied from plasmid pKY102) and normal transformation efficiency, as suggested by Schleif & Wensink (1981).

Table 1: E. coli *strains and plasmids used in this work*[a]

***E. coli* strains**	**Genotype**	**Relevant phenotype and application**	**Reference or origin**
AB1360-12	K12, F$^-$, *pheS12, thi-1, argE3, his-4, proA2, aroD6, lacY1, galK2, mtl-1, xyl-5, tsx-29, supE44?,* λ^-, *str*	*p*-F-Phe^R, Str^R	Hennecke & Böck (1975)
K10-F6	K10, Hfr(Cavalli), *pheS13, rel-1, tonA22, thi,* T2^R	*p*-F-Phe^R	Hennecke & Böck (1975)
RR28	K12, F$^-$, *pheS12, recA, thi-1, leu-6, proA2, hsdS20* ($r_B^-m_B^-$), *lacY1, galK2, ara-14, mtl-1, xyl-5, supE44, rpsL20 (str-20), endA,* λ^-	*p*-F-Phe^R, RecA$^-$, HsdR$^-$M$^-$, Str^R	Hennecke *et al.* (1982)
G1	K12, F$^-$, *pheS76, pps-4, metA90, argH1, proA44, icl-4, str-9, thi,* λ^-	Phe auxotrophic (PheRS with increased K_M for Phe), weakly ts, Str^R	Grüll *et al.* (1979)
NP37	K10, Hfr(Cavalli), *pheS5, rel-1, tonA22, thi,* T2^R	PheRS^{ts}	Eidlic & Neidhardt (1965); Böck & Neidhardt (1967); Comer & Böck (1976)
KA2	K10, Hfr(Cavalli), *pheS5, rel-1?, tonA22, thi,* T2^R Δ(*srlR-recA*)*306*::Tn*10*	PheRS^{ts}, RecA$^-$, Tet^R	this work; P1 transductant of NP37
HB101	K12, F$^-$, *recA13, thi-1, leu-6, proA2, hsdS20* ($r_B^-m_B^-$), *lacY1, galK2, ara-14, mtl-1, xyl-5, supE44, rpsL20 (str-20), endA,* λ^-	Str^R	Boyer & Roulland-Dussoix (1969); Sambrook *et al.* (1989)

JM101	K12, *supE, thi,* Δ(*lac-proAB*), F'[*traD36, proAB*$^+$, *lacI*^{q}, *lacZ*Δ*M15*]	host for single-strand DNA isolation	Yanisch-Perron *et al.* (1985)
JM109	K12, *recA*$^-$, *endA1, gyrA96, thi, hsdR17, supE44, relA*1, λ$^-$, Δ(*lac-proAB*), F'[*traD36, proAB*$^+$, *lacI*^{q}, *lacZ*Δ*M15*]	RecA$^-$, HsdR$^-$M$^+$, host for cloning	Yanisch-Perron *et al.* (1985)
TG1	K12, *hsd*Δ*5, thi, supE,* Δ(*lac-proAB*), F'[*traD36, proAB*$^+$, *lacI*^{q}, *lacZ*Δ*M15*]	HsdR$^-$M$^-$, host for single-strand DNA isolation	Sambrook *et al.* (1989); purchased from Amersham (UK)

Vectors	**Relevant properties**	**Reference or origin**
pACYC184	Cam^R, Tet^R; a derivative of plasmid P15A	Chang & Cohen (1978)
pBluescript KS (+) (=pBLS)	Amp^R; a derivative of pUC19 carrying a polylinker and the intergenic region of phage f1	Short *et al.* (1988); purchased from Stratagene (San Diego CA, USA)
pUC19	Amp^R; a derivative of pBR322 with a polylinker	Yanisch-Perron *et al.* (1985); purchased from Pharmacia (Uppsala, Sweden)

Plasmids	Relevant properties	Vector	Reference or origin
pHE3	Cam^R, *pheS* (wild type)	pACYC184	Hennecke *et al.* (1982)
pHWO	Amp^R, Tet^R; containing *E. coli thrS, infC, rpmI, rplT, pheS, pheT* on 10.25 kb *Eco*RI/*Hin*dIII fragment	pBR322	Elhardt *et al.* (1982); [=pB1 in Plumbridge *et al.* (1980)]
pKSC-X	Amp^R; containing *thrS, infC, rpmI, rplT, pheS*, pheT* on 10.25 kb *Eco*RI/*Hin*dIII fragment from *E. coli pheS* mutant strains	pUC19	this work (Fig. 1); '-X' stands for *pheS* genes from strains K10-F6 (-K), AB1360-12 (-A), G1 (-G) and wild type (-W, *ex* pHWO)
pKSB1-X	Amp^R; containing *pheS** on 1181 bp *Sma*I/*Hin*dII fragment from corresponding pKSC-X; orientation: *pheS** in phase with *lac* promoter	pBluescript KS (+)	this work (Fig. 2); '-X' stands for strain abbreviations (see pKSC-X)
pKSB2-X	as pKSB1-X, but *pheS** in opposite orientation (not expressed from *lac* promoter)	pBluescript KS (+)	this work (Fig. 2); '-X' stands for strain abbreviations (see pKSC-X)
pKSB1-M4S	derived from pKSB1-W; contains *p*-F-Phe^R *pheS* mutation S294	pBluescript KS (+)	this work; *in vitro* mutagenized pKSB1-W
pKSB1-D98	derived from pKSB1-K; contains $pheS^{ts}$	pBluescript KS (+)	this work; S294 replaced by wild type sequence

[a] An asterisk (*) indicates the presence of different *pheS* alleles. Gene symbols are according to Bachmann (1990). Amp, ampicillin; Cam, chloramphenicol; Str, streptomycin; Tet; tetracyclin.

Table 2: *Oligonucleotides used for sequencing, mutagenesis or PCR*

Name	Size	Sequence	Position in *pheS*[a]	Application
M13u (universal)	17mer	5'-GTAAAACGACGGCCAGT-3'	anneals to vector[b]	sequencing with [α-^{32}P]dATP
PEKA7	15mer	5'-CGAACCAGTGTCACC-3'	-36 to -22	sequencing with [α-^{32}P]dATP
PEKA8	15mer	5'-GGCGCTGAATGCGGC-3'	207 to 221	sequencing with [α-^{32}P]dATP
PEKA8N	18mer	5'-CTGAATGCGCGTAAAGCG-3'	211 to 228	sequencing with [α-^{32}P]dATP
PEKA9	15mer	5'-CGCGCGCTGACCACG-3'	452 to 466	sequencing with [α-^{32}P]dATP
PEKA10	15mer	5'-GCAGATTCGCTTCCG-3'	717 to 731	sequencing with [α-^{32}P]dATP
PEKA11	16mer	5'-GCGTTTCCTCAAACAG-3'	960 to 975	sequencing with [α-^{32}P]dATP
-36M13 (dye)[c]	18mer	5'-TCCCAGTCACGACGTTGT-3'	anneals to vector[b]	automated sequencing
PEKA2S (dye)[c]	20mer	5'-CGCGTAACCACAGTTCACTG-3'	1026 to 1007	automated sequencing
PEKA3S (dye)[c]	20mer	5'-GACGTCCACTTCTGCAGAAG-3'	777 to 758	automated sequencing
PEKA4S (dye)[c]	20mer	5'-GCGGATCTGTACGCCAGAGG-3'	528 to 509	automated sequencing
PEKA5S (dye)[c]	20mer	5'-CGTTTCCGCCGCCAGACGCG-3'	273 to 254	automated sequencing
PEKA6S (dye)[c]	20mer	5'-GTTCTGCGAGATGTGACATG-3'	19 to -1	automated sequencing
MA-S	20mer	5'-CTCTGGTTTCTCCTTCGGGA-3'[d]	870 to 889	site-directed mutagenesis
PSCA	22mer	5'-GGGATAGGCTCTAAGTCCAACG-3'	-56 to -35	polymerase chain reaction
PSCZ	22mer	5'-TCGCCGGGTTCACCCATTCGCG-3'	1044 to 1023	polymerase chain reaction

[a] Numbering of nucleotides refers to the A in the start codon of the *pheS* reading frame (see Figs. 3 and 5; Fayat *et al.*, 1983).

[b] Universal M13 primers anneal 5' of polylinker to pBLS: M13u from 50 to 34 and -36M13 from 66 to 49 bases distant to the *Sac*I site.

[c] Each name represents a set of 4 dye-primers having different fluorophores attached to their 5'-ends (Smith *et al.*, 1986, Connell *et al.*, 1987).

[d] The underlined nucleotide indicates the point mutation introduced.

3.3. *In vivo* plate tests

In vivo plate tests for thermoresistance (*e. g.* complementation of strain KA2) were performed by streaking out single colonies on LB-agar plates containing appropriate antibiotics. Growth was recorded after 24 and 48 h of incubation at 30°C or 40°C. When full expression of plasmid genes from the *lac* promoter was required, IPTG was added to 0.5 mM final concentration. Plate tests for the analysis of Phe-analogue resistance were done on minimal medium plates containing 100 mM KH_2PO_4, 50 mM Na_2HPO_4, 12 mM $(NH_4)_2SO_4$, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$ and 0.4 % D-glucose in 1.5 % agar. The medium was supplemented with thiamine-HCl (5 µg/ml) and, when required, with amino acids (50 µg/ml), other vitamins (5 µg/ml), ampicillin (150 µg/ml) and IPTG (0.25 mM). D,L-*p*-F-Phe (Sigma Chemicals, St. Louis MO, USA) was present at 200 µg/ml. Growth at 28°C or 37°C was monitored over a period of 4 days. All tests were performed in duplicate and repeated at least once.

3.4. Recombinant DNA techniques

Restriction endonucleases and DNA modifying enzymes were purchased from Boehringer (Mannheim, Germany) or Pharmacia (Uppsala, Sweden) and used as recommended by the supplier. [α-^{32}P]dATP (3000 Ci/mmol) was from Amersham (UK). Cloning, transformation, plasmid isolation and nick-translation were performed using standard procedures (Maniatis *et al.*, 1982; Sambrook *et al.*, 1989). For Southern-blot analyses (Southern, 1975), restriction enzyme-digested chromosomal DNA was transferred onto Hybond-N membranes (Amersham, UK) and crosslinked by UV light for 3 min. The radiolabelled probe was nick-translated plasmid pHE3; the vector alone (pACYC184) served as control. Blotting, hybridization and washing conditions were essentially as described by Maniatis *et al.* (1982). Oligonucleotides were purified on a 12 % polyacrylamide gel as described (Sambrook *et al.*, 1989). Dye-primers were further chromatographed on C-18 Sep-Pak cartridges (Waters Associates, Milford MA, USA), whereas Sephadex-G25 columns (NAP-10; Pharmacia, Uppsala, Sweden) were used for the other oligonucleotides. Oligonucleotide-directed mutagenesis was carried out according to the *in vitro* method of Nakamaye & Eckstein (1986) using the kit and detailed protocol from Amersham (UK). The mutagenesis strategy is described in detail in the accompanying paper (Kast *et al.*, 1991).

3.5. Cloning of mutant *pheS* genes

Chromosomal DNA was isolated from three mutant *E. coli* strains according to a modification (additional proteinase K treatment) of the method of Schleif & Wensink (1981) and then digested with *Eco*RI and *Hin*dIII. The presence of the 10.25 kb *Eco*RI/*Hin*dIII fragment carrying *pheS* (see Fig. 1(a)) was verified by Southern blot hybridization with the radiolabelled *pheS* plasmid pHE3. DNA fragments from the 10 kb region were cut out from a preparative agarose gel and ligated to a pUC19 vector digested with *Eco*RI and *Hin*dIII. After transformation of the thermosensitive *E. coli* strain NP37 (*pheS*^{ts}), selection at high temperature (40°C) on ampicillin plates allowed only growth of cells containing *pheS* on plasmids (plasmids pKSC-X; see Table 1).

3.6. DNA sequence analysis

The *pheS* DNA was sequenced on both strands with the dideoxy chain termination method of Sanger *et al.*, (1977). A set of specific primers (Table 2) allowed all mutant genes to be sequenced in parallel to the one of the wild type as shown later in Fig. 3. Single-strand DNA of Bluescript plasmids (pKSB1/2) in strain TG1 or JM101 was isolated using a scale-up of the method descibed by Stratagene (San Diego CA, USA) which is based upon superinfection with a helper phage (M13-VCS).

The sequencing reactions were carried out with the Klenow fragment of DNA polymerase I (Pharmacia, Uppsala, Sweden) or Sequenase^{TM} (a modified T7 DNA polymerase; Tabor & Richardson, 1987; United States Biochemical Corp., Cleveland OH, USA). 2'-deoxy-GTP was generally replaced by 7-deaza-2'-deoxy-GTP (Boehringer, Mannheim, Germany); the other nucleotides were purchased from Pharmacia (Uppsala, Sweden). The *pheS* coding strand was sequenced with [α-^{32}P]dATP (3000 Ci/mmol; Amersham, UK) by the conventional methods described in Sambrook *et al.*, (1989), which were improved by the method of uniform labelling by Tsang & Bentley (1988). For the opposite strand, *pheS*-specific 5'-end-labelled dye-primers were used that allowed automatized sequencing (Smith *et al.*, 1986, Connell *et al.*, 1987) with the DNA Sequencer model 370A, following protocols provided by the manufacturer (Applied Biosystems, Foster City CA, USA). For double-stranded DNA sequencing, 1 pmol of purified plasmid DNA was denatured and subsequently sequenced with [α-^{32}P]dATP and Sequenase^{TM} (Zhang *et al.*,

1988). Computer-assisted sequence alignments were done using programs of the University of Wisconsin Genetics Computer Group (UWGCG; sequence analysis software package release 6.2; Madison WI, USA).

3.7. Polymerase chain reaction (PCR)

PCR (Saiki *et al.*, 1988) was performed in a DNA Thermal Cycler (Perkin-Elmer Cetus, Norwalk CT, USA). The reactions were carried out asymmetrically (Gyllensten & Erlich, 1988; Shyamala & Ames, 1989) by using different concentrations of the primers flanking the complete *pheS* gene (Table 2; see later in Fig. 6). The reaction mixture contained 50 ng of bacterial chromosomal DNA as template, 50 mM KCl, 10 mM Tris-HCl, pH 8.5, 2 mM $MgCl_2$, 100 µg/ml gelatine and 200 µM each of dATP, dCTP, dGTP and dTTP in a volume of 50 µl. The primers were present at 500 nM (primer 'PSCZ') or 10 nM (primer 'PSCA'). After denaturation at 95°C for 10 min, 2 U of *Taq* DNA Polymerase (Perkin-Elmer Cetus, Norwalk CT, USA) and then 70 µl mineral oil (Sigma Chemicals, St Louis MO, USA) were added. Amplification was performed by 35 cycles of denaturation (1 min at 94°C), primer annealing (2 min at 67°C) and primer extension (1 min at 72°C), the last extension step lasting 10 min. The PCR products were analysed on conventional 1.5 % ethidium bromide-stained agarose gels (Sambrook *et al.*, 1989) and purified from excess primers and nucleotides by ultrafiltration employing a Centricon-30 microconcentrator (Amicon, Danvers MA, USA). Sequencing of the generated single-stranded DNA was accomplished with a third primer ('PEKA8N') using Sequenase™ as described above (see also Figs. 6 and 7).

3.8. Preparation of protein extracts

The procedure used was essentially that described by Grüll *et al.* (1979). A bacterial culture was grown at 30°C by vigorous shaking to an optical density (at 600 nm wavelength) of 2.0 in 300 ml rich medium (1 % bactotryptone, 0.5 % yeast extract, 0.2 % D-glucose and appropriate antibiotics; 0.5 mM IPTG was present if plasmid-encoded *pheS* had to be expressed). The cells were washed in 125 ml PRS buffer (0°C; 20 mM Tris-HCl, pH 7.5, 0.2 mM EDTA, 6 mM β-mercaptoethanol, 30 mM NH_4Cl, 10 mM $MgCl_2$), centrifuged again (5'000 x g, 20 min, 4°C) and resuspended in 8 ml PRS buffer containing 1 µg/ml DNAse I (Boehringer, Mannheim, Germany). Then, the cells were disrupted by two passages through a

French pressure cell (Aminco, Silver Spring MD, USA) at 120 MPa (0°C). The suspension was centrifuged at 12'000 x g for 20 min (4°C). Afterwards, the supernatant was freed from ribosomes by ultracentrifugation at 140'000 x g for 2 h (2°C, SW 55Ti rotor, Beckman Instruments, Palo Alto CA, USA). The resulting S140 extract was adjusted to 50 % glycerol and stored at -20°C. Protein concentration was determined in 4 parallel samples by a modification of the Folin phenol procedure (Lowry *et al.*, 1951) with bovine serum albumin (A-7906; Sigma Chemicals, St Louis MO, USA) and a reference extract as standards. The assays were repeated several times and the calculated values (hyperbolic calibration curve; Peterson, 1983) typically had a standard deviation of 5 % or less.

3.9. Determination of PheRS activity *in vitro*

The PheRS assay was based on the methods given by Kosakowski & Böck (1970) and Comer & Böck (1976). Enzymatic activity was measured by attachment of L-[^{14}C]Phe to tRNA. The reaction mixture contained (in a total volume of 250 µl) 100 mM Tris-HCl, pH 7.5, 10 mM KCl, 10 mM $MgCl_2$, 2 mM reduced glutathione, 2 mM adenosine triphosphate, 2.4 mg/ml *E. coli* tRNA (from strain MRE600; Boehringer, Mannheim, Germany) and 20 µM L-phenyl-[1-^{14}C]alanine (20 mCi/mmol; Amersham, UK). After addition of S140 protein extract (usually 4-8 µg, suitably diluted in 20 mM Tris-HCl, pH 7.5, 0.2 mM EDTA, 6 mM β-mercaptoethanol, 50 µg/ml bovine serum albumin), the mixtures were incubated at 28°C for various time intervals. The reaction was then stopped by chilling and precipitation of tRNA with 2 ml 10 % trichloroacetic acid containing 1 mg/ml D,L-Phe (0°C). The precipitates were collected on glass-fiber filters (GF/C, Whatman International, Maidstone, UK) and washed with 2 x 5 ml of the trichloroacetic acid solution and 5 ml 70 % ethanol. Radioactivity on filters was determined by liquid scintillation counting (Packard Tricarb 2200; Packard Instrument Company, Downers Grove IL, USA). For incorporation of *p*-F-Phe, 40 µM *p*-fluoro-D,L-[U-^{14}C]phenylalanine (4 mCi/mmol; CIS, Gif-sur-Yvette, France) was used instead of L-[^{14}C]Phe. Initial velocities were determined graphically from at least 5 independent data points taken from 0 to 10 min; background slopes (without enzyme) were subtracted. All test series were repeated at least once, the deviations of initial velocities usually being less than 5 %.

4. Results

4.1. Cloning of mutant *pheS* genes

To determine the mutations responsible for the altered substrate binding phenotypes of PheRS, the mutant *pheS* gene regions (10.25 kb *Eco*RI/*Hin*dIII fragments, Fig. 1(a); Fayat *et al.*, 1983) were cloned from chromosomal DNA isolated from the following strains: the two spontaneously obtained *p*-F-Phe^R strains AB1360-12 and K10-F6, and the ethylmethanesulfonate-mutagenized Phe auxotroph G1 (see Table 1). The cloning was achieved by complementation of the *pheS* mutant strain NP37 (thermosensitive PheRS) as described in Materials and Methods. All of 7 analyzed candidate clones contained plasmids of the expected size, and their restriction maps corresponded to that described by Fayat *et al.* (1983). The plasmids were named pKSC-A, pKSC-K and pKSC-G containing the fragments from strains AB1360-12, K10-F6 and G1, respectively. pKSC-W was constructed by cloning the corresponding wild-type *Eco*RI/*Hin*dIII fragment from plasmid pHWO (Table 1) into the vector pUC19 in analogy to the procedure for the mutant genes.

4.2. Identification of point mutations in the cloned *pheS* genes

As shown in Fig. 2, the *pheS*-containing *Sma*I/*Hin*dII fragments from the pKSC plasmids were subcloned in both orientations into the sequencing and expression vector pBluescript KS (+). The pKSB**1**-A, -K, -G, and -W series (using similar plasmid designations as above) allowed the production of single-strand DNA from the coding strand. In addition, expression of the corresponding *pheS* genes from the *lac* promoter was possible. With the other series of plasmids (pKSB**2**-A, -K, -G and -W), the non-coding strand could be obtained in single-stranded form (Fig. 2). The sequencing strategy described in Fig. 3 was based on a set of *pheS*-specific oligonucleotide primers (listed in Table 2) that were positioned at distances of about 250 bp from each other. With this strategy, all 3 mutant genes could be sequenced parallel to the wild-type *pheS* gene on both strands. Examples of identified point mutations with the two applied sequencing methods are shown in Fig. 4, and an overview of all nucleotide exchanges relative

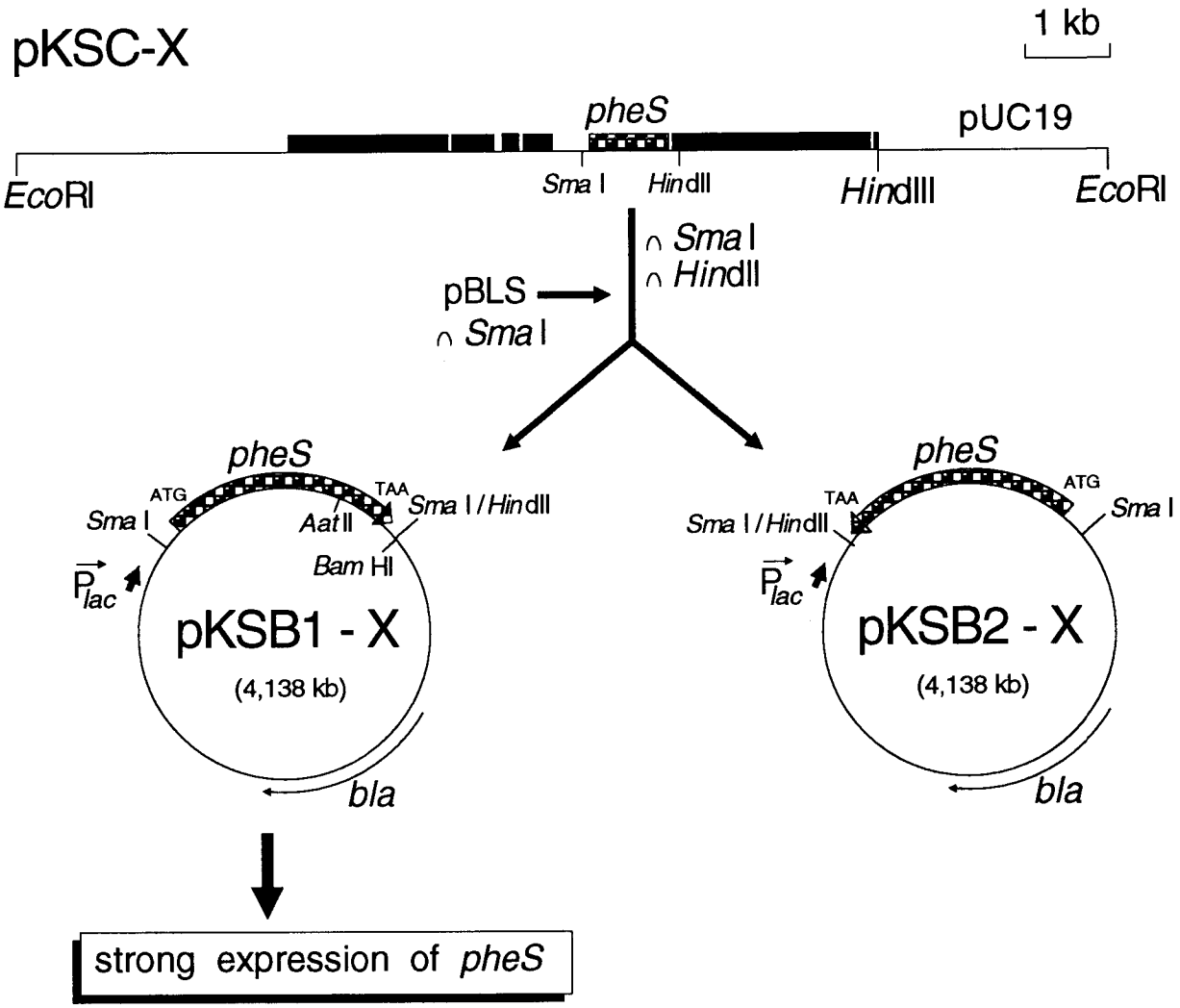


Figure 2. Subcloning of *pheS* genes. The pKSC-X plasmids obtained after cloning of the chromosomal *pheST* region were digested with *Sma*I and *Hin*dII and the 1181 bp blunt-end *pheS* fragment was isolated and ligated to a *Sma*I-linearized pBluescript KS (+) vector (pBLS). Both insert orientations (pKSB1-X and pKSB2-X) were required to allow the isolation of both *pheS* single strands. The pKSB1-X plasmid type could be utilised for high expression of *pheS* from the *lac* promoter (P_{lac}) when fully induced with IPTG. *Aat*II and *Bam*HI restriction sites shown in pKSB1-X were used for further clonings (see text). The orientation of the *pheS* genes is indicated by start (ATG) and stop (TAA) codons. *bla*, gene for β-lactamase conferring ampicillin resistance. The '-X' needs to be replaced by other letters to indicate the source of the *pheS* gene, as follows: A, AB1360-12; K, K10-F6; G, G1; W, wild type).

to the published wild-type sequence is presented in Fig. 5.

At codon position 74, an arginine triplett was found instead of the one for alanine published by Fayat *et al.* (1983). The fact that this deviation was present in all 4 independent genes isolated from different strains suggests, that the former ala-

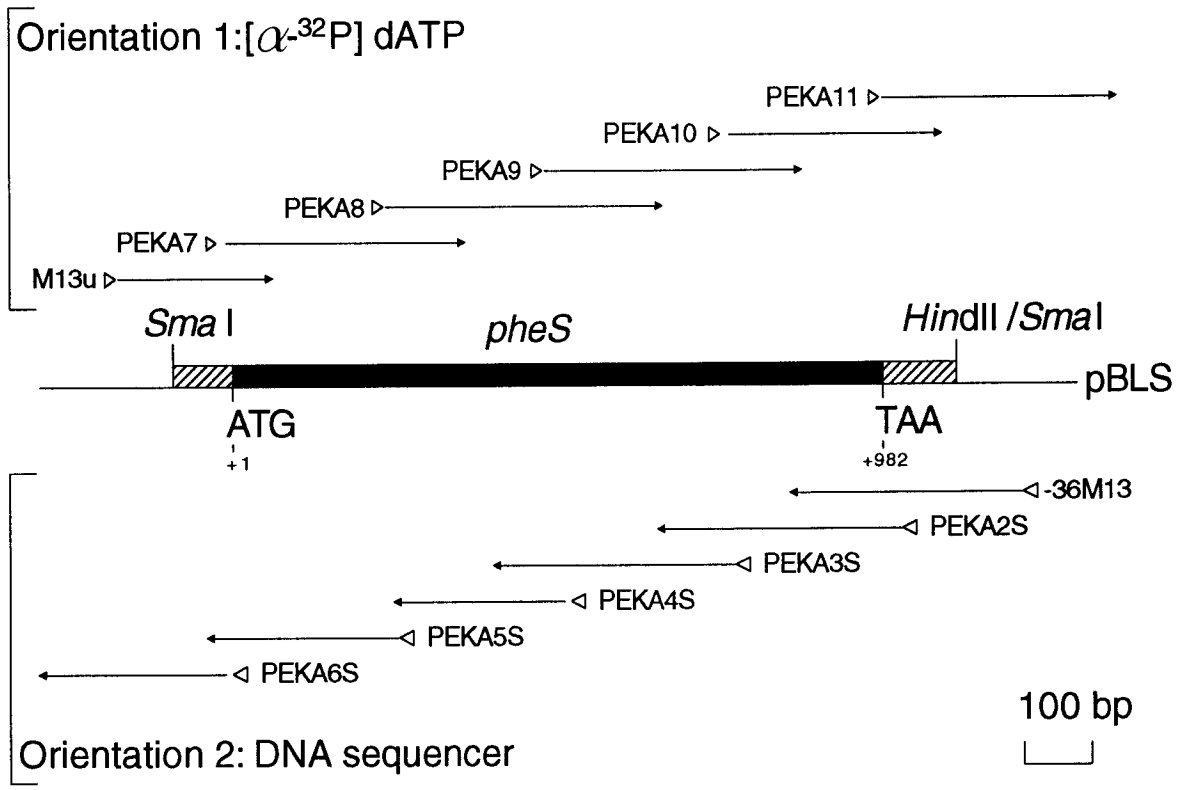
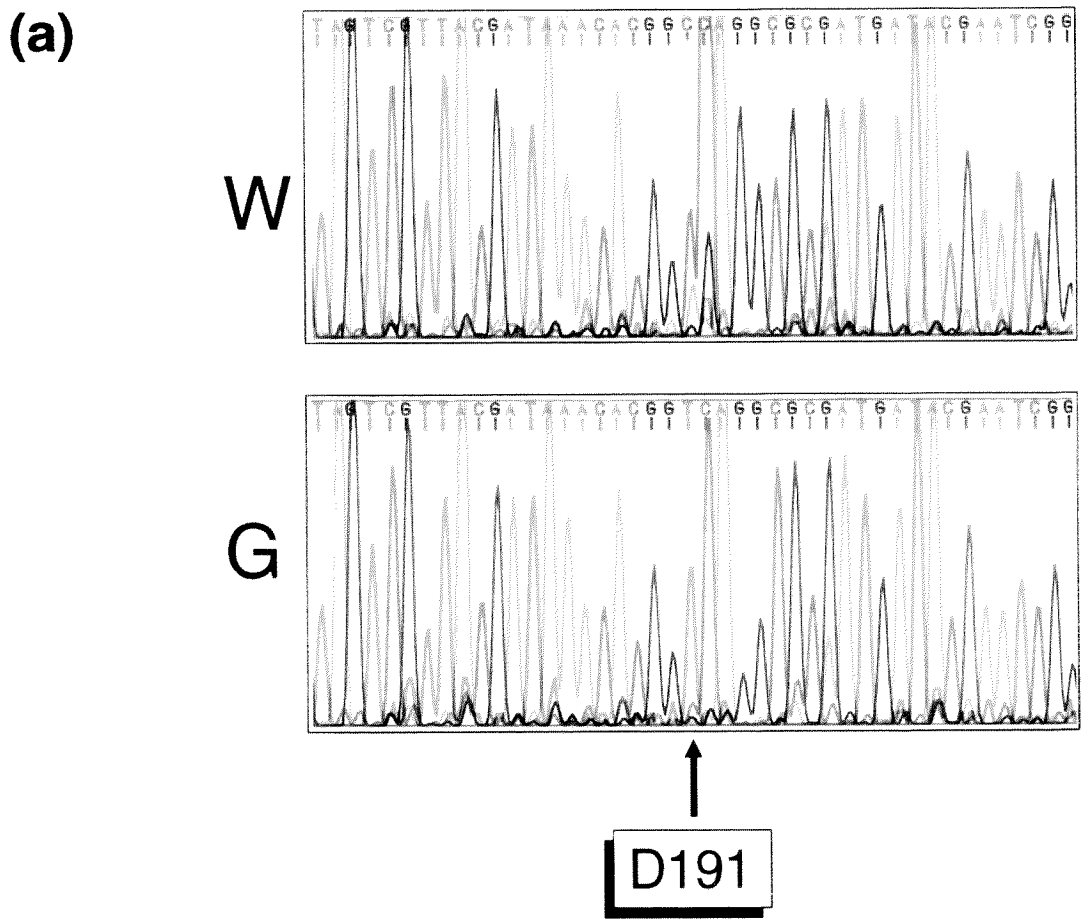
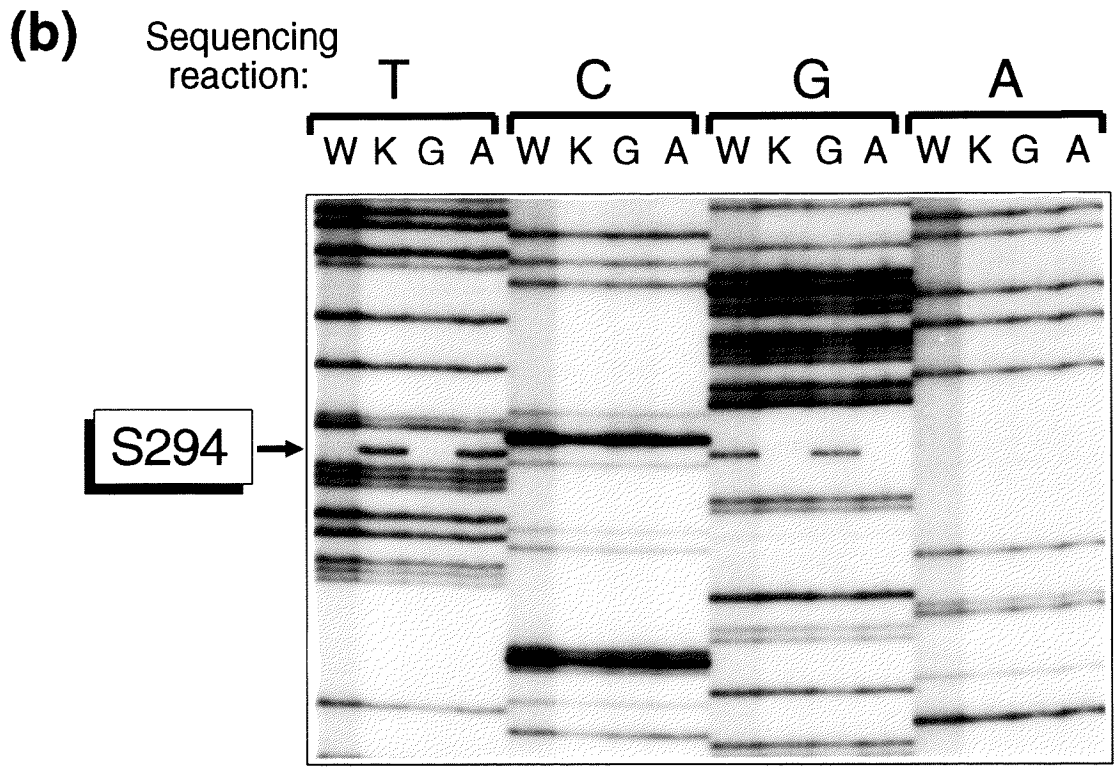


Figure 3. Sequencing strategy. Oligonucleotides (open triangles; listed in Table 2) were synthesized according to the available *pheS* wild-type sequence. When used as primers, they cover, at about 250 bp apart from one another, the complete gene sequence on both strands. This allowed sequencing of mutant and wild-type genes in parallel. The range of each readable sequence is indicated. The plasmids used are described in Fig. 2 and Table 1. Single-strand DNA from pKSB1-X (orientation1) was sequenced with radiolabelled nucleotides, pKSB2-X (orientation 2) with fluorescent primers and automated detection. The filled and hatched bars indicate the coding and the non-coding regions of *pheS*, respectively.

Figure 4. Examples of point mutations. (a) A plot of the sequence around nucleotide 572 determined by automated sequencing. The individual bases are indicated by differentially colored peaks (T, red; C, blue; G, black and A, green). The bottom panel (G: *pheS* from strain G1) shows a T at position 572 (non-coding strand) as opposed to the C in the wild type (W; upper panel). The corresponding exchange in the coding strand was a G to A transition representing mutation D191 which is responsible for the increase of the K_M for Phe in PheRS of strain G1. **(b)** Autoradiogram showing the sequences around nucleotide position 880. All sequencing reactions for the same base specificity were loaded in adjacent lanes in order to directly compare wild-type (W), K10-F6 (K), G1 (G) and AB1360-12 (A) *pheS* sequencing patterns. Mutation S294, consisting of a G to T exchange, is clearly visible in *pheS* from the two *p*-F-Phe^R strains (K and A).


(a)
W
G
D191


(b)
Sequencing reaction:
T
C
G
A
W K G A W K G A W K G A W K G A
S294

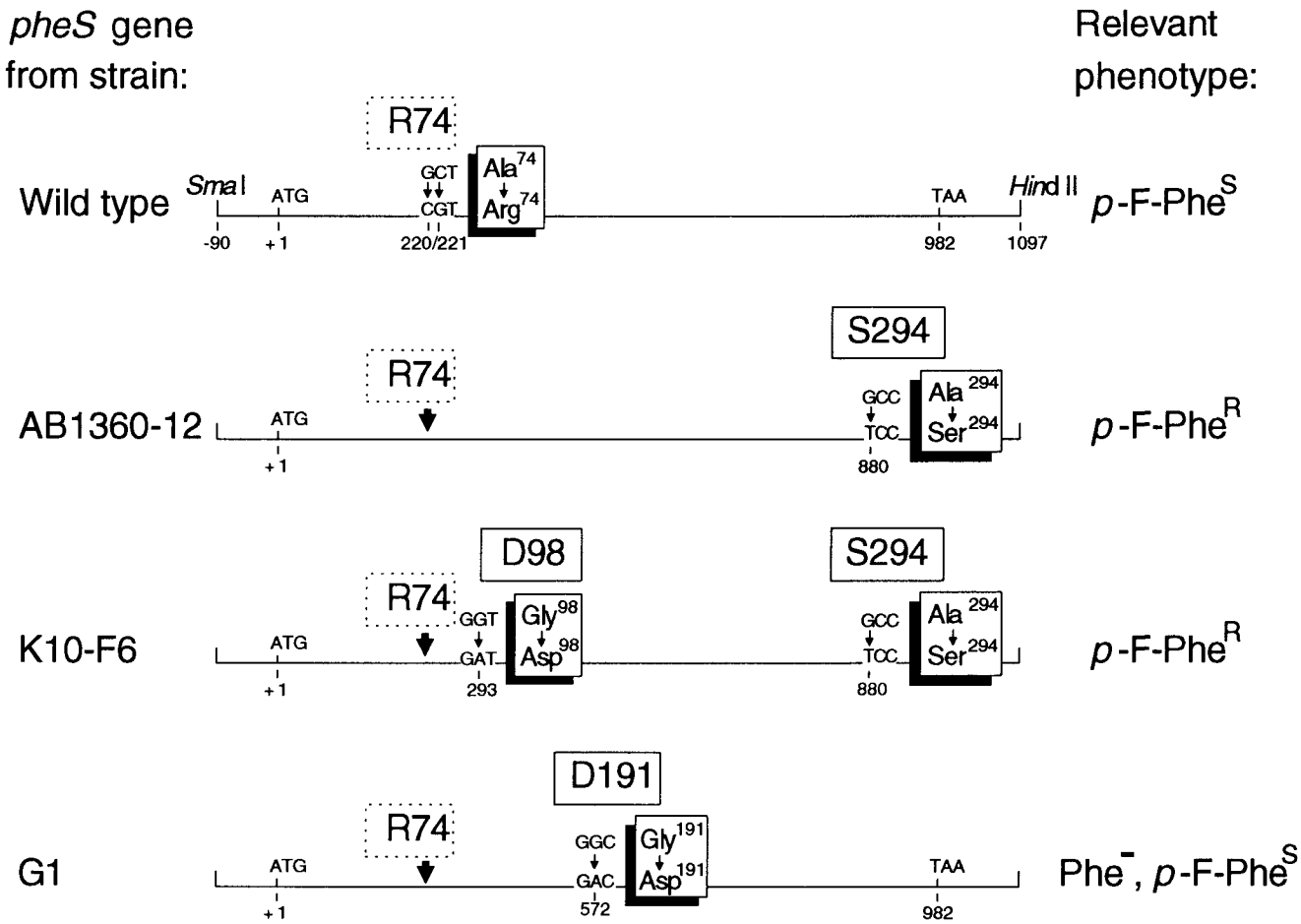


Figure 5. Overview of identified point mutations. Numbers below the gene indicate nucleotide positions relative to the A in the start codon. Numbers associated with amino acids indicate their position in the α subunit of PheRS. R74 was found in all 4 sequenced genes indicating an error in the previously published *pheS* sequence. The other mutations (boxed with solid line) were real deviations from the wild-type sequence. Phe$^-$ indicates the Phe auxotrophic phenotype of strain G1.

nine codon was due to an error in the previous sequence determination (C/G-switch at nucleotide positions 220/221). The Arg fits better into the sequence alignment with the homologous PheRS small subunits from yeast cytoplasm, yeast mitochondria and *Bacillus subtilis*, in which either Arg or another basic amino acid (Lys) is present at that position (Brakhage *et al.*, 1990). Although the nucleotide switch caused mispairings of the two 3'-terminal bases of the designed oligonucleotide PEKA8, we were nevertheless able to use it as a sequencing primer for the Klenow enzyme, albeit at lower efficiency. This demonstrated that a perfect match at the 3'-end of the oligonucleotide (where primer extension starts) is not required for sequencing.

The *pheS* mutation of strain G1 (Fig. 4(a); mutation D191), consisted of a G to A transition (at nucleotide 572) leading to a glycine-to-aspartic acid exchange at position 191 in the α subunit. It causes the Phe auxotrophic phenotype by increasing the K_M for Phe in PheRS (Grüll *et al.*, 1979). We did observe an impaired growth of strain RR28 containing pKSB1-G on minimal medium (+IPTG), but only at high temperature. pKSB2-G or plasmids with wild-type *pheS* showed no growth interference (data not shown). In addition, *pheS* from G1 could not complement the thermosensitive strain KA2 at 37°C or 40°C, unless *pheT* was also present on the plasmid. These data confirmed that pKSB1-G expressed an α subunit that was inefficient and/or thermosensitive; both phenomena were demonstrated previously by Grüll *et al.* (1979) for G1 PheRS *in vitro* and *in vivo*. We did not further examine this mutation for reasons discussed below.

The mutation found in both *pheS* genes from the independently isolated *p*-F-Phe resistant strains AB1360-12 and K10-F6 consisted of a G to T transversion (nucleotide 880; Fig. 4(b)), leading to an alanine-to-serine exchange at amino acid position 294. This mutation, termed 'S294', which resided at the C-terminus of the α subunit (and thus quite distant of D191), was the only deviation from the wild-type sequence in *pheS* from AB1360-12. This indicated that it was exclusively responsible for the observed *p*-F-Phe resistance phenotype. However, sequence analysis of pKSB1-K and pKSB2-K showed an additional mutation at codon position 98: the nature and the consequences of this Gly-to-Asp replacement (D98; G to A transition at nucleotide 293) will be treated in sections 4.3. and 4.4.

A change in the affinity for a substrate can be caused by any distortion of the three-dimensional structure due to alterations that are distant to the substrate binding site. However, the mutation leading to an exclusion of the substrate analogue *p*-F-Phe from the enzymatic reaction was expected to be rather specific and to point more directly to the Phe binding site. Therefore, we decided to concentrate our investigations on position 294.

4.3. The Ala-to-Ser exchange at position 294 is responsible for the *p*-F-Phe^R phenotype *in vivo* and *in vitro*

To prove unambiguously that S294 is responsible for the *p*-F-Phe^R phenotype, a series of *in vivo* and *in vitro* tests were carried out to compare the parental strains with strains carrying the cloned *pheS* genes on plasmids. In addition, the S294 mutation was created again in a wild-type gene by oligonucleotide-directed

mutagenesis yielding plasmid pKSB1-M4S (the mutagenesis is described in detail in the accompanying paper by Kast *et al.*, 1991).

Initial attempts to assay the *pheS* plasmids in the NP37 host (PheRS^{ts}) were not successful because of a high background due to recombination with the chromosomal gene. Therefore, strain KA2, a RecA$^-$ variant of NP37, was constructed by generalized P1 transduction. The pKSB1 plasmids were transformed into KA2, in which the *pheS* genes were fully expressed by addition of IPTG, and the PheRS phenotype was assayed by *in vivo* growth tests (on agar plates) as well as *in vitro* (protein extraction and PheRS assay) as described in Materials and Methods. The results are compiled in Table 3, together with those obtained with the parental strains. In addition, Table 3 contains data from the pKSB1 plasmids in the *p*-F-Phe^R strain RR28. In this background, the presence of a wild-type *pheS* allele causes the *p*-F-Phe^R strain to become sensitive to the Phe-analogue. This dominance of the wild-type gene is due to deleterious *p*-F-Phe incorporation into cellular protein (Fangman & Neidhardt, 1964; Hennecke *et al.*, 1982, and references therein). Growth failure of transformed RR28 on *p*-F-Phe-containing plates, therefore, is an indication for a wild-type *pheS* gene present on the plasmid. (Use of this feature was made previously in the development of a cloning system allowing for a direct selection of inserted DNA in the vector pHE3; Hennecke *et al.*, 1982).

The results in Table 3 showed that strains containing plasmid pKSB1-A behaved like the *p*-F-Phe^R parental strains AB1360-12 and K10-F6 on Phe-analogue plates (*i. e.* growth) and in the PheRS assay (*i. e.* no significant analogue incorporation). Identical results were obtained with plasmid pKSB1-M4S, which proved that the mutation S294 was solely responsible for the *p*-F-Phe^R phenotype. The plasmid pKSB1-W, as expected, conferred *p*-F-Phe sensitivity to RR28, and the PheRS values obtained *in vitro* agreed with those of HB101, the wild-type reference strain. These data demonstrated that the test system for the plasmidial genes could simulate the chromosomal situation: although the α subunit was overproduced, the PheRS activities showed chromosomal levels due to a limiting amount of β subunit that was apparently expressed normally from the chromosomal *pheT* gene. This was confirmed in the accompanying paper (Kast *et al.*, 1991), and will be discussed there in more detail. It is worth mentioning here that the PheRS^{ts} strain KA2 showed virtually no background activity *in vitro* at 28°C, the permissive temperature for growth (Table 3). This had already been observed previously for NP37 (Eidlic & Neidhardt, 1965) and was ascribed to decreased stability of the mutant enzyme under *in vitro* conditions (Neidhardt, 1966; Böck, 1968).

Table 3: *PheRS activity of* pheS *mutants tested* in vivo *and* in vitro[a]

A

Strains carrying chromosomal *pheS* only	*In vivo* growth tests[b]				PheRS activity[d]	
	Rich medium (LB)		MM[c], 28°C			
	30°C	40°C	− *p*-F-Phe	+ *p*-F-Phe	[^{14}C]Phe	*p*-F-[^{14}C]Phe
HB101 (wild type)	++	++	+[c]	0	3.08	1.53
AB1360-12 (*p*-F-Phe^R)	++	±[c]	++	++[c]	2.54	≤0.03
K10-F6 (*p*-F-Phe^R)	++	++	++	+	2.48	≤0.09
RR28 (*p*-F-Phe^R)	++	++	++	+	n.t.[e]	n.t.[e]
KA2 (PheRS^{ts})	++	0	n.t.[e]	n.t.[e]	≤0.10	≤0.08

B

Plasmids carrying *pheS* (expected phenotype)	*In vivo* growth tests[b]				PheRS activity in plasmid-bearing KA2 strain[d]	
	in strain KA2 (PheRS^{ts}) LB, Amp, IPTG		in strain RR28 (*p*-F-Phe^R, recessive) MM[c], Amp, IPTG; 37°C			
	30°C	40°C	− *p*-F-Phe	+ *p*-F-Phe	[^{14}C]Phe	*p*-F-[^{14}C]Phe
pBLS (vector)	++	0	++	+	≤0.02	0.00
pKSB1-W (wild type)	++	++	++	−	3.20	1.71
pKSB1-A (*p*-F-Phe^R)	++	++	++	+	2.48	≤0.09
pKSB1-M4S (*p*-F-Phe^R)	++	++	++	+	2.49	≤0.13
pKSB1-K (*p*-F-Phe^R)	++	0	++	±	≤0.01	≤0.07

[a] All tests were carried out as described in Materials and Methods.

[b] Growth of single colonies is specified from good growth to no growth at all, in the order: ++ > + > ± > − > 0. Amp, ampicillin.

[c] The minimal medium (MM) of each strain was supplemented according to its requirements (Table 1). Individual strains may behave differently *in vivo* due to their dissimilar genetic backgrounds.

[d] Activity is in nmol Phe or *p*-F-Phe attached to tRNA per min and mg protein in S140 extracts at 28°C.

[e] n.t.: not tested.

Whilst the wild-type and mutant S294 *pheS* genes could complement KA2, this was surprisingly not possible with plasmid pKSB1-K carrying *pheS* from K10-F6, and neither Phe nor *p*-F-Phe incorporation were detected *in vitro* (Table 3). The presence of this gene (supposed to be *p*-F-Phe^R) on a high copy number vector did impair, but not abolish, growth of RR28 on *p*-F-Phe plates (Table 3). It appeared reasonable to attribute these phenotypes to the mutation D98, which distinguished pKSB1-K from pKSB1-A. But the striking discrepancy between the phenotypes of pKSB1-K and its parent strain K10-F6 still requested further clarification. Was D98 a secondary mutation picked up during the cloning process?

4.4. The Gly-to-Asp exchange at position 98 of the PheRS α subunit is responsible for the thermosensitivity of strain NP37

As mentioned, the mutation D98 was obviously the cause for the lack of complementing ability of pKSB1-K. To prove this, the D98 mutation was separated from mutation S294 by replacing in pKSB1-K a 326 bp *Aat*II/*Bam*HI fragment (Fig. 2) carrying S294 by the corresponding fragment from the wild-type plasmid pKSB1-W. Unlike pKSB1-K, the resulting plasmid pKSB1-D98 produced a *p*-F-Phe^S phenotype by preventing any growth of transformed RR28 on *p*-F-Phe plates at 30°C (data not shown). This demonstrated (i) that removal of mutation S294 abolished the *p*-F-Phe^R trait and (ii) that a *pheS* gene carrying mutation D98 encodes a PheRS α subunit which is active at low temperature *in vivo*. pKSB1-D98 failed to complement KA2 at 40°C which, together with the above observations, proved that D98 alone is responsible for a thermosensitive and *p*-F-Phe^S α subunit.

The inability of pKSB1-K to complement a PheRS^{ts} strain apparently contradicted the fact that the corresponding *pheS* gene could be cloned by complementation of NP37. We therefore analyzed again the parental plasmid pKSC-K(1) and, in parallel, a similar plasmid pKSC-K(2) from an independent clone obtained during cloning of K10-F6 *pheS*. Both plasmids could complement NP37, but NP37 transformed by pKSC-K(1) showed a slightly reduced growth at 40°C. Direct double-stranded sequencing around codon position 98 (using primer PEKA4S) showed indeed that the parental plasmid of pKSB1/2-K contained mutation D98 which was not found in pKSC-K(2). On one hand, this result proved that D98 within pKSC-K(1) did allow complementation of a PheRS^{ts} strain, probably due to the simultaneous presence of a *pheT* gene on the plasmid leading to large amounts of

(less efficient) PheRS (gene dosage effect; see Discussion). On the other hand, it was now conceivable that D98 did not originate from strain K10-F6 but was acquired later. This would explain the phenotypic differences between K10-F6 and KA2 when transformed with pKSB1-K (Table 3).

Two features associated with the *pheS* gene from pKSB1-K led us to suspect that D98 was the *pheS* mutation of NP37: (i) D98 caused a thermosensitive PheRS, like that of NP37, and (ii) the mutant *pheS* gene was propagated in the $RecA^+$ host NP37 during cloning. To verify this hypothesis, the region around codon 98 was sequenced directly after PCR-mediated amplification of the chromosomal *pheS* genes from strains NP37, K10-F6, AB1360-12 and HB101. The strategy used is displayed in Fig. 6. The polymerase chain reactions were carried out asymmetrically, and the sequencing of the purified single-stranded products was done with a third primer, by means of which only correctly amplified DNA was sequenced (Wrischnik *et al.*, 1987).

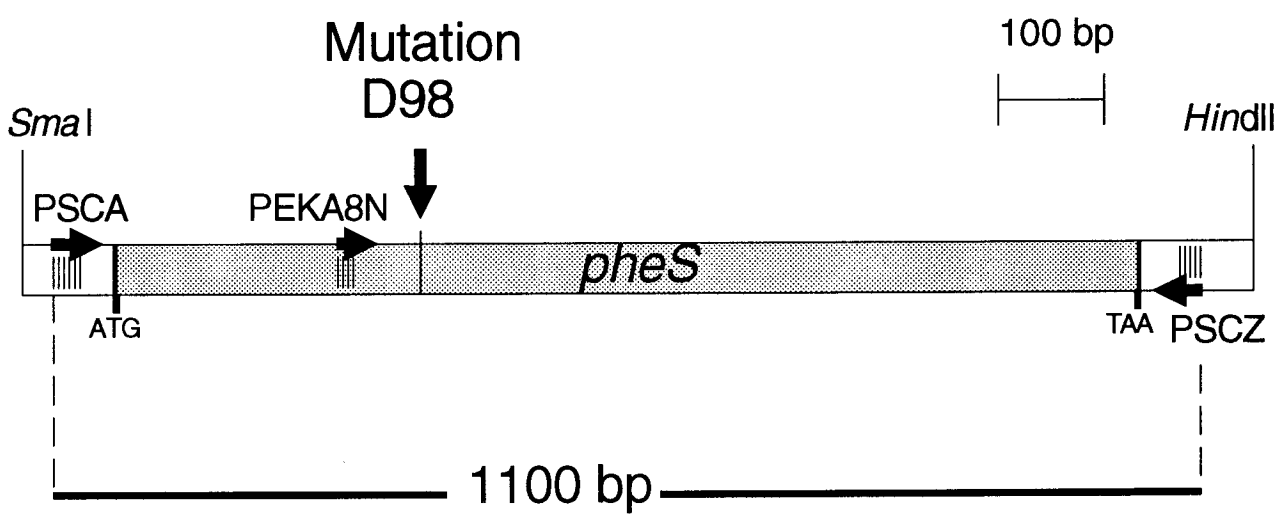


Figure 6. Strategy for the direct sequencing of the region around *pheS* codon 98 using PCR-amplified chromosomal DNA. Two 22mer oligonucleotide primers (PSCA and PSCZ; Table 2) were designed to flank the coding region of *pheS*. PSCZ was present in a 50-fold excess in the PCR reaction, leading to production of single-strand DNA of 1100 nucleotides, which, after depletion of PSCA, appeared in addition to the 1100 bp double-strand fragment. The single-stranded PCR products could be sequenced with a third primer (PEKA8N) located close to the position of mutation D98.

The results (Fig. 7) showed that strain NP37 indeed carries mutation D98, which is obviously responsible for the PheRS^{ts} phenotype of NP37 (and its subsequent derivative KA2). In addition, the absence of D98 in strain K10-F6 proved that this mutation in *pheS* on plasmid pKSC-K (and subsequently on pKSB1/2-K) must have been picked up from NP37 (RecA^{+}!) during cloning in this host strain, probably via homologous recombination. Neither strain AB1360-12, which showed a slight thermosensitivity *in vivo* (Table 3), nor the wild-type control strain HB101 possessed mutation D98 (Fig. 7).

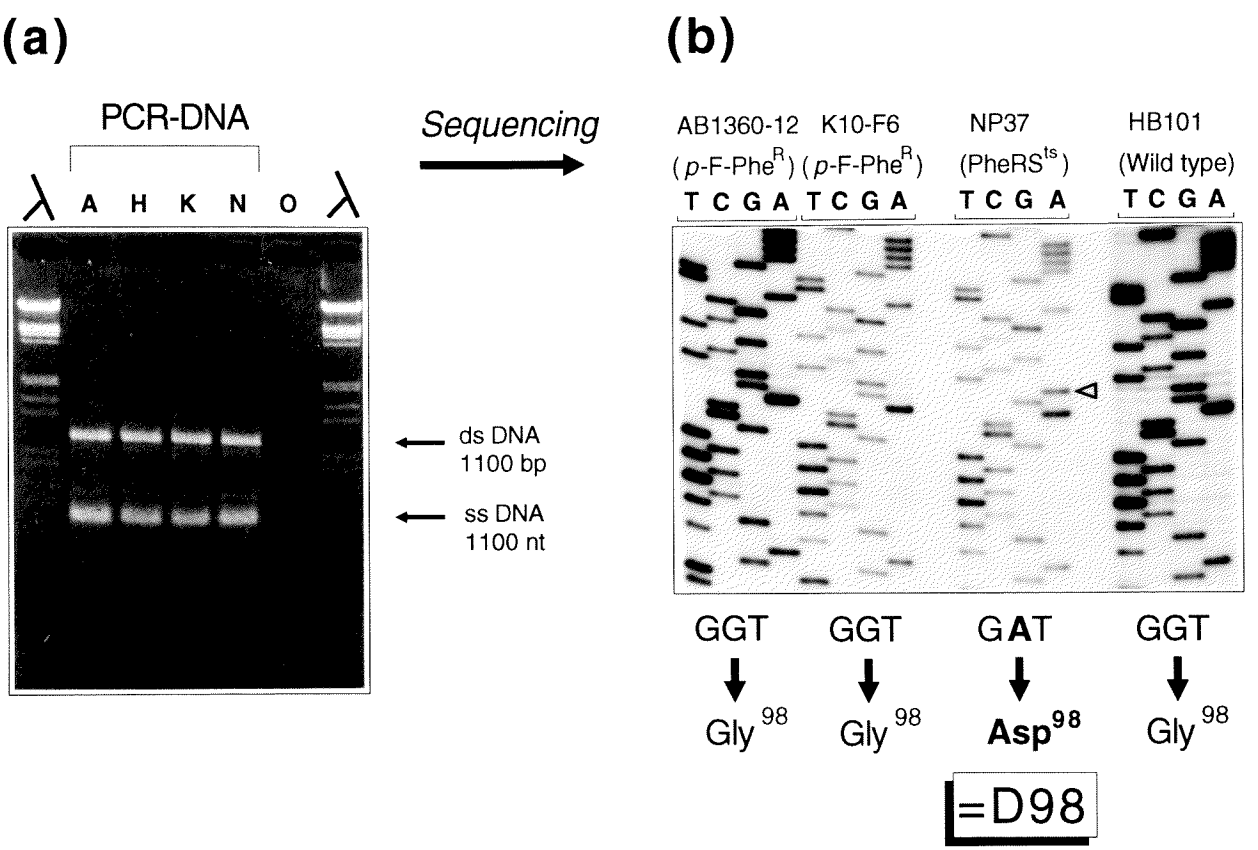


Figure 7. Sequencing of PCR-generated DNA to confirm the presence of mutation D98. (a) Ethidium bromide-stained agarose gel with PCR products. It shows amplified double-stranded as well as single-stranded DNA corresponding to the *pheS* fragments from different strains. Chromosomal DNA templates were from strains AB1360-12 (A), K10-F6 (K), NP37 (N) and the wild type, HB101 (H). In the negative control (0), no template was added. Lanes labelled with λ indicate the size marker (λ-DNA digested with *Eco*RI and *Hin*dIII). nt, nucleotides. **(b)** Sequencing ladders obtained from single-stranded PCR products (see (a)) that were sequenced with PEKA8N as primer. The codons corresponding to amino acid position 98 of the PheRS α subunit are noted below each sequence. Mutation D98 is indicated directly in the sequence by an arrowhead.

5. Discussion

5.1. The basis of thermosensitivity in NP37

The PheRS^{ts} mutation D98 of *E. coli* strain NP37 was cloned accidentally by a recombination event while cloning the *pheS* gene from another strain in the NP37 background. That the RecA$^+$ host NP37 possesses high recombination activity became already apparent by the large background in the *in vivo* plate tests (data not shown), which disappeared when *recA* was deleted (in strain KA2). Hennecke (1976) showed that in RecA$^+$ strains, *pheS* mutations were efficiently transferred to F' plasmids carrying the corresponding genomic regions (homogenotization).

The experiences from cloning the *pheS^{ts}* mutation D98 on plasmid pKSC-K showed that it should be possible in general to clone a variety of defective alleles by complementation of the very same mutation, making use of increased expression and gene dosage on multicopy plasmids. Such a cloning strategy is especially important in cases where no information about the gene structure is available. Other examples in which mutant PheRS phenotypes could be reversed by an increase of the mutant gene dosage were the complementation of strain G1 as well as an *E. coli* strain with a thermosensitive β subunit of PheRS (Grüll *et al.*, 1979). Similarly, the amplification of the genes for an unstable MetRS (from a methionine auxotrophic yeast mutant; Chatton *et al.*, 1987) or a thermosensitive AlaRS (Jasin & Schimmel, 1984) allowed to restore cell viability at high temperature.

Previous attempts to detect the tetrameric PheRS in NP37 crude extracts failed; instead, inactive PheRS protein corresponding to the size of an αβ dimer was found (Böck, 1968; Comer & Böck, 1976). From these results, the cause for the temperature-sensitive behaviour was supposed to be the facilitated disaggregation of mutant PheRS due to weakened subunit interactions. This behaviour, which is critical *in vivo* only at elevated temperature, may be brought about by increased electrostatic repulsion between subunits (or αβ dimers) at their interface. Such a hypothesis would provide a mechanistic explanation for mutation D98 which introduces (besides a size change) a negative charge (Gly-to-Asp replacement). Dissociation of aminoacyl-tRNA synthetase subunits by introduction of charged residues has been observed in TyrRS (Jones *et al.*, 1985) and in two ther-

mosensitive AlaRS mutants, in which also Gly-to-Asp exchanges occurred that were located in the 'oligomerization domain' (Jasin *et al.*, 1985; Schimmel, 1987). By analogy, mutation D98, which mapped to a non-conserved region in the N-terminal half of the PheRS α subunit, might identify an amino acid in the contact area between subunits or in a domain responsible for oligomerization (Jasin *et al.*, 1984; Schimmel, 1987). In this context, it would be of interest to map other PheRS^{ts} α mutations or analyze second site revertants of those mutants (Böck, 1968).

A different explanation for the NP37 phenotype may be derived from the observation that thermosensitivity of this strain was reversed by increasing the tRNA^{Phe} gene dosage (Caillet *et al.*, 1983; Schwartz *et al.*, 1983). Kinetic analysis of PheRS from NP37 indicated that it had a 20-fold increase in the K_M for tRNA^{Phe}, primarily due to a higher dissociation constant (Goodman & Schwartz, 1988). An increased tRNA^{Phe} concentration could compensate for this and might result in the formation of an active tRNA-PheRS complex *in vivo*, whereby the tRNA contributes to stabilization of the tetrameric structure. Stabilization of the oligomeric structure by high substrate concentrations was also demonstrated for a TyrRS mutant (Jones *et al.*, 1985). The impaired K_M of the NP37 PheRS was thus either a consequence of the aforementioned, weakened subunit interactions, or it implies that the amino acid at position 98 is directly involved in tRNA binding. In the latter case, D98 would decrease the affinity for the tRNA by a similar charge repulsion mechanism as described above. So far, affinity labelling experiments did not help to identify amino acid residues in the *E. coli* PheRS α subunit involved in tRNA binding (Hountondji *et al.*, 1987; Khodyreva, *et al.*, 1985).

5.2. From PheRS mutations to the Phe binding site?

Both PheRS mutations S294 and D191 affect Phe binding; therefore, their positions might identify regions involved in forming the Phe binding site. Because of the additional weak thermosensitivity of mutant PheRS carrying mutation D191, we are aware of the possibility that secondary structural distortions caused the impaired affinity for the substrate. In other aminoacyl-tRNA synthetases, examples of K_M mutations within established or probable substrate binding sites are known (Schimmel, 1987; Fersht, 1987; Clarke *et al.*, 1988). Eriani *et al.* (1990*b*) showed that an *E .coli* ArgRS mutant, possessing an increased K_M for ATP and a reduced reaction rate, carries an alteration immediately adjacent to the HIGH consensus

sequence, the ATP binding site in class I aminoacyl-tRNA synthetases (Burbaum *et al.*, 1990). In a yeast MetRS mutant, Chatton *et al.* (1987) found the cause for an elevated K_M for methionine and for thermosensitivity to be a point mutation in the C-terminal half of the nucleotide binding fold that was inferred from the homologous *E. coli* MetRS crystal structure (Brunie *et al.*, 1990). By analogy with the other two crystallized class I synthetases (Brick *et al.*, 1989; Rould *et al.*, 1989), the mutation is probably in the Met binding site (Brunie *et al.*, 1990). Interestingly, the identified amino acid replacement consisted of a Gly-to-Asp exchange, as it is the case with mutation D191 in PheRS of strain G1.

As a member of class II aminoacyl-tRNA synthetases, the PheRS α subunit contains all three conserved sequence motifs which define this group (Eriani *et al.*, 1990*a*), but none of the typical sequences for class I enzymes. It was proposed that the three motifs are part of a super-secondary structure and constitute a functional domain analogous to the nucleotide binding fold in class I synthetases (Eriani *et al.*, 1990*a*). Interestingly, the K_M mutation D191 (Gly-to-Asp) maps within motif 2, only 4 residues away from a totally conserved arginine. In eight out of 17 aligned wild-type sequences, a Gly is present at that position, whereas the yeast cytoplasmic PheRS has an Asp residue there (Eriani *et al.*, 1990*a*). In this context it is worth mentioning that the yeast enzyme displays a five-fold higher K_M for Phe as compared to the *E. coli* enzyme (Gabius *et al.*, 1983). The increase of the K_M in the D191 mutant PheRS (from strain G1) was 15-fold relative to the wild type (Grüll *et al.*, 1979). It is tempting to speculate, therefore, that mutation D191 identifies motif 2 as being part of the amino acid binding site in PheRS, and more generally in class II synthetases. This is supported by the crystal structure of SerRS (Cusack *et al.*, 1990), the only class II synthetase structure published up to now: although the absence of substrates in the crystal did not permit a definitive identification of the binding sites, a large cavity within the globular domain of the enzyme was assumed to be the active site. Motif 2 amino acids, and more precisely those corresponding to the PheRS α subunit residues adjacent to position 191, form one inside wall of this cavity and may well be involved in substrate binding.

S294, the PheRS α subunit mutation responsible for *p*-F-Phe resistance, was thought to point more directly to amino acids in contact with the substrate Phe (see section 5.3.). The mutation mapped at the C-terminus of the PheRS α subunit precisely within motif 3, another conserved region in class II synthetases (Fig. 8; Eriani *et al.*, 1990*a*). A closer inspection of motif 3 around position 294 showed a particularly high degree of conservation within the known PheRS sequences from different organisms. In Fig. 8, the corresponding regions of the four previously se-

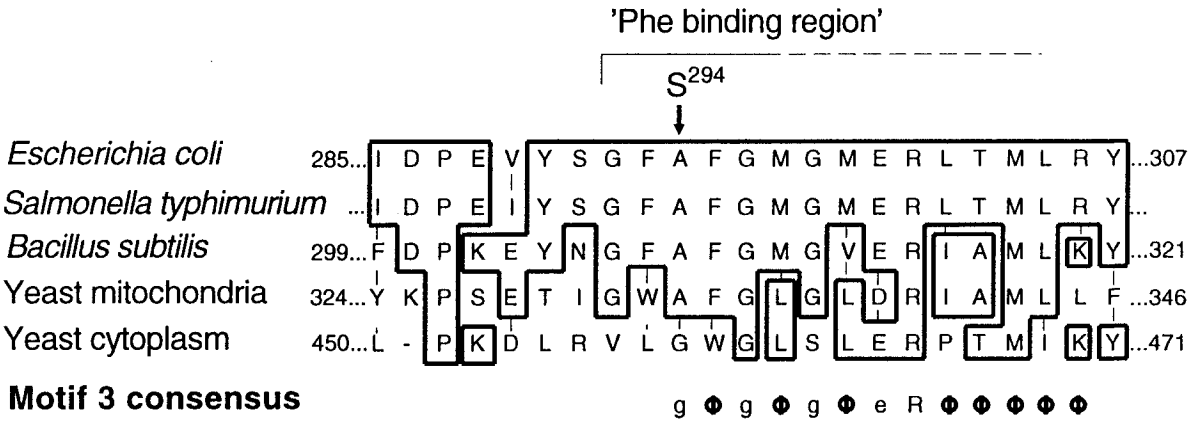


Figure 8. Alignment of PheRS small subunit sequences encoded by the *pheS* region around the position of mutation S294. The *pheS* gene products from *E. coli* (Fayat *et al.*, 1983), *B. subtilis* (Brakhage *et al.*, 1990), yeast mitochondria (Koerner *et al.*, 1987) and yeast cytoplasm (Sanni *et al.*, 1988) were aligned using the UWGCG program CLUSTAL (Higgins & Sharp, 1988). Furthermore, we included a corresponding protein sequence derived from a partial *pheS* sequence of *S. typhimurium* (B. Keller and P. Kast, unpublished work). Numbers indicate the first and the last residues shown here for the respective sequences. The one-letter code for amino acids is used. Identical residues are boxed; thin lines connect related amino acids with a pairwise value of at least 0.7 according to the Dayhoff mutation data matrix (Dayhoff *et al.*, 1983), rescaled as indicated in the UWGCG sequence analysis software package, release 6.2. In addition, the motif 3 consensus sequence of class II synthetases is aligned (adapted from Eriani *et al.*, 1990*a*). Lower case letters denote predominantly occurring residues; the arginine is completely conserved; Φ stands for hydrophobic amino acids if they are present in more than 70 % of the sequences examined by Eriani *et al.* (1990*a*).

quenced small PheRS subunits from *E. coli*, yeast cytoplasm and mitochondria, and *B. subtilis* are aligned. Furthermore, we added the equivalent sequence from *Salmonella typhimurium*, a close relative of *E. coli*. Its *pheS* gene was cloned recently in our laboratory by a strategy analogous to that described in this work for cloning the mutant *pheS* genes (B. Keller and P. Kast, unpublished results). The presence of the *p*-F-Phe^R mutation within a PheRS-specific conserved region prompted us to localize there at least part of the Phe binding site, an assumption that was further validated in the accompanying paper (Kast *et al.*, 1991). We thus propose that motif 3, together with motif 2, might constitute a component of the amino acid substrate binding domain in class II aminoacyl-tRNA synthetases. Again, this proposal is in agreement with the topology of the homologous sequence in the structure of SerRS (Cusack *et al.*, 1990). The position corresponding to the motif 3 residue 294 of the PheRS α subunit participates in a β-sheet building the bottom of the putative active site. Both positions corresponding to the PheRS α

subunit amino acids 191 and 294 come to lie close together in the tertiary structure of SerRS, which seems to support our assumption about their cooperation in amino acid binding. Of course, to extrapolate the assignment of structural elements from SerRS to specific positions in PheRS is only valid, if the secondary and tertiary structure is as conserved among class II synthetases as it was found among class I members (Burbaum *et al.*, 1990). The completion of the crystal structure of class II AspRS from yeast (complexed with tRNA) should make this clearer (Ruff *et al.*, 1988; Cusack *et al.*, 1990). At any rate, one should be aware that the situation with PheRS may be more complicated because the additional β subunit (which shows no homology to other synthetases) seems to be involved in tRNA binding in *E. coli* (Khodyreva *et al.*, 1985; Hountondji *et al.*, 1987) and yeast (Sanni *et al.*, 1991) and might contribute to an active site located at the αβ interface (Baltzinger *et al.*, 1979; Khodyreva *et al.*, 1985). The functional importance of the conserved motifs 2 and 3 was demonstrated by mutations in yeast and *E. coli* AspRS (Prevost *et al.*, 1989; Eriani *et al.*, 1990*a*; Eriani *et al.*, 1990*c*), in *E. coli* AsnRS (Anselme & Härtlein, 1991) and in yeast PheRS (Sanni *et al.*, 1990) which dramatically impaired enzyme activity.

From an increased K_M value for ATP in AsnRS mutated in motif 3, it was recently proposed that motif 3 might participate in ATP binding in class II synthetases (Anselme & Härtlein, 1991). This would not exclude a simultaneous involvement of this sequence in amino acid substrate binding since amino acid adenylate formation requires both substrates to come close together, and their binding sites may therefore be immediately adjacent, as exemplified, for instance, for TyrRS (Brick *et al.*, 1989).

5.3. A model for the mechanism of *p*-F-Phe resistance

The fact, that in *p*-F-Phe^R strains the slightly bigger substrate analogue *p*-F-Phe is excluded from the enzymatic reaction due to the replacement of an alanine residue by serine led to further considerations concerning the mechanisms involved in specific substrate recognition. In Fig. 9, we present a simple model that explains the exclusion of the substrate analogue mainly for steric reasons. The principal assumption is that position 294 of the PheRS α subunit makes direct contact with the *para*-position of the aromatic ring of Phe. In wild-type PheRS, the Ala^{294} would thus allow tight fitting of the analogue (and the natural substrate) into the binding pocket (Fig. 9, left side). By contrast, Ser^{294} with its hydroxyl group

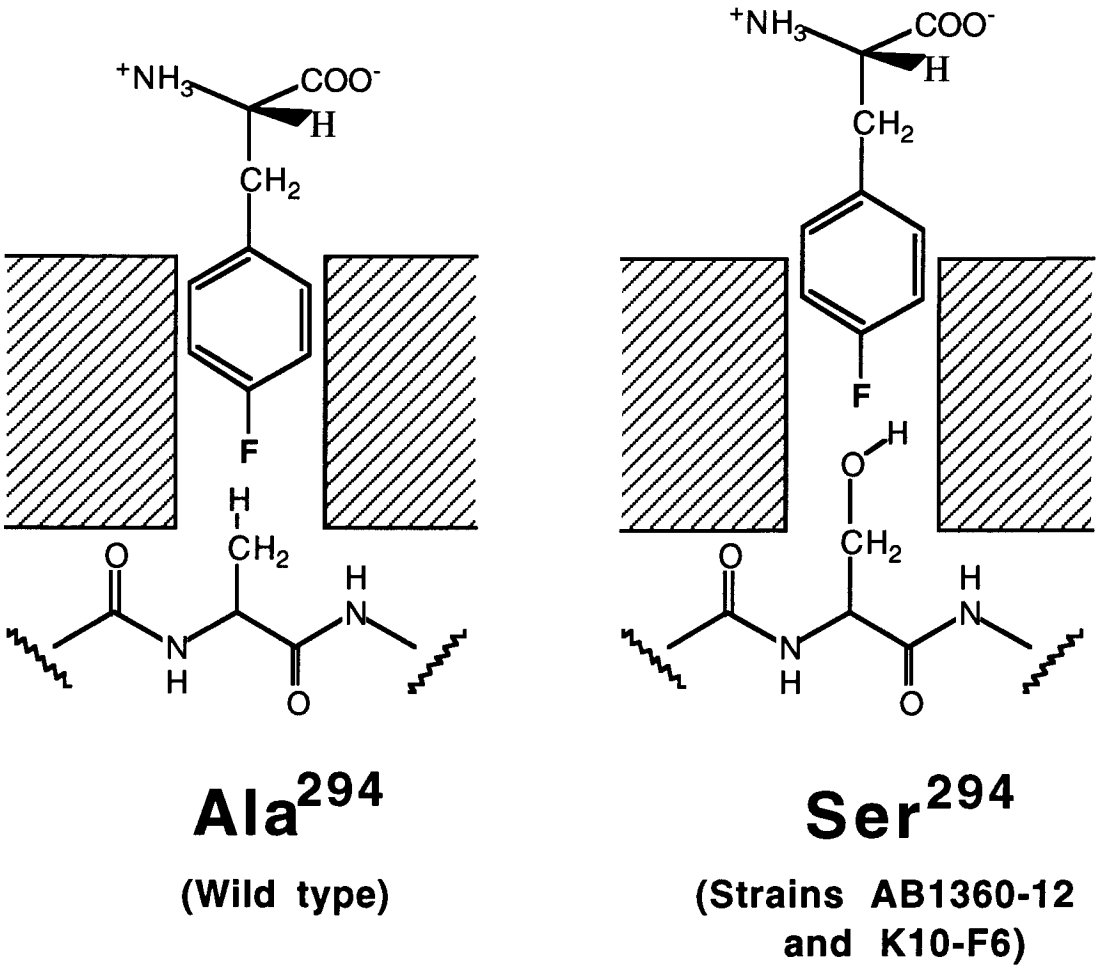


Figure 9. A model that might explain the mechanism of *p*-F-Phe resistance. In a *p*-F-Phe sensitive strain (wild type, on the left), the substrate analogue can enter the Phe binding pocket. In a strain resistant to the analogue, the mutation leading to the replacement of Ala^{294} by Ser^{294} (on the right) prevents *p*-F-Phe from productive binding, possibly due to steric interference of the additional OH-group with the F at the *para*-position of the aromatic ring.

would prevent a complete entering of *p*-F-Phe into the binding site due to steric hindrance, whereas Phe still fits. The OH-group might even establish a F-H hydrogen bond, thereby further displacing the analogue from the catalytically productive binding position (Fig. 9, right side).

The question of whether the difference in van der Waals radii (Bondi, 1964) between H (1.20 Å) and F (1.47 Å) and the potential of the electronegative F to form hydrogen bonds are sufficient to explain the observed phenotype, remains unanswered. Our model does not account for the differences in electron distribution in the aromatic ring systems between the analogue and Phe, although this

could also be a cause for differentiation and discrimination. Previous binding studies with substrate analogues, however, indicated that steric effects are clearly dominant in determining amino acid specificity in the PheRS hydrophobic pocket (Santi & Danenberg, 1971). If, as suggested, Ala^{294} is part of the binding site for Phe, the flanking aromatic (or at least hydrophobic) residues (see Fig. 8) could also participate in the binding of the aromatic ring of the substrate. The results from oligonucleotide-directed mutagenesis experiments presented in the accompanying paper (Kast *et al.*, 1991) provide evidence in favour of our model for the basis of *p*-F-Phe resistance, but speak against the involvement of residues 293 and 295 in substrate binding. Three-dimensional structural information on PheRS is now needed to elucidate the real topology of the binding site. The generation of crystals of *Thermus thermophilus* PheRS (Chernaya *et al.*, 1987; Ankilova *et al.*, 1988) suitable for X-ray structural analysis was a promising step towards this goal, and, therefore, we are currently engaged in determining the nucleotide sequence of the *T. thermophilus pheST* genes.

Acknowledgements

The valuable and expert assistance of Claudia Wehrli is gratefully acknowledged. We thank Barbara Keller for cloning the *pheS* gene from *S. typhimurium*. A. Birkmann is thanked for providing us with a P1kc lysate and K. Girgenrath and L. Dürselen for scintillation counting. This work was supported by a grant from the Federal Institute of Technology Zürich.

6. References

Ankilova, V. N., Reshetnikova, L. S., Chernaya, M. M. & Lavrik, O. I. (1988). Phenylalanyl-tRNA Synthetase from *Thermus thermophilus* HB8. Purification and Properties of the Crystallizing Enzyme. *FEBS Lett.* **227**, 9-13.

Anselme, J. & Härtlein, M. (1991). Tyr-426 of the *Escherichia coli* Asparaginyl-tRNA Synthetase, an Amino Acid in a C-Terminal Conserved Motif, Is Involved in ATP Binding. *FEBS Lett.* **280**, 163-166.

Bachmann, B. J. (1990). Linkage Map of *Escherichia coli* K-12, Edition 8. *Microbiol. Rev.* **54**, 130-197.

Baltzinger, M. & Holler, E. (1982). Catalytic Mechanism of Phenylalanyl-tRNA Synthetase of *Escherichia coli* K10. Conformational Change and $tRNA^{Phe}$ Phenylalanylation Are Concerted. *Biochemistry*, **21**, 2467-2476.

Baltzinger, M., Fasiolo, F. & Remy, P. (1979). Yeast Phenylalanyl-tRNA Synthetase. Affinity and Photoaffinity Labelling of the Stereospecific Binding Sites. *Eur. J. Biochem.* **97**, 481-494.

Blow, D. M., Bhat, T. N., Metcalfe, A., Risler, J.L., Brunie, S. & Zelwer, C. (1983). Structural Homology in the Amino-Terminal Domains of Two Aminoacyl-tRNA Synthetases. *J. Mol. Biol.* **171**, 571-576.

Böck, A. (1968). Relation between Subunit Structure and Temperature-Sensitivity of Mutant Phenylalanyl RNA Synthetases of *Escherichia coli*. *Eur. J. Biochem.* **4**, 395-400.

Böck A. & Neidhardt, F. C. (1967). Genetic Mapping of Phenylalanyl-sRNA Synthetase in *Escherichia coli*. *Science*, **157**, 78-79.

Bondi, A. (1964). van der Waals Volumes and Radii. *J. Phys. Chem.* **68**. 441-451.

Boyer, H. W. & Roulland-Dussoix, D. (1969). A Complementation Analysis of the Restriction and Modification of DNA in *Escherichia coli*. *J. Mol. Biol.* **41**, 459-472.

Brakhage, A. A., Wozny, M. & Putzer, H. (1990). Structure and Nucleotide Sequence of the *Bacillus subtilis* Phenylalanyl-tRNA Synthetase Genes. *Biochimie*, **72**, 725-734. - Erratum (1991): *Biochimie*, **73**, 127.

Brick, P., Bhat, T. N. & Blow, D. M. (1989). Structure of Tyrosyl-tRNA Synthetase Refined at 2.3 Å Resolution. Interaction of the Enzyme with the Tyrosyl Adenylate Intermediate. *J. Mol. Biol.* **208**, 83-98.

Brunie, S., Zelwer, C. & Risler, J.-L. (1990). Crystallographic Study at 2.5 Å Resolution of the Interaction of Methionyl-tRNA Synthetase from *Escherichia coli* with ATP. *J. Mol. Biol.* **216**, 411-424.

Burbaum, J. J., Starzyk, R. M. & Schimmel, P. (1990). Understanding Structural Relationships in Proteins of Unsolved Three-Dimensional Structure. *Proteins*, **7**, 99-111.

Caillet, J., Plumbridge, J. A., Springer, M., Vacher, J., Delamarche, C., Buckingham, R. H. & Grunberg-Manago, M. (1983). Identification of Clones Carrying an *E. coli* $tRNA^{Phe}$ Gene by Suppression of Phenylalanyl-tRNA Synthetase Thermosensitive Mutants. *Nucl. Acids Res.* **11**, 727-736.

Chang, A. C. Y. & Cohen, S. N. (1978). Construction and Characterization of Amplifiable Multicopy DNA Cloning Vehicles Derived from the P15A Cryptic Miniplasmid. *J. Bacteriol.* **134**, 1141-1156.

Chatton, B., Winsor, B., Boulanger, Y. & Fasiolo, F. (1987). Cloning and Characterization of the Yeast Methionyl-tRNA Synthetase Mutation *mes1*. *J. Biol. Chem.* **262**, 15094-15097.

Chernaya, M. M., Korolev, S. V., Reshetnikova, L. S. & Safro, M. G. (1987). Preliminary Crystallographic Study of the Phenylalanyl-tRNA Synthetase from *Thermus thermophilus* HB8. *J. Mol. Biol.* **198**, 555-556.

Cherniack, A. D., Garriga, G., Kittle, Jr., J. D., Akins, R. A. & Lambowitz, A. M. (1990). Function of Neurospora Mitochondrial Tyrosyl-tRNA Synthetase in RNA Splicing Requires an Idiosyncratic Domain Not Found in Other Synthetases. *Cell*, **62**, 745-755.

Clarke, N. D., Lien, D. C. & Schimmel, P. (1988). Evidence from Cassette Mutagenesis for a Structure-Function Motif in a Protein of Unknown Structure. *Science*, **240**, 521-523.

Comer, M. M. & Böck, A. (1976). Genes for the α and β Subunits of the Phenylalanyl-Transfer Ribonucleic Acid Synthetase of *Escherichia coli*. *J. Bacteriol.* **127**, 923-933.

Connell, C., Fung, S., Heiner, C., Bridgham, J., Chakerian, V., Heron, E., Jones, B., Menchen, S., Mordan, W., Raff, M., Recknor, M., Smith, L., Springer, J., Woo, S. & Hunkapiller, M. (1987). Automated DNA Sequence Analysis. *Biotechniques*, **5**, 342-348.

Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N. & Leberman, R. (1990). A Second Class of Synthetase Structure Revealed by X-Ray Analysis of *Escherichia coli* Seryl-tRNA Synthetase at 2.5 Å. *Nature (London)*, **347**, 249-255.

Dayhoff, M. O., Barker, W. C. & Hunt, L. T. (1983). Establishing Homologies in Protein Sequences. *Methods Enzymol.* **91**, 524-545.

de Duve, C. (1988). The Second Genetic Code. *Nature (London)*, **333**, 117-118.

Dessen, P., Ducruix, A., May, R. P. & Blanquet, S. (1990). Low-Resolution Structure of the Tetrameric Phenylalanyl-tRNA Synthetase from *Escherichia coli*. A Neutron Small-Angle Scattering Study of Hybrids Composed of Protonated and Deuterated Protomers. *Biochemistry*, **29**, 3039-3045.

Edwards, H. & Schimmel, P. (1990). A Bacterial Amber Suppressor in *Saccharomyces cerevisiae* Is Selectively Recognized by a Bacterial Aminoacyl-tRNA Synthetase. *Mol. Cell. Biol.* **10**, 1633-1641.

Eidlic, L. & Neidhardt, F. C. (1965). Protein and Nucleic Acid Synthesis in Two Mutants of *Escherichia coli* with Temperature-Sensitive Aminoacyl Ribonucleic Acid Synthetases. *J. Bacteriol.* **89**, 706-711.

Elhardt, D., Wirth, R. & Böck, A. (1982). Regulation of Formation of Threonyl-tRNA Synthetase, Phenylalanyl-tRNA Synthetase and Protein Synthesis Initiation Factor 3 from *Escherichia coli in Vivo* and *in Vitro*. *Eur. J. Biochem.* **123**, 477-482.

Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990*a*). Partition of tRNA Synthetases into Two Classes Based on Mutually Exclusive Sets of Sequence Motifs. *Nature (London)*, **347**, 203-206.

Eriani, G., Dirheimer, G. & Gangloff, J. (1990*b*). Structure - Function Relationship of Arginyl-tRNA Synthetase from *Escherichia coli*: Isolation and Characterization of the *argS* Mutation MA5002. *Nucl. Acids Res.* **18**, 1475-1479.

Eriani, G., Dirheimer, G. & Gangloff, J. (1990*c*). Aspartyl-tRNA Synthetase from *Escherichia coli*: Cloning and Characterisation of the Gene, Homologies of Its

Translated Amino Acid Sequence with Asparaginyl- and Lysyl-tRNA Synthetases. *Nucl. Acids Res.* **18**, 7109-7118.

Fangman, W. L. & Neidhardt, F. C. (1964). Protein and Ribonucleic Acid Synthesis in a Mutant of *Escherichia coli* with an Altered Aminoacyl Ribonucleic Acid Synthetase. *J. Biol. Chem.* **239**, 1844-1847.

Fasiolo, F. & Fersht, A. R. (1978). The Aminoacyladenylate Mechanism in the Aminoacylation Reaction of Yeast Phenylalanyl-tRNA Synthetase. *Eur. J. Biochem.* **85**, 85-88.

Fasiolo, F., Sanni, A., Potier, S., Ebel, J. P. & Boulanger, Y. (1989). Identification of the Major $tRNA^{Phe}$ Binding Domain in the Tetrameric Structure of Cytoplasmic Phenylalanyl-tRNA Synthetase from Baker's Yeast. *FEBS Lett.* **242**, 351-356.

Fayat, G., Blanquet, S., Dessen, P., Batelier, G. & Waller, J.-P. (1974). The Molecular Weight and Subunit Composition of Phenylalanyl-tRNA Synthetase from *Escherichia coli* K-12. *Biochimie,* **56**, 35-41.

Fayat, G., Mayaux, J.-F., Sacerdot, C., Fromant, M., Springer, M., Grunberg-Manago, M. & Blanquet, S. (1983). *Escherichia coli* Phenylalanyl-tRNA Synthetase Operon Region. Evidence for an Attenuation Mechanism. Identification of the Gene for the Ribosomal Protein L20. *J. Mol. Biol.* **171**, 239-261.

Fersht, A. R. (1987). Dissection of the Structure and Activity of the Tyrosyl-tRNA Synthetase by Site-Directed Mutagenesis. *Biochemistry,* **26**, 8031-8037.

Fersht, A. R., Knill-Jones, J. W., Bedouelle, H. & Winter, G. (1988). Reconstruction by Site-Directed Mutagenesis of the Transition State for the Activation of Tyrosine by the Tyrosyl-tRNA Synthetase: A Mobile Loop Envelopes the Transition State in an Induced-Fit Mechanism. *Biochemistry,* **27**, 1581-1587.

Freist, W. (1989). Mechanisms of Aminoacyl-tRNA Synthetases: A Critical Consideration of Recent Results. *Biochemistry,* **28**, 6787-6795.

Gabius, H.-J., von der Haar, F. & Cramer, F. (1983). Evolutionary Aspects of Accuracy of Phenylalanyl-tRNA Synthetase. A Comparative Study with Enzymes from *Escherichia coli, Saccharomyces cerevisiae, Neurospora crassa,* and Turkey Liver Using Phenylalanine Analogues. *Biochemistry,* **22**, 2331-2339.

Ghosh, G., Pelka, H. & Schulman, L. H. (1990). Identification of the tRNA Anticodon Recognition Site of *Escherichia coli* Methionyl-tRNA Synthetase.

Biochemistry, **29**, 2220-2225.

Goodman, R. & Schwartz, I. (1988). Kinetic Analysis of an *E. coli* Phenylalanine-tRNA Synthetase Mutant. *Nucl. Acids Res.* **16**, 7477-7486.

Grüll, J. M., Hennecke, H., Fröhler, J., Thomale, J., Nass, G. & Böck, A. (1979). *Escherichia coli* Mutants Overproducing Phenylalanyl- and Threonyl-tRNA Synthetase. *J. Bacteriol.* **137**, 480-489.

Grunberg-Manago, M. (1987). Regulation of the Expression of Aminoacyl-tRNA Synthetases and Translation Factors. In *Escherichia coli and Salmonella typhimurium*: *Cellular and Molecular Biology* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M. & Umbarger, H. E., eds), vol. 2, pp. 1386-1409, Am. Soc. Microbiol., Washington, DC.

Gyllensten, U. B. & Erlich, H. A. (1988). Generation of Single-Stranded DNA by the Polymerase Chain Reaction and Its Application to Direct Sequencing of the *HLA-DQA* Locus. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 7652-7656.

Hanke, T., Bartmann, P., Hennecke, H., Kosakowski, H. M., Jaenicke, R., Holler, E. & Böck, A. (1974). L-Phenylalanyl-tRNA Synthetase of *Escherichia coli* K-10. A Reinvestigation of Molecular Weight and Subunit Structure. *Eur. J. Biochem.* **43**, 601-607.

Hennecke, H. (1976). Use of Mutant Enzymes to Demonstrate the Presence of Two Active Sites on Phenylalanyl-tRNA Synthetase from *Escherichia coli*. *FEBS Lett.* **72**, 182-186.

Hennecke, H. & Böck, A. (1975). Altered α Subunits in Phenylalanyl-tRNA Synthetases from *p*-Fluorophenylalanine-Resistant Strains of *Escherichia coli*. *Eur. J. Biochem.* **55**, 431-437.

Hennecke, H., Günther, I. & Binder, F. (1982). A Novel Cloning Vector for the Direct Selection of Recombinant DNA in *E. coli*. *Gene*, **19**, 231-234.

Hennecke, H., Springer, M. & Böck, A. (1977). A Specialized Transducing λ Phage Carrying the *Escherichia coli* Genes for Phenylalanyl-tRNA Synthetase. *Mol. Gen. Genet.* **152**, 205-210.

Herbert, C. J., Labouesse, M., Dujardin, G. & Slonimski, P. P. (1988). The NAM2 Proteins from *S. cerevisiae* and *S. douglasii* Are Mitochondrial Leucyl-tRNA Synthetases, and Are Involved in mRNA Splicing. *EMBO J.* **7**, 473-483.

Higgins, D. G. & Sharp, P. M. (1988). CLUSTAL: A Package for Performing Multiple Sequence Alignment on a Microcomputer. *Gene*, **73**, 237-244.

Hountondji, C., Blanquet, S. & Lederer, F. (1985). Methionyl-tRNA Synthetase from *Escherichia coli*: Primary Structure at the Binding Site for the 3'-End of $tRNA_f^{Met}$. *Biochemistry*, **24**, 1175-1180.

Hountondji, C., Dessen, P. & Blanquet, S. (1986). Sequence Similarities among the Family of Aminoacyl-tRNA Synthetases. *Biochimie*, **68**, 1071-1078.

Hountondji, C., Schmitter, J.-M., Beauvallet, C. & Blanquet, S. (1987). Affinity Labelling of *Escherichia coli* Phenylalanyl-tRNA Synthetase at the Binding Site for $tRNA^{Phe}$. *Biochemistry*, **26**, 5433-5439.

Hountondji, C., Schmitter, J.-M., Fukui, T., Tagaya, M. & Blanquet, S. (1990). Affinity Labelling of Aminoacyl-tRNA Synthetases with Adenosine Triphosphopyridoxal: Probing the Lys-Met-Ser-Lys-Ser Signature Sequence as the ATP-Binding Site in *Escherichia coli* Methionyl- and Valyl-tRNA Synthetases. *Biochemistry*, **29**, 11266-11273.

Ihara, M., Oda, Y. & Yamamoto, K. (1985). Convenient Construction of Strains Useful for Transducing *recA* Mutations with Bacteriophage P1. *FEMS Microbiol. Lett.* **30**, 33-35.

Jasin, M. & Schimmel, P. (1984). Deletion of an Essential Gene in *Escherichia coli* by Site-Specific Recombination with Linear DNA Fragments. *J. Bacteriol.* **159**, 783-786.

Jasin, M., Regan, L. & Schimmel, P. (1984). Dispensable Pieces of an Aminoacyl tRNA Synthetase Which Activate the Catalytic Site. *Cell*, **36**, 1089-1095.

Jasin, M., Regan, L. & Schimmel, P. (1985). Two Mutations in the Dispensable Part of Alanine tRNA Synthetase Which Affect the Catalytic Activity. *J. Biol. Chem.* **260**, 2226-2230.

Jones, D. H., McMillan, A. J. & Fersht, A. R. (1985). Reversible Dissociation of Dimeric Tyrosyl-tRNA Synthetase by Mutagenesis at the Subunit Interface. *Biochemistry*, **24**, 5852-5857.

Kast, P., Wehrli, C. & Hennecke, H. (1991). Amino Acid Substrate Specificitiy of *Escherichia coli* Phenylalanyl-tRNA Synthetase Altered by Site-Directed Mutagenesis. *J. Mol. Biol.* (submitted); see Chapter III.

Khodyreva, S. N., Moor, N. A., Ankilova, V. N. & Lavrik, O. I. (1985). Phenylalanyl-tRNA Synthetase from *E. coli* MRE-600: Analysis of the Active Site Distribution on the Enzyme Subunits by Affinity Labelling. *Biochim. Biophys. Acta,* **830**, 206-212.

Koerner, T. J., Myers, A. M., Lee, S. & Tzagoloff, A. (1987). Isolation and Characterization of the Yeast Gene Coding for the α Subunit of Mitochondrial Phenylalanyl-tRNA Synthetase. *J. Biol. Chem.* **262**, 3690-3696.

Kosakowski, M. H. J. E. & Böck, A. (1970). The Subunit Structure of Phenylalanyl-tRNA Synthetase of *Escherichia coli. Eur. J. Biochem.* **12**, 67-73.

Labouesse, M. (1990). The Yeast Mitochondrial Leucyl-tRNA Synthetase Is a Splicing Factor for the Excision of Several Group I Introns. *Mol. Gen. Genet.* **224**, 209-221.

Leatherbarrow, R. J., Fersht, A. R. & Winter, G. (1985). Transition-State Stabilization in the Mechanism of Tyrosyl-tRNA Synthetase Revealed by Protein Engineering. *Proc. Nat. Acad. Sci., U.S.A.* **82**, 7840-7844.

Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951). Protein Measurement with the Folin Phenol Reagent. *J. Biol. Chem.* **193**, 265-275.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Masters, M. (1985). Generalized Transduction. In *Genetics of Bacteria* (Scaife, J., Leach, D. & Galizzi, A., eds), pp. 197-215, Academic Press, London.

Mechulam, Y., Fayat, G. & Blanquet, S. (1985). Sequence of the *Escherichia coli pheST* Operon and Identification of the *himA* Gene. *J. Bacteriol.* **163**, 787-791.

Meinnel, T., Mechulam, Y., Dardel, F., Schmitter, J.-M., Hountondji, C., Brunie, S., Dessen, P., Fayat, G. & Blanquet, S. (1990). Methionyl-tRNA Synthetase from *E. coli* - A Review. *Biochimie*, **72**, 625-632.

Meinnel, T., Mechulam, Y., Le Corre, D., Panvert, M., Blanquet, S. & Fayat, G. (1991). Selection of Suppressor Methionyl-tRNA Synthetases: Mapping the tRNA Anticodon Binding Site. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 291-295.

Mellot, P., Mechulam, Y., Le Corre, D., Blanquet, S. & Fayat, G. (1989). Identification of an Amino Acid Region Supporting Specific Methionyl-tRNA

Synthetase : tRNA Recognition. *J. Mol. Biol.* **208**, 429-443.

Miller, J. H. (1972). *Experiments in Molecular Genetics,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Moras, D. (1990). Synthetases Gain Recognition. *Nature (London)*, **344**, 195-197.

Nagel, G. M., Johnson, M. S., Rynd, J., Petrella, E. & Weber, B. H. (1988). Glycyl-tRNA Synthetase of *Escherichia coli*: Immunological Homology with Phenylalanyl-tRNA Synthetase. *Arch. Biochem. Biophys.* **262**, 409-415.

Nakamaye, K. L. & Eckstein, F. (1986). Inhibition of Restriction Endonuclease *Nci*I Cleavage by Phosphorothioate Groups and Its Application to Oligonucleotide-Directed Mutagenesis. *Nucl. Acids Res.* **14**, 9679-9698.

Neidhardt, F. C. (1966). Roles of Amino Acid Activating Enzymes in Cellular Physiology. *Bacteriol. Rev.* **30**, 701-719.

Normanly, J. & Abelson, J. (1989). tRNA Identity. *Ann. Rev. Biochem.* **58**, 1029-1049.

Perona, J. J., Swanson, R. N., Rould, M. A., Steitz, T. A. & Söll, D. (1989). Structural Basis for Misaminoacylation by Mutant *E. coli* Glutaminyl-tRNA Synthetase Enzymes. *Science*, **246**, 1152-1154.

Peterson, G. L. (1983). Determination of Total Protein. *Methods Enzymol.* **91**, 95-119.

Plumbridge, J. A., Springer, M., Graffe, M., Goursot, R. & Grunberg-Manago, M. (1980). Physical Localisation and Cloning of the Structural Gene for *E. coli* Initiation Factor IF3 from a Group of Genes Concerned with Translation. *Gene,* **11**, 33-42.

Prevost, G., Eriani, G., Kern, D., Dirheimer, G. & Gangloff, J. (1989). Study of the Arrangement of the Functional Domains along the Yeast Cytoplasmic Aspartyl-tRNA Synthetase. *Eur. J. Biochem.* **180**, 351-358.

Rapaport, E., Yogeeswaran, G., Zamecnik, P. C. & Remy, P. (1985). Covalent Modification of Phenylalanyl-tRNA Synthetase with Phenylalanine during the Amino Acid Activation Reaction Catalyzed by the Enzyme. *J. Biol. Chem.* **260**, 9509-9512.

Renaud, M., Fasiolo, F., Baltzinger, M., Boulanger, Y. & Remy, P. (1982). Affinity

Labelling of Yeast Phenylalanyl-tRNA Synthetase with a 3'-Oxidised $tRNA^{Phe}$. Isolation and Sequence of the Labelled Peptide. *Eur. J. Biochem.* **123**, 267-274.

Richmond, M. H. (1962). The Effect of Amino Acid Analogues on Growth and Protein Synthesis in Microorganisms. *Bacteriol. Rev.* **26**, 398-420.

Rossmann, M. G., Moras, D. & Olsen, K. W. (1974). Chemical and Biological Evolution of a Nucleotide-Binding Protein. *Nature (London),* **250**, 194-199.

Rould, M. A., Perona, J. J., Söll, D. & Steitz, T. A. (1989). Structure of *E. coli* Glutaminyl-tRNA Synthetase Complexed with $tRNA^{Gln}$ and ATP at 2.8 Å Resolution. *Science*, **246**, 1135-1142.

Ruff, M., Cavarelli, J., Mikol, V., Lorber, B., Mitschler, A., Giege, R., Thierry, J. C. & Moras, D. (1988). A High Resolution Diffracting Crystal Form of the Complex between Yeast $tRNA^{Asp}$ and Aspartyl-tRNA Synthetase. *J. Mol. Biol.* **201**, 235-236.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988). Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase. *Science,* **239**, 487-494.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual* (2^{nd} ed), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA Sequencing with Chain-Terminating Inhibitors. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463-5467.

Sanni, A., Mirande, M., Ebel, J.-P., Boulanger, Y., Waller, J.-P. & Fasiolo, F. (1988). Structure and Expression of the Genes Encoding the α and β Subunits of Yeast Phenylalanyl-tRNA Synthetase. *J. Biol. Chem.* **263**, 15407-15415.

Sanni, A., Walter, P., Ebel, J.-P. & Fasiolo, F. (1990). Construction of a *FRS1-FRS2* Operon Encoding the Structural Genes for the α and β Subunits of Yeast Phenylalanyl-tRNA Synthetase and Its Use in Deletion Analysis. *Nucl. Acids Res.* **18**, 2087-2092.

Sanni, A., Hountondji, C., Blanquet, S., Ebel, J.-P., Boulanger, Y. & Fasiolo, F. (1991). Interaction of the $tRNA^{Phe}$ Acceptor End with the Synthetase Involves a Sequence Common to Yeast and *Escherichia coli* Phenylalanyl-tRNA Synthetases. *Biochemistry*, **30**, 2448-2453.

Santi, D. V. & Danenberg, P. V. (1971). Phenylalanyl Transfer Ribonucleic Acid Synthetase from *Escherichia coli*. Analysis of the Phenylalanine Binding Site. *Biochemistry*, **10**, 4813-4820.

Schimmel, P. (1987). Aminoacyl tRNA Synthetases: General Scheme of Structure-Function Relationships in the Polypeptides and Recognition of Transfer RNAs. *Ann. Rev. Biochem.* **56**, 125-158.

Schimmel, P. R. & Söll, D. (1979). Aminoacyl-tRNA Synthetases: General Features and Recognition of Transfer RNAs. *Ann. Rev. Biochem.* **48**, 601-648.

Schleif, R. F. & Wensink, P. C. (1981). *Practical Methods in Molecular Biology,* Springer-Verlag, New York.

Schwartz, I., Klotsky, R.-A., Elseviers, D., Gallagher, P. J., Krauskopf, M., Siddiqui, M. A. Q., Wong, J. F. H. & Roe, B. A. (1983). Molecular Cloning and Sequencing of *pheU*, a Gene for *Escherichia coli* $tRNA^{Phe}$. *Nucl. Acids Res.* **11**, 4379-4389.

Short, J. M., Fernandez, J. M., Sorge, J. A. & Huse, W. D. (1988). λ ZAP: a Bacteriophage λ Expression Vector with *in Vivo* Excision Properties. *Nucl. Acids Res.* **16**, 7583-7600.

Shyamala, V. & Ames, G. F.-L. (1989). Amplification of Bacterial Genomic DNA by the Polymerase Chain Reaction and Direct Sequencing after Asymmetric Amplification: Application to the Study of Periplasmic Proteases. *J. Bacteriol.* **171**, 1602-1608.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H. & Hood, L. E. (1986). Fluorescence Detection in Automated DNA Sequence Analysis. *Nature (London)*, **321**, 674-679.

Söll, D. & Schimmel, P. R. (1974). Aminoacyl-tRNA Synthetases. In *The Enzymes* (Boyer, P. D., ed), vol. 10, pp. 489-538, Academic Press, London.

Southern, E. M. (1975). Detection of Specific Sequences Among DNA Fragments Separated by Gel Electrophoresis. *J. Mol. Biol.* **98**, 503-517.

Starzyk, R. M., Webster, T. A. & Schimmel, P. (1987). Evidence for Dispensable Sequences Inserted into a Nucleotide Fold. *Science,* **237**, 1614-1618.

Tabor, S. & Richardson, C. C. (1987). DNA Sequence Analysis with a Modified Bacteriophage T7 DNA Polymerase. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 4767-4771.

Toth, M. J. & Schimmel, P. (1990*a*). Deletions in the Large (β) Subunit of a Hetero-Oligomeric Aminoacyl-tRNA Synthetase. *J. Biol. Chem.* **265**, 1000-1004.

Toth, M. J. & Schimmel, P. (1990*b*). A Mutation in the Small (α) Subunit of Glycyl-tRNA Synthetase Affects Amino Acid Activation and Subunit Association Parameters. *J. Biol. Chem.* **265**, 1005-1009.

Tsang, T. C. & Bentley, D. R. (1988). An Improved Sequencing Method Using Sequenase^{TM} that Is Independent of Template Concentration. *Nucl. Acids Res.* **16**, 6238.

Wada, A. & Sako, T. (1987). Primary Structures of and Genes for New Ribosomal Proteins A and B in *Escherichia coli. J. Biochem.* **101**, 817-820.

Waldrop, M. M. (1989). The Structure of the "Second Genetic Code". *Science*, **246**, 1122.

Webster, T. A., Lathrop, R. H. & Smith, T. F. (1987). Prediction of a Common Structural Domain in Aminoacyl-tRNA Synthetases through Use of a New Pattern-Directed Inference System. *Biochemistry,* **26**, 6950-6957.

Weiner, A. M. & Maizels, N. (1987). tRNA-Like Structures Tag the 3' Ends of Genomic RNA Molecules for Replication: Implications for the Origin of Protein Synthesis. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 7383-7387.

Wrischnik, L. A., Higuchi, R. G., Stoneking, M., Erlich, H. A., Arnheim, N. & Wilson, A. C. (1987). Length Mutations in Human Mitochondrial DNA: Direct Sequencing of Enzymatically Amplified DNA. *Nucl. Acids Res.* **15**, 529-542.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 Phage Cloning Vectors and Host Strains: Nucleotide Sequence of the M13mp18 and pUC19 Vectors. *Gene,* **33**, 103-119.

Zhang, H., Scholl, R., Browse, J. & Somerville, C. (1988). Double Stranded DNA Sequencing as a Choice for DNA Sequencing. *Nucl. Acids Res.* **16**, 1220.

CHAPTER III

Amino Acid Substrate Specificity of *Escherichia coli* Phenylalanyl-tRNA Synthetase Altered by Site-Directed Mutagenesis

(Submitted to: *Journal of Molecular Biology*)

Peter Kast, Claudia Wehrli and Hauke Hennecke

Mikrobiologisches Institut, Eidgenössische Technische Hochschule,
Schmelzbergstrasse 7,
CH-8092 Zürich, Switzerland

Running title: Changing Phenylalanyl-tRNA Synthetase Specificity

Keywords: aromatic binding site / oligonucleotide-directed mutagenesis / phenylalanine analogues / phenylalanyl-tRNA synthetase / substrate specificity

Non-standard abbreviations

aa: amino acid; bp: basepair(s); 3-F-Tyr: 3-fluoro-tyrosine; IPTG: isopropyl-β-D-thiogalacto-pyranoside; kb: kilobasepair(s); *p*-F-Phe, *p*-Cl-Phe and *p*-Br-Phe: *para*-fluoro-, *para*-chloro- and *para*-bromo-phenylalanines; SDS: sodium dodecylsulfate; XxxRS: aminoacyl-tRNA synthetase for amino acid Xxx in 3-letter-code.

1. Abstract

In the accompanying report, an alanine-to-serine exchange at position 294 in the α subunit of phenylalanyl-tRNA synthetase (PheRS) was identified which renders *Escherichia coli* resistant to *p*-fluoro-phenylalanine (*p*-F-Phe). The site affected by mutation, which mapped to motif 3, a region generally conserved in class II aminoacyl-tRNA synthetases, was proposed to be part of the phenylalanine (Phe) binding pocket of PheRS. To test this hypothesis, we replaced the alanine at position 294 as well as the two flanking phenylalanines (positions 293 and 295) by a number of selected other amino acids. An expression system was designed to assay for the mutated gene products at the chromosomal expression level. *In vivo* and *in vitro* results demonstrated that Phe^{293} and Phe^{295} are not directly involved in substrate binding, but replacements of those residues affected PheRS stability. A glycine at position 295 yielded an enzyme with drastically reduced activity emphasizing the importance of conserved motif 3 residues. Exchanges of the amino acid at position 294 affected binding of Phe as observed by alterations in the apparent Michaelis-Menten constants. In addition, certain mutants showed pronounced changes in specificity towards amino acid substrate analogues. The results support a model according to which the residue at position 294 determines substrate specificity by interacting with the *para*-position of the amino acid's aromatic ring. Of particular interest was the Gly^{294} PheRS in which presumably an enlarged cavity for the *para*-position of the aromatic ring allowed an increased aminoacylation of tRNA with *p*-F-Phe. Moreover, the larger *para*-chloro- and *para*-bromo- derivatives of Phe could interact with this enzyme *in vitro* and became highly toxic *in vivo*. The possible exploitation of the Gly^{294} mutant PheRS for the incorporation of non-proteinogenic amino acids into proteins is discussed.

2. Introduction

In the preceding paper (Kast & Hennecke, 1991), a mutation was characterized that led to resistance of *Escherichia coli* against *p*-fluoro-phenylalanine (*p*-F-Phe). The mutation was located in *pheS*, the gene for the α subunit of the tetrameric ($\alpha_2\beta_2$) phenylalanyl-tRNA synthetase (PheRS) and consisted of an alanine-to-serine exchange at amino acid position 294. The mutation (S294) mapped within motif 3, one of three regions sharing significant homology with other class II aminoacyl-tRNA synthetases (Eriani *et al.*, 1990*a*). Based on the specific phenotype of the resistance mutation as well as on structural considerations, we proposed that residue 294, and hence motif 3, participate in the formation of the phenylalanine binding site on PheRS (Kast & Hennecke, 1991). This hypothesis was further tested in the present work.

One aspect in this report deals with the idea that specific aromatic residues in the enzyme might be involved in the binding of phenylalanine. Although the tertiary structure of PheRS and the topology of its binding sites are not known, an examination of the phenylalanine binding site, in particular that for the aromatic ring, appears possible because the amino acid at position 294 may be in contact with the *para*-position of the substrate's aromatic ring (Kast & Hennecke, 1991). The two amino acids flanking position 294 are phenylalanines (Fig. 1; Fayat *et al.*, 1983). As a working hypothesis we considered a model in which the aromatic rings of residues 293 and 295 could interact with the aromatic side chain of the substrate phenylalanine. The model implied that residues 293 to 295 form part of a hydrophobic pocket for phenylalanine.

Perpendicularly interacting aromatic residues may contribute substantially to the stabilization of a protein structure (Burley & Petsko, 1985; Singh & Thornton, 1985; Burley & Petsko, 1986). Apart from this, aromatic amino acids are often found to participate in the formation of a hydrophobic pocket designed to bind an aromatic substrate (Hangauer *et al.*, 1984). Aromatic binding sites have been determined in carboxypeptidase A (Hartsuck & Lipscomb, 1971), thermolysin (Kester & Matthews, 1977; Weaver *et al.*, 1977), chymotrypsin (Steitz *et al.*, 1969; Ringe *et al.*, 1985), subtilisin (Robertus *et al.*, 1972; Estell *et al.*, 1986), Trp repressor (Zhang *et al.*, 1987) and tyrosyl-tRNA synthetase (Brick & Blow, 1987; Brick *et al.*, 1989). In all cases, the binding pocket for the aromatic ring is mainly formed by hydrophobic amino acids including aromatic residues. The excellent conservation of

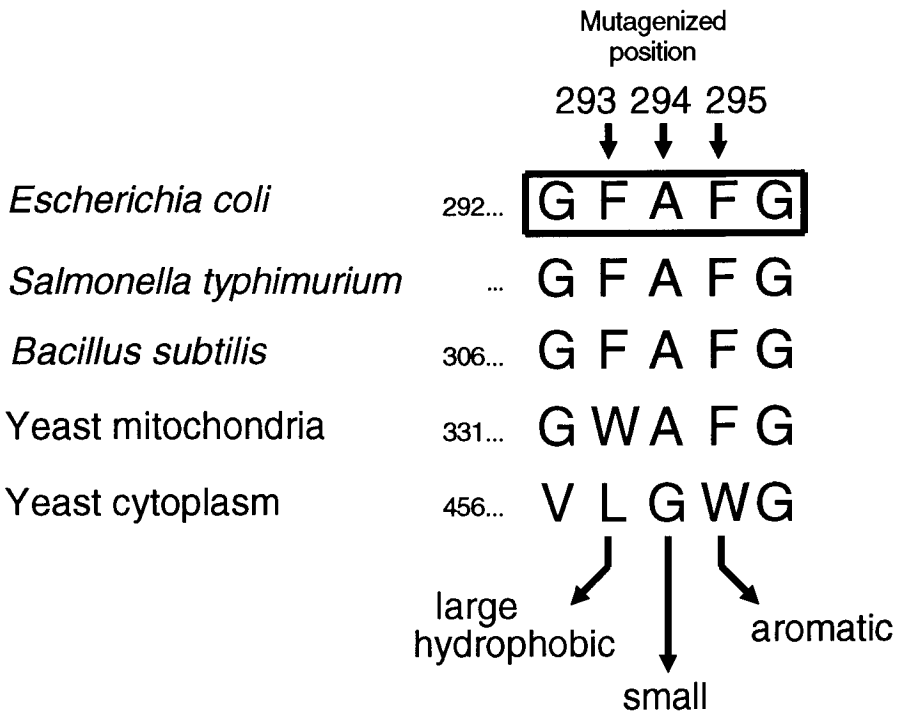


Figure 1. Consensus features of aligned amino acid sequences of PheRS small subunits from different organisms around position 294 (*E. coli*). Amino acids are written in one-letter-code. Numbering of sequences and references are as indicated in the preceding paper (Kast & Hennecke, 1991). Residues of *E. coli* PheRS mutagenized in this work are specified by arrows.

the region around position 294 in PheRS of other organisms (Fig. 1; Kast & Hennecke, 1991) further prompted us to investigate the effects that occurred after replacing the phenylalanines at positions 293 and 295 by other amino acids.

A second aspect in the analysis of the phenylalanine binding site concerned the substrate specificity. The specificity of an enzyme for a substrate is given by the structural complementarity of the binding site to that substrate in its transition state of the enzymatic reaction (Fersht, 1985*b*). An exchange of amino acid residues involved in substrate binding may therefore alter the substrate specificity of that enzyme. Experiments of that kind were performed extensively with serine proteases of known tertiary structure such as trypsin (Craik *et al.*, 1985) and subtilisin (Estell *et al.*, 1986; Wells *et al.*, 1987*a*; Wells *et al.*, 1987*b*). The results confirmed that optimal and specific fitting of substrates into binding pockets depend on suitable electrostatic, hydrophobic and steric interactions as well as on hydrogen bonding (Fersht, 1985*b*). Whilst these studies dealt with proteases of comparably poor specificity (Bosshard, 1976), the fidelity of substrate recognition and specifici-

ty of aminoacylation in the 20 aminoacyl-tRNA synthetases must be much higher in order to maintain the accuracy of cellular translation. Moreover, elaborate editing and proofreading mechanisms have evolved in this enzyme family to avoid frequent misincorporation of amino acids (Fersht, 1985*b*; Freist, 1989; Jakubowski, 1990). Therefore, it seemed more difficult to alter the substrate specificity of aminoacyl-tRNA synthetases simply by changing single amino acid residues. Nevertheless, aminoacyl-tRNA synthetase mutants with altered substrate specificity exist. In the preceding paper (Kast & Hennecke, 1991), a PheRS mutation was described that caused exclusion of the substrate analogue *p*-F-Phe from the enzymatic reaction. In the present work, we wished to test whether the substrate specificity of PheRS can be changed further by replacing the residue at position 294 with other amino acids. The results obtained supported the notion that residue 294 is involved in the Phe binding site, and showed that the mechanism postulated for discrimination against *p*-F-Phe (Kast & Hennecke, 1991) could be extended to other substrates.

3. Materials and Methods

3.1. Bacterial strains and plasmids

The bacterial strains used in this work and the plasmids pKSB1-W, pKSB1-M4S, pKSB2-W and the vector pBLS were described in the accompanying paper (Kast & Hennecke, 1991). Other plasmids constructed in this work are (depending on the insert orientation) of the types pKSB1-MnZ or pKSB2-MnZ (Fig. 2; analogous to the pKSB1 and pKSB2 -plasmids described by Kast & Hennecke, 1991), in which n indicates the last digit of the mutagenized position (293, 294 or 295) of the PheRS α subunit and Z denotes the newly introduced amino acid in one-letter code. These plasmids are listed in Table 1 together with the oligonucleotides used to create the mutation.

3.2. Substrates and analogues

The amino acid substrates and phenylalanine analogues used are shown later in Fig. 7. *p*-Fluoro-L-phenylalanine (used for data in Table 5) and proteinogenic amino acids were from Fluka Chemie AG (Buchs, Switzerland), and *p*-fluoro-D,L-phenylalanine (*p*-F-Phe) from Sigma Chemicals (St. Louis MO, USA). *o*- and *m*-Fluoro-D,L-phenylalanines, 3-fluoro-D,L-tyrosine and all six di-fluoro substituted D,L-phenylalanines were kindly provided by P. Babczinski (Bayer AG, Leverkusen, Germany). *p*-Deutero-, *p*-chloro-, *m*-chloro- and *p*-bromo- derivatives of L-phenylalanine and β-thien-2-yl-L-alanine were synthesized by Sergio Cantoreggi from D. Seebach's group (Laboratorium für Organische Chemie, ETH Zürich; Seebach *et al.*, 1989). ^{14}C-labelled substrates were described by Kast & Hennecke (1991).

3.3. Microbiological and recombinant DNA techniques, and DNA sequence analysis

The procedures and materials used (*e. g.* sequencing primers PEKA10 and PEKA2S) were described by Kast & Hennecke (1991).

Table 1: *Oligonucleotides and corresponding mutations*

Oligonucleotide (-mixture)			Mutated codons[a](mutation)[b]	aa-Sequence[a]	Plasmid
Name	**Size**	**Sequence[a]**		**292 293 294 295 296[b]**	**pKSB1/2-**
Wild-type			[877]-TTC GCC TTC-[885]	...GlyPheAlaPheGly...	-W
MA-ST	20mer	5'-CTCTGGTTTC(T/A)CCTTCGGGA-3'	GCC→TCC (880)	...GlyPheSerPheGly...	-M4S[c]
			GCC→ACC (880)	...GlyPheThrPheGly...	-M4T
MA-GV	20mer	5'-CTCTGGTTTCG(G/T)CTTCGGGA-3'	GCC→GGC (881)	...GlyPheGlyPheGly...	-M4G
			GCC→GTC (881)	...GlyPheValPheGly...	-M4V
MA-C	20mer	5'-CTCTGGTTTCTGCTTCGGGA-3'	GCC→TGC (880/1)	...GlyPheCysPheGly...	-M4C
MF1-GV	20mer	5'-TACTCTGGTG(G/T)CGCCTTCGG-3'	TTC→GGC (877/8)	...GlyGlyAlaPheGly...	-M3G
			TTC→GTC (877)	...GlyValAlaPheGly...	-M3V
MF1-YW	20mer	5'-TACTCTGGTT(A/G)(C/G)GCCTTCGG-3'	TTC→TAC (878)	...GlyTyrAlaPheGly...	-M3Y
			TTC→TGG (878/9)	...GlyTrpAlaPheGly...	-M3W
			TTC→TGC (878)	...GlyCysAlaPheGly...	-M3C
MF2-GV	20mer	5'-GGTTTCGCCG(G/T)CGGGATGGG-3'	TTC→GGC (883/4)	...GlyPheAlaGlyGly...	-M5G
			TTC→GTC (883)	...GlyPheAlaValGly...	-M5V
MF2-YW	20mer	5'-GGTTTCGCCT(A/G)(C/G)GGGATGGG-3'	TTC→TAC (884)	...GlyPheAlaTyrGly...	-M5Y
			TTC→TGG (884/5)	...GlyPheAlaTrpGly...	-M5W
			TTC→TGC (884)	...GlyPheAlaCysGly...	-M5C
			TTC→TAG (884/5)	...GlyPheAla	-M5
2MF-GV	26mer	5'-TACTCTGGTG(T/G)CGCCG(T/G)CGGGATGGG-3'	TTC→GGC, TTC→GGC (877/8, 883/4)	...GlyGlyAlaGlyGly...	-M3G5G
			TTC→GGC, TTC→GTC (877/8, 883)	...GlyGlyAlaValGly...	-M3G5V
			TTC→GTC, TTC→GGC (877, 883/4)	...GlyValAlaGlyGly...	-M3V5G
			TTC→GTC, TTC→GTC (877, 883)	...GlyValAlaValGly...	-M3V5V
2MF-W	26mer	5'-TACTCTGGTTGGGCCTGGGGGATGGG-3'	TTC→TGG, TTC→TGG (878/9, 884/5)	...GlyTrpAlaTrpGly...	-M3W5W

[a] Underlined nucleotides or amino acids indicate the point mutations introduced.

[b] Numbering of nucleotides and amino acids is as described in Kast & Hennecke (1991).

[c] Plasmid pKSB1-M4S was used previously (Kast & Hennecke, 1991).

3.4. *In vivo* plate tests

In vivo tests for complementation of strain KA2 at high temperature and analysis of Phe-analogue resistance of RR28 strains carrying a *pheS* gene on a plasmid were done as described in the preceding paper (Kast & Hennecke, 1991), unless stated otherwise. In the plate tests displayed in Table 5, the L-enantiomers of the substrates were used at a concentration of 0.55 mM which corresponded to that of the L-form in the standard test with racemic *p*-F-Phe (200 µg/ml).

3.5. Oligonucleotide-directed mutagenesis

To change single or multiple codons in wild-type *pheS*, the oligonucleotide-directed mutagenesis method of Nakamaye & Eckstein (1986) was applied; it makes use of *in vitro* selection of the mutagenized strand, synthesized in the presence of a thionucleotide which renders this strand resistant against degradation by specific restriction endonucleases. Mutagenic oligonucleotides (Table 1) were synthesized in an Applied Biosystems DNA Synthesizer model 380B with mixed base addition at suitable positions to create up to four different mutations in one redundant synthesis run. Purification of oligonucleotides was carried out on a Pharmacia FPLC device by using a reversed phase column (ProRPC HR 5/10; Pharmacia, Uppsala, Sweden) and a gradient in 100 mM triethylamine acetate, pH 7.0, from 10 % to typically 20 % acetonitrile. Under these conditions, the individual oligonucleotides present in a mixture could usually be separated from each other due to sequence-specific differences in hydrophobicity. Mutagenesis reactions were carried out with pools of oligonucleotides or, if needed, with single FPLC fractions by using the kit (version 1) and detailed protocols from Amersham (UK). As template, single strand DNA of plasmid pKSB2-W (containing the wild-type *pheS* gene) was used which was obtained after helper phage superinfection of plasmid-carrying TG1 cells (Kast & Hennecke, 1991).

From candidate plasmid clones with mutated *pheS* genes, single strand DNA was isolated and analyzed by sequencing the mutagenized region with primer PEKA10. The average yield of clones containing the desired mutation was slightly above 50 %. Correctly mutated fragments (174 bp *Aat*II/*Bst*BI) were excised from the mutagenized plasmid (pKSB2-type) and inserted into the expression plasmid (pKSB1-type) as shown in Fig. 2. The absence of secondary mutations that could have occurred during mutagenesis and cloning was checked by total sequencing

of the transferred 174 bp fragment using primer PEKA2S.

3.6. Preparation and analysis of protein extracts

Preparation of protein S140 extracts and determination of protein concentration by the Lowry method were performed as described in the accompanying paper (Kast & Hennecke, 1991). The content of PheRS β and α subunits was monitored by Western-blot analysis (Burnette, 1981) after separation of protein samples on a 10 % polyacrylamide gel containing 0.1 % SDS (Laemmli, 1970). Proteins were either stained with Coomassie Blue or transferred to a nitrocellulose membrane (Hybond-C; Amersham, UK). For gel electrophoresis and blotting, the Mini Protean II Dual Slab Cell and the corresponding transfer cell from Bio-Rad (Richmond CA, USA) were employed. As primary antibodies, rabbit antisera (γ-immunoglobulin fraction) raised against wild-type PheRS or islolated α subunits (Hennecke *et al.*, 1977) were used. The immunological staining of the membrane was done essentially as described in the Immun-Blot Assay protocol (1988) from Bio-Rad (Richmond CA, USA) using goat anti-rabbit IgG/alkaline phosphatase conjugate as secondary antibody. As standards, a prestained protein size marker (Bio-Rad; low-range protein standard, No 161-0305) and purified *E. coli* PheRS protein (kindly provided by H. Sternbach, Max-Planck-Institut, Göttingen, Germany) were used. Stained gels and membranes were scanned in a Molecular Dynamics Computing Densitometer model 300A (Sunnyvale CA, USA).

3.7. Determination of PheRS activity *in vitro*

The estimated errors of the individual assays described below are indicated in percent of the normal wild-type value. The standard assay was described by Kast & Hennecke (1991). Initial velocities of aminoacylation (expressed in nmol Phe attached to tRNA per min and mg protein) in S140 extracts from KA2 cells harbouring mutant *pheS* plasmids were determined at 28°C, and also at 40°C and 45°C (with or without preincubation of extracts at 1 mg/ml protein concentration for 20 min at 45°C). Values of certain thermosensitive mutants (marked in the Figures) could not be determined accurately (within 5-10 % error) at high temperatures. For rough approximations of apparent K_M values, Phe concentrations of 4 μM, 20 μM and 80 μM were used. Each data point was measured twice (incubation for 5 or 10 min at 28°C) and individual backgrounds were subtracted for each Phe concentra-

tion. The results were analyzed by Lineweaver-Burk plots (Fersht, 1985*a*); the error was estimated to be 25 %. From these K_M data, approximate apparent k_{cat} values were deduced from the Michaelis-Menten equation

$$k_{cat}=v_i\,(K_M + [S])\,/\,[E]_o\,[S] \qquad \text{(Fersht, 1985}a\text{)}$$

using initial velocities (v_i) at 28°C, phenylalanine concentration [S]=20 µM and a total enzyme concentration $[E]_o$ which was estimated by comparison (of the β subunit band) to known PheRS standards on Western blots.

The effect of a change in substrate specificity of mutant PheRSs was assessed by performing the standard assay (5 or 10 min incubation at 28°C) in the presence of a 50-fold excess (corresponding to 1 mM of the L-enantiomer) of one or more potentially competing substrates. Differential competition effects of *p*-deutero-L-phenylalanine compared to the natural substrate were analyzed by titrating both substances over a concentration range of 0.5 µM to 1 mM. All competition assays were carried out at least twice. The value of remaining [^{14}C]Phe incorporation was compared to the activity in the absence of a potentially competing substrate (=100 %). Deviations among independent measurements were generally less than 10-15 % of the value for non-competed incorporation.

4. Results

4.1. Experimental strategies for oligonucleotide-directed mutagenesis, and precautions made in the subsequent analysis of mutants

The purpose was to specifically exchange the amino acids in positions 293, 294 and 295 of the PheRS α subunit, a region suspected to be in close contact with the aromatic ring of the phenylalanine substrate (see Introduction). The decision for these amino acid replacements was also based on the similarity in that region within PheRS sequences from different origins. Figure 1 shows the consensus features of residues at positions 293, 294 and 295, being large hydrophobic, small, and aromatic, respectively. Consequently, the amino acid replacements in *E. coli* PheRS were designed to systematically test and alter this conserved pattern. The strategy for mutagenesis and subsequent analysis of mutant α subunits is presented in Fig. 2. Mutagenic oligonucleotides (or oligonucleotide mixtures), the resulting mutant *pheS* gene plasmids and the exchanged amino acids are listed in Table 1.

To measure the activity of mutated PheRS *in vivo* and *in vitro*, two peculiarities of the protein were taken into account: (i) Since PheRS is an essential enzyme, mutant plasmids had to be transformed into a conditional mutant. The thermosensitive *E. coli* strain KA2 was chosen as it showed no PheRS activity *in vitro* even at 28°C (Kast & Hennecke, 1991). (ii) To avoid an interference of the chromosomally encoded thermosensitive α subunit of strain KA2 in the assays, the gene dosage for the proteins to be tested was drastically increased by introducing the *pheS* genes on a multicopy plasmid; in addition, the genes were expressed from an induced *lac* promoter. Nevertheless, the plasmidial *pheS* expression did not affect the amount of PheRS holoenzymes produced because both α and β subunits were needed to form an active enzyme and the chromosomally encoded β subunit was limiting (see Discussion).

Specific enzyme activities in the aminoacylation tests were calculated relative to protein concentrations in the S140 extracts. This was justified on the basis of the high reproducibility of specific activity values (Kast & Hennecke, 1991) and was further validated by Western-blot analyses (Fig. 3) which showed that the con-

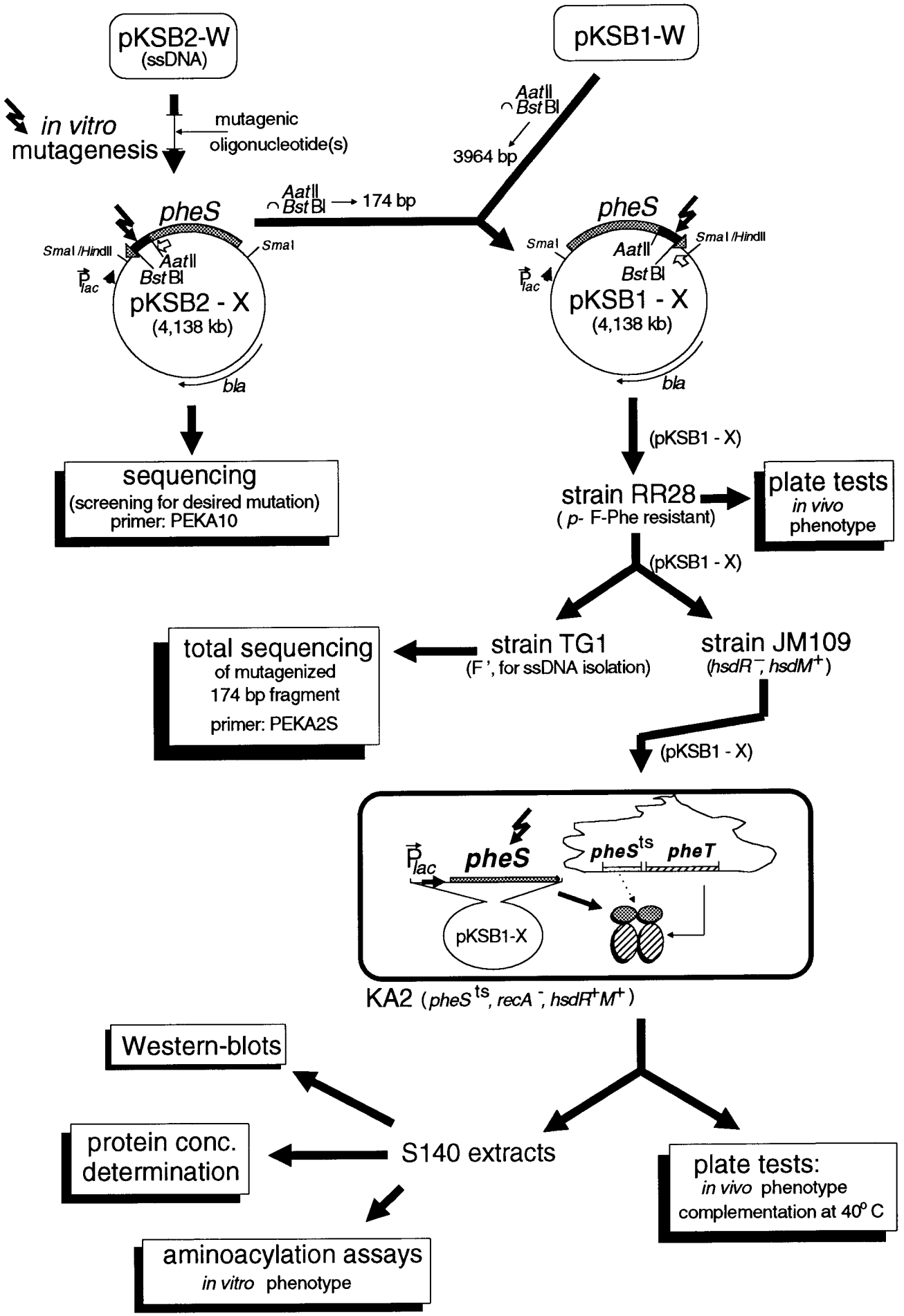

pKSB2-W
(ssDNA)
pKSB1-W
in vitro
mutagenesis
mutagenic
oligonucleotide(s)
AatII
BstBI
3964 bp
AatII
BstBI
174 bp
pheS
SmaI /HindII
SmaI
AatII
BstBI
Plac
pKSB2 - X
(4,138 kb)
bla
pheS
SmaI
SmaI /HindII
AatII
BstBI
Plac
pKSB1 - X
(4,138 kb)
bla
sequencing
(screening for desired mutation)
primer: PEKA10
(pKSB1 - X)
strain RR28
(p- F-Phe resistant)
plate tests
in vivo
phenotype
(pKSB1 - X)
total sequencing
of mutagenized
174 bp fragment
primer: PEKA2S
strain TG1
(F ', for ssDNA isolation)
strain JM109
(hsdR⁻, hsdM⁺)
(pKSB1 - X)
Plac
pheS
pheS^ts
pheT
pKSB1-X
KA2 (pheS^ts, recA⁻, hsdR⁺M⁺)
Western-blots
protein conc.
determination
S140 extracts
plate tests:
in vivo phenotype
complementation at 40° C
aminoacylation assays
in vitro phenotype

Figure 2. Experimental strategy for oligonucleotide-directed mutagenesis of *pheS* genes and subsequent analysis of mutant PheRSs. Single strand DNA of pKSB2-W carrying the wild-type *pheS* gene was mutagenized using mutagenic oligonucleotides listed in Table 1 (see Materials and Methods). The resulting plasmids (pKSB2-X; in which X stands for the different mutants in Table 1) were isolated as single strands from strain TG1 and sequenced to screen for mutations in the region around codon 294. From plasmids containing the desired mutations, a 174 bp *Aat*II/*Bst*BI fragment (marked by a flash arrow) was excised and used to replace the corresponding fragment in the wild-type *pheS* plasmid pKSB1-W. Strain RR28 was transformed with the ligation mixture. The resulting constructs (pKSB1-X) were suited to express the mutant *pheS* genes from the *lac* promoter (P_{lac}) of pBluescript after addition of the inducer IPTG. Plasmids isolated from strain RR28 were transformed into strain TG1 and subsequently sequenced. In parallel, the pKSB1-X plasmids were modified by passage through the *Eco*K modification proficient strain JM109 and finally introduced into the tester strain KA2 which carries a *pheS* gene for a thermosensitive PheRS (Kast & Hennecke, 1991). The *in vivo* phenotypes of RR28 and KA2 cells transformed with pKSB1-X were analyzed on agar plates. Tests in RR28 allowed conclusions about substrate specificity; growth of transformed KA2 at high temperature indicated that the α subunit, encoded predominantly by the plasmidial *pheS* gene, was active (see Results). Western blots and *in vitro* assays of mutant PheRS activities were carried out with S140 protein extracts of KA2 transformed with pKSB1-X. An arrow in front of the mutagenized fragment indicates the position of the sequencing primers. *bla*, gene for β-lactamase conferring resistance to ampicillin; ts, thermosensitive; ssDNA, single stranded DNA. Only the relevant genotypes/phenotypes for the strains used are indicated; a complete listing is provided in the preceding paper (Kast & Hennecke, 1991).

centration of (activity limiting) β subunit was essentially the same in all S140 extracts, irrespective of strain type and presence or absence of (mutant) *pheS* genes on plasmids. By correlating the scanned intensities of β bands from S140 extracts with those of the PheRS standards, a typical concentration of 0.029 nmol PheRS/mg S140 extract was obtained (M_r=248'000; see Kast & Hennecke, 1991). Based on this and on the determinations of apparent K_M values (see next section), the calculated apparent k_{cat} for aminoacylation by wild-type PheRS was about 150/min (at 28°C). This value is in good agreement with the previously determined k_{cat} for aminoacylation by purified *E. coli* PheRS of 228/min (at 25°C, analogous assay conditions; Holler, 1980).

Figure 3 shows examples of S140 extracts, separated on 10 % SDS polyacrylamide gels, after Coomassie Blue and immuno-staining. Surprisingly, the α subunits of PheRS were not present at the expected high level when compared to controls without plasmid *pheS* genes. The amount of α subunit varied between 1.1 and 8.3 times the value of the host strain KA2. This phenomenon was observed with three different antisera raised against the α subunit or the PheRS holoenzyme (data not shown) and was supposed to be due to degradation of excess free α subunits. The relative α subunit levels did neither correlate with the PheRS activity/thermosensitivity in the S140 extracts nor to the mutation type, *e. g.* the mutant

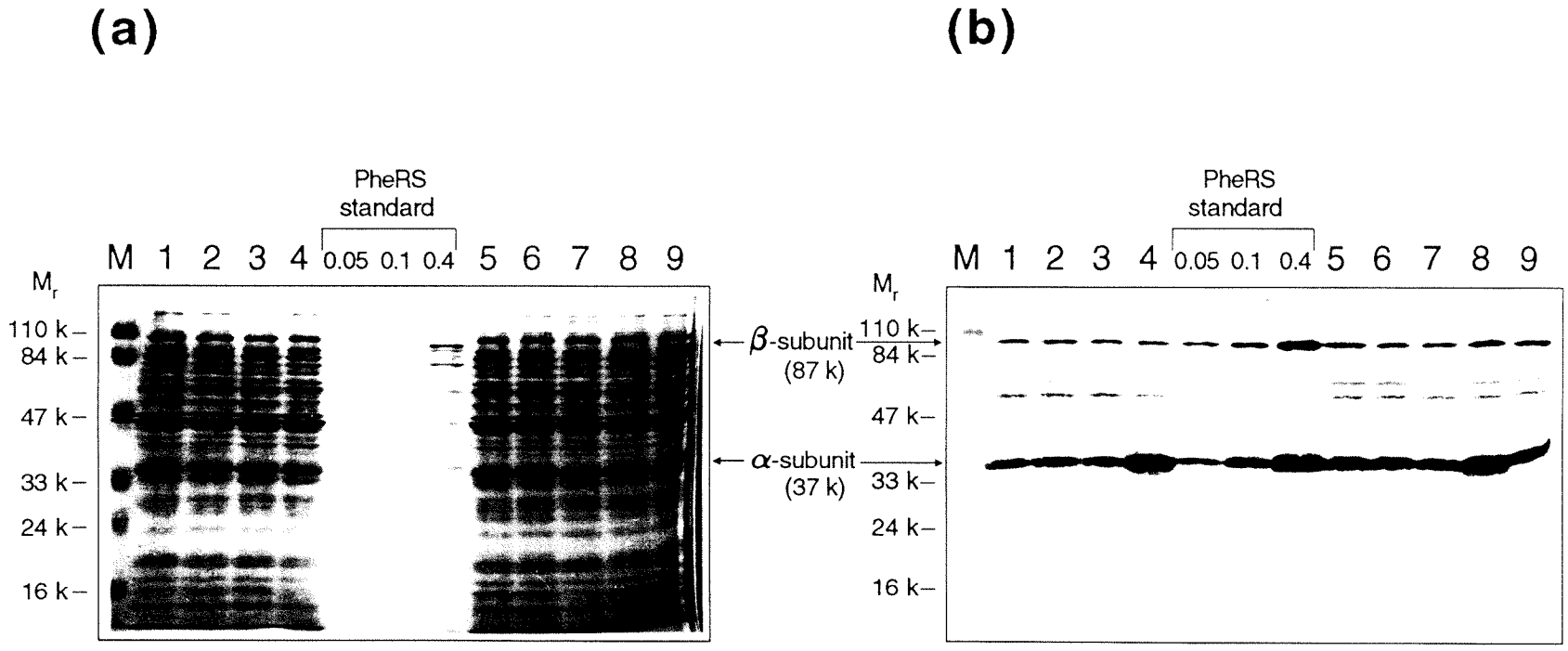


Figure 3. S140 cell extracts analyzed on 10 % SDS polyacrylamide gels. Coomassie Blue stained gel **(a)** and immuno-stained membrane **(b)** using anti-wild-type PheRS antibodies. Each sample contained 5 µg of total protein. Arrows indicate the positions of PheRS α and β subunits as deduced from a (prestained) molecular weight marker (lane 'M'; relative molecular weights are given on the left of the gel and membrane) and from partially purified PheRS. Individual subunit concentrations were estimated by comparisons to the known PheRS standards (0.05, 0.1 and 0.4 µg protein loaded) after densitometric scanning of gel and membrane bands. Lane 1, HB101 (wild-type strain); lane 2, KA2 (without plasmid); lane 3, KA2/pBLS (vector); lane 4, KA2/pKSB1-W (wild-type *pheS* on plasmid); lanes 5 and 6, KA2/pKSB1-M4S (Ser^{294}; 2 independent isolates); lane 7, KA2/pKSB1-M4G (Gly^{294}); lane 8, KA2/pKSB1-M4C (Cys^{294}); lane 9, KA2/pKSB1-M4T (Thr^{294}).

variants Val^{294} (low activity) and Trp^{293} (high activity) had overproduction factors of 8.3 and 1.3, respectively (see Discussion, section 5.1.).

To rule out the presence of significant amounts of hybrid enzymes (containing one chromosomally and one plasmid encoded α subunit; Hennecke, 1976), thermoinactivation tests were carried out. Hybrid tetramers containing the PheRS α subunit of the host strain KA2 were previously shown to be completely inactivated by preincubation at 45°C prior to the enzyme assay at 28°C (Hennecke, 1976), whereas the activity of the mutants described in this work was not (data not shown). Additional substrate competition experiments carried out at 45°C (after preincubation at that temperature; data not shown) with several of the partially thermolabile mutants gave results identical to those obtained at 28°C (see below).

4.2. Replacing phenylalanine at position 293

To test the hypothesis of an aromatic interaction between the presumed amino acid binding site of PheRS and the side chain of the phenylalanine substrate, Phe^{293} was exchanged against Gly, Cys, Val, Tyr or Trp. All mutants were active in complementing the thermosensitive strain KA2 (Fig. 4(a)), and *in vitro* activity at 28°C showed a maximal reduction of 30 % for the Gly^{293} and Trp^{293} variants relative to the wild-type Phe^{293} (Fig. 4(b)).

With a wild-type *pheS* gene in KA2, we obtained a K_M value for Phe of 7 µM which is in excellent agreement with the value of 6.6 µM obtained previously under similar conditions for a wild-type strain (Grüll *et al.*, 1979). This result was further confirmed with our PheRS wild-type strain HB101 (data not shown). Figure 4(c) shows that K_M values for Phe were not significantly affected by the mutations. Together with the other data in Fig. 4, this suggests that Phe^{293} (or an aromatic residue at that position) is not important for phenylalanine binding or PheRS activity.

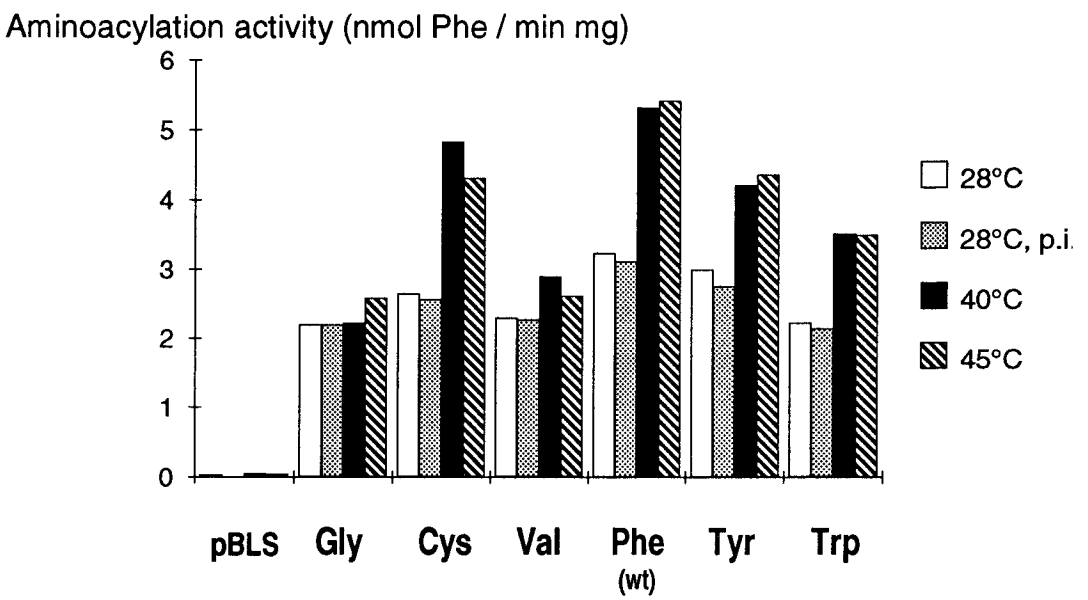
At higher temperature, the reduction of *in vitro* enzyme activities amounted to 58 % (Gly^{293}), as compared to the wild-type level (Fig. 4(b)). With the exception of Cys^{293}, there may be a tendency to favour a residue with a surface area comparable to Phe (especially at high temperature). The slightly thermosensitive behaviour of some mutants was not due to irreversible thermal denaturation, since preincubation of S140 extracts at 45°C (and subsequent cooling to 0°C) did not alter activity at 28°C (Fig. 4(b)). This feature was also apparent with all mutants analyzed below.

Figure 4. Phenotypic effects of amino acid exchanges at position 293 of the PheRS α subunit. The amino acid variants are arranged in the order of increasing surface area of the residue at the corresponding mutated position (Rose *et al.*, 1985). wt, wild type; pBLS, negative control (vector only). **(a)** The ability of plasmid borne *pheS* variants to complement the thermosensitive strain KA2 was scored from good (++), normal (+), reduced (±), marginal (-) growth to no growth at all (0). **(b)** Initial velocities of Phe aminoacylation by S140 extracts from KA2 cells carrying mutant *pheS* plasmids. The standard assay was carried out at the temperatures indicated. The values '28°C, p.i.' were obtained with S140 extracts preincubated at 45°C for 20 min and then cooled to 0°C prior to assaying at 28°C. The errors are estimated to be 5-10 % relative to the wild-type values. **(c)** Apparent k_{cat} and K_M values at 28°C calculated as described in Materials and Methods. For determination of k_{cat}, the average of initial velocities at 28°C (with and without preincubation) from (b) was taken. The values represent only rough approximations (estimated errors 25 %).

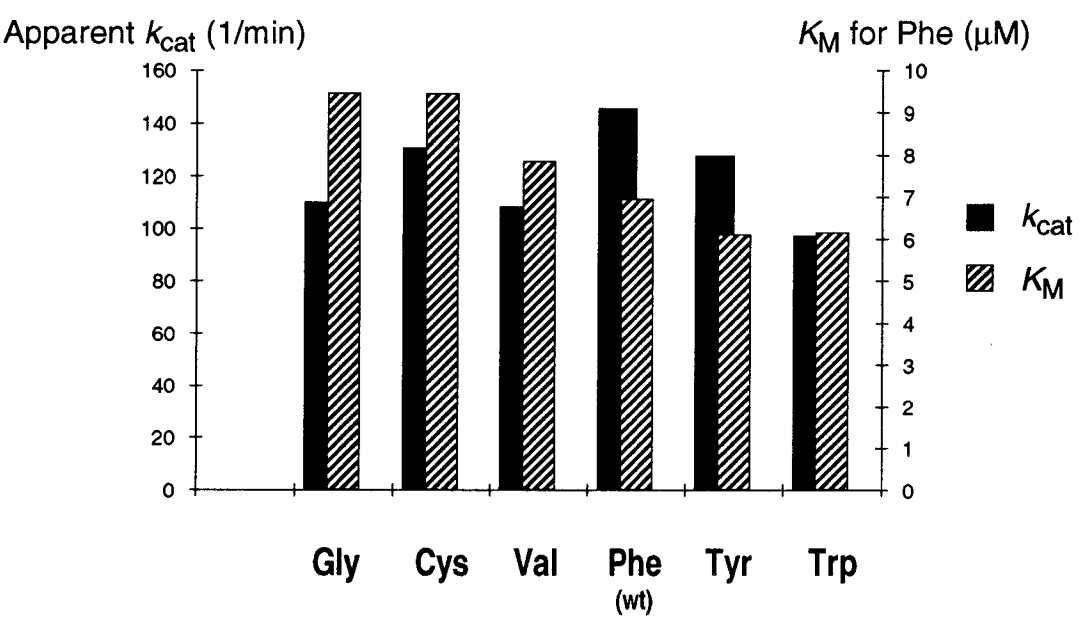
(a) *In vivo* activity: complementation of KA2 at 40°C

Postition 293	pBLS (vector)	Gly	Cys	Val	Phe (wt)	Tyr	Trp
Growth at 40 °C: rich medium; ampicillin	0	++	++	++	++	++	++

(b) Initial velocities of aminoacylation at different temperatures



(c) Approximate kinetic parameters at 28°C



4.3. Replacing phenylalanine at position 295

It was also tested, whether the conserved aromatic residue at position 295 (again a phenylalanine; Fig. 1) was essential for PheRS activity. The data in Fig. 5 demonstrate that the effect of point mutations at this position was more pronounced than with Phe^{293}. Mutant Gly^{295} could not complement strain KA2 (Fig. 5(a)) and showed low PheRS activity *in vitro* (19 %, 3 % and 0 % of wild-type values at 28°C, 40°C and 45°C, respectively; Fig. 5(b)). Mutations to Cys^{295} and Val^{295}, both non-aromatic residues, allowed the formation of complementing α subunits and exhibited nearly wild-type activity at 28°C (Fig. 5(a,b)). Exchanging Phe^{295} against Trp represents a fully tolerated alternative for *E. coli* PheRS. The other aromatic exchange to Tyr^{295} yielded wild-type activity *in vivo* and *in vitro* at 28°C, but showed, like Cys^{295} and Val^{295} variants, a strong thermosensitive behaviour at 40°C and 45°C (reduction of more than 67 % and 81 %, respectively, relative to wild type; Fig. 5(a,b)). The K_M values were not changed (Fig. 5(c)). Together with the fact that Phe^{295} could be replaced by other aromatic or hydrophobic residues without loosing enzyme activity (at 28°C), this spoke against a direct involvement of position 295 in substrate binding.

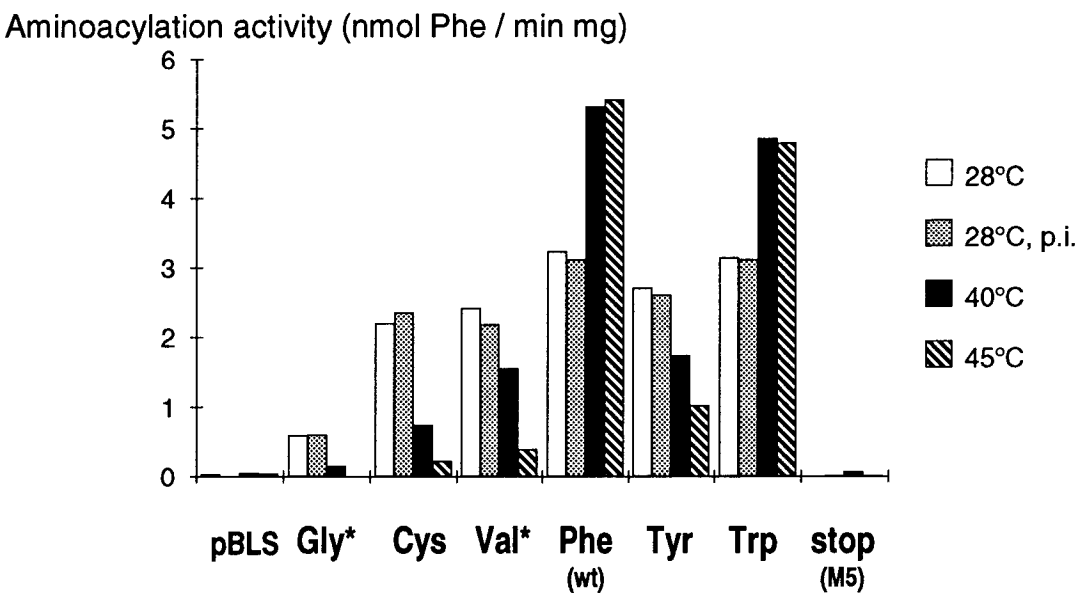
Insertion of a nonsense codon at position 295 (mutant M5) abolished PheRS activity completely (Fig. 5(a,b)). The absence of crossreacting material corresponding to this truncated α subunit in the S140 extract (not shown) suggested that this defective protein was rapidly degraded.

Figure 5. Phenotypic effects of amino acid exchanges at position 295 of the PheRS α subunit. For description of the Figure see legend to Fig. 4. An asterisk next to the replaced amino acid (in (b)) indicates that initial velocity of aminoacylation at 40°C and 45°C could not be determined accurately due to thermal denaturation during the assay. In these cases, the remaining aminoacylation velocity after 2 min was scored.

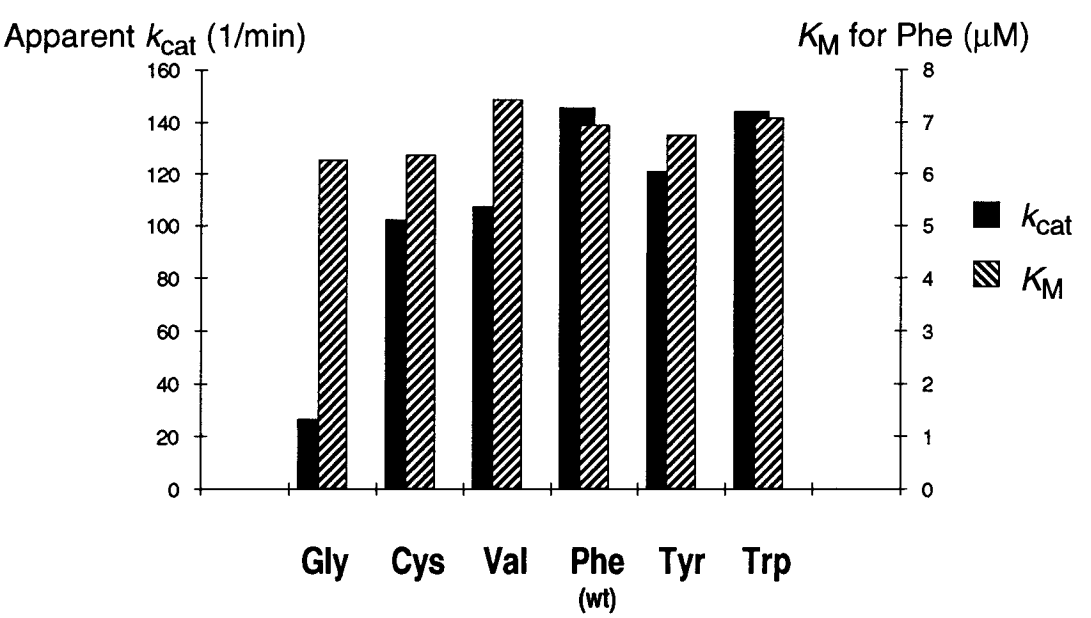
(a) *In vivo* activity: complementation of KA2 at 40°C

Postition 295	pBLS (vector)	Gly	Cys	Val	Phe (wt)	Tyr	Trp	stop (M5)
Growth at 40°C rich medium; ampicillin	0	–	++	++	++	++	++	0

(b) Initial velocities of aminoacylation at different temperatures



(c) Approximate kinetic parameters at 28°C



4.4. Replacing both phenylalanines at positions 293 and 295

To rule out the possibility that one of the two aromatic residues alone was sufficient for binding of the phenylalanine substrate, both Phe^{293} and Phe^{295} were exchanged simultaneously. The data in Fig. 6 confirmed that a glycine at position 295 was not tolerated as already seen in the previous section (of Fig. 5). Activity values were even more reduced *in vitro*, when Phe^{293} was additionally replaced by Gly or Val (to less than 12 %, 0 % and 0 % at 28°C, 40°C and 45°C, respectively; Fig. 6(b)). The fact that the double mutants Val^{293}/Val^{295} and Gly^{293}/Val^{295} could complement KA2 (Fig. 6(a)) and were active at 28°C (albeit with reduced rate) proved that aromatic residues at positions 293 and 295 are not at all required for PheRS activity and substrate binding. There were no significant alterations in the apparent K_M values for the double mutants, perhaps with the slight exception of the Trp^{293}/Trp^{295} variant (Fig. 6(c)).

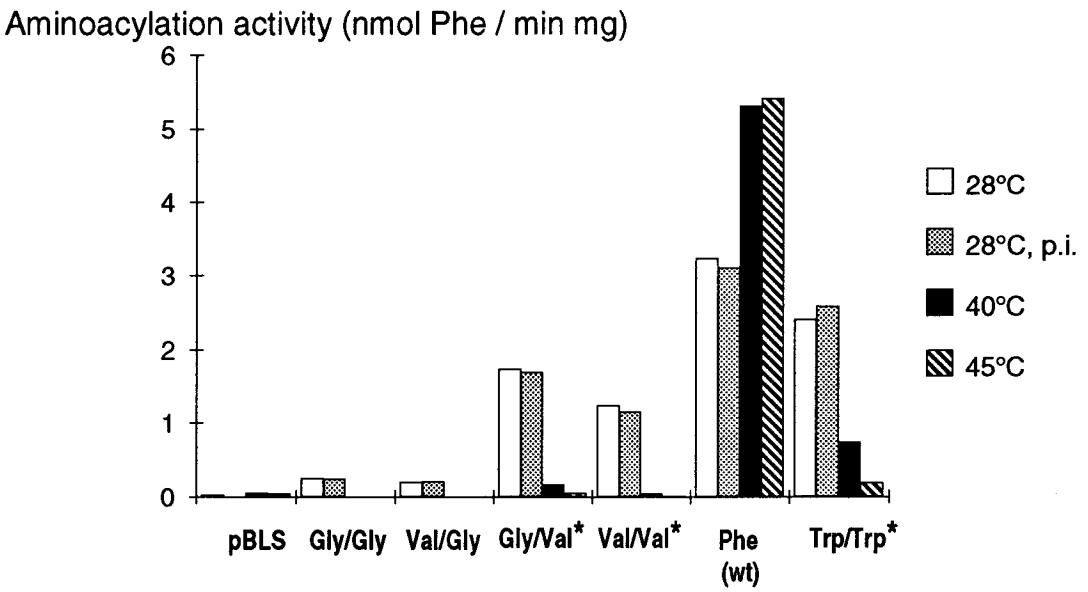
Taken together, the results described in this and the previous two sections make it seem unlikely that Phe^{293} and Phe^{295} interact directly with the phenylalanine substrate. However, the strongly reduced activities at 40°C and 45°C, also apparent with the double mutants (Fig. 6(b)), point to their role in maintaining and stabilizing the protein structure (see Discussion). This hypothesis is supported by the observation that certain mutants exhibited a pronounced decline of aminoacylation with time (data not shown). This effect appeared exclusively in assays carried out at high temperatures and was seen for all mutants having Gly or Val at position 295 as well as for the Trp^{293}/Trp^{295} double mutant.

Figure 6. Phenotypic effects of simultaneous amino acid exchanges both at position 293 and 295 of the PheRS α subunit. For description of the Figure see legends to Figs. 4 and 5.

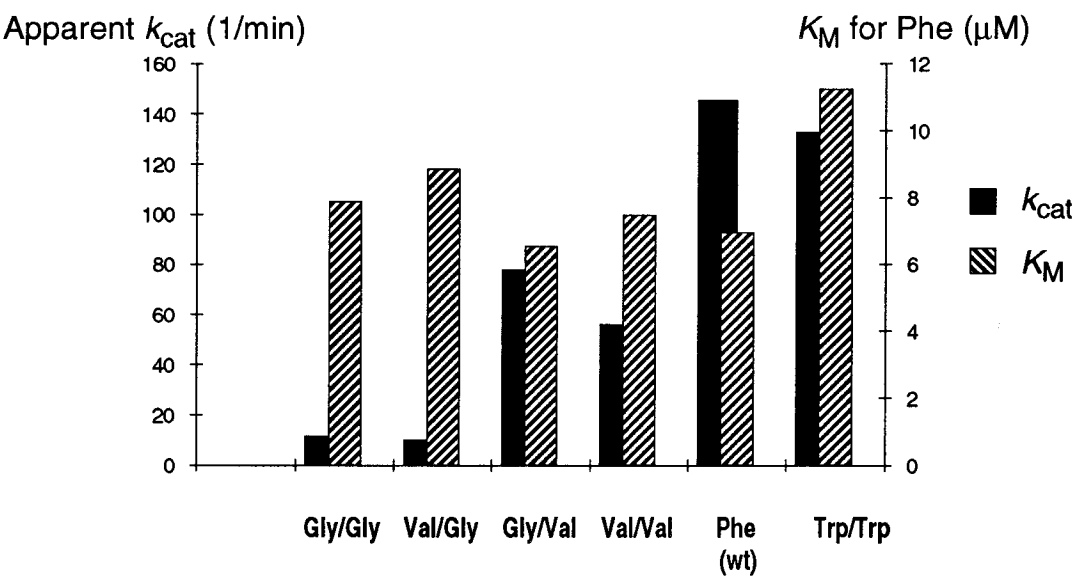
(a) *In vivo* activity: complementation of KA2 at 40°C

Position **293/295**	**pBLS (vector)**	**Gly/Gly**	**Val/Gly**	**Gly/Val**	**Val/Val**	**Phe (wt)**	**Trp/Trp**
Growth at 40 °C: rich medium; ampicillin	0	–	–	++	++	++	++

(b) Initial velocities of aminoacylation at different temperatures



(c) Approximate kinetic parameters at 28°C



4.5. Do replacements at positions 293 and 295 change the substrate specificity?

An independent means to screen for phenylalanine binding site mutants is the examination of the substrate (or substrate analogue) spectrum of the different PheRS variants. The plasmids carrying the mutant *pheS* genes were analyzed in *E. coli* strain RR28 for their ability to confer resistance or sensitivity to *p*-F-Phe (Table 2A). This *in vivo* test system was described in detail in the accompanying report (Kast & Hennecke, 1991); it is based on the fact that RR28 does not incorporate this toxic Phe analogue into cellular protein due to its mutant PheRS (mutation Ser^{294}) and therefore grows on *p*-F-Phe-containing plates. Table 2A shows that *pheS* genes encoding the Gly^{295} mutant form (single or double mutations) or the truncated M5 variant allowed growth of transformed RR28. This indicated that those α subunits were inefficient or inactive in RR28, similarly as in strain KA2. On the other hand, all mutants previously shown to possess a KA2-complementing *pheS* gene prevented growth of RR28. It was concluded, therefore, that they behave like the wild type with respect to sensitivity to *p*-F-Phe by accepting the analogue as substrate. Table 2B confirms these results by showing that the ratios of *p*-F-[^{14}C]Phe/[^{14}C]Phe incorporation into tRNA corresponded to that for wild-type PheRS.

Figure 7 shows a selection of phenylalanine analogues and other substrates that were analyzed *in vitro* for competition with [^{14}C]Phe incorporation catalysed by the mutant PheRSs. To limit the number of assays, we pooled substrates believed to not compete with Phe in the wild-type PheRS reaction (Fig. 7). When a pool exhibited competition with Phe in the test with a particular mutant, individual members of this pool were then tested separately. Table 2C shows relative remaining [^{14}C]Phe incorporation in the presence of a 50-fold excess of the L-enantiomer(s).

The effects of the presence of the different substrates and analogues on wild-type PheRS activity can be summarized as follows: Among the 19 non-cognate proteinogenic amino acids, only Tyr showed a slight tendency to compete with Phe

Figure 7. Substrates and substrate analogues tested with PheRS mutant enzymes. Competition tests were carried out in the presence of a 50-fold excess (*i. e.* 1 mM) with respect to the relevant L-enantiomer (displayed in the Figure) as described in Materials and Methods. Dashed boxes group amino acids that did not compete with wild-type PheRS and were therefore tested in pools (each substance at 1 mM concentration). *p*-Deutero-phenylalanine was only tested with the wild-type and the Ser^{294} mutant enzyme (see text).

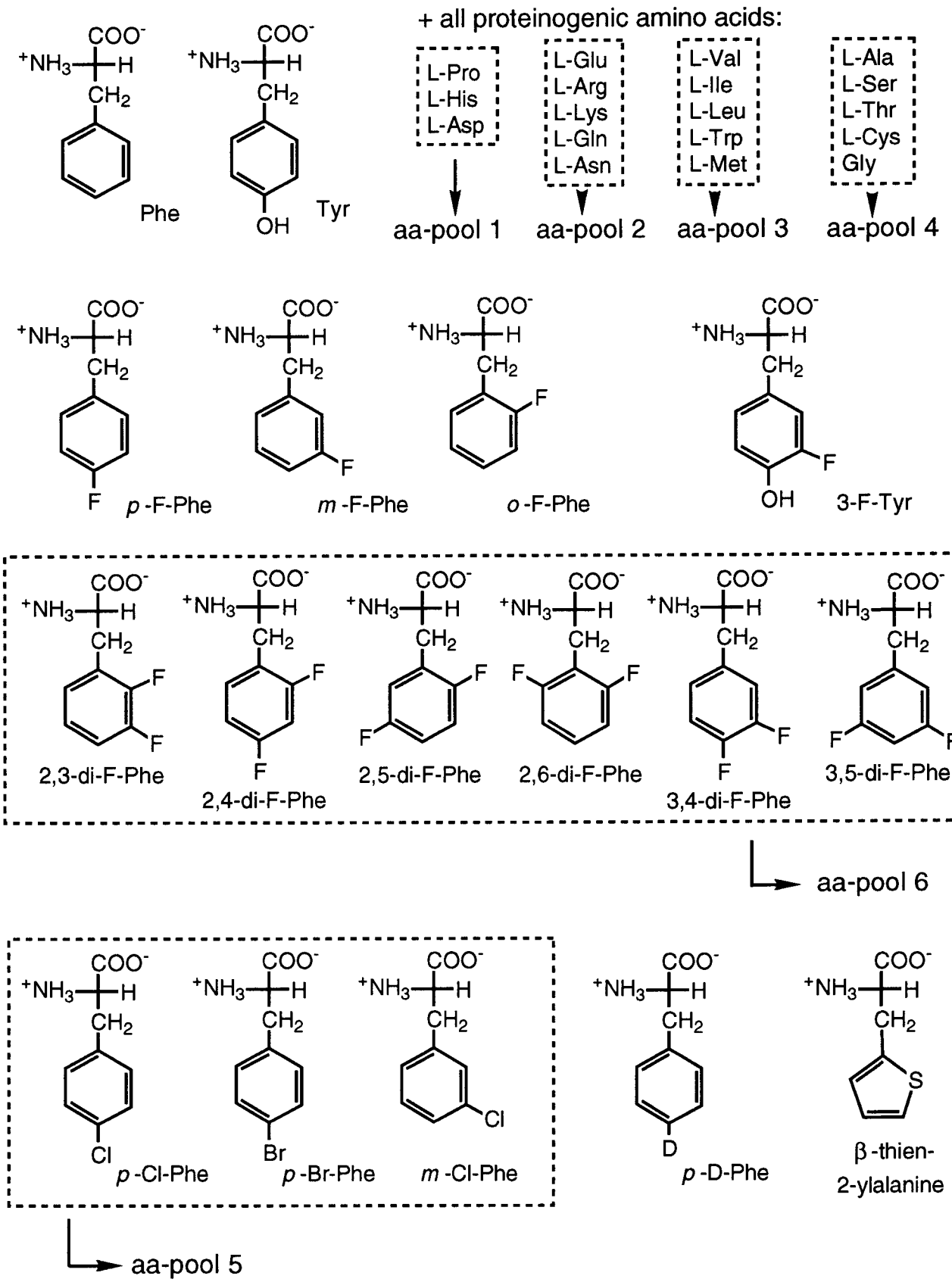

+ all proteinogenic amino acids:
Phe
Tyr
L-Pro
L-His
L-Asp
L-Glu
L-Arg
L-Lys
L-Gln
L-Asn
L-Val
L-Ile
L-Leu
L-Trp
L-Met
L-Ala
L-Ser
L-Thr
L-Cys
Gly
aa-pool 1
aa-pool 2
aa-pool 3
aa-pool 4
p-F-Phe
m-F-Phe
o-F-Phe
3-F-Tyr
2,3-di-F-Phe
2,4-di-F-Phe
2,5-di-F-Phe
2,6-di-F-Phe
3,4-di-F-Phe
3,5-di-F-Phe
aa-pool 6
p-Cl-Phe
p-Br-Phe
m-Cl-Phe
p-D-Phe
β-thien-2-ylalanine
aa-pool 5

Table 2: *Effect of substrates and substrate analogues on position 293/295 mutant PheRS activities*

	Wild type	Position 293 mutants					Position 295 mutants						Position 293/295 mutants				
A *In vivo* growth tests: mutant plasmids in strain RR28 (*p*-F-Phe resistant)[a]																	
Minimal medium, 37°C	Phe$^{293/295}$	Gly	Cys	Val	Tyr	Trp	Gly	Cys	Val	Tyr	Trp	stop (M5)	Gly/Gly	Val/Gly	Gly/Val	Val/Val	Trp/Trp
– *p*-F-Phe	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
+ *p*-F-Phe	–	–	–	–	–	–	±	–	–	–	–	+	±	±	–	–	–
B *In vitro p*-F-Phe / Phe incorporation[b]																	
Ratio	Phe$^{293/295}$	Gly	Cys	Val	Tyr	Trp	Gly	Cys	Val	Tyr	Trp	stop (M5)	Gly/Gly	Val/Gly	Gly/Val	Val/Val	Trp/Trp
p-F-Phe/Phe	0.54	0.46	0.42	0.45	0.47	0.50	0.31[c]	0.59	0.43	0.48	0.40	n.t.[c]	n.t.[c]	n.t.[c]	0.38	0.49	0.58

C *In vitro* competition of Phe incorporation by substrate analogues (in %)[d]

Substance	Phe[293/295]	Gly	Cys	Val	Tyr	Trp	Gly[c]	Cys	Val	Tyr	Trp	stop (M5)	Gly/Gly[c]	Val/Gly[c]	Gly/Val	Val/Val	Trp/Trp
Phe[e]	6	5	9	7	8	0	12	3	1	5	3	n.t.[c]	9	7	3	8	5
Tyr	93	97	91	103	88	88	91	88	91	90	91	n.t.[c]	85	87	92	89	94
aa-pool 1	94	98	101	103	99	96	99	98	104	100	101	n.t.[c]	96	100	97	102	93
aa-pool 2	101	99	105	107	107	99	108	103	109	102	106	n.t.[c]	91	123	100	104	98
aa-pool 3	97	98	100	102	100	96	100	100	99	99	100	n.t.[c]	94	106	96	98	89
aa-pool 4	100	98	103	116	104	99	104	101	106	103	103	n.t.[c]	99	119	99	103	91
p-F-Phe	26	23	28	40	28	21	40	30	34	30	31	n.t.[c]	24	37	29	36	34
m-F-Phe	53	51	61	65	50	47	63	59	56	52	55	n.t.[c]	49	52	47	55	62
o-F-Phe	71	78	82	92	68	68	72	71	75	69	73	n.t.[c]	73	93	71	72	79
3-F-Tyr	105	98	107	109	104	99	94	97	96	97	100	n.t.[c]	101	93	95	97	98
β-thien-2-ylalanine	57	60	70	69	65	58	58	71	62	59	68	n.t.[c]	50	62	61	61	75
aa-pool 5	97	95	98	105	106	108	106	101	108	106	98	n.t.[c]	101	107	98	102	100
aa-pool 6	85	74	82	84	83	85	81	79	85	79	81	n.t.[c]	73	91	76	81	81

[a] Conditions for the growth tests are described in detail in the accompanying paper (Kast & Hennecke, 1991). Single colonies were analyzed for growth: ++, good; +, normal; ±, reduced; –, marginal growth. The negative control (vector pBLS alone) gave a result identical to the position 295 stop codon insertion (M5).

[b] The ratio was obtained by using the corresponding initial velocities of aminoacylation at 28°C. The assays were as described in the accompanying report (Kast & Hennecke, 1991).

[c] Values that could not (n. t. = not tested) or only inaccurately be determined (error of 20 %) due to low absolute PheRS activities.

[d] The competition assays were carried out as described in Materials and Methods with an estimated error of 10 % - 15 %.

[e] Isotope dilution effect (positive control) by addition of a 50-fold excess of unlabelled Phe.

(about 90-95 % remaining Phe incorporation). This is consistent with a previous report (Gabius *et al.*, 1983*a*) in which activation and aminoacylation of Tyr by *E. coli* PheRS was quantified. Competition by *p*-F-Phe (a good PheRS substrate as manifested in Tables 2A & 2B) lowered Phe incorporation drastically to 25-30 %. *m*-F-Phe and *o*-F-Phe caused reductions to 50-60 % and 65-75 %, respectively. Both values were close to those obtained previously in similar competition experiments (Hennecke & Böck, 1975). β-Thien-2-ylalanine addition reduced Phe incorporation to 55-65 % (Table 2C). This substrate analogue was shown to be reasonably well aminoacylated by *E. coli* PheRS (Gabius *et al.*, 1983*a*). It was speculated by Gabius *et al.* (1983*a*) that the loss of binding energy due to the reduced ring size could be compensated partially by increased dispersion forces of the ring sulphur (Fersht & Dingwall, 1979). The latter four substrate analogues are known to exert their toxic action *in vivo* by being incorporated into cellular protein instead of Phe (Richmond, 1962; Shive & Skinner, 1963; Hortin & Boime, 1983). 3-F-Tyr and pool group 5 (*p*-Cl-Phe, *m*-Cl-Phe, *p*-Br-Phe) did not compete with the cognate substrate at the concentrations used. The large pool group 6, containing all di-fluoro substituted phenylalanines, showed a remaining Phe incorporation of around 80 %. As will be demonstrated later (Table 4) this effect was not due to a single substance but rather to small contributions of individual components.

The data obtained from the competition tests with the mutant α subunits did not reveal a significant shift in amino acid substrate specificity (Table 2C). This, together with the results from Tables 2A & 2B, again confirms that positions 293 and 295 are not directly part of a phenylalanine specificity pocket.

4.6. Mutating the alanine at position 294 affects Phe binding

In the preceding paper (Kast & Hennecke, 1991), an alanine-to-serine exchange at position 294 of a mutant PheRS α subunit was shown to be responsible for exclusion of the substrate analogue *p*-F-Phe from the enzymatic reaction. We therefore proposed the existence of steric interactions of residue 294 with the *para*-position of the aromatic ring of the substrate phenylalanine. To test this assumption directly, Ala^{294} was replaced by smaller or larger amino acids.

Figure 8(a) shows that residues with volumes larger than serine (Cys, Thr, Val) exhibited reduced KA2 complementation activities. The colonies containing the mutated gene for the Val^{294} enzyme were tiny at 40°C and apparently dead because, when restreaked, they were unable to grow even at 30°C. *In vitro* assays confirmed the drop in PheRS activity for the mutants Cys^{294}, Thr^{294} and Val^{294}, which was most pronounced at high temperature (Fig. 8(b)). The values were at 28°C: 25-30 %, at 40°C: 1-8 % and at 45°C: 0-2 % relative to wild type. In addition, these three mutants showed a sharp decline in aminoacylation with time *in vitro*, but only at high temperatures (data not shown). This behaviour, already described for certain position 295 mutants (section 4.4.), points to reduced protein stability (see Discussion). In contrast to variants Cys^{294}, Thr^{294} and Val^{294}, the decrease of aminoacylation with time was not observed with mutants Gly^{294} and Ser^{294} (data not shown). Figure 8(b) shows that both mutants were less active at high temperatures (40°C and 45°C) but the residual activity was still at least 31 % (Gly^{294}; 45°C). The activity at 28°C was only reduced to about 75 % of the wild-type level. It seemed as if those two mutations did not significantly destabilize the protein.

Significant differences appeared when the kinetic parameters of the position 294 variants were determined at 28°C (Fig. 8(c)). The binding affinity for phenylalanine was lowered in all residue 294 mutants, as indicated by increases in the apparent K_M values relative to the wild-type Ala^{294} value (in the following defined as 100 %). Removing the methyl group of Ala^{294} resulted in a reduction of the apparent reaction rate k_{cat} to 85 %, whereas the apparent K_M was increased to 164 % (mutant Gly^{294}). An increase in size of the side chain of Ala^{294} to Ser also caused an increase in K_M, even by a factor of four. Obviously, for an optimal binding affinity for Phe, the Ala side chain at position 294 was necessary; a smaller group (Gly^{294}) proved to be less efficient. The same was true for the larger amino acid Ser, which may sterically interfere with optimal binding of phenylalanine (see Discussion). This argument also seemed to be valid for the larger side chain vari-

ants Cys^{294}, Thr^{294} and Val^{294} through which the K_M was increased to 247 %, 179 % and 117 % of the wild-type value, respectively (Fig. 8(c)).

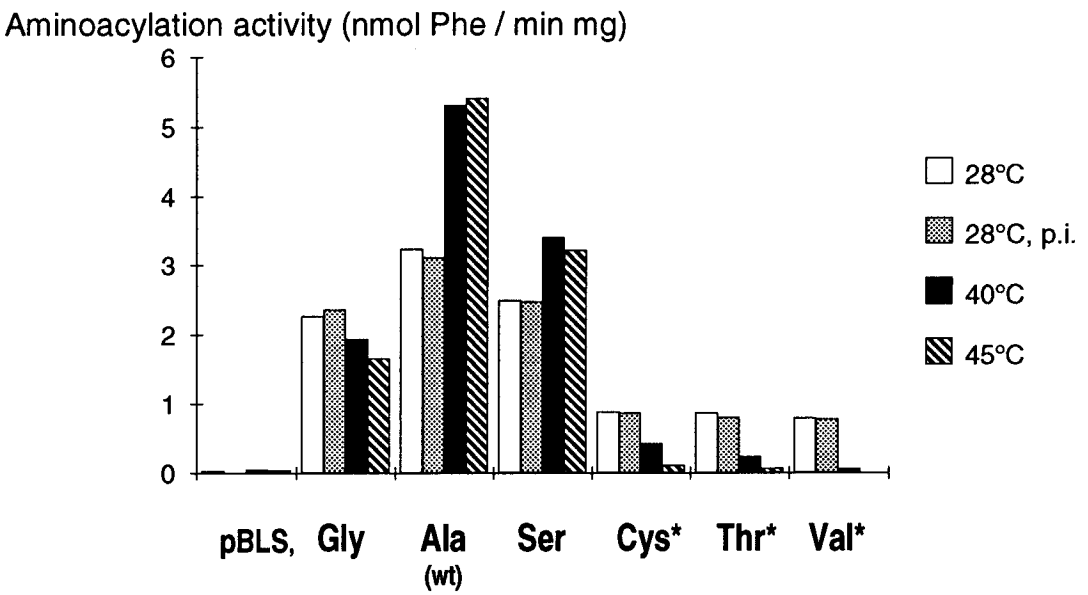
Although k_{cat} of mutant Ser^{294} was increased to 142 %, the catalytic efficiency (k_{cat}/K_M) was reduced to 34 % of the wild-type level. As shown in Fig. 8(c) for mutants Cys^{294}, Thr^{294} and Val^{294}, the apparent k_{cat} values were lowered to 29 %, 31 % and 18 % yielding catalytic efficiencies of 12 %, 18 % and 16 %, respectively. Obviously, the decrease in k_{cat} became dramatic as soon as a residue with a volume larger than Ser was inserted. A lower k_{cat} could indicate either a general distortion of the overall conformation of the catalytic site or possible protein stability problems under our assay conditions (see Discussion).

Figure 8. Phenotypic effects of amino acid exchanges at position 294 of the PheRS α subunit. The amino acid variants are arranged according to increasing volume (Chothia, 1984). For further description of the Figure see legends to Figs. 4 and 5.

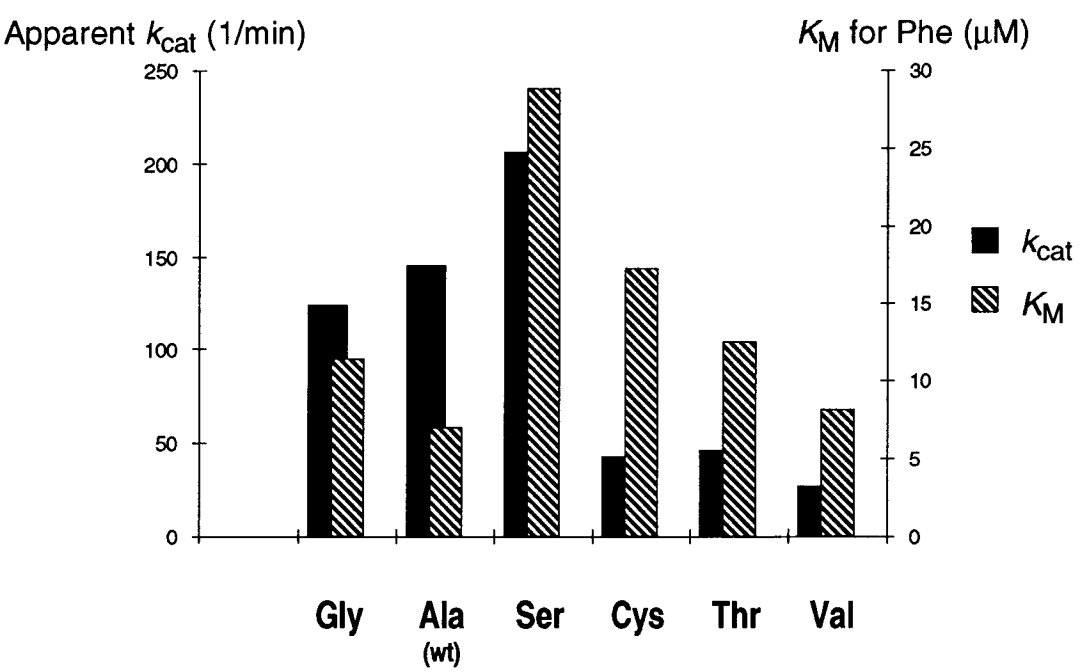
(a) *In vivo* activity: complementation of KA2 at 40°C

Postition 294	pBLS (vector)	Gly	Ala (wt)	Ser	Cys	Thr	Val
Growth at 40°C: rich medium; ampicillin	0	++	++	++	+	+	±

(b) Initial velocities of aminoacylation at different temperatures



(c) Approximate kinetic parameters at 28°C



4.7. Position 294 is a major specificity determinant

The differential effects of position 294 mutants on phenylalanine binding confirmed the idea that this residue may be in contact with the amino acid substrate. To obtain more details about the interactions with the substrate, effects of other substrates and substrate analogues on activity of residue 294 mutants were tested. The experiments were carried out with the substances shown in Fig. 7 and using assays described in section 4.5.

(a) p-*F-Phe incorporation and competition*

Initial velocities (at 28°C) of incorporation of [^{14}C]Phe and *p*-F-[^{14}C]Phe into tRNA were compared for the different position 294 mutants. The ratios of *p*-F-[^{14}C]Phe/[^{14}C]Phe incorporation (Fig. 9) revealed the following characteristics: mu-

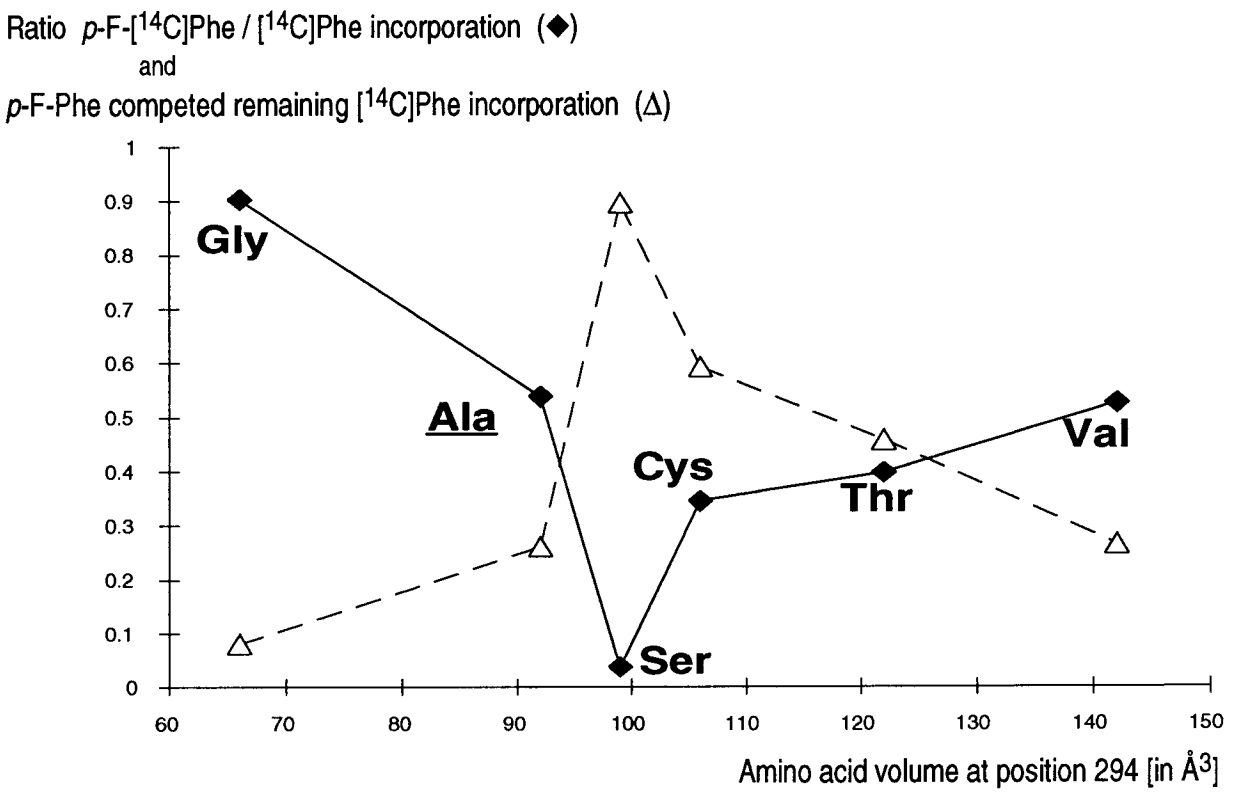


Figure 9. Relative *p*-F-Phe / Phe specificities of position 294 PheRS mutants. The ratios of initial velocities in independent aminoacylation assays with *p*-F-[^{14}C]Phe or [^{14}C]Phe as substrates are displayed together with the remaining [^{14}C]Phe incorporation in the presence of excess competing, unlabelled *p*-F-Phe (no competitor added equals 1). The assays were carried out at 28°C. The values are plotted against the volume of the corresponding position 294 amino acid (Chothia, 1984). The wild-type enzyme (Ala^{294}) is underlined.

tant Gly^{294} (ratio *p*-F-Phe/Phe = 0.90) incorporated *p*-F-Phe essentially as well as the cognate substrate Phe, in contrast to the wild type (Ala^{294}; ratio = 0.53). As expected (Kast & Hennecke, 1991), the mutant Ser^{294} did not incorporate *p*-F-Phe at a significant level (ratio ≤ 0.04). Mutants Cys^{294} and, to a lesser extent, Thr^{294} showed intermediate ratios (0.34 and 0.40, respectively) that ranged between wild type (Ala^{294}) and *p*-F-Phe-excluding (Ser^{294}) variants; the differences to the wild type, however, were at the limit of significance. With a ratio of 0.53, the Val^{294} mutant exhibited again wild-type specificity (Fig. 9).

This pattern of relative *p*-F-Phe/Phe specificity is exactly reflected by competition experiments in which [^{14}C]Phe incorporation was competed by a 50-fold excess of *p*-F-Phe (Fig. 9, Table 3B). The largest competition appeared in the Gly^{294} mutant which was in the range of the isotope dilution effect, *i. e.* the value obtained by adding the same concentration of unlabelled Phe to the assay (Tables 2 and 3). In the Ser^{294} mutant, *p*-F-Phe did not significantly compete with Phe. Also for the other position 294 variants, the values of remaining Phe incorporation correlated well with the corresponding ratios of *p*-F-Phe/Phe aminoacylation (Fig. 9).

These results were further confirmed by *in vivo* experiments (Table 3A). *E. coli* strain RR28 (*p*-F-Phe resistant) showed some marginal growth on *p*-F-Phe-containing plates when a wild-type *pheS* gene (encoding the Ala^{294} variant) was present on a plasmid. In contrast, the presence of a Gly^{294} mutant α subunit prevented any trace of growth. This *p*-F-Phe 'supersensitive' phenotype is explained by an enhanced detrimental *p*-F-Phe incorporation as compared to wild-type PheRS, which paralleled the *in vitro* results (see above). The intermediate growth of cells carrying genes encoding the Cys^{294} and Thr^{294} mutant forms could be rationalized by the partial exclusion of *p*-F-Phe from the enzymatic reaction (Fig. 9; Table 3B). An alternative explanation is that their PheRSs have essentially wild-type specificity but at a reduced activity level. This seems to be the case for the Val^{294} enzyme (Tables 3A & 3B). A model for the interaction of *p*-F-Phe with the residue at position 294 will be discussed later.

(b) Effects of other substrates and substrate analogues

The results of competition experiments with position 294 mutants, using the substrates and analogues of Fig. 7, are presented in Table 3B. The spectrum of competition of the proteinogenic amino acids corresponded to wild-type PheRS (section 4.5.) with two remarkable exceptions. Tyrosine interfered substantially with aminoacylation by the Gly^{294} and the Ser^{294} mutants. The effect of Tyr in mu-

Table 3: *Effect of substrates and substrate analogues on position 294 mutant PheRS activities*

	Wild type	Position 294 mutant α subunits					Vector
A *In vivo* growth tests: mutant plasmids in strain RR28 (*p*-F-Phe resistant)[a]							
Minimal medium, 37°C	Ala^{294}	Gly^{294}	Ser^{294}	Cys^{294}	Thr^{294}	Val^{294}	pBLS
- *p*-F-Phe	++	++	++	++	++	++	++
+ *p*-F-Phe	–	0	+	±	±	±	+

B *In vitro* competition of Phe incorporation by substrate analogues (in %)[b]

Substance	Ala^{294}	Gly^{294}	Ser^{294}	Cys^{294}	Thr^{294}	Val^{294}
Phe[c]	6 %	7 %	5 %	10 %	8 %	11 %
Tyr	93 %	53 %	51 %	92 %	89 %	91 %
aa-pool 1	94 %	89 %	89 %	106 %	98 %	95 %
aa-pool 2	101 %	91 %	100 %	109 %	98 %	101 %
aa-pool 3	97 %	88 %	98 %	104 %	94 %	96 %
aa-pool 4	100 %	94 %	96 %	107 %	101 %	94 %
p-F-Phe	26 %	8 %	90 %	59 %	46 %	27 %
m-F-Phe	53 %	34 %	45 %	65 %	59 %	49 %
o-F-Phe	71 %	63 %	73 %	85 %	77 %	65 %
3-F-Tyr	105 %	89 %	88 %	100 %	102 %	101 %
β-thien-2-ylalanine	57 %	69 %	38 %	62 %	70 %	53 %
aa-pool 5	97 %	20 %	96 %	101 %	106 %	108 %
aa-pool 6	85 %	41 %	82 %	84 %	86 %	80 %

[a] See legend to Table 2. Growth of single colonies is specified from good growth to no growth at all, in the order: ++ > + > ± > – > 0.

[b] The competition assay was carried out as described in Materials and Methods with an estimated error of 10% -15 %.

[c] Isotope dilution effect (positive control) by addition of a 50-fold excess of unlabelled Phe.

tants Gly^{294} and Ser^{294} is paralleled by a slight competition (15-20 %) seen with 3-F-Tyr (Table 3B). These effects may be explained by interactions between the *p*-hydroxyl groups of the substrate analogues (Tyr or 3-F-Tyr) with the Ser^{294} or Gly^{294} variants (see Discussion).

Other deviations from the wild-type pattern of competition by the individual substances (see section 4.5.) are observed with β-thien-2-ylalanine in the Ser^{294} mutant (Table 3B). The increased competition effect could be interpreted to result from favourable H-bonding of the hydroxyl group with the ring sulfur. Another explanation implies that the reduced ring size of the analogue is accommodated better in the smaller binding cavity of the Ser^{294} variant (see Discussion) due to increased van der Waals interactions. In this context, and by considering the marked differences in *p*-F-Phe competition between wild-type and Ser^{294} enzymes (Fig. 9; Table 3), we wished to explore the limits of substrate discrimination in both variants. The fact that Phe binding is somewhat hampered by the Ser^{294} mutation (section 4.6.) prompted us to analyze the marginally smaller *p*-deutero-L-Phe (a C-^{2}H bond may be shorter by up to 0.007 Å (or 0.6 %) than a C-^{1}H bond; Bartell *et al.*, 1961; Mitchell & Phillips, 1975). Remaining L-[^{14}C]Phe incorporation was measured in the presence of varying concentrations of either unlabelled L-Phe or unlabelled *p*-deutero-L-Phe. The tests were done in parallel with the wild-type and the Ser^{294} enzymes. From the titration curves in these competition assays (not shown) no differences were found between the two substances with both PheRS variants. Thus, the enzymes could not discriminate between Phe and *p*-^{2}H-Phe within the accuracy of the experiment (±5 %).

The most striking differential competition effect was found with pool groups 5 and 6 in the case of the Gly^{294} enzyme (Table 3B). Subsequently, each member of these pool groups was tested separately for its ability to compete with Phe. The results are summarized in Table 4. In contrast to the wild-type enzyme, which was not markedly inhibited by the substrate analogues, the addition of *p*-Cl-Phe caused a reduction of Phe incorporation catalysed by the Gly^{294} PheRS of more than 70 %. Even a Phe derivative with a large bromo-substituent at the *para*-position could interact with the mutated binding site (competition effect of more than 50 %; Table 4). Di-fluoro-derivatives, having one substitution at the *para*-position, showed also increased competition with the cognate substrate in the Gly^{294} variant, as compared to the wild-type enzyme. The relaxed specificity of the Gly^{294} mutant appeared to be directed mainly to the *para*-position of the amino acid substrate; this was indicated, for instance, by the missing effect of *m*-Cl-Phe (Table 4).

Table 4: *Competition effects by individual members of amino acid pool groups 5 and 6 on PheRS mutant Gly^{294} and the wild type*[a]

Competing substrate analogue	% Remaining Phe incorporation in PheRS variant:	
	Gly^{294}	Ala^{294} (wild type)
none	100 %	100 %
Phe	7 %	3 %
para-chloro-Phe	28 %	105 %
para-bromo-Phe	46 %	104 %
meta-chloro-Phe	99 %	102 %
2,3-di-fluoro-Phe	92 %	99 %
2,4-di-fluoro-Phe	80 %	97 %
2,5-di-fluoro-Phe	98 %	98 %
2,6-di-fluoro-Phe	103 %	102 %
3,4-di-fluoro-Phe	56 %	94 %
3,5-di-fluoro-Phe	93 %	99 %

[a] The competition tests were carried out as described in Materials and Methods. Each number is the average of four independent assays incubated for 5 min at 28°C. Deviations were in the range of 5 %.

(c) In vivo *phenotypes of PheRS mutants: growth on different Phe analogue-containing plates*

It became of interest to test whether the ability of certain substrate analogues to compete with Phe is of significance *in vivo*. Therefore, the growth of *E. coli* strains carrying plasmids with wild-type and mutant *pheS* genes was analyzed on different analogue-containing agar plates. The results in the *p*-F-Phe resistant strain RR28 are shown in Table 5. An analogous growth pattern was also observed using the thermosensitive PheRS strain KA2 (data not shown).

For the *para*-halogenated Phe derivatives, the results *in vivo* (Table 5) matched exactly those from competition experiments *in vitro* (Tables 3B & 4). The

Table 5: In vivo *phenotypes of PheRS mutants grown in the presence of Phe analogues*

Strain/plasmid	Amino acid at position 294 of α subunit, encoded by:		Growth on supplemented minimal medium[a], ampicillin, IPTG at 37°C with addition of (0.55 mM end concentration):				
	Chromosome	Plasmid	–	L-*p*-F-Phe	L-*p*-Cl-Phe	L-*p*-Br-Phe	L-Tyr
RR28	Ser	–	0	0	0	0	0
RR28/pBLS	Ser	–	++	+	++	++	++
RR28/pKSB1-W[b]	Ser	Ala[b]	++	–	++	++	++
RR28/pKSB1-M4G	Ser	Gly	++	0	0	0	++
RR28/pKSB1-M4S	Ser	Ser	++	+	++	++	++

[a] Medium and conditions for the growth tests were described in detail in the preceding paper (Kast & Hennecke, 1991). Growth of single colonies is characterized from good (++), normal (+), marginal (–) growth to no growth at all (0).

[b] Plasmid with wild-type *pheS* gene.

strain carrying wild-type *pheS* was sensitive only to *p*-F-Phe, an analogue shown to be incorporated into $tRNA^{Phe}$ and subsequently into (defective) proteins (Fig. 9; Richmond, 1962; Fangman & Neidhardt, 1964; Kast & Hennecke, 1991). If the Gly^{294} mutant gene was present, the strain showed extreme sensitivity not only to *p*-F-Phe but also to the *p*-Cl-Phe and *p*-Br-Phe analogues (Table 5). This indicated that a PheRS which contains Gly^{294} mutant α subunits possesses a broadened substrate range *in vivo* (see Discussion).

The addition of tyrosine to agar plates showed no growth inhibition, neither in cells carrying wild-type *pheS*, nor in those with genes for Gly^{294} and Ser^{294} mutant α subunits (Table 5). Obviously, there was no significant erroneous incorporation of Tyr instead of Phe into cellular proteins (which would certainly have been lethal), even though Tyr was found to compete with Phe *in vitro* in the latter two PheRS mutants (Table 3B). Explanations for this behaviour will be given in the Discussion.

5. Discussion

5.1. A method for assaying mutated genes for essential proteins at normal cellular levels

(a) The chromosomally encoded β subunit is limiting

The activity of the mutant PheRS α subunits was measured in S140 extracts of a host strain with a thermosensitive PheRS (KA2: lesion in *pheS*; Kast & Hennecke, 1991) which carried the mutated *pheS* genes on a multicopy expression vector. The high gene dosage of plasmidial *pheS*, used to outcompete potential effects coming from the chromosomal copy, did not influence the PheRS activity level. PheRS was shown to be active only if both α and β subunits were present (Hennecke & Böck, 1975). Therefore, the chromosomally encoded β subunits should limit the concentration of functional oligomeric enzymes. That this assumption was correct could be demonstrated by comparing PheRS activities of independent S140 extracts. Aminoacylation values in protein extracts from three different strains were identical to the activities obtained with the host strain KA2 that carried the corresponding *pheS* genes on plasmids (Kast & Hennecke, 1991).

The synthesis of the activity-limiting β subunit was not influenced by the *pheS* gene dosage. This was shown directly by Western-blot analyses which demonstrated a constant relation between S140 protein content and amount of β subunit for all mutants, irrespective of the presence or absence of *pheS* plasmids. Analogous results regarding the expression of the *pheST* operon have been reported earlier (Springer *et al.*, 1983; Springer *et al.*, 1985).

(b) Expression of PheRS α subunit variants

The low amount of α subunits seen in Western-blots after overexpression of plasmid-encoded *pheS* gene products was unexpected. The observed concentrations ranged from 1.1 to 8.3 times that of the host protein and contrasted with overproduction factors (more than 100 fold) of other proteins using similar systems (*e. g.* by Eriani *et al.*, 1990*b*). An artifact in the Western-blot analyses caused by differential binding of anti-wild-type-PheRS antiserum to different mutant α subunits appeared unlikely since 3 independently obtained antisera showed similar cross-reactions. Furthermore, even the wild-type α subunit was present in quite a small

amount (Fig. 3 (a,b)). The low amount of overexpressed *pheS* gene products in S140 extracts might rather be exlained by their degradation, once they are not complexed with β subunits. Free α subunits might be recognized as abnormal cellular components and could induce degradative processes (Goff & Goldberg, 1985). A low amount of PheRS α subunit, when expressed without β, was already observed in earlier reports (Plumbridge *et al.*, 1980; Plumbridge & Springer, 1982).

That the inactive, truncated mutant form M5 (33 amino acids missing at the C-terminus) was not visible on Western-blots may also be ascribed to rapid degradation. This could be due to inherent instability of the protein or to its inability to form a stabilized complex with the β subunit. All other mutants, however, were shown to be able to compete with the chromosomally encoded α subunit in strain RR28 to engage with the β subunit in the formation of a holoenzyme (Tables 2A & 3A). A correlation between high aminoacylation activity and relatively high content of overproduced PheRS α subunit in S140 extracts could not be found (see Results, section 4.1.). This indicates that the absence of excess free α subunits does not interfere with the determination of mutant PheRS activities, as long as the stoichiometric relation of α to β subunits is not below one.

(c) Background activity of the host α subunit

The host *pheS* gene products were thought to represent a minority as compared to the plasmid-expressed α subunits. However, the appearance of degradative processes and the existence of low activity mutants made an examination of background activities necessary. Host-encoded $\alpha_2\beta_2$ homotetramers had no activity *in vitro* at 28°C (Kast & Hennecke, 1991). The presence of a considerable percentage of hybrid tetramers (containing one chromosome- and one plasmid-encoded α subunit) could be ruled out from thermoinactivation tests (as described in Results, section 4.1.). The results from the *p*-F-Phe incorporation and competition studies also argued against significant amounts of hybrid PheRS. Hybrids, containing one thermosensitive and one *p*-F-Phe excluding α subunit, were shown to incorporate the analogue (Hennecke, 1976). By contrast, *p*-F-Phe was not a substrate for PheRS in S140 extracts of KA2 host cells which overproduced the *p*-F-Phe resistant Ser^{294} α subunit by a factor of 2 (Fig. 3). Although the influence of hybrid enzymes in cell extracts with comparably high aminoacylation activities may be neglected, a contribution to the low values of some thermosensitive mutants (*e. g.* those containing Gly^{295}) cannot be excluded. Thus, the activities obtained in those cases may be overestimates.

(d) General applicability of the expression method

The method described for assaying *in vitro* mutagenized essential genes at the normal cellular level represents an easy alternative to the chromosomal integration of these mutant genes. In the approach used here, the resident chromosomal copy is outcompeted by high expression of the mutant gene on a high copy number replicon. Nevertheless, the activity level appears to stay in the normal cellular range because other subunits, essential for the activity of the enzyme, are provided by the chromosome, and because the normal regulation is not affected. This strategy should be applicable to many multisubunit enzymes (essential or non-essential ones) whose activity is dependent on the presence of all subunits. The availability of (conditional) mutant strains is an advantage, but the assay could do without, provided there are means to quantitate background activity.

5.2. The role of the aromatic residues at positions 293 and 295 in the α subunit

In this work, the two phenylalanine residues flanking position 294 of the PheRS α subunit were replaced by other amino acids. From *in vivo* and *in vitro* results it became thus clear that aromatic interactions between the substrate phenylalanine and those residues are not essential for activity. Determination of apparent K_M values and competition studies with substrate analogues suggested that the side chains of the two residues 293 and 295 are not at all involved in Phe binding.

Residues 293 to 295 are part of motif 3, a generally conserved region of class II aminoacyl-tRNA synthetases (Eriani *et al.*, 1990*a*). Its importance in *E. coli* PheRS was clearly demonstrated with the Phe^{295}-to-Gly^{295} exchange, which showed a dramatic loss of activity *in vitro* and *in vivo*. However, the other mutant α subunits (excluding the deletion variant M5), were able to complement strain KA2 at high temperature, although some had only low activity *in vitro*. That those mutations in the rather well conserved PheRS region (Fig. 1) were tolerated *in vivo* might be explained by the observation that PheRS is normally synthesized in excess (Plumbridge & Springer, 1982), and that the applied assay conditions did not require maximal PheRS activity.

When the enzymes were tested *in vitro*, a drastic thermosensitive behaviour was apparent with several PheRS mutants, indicating a general destabilization of the protein structure (Grütter *et al.*, 1979; Matthews *et al.*, 1987; Menéndez-Arias

& Argos, 1989; Toth & Schimmel, 1990). At elevated temperatures, the contribution of residue 295 to the stabilization of the protein structure was clearly more important than that of position 293. Residue 295 seems to cause stabilization via hydrophobic interactions: The hydrophobic Trp could fully replace Phe^{295}, whereas more polar and/or smaller residues produced only unfavourable interactions. This agrees very well with the good conservation of this position in different organisms (Fig. 1), the only deviation from Phe being Trp in yeast cytoplasmic PheRS. The less conserved position 293 seems to tolerate polar residues (Tyr, Cys). Interestingly, replacing both Phe residues by Trp strongly reduced enzyme stability, even though each substitution by itself had little detrimental effects. This may point to an interaction between the two positions 293 and 295. In the double mutant, the two bulky Trp residues might hinder each other in providing good stabilization to the enzyme and might even distort the conformation of the substrate binding pocket. The latter interpretation would explain why the Trp^{293}/Trp^{295} variant had a slightly increased K_M value for Phe.

Thermosensitivity of the PheRS from strain NP37 (the parent strain of KA2; Kast & Hennecke, 1991) was speculated to be due to (reversible) dissociation into αβ dimers (Böck, 1968; Comer & Böck, 1976; Kast & Hennecke, 1991). From an analysis of reversion frequencies of thermosensitive PheRS mutants it was proposed that many sites could be involved in subunit interactions (Böck, 1968). A model of the quarternary structure of PheRS based on neutron small-angle scattering also proposed extended contact areas between the subunits (Dessen *et al.*, 1990). Therefore, it may be possible that the reversibly thermosensitive PheRS mutants constructed in this work are affected in the association parameters of the subunits, as discussed for other thermosensitive mutant proteins (Jasin *et al.*, 1985; Collier & Johnson, 1990; Kast & Hennecke, 1991). Residues 293 and 295 might thus be located close to a subunit interface where they could interact with another subunit.

This would also place residue 294, assumed to be involved in amino acid substrate binding (see below), near a subunit contact area. Such a model would be in agreement with biochemical affinity labelling data. Using a Phe affinity label, Khodyreva *et al.* (1985) could locate the Phe binding site on the α subunit, which is consistent with our results. However, when reactive ATP and aminoacyladenylate analogues were used, both α and β subunits were labelled, indicating that the corresponding binding sites were close to the subunit interface. The crosslinking sites for different reactive derivatives of $tRNA^{Phe}$ were all confined to the β subunit. It was thus proposed that the active site of *E. coli* PheRS is located at the αβ subunit

interface (Khodyreva *et al.*, 1985; Hountondji *et al.*, 1987). For yeast and turkey liver PheRS, the catalytic site (and especially the Phe binding site) was also suggested to be located at the αβ contact region (Baltzinger *et al.*, 1979; Renaud *et al.*, 1982; Fasiolo *et al.*, 1989; Gabius *et al.*, 1983*b*). Based on an analysis of *E. coli* GlyRS α and β subunit mutants, Toth & Schimmel (1990) proposed the involvement of both subunits in the formation of the site for adenylate synthesis. GlyRS is the other aminoacyl-tRNA synthetase having an $\alpha_2\beta_2$ subunit structure, and was shown to be immunologically related to *E. coli* PheRS (Nagel *et al.*, 1988). Interestingly, sequence homologies between both enzymes are restricted to the presence of motif 3 in the small subunit and, hence, to the region analyzed in this work.

5.3. Changing the substrate specificity by mutating Ala^{294}

The most interesting effects were obtained, when position 294 was mutagenized. From K_M determinations and competition with substrate analogues it seemed obvious that the binding of the amino acid substrate was affected. Although variations in apparent K_M values for the cognate substrate were only within one order of magnitude, they are comparable to the factor of 2 obtained after mutating one of the specificity-determining residues in TyrRS (Tyr^{34} to Phe^{34}; Fersht *et al.*, 1985). With PheRS, wild-type like activity was only obtained, when the small amino acids Gly, Ala and Ser were present at position 294; Cys, Thr and Val led to a strong decrease in aminoacylation at 28°C and showed a strong thermosensitive behaviour at 40°/45°C. A drop in activity was also observed *in vivo*, when the amino acid volume at position 294 was larger than Ser. It is possible that a critical size increase above Ser sterically alters the topology of the binding site and leads to a (reversible) destabilization of the protein structure, as discussed in the previous section. In this context it should be recalled that a small amino acid is conserved at this position in different PheRSs (Fig. 1). Three position 294 mutations will be discussed in more detail in the following.

(a) Cys at position 294

The mutant Cys^{294} showed a tendency for exclusion of *p*-F-Phe. This phenotype corroborates that of the structurally similar Ser^{294} derivative and can be explained by a mechanism analogous to that proposed in the preceding paper (Kast & Hennecke, 1991). The introduced Cys residue could become useful for active

site modifications with reactive substrate analogues (Khodyreva *et al.*, 1985; Hountondji *et al.*, 1987), especially with respect to proving any direct contacts between the *para*-position of the substrate's aromatic ring and residue 294 (see below). An interesting approach for further testing the steric requirements of the Phe binding site in PheRS could be to covalently attach groups of varying sizes to the (reactive) Cys^{294} (Bech & Breddam, 1988). Furthermore, once suitable crystals of *E. coli* PheRS are obtained, heavy metal atom derivatives at the Cys^{294} residue could greatly aid in the structural determination by X-ray diffraction analysis (Rould *et al.*, 1989).

(b) Ser at position 294

The Ser^{294} mutation, which we identified in a *p*-F-Phe resistant strain, is known to exclude this substrate analogue from the enzymatic reaction. As discussed before (Kast & Hennecke, 1991), this is thought to be due to steric and/or polar interactions between Ser^{294} and the fluoro substituent of the Phe analogue. In a simple model presented in Fig. 10(a) we assume that the hydroxyl group of Ser^{294} prevents complete entering of *p*-F-Phe into the substrate binding pocket. The K_M increase for Phe in the Ser^{294} mutant, as compared to the wild type, is also consistent with the model: the Ser residue might sterically interfere with optimal binding of phenylalanine. This is in agreement with the interpretation of results from subtilisin variants with Ala or Ser substitutions in the substrate specificity pocket (Estell *et al.*, 1986). However, the binding affinity of PheRS for the slightly smaller *p*-deutero-Phe was, compared to Phe, not improved in the Ser^{294} mutant. This showed up the lower limit of the enzyme's capability to significantly discriminate between sterically comparable substrates.

In contrast to *p*-F-Phe, tyrosine significantly competed with the incorporation of the cognate substrate Phe by the Ser^{294} enzyme. A possible explanation is that Tyr forms H-bonds with the hydroxyl group of Ser^{294}, thereby stabilizing the binding of Tyr in the Phe recognition site (Fig. 10(b)). H-bonding between enzyme residues and the *para*-hydroxyl group of Tyr is known to play an essential role in specific binding and recognition of the cognate substrate (and selection against Phe) in TyrRS (Fersht *et al.*, 1985; Brick & Blow, 1987; Freist & Sternbach, 1988). In the case of PheRS, it could be argued that for steric reasons an entering of Tyr into the binding pocket of the Ser^{294} variant is incomplete (Fig. 10(b)), thus leading to non-productive binding. This assumption is in agreement with growth tests on agar plates containing Tyr. Although the external addition of Tyr raises the cellular Tyr concentration at least 3 to 5 times that of Phe (Brown, 1970), thereby enhancing

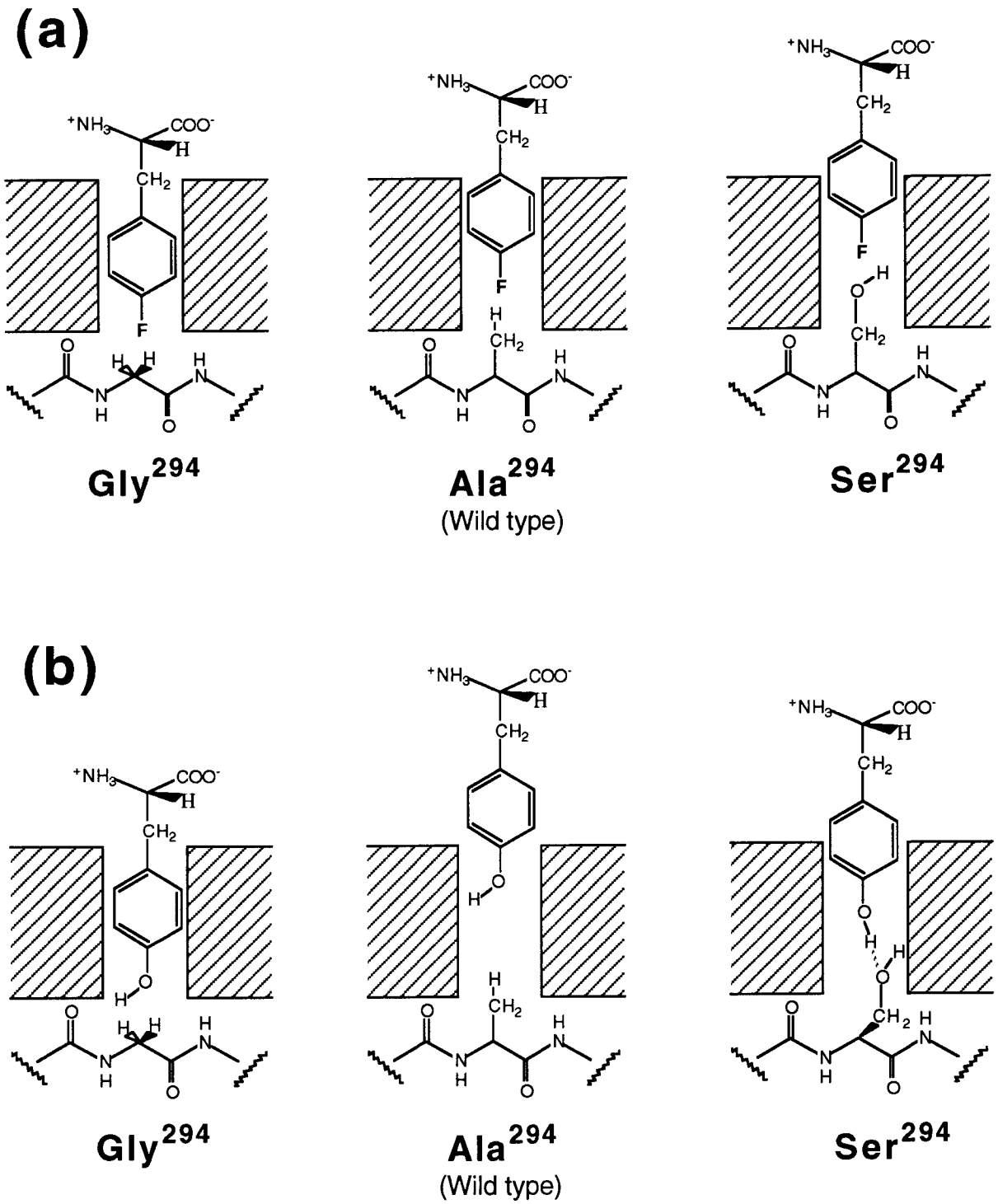


Figure 10. Hypothetical model for interactions between amino acid substrates and position 294 residues in PheRS α subunit variants. (a) The substrate analogue *p*-F-Phe is accommodated best in the binding cavity of the Gly^{294} variant, whereas Ser^{294} does not allow complete entering of *p*-F-Phe into the binding pocket. **(b)** The *para*-hydroxyl group of the noncognate amino acid Tyr prevents binding to the hydrophobic pocket of the wild-type PheRS. However, interactions of Tyr with the Gly^{294} and Ser^{294} variants may be possible due to an enlarged binding cavity and the potential to form H-bonds, respectively.

the probability of misincorporations, the normal growth on Tyr plates of cells producing the Ser^{294} mutant PheRS suggests that significant misacylation of $tRNA^{Phe}$ with Tyr does not occur.

Besides a non-productive binding, there is another explanation why the competing substrate Tyr is not incorporated by PheRS. As with other aminoacyl-tRNA synthetases, the evolution of PheRS has led to intrinsic mechanisms that help discriminate against non-cognate substrates. If the initial recognition of the amino acid by the binding site fails, hydrolytic proofreading steps come into action and thus improve the accuracy of aminoacylation (Freist, 1989). The proofreading system of *E. coli* PheRS (Gabius *et al.*, 1983*a*), may be able to cope with enhanced Tyr misactivation by the Ser^{294} mutant, and could therefore account for the missing growth inhibition on Tyr-containing plates, too.

(c) Gly at position 294

In line with the model proposed in the preceding paper (Kast & Hennecke, 1991), one would predict the formation of an enlarged cavity for accomodating the *para*-position of the aromatic ring of the amino acid substrate when Ala^{294} is replaced by a Gly residue. In fact, from independent *in vitro* and *in vivo* tests it emerged that, in contrast to the wild type, phenylalanine analogues with *para*-substitutions larger than fluorine could interact with the Gly^{294} enzyme.

As illustrated in Fig. 10(b), a tyrosine substrate will not enter the Phe binding pocket in wild-type PheRS due to the lack of space for accomodation of the (possibly hydrated) *p*-hydroxyl group. A very similar situation of steric repulsion of Tyr was described by Estell *et al.* (1986). In the Gly^{294} variant, however, competition of Phe aminoacylation by Tyr was strongly enhanced. From this we infer that Tyr is now accomodated in the large cavity at the bottom of the presumptive binding pocket as a result of the absence of the methyl group from the former Ala residue. That cells producing the Gly^{294} mutant exhibited no impaired growth on Tyr-containing plates could be explained by reasons already given for the Ser^{294} variant (see above).

Further evidence for an increased cavity for the *para*-position of the amino acid substrate in the Gly^{294} enzyme came from *p*-F-Phe incorporation experiments, and from competition studies with different *para*-halogenated analogues. Under our test conditions, the mutant enzyme, as opposed to the wild type, had lost the ability to distinguish *p*-F-Phe from the natural substrate. Interestingly, the yeast cytoplasmic PheRS, which happens to have a Gly residue at the corre-

sponding sequence position (Fig. 1), also discriminates much poorer against *p*-F-Phe than the wild-type *E. coli* PheRS: the ratio $(k_{cat}/K_M)_{Phe}$ to $(k_{cat}/K_M)_{p\text{-F-Phe}}$ is 4.5 times higher for the *E. coli* enzyme (Gabius *et al.*, 1983*a*).

The model shown in Fig. 10(a) explains the complete loss of discrimination between Phe and *p*-F-Phe again by the larger binding cavity present in the Gly^{294} enzyme. The absence of the methyl group could now allow an even better entering of *p*-F-Phe into the mutant binding pocket as compared to wild-type PheRS. Consistent with this model is the fact that the spectrum of substrates for the Gly^{294} variant was extended to Phe analogues with *para*-substitutents larger than fluorine (van der Waals radius: 1.47 Å; Bondi, 1964): Unlike the wild type, mutant Gly^{294} allowed *p*-Cl-Phe and *p*-Br-Phe (van der Waals radii: Cl, 1.77 Å; Br, 1.92 Å) to considerably compete with Phe. In addition, those substances, like the *para*-fluorinated Phe, completely prevented growth of strains carrying the gene for mutant Gly^{294}. It was clear that the Ala^{294}-to-Gly^{294} exchange in the PheRS α subunit alone must be responsible for the observed hypersensitive phenotypes. Therefore, the *para*-halogenated Phe analogues have to exert their toxicity *in vivo* by interacting with the Gly^{294} mutant PheRS, an interaction already shown to occur *in vitro* (Table 4). For reasons of analogy to *p*-F-Phe and parallelism to other systems (*e. g.* incorporation of furanomycin by IleRS; Kohno *et al.*, 1990), it seems highly probable that the growth inhibition by *p*-Cl-Phe and *p*-Br-Phe was caused by a detrimental incorporation of those analogues into cellular protein. However, protein incorporation studies *in vivo* (Fangman & Neidhardt, 1964; Sykes *et al.*, 1974; Koide *et al.*, 1988) or *in vitro* (Kohno *et al.*, 1990) will be necessary to support this contention (Hortin & Boime, 1983). Furthermore, radiolabelled analogues (as in the case of *p*-F-Phe) and purified Gly^{294} PheRS will be useful tools in future work to definitively prove the incorporation of those analogues into $tRNA^{Phe}$ (Igloi *et al.*, 1979).

A change in substrate specificity created by protein engineering and/or chemical modification was achieved for the proteases subtilisin (Estell *et al.*, 1986), trypsin (Craik, *et al.*, 1985) and carboxypeptidase Y (Bech & Breddam, 1988). Similar to the position 294 PheRS mutants, it emerged from these studies, that substrate specificity was largely determined by steric complementarity. In particular, the example of subtilisin resembled the findings from our own work. Replacing the Gly residue at the bottom of subtilisin's substrate binding pocket by bigger amino acids resulted in a change in specificity towards smaller substrates (Estell *et al.*, 1986). Even in the 'superspecific' aminoacyl-tRNA synthetases (as opposed to proteases; Bosshard, 1976), amino acid specificity could be altered by exchanging

binding pocket residues. In TyrRS, Fersht *et al.* (1985) substituted one of the two amino acids that determine specificity for Tyr through hydrogen bonding to the substrate's *para*-hydroxyl group (Brick & Blow, 1987). The mutation lowered the discrimination of Phe relative to Tyr in the activation reaction by one order of magnitude; yet the biological effect may be small because the relative specificity for activation of the cognate substrate Tyr compared to Phe was still higher (by a factor of 10^4) in the mutant enzyme. In addition, Tyr specificity might be further enhanced in the transfer reaction to $tRNA^{Tyr}$ by proofreading mechanisms (Freist & Sternbach, 1988; Freist, 1989). The substrate specificity mutant Gly^{294} of PheRS (and also Ser^{294}), however, showed a very distinct phenotype with some non-proteinogenic substrate analogues in the *in vivo* growth tests. Apparently, the proofreading system of PheRS (see above) did not prevent incorporation of those substances. This was in agreement with the idea that proofreading mechanisms were optimized solely to discriminate against noncognate substrates that occur naturally in the cell, but may tolerate incorporation of xenobiotic analogues under experimental conditions (Wilson & Hatfield, 1984).

Due to the enlarged binding cavity, the Gly^{294} enzyme tolerates *para*-substituents at the Phe analogues that are at least as big as bromine. This opens up a new perspective for the *in vivo* synthesis of polypeptides containing non-proteinogenic amino acids. The Gly^{294} mutant PheRS, for instance, may be used to incorporate a variety of Phe analogues into proteins of pharmaceutical interest such as polypeptide growth factors (James & Bradshaw, 1984) or endocrine peptide hormones (Douglass *et al.*, 1984; Lynch & Snyder, 1986). The resulting artificial polypeptides could have novel pharmacologically interesting properties. We are currently exploring ways to establish an *in vivo* system for the incorporation of amino acid analogues into target proteins with the help of our PheRS mutants. Different approaches for incorporating non-proteinogenic amino acids into protein were tested out previously with varying yields of analogue substitution (Morris & Schlesinger, 1972; Sykes *et al.*, 1974; Lu *et al.*, 1976; Pines *et al.*, 1981; Koide *et al.*, 1988; Kohno *et al.*, 1990). However, these authors worked with wild-type enzymes that use only a limited spectrum of amino acid analogues as substrates (Hortin & Boime, 1983). Our examples with the PheRS mutants Ser^{294} and Gly^{294} demonstrated that aminoacyl-tRNA synthetases can be specifically 'tailored' by mutation to use substrates of interest.

5.4. Summing up

With the set of amino acid exchanges introduced at positions 293, 294 and 295 of the α subunit of *E. coli* PheRS, we obtained a survey on the importance of these residues in a well conserved region. From mutations that negatively affected enzyme activity (especially at position 295) it emerged that this region is essential for aminoacylation by *E. coli* PheRS and implicated the importance of the generally conserved sequence motif 3 of class II aminoacyl-tRNA synthetases (Eriani *et al.*, 1990*a*). In the preceding report (Kast & Hennecke, 1991) we proposed a general participation of motif 3 in amino acid substrate binding. Although it was shown here that residues 293 and 295 of PheRS were not directly involved in binding of phenylalanine, they were important for protein stability, which may thus be another possible role of motif 3. However, replacements at position 294 and their effects on the use of various substrates strongly suggested that this position is involved in determining substrate specificity. In addition, it was demonstrated that even an aminoacyl-tRNA synthetase, normally designed for very strict substrate discrimination, can be manipulated by single amino acid replacements to use a wider substrate spectrum.

We are greatly indebted to S. Cantoreggi and D. Seebach for the chemical synthesis of several Phe analogues used in this work and for the suggestion to test *p*-deutero-Phe. We also wish to thank P. Babczinski for sending us other Phe derivatives, H. Sternbach for providing purified PheRS, T. Schmidheini for FPLC purification of oligonucleotides, and K. Girgenrath and L. Dürselen for scintillation counting. This work was supported by a grant from the Federal Institute of Technology Zürich.

6. References

Baltzinger, M., Fasiolo, F. & Remy, P. (1979). Yeast Phenylalanyl-tRNA Synthetase. Affinity and Photoaffinity Labelling of the Stereospecific Binding Sites. *Eur. J. Biochem.* **97**, 481-494.

Bartell, L. S., Kuchitsu, K. & DeNeui, R. J. (1961). Mean and Equilibrium Molecular Structures of Methane and Deuteromethane as Determined by Electron Diffraction. *J. Chem. Phys.* **35**, 1211-1218.

Bech, L. M. & Breddam, K. (1988). Chemical Modifications of a Cysteinyl Residue Introduced in the Binding Site of Carboxypeptidase Y by Site-Directed Mutagenesis. *Carlsberg Res. Commun.* **53**, 381-393.

Böck, A. (1968). Relation between Subunit Structure and Temperature-Sensitivity of Mutant Phenylalanyl RNA Synthetases of *Escherichia coli*. *Eur. J. Biochem.* **4**, 395-400.

Bondi, A. (1964). van der Waals Volumes and Radii. *J. Phys. Chem.* **68**. 441-451.

Bosshard, H. R. (1976). Theories of Enzyme Specificity and Their Application to Proteases and Aminoacyl-Transfer RNA Synthetases. *Experientia,* **32**, 949-963.

Brick, P. & Blow, D. M. (1987). Crystal Structure of a Deletion Mutant of a Tyrosyl-tRNA Synthetase Complexed with Tyrosine. *J. Mol. Biol.* **194**, 287-297.

Brick, P., Bhat, T. N. & Blow, D. M. (1989). Structure of Tyrosyl-tRNA Synthetase Refined at 2.3 Å Resolution. Interaction of the Enzyme with the Tyrosyl Adenylate Intermediate. *J. Mol. Biol.* **208**, 83-98.

Brown, K. D. (1970). Formation of Aromatic Amino Acid Pools in *Escherichia coli* K-12. *J. Bacteriol.* **104**, 177-188.

Burley, S. K. & Petsko, G. A. (1985). Aromatic-Aromatic Interaction: A Mechanism of Protein Structure Stabilization. *Science,* **229**, 23-28.

Burley, S. K. & Petsko, G. A. (1986). Dimerization Energetics of Benzene and Aromatic Amino Acid Side Chains. *J. Am. Chem. Soc.* **108**, 7995-8001.

Burnette, W. N. (1981). "Western Blotting": Electrophoretic Transfer of Proteins

from Sodium Dodecyl Sulfate-Polyacrylamide Gels to Unmodified Nitrocellulose and Radiographic Detection with Antibody and Radioiodinated Protein A. *Anal. Biochem.* **112**, 195-203.

Chothia, C. (1984). Principles That Determine the Structure of Proteins. *Ann. Rev. Biochem.* **53**, 537-572.

Collier, C. D. & Johnson, J. R. (1990). The *Escherichia coli* K-12 *metJ193* Allele Contains a Point Mutation Which Alters the Hydrophobic Pocket Responsible for *in Vitro* Binding of *S*-Adenosylmethionine: Effects on Cell Growth and Induction of *met* Regulon Expression. *J. Bacteriol.* **172**, 3918-3924.

Comer, M. M. & Böck, A. (1976). Genes for the α and β Subunits of the Phenylalanyl-Transfer Ribonucleic Acid Synthetase of *Escherichia coli*. *J. Bacteriol.* **127**, 923-933.

Craik, C. S., Largman, C., Fletcher, T., Roczniak, S., Barr, P. J., Fletterick, R. & Rutter, W. J. (1985). Redesigning Trypsin: Alteration of Substrate Specificity. *Science,* **228**, 291-297.

Dessen, P., Ducruix, A., May, R. P. & Blanquet, S. (1990). Low-Resolution Structure of the Tetrameric Phenylalanyl-tRNA Synthetase from *Escherichia coli*. A Neutron Small-Angle Scattering Study of Hybrids Composed of Protonated and Deuterated Protomers. *Biochemistry*, **29**, 3039-3045.

Douglass, J., Civelli, O. & Herbert, E. (1984). Polyprotein Gene Expression: Generation of Diversity of Neuroendocrine Peptides. *Ann. Rev. Biochem.* **53**, 665-715.

Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990*a*). Partition of tRNA Synthetases into Two Classes Based on Mutually Exclusive Sets of Sequence Motifs. *Nature (London)*, **347**, 203-206.

Eriani, G., Dirheimer, G. & Gangloff, J. (1990*b*). Structure - Function Relationship of Arginyl-tRNA Synthetase from *Escherichia coli*: Isolation and Characterization of the *argS* Mutation MA5002. *Nucl. Acids Res.* **18**, 1475-1479.

Estell, D. A., Graycar, T. P., Miller, J. V., Powers, D. B., Burnier, J. P., Ng, P. G. & Wells, J. A. (1986). Probing Steric and Hydrophobic Effects on Enzyme-Substrate Interactions by Protein Engineering. *Science,* **233**, 659-663.

Fangman, W. L. & Neidhardt, F. C. (1964). Protein and Ribonucleic Acid Synthesis in a Mutant of *Escherichia coli* with an Altered Aminoacyl Ribonucleic Acid

Synthetase. *J. Biol. Chem.* **239**, 1844-1847.

Fasiolo, F., Sanni, A., Potier, S., Ebel, J. P. & Boulanger, Y. (1989). Identification of the Major $tRNA^{Phe}$ Binding Domain in the Tetrameric Structure of Cytoplasmic Phenylalanyl-tRNA Synthetase from Baker's Yeast. *FEBS Lett.* **242**, 351-356.

Fayat, G., Mayaux, J.-F., Sacerdot, C., Fromant, M., Springer, M., Grunberg-Manago, M. & Blanquet, S. (1983). *Escherichia coli* Phenylalanyl-tRNA Synthetase Operon Region. Evidence for an Attenuation Mechanism. Identification of the Gene for the Ribosomal Protein L20. *J. Mol. Biol.* **171**, 239-261.

Fersht, A. (1985*a*). In *Enzyme Structure and Mechanism* (2^{nd} ed), pp. 98-120, Freeman, New York.

Fersht, A. (1985*b*). In *Enzyme Structure and Mechanism* (2^{nd} ed), pp. 293-368, Freeman, New York.

Fersht, A. R. & Dingwall, C. (1979). Cysteinyl-tRNA Synthetase from *Escherichia coli* Does Not Need an Editing Mechanism to Reject Serine and Alanine. High Binding Energy of Small Groups in Specific Molecular Interactions. *Biochemistry*, **18**, 1245-1249.

Fersht, A. R., Shi, J.-P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y. & Winter, G. (1985). Hydrogen Bonding and Biological Specificity Analysed by Protein Engineering. *Nature (London)*, **314**, 235-238.

Freist, W. (1989). Mechanisms of Aminoacyl-tRNA Synthetases: A Critical Consideration of Recent Results. *Biochemistry*, **28**, 6787-6795.

Freist, W. & Sternbach, H. (1988). Tyrosyl-tRNA Synthetase from Baker's Yeast. Order of Substrate Addition, Discrimination of 20 Amino Acids in Aminoacylation of $tRNA^{Tyr}$-C-C-A and $tRNA^{Tyr}$-C-C-A ($3'NH_2$). *Eur. J. Biochem.* **177**, 425-433.

Gabius, H.-J., von der Haar, F. & Cramer, F. (1983*a*). Evolutionary Aspects of Accuracy of Phenylalanyl-tRNA Synthetase. A Comparative Study with Enzymes from *Escherichia coli*, *Saccharomyces cerevisiae*, *Neurospora crassa*, and Turkey Liver Using Phenylalanine Analogues. *Biochemistry*, **22**, 2331-2339.

Gabius, H.-J., von der Haar, F. & Cramer, F. (1983*b*). Purification by Salting-Out

Chromatography and Properties of Phenylalanyl-tRNA Synthetase from Turkey Liver. *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 71-81.

Goff, S. A. & Goldberg, A. L. (1985). Production of Abnormal Proteins in *E. coli* Stimulates Transcription of *lon* and Other Heat Shock Genes. *Cell*, **41**, 587-595.

Grüll, J. M., Hennecke, H., Fröhler, J., Thomale, J., Nass, G. & Böck, A. (1979). *Escherichia coli* Mutants Overproducing Phenylalanyl- and Threonyl-tRNA Synthetase. *J. Bacteriol.* **137**, 480-489.

Grütter, M. G., Hawkes, R. B. & Matthews, B. W. (1979). Molecular Basis of Thermostability in the Lysozyme from Bacteriophage T4. *Nature (London)*, **277**, 667-669.

Hangauer, D. G., Monzingo, A. F. & Matthews, B. W. (1984). An Interactive Computer Graphics Study of Thermolysin-Catalyzed Peptide Cleavage and Inhibition by *N*-Carboxymethyl Dipeptides. *Biochemistry*, **23**, 5730-5741.

Hartsuck, J. A. & Lipscomb, W. N. (1971). Carboxypeptidase A. In *The Enzymes* (Boyer, P. D., ed), vol. 3, pp. 1-56, Academic Press, London.

Hennecke, H. (1976). Use of Mutant Enzymes to Demonstrate the Presence of Two Active Sites on Phenylalanyl-tRNA Synthetase from *Escherichia coli*. *FEBS Lett.* **72**, 182-186.

Hennecke, H. & Böck, A. (1975). Altered α Subunits in Phenylalanyl-tRNA Synthetases from *p*-Fluorophenylalanine-Resistant Strains of *Escherichia coli*. *Eur. J. Biochem.* **55**, 431-437.

Hennecke, H., Walther, I. & Franz, H. (1977). Immunochemical Studies on Phenylalanyl-tRNA Synthetase from *Escherichia coli*. *Hoppe-Seyler's Z. Physiol. Chem.* **358**, 197-208.

Holler, E. (1980). Kinetics of Anticooperative Binding of Phenylalanyl-tRNA^{Phe} and tRNA^{Phe} to Phenylalanyl-tRNA Synthetase of *Escherichia coli* K10. *Biochemistry*, **19**, 1397-1402.

Hortin, G. & Boime, I. (1983). Applications of Amino Acid Analogs for Studying Co- and Posttranslational Modifications of Proteins. *Methods Enzymol.* **96**, 777-784.

Hountondji, C., Schmitter, J.-M., Beauvallet, C. & Blanquet, S. (1987). Affinity

Labelling of *Escherichia coli* Phenylalanyl-tRNA Synthetase at the Binding Site for $tRNA^{Phe}$. *Biochemistry*, **26**, 5433-5439.

Igloi, G. L., von der Haar, F. & Cramer, F. (1979). Experimental Proof for the Misactivation of Amino Acids by Aminoacyl-tRNA Synthetases. *Methods Enzymol.* **59**, 282-291.

Jakubowski, H. (1990). Proofreading *in Vivo*: Editing of Homocysteine by Methionyl-tRNA Synthetase in *Escherichia coli*. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 4504-4508.

James, R. & Bradshaw, R. A. (1984). Polypeptide Growth Factors. *Ann. Rev. Biochem.* **53**, 259-292.

Jasin, M., Regan, L. & Schimmel, P. (1985). Two Mutations in the Dispensable Part of Alanine tRNA Synthetase Which Affect the Catalytic Activity. *J. Biol. Chem.* **260**, 2226-2230.

Kast, P. & Hennecke, H. (1991). Phenylalanyl-tRNA Synthetase Mutants of *Escherichia coli* Analyzed at the Molecular Level. *J. Mol. Biol.* (submitted); see Chapter II.

Kester, W. R. & Matthews, B. W. (1977). Crystallographic Study of the Binding of Dipeptide Inhibitors to Thermolysin: Implications for the Mechanism of Catalysis. *Biochemistry*, **16**, 2506-2516.

Khodyreva, S. N., Moor, N. A., Ankilova, V. N. & Lavrik, O. I. (1985). Phenylalanyl-tRNA Synthetase from *E. coli* MRE-600: Analysis of the Active Site Distribution on the Enzyme Subunits by Affinity Labelling. *Biochim. Biophys. Acta,* **830**, 206-212.

Kohno, T., Kohda, D., Haruki, M., Yokoyama, S. & Miyazawa, T. (1990). Nonprotein Amino Acid Furanomycin, Unlike Isoleucine in Chemical Structure, Is Charged to Isoleucine tRNA by Isoleucyl-tRNA Synthetase and Incorporated into Protein. *J. Biol. Chem.* **265**, 6931-6935.

Koide, H., Yokoyama, S., Kawai, G., Ha, J.-M., Oka, T., Kawai, S., Miyake, T., Fuwa, T. & Miyazawa, T. (1988). Biosynthesis of a Protein Containing a Nonprotein Amino Acid by *Escherichia coli*: L-2-Aminohexanoic Acid at Position 21 in Human Epidermal Growth Factor. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 6237-6241.

Laemmli, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the

Head of Bacteriophage T4. *Nature (London)*, **227**, 680-685.

Lu, P., Jarema, M., Mosser, K. & Daniel, Jr., W. E. (1976). *lac* Repressor: 3-Fluorotyrosine Substitution for Nuclear Magnetic Resonance Studies. *Proc. Nat. Acad. Sci., U.S.A.* **73**, 3471-3475.

Lynch, D. R. & Snyder, S. H. (1986). Neuropeptides: Multiple Molecular Forms, Metabolic Pathways, and Receptors. *Ann. Rev. Biochem.* **55**, 773-799.

Matthews, B. W., Nicholson, H. & Becktel, W. J. (1987). Enhanced Protein Thermostability from Site-Directed Mutations that Decrease the Entropy of Unfolding. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 6663-6667.

Menéndez-Arias, L. & Argos, P. (1989). Engineering Protein Thermal Stability. Sequence Statistics Point to Residue Substitutions in α Helices. *J. Mol. Biol.* **206**, 397-406.

Mitchell, P. J. & Phillips, L. (1975). Evidence That a C-^{2}H Bond Has a Smaller Steric Requirement Than a C-^{1}H Bond. *J. Chem. Soc. Chem. Commun.* **1975**, 908-909.

Morris, H. & Schlesinger, M. J. (1972). Effects of Proline Analogues on the Formation of Alkaline Phosphatase in *Escherichia coli*. *J. Bacteriol.* **111**, 203-210.

Nagel, G. M., Johnson, M. S., Rynd, J., Petrella, E. & Weber, B. H. (1988). Glycyl-tRNA Synthetase of *Escherichia coli*: Immunological Homology with Phenylalanyl-tRNA Synthetase. *Arch. Biochem. Biophys.* **262**, 409-415.

Nakamaye, K. L. & Eckstein, F. (1986). Inhibition of Restriction Endonuclease *Nci*I Cleavage by Phosphorothioate Groups and Its Application to Oligonucleotide-Directed Mutagenesis. *Nucl. Acids Res.* **14**, 9679-9698.

Pines, M., Rosenthal, G. A. & Applebaum, S. W. (1981). *In Vitro* Incorporation of L-Canavanine into Vitellogenin of the Fat Body of the Migratory Locust *Locusta migratoria migratorioides*. *Proc. Nat. Acad. Sci., U.S.A.* **78**, 5480-5483.

Plumbridge, J. A. & Springer, M. (1982). *Escherichia coli* Phenylalanyl-tRNA Synthetase Operon: Characterization of Mutations Isolated on Multicopy Plasmids. *J. Bacteriol.* **152**, 650-660.

Plumbridge, J. A., Springer, M., Graffe, M., Goursot, R. & Grunberg-Manago, M. (1980). Physical Localisation and Cloning of the Structural Gene for *E. coli* Initiation Factor IF3 from a Group of Genes Concerned with Translation.

Gene, **11**, 33-42.

Renaud, M., Fasiolo, F., Baltzinger, M., Boulanger, Y. & Remy, P. (1982). Affinity Labelling of Yeast Phenylalanyl-tRNA Synthetase with a 3'-Oxidised $tRNA^{Phe}$. Isolation and Sequence of the Labelled Peptide. *Eur. J. Biochem.* **123**, 267-274.

Richmond, M. H. (1962). The Effect of Amino Acid Analogues on Growth and Protein Synthesis in Microorganisms. *Bacteriol. Rev.* **26**, 398-420.

Ringe, D., Seaton, B. A., Gelb, M. H. & Abeles, R. H. (1985). Inactivation of Chymotrypsin by 5-Benzyl-6-Chloro-2-Pyrone: ^{13}C NMR and X-Ray Diffraction Analyses of the Inactivator-Enzyme Complex. *Biochemistry,* **24**, 64-68.

Robertus, J. D., Alden, R. A., Birktoft, J. J., Kraut, J., Powers, J. C. & Wilcox, P. E. (1972). An X-Ray Crystallographic Study of the Binding of Peptide Chloromethyl Ketone Inhibitors to Subtilisin BPN'. *Biochemistry,* **11**, 2439-2449.

Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H. & Zehfus, M. H. (1985). Hydrophobicity of Amino Acid Residues in Globular Proteins. *Science*, **229**, 834-838.

Rould, M. A., Perona, J. J., Söll, D. & Steitz, T. A. (1989). Structure of *E. coli* Glutaminyl-tRNA Synthetase Complexed with $tRNA^{Gln}$ and ATP at 2.8 Å Resolution. *Science*, **246**, 1135-1142.

Seebach, D., Dziadulewicz, E., Behrendt, L., Cantoreggi, S. & Fitzi, R. (1989). Synthesis of Nonproteinogenic (*R*)- or (*S*)-Amino Acids Analogues of Phenylalanine, Isotopically Labelled and Cyclic Amino Acids from *tert*-Butyl 2-(*tert*-Butyl)-3-methyl-4-oxo-1-imidazolidinecarboxylate (Boc-BMI). *Liebigs Ann. Chem.* **1989**, 1215-1232.

Shive, W. & Skinner, C. G. (1963). Amino Acid Analogues. In *Metabolic Inhibitors* (Hochster, R. M. & Quastel, J. H., eds), vol. 1, pp. 1-73, Academic Press, London.

Singh, J. & Thornton, J. M. (1985). The Interaction between Phenylalanine Rings in Proteins. *FEBS Lett.* **191**, 1-6.

Springer, M., Mayaux, J.-F., Fayat, G., Plumbridge, J. A., Graffe, M., Blanquet, S. & Grunberg-Manago, M. (1985). Attenuation Control of the *Escherichia coli* Phenylalanyl-tRNA Synthetase Operon. *J. Mol. Biol.* **181**, 467-478.

Springer, M., Trudel, M., Graffe, M., Plumbridge, J., Fayat, G., Mayaux, J.-F., Sacerdot, C., Blanquet, S. & Grunberg-Manago, M. (1983). *Escherichia coli* Phenylalanyl-tRNA Synthetase Operon Is Controlled by Attenuation *in Vivo*. *J. Mol. Biol.* **171**, 263-279.

Steitz, T. A., Henderson, R. & Blow, D. M. (1969). Structure of Crystalline α-Chymotrypsin. *J. Mol. Biol.* **46**, 337-348.

Sykes, B. D., Weingarten, H. I. & Schlesinger, M. J. (1974). Fluorotyrosine Alkaline Phosphatase from *Escherichia coli*: Preparation, Properties, and Fluorine-19 Nuclear Magnetic Resonance Spectrum. *Proc. Nat. Acad. Sci., U.S.A.* **71**, 469-473.

Toth, M. J. & Schimmel, P. (1990). A Mutation in the Small (α) Subunit of Glycyl-tRNA Synthetase Affects Amino Acid Activation and Subunit Association Parameters. *J. Biol. Chem.* **265**, 1005-1009.

Weaver, L. H., Kester, W. R. & Matthews, B. W. (1977). A Crystallographic Study of the Complex of Phosphoramidon with Thermolysin. A Model for the Presumed Catalytic Transition State and for the Binding of Extended Substrates. *J. Mol. Biol.* **114**, 119-132.

Wells, J. A., Cunningham, B. C., Graycar, T. P. & Estell, D. A. (1987*a*). Recruitment of Substrate-Specificity Properties from One Enzyme into a Related One by Protein Engineering. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 5167-5171.

Wells, J. A., Powers, D. B., Bott, R. R., Graycar, T. P. & Estell, D. A. (1987*b*). Designing Substrate Specificity by Protein Engineering of Electrostatic Interactions. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 1219-1223.

Wilson, M. J. & Hatfield, D. L. (1984). Incorporation of Modified Amino Acids into Proteins *in Vivo*. *Biochim. Biophys. Acta*, **781**, 205-215.

Zhang, R.-G., Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z. & Sigler, P. B. (1987). The Crystal Structure of *trp* Aporepressor at 1.8 Å Shows How Binding Tryptophan Enhances DNA Affinity. *Nature (London)*, **327**, 591-597.

CHAPTER IV

Design of Systems for the Incorporation of Non-Proteinogenic Amino Acids into Proteins

1. Introduction

1.1. PheRS mutant Gly^{294}, an enzyme with relaxed substrate specificity

In Chapter III, a mutant phenylalanyl-tRNA synthetase (PheRS) gene constructed *in vitro* was described whose product deviated from the *Escherichia coli* wild-type enzyme by a glycine residue at position 294 instead of alanine. The mutant PheRS (referred to as Gly^{294} variant) showed enhanced aminoacylation of tRNA with the substrate analogue *para*-fluoro-phenylalanine (*p*-F-Phe) as compared to the wild-type enzyme. In fact, the analogue was used as efficiently as the cognate substrate phenylalanine (Phe). In addition, other substrate analogues like the *para*-substituted chloro- and bromo- derivatives of Phe, 3,4-di-fluoro-Phe and even the proteinogenic amino acid tyrosine were able to drastically compete with Phe incorporation by the Gly^{294} enzyme. From these results and from the toxic effects of these compounds *in vivo* it was concluded that the absence of the methyl group at position 294 enlarged the binding cavity for the *para*-position of the aromatic amino acid substrate, thereby broadening the substrate spectrum of the enzyme (Chapter III). It seems reasonable to assume that the Gly^{294} variant may also attach other *para*-substituted Phe analogues to $tRNA^{Phe}$. Once aminoacylated, the substrate analogues will end up with a high probability in cellular proteins

---

Non-standard abbreviations

Amp: ampicillin; bp: basepair(s); Cam: chloramphenicol; *goi*: **g**ene **o**f **i**nterest; IFN-α2: human interferon-α2; IPTG: isopropyl-β-D-thiogalactopyranoside; kb: kilobasepair(s); *p*-Cl-Phe: *p*-chlorophenylalanine; *p*-F-Phe: *p*-fluoro-phenylalanine; ts: thermosensitive; XxxRS: aminoacyl-tRNA synthetase for amino acid Xxx in three-letter-code.

(Hortin & Boime, 1983; Kohno *et al.*, 1990).

These features provide an excellent basis for the use of the Gly^{294} PheRS, in an appropriate genetic constellation, to deliberately incorporate various Phe analogues into polypeptides *in vivo*. This, in turn, opens up a perspective for the production of artificial proteins possessing a variety of new constituents and properties. In this Chapter I will outline the theoretical possibilities and strategies for Phe analogue incorporation into protein by making use of mutant Gly^{294}. Possible applications will also be discussed.

1.2. Previous work on amino acid analogue incorporation into proteins

It is well known that many amino acid analogues exert their toxicity due to the formation of defective cellular proteins which contain those analogues instead of the cognate amino acid (Richmond, 1962; Shive & Skinner, 1963; Fowden *et al.*, 1967; Hortin & Boime, 1983; Wilson & Hatfield, 1984; Kohno *et al.*, 1990). This fact was utilized in a variety of protein structure-function studies (Hortin & Boime, 1983): Analogue incorporation into proteins has been applied for the analysis of proteolytic processing events, including viral polypeptide maturation and cleavage of signal sequences (Wilson & Hatfield, 1984). Substitution of tyrosine by 3-fluorotyrosine allowed an investigation into the chemical environment at the individual substitution sites by ^{19}F nuclear magnetic resonance spectroscopy. Such studies, carried out with *E. coli* alkaline phosphatase (Sykes *et al.*, 1974) and *lac* repressor (Lu *et al.*, 1976; Jarema *et al.*, 1981), were also suited to uncover specific conformational changes upon ligand binding. Other reports dealt with effects on enzyme function and subunit assembly caused by replacement of amino acid residues with analogues (Richmond, 1962; Fowden *et al.*, 1967). Examples of enzymes studied are alkaline phosphatase (Schlesinger *et al.*, 1969; Morris & Schlesinger, 1972) and aspartate transcarbamylase (Gueguen *et al.*, 1980) from *E. coli*.

In some studies, the incorporation of amino acid analogues was accomplished simply by adding the analogue to the growth medium (Richmond, 1962; Richmond, 1963; Fangman & Neidhardt, 1964; Koide *et al.*, 1988). A more elaborate methodology made use of bacterial strains that were auxotrophic for the natural amino acid to be replaced, thereby increasing the substitution yield: The cells were first grown in a medium containing the required amino acid. In the second stage, the incubation was continued in a medium in which the natural amino acid

was replaced by its analogue. This procedure led to a depletion of the endogenous pool of the natural, competing amino acid. If at the same time the synthesis of the protein of interest was induced, it was expected that mainly protein containing the analogue would be formed (Schlesinger *et al.*, 1969). By using this strategy, replacement efficiencies of 50 % to 80 % were reported (Gueguen *et al.*, 1980). To circumvent the necessity of changing the growth medium, Sykes *et al.* (1974) adjusted the concentration of the supplementing natural amino acid such that it was depleted by the time of induction of the synthesis of the protein into which the analogue was to be incorporated. At the same time, the analogue (3-fluorotyrosine) was added, too. The substitution yield in that system was 73 %. Another system proposed by Lu *et al.* (1976) was based on the fact that the cellular aromatic amino acid biosynthetic pathway can be repressed and feedback-inhibited by its end products (see also Fig. 1). The endogenous synthesis of Tyr was down-regulated by adding Phe and Trp to the growth medium and, at a later stage, 3-fluorotyrosine, the analogue to be incorporated into the *lac* repressor. At the time of analogue addition, the synthesis of the *lac* repressor was drastically enhanced by heat induction of a prophage carrying the corresponding *lacI* gene. In contrast to the system designed by Sykes *et al.* (1974), this strategy allowed the use of a prototrophic strain (Tyr^+) without having the disadvantage of substantial endogenous synthesis of the natural competitor Tyr. The efficiency of substitution by the analogue reached 90 %; in addition, an increased absolute yield of the synthesized target protein was obtained (Lu *et al.*, 1976).

Incorporation studies with the help of chemically aminoacylated tRNAs have allowed the *in vitro* synthesis of (tiny amounts of) proteins containing a wide variety of amino acid analogues (Bain *et al.*, 1991; Robertson *et al.*, 1991). In contrast, the spectrum of amino acid analogues that could be incorporated into proteins *in vivo* was up to now restricted to those tolerated as substrates by wild-type aminoacyl-tRNA synthetases (Hortin & Boime, 1983; Kohno *et al.*, 1990). The proposed exploitation of the PheRS mutant Gly^{294} in a suitable system will perhaps demonstrate that a genetically engineered aminoacyl-tRNA synthetase can be used to extend the range of *in vivo* incorporated amino acid analogues.

2. General strategies in the design of phenylalanine analogue incorporation systems

2.1. Test experiments

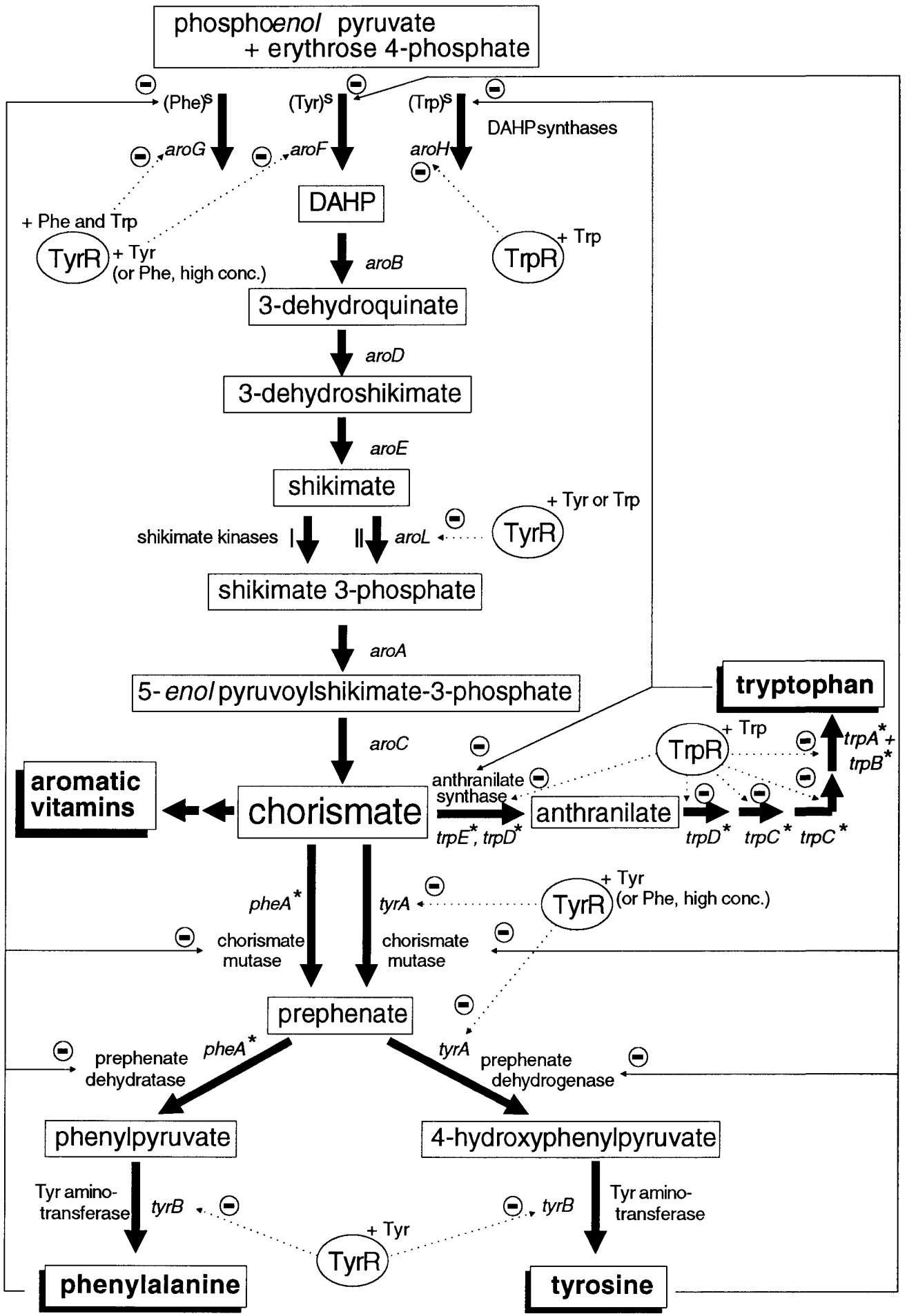
Before making extensive efforts towards the construction of sophisticated amino acid analogue incorporation systems, two types of alternative test experiments ought to be carried out first. (i) Purified Gly^{294} PheRS will be needed to prove the incorporation of individual Phe analogues into tRNA *in vitro*, an important prerequisite for the later insertion into protein. (ii) More directly, a simplified *in vivo* incorporation system should be established to test whether analogue-containing proteins can be synthesized by means of the relaxed specificity of the Gly^{294} PheRS variant. In the following, a concrete example for such a system will be described which will quickly lead to results without asking for high yields of analogue substitution.

(a) Choice of the Phe analogue

The Phe analogue that will be chosen for incorporation is *p*-Cl-Phe. There is evidence (Chapter III) that this amino acid is incorporated only by the Gly^{294} enzyme and not by the wild-type PheRS (which will be used as a control). The proposed incorporation system follows the strategy by Lu *et al.* (1976) which makes use of a Phe prototrophic *E. coli* strain (see Introduction). Endogenous Phe bio-

---

Figure 1. Pathway for the biosynthesis of aromatic amino acids and its regulation in *E. coli*. Each bold face arrow designates an enzymatic reaction in the anabolism of aromatic amino acids. The arrows between chorismate and aromatic vitamins, however, stand for multiple steps not further considered here. Thin lines and arrows point to enzymes which are feedback-inhibited by the end products of the pathways. Dotted lines indicate regulation by repression of genes for biosynthetic enzymes. The effector(s) which associate(s) with the repressor proteins TyrR or TrpR are indicated. Recently it was discovered that the PheR gene does not encode a repressor protein, but is identical to one of the two cellular tRNA^{Phe} genes (Gavini & Davidson, 1990; Pittard *et al.*, 1990). This suggests that attenuation is the exclusive type of regulation of *pheA* (Pittard, 1987). Regulation of gene expression by attenuation (concerned genes are marked by an asterisk) was also observed for the *trp* operon (mediated by tRNA^{Trp}), in addition to repression. Shikimate kinase I is constitutively expressed. Only the regulated key enzymes are shown by name. DAHP, 3-deoxy-*D*-*arabino*-heptulosonate 7-phosphate; (Xxx)^{S} denotes the DAHP synthase isoenzyme sensitive to amino acid Xxx (three-letter-code). The data in this Figure come from reviews by Herrmann (1983), Garner & Herrmann (1983), Camakaris & Pittard (1983), Somerville (1983) and Pittard (1987).


phosphoenol pyruvate
+ erythrose 4-phosphate
(Phe)S
(Tyr)S
(Trp)S
DAHP synthases
aroG
aroF
aroH
DAHP
+ Phe and Trp
TyrR
+ Tyr
(or Phe, high conc.)
TrpR
+ Trp
aroB
3-dehydroquinate
aroD
3-dehydroshikimate
aroE
shikimate
+ Tyr or Trp
shikimate kinases I
II
aroL
TyrR
shikimate 3-phosphate
aroA
5-enol pyruvoylshikimate-3-phosphate
tryptophan
aroC
TrpR
+ Trp
trpA*+
trpB*
aromatic vitamins
chorismate
anthranilate synthase
trpE*, trpD*
anthranilate
trpD*
trpC*
trpC*
pheA*
tyrA
TyrR
+ Tyr
(or Phe, high conc.)
chorismate mutase
chorismate mutase
prephenate
prephenate dehydratase
pheA*
tyrA
prephenate dehydrogenase
phenylpyruvate
4-hydroxyphenylpyruvate
Tyr amino-transferase
tyrB
tyrB
Tyr amino-transferase
+ Tyr
TyrR
phenylalanine
tyrosine

synthesis can be suppressed by adding Tyr and Trp to the growth medium. These aromatic amino acids should repress and feedback-inhibit various enzymes in the aromatic amino acid biosynthetic pathway which are also used for Phe production. Furthermore, the later addition of the Phe analogue *p*-Cl-Phe may directly inhibit the Phe sensitive enzymes of this pathway and, therefore, further lower the amount of endogenous Phe, which would otherwise compete with incorporation of the analogue. The regulation of the aromatic amino acid biosynthetic pathway by its endproducts is illustrated in Fig. 1.

(b) Choice of the target protein

Human interferon-α2 (IFN-α2) will be chosen as the protein into which the analogue is to be incorporated. Interferon is the designation of members of a large family of structurally related proteins produced by vertebrate cells *e. g.* after virus infection (Hiscott *et al.*, 1984; Weissmann & Weber, 1986). One of the various cellular processes modulated by interferons is the interference with virus proliferation (Stewart, 1981). Several features make IFN-α2 suitable for our test system: (i) The (mature) protein with its 165 amino acid residues (Streuli *et al.*, 1980) is in the size range of proteins such as human epidermal growth factor with 49 amino acids (Gregory, 1975) and *E. coli* alkaline phosphatase with 471 amino acids (Chang *et al.*, 1986) that were already used for detailed analogue incorporation studies by Koide *et al.* (1988) and Sykes *et al.* (1974), respectively. (ii) The IFN-α2 gene contains 10 Phe codons which could direct Phe analogue incorporation. (iii) IFN-α2 belongs to the class of pharmaceutically interesting proteins which could become the primary targets for analogue incorporation by the systems proposed in this Chapter. (iv) The activity of the natural and of analogue-containing IFN-α2 can be compared with the help of a well established biological assay (Nagata *et al.*, 1980; Weber *et al.*, 1987). (v) IFN-α2 has successfully been overexpressed, and purification procedures are available (Palva *et al.*, 1983; Valenzuela *et al.*, 1985; Thatcher & Panayotatos, 1986; Schein & Noteborn, 1988).

(c) Elements of the test system

Figure 2 displays the characteristic elements of the proposed test system. The *E. coli* strain KA2 (carrying the *pheS* allele for the production of a thermosensitive PheRS α subunit; see Chapter II) is used as host for two different compatible plasmids.

The plasmid pHE3-M4G (P15A origin of replication) carries the *pheS* gene for

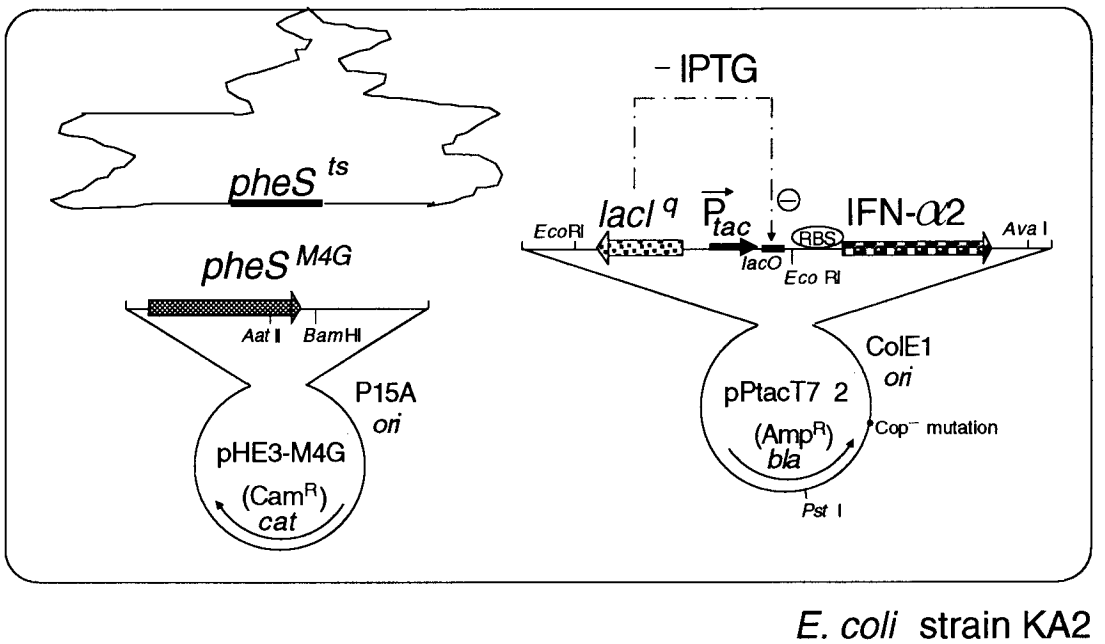


Figure 2. A simple Phe analogue incorporation test system. The functional elements of this test system are described in detail in the text. Key restriction sites used to construct the plasmids are included. *pheS*^{ts} and *pheS*^{M4G} encode the thermosensitive and relaxed-specificity α subunits of PheRS, respectively. The types of replication origins (*ori*) are indicated. Resistance against chloramphenicol (Cam^R) and ampicillin (Amp^R) are mediated by chloramphenicol acetyltransferase and β-lactamase, encoded by *cat* and *bla*, respectively. Plasmid pPtacT7α2 contains the Cop$^-$ mutation, the efficient T7 gene *1.1* ribosomal binding site (RBS) and the IFN-α2 gene of pPIT7α2 (Thatcher & Panayotatos; 1986) allowing for high expression of IFN-α2. The interferon gene is expressed from a *tac* promoter which can be blocked (in the absence of the inducer substance IPTG) at the operator site (*lacO*) by constitutively expressed *lac* repressor (encoded by *lacI*^{q}).

the Gly^{294} variant. It can be constructed easily by replacing the 951 bp *Aat*II/*Bam*HI fragment of pHE3 (Hennecke *et al.*, 1982) by the 326 bp *Aat*II/*Bam*HI fragment from the *pheS* expression plasmid pKSB1-M4G (see Chapter III and Fig. 2 in Chapter II). A similar construct (termed 'pHE3-W') which carries the wild-type *pheS* gene will be obtained by using the corresponding *Aat*II/*Bam*HI fragment from plasmid pKSB1-W (Chapter II). Although not always mentioned explicitly, all tests should be done with the wild-type *pheS* plasmid (pHE3-W) in parallel as negative control (*i. e.* no *p*-Cl-Phe incorporation).

The second plasmid pPtacT7α2 is a ColE1 derivative and expresses IFN-α2. Unfortunately, the plasmid used earlier for high expression of IFN-α2 (pPIT7α2; Thatcher & Panayotatos, 1986; Schein & Noteborn, 1988) does not allow an inducible production of the target protein. However, inducibility is required for efficient

incorporation of analogues into proteins, since the analogue has to be added at a late stage of bacterial growth. Therefore, the constitutive promoter of pPIT7α2 (PI of pBR322; Panayotatos *et al.*, 1983; Thatcher & Panayotatos, 1986) should be replaced by an inducible one. This will be achieved easily by replacing the 0.26 kb *Eco*RI fragment of pPIT7α2, carrying PI (H. Weber, personal communication), by the 1.54 kb *EcoR*I/*Nru*I fragment from the vector pJF118EH (Fürste *et al.*, 1986), after providing the latter with *EcoR*I linkers. The fragment from pJF118EH contains the IPTG-inducible, strong *tac* promoter (Amann *et al.*, 1983) and a constitutively expressed *lac* repressor gene (*lacI*^{q}) allowing for a tight repression of the system, when IPTG is absent.

The last cloning steps in these constructions should be done in strain JM109 ($HsdR^-M^+$, see Chapter II) to obtain *Eco*K-modified DNA before transforming KA2 ($HsdR^+M^+$) with the plasmids.

(d) Strategy for Phe analogue incorporation

The following protocol is a recommendation how incorporation of *p*-Cl-Phe into IFN-α2 might be achieved. In stage I, the host strain KA2 harbouring both plasmids is grown at 40°C in the glucose-minimal medium (Chapter II) containing in addition 1 mM Trp, 1 mM Tyr, 5 µg/ml thiamine-HCl, 150 µg/ml ampicillin and 20 µg/ml chloramphenicol. When an optical density (A_{550}) of about 1.0 is reached, the Phe analogue *p*-Cl-Phe is added (*e. g.* at 2 mM final concentration of the L-form) and the IFN-α2 gene is induced by addition of IPTG at 2 mM (stage II). The PheRS molecules present in the cell should now contain predominantly the α subunits of the Gly^{294} type due the multicopy effect of the plasmidial *pheS* gene. PheRS containing two chromosomally-encoded thermosensitive α subunits is inactive at 40°C (see Chapters II and III). This ensures that charging of $tRNA^{Phe}$ is mainly achieved by the Gly^{294} variant which recognizes *p*-Cl-Phe as a substrate. After further incubation at 40°C for 3-6 h, the cells are harvested and the IFN-α2 is isolated and purified (stage III) using established protocols (*e. g.* by Valenzuela *et al.*, 1985). The analogue substitution grade (see below) as well as the biological activity can now be determined (Nagata *et al.*, 1980; Weber *et al.*, 1987).

2.2. Verification of analogue incorporation into proteins

There are several possibilities to check whether an analogue has been incorporated into protein (Hortin & Boime, 1983). However, it should be mentioned here

that some of the procedures described will only provide indirect evidences for a successful incorporation.

(a) Amino acid analysis

The most direct approach consists of analyzing the content of individual amino acids after hydrolysis of the purified protein. This procedure gives direct information about the substitution yield (ratio of detected analogue / [analogue + natural amino acid]). This method was applied to determine analogue incorporation in a variety of studies (Sykes *et al.*, 1974; Lu *et al.*, 1976; Gueguen *et al.*, 1980; Pines *et al.*, 1981; Koide *et al.*, 1988). Amino acid analyses can be carried out with amounts as low as 0.1 to 5 µg of a 25 kDa protein (corresponding to 4 to 200 pMol; Ozols, 1990).

(b) Radiolabelled amino acid analogues

Another strategy could be applied if the analogue is available in a radiolabelled form. This allows a direct demonstration of the incorporation into total protein (Fangman & Neidhardt, 1964; Lu *et al.*, 1976) and may even permit the analysis of analogue incorporation into a specifically induced protein without the need to purify it (Morris & Schlesinger, 1972). The latter strategy involves a simple separation of a crude protein extract on a polyacrylamide gel and subsequent autoradiography (perhaps in conjunction with a Western blot to immunologically identify the correct band). However, it cannot be ruled out *a priori* that the analogue was metabolized *in vivo* into another amino acid which, when incorporated, would also give rise to labelled protein. To solve this uncertainty, Hortin & Boime (1983) suggested to hydrolyse the labelled total protein and to confirm that the original radiolabelled amino acid analogue was still present, as exemplified, for instance, by Pines *et al.* (1981). However, in the test system proposed here this control may not be necessary: if the label was incorporated only in cells having the Gly^{294} PheRS but not in cells having the wild-type enzyme (negative control) metabolic conversion of the analogue into other natural amino acids did probably not occur. The main disadvantage of this very fast method is that it does not give information on the substitution yield.

(c) Radiolabelled phenylalanine

If the radiolabelled analogue is not available, incorporation experiments similar to those just mentioned may also be carried out with trace amounts of labelled

phenylalanine. This approach, however, is rather indirect, because it takes the reduced Phe incorporation as a measure to determine analogue incorporation. Again, the comparison of assays with the Gly^{294} mutant and the wild-type PheRS would be useful. To exclude a general diminution of protein synthesis under conditions of analogue incorporation, control experiments with a labelled amino acid other than Phe should also be carried out (Gale & Folkes, 1953).

(d) Analysis of incorporation in vitro

If analogue incorporation into a specified protein fails *in vivo*, it may be of interest to carry out *in vitro* incorporation studies. This can be done by using a prokaryotic *in vitro* transcription/translation system. In the presence of excess Gly^{294} or wild-type PheRS, different *in vitro* experiments will then be carried out, in which the amino acid Phe is (A) missing, (B) present or (C) replaced by the analogue (Kohno *et al.*, 1990). The DNA template-directed proteins synthesized *in vitro* (labelled with [^{35}S]methionine) can then be analyzed on polyacrylamide gels.

2.3. Some complications and potential pitfalls

The application of an amino acid analogue to a growing bacterial cell may produce several unpredicted and undesired effects, *e. g.* a low absolute yield of target protein compared to the yield without the analogue or low substitution ratios. The causes of some of these effects are discussed in the following, and possible counter-measures are proposed.

(a) Uptake of the amino acid analogue into the cell may be insufficient

The yield of substitution of a natural amino acid by its analogues may be low because the analogue is not efficiently transported into the cell. This problem could be due to the repression of the uptake system by amino acids present in the growth medium. In the example of the aforementioned test system, the aromatic amino acids Trp and Tyr were proposed to be added to the medium to inhibit and repress the Phe biosynthetic enzymes. However, these amino acids could also impair the activity of the *aroP*-encoded general aromatic amino acid uptake system by competitive inhibition (Brown, 1970) and repression (Whipp & Pittard, 1977; Camakaris & Pittard, 1983). Cells grown in the absence of exogenous amino acids have the capacity to import 80-90 % of added aromatic amino acids via this transport system (Brown, 1970), which is also largely responsible for the uptake of aro-

matic amino acid analogues (*e. g. p*-F-Phe and β-thien-2-ylalanine; Brown, 1970). The use of a constitutively derepressed *aroP* operator mutant strain (Pittard, 1987) should improve the uptake rates. However, it may already be sufficient to omit the aromatic amino acids from the medium. In this case, alternative ways to diminish the endogenous Phe production could be applied, if required (see section 3).

(b) Vital cellular proteins may be inactivated after analogue incorporation

After addition to the culture, the analogue will not only be incorporated into the target protein, whose synthesis is induced simultaneously, but also into other cellular proteins that are still produced. If functions essential for protein biosynthesis are impaired by this process, the yield of the desired product might be low (Sykes *et al.*, 1974; Koide *et al.*, 1988). The problem might be overcome by lowering the analogue concentration to a non-toxic level; however, this could also result in a reduction of the substitution yield. Alternatively, one might think of establishing conditions that prevent any incorporation of the analogue into proteins other than the desired target protein (see strategies presented in section 3).

(c) Analogue-substituted proteins may induce a heat shock-like response

It is known that accumulation of abnormal polypeptides in the cytoplasm of an *E. coli* cell may lead to induction of heat-shock genes (Goff & Goldberg, 1985). As a consequence of this stress response, transcription of genes with non-heat-shock promoters is depressed (Lindquist, 1986). One of the synthesized heat-shock proteins, the Lon protease (also termed 'La', encoded by the *lon* gene; Phillips *et al.*, 1984), leads to increased degradation of abnormal proteins (Goff & Goldberg, 1985; Lindquist, 1986). These heat-shock effects may reduce the yield of analogue-substituted target proteins (Koide *et al.*, 1988). Furthermore, a generally increased proteolysis in the cell (Mizusawa & Gottesman, 1983) would replete the pool of natural amino acids. The endogenous amino acid could then again compete with the analogue for being incorporated into tRNA (Gueguen *et al.*, 1980).

To prevent the unfavourable induction of the heat-shock response, the *in vivo* incorporation could be carried out in mutant strains with a defective heat-shock system (*e. g.* in an *rpoH* (*hptR*) mutant; Grossman *et al.*, 1984). Such strains show a decreased proteolytic activity (Goff *et al.*, 1984; Maurizi *et al.*, 1985); however, they are thermosensitive and preclude the use of systems inducible by a tempera-

ture upshift (Gottesman, 1990). Other solutions rely on mutations that knock out the heat-shock induced proteases (*e. g.* in *lon* and *dnaJ*; Gottesman, 1990). Alternatively, precautions could be taken to avoid the accumulation of aberrant polypeptides in the cell. This can be achieved (i) if the analogue-substituted protein is secreted immediately after its synthesis, and (ii) if transcription of other genes is repressed. Secretion of the target protein protects it from general degradation by cytoplasmic proteases and may also simplify the purification procedure (see section 3.5.). In addition, rapid secretion may reduce potential cytotoxic effects of certain target proteins. Koide *et al.* (1988) described a strategy for avoiding the heat shock-like response by making use of the regulation and secretion system of the *phoA* gene product. Another strategy to escape from the unfavourable heat-shock conditions will be described in section 3 and is based on applying the specific T7 RNA polymerase transcription system.

3. Proposals for phenylalanine analogue incorporation systems *in vivo*: Four variations on a theme

In this section, four alternative systems for the production of analogue-substituted target protein are described. The strategies are designed for high substitution yields, *i. e.* a nearly complete substitution of Phe by Phe analogues. This would simplify purification, especially in cases in which the proteins are encoded by genes that have more than one Phe codon. In addition, the overall yield of the protein should be high. In the proposed systems, some of the possible difficulties mentioned in the previous section (2.3.) are considered. While it is clear that the results and experiences from the prior test experiments (described in section 2.1.) must be awaited first, the strategies presented in the following may nevertheless be regarded as useful guidelines towards the development of an optimal system. The necessary constructions of the individual strains and plasmids will not be described in detail.

3.1. Version A

(a) Description of the functional elements of the system (Fig. 3)

Plasmid 1 (contains the target gene)

The gene of interest (*goi*) is under the control of a T7 promoter and resides on a multicopy plasmid (ColE1 derivative) containing restriction sites (MCS) suitable for cloning (Fig. 3). In addition, transcription is controlled by the *lac* repressor system (see below). Expression vectors with the T7 promoter and a highly efficient ribosome binding site (from T7 gene *10*; Olins & Rangwala, 1989; Olins & Rangwala, 1990) have been described which could serve as a basis for construction of plasmid 1 (Rosenberg *et al.*, 1987; Studier *et al.*, 1990). To facilitate easy purification of the *goi* gene product, *goi* could be fused to the coding region of a peptide which allows secretion and/or isolation of the fusion protein by affinity methods (*e. g.* antibody precipitations or ligand affinity columns; see also section 3.5.).

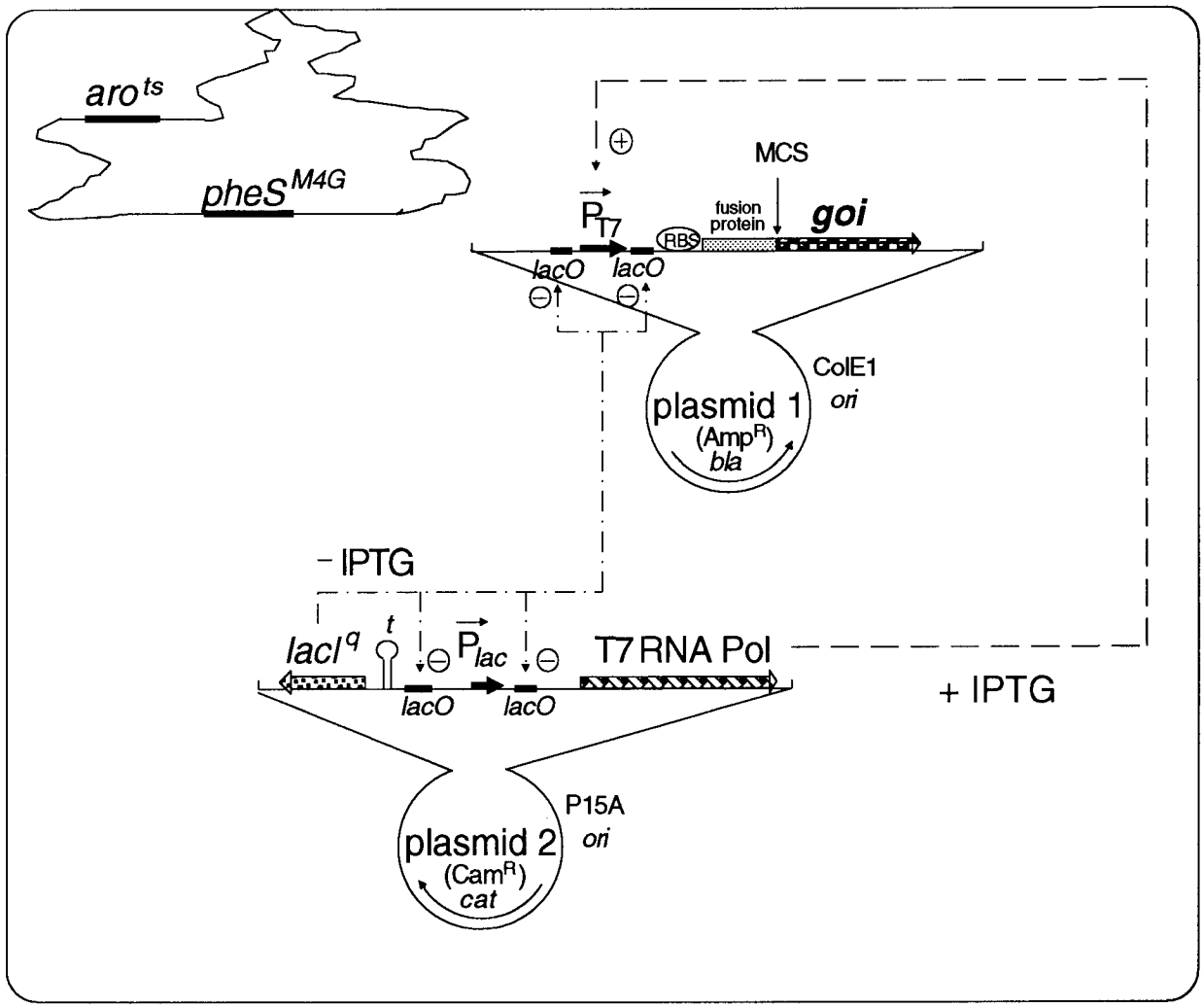


Figure 3. A sophisticated Phe analogue incorporation system, version A. The system is described in detail in the text and in Table 1. *aro*^{ts} encodes a thermosensitive enzyme of the aromatic amino acid biosynthetic pathway. *pheS*^{M4G} encodes the PheRS α subunit which confers relaxed substrate specificity. The origin of replication (*ori*) of each plasmid is indicated. *bla* and *cat* encode β-lactamase and chloramphenicol acetyltransferase conferring resistance to ampicillin (Amp^R) and chloramphenicol (Cam^R), respectively. *lacO* shows the location of the *lac* operators; *t* indicates a transcription terminator structure for the bacterial RNA polymerase. Locations and orientations of the T7 and *lac* promoters (P_{T7} and P_{lac}) are shown as well as of the genes *lacI*^{q} (encoding the constitutively high expressed *lac* repressor), *goi* (encoding the 'gene of interest') and the gene encoding T7 RNA polymerase. To facilitate purification, one could express *goi* as a fusion protein, if it is inserted into the multiple cloning site (MCS) in frame with a suitable coding region. RBS, ribosomal binding site.

Broken lines show some interactions between functional elements on the plasmids. In the repressed state, *lac* repressor blocks the *lac* and T7 promoters. This repression is relieved after induction with IPTG, and T7 RNA polymerase starts transcribing *goi*.

Plasmid 2

The system contains a second plasmid (*e. g.* a P15A derivative; Chang & Cohen, 1978) compatible with plasmid 1. It carries the genes for the T7 RNA polymerase (Tabor & Richardson, 1985) and a highly and constitutively expressed *lac* repressor (*lacI*^{q}; Calos, 1978; or even *lacI*^{q1}; Beyreuther, 1978). The RNA polymerase of bacteriophage T7 specifically transcribes genes (here *goi*) controlled by a phage T7 promoter (Tabor & Richardson, 1985; Studier *et al.*, 1990). This transcription proceeds even in the presence of rifampicin, which binds to the β subunit of the host bacterial RNA polymerase and thereby inhibits it (Scaife, 1976; Wehrli *et al.*, 1976). Hence, addition of Rifampicin after synthesis of the T7 RNA polymerase guarantees the exclusive transcription of *goi* from plasmid 1 (Tabor & Richardson, 1985).

In the absence of the analogue, *goi* ought not to be expressed. The *lac* repressor/operator system takes care of regulating *goi* expression. Transcription can then be induced by addition of the inducer substance IPTG which, by binding to the *lac* repressor, prevents it from occupying the *lac* operator site(s) near the promoter (Beckwith, 1987). In the uninduced state, *goi* expression is prevented due to occlusion of the promoters of the T7 RNA polymerase gene **and** of *goi* by *lac* repressor molecules binding to *lac* operator sites (*lacO*) (Fig. 3). This 'double repression' of *goi* expression is recommended since it is known that basal (uninduced) levels of T7 RNA polymerase can lead to considerable (occasionally lethal) transcription from a T7 promoter (Tabor & Richardson, 1985; Studier *et al.*, 1990). Repression mediated by the *lac* repressor/operator system has already been applied successfully to modulate transcription from a T7 promoter (Studier *et al.*, 1990). An optimal relative position of the *lac* operator to the transcription start site may be useful for maximum repression of the promoter (Elledge & Davis, 1989). In the case presented here, the *lac* operators are designed to occur in tandem (Fig. 3). This should further improve repression efficiency, since it is known that the *lac* repressor acts cooperatively at appropriately spaced operators (Besse *et al.*, 1986; Mossing & Record, 1986; Krämer *et al.*, 1987; Krämer *et al.*, 1988; Gralla, 1990). A transcription terminator (for the host bacterial RNA polymerase) in front of the T7 RNA polymerase promoter should further prevent undesirable basal transcription.

Host strain

Both plasmids are harboured by a host strain which provides the relaxed substrate-specificity PheRS variant Gly^{294} whose α subunit is encoded by the chro-

mosomal *pheS*^{M4G} gene. In addition, this strain carries a mutation in *pheA* leading to a thermosensitive Phe biosynthesis (see Fig. 1). Raising the cultivation temperature will cause an interruption of Phe biosynthesis and thus lead to drainage of the endogenous Phe pool. This decreases the intracellular ratio of Phe to Phe analogue (added externally) and therefore favours analogue incorporation.

Other useful phenotypes of the host strain include the RecA$^-$ marker to avoid homologous recombination events and the absence of a host-specific restriction endonuclease system (HsdR$^-$) to facilitate cloning. Mutations in *lon, clpA, dnaJ* and *ompT* may help decrease proteolytic activity and stabilize the target protein during synthesis and purification (Grodberg & Dunn, 1988; Gottesman, 1990; Studier *et al.*, 1990). Practical solutions for the construction of suitable *E. coli* host strains with diminished protease activities are given by Gottesman (1990).

(b) Successive steps in the incorporation system, version A

These are listed in Table 1. For additional information consult Fig. 3.

(c) Comments to version A

Two features of the proposed incorporation system potentially induce a heat-shock response: (i) the accumulation of abnormal polypeptides in the cell and (ii) the temperature upshift used for shutting down Phe biosynthesis (Lindquist, 1986). However, the strategy employed here to specifically express *goi* should circumvent some of the disturbances caused by the heat-shock response (discussed in section 2.3.c.): The use of the T7 RNA polymerase for target gene expression compensates for a general decrease in transcription activity from normal promoters as observed for the host RNA polymerase (Lindquist, 1986). Furthermore, the addition of rifampicin blocks the expression of heat-shock genes. The use of heat-shock protease deficient host strains (*lon, dnaJ*; Gottesman, 1990) should help eliminate undesired degradation of the target protein.

The main drawback of this system may be the incorporation of Phe analogues into T7 RNA polymerase. Substitution of most or all of the 37 Phe residues (Moffatt *et al.*, 1984) by the analogue might lead to impaired transcription efficiency or accuracy. The following alternative versions are designed to reduce analogue incorporation into T7 RNA polymerase. Another problem might arise, if the expression of this enzyme is not sufficiently repressed by the *lac* repressor/operator system, thus causing uncontrolled *goi* expression. This could further lead to plasmid loss and even cell death (Studier *et al.*, 1990). The versions proposed below may

Table 1: *Succession of stages of the incorporation system, version A (cf. Fig. 3)*

Stage	Action	State of the cell/immediate reaction	Subsequent effects/result
I	Growth in minimal medium, (+ glucose, Amp, Cam, 30°C) to late exponential phase	*lac* repressor blocks T7 RNA polymerase transcription/action	Target gene *goi* not expressed
II	Temperature shift 30° → 42°C	Phe biosynthesis (ts) stops	Drainage of cellular Phe pool
III	After X min: addition of IPTG and the Phe analogue	1.) IPTG inactivates *lac* repressor	Expression of T7 RNA polymerase from P_{lac} and deblocking of P_{T7} → Expression of target gene *goi*
		2.) Charging of $tRNA^{Phe}$ with Phe analogues only	Incorporation of analogues instead of Phe into proteins
IV	Addition of Rifampicin	Host RNA polymerase blocked	a) All transcription activity concentrated on T7 promoter b) Protein biosynthetic apparatus protected against adverse effects of Phe analogue incorporation
V	After 2-6 h: harvesting of the culture and purification of the target protein		

solve this problem since they use different means of controlling T7 RNA polymerase activity.

3.2. Version B

(a) Description of the functional elements of the system (Fig. 4)

Plasmid 1 (contains the target gene)

Version B (Fig. 4) is designed to further reduce the basal level of *goi* transcription in the uninduced state. Most features of plasmid 1 are as described in section 3.1.a. (Fig. 3). What differs from the previous system is the presence on plasmid 1 of the *lacI*^{q} gene and, additionally, the strong and inducible λ P_L promoter (see below). This promoter is orientated in such a way that, if induced, *goi* antisense mRNA will be produced.

Antisense RNA was found to play a key role in some regulation systems (Green *et al.*, 1986; Simons, 1988). Inhibition of gene expression by antisense RNA was observed both at the transcriptional and translational levels. An example for the latter case is the masking of the ribosome binding site on IS*10* transposase mRNA by an antisense transcript, thereby preventing translation (Simons & Kleckner, 1983; Simons & Kleckner, 1988). These natural antisense systems, however, contain specialized secondary structural features which seem to be essential for antisense inhibition (Simons, 1988). Whether or not a given artificial antisense RNA will function by those mechanisms is difficult to predict (Inouye, 1988). Nevertheless, it has been possible to demonstrate inhibition of target gene expression by artificial antisense RNAs, at least in cases in which the antisense RNA was in large excess over the target RNA (Coleman *et al.*, 1984; Pestka *et al.*, 1984; Green *et al.*, 1986; van der Krol *et al.*, 1988; Hélène & Toulmé, 1990). In addition to inhibitory effects of transcribed antisense RNA, convergent transcription *per se* may interfere with gene expression. The causes of this phenomenon, referred to as 'transcriptional interference', are probably collisions of transcribing polymerases (Elledge & Davis, 1989). An antisense promoter has already been applied successfully to reduce the basal expression levels from a repressed T7 promoter (Studier *et al.*, 1990).

But why should one reduce *goi* expression by antisense transcription? The rationale behind this concept is to allow for a time period during which T7 RNA polymerase is synthesized in the absence of the Phe analogue. During that phase,

goi expression should still be repressed, since synthesis of phenylalanine-containing target protein must be avoided. This is achieved in version B by the onset of antisense RNA transcription from the λ P_L promoter and in parallel by repression of the T7 promoter with *lac* repressor (Fig. 4). After addition of the amino acid analogue, *goi* expression is induced by adding IPTG as well as rifampicin (which then blocks antisense transcription).

Plasmid 2

The second plasmid is based on the same vector and T7 RNA polymerase gene as plasmid 2 in the preceding version A (section 3.1.a). Deviating from that system is, however, the regulation of T7 RNA polymerase gene expression. To hold down the T7 RNA polymerase levels, the very tightly repressible lambda P_L regulation system is used (Denhardt & Colasanti, 1988; Brosius, 1988). The function of the λ P_L promoter region (available from Pharmacia, Uppsala, Sweden) was described by Friedman & Gottesman (1983). The antitermination modules (*nutL*, *N* and t_{L1}) have been analyzed by Drahos & Szybalski (1981).

The λ P_L promoter is repressed at low temperatures by the plasmid-encoded λ *c*I857 repressor (Tsurimoto *et al.*, 1982) which is provided from a constitutive promoter (*e. g.* the *lacI*^{q} promoter or P_{con}; Calos (1978) and Deuschle *et al.* (1986), respectively). This thermosensitive repressor binds to the λ operator site(s) (o_L), thereby preventing transcription of gene *N* and the T7 RNA polymerase gene. A transcription terminator upstream of the λ P_L promoter also reduces background transcription activity originating outside of λ P_L. Occasional transcription starting at the repressed λ P_L promoter is terminated with 80-90 % efficiency (Drahos & Szybalski, 1981) at the rho-dependent t_{L1} terminator in front of the polymerase gene. Thus, both terminators serve to further lower the basal T7 RNA polymerase expression level.

Induction of the λ P_L promoter by raising the temperature to 42°C leads to expression of gene *N* while transcription still terminates at t_{L1}. Gene *N* encodes an antiterminator protein. As soon as a certain cellular level of antiterminator is available, it enables the host RNA polymerase to read through the transcription terminators (as t_{L1}) downstream of *nutL* (Friedman & Gottesman, 1983; Roberts, 1988). This allows full expression of the T7 RNA polymerase after a delay of a few minutes (Luzzati, 1970; Kourilsky *et al.*, 1971; Court *et al.*, 1980).

Table 2: *Succession of stages of the incorporation system, version B (cf. Fig. 4)*

Stage	Action	State of the cell/immediate reaction	Subsequent effects/results
I	Growth in minimal medium, (+ glucose, Amp, Cam, 30°C) to late exponential phase	-T7 RNA polymerase expression blocked by λ cI^{ts} repressor and termination at t_{L1} -*lac* repressor blocks T7 promoter	Target gene *goi* not expressed
II	Temperature shift 30° → 42°C	1.) Phe biosynthesis (ts) interrupted	Drainage of cellular Phe pool
		2.) λ repressor (ts) inactivated	λ P_L promoters switched on: a) Expression of antiterminator (λ gene *N*) → read-through at terminator t_{L1} → transcription of T7 RNA polymerase gene b) *goi* expression remains still blocked by - bound *lac* repressor - strong antisense transcription from induced λ P_L
III	After X min: addition of Rifampicin, IPTG and the Phe analogue	1.) Host RNA polymerase blocked	a) *goi* antisense transcription (=inhibition) stopped b) All transcription activity concentrated on T7 promoter c) Protein biosynthetic apparatus protected against adverse effects of Phe analogue incorporation
		2.) IPTG inactivates *lac* repressor	Expression of *goi* by T7 RNA polymerase
		3.) Charging of $tRNA^{Phe}$ with Phe analogues only	Incorporation of Phe analogues exclusively into the target protein encoded by the only gene transcribed: *goi*
IV	After 2-6 h: harvesting of the culture and purification of the target protein		

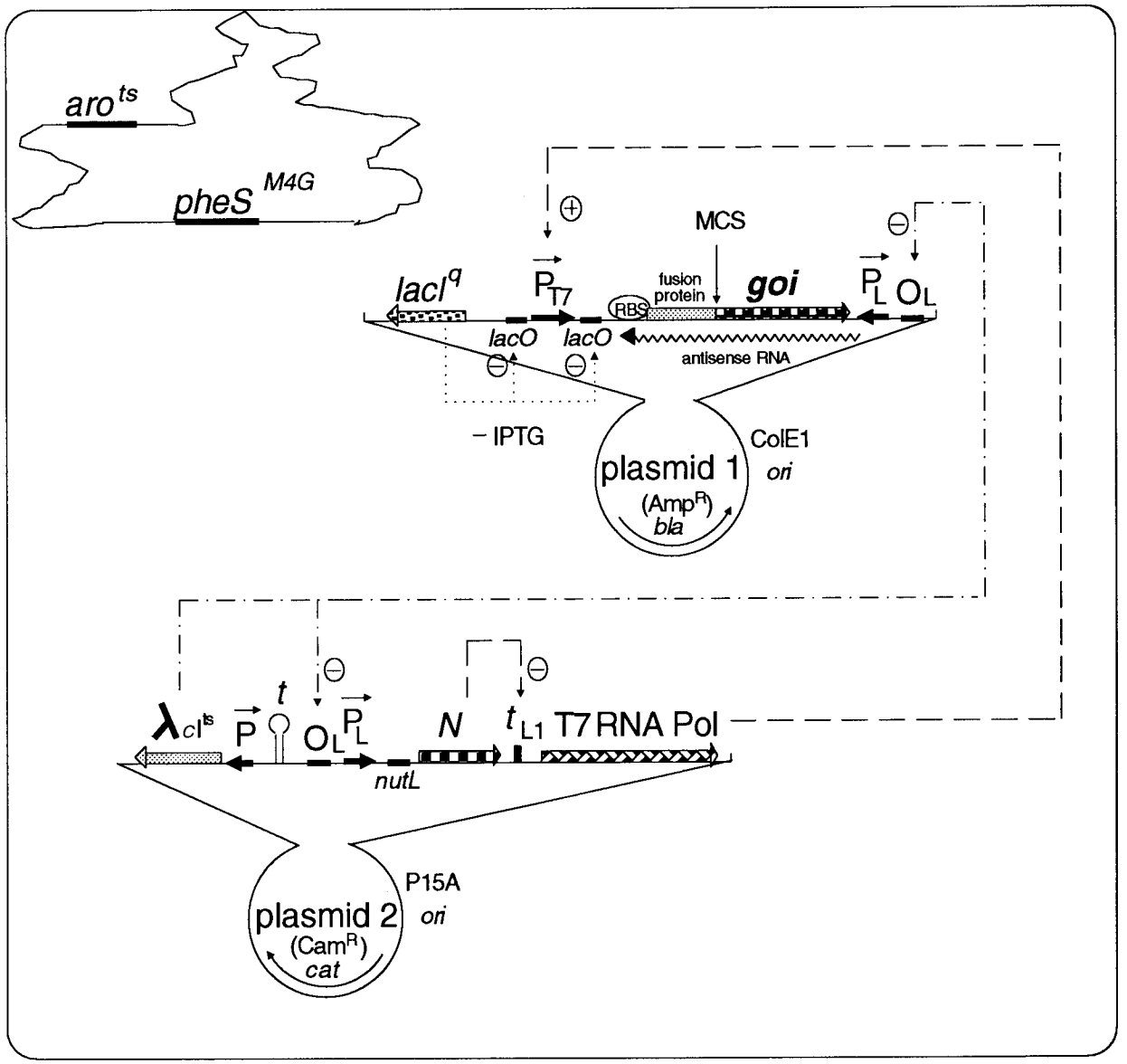


Figure 4. A sophisticated Phe analogue incorporation system, version B. The system is described in detail in the text and in Table 2. In the repressed state (at low temperature), the constitutively produced thermosensitive λ repressor *c*I857 (λ *c*I^{ts}) blocks transcription from the λ promoters (λ P_L). This prevents transcription of the genes for the antiterminator protein (encoded by *N*) and T7 RNA polymerase. Background transcription from λ P_L in front of *N* is terminated with high frequency at the terminator t_{L1}. Induction by a temperature upshift inactivates λ *c*I^{ts} and therefore causes transcription at the λ promoters. This leads to production of the antiterminator which enables the host RNA polymerase, after passing *nutL* (*N*-utilization site), to read through t_{L1}, causing transcription of the T7 RNA polymerase gene. At the same time, antisense transcription designed to interfere with *goi* expression starts from the λ P_L promoter on plasmid 1. Repression of the T7 promoter by *lac* repressor further prevents expression of *goi* until IPTG and rifampicin are added together with the Phe analogue. Symbols and features not explained here are described in the legend to Fig. 3.

Host strain

The host strain for version B is identical to the strain described for version A (section 3.1.a.).

(b) Successive steps in the incorporation system, version B

These are listed in Table 2. For additional information consult Fig. 4.

(c) Comments to version B

The system of version B is an improvement over version A in that more care is taken to avoid basal transcription of *goi* in the uninduced state. In addition, analogue incorporation into T7 RNA polymerase will be diminished by allowing a phase during which its gene is expressed in the absence of the analogue. At the same time *goi* expression is inhibited by repression (*lac* repressor) and convergent antisense transcription. Finally, target protein synthesis is fully induced in the presence of the amino acid analogue by addition of rifampicin and IPTG. The antisense promoter is designed here to be induced together with the expression of the T7 RNA polymerase gene. This has the advantage that there may be no deleterious influence of massive transcription from the antisense promoter on bacterial growth or plasmid stability during phase I (Table 2). If it turns out that the uninduced level of *goi* expression is still too high, a replacement of the antisense λ P_L promoter by a constitutive strong promoter could be considered (*e. g.* by P_{A1}; Deuschle *et al.*, 1986). This should further lower the basal *goi* expression levels by maintaining constant antisense transcription inhibition. Any interference with plasmid replication may be circumvented by placing a transcription terminator between promoter and origin of replication (Gentz *et al.*, 1981; Stueber & Bujard, 1982; Brown *et al.*, 1990).

The crucial point of this (and the other) proposed system(s) is a precise timing at which one stage is switched to the next. For instance, it seems important that T7 RNA polymerase synthesis is induced as long as there is still sufficient Phe in the cell. Systematic empirical tests and subsequent analyses of intermediates (*e. g.* the amount of *goi* mRNA) will be needed to work out an optimal timing.

3.3. Version C

(a) Description of the functional elements of the system (Fig. 5)

Plasmid 1 (contains the target gene)

The plasmid containing the target gene *goi* is identical to that described in version B (section 3.2.a.)

Plasmid 2

As version C uses a host with a *pheS*^{ts} mutation on the chromosome, the *pheS* allele *pheS*^{M4G} (providing relaxed substrate specificity to PheRS) has to reside on the plasmid. Previous experiments (see Chapter III) have shown that expression of *pheS* alone may not lead to strong overexpression of PheRS α subunits. To make sure that most PheRS molecules contain *pheS*^{M4G}-encoded α subunits by the time of Phe analogue addition, the *pheT* gene (encoding the PheRS β subunit) is present as well (Fig. 5). Coexpression of *pheS* and *pheT* should lead to considerable PheRS synthesis (Plumbridge & Springer, 1982; Fayat *et al.*, 1983). To avoid deleterious effects of overproduced PheRS (with relaxed specificity) on growth of the cells, the plasmidial *pheST* operon is expressed just before Phe analogue incorporation starts. This is achieved by placing the genes under the control of the inducible λ P_L promoter. The other features of plasmid 2 are identical to the corresponding plasmid described for version B in section 3.2.a. Expression of the T7 RNA polymerase gene by temperature upshift will at the same time lead to the synthesis of relaxed-specificity PheRS (Fig. 5; Table 3).

Host strain

In contrast to versions A and B, the host strain for version C does not need to contain a chromosomal *pheS*^{M4G} mutant gene responsible for the relaxed-specificity PheRS. Instead, it carries the *pheS* mutation D98 from strain NP37 (or KA2) leading to a thermosensitive enzyme (see Chapter II). By raising the cultivation temperature, the PheRS enzymes encoded by the chromosomal *pheST* genes are inactivated (as discussed in Chapter III). Charging of $tRNA^{Phe}$ will then be accomplished by the plasmid-encoded PheRS. Other suitable markers of the host strain are as described for version A (section 3.1.a.).

Table 3: *Succession of stages of the incorporation system, version C (cf. Fig. 5)*

Stage	Action	State of the cell/immediate reaction	Subsequent effects/results
I	Growth in minimal medium, (+ glucose, Amp, Cam, 30°C) to late exponential phase	-T7 RNA polymerase expression blocked by λ cI^{ts} repressor and termination at t_{L1} -*lac* repressor blocks T7 promoter	Target gene *goi* not expressed
II	Temperature shift 30° → 42°C	1.) Phe biosynthesis (ts) interrupted	Drainage of cellular Phe pool
		2.) λ repressor (ts) inactivated	λ P_L promoters switched on: a)Synthesis of relaxed-specificity PheRS b)Expression of antiterminator (λ gene *N*) → read-through at terminator t_{L1} → transcription of T7 RNA polymerase gene c)*goi* expression remains still blocked by - bound *lac* repressor - strong antisense transcription from induced λ P_L
		3.) PheRS containing chromosomal $pheS^{ts}$ encoded α subunits is inactivated	$tRNA^{Phe}$ is charged by plasmid-encoded PheRS
III	After X min: addition of Rifampicin, IPTG and the Phe analogue	1.)Host RNA polymerase blocked	a)*goi* antisense transcription (=inhibition) stopped b)All transcription activity concentrated on T7 promoter c)Protein biosynthetic apparatus protected against adverse effects of Phe analogue incorporation
		2.)IPTG inactivates *lac* repressor	Expression of *goi* by T7 RNA polymerase
		3.)Charging of $tRNA^{Phe}$ with Phe analogues only	Incorporation of Phe analogues exclusively into the target protein encoded by the only gene transcribed: *goi*
IV	After 2-6 h: harvesting of the culture and purification of the target protein		

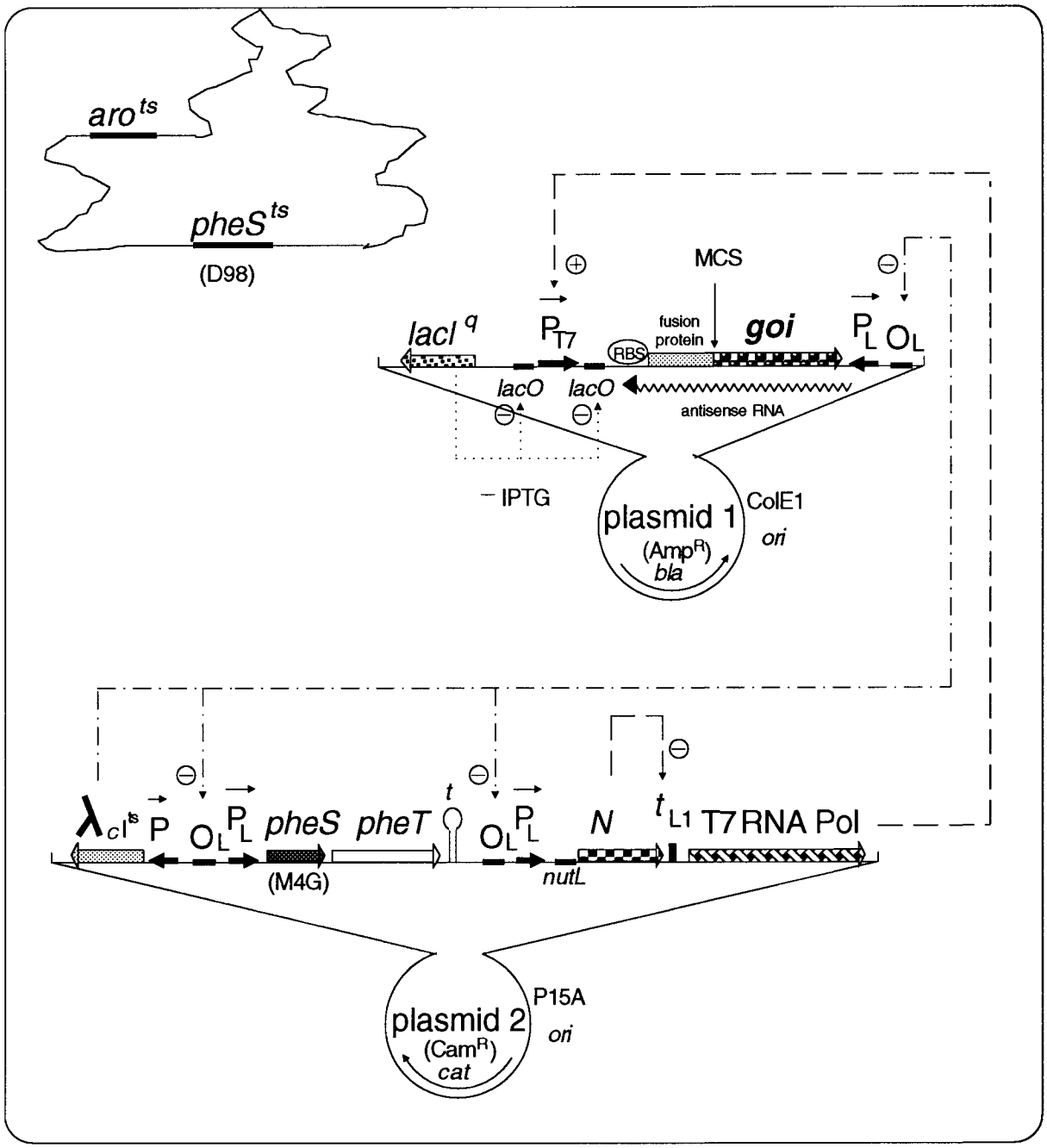


Figure 5. A sophisticated Phe analogue incorporation system, version C. The system is described in detail in the text and in Table 3. The individual elements of this Figure are essentially as specified in the legend to Fig. 4, except that the gene for the PheRS α subunit causing relaxed substrate specificity does not reside on the host chromosome. Instead, a PheRS thermosensitive strain (*pheS*^{ts}) is used which carries the mutation D98. Shortly before Phe analogue incorporation is required, the PheRS genes *pheS*^{M4G} (encoding the α subunit variant Gly^{294}) and *pheT* (for the β subunit) on plasmid 2 are induced by a temperature upshift. At the same time, the chromosomally-encoded PheRS is thermoinactivated. Therefore, tRNA^{Phe} is now predominantly charged by the relaxed-specificity PheRS variant Gly^{294}, leading to Phe analogue incorporation into the target protein encoded by *goi*.

(b) Successive steps in the incorporation system, version C

These are listed in Table 3. For additional information consult Fig. 5.

(c) Comments to version C

Version C differs from version B by the fact that the relaxed-specificity PheRS is encoded by plasmid 2 instead of by the host chromosome. This has on one hand the advantage of simplifying strain constructions in that the *pheS*^{M4G} allele needs not to be integrated at the corresponding chromosomal *pheS* locus whilst a *pheS*^{ts} mutation is already present on the chromosome. On the other hand, the presence of *pheST* on the plasmid complicates the system. It may also be more difficult to construct or handle plasmid 2. Other features of version C are as discussed in section 3.2.c.

3.4. Version D

(a) Description of the functional elements of the system (Fig. 6)

Plasmid 1 (contains the target gene)

The plasmid carrying *goi* is a simplified version of plasmid 1 described in section 3.1.a. (version A). It contains the same signals to express *goi* but the T7 promoter is not repressible.

Plasmid 2

The second plasmid is similar to the one described for version C (see section 3.3.a.). In this case, however, it is of primary importance that the basal (repressed) level of transcription of the genes controlled by the λ P_L promoters is extremely low. A tight repression may be brought about by inserting transcription terminators in front of the promoters to avoid transcription originating from upstream sequences. Another possibility is the construction of highly efficient artificial operator sites in λ P_L (Gussin *et al.*, 1983). Once induced by a temperature upshift, the plasmidial *pheS*^{M4G}*pheT* genes are expressed from the first λ P_L promoter. Transcription originating from the second λ promoter proceeds through gene *N* and halts at efficient terminators in front of the T7 RNA polymerase gene. Only when enough antiterminator has accumulated, this termination is overcome and the polymerase can be synthesized which in turn starts transcribing *goi*. The delay in T7

RNA polymerase synthesis is a crucial feature of this system (see below).

Host strain

The host strain matches with the criteria given for the strain described in version A (section 3.1.a.) except that it carries a *pheS* allele conferring resistance to the Phe analogue used. In a first stage, this allows growth of host cells in the presence of the analogue. When *p*-Cl-Phe is used as an analogue, for instance, the wild-type α subunit variant (Ala^{294}) would be sufficient to prevent incorporation of this analogue (Chapter III). In case of *p*-F-Phe being the analogue, the Ser^{294} mutant α subunit (encoded by *pheS* with mutation S294; see Chapter II) has to be present. In both cases, it would be an advantage if the corresponding *pheS*-encoded α subunits happened to be thermosensitive as well. This would allow thermoinactivation of the chromosomally-encoded PheRS at the stage during which Phe analogues have to be incorporated into the target protein (see below). When *p*-Cl-Phe is the analogue, the already existing chromosomal $pheS^{ts}$ allele carrying mutation D98 (described in section 3.3.a.) could be used. In the case of *p*-F-Phe, a strain carrying the *pheS* allele with mutations D98 and S294 would have to be constructed.

(b) Successive steps in the incorporation system, version D

These are listed in Table 4. For additional information consult Fig. 6.

(c) Comments to version D

Although the substrate analogue is meant to be incorporated into the target protein only, it is present during all growth stages. This characteristic feature of version D may be an advantage if the Phe analogue has a poor solubility in water, as is often the case for aromatic amino acids. In order to allow growth of cells in the presence of the analogue, conditions must be established preventing a detrimental incorporation into cellular protein in stage I. This can be achieved by using a PheRS variant which excludes the amino acid analogue from the enzymatic reaction. In some cases the wild-type enzyme having a more narrow substrate range than mutant Gly^{294} will suffice. For *p*-F-Phe incorporation, the enzyme of choice is the PheRS variant containing the α subunit encoded by the *pheS* allele with mutation S294 (Chapter II).

Table 4: *Succession of stages of the incorporation system, version D (cf. Fig. 6)*

Stage	Action	State of the cell/immediate reaction	Subsequent effects/results
I	Growth in minimal medium + Phe analogue (+ glucose, Amp, Cam, 30°C) to late exponential phase	-T7 RNA polymerase expression blocked by λ *cI*^{ts} repressor and transcription termination -λ *cI*^{ts} repressor prevents expression of genes for relaxed-specificity PheRS	- Target gene *goi* not expressed - Analogue-resistant PheRS prevents Phe analogue incorporation
II	Temperature shift 30° → 42°C	1.) Phe biosynthesis (ts) interrupted	Drainage of cellular Phe pool
		2.) λ repressor (ts) inactivated	λ P_L promoters switched on: a)Synthesis of relaxed-specificity PheRS b)Expression of antiterminator (λ gene *N*) → (delayed) read-through at terminators around t_{L1} → (delayed) transcription of T7 RNA polymerase gene → (delayed) expression of target gene *goi*
		3.) PheRS containing chromosomal *pheS*^{ts}-encoded α subunits is inactivated	$tRNA^{Phe}$ is charged by plasmid-encoded PheRS only → transition from Phe to Phe analogue incorporation
III	After X min: addition of Rifampicin	-Host RNA polymerase blocked	a)All transcription activity concentrated on T7 promoter b)Protein biosynthetic apparatus protected against adverse effects of Phe analogues, since they are exclusively incorporated into the protein encoded by the only gene transcribed: *goi*
IV	After 2-6 h: harvesting of the culture and purification of the target protein		

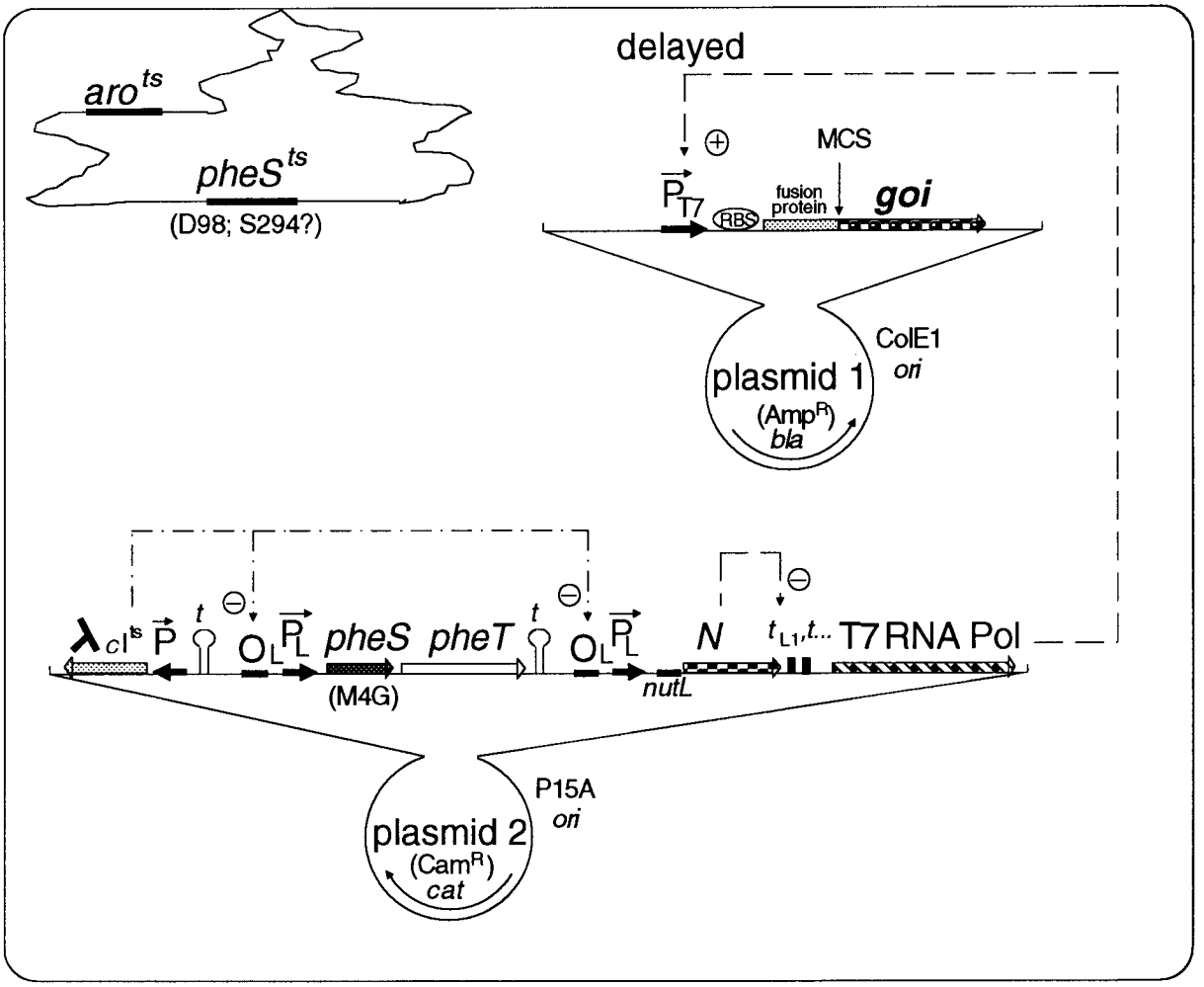


Figure 6. A sophisticated Phe analogue incorporation system, version D. The system is described in detail in the text and in Table 4. The elements in this Figure are explained in the legends to Figs. 4 and 5. Instead of preventing background transcription from the T7 promoter directly by repression and antisense transcription, more effort is put into lowering T7 RNA polymerase synthesis in the uninduced state. This is accomplished by inserting more transcription terminators in front of the polymerase gene, which have to be overcome with the help of the antiterminator (*N* gene product).

The peculiarity of this system is the shift in substrate specificity of PheRS, allowing growth of the cells in the presence of the Phe analogue. At the beginning of cell growth, only the chromosomally-encoded PheRS is expressed, which prevents deleterious incorporation of the analogue into cellular protein. In case that *p*-F-Phe is to be incorporated into the target protein, a *pheS* allele carrying mutation S294 (*p*-F-Phe resistance phenotype) ought to be used. Raising the temperature leads to expression of the genes for the relaxed-specificity PheRS variant Gly^{294} and to inactivation of chromosomally-encoded thermosensitive PheRS. In addition, the cellular Phe pool becomes exhausted. After a delay of a few minutes, the newly formed antiterminator permits transcription of the T7 RNA polymerase gene and leads to *goi* expression. Since $tRNA^{Phe}$ is now charged with the prevailing Phe analogue (by the Gly^{294} variant PheRS), the analogue will be incorporated into the target protein instead of Phe.

At the time of expression of *goi*, the analogue-excluding PheRS has to be replaced by the relaxed-specificity enzyme which is encoded by the plasmid (as in version C). To ensure that the specificity switch in PheRS activity occurs before *goi* is expressed, transcription of the T7 RNA polymerase gene is delayed relative to the plasmidial *pheS*^{M4G}*pheT* expression. This is accomplished by the termination-antitermination mechanism involving gene *N* as described above. During the resulting lag of *goi* expression, tRNA^{Phe} can be charged with the Phe analogue to obtain maximal analogue substitution yields.

3.5. Isolation and purification of the target protein

For the purification of the analogue-substituted proteins out of a bacterial culture, a high absolute yield is of great advantage. Although the systems described above are designed for high overproduction, one should be aware of the fact that a variety of factors determine the final yield. Apart from the genetic constellation of the system (plasmids and host strains), these include effects of the individual target proteins on the cells as well as of specific fermentation conditions (Balbas & Bolivar, 1990). Therefore, it may pay off to optimize at least the cultivation parameters, before one invests much effort in isolating proteins of poor yield.

It is, of course, not possible to give a general scheme for the isolation and purification of the target proteins. Each protein has to be treated individually, according to its characteristic physico-chemical and biological properties. Certain general strategies might simplify the isolation and purification procedures such as the secretion of the target protein (Stader & Silhavy, 1990) or the fusion of it to other easily purifiable proteins (Josephson & Bishop, 1988; Smith & Johnson, 1988; Nilsson & Abrahmsén, 1990; Uhlén & Moks, 1990).

4. Applications

The analogue incorporation systems proposed in section 3 are of interest in two respects. First, new experience will be gained concerning the assembly of individual control elements for gene expression. It will be examined, if and how different regulatory circuits can act together in more complex systems. The results of these experiments could thus have an impact on the 'genetic engineering' of other gene expression systems.

Second, a principal aspect of the designed strategies is to put efficient analogue incorporation to work *in vivo*. Amino acid analogue-containing proteins are of interest both for research purposes as well as for potential pharmaceutical application. Examples for the use of amino acid analogues in proteins for structure-function studies and as reporter groups for nuclear magnetic resonance spectroscopy have been listed in section 1.2. Besides these rather analytical applications, one can exploit the potential to incorporate amino acid analogues in order to intentionally produce novel proteins. It seems reasonable to speculate that the replacement of one type of the 20 naturally encoded amino acids (plus Se-Cys; Leinfelder *et al.*, 1988; Söll, 1988; Böck *et al.*, 1991) by an analogue may create proteins with uniquely modified functions.

4.1. Potential target proteins of pharmaceutical interest

Proteins and small peptides of pharmaceutical interest are among the primary targets for amino acid analogue incorporation by the types of systems described in this Chapter. Substitution of natural amino acid residues by analogues could result in altered biological and physiological properties: The turnover rate of a peptide drug could be different in the analogue substituted variant due to increased or decreased resistance towards proteolytic degradation (Wilson & Hatfield, 1984). The modification in the protein structure could alter the receptor recognition. This may lead to a narrowed or broadened target range of action for a peptide pharmaceutical.

Examples of target proteins suitable for analogue incorporation can be found among polypeptide growth factors (James & Bradshaw, 1984) and endocrine peptide hormones (Douglass *et al.*, 1984; Lynch & Snyder, 1986). The activities of

analogue-substituted peptide hormones may be altered, as exemplified for *p*-F-Phe-containing angiotensin II or bradykinin (Vine *et al.*, 1973). Koide *et al.* (1988) compared the biological activities of a native and an analogue-substituted form of human epidermal growth factor. Many other proteins of these classes, which can be produced by recombinant DNA technology in bacteria (Josephson & Bishop, 1988; Uhlén & Moks, 1990), should also be suitable as targets in an analogue incorporation system.

4.2. Further perspectives

The incorporation systems described in section 3 are designed to incorporate Phe analogues, because of the availability of the relaxed-specificity PheRS variant Gly^{294}. Therefore, the native target proteins have to contain at least one Phe residue in order to allow production of novel polypeptide variants. Slightly modified versions of these systems, however, could be used for incorporation of analogues of other amino acids. In such cases, one could make use of the natural misincorporation spectrum of some aminoacyl-tRNA synthetases (Hortin & Boime, 1983). Alternatively, it could also be attempted to find broad specificity mutants of those enzymes and utilize them in a similar manner as proposed here for the PheRS mutant Gly^{294}.

5. References

Amann, E., Brosius, J. & Ptashne, M. (1983). Vectors Bearing a Hybrid *trp-lac* Promoter Useful for Regulated Expression of Cloned Genes in *Escherichia coli*. *Gene*, **25**, 167-178.

Bain, J. D., Wacker, D. A., Kuo, E. E. & Chamberlin, A. R. (1991). Site-Specific Incorporation of Non-Natural Residues into Peptides: Effect of Residue Structure on Suppression and Translation Efficiencies. *Tetrahedron*, **47**, 2389-2400.

Balbas, P. & Bolivar, F. (1990). Design and Construction of Expression Plasmid Vectors in *Escherichia coli*. *Methods Enzymol.* **185**, 14-37.

Beckwith, J. (1987). The Lactose Operon. In *Escherichia coli and Salmonella typhimurium*: *Cellular and Molecular Biology* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M. & Umbarger, H. E., eds), vol. 2, pp. 1444-1452, Am. Soc. Microbiol., Washington, DC.

Besse, M., von Wilcken-Bergmann, B. & Müller-Hill, B. (1986). Synthetic *lac* Operator Mediates Repression through *lac* Repressor When Introduced Upstream and Downstream from *lac* Promoter. *EMBO J.* **5**, 1377-1381.

Beyreuther, K. (1978). Chemical Structure and Functional Organization of *lac* Repressor from *Escherichia coli*. In *The Operon* (Miller, J. H. & Reznikoff, W. S., eds), pp. 123-154, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Böck, A., Forchhammer, K., Heider, J., Leinfelder, W., Sawers, G., Veprek, B. & Zinoni, F. (1991). Selenocysteine: the 21st Amino Acid. *Mol. Microbiol.* **5**, 515-520.

Brosius, J. (1988). Expression Vectors Employing Lambda-, *trp*-, *lac*-, and *lpp*-Derived Promoters. In *Vectors. A Survey of Molecular Cloning Vectors and Their Uses* (Rodriguez, R. L. & Denhardt, D. T., eds), pp. 205-225, Butterworth Publishers, Stoneham, MA.

Brown, K. D. (1970). Formation of Aromatic Amino Acid Pools in *Escherichia coli* K-12. *J. Bacteriol.* **104**, 177-188.

Brown, S., Ferm, J., Woody, S. & Gussin, G. (1990). Selection for Mutations in the

P_R Promoter of Bacteriophage Lambda. *Nucl. Acids Res.* **18**, 5961-5967.

Calos, M. P. (1978). DNA Sequence for a Low-Level Promoter of the *lac* Repressor Gene and an 'Up' Promoter Mutation. *Nature (London),* **274**, 762-765.

Camakaris, H. & Pittard, J. (1983). Tyrosine Biosynthesis. In *Amino Acids: Biosynthesis and Genetic Regulation* (Herrmann, K. M. & Somerville, R. L., eds), pp. 339-350, Addison-Wesley Publishing Company, London.

Chang, A. C. Y. & Cohen, S. N. (1978). Construction and Characterization of Amplifiable Multicopy DNA Cloning Vehicles Derived from the P15A Cryptic Miniplasmid. *J. Bacteriol.* **134**, 1141-1156.

Chang, C. N., Kuang, W.-J. & Chen. E. Y. (1986). Nucleotide Sequence of the Alkaline Phosphatase Gene of *Escherichia coli. Gene,* **44**, 121-125.

Coleman, J., Green, P. J. & Inouye, M. (1984). The Use of RNAs Complementary to Specific mRNAs to Regulate the Expression of Individual Bacterial Genes. *Cell,* **37**, 429-436.

Court, D., Brady, C., Rosenberg, M., Wulff, D. L., Behr, M., Mahoney, M. & Izumi, S. (1980). Control of Transcription Termination: A Rho-Dependent Termination Site in Bacteriophage Lambda. *J. Mol. Biol.* **138**, 231-254.

Denhardt, D. T. & Colasanti, J. (1988). A Survey of Vectors for Regulating Expression of Cloned DNA in *E. coli.* In *Vectors. A Survey of Molecular Cloning Vectors and Their Uses* (Rodriguez, R. L. & Denhardt, D. T., eds), pp. 179-203, Butterworth Publishers, Stoneham, MA.

Deuschle, U., Kammerer, W., Gentz, R. & Bujard, H. (1986). Promoters of *Escherichia coli*: A Hierarchy of *in Vivo* Strength Indicates Alternate Structures. *EMBO J.* **5**, 2987-2994.

Douglass, J., Civelli, O. & Herbert, E. (1984). Polyprotein Gene Expression: Generation of Diversity of Neuroendocrine Peptides. *Ann. Rev. Biochem.* **53**, 665-715.

Drahos, D. & Szybalski, W. (1981). Antitermination and Termination Functions of the Cloned *nut*L, *N*, and t_{L1} Modules of Coliphage Lambda. *Gene,* **16**, 261-274.

Elledge, S. J. & Davis, R. W. (1989). Position and Density Effects on Repression by Stationary and Mobile DNA-Binding Proteins. *Genes Dev.* **3**, 185-197.

Fangman, W. L. & Neidhardt, F. C. (1964). Protein and Ribonucleic Acid Synthesis in a Mutant of *Escherichia coli* with an Altered Aminoacyl Ribonucleic Acid Synthetase. *J. Biol. Chem.* **239**, 1844-1847.

Fayat, G., Fromant, M., Kalogerakos, T. & Blanquet, S. (1983). Effect of the Overproduction of Phenylalanyl- and Threonyl-tRNA Synthetases on $tRNA^{Phe}$ and $tRNA^{Thr}$ Concentrations in *E. coli* cells. *Biochimie*, **65**, 221-225.

Fowden, L., Lewis, D. & Tristram, H. (1967). Toxic Amino Acids: Their Action As Antimetabolites. *Adv. Enzymol.* **29**, 89-163.

Friedman, D. I. & Gottesman, M. (1983). Lytic Mode of Lambda Development. In *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W. & Weisberg, R. A., eds), pp. 21-51, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Fürste, J. P., Pansegrau, W., Frank, R., Blöcker, H., Scholz, P., Bagdasarian, M. & Lanka, E. (1986). Molecular Cloning of the Plasmid RP4 Primase Region in a Multi-Host-Range *tacP* Expression Vector. *Gene*, **48**, 119-131.

Gale, E. F. & Folkes, J. P. (1953). The Assimilation of Amino Acids by Bacteria. The Inhibition of Phenylalanine Incorporation in *Staphylococcus aureus* by Chloramphenicol and *p*-Chlorophenylalanine. *Biochem. J.* **55**, 730-735.

Garner, C. & Herrmann, K. M. (1983). Biosynthesis of Phenylalanine. In *Amino Acids: Biosynthesis and Genetic Regulation* (Herrmann, K. M. & Somerville, R. L., eds), pp. 323-338, Addison-Wesley Publishing Company, London.

Gavini, N. & Davidson, B. E. (1990). The *pheR* Gene of *Escherichia coli* Encodes $tRNA^{Phe}$, Not a Repressor Protein. *J. Biol. Chem.* **265**, 21527-21531.

Gentz, R., Langner, A., Chang, A. C. Y., Cohen, S. N. & Bujard, H. (1981). Cloning and Analysis of Strong Promoters Is Made Possible by the Downstream Placement of a RNA Termination Signal. *Proc. Nat. Acad. Sci., U.S.A.* 78, 4936-4940.

Goff, S. A. & Goldberg, A. L. (1985). Production of Abnormal Proteins in *E. coli* Stimulates Transcription of *lon* and Other Heat Shock Genes. *Cell*, **41**, 587-595.

Goff, S. A., Casson, L. P. & Goldberg, A. L. (1984). Heat Shock Regulatory Gene *hptR* Influences Rates of Protein Degradation and Expression of the *lon* Gene in *Escherichia coli. Proc. Nat. Acad. Sci., U.S.A.* **81**, 6647-6651.

Gottesman, S. (1990). Minimizing Proteolysis in *Escherichia coli*: Genetic Solutions. *Methods Enzymol.* **185**, 119-129.

Gralla, J. D. (1990). Promoter Recognition and mRNA Initiation by *Escherichia coli* $E\sigma^{70}$. *Methods Enzymol.* **185**, 37-54.

Green, P. J., Pines, O. & Inouye, M. (1986). The Role of Antisense RNA in Gene Regulation. *Ann. Rev. Biochem.* **55**, 569-597.

Gregory, H. (1975). Isolation and Structure of Urogastrone and its Relationship to Epidermal Growth Factor. *Nature (London)*, **257**, 325-327.

Grodberg, J. & Dunn, J. J. (1988). *ompT* Encodes the *Escherichia coli* Outer Membrane Protease That Cleaves T7 RNA Polymerase during Purification. *J. Bacteriol.* **170**, 1245-1253.

Grossman, A. D., Erickson, J. W. & Gross, C. A. (1984). The *hptR* Gene Product of *E. coli* Is a Sigma Factor for Heat-Shock Promoters. *Cell*, **38**, 383-390.

Gueguen, P., Padron, M., Perbal, B. & Hervé, G. (1980). Incorporation of Amino Acid Analogs during the Biosynthesis of *Escherichia coli* Aspartate Transcarbamylase. *Biochim. Biophys. Acta*, **615**, 59-69.

Gussin, G. N., Johnson, A. D., Pabo, C. O. & Sauer, R. T. (1983). Repressor and Cro Protein: Structure, Function, and Role in Lysogenization. In *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W. & Weisberg, R. A., eds), pp. 93-121, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Hélène, C. & Toulmé, J.-J. (1990). Specific Regulation of Gene Expression by Antisense, Sense and Antigene Nucleic Acids. *Biochim. Biophys. Acta*, **1049**, 99-125.

Hennecke, H., Günther, I. & Binder, F. (1982). A Novel Cloning Vector for the Direct Selection of Recombinant DNA in *E. coli*. *Gene*, **19**, 231-234.

Herrmann, K. M. (1983). The Common Aromatic Biosynthetic Pathway. In *Amino Acids: Biosynthesis and Genetic Regulation* (Herrmann, K. M. & Somerville, R. L., eds), pp. 301-322, Addison-Wesley Publishing Company, London.

Hiscott, J., Cantell, K. & Weissmann, C. (1984). Differential Expression of Human Interferon Genes. *Nucl. Acids Res.* **12**, 3727-3746.

Hortin, G. & Boime, I. (1983). Applications of Amino Acid Analogs for Studying Co- and Posttranslational Modifications of Proteins. *Methods Enzymol.* **96**, 777-784.

Inouye, M. (1988). Antisense RNA: Its Functions and Applications in Gene Regulation - a Review. *Gene*, **72**, 25-34.

James, R. & Bradshaw, R. A. (1984). Polypeptide Growth Factors. *Ann. Rev. Biochem.* **53**, 259-292.

Jarema, M. C., Lu, P. & Miller, J. H. (1981). Genetic Assignment of Resonances in the NMR Spectrum of a Protein: *lac* Repressor. *Proc. Nat. Acad. Sci., U.S.A.* **78**, 2707-2711.

Josephson, S. & Bishop, R. (1988). Secretion of Peptides from *E. coli*: A Production System for the Pharmaceutical Industry. *Trends Biotechnol.* **6**, 218-224.

Kohno, T., Kohda, D., Haruki, M., Yokoyama, S. & Miyazawa, T. (1990). Nonprotein Amino Acid Furanomycin, Unlike Isoleucine in Chemical Structure, Is Charged to Isoleucine tRNA by Isoleucyl-tRNA Synthetase and Incorporated into Protein. *J. Biol. Chem.* **265**, 6931-6935.

Koide, H., Yokoyama, S., Kawai, G., Ha, J.-M., Oka, T., Kawai, S., Miyake, T., Fuwa, T. & Miyazawa, T. (1988). Biosynthesis of a Protein Containing a Nonprotein Amino Acid by *Escherichia coli*: L-2-Aminohexanoic Acid at Position 21 in Human Epidermal Growth Factor. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 6237-6241.

Kourilsky, P., Bourguignon, M.-F. & Gros, F. (1971). Kinetics of Viral Transcription after Induction of Prophage. In *The Bacteriophage Lambda* (Hershey, A. D., ed), pp. 647-666, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Krämer, H., Amouyal, M., Nordheim, A. & Müller-Hill, B. (1988). DNA Supercoiling Changes the Spacing Requirement of Two *lac* Operators for DNA Loop Formation with *lac* Repressor. *EMBO J.* **7**, 547-556.

Krämer, H., Niemöller, M., Amouyal, M., Revet, B., von Wilcken-Bergmann, B. & Müller-Hill, B. (1987). *lac* Repressor Forms Loops with Linear DNA Carrying Two Suitably Spaced *lac* Operators. *EMBO J.* **6**, 1481-1491.

Leinfelder, W., Zehelein, E., Mandrand-Berthelot, M.-A. & Böck, A. (1988). Gene for a Novel tRNA Species That Accepts L-Serine and Cotranslationally Inserts Selenocysteine. *Nature (London)*, **331**, 723-725.

Lindquist, S. (1986). The Heat-Shock Response. *Ann. Rev. Biochem.* **55**, 1151-1191.

Lu, P., Jarema, M., Mosser, K. & Daniel, Jr., W. E. (1976). *lac* Repressor: 3-Fluorotyrosine Substitution for Nuclear Magnetic Resonance Studies. *Proc. Nat. Acad. Sci., U.S.A.* **73**, 3471-3475.

Luzzati, D. (1970). Regulation of λ Exonuclease Synthesis: Role of the *N* Gene Product and λ Repressor. *J. Mol. Biol.* **49**, 515-519.

Lynch, D. R. & Snyder, S. H. (1986). Neuropeptides: Multiple Molecular Forms, Metabolic Pathways, and Receptors. *Ann. Rev. Biochem.* **55**, 773-799.

Maurizi, M. R., Trisler, P. & Gottesman, S. (1985). Insertional Mutagenesis of the *lon* Gene in *Escherichia coli*: *lon* Is Dispensable. *J. Bacteriol.* **164**, 1124-1135.

Mizusawa, S. & Gottesman, S. (1983). Protein Degradation in *Escherichia coli*: The *lon* Gene Controls the Stability of SulA Protein. *Proc. Nat. Acad. Sci., U.S.A.* **80**, 358-362.

Moffatt, B. A., Dunn, J. J. & Studier, F. W. (1984). Nucleotide Sequence of the Gene for Bacteriophage T7 RNA Polymerase. *J. Mol. Biol.* **173**, 265-269.

Morris, H. & Schlesinger, M. J. (1972). Effects of Proline Analogues on the Formation of Alkaline Phosphatase in *Escherichia coli*. *J. Bacteriol.* **111**, 203-210.

Mossing, M. C. & Record, Jr., M. T. (1986). Upstream Operators Enhance Repression of the *lac* Promoter. *Science*, **233**, 889-892.

Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Escödi, J., Boll, W., Cantell, K. & Weissmann, C. (1980). Synthesis in *E. coli* of a Polypeptide with Human Leukocyte Interferon Activity. *Nature (London)*, **284**, 316-320.

Nilsson, B. & Abrahmsén, L. (1990). Fusions to Staphylococcal Protein A. *Methods Enzymol.* **185**. 144-161.

Olins, P. O. & Rangwala, S. H. (1989). A Novel Sequence Element Derived from Bacteriophage T7 mRNA Acts as an Enhancer of Translation of the *lacZ* Gene in *Escherichia coli*. *J. Biol. Chem.* **264**, 16973-16976.

Olins, P. O. & Rangwala, S. H. (1990). Vector for Enhanced Translation of Foreign

Genes in *Escherichia coli. Methods Enzymol.* **185**, 115-119.

Ozols, J. (1990). Amino Acid Analysis. *Methods Enzymol.* **182**, 587-601.

Palva, I., Lehtovaara, P., Kääriäinen, L., Sibakov, M., Cantell, K., Schein, C. H., Kashiwagi, K. & Weissmann, C. (1983). Secretion of Interferon by *Bacillus subtilis. Gene*, **22**, 229-235.

Panayotatos, N., Fontaine, A. & Truong, K. (1983). A Cop$^-$ Mutation Causes a Specific 200 Fold Increase in Interferon Levels Expressed from the ColE1 Replication Promoters. *J. Cell. Biochem., Suppl.* **7B**, 109.

Pestka, S, Daugherty, B. L., Jung, V., Hotta, K. & Pestka, R. K. (1984). Anti-mRNA: Specific Inhibition of Translation of Single mRNA Molecules. *Proc. Nat. Acad. Sci., U.S.A.* **81**, 7525-7528.

Phillips, T. A., VanBogelen, R. A. & Neidhardt, F. C. (1984). *lon* Gene Product of *Escherichia coli* Is a Heat-Shock Protein. *J. Bacteriol.* **159**, 283-287.

Pines, M., Rosenthal, G. A. & Applebaum, S. W. (1981). *In Vitro* Incorporation of L-Canavanine into Vitellogenin of the Fat Body of the Migratory Locust *Locusta migratoria migratorioides. Proc. Nat. Acad. Sci., U.S.A.* **78**, 5480-5483.

Pittard, A. J. (1987). Biosynthesis of the Aromatic Amino Acids. In *Escherichia coli and Salmonella typhimurium*: *Cellular and Molecular Biology* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M. & Umbarger, H. E., eds), vol. 1, pp. 368-394, Am. Soc. Microbiol., Washington, DC.

Pittard, J., Praszkier, J., Certoma, A., Eggertsson, G., Gowrishankar, J., Narasaiah, G. & Whipp, M. J. (1990). Evidence that There Are Only Two $tRNA^{Phe}$ Genes in *Escherichia coli. J. Bacteriol.* **172**, 6077-6083.

Plumbridge, J. A. & Springer, M. (1982). *Escherichia coli* Phenylalanyl-tRNA Synthetase Operon: Characterization of Mutations Isolated on Multicopy Plasmids. *J. Bacteriol.* **152**, 650-660.

Richmond, M. H. (1962). The Effect of Amino Acid Analogues on Growth and Protein Synthesis in Microorganisms. *Bacteriol. Rev.* **26**, 398-420.

Richmond, M. H. (1963). Random Replacement of Phenylalanine by *p*-Fluorophenylalanine in Alkaline Phosphatase(s) Formed during Biosynthesis by *E. coli. J. Mol. Biol.* **6**, 284-294.

Roberts, J. W. (1988). Phage Lambda and the Regulation of Transcription Termination. *Cell*, **52**, 5-6.

Robertson, S. A., Ellman, J. A. & Schultz, P. G. (1991). A General and Efficient Route for Chemical Aminoacylation of Transfer RNAs. *J. Am. Chem. Soc.* **113**, 2722-2729.

Rosenberg, A. H., Lade, B. N., Chui, D., Lin, S., Dunn, J. J. & Studier, F. W. (1987). Vectors for Selective Expression of Cloned DNAs by T7 RNA Polymerase. *Gene*, **56**, 125-135.

Scaife, J. (1976). Bacterial RNA Polymerases: The Genetics and Control of Their Synthesis. In *RNA Polymerase* (Losick, R. & Chamberlin, M., eds), pp. 207-225, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Schein, C. H. & Noteborn, M. H. M. (1988). Formation of Soluble Recombinant Proteins in *Escherichia coli* Is Favored by Lower Growth Temperature. *Bio/Technology*, **6**, 291-294.

Schlesinger, M. J., Reynolds, J. A. & Schlesinger, S. (1969). Formation and Localization of the Alkaline Phosphatase of *Escherichia coli*. *Ann. N. Y. Acad. Sci.* **166**, 368-379.

Shive, W. & Skinner, C. G. (1963). Amino Acid Analogues. In *Metabolic Inhibitors*, (Hochster, R. M. & Quastel, J. H., eds), vol. 1, pp. 1-73, Academic Press, London.

Simons, R. W. (1988). Naturally Occurring Antisense RNA Control - a Brief Review. *Gene*, **72**, 35-44.

Simons, R. W. & Kleckner, N. (1983). Translational Control of IS*10* Transposition. *Cell*, **34**, 683-691.

Simons, R. W. & Kleckner, N. (1988). Biological Regulation by Antisense RNA in Prokaryotes. *Ann. Rev. Genet.* **22**, 567-600.

Smith, D. B. & Johnson, K. S. (1988). Single-Step Purification of Polypeptides Expressed in *Escherichia coli* as Fusions with Glutathione S-Transferase. *Gene*, **67**, 31-40.

Söll, D. (1988). Enter a New Amino Acid. *Nature (London)*, **331**, 662-663.

Somerville, R. L. (1983). Tryptophan: Biosynthesis, Regulation, and Large-Scale

Production. In *Amino Acids: Biosynthesis and Genetic Regulation* (Herrmann, K. M. & Somerville, R. L., eds), pp. 351-378, Addison-Wesley Publishing Company, London.

Stader, J. A. & Silhavy, T. J. (1990). Engineering *Escherichia coli* to Secrete Heterologous Gene Products. *Methods Enzymol.* **185**, 166-187.

Stewart, W. E., II. (1981). *The Interferon System.* 2nd ed., Springer Verlag, New York.

Streuli, M., Nagata, S. & Weissmann, C. (1980). At Least Three Human Type α Interferons: Structure of α2. *Science*, **209**, 1343-1347.

Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA Polymerase to Direct Expression of Cloned Genes. *Methods Enzymol.* **185**, 60-89.

Stueber, D. & Bujard, H. (1982). Transcription from Efficient Promoters Can Interfere with Plasmid Replication and Diminish Expression of Plasmid Specified Genes. *EMBO J.* **1**, 1399-1404.

Sykes, B. D., Weingarten, H. I. & Schlesinger, M. J. (1974). Fluorotyrosine Alkaline Phosphatase from *Escherichia coli*: Preparation, Properties, and Fluorine-19 Nuclear Magnetic Resonance Spectrum. *Proc. Nat. Acad. Sci., U.S.A.* **71**, 469-473.

Tabor, S. & Richardson, C. C. (1985). A Bacteriophage T7 RNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes. *Proc. Nat. Acad. Sci., U.S.A.* **82**, 1074-1078.

Thatcher, D. R. & Panayotatos, N. (1986). Purification of Recombinant Human IFN-α2. *Methods Enzymol.* **119**, 166-177.

Tsurimoto, T., Hase, T., Matsubara, H. & Matsubara, K. (1982). Bacteriophage Lambda Initiators: Preparation from a Strain That Overproduces the *O* and *P* Proteins. *Mol. Gen. Genet.* **187**, 79-86.

Uhlén, M. & Moks, T. (1990). Gene Fusions for Purpose of Expression: An Introduction. *Methods Enzymol.* **185**, 129-143.

Valenzuela, D., Weber, H. & Weissmann, C. (1985). Is Sequence Conservation in Interferons Due to Selection for Functional Proteins? *Nature (London)*, **313**, 698-700.

van der Krol, A. R., Mol, J. N. M. & Stuitje, A. R. (1988). Antisense Genes in Plants: an Overview. *Gene*, **72**, 45-50.

Vine, W. H., Brueckner, D. A., Needleman, P. & Marshall, G. R. (1973). Synthesis, Biological Activity, and ^{19}F Nuclear Magnetic Resonance Spectra of Angiotensin II Analogs Containing Fluorine. *Biochemistry*, **12**, 1630-1637.

Weber, H., Valenzuela, D., Lujber, G., Gubler, M. & Weissmann, C. (1987). Single Amino Acid Changes That Render Human IFN-α2 Biologically Active on Mouse Cells. *EMBO J.* **6**, 591-598.

Wehrli, W., Handschin, J. & Wunderli, W. (1976). Interaction Between Rifampicin and DNA-Dependent RNA Polymerase of *E. coli*. In *RNA Polymerase* (Losick, R. & Chamberlin, M., eds), pp. 397-412, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Weissmann, C. & Weber, H. (1986). The Interferon Genes. In *Progr. Nucleic Acid Res. Mol. Biol.* (Cohn, W. E. & Moldave, K., eds), vol. 33, pp. 251-300, Academic Press, London.

Whipp, M. J. & Pittard, A. J. (1977). Regulation of Aromatic Amino Acid Transport Systems in *Escherichia coli* K-12. *J. Bacteriol.* **132**, 453-461.

Wilson, M. J. & Hatfield, D. L. (1984). Incorporation of Modified Amino Acids into Proteins *in Vivo*. *Biochim. Biophys. Acta*, **781**, 205-215.

CHAPTER V

Cloning of *pheS* Genes from *Thermus thermophilus* and *Salmonella typhimurium*

1. Introduction

In Chapter II, two different PheRS mutations were analyzed and the resulting amino acid exchanges were localized in the primary structure of the PheRS α subunit. As deduced from their phenotypes, both mutations were assumed to alter the structure of the Phe binding site. For one of these mutations this hypothesis was largely confirmed by the results of site-directed mutagenesis experiments described in Chapter III. The most interesting mutation (Ala→Gly^{294}) changed substrate specificity in a way that was predicted from a model for the putative Phe binding site.

A different and independent approach to localize functionally important regions consists of comparing the primary sequences of PheRS enzymes from as many different origins as possible (see below). It is clear, however, that the ultimate elucidation of the topology of a substrate binding site must await the determination of the three-dimensional structure.

1.1. Phylogenetically conserved sequences may identify functionally important regions

As pointed out in Chapter II, both amino acid exchanges responsible for altered amino acid substrate binding properties of PheRS variants mapped to two

Non-standard abbreviations

bp: basepair(s); EDTA: ethylenediaminetetraacetate; kb: kilobasepair(s); SDS: sodium dodecylsulfate; Tris: tris-(hydroxymethyl)-aminomethane; ts: thermosensitive; XxxRS: aminoacyl-tRNA synthetase for amino acid Xxx in three-letter-code.

generally conserved sequence motifs. The presence of these two consensus sequences (motifs 2 and 3), together with another sequence (motif 1), is characteristic for members of class II aminoacyl-tRNA synthetases (Eriani *et al.*, 1990). The fact that motifs 2 and 3 were affected by the mutations could be interpreted to mean that those motifs are involved in building the amino acid binding site in each individual class II member (Chapter II). While the binding sites for different amino acids in various class II enzymes may share common features, this should even more be true for aminoacyl-tRNA synthetases with specificity for the same amino acid, regardless of the organismic origin. In fact, as emphasized in Chapter II, the amino acid sequence around the motif 3 residue hit by mutation S294 showed a particularly good conservation in all compared PheRS α subunits of different origin.

Apart from the PheRS α subunit of *Escherichia coli* (Fayat *et al.*, 1983), the sequences of the *Bacillus subtilis* (Brakhage *et al.*, 1990) and *Saccharomyces cerevisiae* cytoplasmic (Sanni *et al.*, 1988) and mitochondrial (Koerner *et al.*, 1987) PheRS small subunits were known. To make more sequences available for comparisons, two additional organisms were chosen for cloning and sequencing of their *pheS* genes. One choice was the enterobacterium *Salmonella typhimurium* because this organism is very closely related to *E. coli* (Fig. 1). Therefore, it was rationalized that amino acid differences in the PheRS α subunits of these two organisms might identify residues that are not essential for PheRS function. The second species chosen for the isolation of its PheRS genes was *Thermus thermophilus*. *T. thermophilus* is a Gram-negative, nonsporulating aerobic rod. Growth of this extremely thermophilic eubacterium is optimal at temperatures between 65°C and 72°C, but can continue up to 85°C (Oshima & Imahori, 1974). *T. thermophilus* is only distantly related to *E. coli*. In fact, based on 16S rRNA comparisons, the genus *Thermus* was placed within the second deepest eubacterial branching known so far (Fig. 1; Hartmann *et al.*, 1989). Therefore, it was assumed that the amino acid sequence would show considerable differences in regions not essential for PheRS function whereas invariant residues would be indicative of functionally important sequences. Moreover, the *Thermus* enzyme was expected to have been optimized for thermostability and thus to contain additional amino acid replacements that reflect this property (Argos *et al.*, 1979; Menéndez-Arias & Argos, 1989).

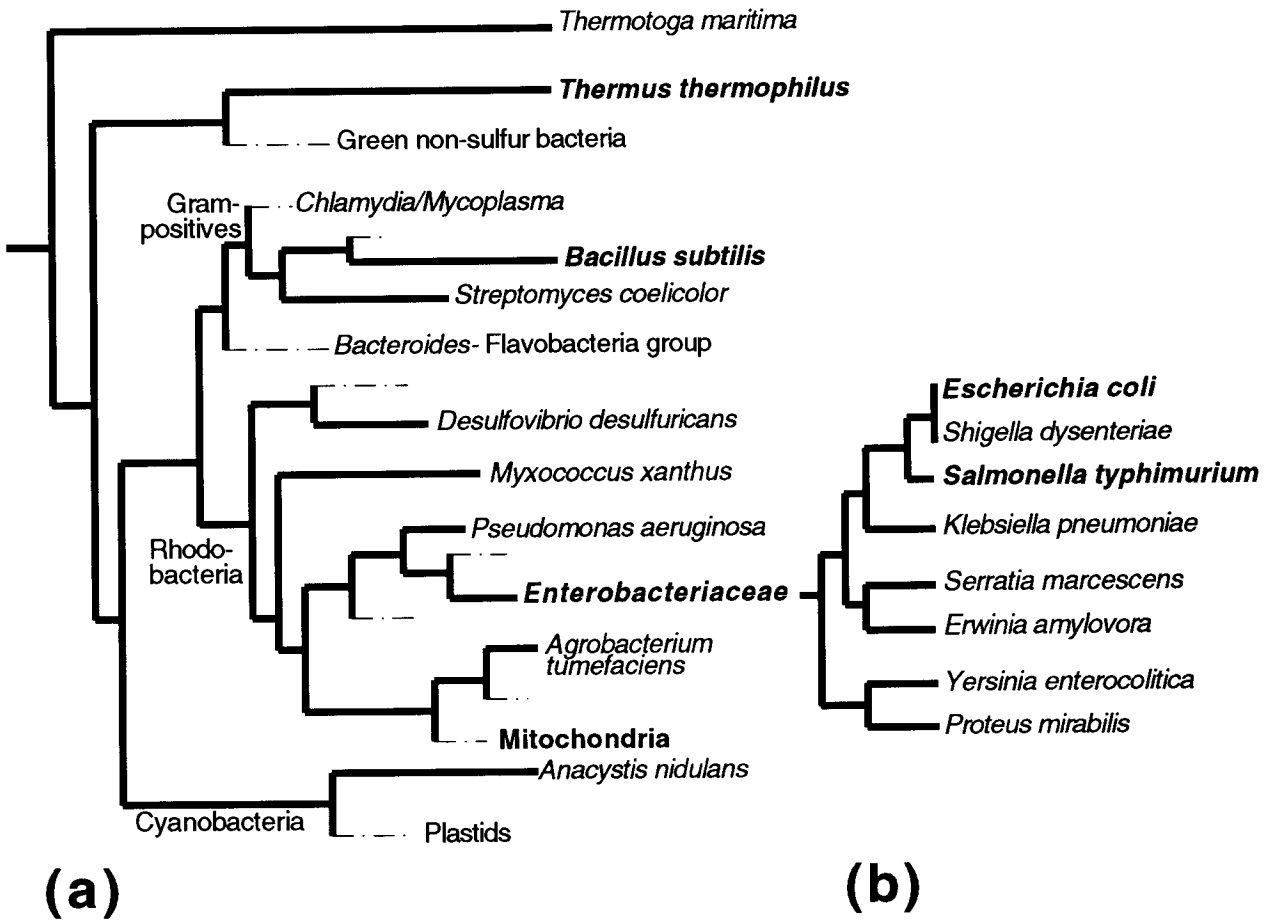


Figure 1. Phylogenetic relationship of prokaryotic organisms used for PheRS comparisons. Only the relevant branches and qualitative phylogenetic distances are shown in this tree. Organisms relevant for the PheRS comparisons treated in this Chapter are typed in bold face letters. **(a)** Simplified phylogenetic tree of eubacteria based on comparisons of complete 16S rRNA sequences. Data were taken from Hartmann *et al.* (1989). **(b)** Phylogenetic relationships among enterobacteria. The tree is based on 16S rRNA oligonucleotide cataloguing data and on features of the aromatic amino acid biosynthetic pathways (Ahmad *et al.*, 1990; simplified version).

1.2. *Thermus thermophilus* PheRS crystals suitable for X-ray analysis are available

The second and principal reason for cloning the PheRS genes from *T. thermophilus* was the availability of PheRS crystals from this organism. These crystals are suitable for X-ray diffraction analysis (Chernaya *et al.*, 1987; Ankilova *et al.*, 1988) at a resolution of at least 3 Å (M. Chernaya, personal communication). The

T. thermophilus enzyme is the only PheRS which has been crystallized up to now, whereas attempts to obtain crystals from the *E. coli* enzyme have failed (Ankilova *et al.*, 1988). It is known that proteins from thermophilic bacteria tend to crystallize more readily than those from mesophiles (Chernaya *et al.*, 1987). In line with this experience, it was recently possible to crystallize the *T. thermophilus* SerRS (Garber *et al.*, 1990*a*) and ThrRS (Garber *et al.*, 1990*b*), aminoacyl-tRNA synthetases of class II (Eriani *et al.*, 1990). The crystals diffract to at least 2.0 Å and 2.4 Å, respectively, and seem appropriate for determination of the tertiary structure by X-ray diffraction analysis. The three-dimensional structure has already been determined for the class I enzyme TyrRS from the thermophilic *Bacillus stearothermophilus* (Brick *et al.*, 1989). The other crystallized aminoacyl-tRNA synthetases with resolved structure include MetRS (class I; Brunie *et al.*, 1990), GlnRS (class I; Rould *et al.*, 1989) and SerRS (class II; Cusack *et al.*, 1990) from *E. coli* (see also Chapter II). The structure of yeast AspRS (class II), whose crystals diffract to 2.7 Å, may soon become published (Ruff *et al.*, 1988).

A clear limitation for the resolution of the tertiary structure of the *T. thermophilus* PheRS is the total lack of knowledge about the primary structure. Therefore, the cloning and sequencing of the *T. thermophilus pheS* and *pheT* genes encoding the α and β subunits of PheRS, respectively, would be a major leap towards the first three-dimensional structure, at high resolution, known for an aminoacyl-tRNA synthetase of the $\alpha_2\beta_2$ type. Finally, co-crystallization of PheRS with the bound substrate phenylalanine may elucidate the definitive location and structure of the Phe binding site.

2. Materials and Methods

Several aspects of this section have been described in more detail by Keller (1990*a*, 1990*b*).

2.1. Bacterial strains and DNA

All *Escherichia coli* strains used here have been described in Chapter II. *Salmonella typhimurium* DB21 is a LT2 (wild-type) derivative which is cured from one of the two cryptic prophages; it was obtained from H. Schmieger (University of Munich, Germany). The type strain *Thermus thermophilus* HB8 (Oshima & Imahori, 1974) was received from L. Reshetnikova (Institute of Molecular Biology, Moscow, USSR). Plasmids pKSB1-W and pHWO have been described (Chapter II). The vectors used for plasmid constructions were pUC19 and pBLS (pBluescript KS(+)), both described in Chapter II, and pUC18 (ampicillin resistant; Norrander *et al.*, 1983; purchased from Pharmacia, Uppsala, Sweden). Plasmids constructed in this work are described in the Results section. Oligonucleotides used only for sequencing were PEKA10 and M13u (listed in Chapter II), and PEKA2 (5'-GCGTAACCACAGTTC-3'; corresponding to the PEKA2S position, Chapter II). The *T. thermophilus*-specific oligonucleotides (TTH1-TTH5; synthesized by A. Savioz with an Applied Biosystems DNA synthesizer model 380B) are presented in detail later (Fig. 5). λ-DNA size markers were from Biofinex (Praroman, Switzerland).

2.2. Media and bacterial growth conditions

The growth conditions for *E. coli* strains were as described in Chapter II. *S. typhimurium* was grown in LB medium at 37°C under identical conditions as *E. coli*. Single colonies of *T. thermophilus* were obtained on 'Th-0 agar plates' containing 0.8 % Bacto-Peptone, 0.4 % yeast extract (both from Oxoid; Basingstoke, UK), 0.3 % NaCl (medium adjusted to pH 7.5 (at 20°C) with 1 N NaOH), and finally 2 % agar (Oshima & Imahori, 1974). Incubation was done at 60°C in a humid chamber for about 40 h. For liquid cultivation, 'Th-A medium' (0.5 % Bacto-Peptone, 0.2 % yeast extract, 0.2 % NaCl; pH 7.0 - 7.2) was inoculated with a single colony, and

overnight incubation was done at 75°C under vigorous shaking.

2.3. Isolation of chromosomal DNA

Chromosomal DNA from *E. coli* strains was isolated from a 100 ml culture grown to stationary phase. The cells were sedimented by centrifugation (4'000 x g, 5 min, 4°C) and resuspended in 10 ml of 50 mM Tris-HCl, pH 8.0, 25 mM EDTA (0°C). Cells were lysed by addition of lysozyme (Sigma Chemicals, St Louis MO, USA) to 1 mg/ml final concentration. After 10 min at room temperature, RNAse A (Sigma Chemicals; preincubated for 10 min at 100°C at a concentration of 10 mg/ml in H_2O) was added to 100 µg/ml final concentration and incubation was continued for 30 min at 37°C. Then, proteinase K (Boehringer, Mannheim, Germany; 20 mg/ml in H_2O, freshly preincubated for 30 min at 37°C) and SDS were added to 1 mg/ml and 0.3 % final concentrations, respectively, and the lysate was further incubated for 30 min at 37°C. After addition of NaCl to 100 mM, two phenol/chloroform (equilibrated in 50 mM Tris-HCl, pH 8.0) extractions and subsequent phase separations by centrifugation (8'000 x g, 5 min, 4°C) were carried out. Then the aqueous phase was extracted twice with chloroform. The DNA was precipitated with 2.2 volumes of 100 % ethanol (10 min 0°C; centrifugation at 8'000 x g, 20 min, 4°C), washed with 70 % ethanol, dried and finally dissolved carefully in 1 ml of 10 mM Tris-HCl, pH 8.0.

For *S. typhimurium* and *T. thermophilus* DNA isolations, the conditions for the RNAse A incubation were modified to 15 min at 30°C.

2.4. Hybridizations with oligonucleotides and heterologous DNA

For Southern blot analyses (Southern, 1975), 2 µg of restriction enzyme-digested chromosomal DNA was transferred onto Hybond-N membranes (Amersham, UK) as described by Sambrook *et al.* (1989) and crosslinked by UV light for 3 min. The prehybridization solution corresponded to the prehybridization fluid described by Maniatis *et al.* (1982), except that salmon sperm DNA was replaced by low fat milk powder (0.4 % final concentration) and sodium phosphate (pH of a 1 M stock solution adjusted to 6.5) was present at 20 mM. The hybridization step was carried out with a fresh aliquot of this solution that also contained the radiolabelled DNA probe. Conditions and manipulations for prehybridization and hybridization were as described previously (Maniatis *et al.*, 1982), except that the

temperature was varied (see below). Oligonucleotides were labelled at their 5'-ends by a polynucleotide kinase reaction (Sambrook *et al.*, 1989). The method of 'random priming' by Feinberg & Vogelstein (1984) was used to label double stranded DNA restriction fragments. Radiolabelled mononucleotides were purchased from Amersham (UK).

All washing steps were carried out in 0.1 % SDS, 6 x SSC (1 x SSC is 0.15 M NaCl, 15 mM Na_3Citrate, pH 7.0). For the primary washing steps, the hybridized membranes were immersed once in 100 ml (1 min) and one to three times in 500 ml (for 1 h each time), partially dried and exposed to an X-ray film (Diagnostic Film XAR2; Kodak, Rochester NY, USA) with an amplifying screen. The temperatures were identical in the prehybridization, hybridization and primary washing steps. The optimal temperature was dependent on the G+C-content and the degree of homology of the probe. For oligonucleotides, the first approximation of the hybridization temperature (T_h) was based on melting point (T_m) calculations:

$T_h \cong T_m - 12°C$; $T_m \cong 0°C + (A/T) \times 2°C + (G/C) \times 4°C$ (Sambrook *et al.*, 1989)

The actual temperatures used are indicated in the concrete examples shown in the Results section. If the hybridization appeared too unspecific after the first exposure, more washing steps at higher temperatures were performed (see Results).

To re-hybridize the same membrane with another DNA probe, the previous probe was washed off by incubation at 45°C for 30 min in 500 ml each of 0.4 M NaOH and then 6 x SSC, 0.1 % SDS, 0.2 M Tris-HCl, pH 7.5.

In order to detect *E. coli* transformants with *T. thermophilus pheS* clones, a colony hybridization was performed with the oligonucleotide TTH1. The transfer of DNA to Hybond-N membranes was done as described in the protocol 'Membrane Transfer and Detection Methods' from Amersham (1986). The hybridization was carried out at 31°C, followed by two washing steps at 34° and 37°C. Radiolabelling of TTH1 and other hybridization and washing conditions were as specified above for Southern blottings.

2.5. Recombinant DNA techniques and DNA sequence analysis

Plasmid isolation, gel electrophoresis, enzymatic DNA modifications, cloning, transformation, purification of oligonucleotides and single and double stranded DNA sequencing were carried out essentially as described or referenced in Chapter II. Computer-assisted sequence analyses (restriction maps, codon usage

calculations, sequence searches and alignments) were done using programs of the University of Wisconsin Genetics Computer Group (UWGCG; sequence analysis software package release 6.2; Madison WI, USA).

3. Results

3.1. Detection of *pheS*- and *pheT*-like genes by Southern blot hybridization

(a) Heterologous hybridizations with E. coli pheS *and* pheT *fragments*

In order to find restriction fragments suitable for cloning of new PheRS genes, Southern blot analyses were carried out with chromosomal DNA from *Salmonella typhimurium* DB21, *Thermus thermophilus* HB8 and, as a positive control, *Escherichia coli* NP37. The choice of restriction enzymes was influenced by the expected average size of the chromosomal fragments and by their usefulness for the intended cloning in the vector pUC19: fragments which possibly contained both PheRS genes were expected to be in the size range of at least 3.5 kb as judged, for instance, from the known *pheST* genes of *E. coli* (see Fig. 2).

Suitable restriction fragments from the *E. coli pheST* operon were taken as hybridization probes (see Fig. 2). The *'pheS'* probe (a 506 bp *Bst*EII/*Bst*BI frag-

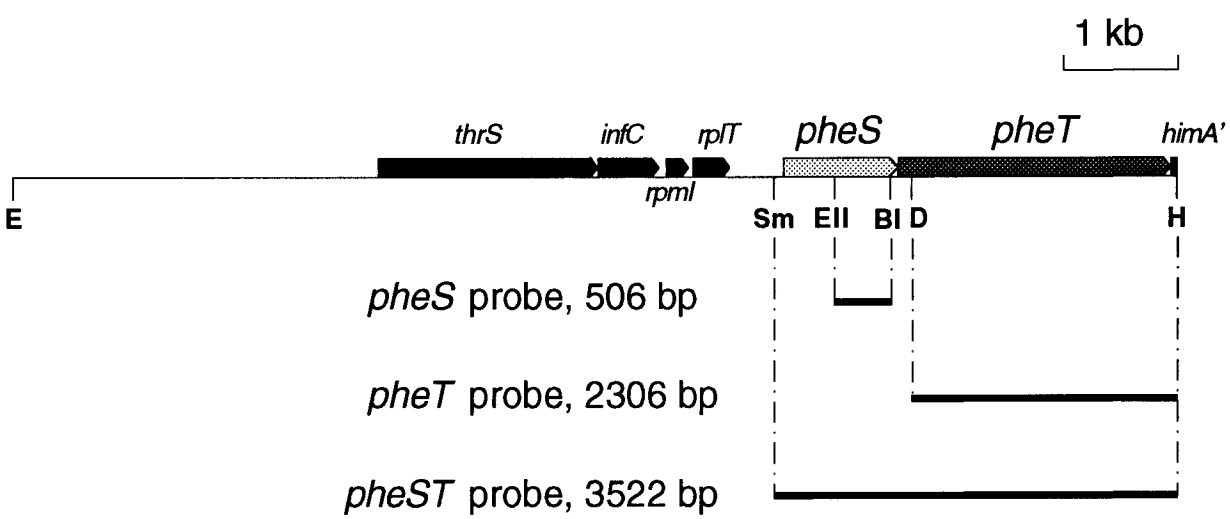


Figure 2. Probes used for heterologous hybridization. The genes on the 10.25 kb *Eco*RI/*Hin*dIII fragment from *E. coli* have been described in Fig. 1 of Chapter II. The fragments used as heterologous hybridization probes are shown as horizontal bars below the genes. Relevant restriction sites are abbreviated as follows: E, *Eco*RI; Sm, *Sma*I; EII; *Bst*EII; BI, *Bst*BI; D, *Dra*III; H, *Hin*dIII.

ment from plasmid pKSB1-W) approximately consisted of the 3'-half of *pheS*. This region was found to be quite well conserved among the *pheS* genes encoding *E. coli* (Fayat *et al.*, 1983), yeast cytoplasmic (Sanni *et al.*, 1988) and yeast mitochondrial (Koerner *et al.*, 1987) PheRS α (small) subunits (DNA sequence alignment not shown). The *'pheST'* and *'pheT'* probes were the 3522 bp *Sma*I/*Hin*dIII and 2306 bp *Dra*III/*Hin*dIII fragments, respectively, from plasmid pHWO (Fig. 2).

(i) Hybridizations with the *E. coli pheS* probe

The results of a Southern blot hybridization with the *pheS* probe are shown in Fig. 3(a,b). Hybridization to restriction enzyme-digested chromosomal DNA from *T. thermophilus* (Fig. 3(a)) gave signals only if it was carried out under low stringency conditions. Knowing the difference in G+C-content between *E. coli* (48-52 %; Ørskov, 1984) and *T. thermophilus* (69 %, Oshima & Imahori, 1974), it was not surprising to see that the hybridization signals were not very specific. For example, there was considerable hybridization of the λ-DNA size standard (Fig. 3(a)). Nevertheless, various hybridizing fragments appeared in the *T. thermophilus* DNA digests. As will be shown later (section 3.1.b.) the fragments containing the *T. thermophilus pheS* gene were actually among them.

Under more stringent conditions, specific hybridization of the heterologous *pheS* probe was observed with *S. typhimurium* DNA (Fig. 3(b)). The signals were nearly as strong as those obtained with the homologous *E. coli* controls (lanes 1 and 8). The sizes of the hybridizing fragments (one fragment per digestion) varied from 3.9 to about 20 kb (lanes 3 to 7).

(ii) Hybridizations with the *E. coli pheST* and *pheT* probes

To identify fragments that contained also the *pheT* gene in addition to *pheS*, the same membranes as used above were freed from the *pheS* label and hybridized a second time to the *pheT* and *pheST* probes described in Fig. 2. *T. thermophilus* DNA did not show clear hybridization signals, whereas very clear bands were obtained again with digested chromosomal DNA from *S. typhimurium* (data shown by Keller, 1990*a*); it turned out that the digestions with *Sal*I and *Sac*I yielded fragments (10.5 kb and 20 kb, respectively) that contained both the *pheS* and *pheT* genes. This was deduced from the fact that the same, single fragment in each of these digestions hybridized with the *pheS*, *pheT* and the *pheST* probes.

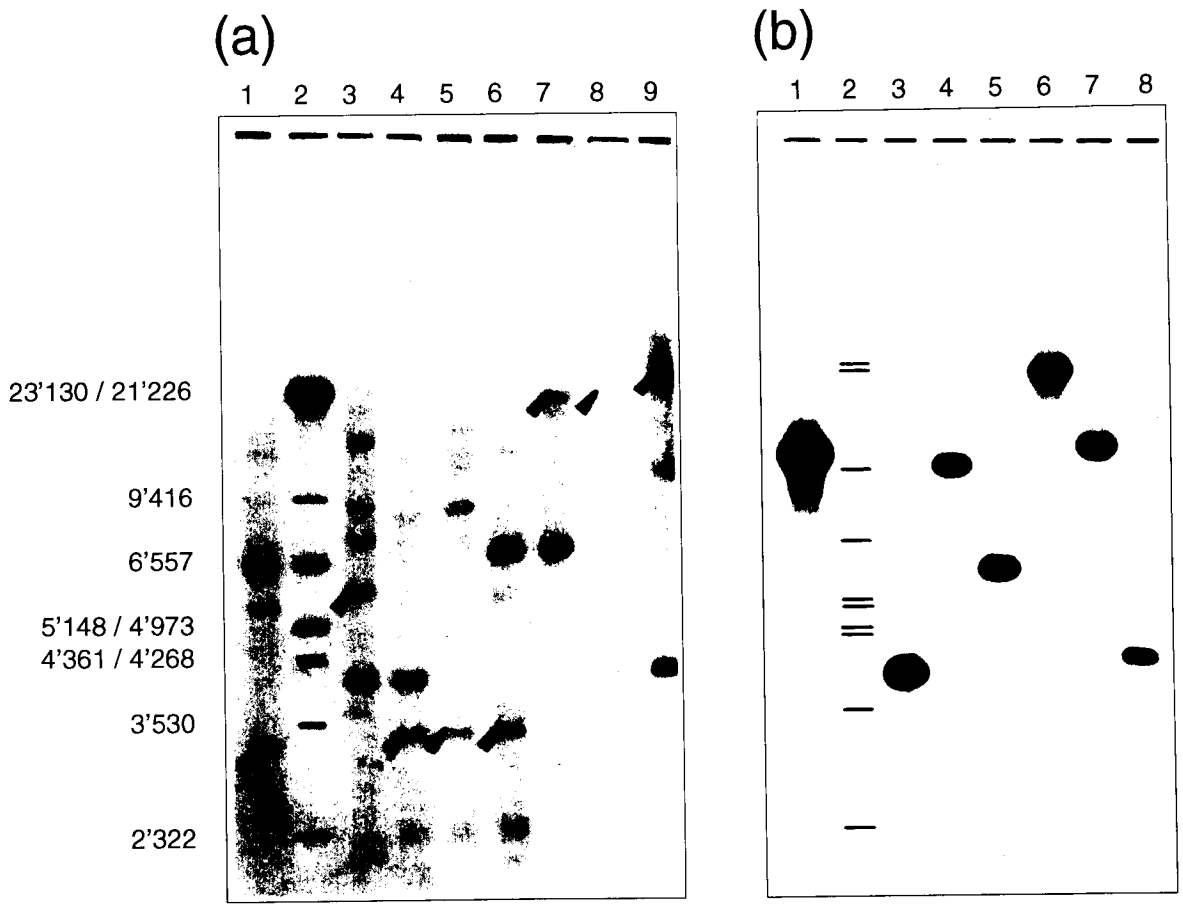


Figure 3. Heterologous hybridizations with the *pheS* probe. The autoradiograms of the blot hybridization with the 506 bp *pheS* probe (carried out at 45°C) are shown. In lane 2 each of (a) and (b), a size standard was loaded (λ-DNA digested with *Eco*RI/*Hin*dIII and with *Hin*dIII; bands hand-drawn in lane 2); the fragment sizes (in bp) are given on the left margin. **(a)** Hybridization of *E. coli pheS* to *T. thermophilus* DNA. The membrane was washed at 51°C. The chromosomal DNA was digested with *Sph*I/*Hin*dIII (lane 1), *Sph*I/*Xba*I (lane 3), *Xba*I/*Kpn*I (lane 4), *Kpn*I (lane 5), *Kpn*I/*Hin*dIII (lane 6), *Hin*dIII (lane 7), *Xba*I/*Hin*dIII (lane 8) and *Xba*I (lane 9). Bands labelled with an arrowhead are mentioned in section 3.1.b. of the text: they will appear again in Fig. 6. **(b)** Hybridization of *E. coli pheS* to *S. typhimurium* DNA. Washing of the membrane was done at 58°C. Lane 1 *(E. coli* NP37 chromosomal DNA cut with *Eco*RI and *Hin*dIII) and lane 8 (linearized plasmid pKSB1-W, diluted) were used for homologous control hybridizations. Chromosomal *S. typhimurium* DNA was digested with *Eco*RI/*Kpn*I (lane 3), *Kpn*I (lane 4), *Sph*I (lane 5), *Sac*I (lane 6) and *Sal*I (lane 7).

(b) Hybridizations with oligonucleotides

(i) Design of oligonucleotides to identify *T. thermophilus pheS*

In order to detect the *T. thermophilus pheS* gene more clearly, five specific oligonucleotides were synthesized. The oligonucleotides matched the following criteria:

(1) They were designed to hybridize to strongly conserved regions within *pheS* genes from different origins. The choice of these regions was based on DNA sequence alignments (not shown) as well as on amino acid sequence alignments (Fig. 4) of the three *pheS* genes or gene products, respectively, which were known at the time when the experiment was initiated.

(2) The oligonucleotides were adapted to the extremely biased codon usage of *T. thermophilus* (Table 1) resulting from the high G+C-content (overall 69 %, and third base up to 96 %; Oshima & Imahori (1974) and Nureki *et al.* (1991), respectively). A codon usage table was established by using the UWGCG program CODONFREQUENCY and all *T. thermophilus* HB8 genes known at that time (Table 1). This table was also in agreement with the codon usage in genes of other *Thermus* species, like *mdh* (*T. flavus*; Nishiyama *et al.*, 1986), the lactate dehydrogenase gene (*T. caldophilus*; Kunai *et al.*, 1986) and *sucD* (*T. aquaticus*; Nicholls *et al.*, 1988).

(3) The recognition site TCGA for the restriction endonucleases produced by the *Thermus* species *T. thermophilus* HB8 (*Tth*HB8I) and *T. aquaticus* YT-1 (*Taq*I) is almost absent in *Thermus* DNA (Kunai *et al.*, 1986; Kushiro *et al.*, 1987). The oligonucleotides were therefore designed such as to avoid this sequence.

Figure 4 shows the four conserved regions in the PheRS α subunits selected for construction of five corresponding oligonucleotide probes TTH1 to TTH5 (Fig. 5).

Figure 4. Alignment of PheRS α subunits used to design the oligonucleotide hybridization probes. The sequences for the PheRS α subunits from yeast mitochondria (Koerner *et al.*, 1987), *E. coli* (Fayat *et al.*, 1983; as corrected in Chapter II) and yeast cytoplasm (Sanni *et al.*, 1988) were aligned using UWGCG sequence alignment programs. Identical amino acids (in one-letter-code) are boxed; hatched bars below the sequences indicate the conserved regions chosen for the construction of oligonucleotides TTH1 to TTH5.

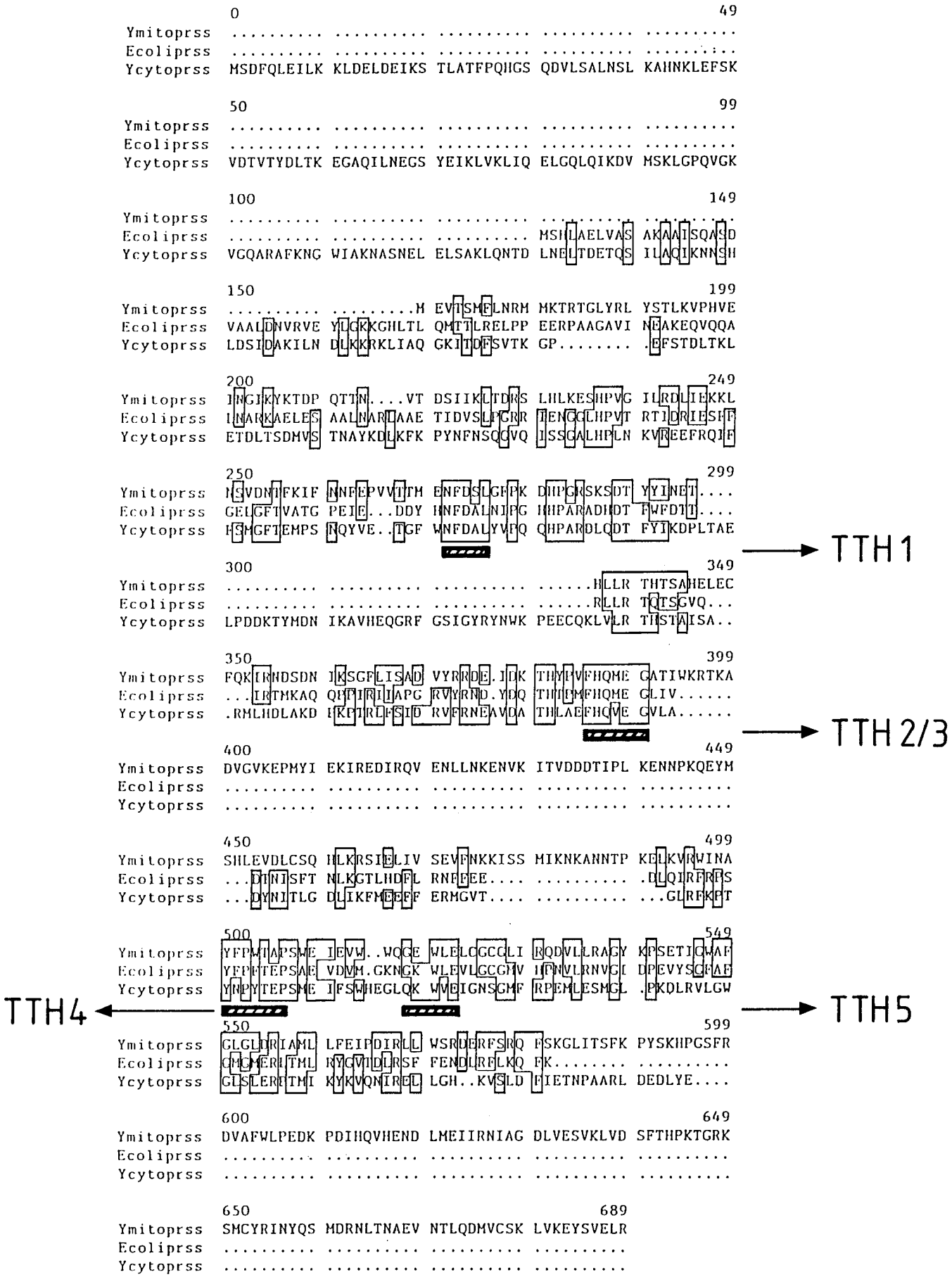

TTH1
TTH2/3
TTH4
TTH5

Table 1: *Codon usage table of* Thermus thermophilus *genes*[a]

Aa	Codon	Number[b]	Fraction[c]	Aa	Codon	Number[b]	Fraction[c]
Gly	GGG	105.00	0.59	Trp	TGG	10.00	1.00
Gly	GGA	13.00	0.07	End	TGA	3.00	0.60
Gly	GGT	10.00	0.06	Cys	TGT	0.00	0.00
Gly	GGC	49.00	0.28	Cys	TGC	4.00	1.00
Glu	GAG	146.00	0.91	End	TAG	0.00	0.00
Glu	GAA	15.00	0.09	End	TAA	2.00	0.40
Asp	GAT	4.00	0.04	Tyr	TAT	4.00	0.08
Asp	GAC	87.00	0.96	Tyr	TAC	46.00	0.92
Val	GTG	129.00	0.68	Leu	TTG	28.00	0.15
Val	GTA	1.00	0.01	Leu	TTA	1.00	0.01
Val	GTT	5.00	0.03	Phe	TTT	20.00	0.30
Val	GTC	54.00	0.29	Phe	TTC	47.00	0.70
Ala	GCG	69.00	0.39	Ser	TCG	11.00	0.17
Ala	GCA	1.00	0.01	Ser	TCA	0.00	0.00
Ala	GCT	9.00	0.05	Ser	TCT	2.00	0.03
Ala	GCC	100.00	0.56	Ser	TCC	26.00	0.39
Arg	AGG	36.00	0.24	Arg	CGG	63.00	0.43
Arg	AGA	1.00	0.01	Arg	CGA	4.00	0.03
Ser	AGT	1.00	0.02	Arg	CGT	7.00	0.05
Ser	AGC	26.00	0.39	Arg	CGC	37.00	0.25
Lys	AAG	67.00	0.94	Gln	CAG	20.00	0.87
Lys	AAA	4.00	0.06	Gln	CAA	3.00	0.13
Asn	AAT	2.00	0.06	His	CAT	0.00	0.00
Asn	AAC	31.00	0.94	His	CAC	42.00	1.00
Met	ATG	37.00	1.00	Leu	CTG	63.00	0.33
Ile	ATA	3.00	0.05	Leu	CTA	3.00	0.02
Ile	ATT	16.00	0.29	Leu	CTT	22.00	0.11
Ile	ATC	37.00	0.66	Leu	CTC	75.00	0.39
Thr	ACG	43.00	0.56	Pro	CCG	41.00	0.32
Thr	ACA	0.00	0.00	Pro	CCA	4.00	0.03
Thr	ACT	0.00	0.00	Pro	CCT	14.00	0.11
Thr	ACC	34.00	0.44	Pro	CCC	69.00	0.54

[a] This codon usage table is based on 5 sequenced *T. thermophilus* HB8 genes, *i. e. leuB* (Kagawa *et al.*, 1984), *tuf1* (Seidler *et al.*, 1987), *trpE* and *trpG* (Sato *et al.*, 1988) and the gene for phosphoglycerate kinase (PGK; Bowen, *et al.*, 1988). Aa, amino acid.

[b] Total number of codons of the respective type scored in the included genes.

[c] Fraction of the abundance of an individual codon compared to that of the sum of all possible codons for the corresponding amino acid.

TTH1

For peptide [S] **N F D A L**

Probable codons: (2) (2) (2) (4) (6) in general

(1) (2) (1) (2) (4) for *T. thermophilus*

E. coli sequence: **AAC TTC GAT GCT CTG**

Thermus variants: **AAC TT[T/C] GAC GC[G/C] T[T/C][G/C]**

TaqI site

TTH1 15-mer **AAC TTT GAC IC[G/C] CTG** (2 variants, 50 % GC)

TTH2/3

For peptide [V] **F H Q M E G**

Probable codons: (2) (2) (2) (1) (2) (4) in general

(2) (1) (1) (1) (1) (2) for *T. thermophilus*

E. coli sequence: **TTC CAT CAG ATG GAA GGT**

Thermus variants: **TT[T/C] CAC CAG ATG GAG GG[G/C]**

TTH2 17-mer **TT[T/C] CAC CAG ATG GAG GG** (2 variants, 53-59% GC)

TTH3 18-mer **[C/G]CC CTC [C/G]AC CTG GTG GAA** (4 variants, 67 % GC)

(=complement of: TTC CAC CAG GT[G/C] GAG GG[G/C])

TTH4

For peptide [N] [Y/W] [A] **Y F P F T E P**

Probable codons: (2)(2)(4) (5) (4) (2)(4) in general

(1)(2)(2) (4) (2) (1)(2) for *T. thermophilus*

E. coli sequence: **TAC TTC CCG TTT ACC GAA CCT**

Thermus variants: **TAC TT[T/C] CC[G/C] T[GG/TT/AC] AC[G/C] GAG CC[G/C]**

TTH4 20-mer **TAC TTC CC[G/C] TII AC[G/C] GAG CC** (4 variants, 61 % GC)

TTH5

For peptide [Q][E] [V] **G K W L E**

Probable codons: (4) (2) (1) (6) (2) in general

(2) (1) (1) (4) (1) for *T. thermophilus*

E. coli sequence: **GGT AAA TGG CTG GAA**

Thermus variants: **GG[G/C] AAG TGG [C/T][G/C] GAG**

TTH5 15-mer **CTC [G/C]AG CCA CT[T/C] [G/C]CC** (8 variants, 67-73 % GC)

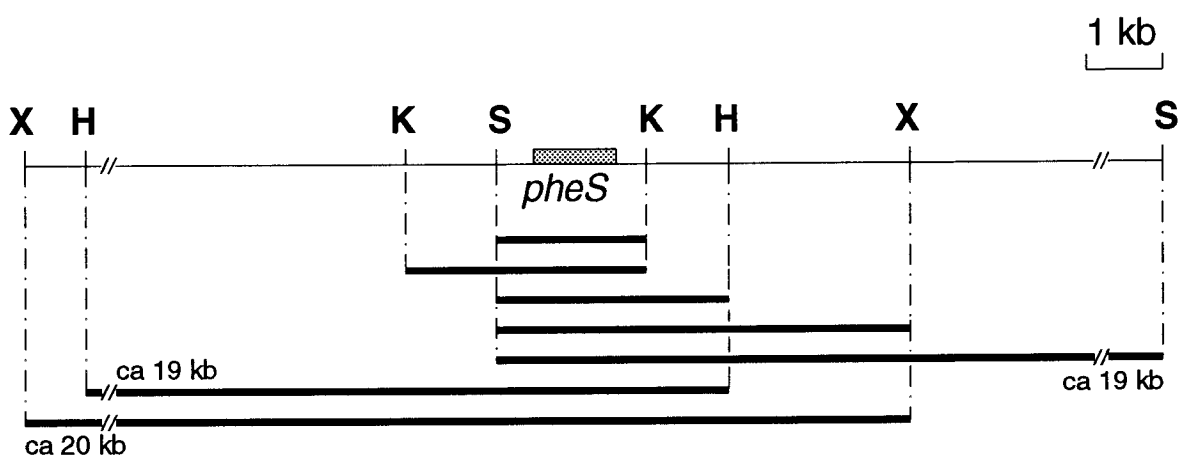
(=complement of: GG[G/C] [A/G]AG TGG CT[G/C] GAG)

Figure 5. Oligonucleotides constructed as hybridization probes to detect *T. thermophilus pheS* DNA. The amino acids of the *E. coli pheS* gene product that are covered by the oligonucleotides are indicated in one-letter-code. Alternative amino acids encountered in the α subunits from yeast cytoplasmic and mitochondrial PheRS (and corresponding nucleotides) are indicated in brackets. The 'probable' codons for *T. thermophilus* were defined as those which make up more than 85 % of the possible codons ('general') for a given amino acid according to Table 1. Nucleotides in the *E. coli pheS* sequence conserved in all three examined *pheS* genes are underlined. Deoxyinosine (I) with the capacity to pair with the bases in the order C > A >> G > T (Kawase *et al.*, 1986) was used at suitable ambiguity positions. TTH2 and TTH3 are alternative variants to the same sequence. All nucleotide sequences are written in the 5'-to-3' direction.

(ii) Hybridizations with oligonucleotides TTH1 to TTH5

The blots probed previously with the *E. coli pheS* and *pheT* fragments (see section 3.1.a.) were now sequentially hybridized with oligonucleotides TTH1 (at 30°C), TTH2 (at 36°C), TTH3 (at 48°C), TTH4 (at 41°C) and TTH5 (at 34°C). The probes did not hybridize to the *E. coli* and presumptive *S. typhimurium pheS* bands, whereas various signals appeared with *T. thermophilus* DNA. The strongest bands were obtained with TTH1 as probe (Fig. 6). The TTH1 signals marked by an arrowhead in Fig. 6 were also visible when the other oligonucleotides were used as hybridization probes (except with TTH3 which is an alternative to TTH2) (data not shown). Furthermore, these bands had been present even in the heterologous hybridization with the *E. coli pheS* probe (marked in Fig. 3(a)). There were no other common signals among the hybridizations with the different oligonucleotides and/or between hybridizations with the oligonucleotides and the *pheS* probe.

The results of the Southern blot analyses allowed us to establish a preliminary partial restriction map of the *pheS* region of *T. thermophilus* as shown in Figure 7.



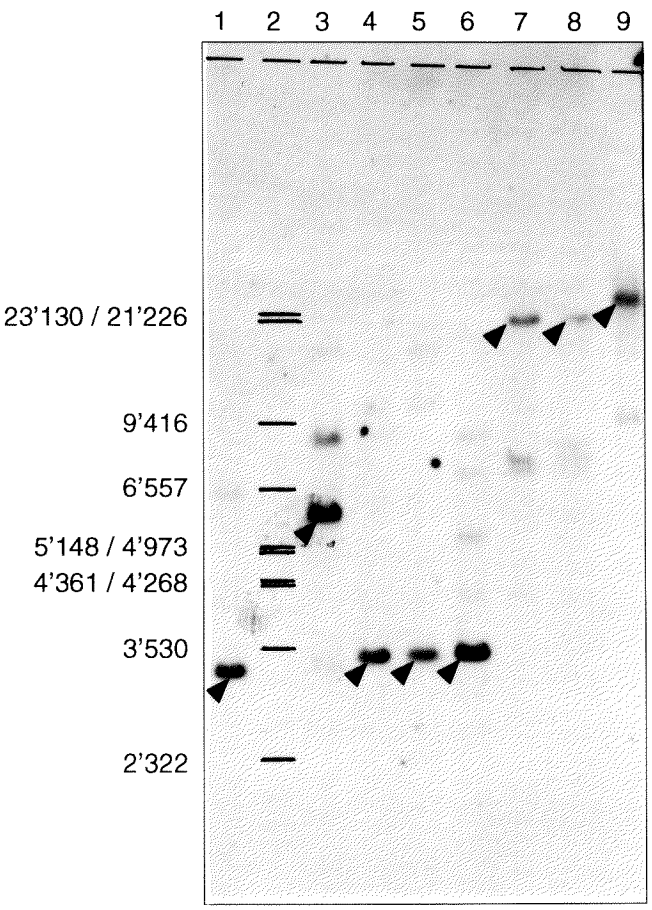


Figure 6. Hybridization of *T. thermophilus* DNA with oligonucleotide TTH1 as probe. The autoradiogram shows the blot hybridized with radiolabelled TTH1 at 30°C and subsequent washing at 34°C. Lane 2 contains the bands from the size strandard (as in Fig. 3), with the fragment sizes (in bp) indicated on the left margin. Chromosomal DNA from *T. thermophilus* was digested with *Sph*I/*Hin*dIII (lane 1), *Sph*I/*Xba*I (lane 3), *Xba*I/*Kpn*I (lane 4), *Kpn*I (lane 5), *Kpn*I/*Hin*dIII (lane 6), *Hin*dIII (lane 7), *Xba*I/*Hin*dIII (lane 8) and *Xba*I (lane 9). Arrowheads point to hybridizing bands (the strongest signals on this blot) which were already visible in Fig. 3(a); these bands are described in more detail in the text.

Figure 7. Preliminary restriction map of the *pheS* region in *T. thermophilus* chromosomal DNA as deduced from Southern blot analyses. The hybridizing fragments used to establish this partial map are indicated by bold lines. The relevant restriction sites are X, *Xba*I; H, *Hin*dIII, K, *Kpn*I and S, *Sph*I.

3.2. Cloning of *pheS* and *pheT* from *Salmonella typhimurium*

(a) Cloning strategy

The strategy was analogous to the one described in Chapter II for cloning the *pheS* genes from *E. coli* mutant strains. The *S. typhimurium* 10.5 kb *Sal*I fragment, identified in hybridization analyses (section 3.1.a.) to contain both the *pheS* and *pheT* genes, was chosen for cloning. *Sal*I-digested *S. typhimurium* DNA from the 10-11 kb size range was cloned into *Sal*I-linearized pUC19 and introduced into *E. coli* strain JM109 ($HsdR^-$, $HsdM^+$) to obtain *Eco*K-modified plasmid DNA. The plasmids from the pool of transformed JM109 cells were isolated and then introduced into the *E. coli* strain KA2, whose *pheS* gene encodes a thermosensitive PheRS α subunit (see Chapter II). Transformed KA2 cells able to grow at 40°C (the non-permissive temperature for KA2) were analyzed and shown to contain the pUC19-based plasmid with a 10.5 kb *Sal*I fragment (named 'pKSC-S1').

(b) Identification of the S. typhimurium pheS *gene and evidence for the presence of* pheT *and* thrS

The presence of *pheS* on the cloned 10.5 kb *S. typhimurium Sal*I DNA fragment of pKSC-S1 was verified by three independent means:

(1) Transformation of the *E. coli* strain KA2 ($pheS^{ts}$) with plasmid pKSC-S1 always resulted in a thermoresistant phenotype, indicating that the plasmid carried a functional *pheS* gene.

(2) The restriction map of the 10.5 kb *Sal*I fragment (Fig. 8) was consistent with the bands seen in the previous Southern blots (section 3.1.a.; *e. g.* the presence of *Kpn*I and *Sph*I sites in the middle of the *pheST* region).

(3) Small sections of plasmid pKSC-S1 were sequenced by applying the oligonucleotide primers PEKA2 and PEKA10 that were used previously for the *E. coli pheS* gene (Chapter II). These two primers were chosen because they allowed sequencing of the region encoding motif 3 (*i. e.* part of the presumed Phe binding site) on both DNA strands. The sequence obtained (Fig. 9) showed an extremely high similarity to the *E. coli pheS* gene (90.2 % and 98.0 % identity on the nucleotide and amino acid sequence levels, respectively).

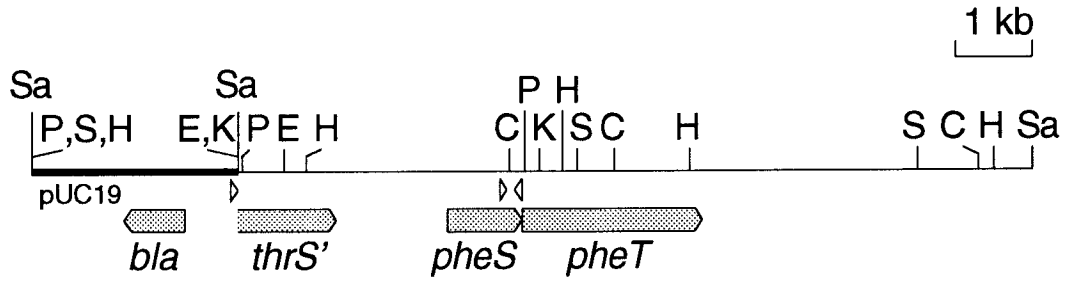


Figure 8. Restriction map of the *pheS* gene region from *S. typhimurium*. Plasmid pKSC-S1 was mapped with *Sal*I (Sa), *Pst*I (P), *Eco*RI (E), *Hin*dIII (H), *Cla*I (C), *Kpn*I (K) and *Sph*I (S). The thick line indicates the vector part (pUC19). Open arrowheads show the position and orientation of the sequencing primers (from left to right): M13u, PEKA10 and PEKA2. *bla* stands for the gene encoding β-lactamase; the other genes are described in the text.

```
                10                  30                  50
                .                   .                   .
S.t.  L E V L G C G M V H P N V L R N V
S.t.  CTGGAAGTGCTGGGCTGCGGGATGGTGCATCCGAATGTGCTGCGTAATGTC
      ||||||||||||||||||||||||||||||||||||| ||| ||||||| ||
E.c.  CTGGAAGTGCTGGGCTGCGGGATGGTGCATCCGAACGTGTTGCGTAACGTT 849
E.c.  L E V L G C G M V H P N V L R N V                    283

                    70                  90
                    .                   .                   .
S.t.  G I D P E I Y S G F A F G M G M E
S.t.  GGTATCGATCCGGAAATCTATTCTGGCTTCGCTTTTGGTATGGGGATGGAA
      || ||||| |||||| | || ||||| ||||| || || ||||||||||||
E.c.  GGCATCGACCCGGAAGTTTACTCTGGTTTCGCCTTCGGGATGGGGATGGAG 900
E.c.  G I D P E V Y S G F A F G M G M E                    300

            110                 130                 150
            .                   .                   .
S.t.  R L T M L R Y G V T D L R S F F E
S.t.  CGTCTGACCATGTTGCGTTACGGCGTCACCGACCTGCGTTCATTCTTCGAA
      |||||||| ||||||||||||||||||||||||||||||||||||||||||
E.c.  CGTCTGACTATGTTGCGTTACGGCGTCACCGACCTGCGTTCATTCTTCGAA 951
E.c.  R L T M L R Y G V T D L R S F F E                    317
```

Figure 9. Comparison of partial *pheS* DNA and derived amino acid sequences between *S. typhimurium* and *E. coli*. The nucleotide sequence of *pheS* from *S. typhimurium* (S.t.) obtained with primers PEKA2 and PEKA10 is shown. The corresponding *E. coli* (E.c.) *pheS* gene region (Fayat *et al.*, 1983) is aligned. Identical nucleotides are connected by vertical lines. The only deviating amino acid between the two derived protein sequences (above and below the corresponding nucleotide sequence) is highlighted in bold italic. Numbers to the right of the *E. coli* sequences indicate sequence positions according to Fayat *et al.* (1983).

These experiments proved that the cloning of the *pheS* gene from *S. typhimurium* was successful. Also, it was possible to determine the position and orientation of *pheS* on the plasmid pKSC-S1 restriction map (Fig. 8). Moreover, hybridization data suggested that the entire *pheT* gene was present immediately downstream of *pheS* (Fig. 8).

Double stranded sequencing with the univeral primer M13u across the pUC19 polylinker into the *Sal*I insert (Fig. 8) yielded a sequence (Keller, 1990*b*) which showed strong similarity to the *E. coli thrS* gene for threonyl-tRNA synthetase (identity 87.7 %). The amino acid sequence derived from the determined DNA sequence was identical to residues 187 to 261 of the *E. coli* threonyl-tRNA synthetase (Mayaux *et al.*, 1983), except that a His residue was present instead of Arg^{195}. (The codon for an amino acid that corresponds to *E. coli* ThrRS residue 207 could not be determined unambiguously.) The 5'-end of *thrS* was not contained on plasmid pKSC-S1. The 3'-end of *thrS* was found at about the same distance from *pheS* as in *E. coli* (Fig. 8 and Fig. 2).

3.3. Cloning of *pheS*- and *pheT*-like regions from *Thermus thermophilus*

(a) Cloning strategy

For the cloning of the *T. thermophilus pheS* gene, two of the genomic restriction fragments identified by Southern blot hybridization to carry *pheS*-like sequences (section 3.1.b.) were chosen. The fragments of around 3.2 kb of *Kpn*I-digested *T. thermophilus* DNA and of about 5.5 kb in *Sph*I/*Xba*I double digestions (see Figs. 6 & 7) were excised from a low-melting agarose gel. The isolated DNA was ligated with analogously digested pUC18 and pUC19 vectors and then used to transform *E. coli* strain JM109.

It was first attempted to clone the *T. thermophilus pheS* gene by complementation of the thermosensitive *E. coli* strain KA2, as described for the cloning of *pheS* from *S. typhimurium* (section 3.2.a.). Transformation of the plasmid pool isolated from the JM109 cells, however, did not yield KA2 colonies able to grow at high temperature (40°C).

The following approach turned out to be successful. The transformed JM109 cells were screened for the presence of cloned *T. thermophilus pheS* sequences by colony hybridization. The radiolabelled oligonucleotide TTH1 was used as the

hybridization probe because it showed the strongest hybridization to the presumptive *T. thermophilus pheS* bands (section 3.1.b.; Fig. 6). TTH1 did not hybridize to *E. coli pheS* DNA (see above) and the background hybridization to untransformed JM109 colonies (negative control) was indeed clearly lower than to *T. thermophilus* colonies (positive control; data not shown). Moreover, it was expected that a plasmidial *pheS* gene should yield much stronger hybridization signals due to the multicopy effect. In fact, strongly hybridizing colonies were found among JM109 cells transformed with both *Kpn*I and *Sph*I/*Xba*I fragment-carrying plasmids.

(b) Identification of the T. thermophilus pheS-*like gene*

The plasmid DNA of some of the positive clones was isolated and examined by restriction analysis. One clone of each type containing either a 3.2 kb *Kpn*I fragment in pUC19 (pKST-K19U) or a 5.5 kb *Sph*I/*Xba*I insertion in pUC18 (pKST-SX18U) or pUC19 (pKST-SX19U) was selected for further restriction mappings (Fig. 10). All three *T. thermophilus* DNA inserts contained internal restriction sites that were expected from the preliminary map based on the Southern blot results (Fig. 7); *i. e.* an *Sph*I site in pKST-K19U and a *Kpn*I and a *Hin*dIII site each in pKST-SX18U and pKST-SX19U. Using other restriction endonucleases, it became clear that the cloned fragments overlapped (Fig. 10), as predicted from Fig. 7. The results from the restriction mapping confirmed that the cloned fragments corresponded to those that appeared in the previous Southern hybridizations.

The plasmids pKST-K19U, pKST-SX18U and pKST-SX19U were partially sequenced (in double stranded form) using the oligonucleotide TTH1 as primer (Fig. 10). In all cases, the same DNA sequence was obtained (shown by Keller, 1990*b*). The total G+C-content of the readable 201 nucleotide stretch was 69 %, and even 90 % for the third position in codons, as is typical for *T. thermophilus* genes. The derived amino acid sequence showed a clear similarity (37 % identity) to the *E. coli* PheRS α subunit (positions 147 to 211; Fayat *et al.*, 1983). Furthermore, the similarity was most pronounced in regions already known to be well conserved among PheRS α subunits from different origins (preliminary sequence alignments shown by Keller, 1990*b*). It was thus likely that at least part of the *pheS* gene from *T. thermophilus* had been cloned. The determined nucleotide sequence allowed us to locate the *pheS* gene on the cloned fragments with the help of the *Pvu*II site also encountered in the restriction map of Fig. 10. From the position and presumptive size of the gene (1.1 kb; deduced from M_r=40'000 for the α subunit; Ankilova *et al.*, 1988), it became evident that the entire *pheS* gene must be present on all three cloned fragments (Fig. 10).

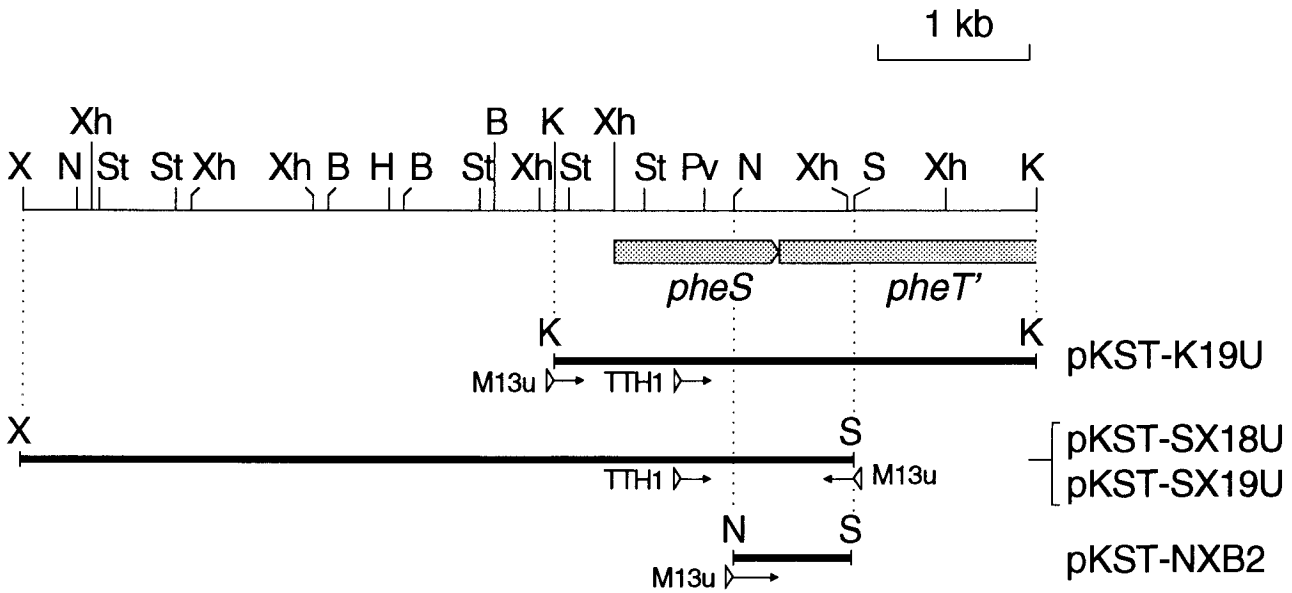


Figure 10. Restriction map of the cloned *pheS* gene region from *T. thermophilus*. The overlapping *T. thermophilus* DNA fragments from plasmids pKST-K19U, pKST-SX18U and pKST-SX19U were mapped with *Xba*I (X), *Nco*I (N), *Xho*I (Xh), *Stu*I (St), *Bam*HI (B), *Hin*dIII (H), *Kpn*I (K), *Pvu*II (Pv) and *Sph*I (S). Plasmids pKST-SX18U and pKST-SX19U are distinguished only by the orientation of the polylinker. The locations of *pheS* and the incomplete *pheT* gene are shown, as determined by sequencing (tiny arrows) using the indicated primers (open arrowheads).

In order to determine more of the *pheS* sequence, in particular the region supposed to encode part of the Phe binding site (see Introduction), a suitable restriction fragment was subcloned into pBLS (Keller, 1990*b*); this vector allowed the production of single stranded DNA for sequencing (see Chapter II). The DNA sequence (305 readable nucleotides; Keller, 1990*b*) obtained from the resulting plasmid pKST-NXB2 (Fig. 10) showed again the strong G+C-bias typical for *T. thermophilus* DNA. The deduced amino acid sequence showed 48 % identity to positions 220 to 311 in the *E. coli* PheRS α subunit (Fayat *et al.*, 1983). Those parts of the preliminary *T. thermophilus* sequence relevant for this work will be discussed later in Figs. 12 and 13; more sequence information is given by Keller (1990*b*).

(c) Identification of part of a T. thermophilus pheT-*like gene*

It was of interest to find out whether the *pheT* gene encoding the PheRS β subunit was also present on the cloned *T. thermophilus* fragments. Double stranded sequence determinations were carried out using the universal primer (M13u) and plasmids pKST-K19U and pKST-SX18U. The results from sequencing

of pKST-K19U (~200 bases, not shown) from the *Kpn*I site into the *pheS* upstream region (Fig. 10) did not reveal significant similarities to the *pheST* regions from *E. coli* (Fayat *et al.*, 1983; Mechulam *et al.*, 1985) and *Bacillus subtilis* (Brakhage *et al.*, 1990). However, the sequence of about 180 nucleotides (Keller, 1990*b*) determined from plasmid pKST-SX18U (reading from the *Sph*I site towards the 3'-end of *pheS*; Fig. 10) allowed us to identify the *pheT*-like gene of *T. thermophilus*. This was achieved through comparisons of the deduced (preliminary) amino acid sequence (Keller, 1990*b*) with the PheRS β subunits from other organisms. As shown in Fig. 11 for the best-matching *T. thermophilus* fragment, there was a clear similarity of the *T. thermophilus* sequence to a region well conserved between *E. coli* and *B. subtilis* β subunits; in contrast, the three sequences showed low similarity to the corresponding segment of PheRS from yeast cytoplasm (alignment according to Brakhage *et al.*, 1990). The approximate position of *pheT* on the map in Fig 10, as determined from these sequencing data, corresponded quite well to the *pheST* gene arrangements in *E. coli* and *B. subtilis* (Fayat *et al.*, 1983; Brakhage *et al.*, 1990). Assuming a size of 2.5 kb for *pheT* (M_r=92'000 for the β subunit; Ankilova *et al.*, 1988), it emerged that the gene was not completely contained on the cloned plasmids; the 3'-end of *pheT* must lie at about 0.8 kb downstream of the *pheT*-internal *Kpn*I site in pKST-K18U (Fig. 10).

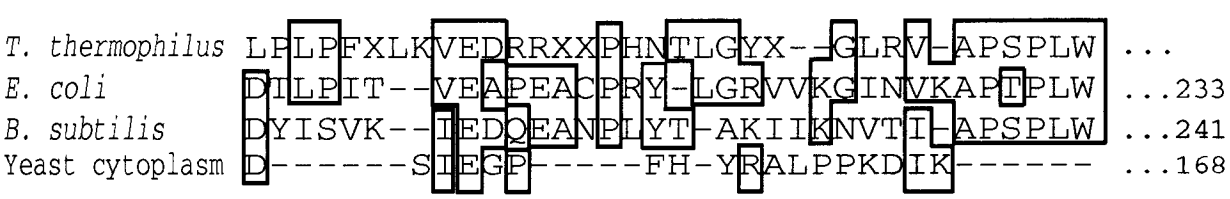


Figure 11. Alignment of partial amino acid sequences of PheRS β subunits derived from *pheT* genes from different origins. Identical amino acid residues (in one-letter-code) are boxed. The preliminary partial amino acid sequence of the *Thermus thermophilus* PheRS β subunit shown here still contains undetermined positions (denoted by 'X'). The alignment of sequences from *Escherichia coli*, *Bacillus subtilis* and yeast cytoplasm was taken from Brakhage *et al.* (1990). Numbers to the right of the published sequences identify residue positions within the respective PheRS β subunits.

4. Discussion

4.1. Cloning of the *pheS* and *pheT* genes from *Salmonella typhimurium*

The cloning of the genes for *S. typhimurium* PheRS was achieved by the selection of plasmid clones able to complement the thermosensitive *E. coli* PheRS α subunit mutant KA2. In a similar way, Brakhage *et al.* (1989) cloned the *pheS* and *pheT* genes from *B. subtilis* (which is far more distantly related to *E. coli* than *S. typhimurium*; Fig. 1). The cloned 10.5 kb *Sal*I fragment from *S. typhimurium* DNA that complemented KA2 was shown by hybridization analysis to contain both *pheS* and *pheT* sequences. From restriction mapping, sequence determinations and comparisons to the *E. coli pheST* region it can be assumed that both genes are entirely present on the cloned fragment. *In vitro* and *in vivo* subunit mixing experiments (Hennecke & Böck, 1975; Hennecke, 1976; Hennecke *et al.*, 1977) can now be carried out to elucidate whether hybrid enzymes consisting of PheRS α and β subunits originating from *S. typhimurium* and *E. coli* genes will be functional. For example, as suggested by Keller (1990*b*), it could be tested whether the presence of the *S. typhimurium pheS* gene alone (*e. g.* the 3.4 kb *Hin*dIII fragment in Fig. 8) can complement strain KA2.

The hybridizations of *S. typhimurium* DNA with the *E. coli* probes resulted in clear signals that were only slightly weaker than those observed in homologous control hybridizations. This already reflected a strong similarity of the *pheST* genes from both organisms which was further confirmed by partial sequence analysis of *S. typhimurium pheS*. The high level of nucleic acid sequence identity (90.2 %) even allowed the use of sequencing primers originally designed for *E. coli*. Similar high degrees of sequence conservation between *S. typhimurium* and *E. coli* genes were reported earlier (*e. g.* by Yanofsky & vanCleemput, 1982; Higgins & Hillyard, 1988; Smith *et al.*, 1990). The same was found in this work for the compared parts of the *thrS* genes from both organisms (87.7 % identity).

Li *et al.* (1989) determined the *S. typhimurium himA* nucleotide sequence to be 95.7 % identical to the *E. coli himA* gene. Furthermore, these authors noticed that the gene order *pheT-himA* was conserved in both organism. In the present work it was discovered that the *S. typhimurium thrS* gene is located on the 5'-side

of the *pheST* genes, at the same distance and in the same orientation as in *E. coli* (Fayat *et al.*, 1983). This confirmed the idea that the gene organization in the region around the *pheST* operon is highly conserved in both organisms.

4.2. Cloning of the *pheS* gene from *Thermus thermophilus*

(a) How good were the hybridization probes?

The restriction fragments of *T. thermophilus* DNA identified by Southern blot analyses (the only common bands hybridizing with different probes) were subsequently shown by partial sequencing of the cloned fragments to contain the *pheS* gene. These *pheS*-carrying fragments hybridized (albeit very weakly) to the heterologous *E. coli pheS* probe and, with varying strenghts, to four of the five oligonucleotides. By adapting the oligonucleotide sequence to the codon usage of *T. thermophilus*, a switch in hybridization specificity towards the *T. thermophilus pheS* gene was obtained. This was clearly reflected by the lack of hybridization to the *pheS* DNA from *E. coli* and *S. typhimurium*.

At the beginning of this work, neither *pheS* gene nor PheRS protein sequence information was available for *T. thermophilus*. Therefore, the oligonucleotides had to be designed solely on the basis of conserved regions in three other known PheRS small subunit sequences (Fig. 4). With the recent publication of the PheRS sequence from *B. subtilis* (Brakhage *et al.*, 1990) it became evident that its α subunit sequences corresponding to oligonucleotides TTH1 to TTH4 were quite well conserved, too. Moreover, it turned out later that TTH1 and TTH2/3 mapped to motifs 1 and 2, respectively, which represent two of the three class II aminoacyl-tRNA synthetase signature sequences (Eriani *et al.*, 1990).

Having cloned and partially sequenced the *T. thermophilus pheS*-like gene, it was of interest now to examine how close prediction and reality came in the case of the designed oligonucleotide probes. While the primer annealing site for oligonucleotide TTH1 was not determined, the strong and nearly exclusive hybridization of TTH1 to the *T. thermophilus pheS* DNA and its success as a *pheS*-specific sequencing primer justified its choice as a probe. The sites covered by TTH2 to TTH5 have been sequenced. It turned out (Fig. 12) that there was a perfect match for one out of the eight variants of TTH5. The better one of the two TTH2 variants contained only one mismatch. At least one mispaired base and two mismatches were apparent for the best of the four TTH3 and TTH4 variants, respectively. The much clearer results in hybridizations with TTH2 as compared to TTH3 (each con-

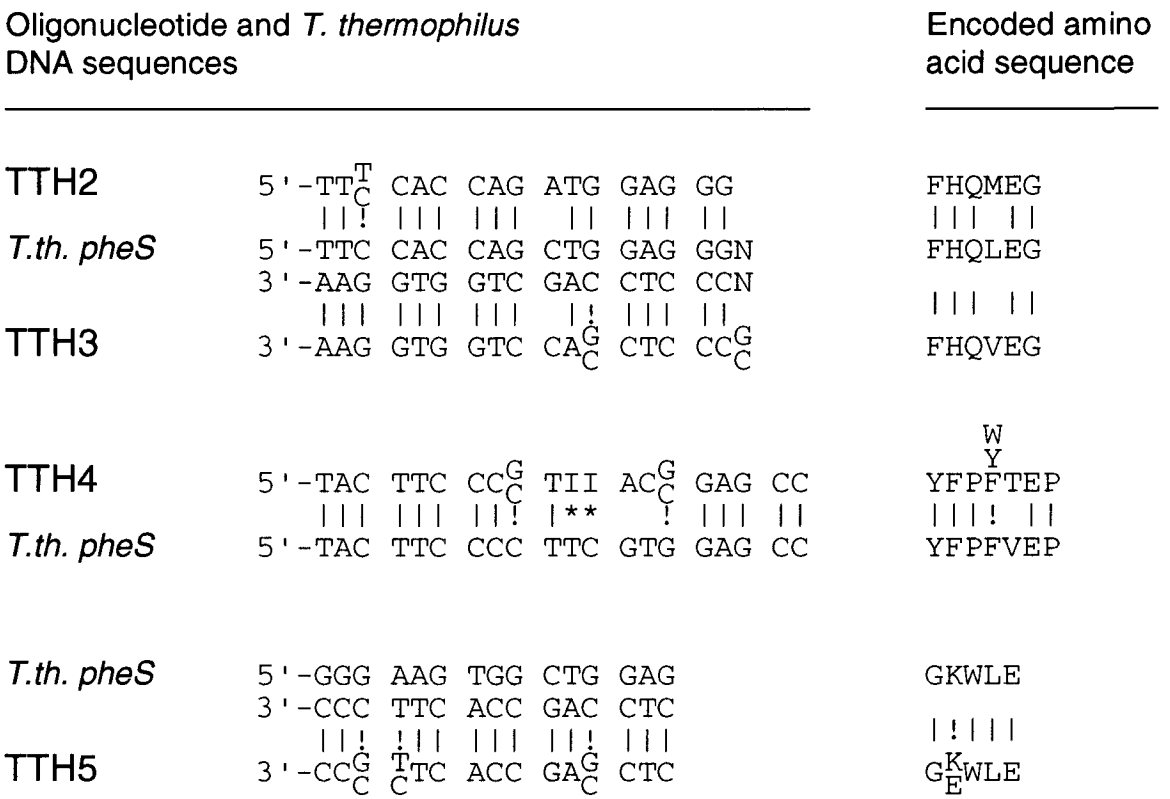


Figure 12. Comparisons of oligonucleotide sequences with the corresponding annealing sites in *T. thermophilus pheS* DNA. The oligonucleotides TTH2 to TTH5 were described in Fig. 5. The *T. thermophilus pheS* sequences (Keller, 1990*b*) were determined as explained in the text. Identical bases and amino acids (in one-letter-code) are connected with vertical lines. An exclamation mark relates to positions where other variants occur besides the perfect match. Asterisks indicate pairings with inosine (I); N denotes a still undetermined nucleotide.

taining one mismatch at the same position) might either be explained by the fewer imperfect TTH2 variants or by the one still undetermined nucleotide in that region of the *T. thermophilus pheS* sequence (Fig. 12). Perhaps there are also conformational differences between the two groups of oligonucleotide variants which might contribute to the observed differential behaviour.

As demonstrated in Fig. 12, almost all of the amino acids derived from the established *T. thermophilus pheS* sequences at the positions of oligonucleotides TTH2 to TTH5 corresponded to the predictions made from the alignment of three known PheRS α subunits (Fig. 4). *A posteriori*, this fully validated the strategy used to design the hybridization probes.

(b) Some features of the preliminary T. thermophilus pheS *and* pheT *sequences*

The partial sequence analysis of the cloned *T. thermophilus* DNA fragments suggested the presence of *pheS* and a part of *pheT*. The codon usage and high G+C-content in all determined sequences (Keller, 1990*b*) corresponded to that of typical *T. thermophilus* genes (Table 1; Oshima & Imahori, 1974; Nureki *et al.*, 1991). The few derived amino acid sequences for the PheRS α subunit were shown to be 37 % to 48 % identical to that of the *E. coli* protein. This is higher than the amino acid sequence identity of 27 % between MetRS from *T. thermophilus* and the corresponding enzyme of *E. coli* (Nureki *et al.*, 1991). However, amino acid identities of up to 72 % between proteins from these two organisms may occur (*e. g.* in elongation factor Tu; Seidler *et al.*, 1987).

As pointed out by Keller (1990*b*), the preliminary sequence comparisons between the *E. coli* and *T. thermophilus* PheRS α subunits revealed characteristic amino acid differences typically found between mesophilic and thermophilic proteins. A main contribution to stabilization of thermophilic proteins seems to stem from an increased hydrophobicity and a decreased flexibility in α-helical regions (Menéndez-Arias & Argos, 1989). As deduced from studies with homologous mesophilic and thermophilic enzymes, this can be achieved by specific exchanges of single amino acids (*e. g.* Ser-to-Ala) at different parts of the polypeptide chain (Argos *et al.*, 1979; Menéndez-Arias & Argos, 1989). In fact, the few derived protein sequences in the *T. thermophilus pheS* gene product contained at least 17 amino acid exchanges of the types listed by Argos *et al.* (1979) and Menéndez-Arias & Argos (1989), when compared to the mesophilic *E. coli* counterpart (Keller, 1990*b*). It is noteworthy that the only Cys residue of the *E. coli* PheRS α subunit (Cys^{272}) was replaced by alanine in the *T. thermophilus* enzyme (Keller, 1990*b*). A drastically lowered content of cysteines has been observed in many *Thermus* proteins when compared to the mesophilic homologs (*e. g.* by Kagawa *et al.*, 1984; Kushiro *et al.*, 1987; Sato *et al.*, 1988; Koyama & Furukawa, 1990; Nureki *et al.*, 1991). It was speculated that the absence of the easily oxidizable thiol groups may increase thermostability of proteins (Mozhaev & Martinek, 1984). Other amino acid differences may be caused by the higher G+C-content in *T. thermophilus* genes which implies the preferred use of G+C-rich codons (Kagawa *et al.*, 1984; Kushiro *et al.*, 1987; Bowen *et al.*, 1988).

The fact that the Cys^{272} residue of the *E. coli* PheRS α subunit was not conserved in the corresponding proteins from *T. thermophilus*, *Bacillus subtilis* and

yeast cytoplasm (Brakhage *et al.*, 1990) suggests that this residue is not essential for PheRS function. This could be exploited when the Cys^{294} PheRS mutant α subunit is chemically derivatized at the reactive SH group to probe for further changes in substrate specificity (Chapter III). For such experiments, it might be advantageous to remove the inherent Cys residue at position 272.

(c) Perspectives of further work on the T. thermophilus pheS/pheT *genes*

The complete sequencing of *pheS* and *pheT* of *T. thermophilus* requires the cloning of the missing part of the *pheT* gene. The first step towards this goal will involve the identification of a suitable restriction fragment carrying the whole *pheT* gene by Southern blot hybridization, using part of the already cloned *pheST* region (Fig. 10) as homologous probe. These experiments are currently underway.

The subsequent cloning can be achieved via colony hybridization or - if both *pheS* and *pheT* are present on the sought fragment - perhaps by complementation of the *E. coli* strain KA2. The latter approach failed to give colonies in the attempts made in this work. However, this may be explained by the fact that the fragments to be cloned did not contain both complete PheRS genes, and functional hybrid enzymes containing the *E. coli* PheRS β subunits may not have been formed or active.

Complementation of *E. coli* mutants by *Thermus* genes, however, is principally possible as demonstrated with *trpE* (Sato *et al.*, 1988), *trpA* and *trpB* (Koyama & Furukawa, 1990) and *leuB, leuC* and *leuD* (Croft *et al.*, 1987). This was probably successful owing to the similarity of gene expression signals between both organisms. Some *T. thermophilus* genes were shown to carry potential promoter sequences similar to the *E. coli* -35/-10 consensus (Croft *et al.*, 1987; Sato *et al.*, 1988; Yakhnin *et al.*, 1990), and (weak) gene expression from putative *T. thermophilus* promoters in *E. coli* was observed (Croft *et al.*, 1987; Seidler *et al.*, 1987; Nureki *et al.*, 1991). In addition, *E. coli*-like Shine-Dalgarno sequences (Shine & Dalgarno, 1974) were encountered in front of virtually all sequenced *Thermus* genes (*e. g.* Nishiyama *et al.*, 1986; Seidler *et al.*, 1987; Bowen *et al.*, 1988; Nicholls *et al.*, 1988; Nicholls *et al.*, 1990; Nureki *et al.*, 1991). The significance of this sequence in *Thermus* is accentuated by the fact that that part of the *E. coli* 16S rRNA 3'-end, which interacts with the ribosome binding site on mRNA, is identical in *T. thermophilus* (Hartmann *et al.*, 1989).

Complementation tests with cloned *T. thermophilus pheST* or *pheT* genes

may be carried out with *E. coli* strains possessing *pheS* or *pheT* mutations that determine thermosensitive PheRSs (Hennecke *et al.*, 1977). Furthermore, *in vitro* enzyme assays with protein extracts of *E. coli* cells producing *T. thermophilus* PheRS should reveal thermoresistant Phe aminoacylation activity (Ankilova *et al.*, 1988). Overproduction of PheRS from *T. thermophilus* in a suitable expression system in *E. coli* (Nureki *et al.*, 1991) may finally yield large amounts of enzyme for structural analyses (see below).

Up to now, only one other aminoacyl-tRNA synthetase gene from the genus *Thermus* has been cloned. The recently published sequence of *T. thermophilus* MetRS (Nureki *et al.*, 1991) revealed that the regulation of gene expression (probably by attenuation) is different to *E. coli* MetRS expression. The complete sequencing of the *T. thermophilus* PheRS gene region should help elucidate whether these genes also constitute an operon controlled by attenuation as in *E. coli* (Fayat *et al.*, 1983; Springer *et al.*, 1983; Springer *et al.*, 1985). Gene expression by attenuation has already been postulated for other *T. thermophilus* genes (Croft *et al.*, 1987; Sato *et al.*, 1988). In this context it should be mentioned, however, that the *pheST* genes from *B. subtilis* are not regulated by the classical attenuation mechanism (Brakhage *et al.*, 1990). The regulation of *T. thermophilus* *pheST* expression may even be analyzed in the homologous background, since transformation of *T. thermophilus* is possible (Koyama & Furukawa, 1990) and plasmid vectors for this organism have recently been developed (Koyama *et al.*, 1990*a,b*).

The major incentive to cloning the *T. thermophilus* PheRS genes was the prospect that the protein sequence will then become available which in turn will allow an elaboration of the three-dimensional structure of this enzyme. Until now, only low resolution structural analyses were possible with *E. coli* PheRS which led to a tentative model of relative subunit arrangements (Dessen *et al.*, 1990; Chapter II). A much better three-dimensional structure can be expected from the combination of data from X-ray diffraction analysis of *T. thermophilus* PheRS crystals (Ankilova *et al.*, 1988) with the amino acid sequence information of PheRS α and β subunits provided by our laboratory. Moreover, Cys residues, introduced by protein engineering into *T. thermophilus* PheRS, may serve to create heavy atom derivatives (Rould *et al.*, 1989) and thus facilitate structural analyses. The approach with *T. thermophilus* PheRS may finally permit the determination of the first high resolution structure of an $\alpha_2\beta_2$ aminoacyl-tRNA synthetase.

4.3. The PheRS sequence that determines phenylalanine specificity is well conserved in *S. typhimurium* and *T. thermophilus*

Another rationale for the cloning of *pheS* genes from *S. typhimurium* and *T. thermophilus* was the idea that a comparative sequence analysis of PheRS α subunits from different organisms would identify functionally important regions. Of special interest was the section of the α subunit thought to be involved in the binding of phenylalanine (Chapter III). Therefore, the preliminary sequence analyses of the two *pheS* genes cloned here aimed at that region. Figure 13 shows an alignment of PheRS α subunit sequences from different origins around the presumptive Phe binding site, which coincides with the generally conserved motif 3 of class II aminoacyl-tRNA synthetases (Eriani *et al.*, 1990). The similarity of the *E. coli* sequence to the *S. typhimurium* region was so strong that the latter sequence was of little help in identifying less conserved and thereby less important residues. However, a close inspection of conserved amino acids within the corresponding sequence from *T. thermophilus* in Fig. 13 revealed a perfect conservation of the F-A-F sequence analyzed previously by mutagenesis (see Chapter III). This provides an independent support for the hypothesis that the analyzed region is an important, PheRS-specific site which may determine specificity for phenylalanine (Chapter III).

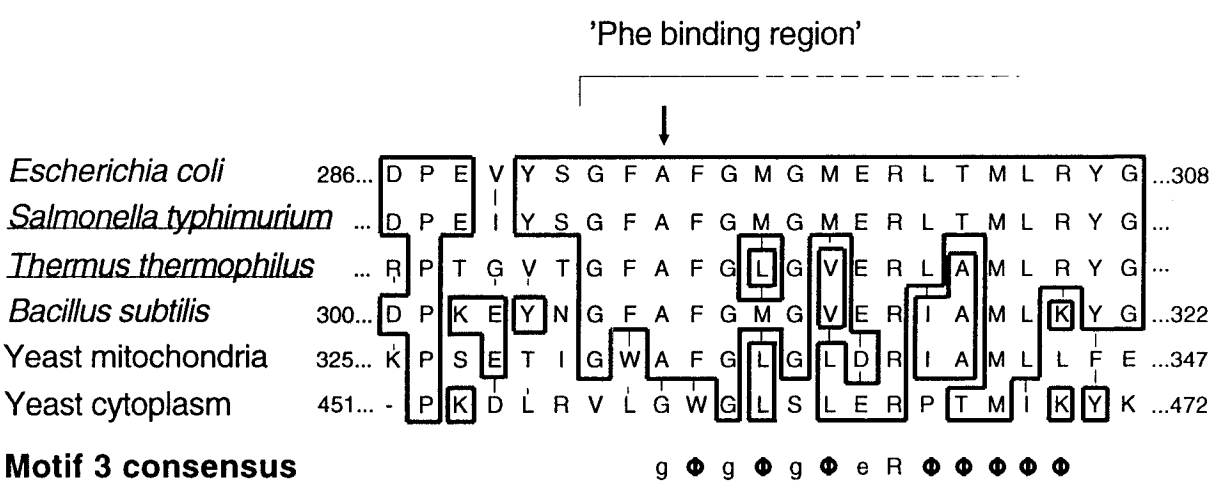


Figure 13. Alignment of PheRS small subunit sequences in a region which presumably interacts with phenylalanine. The Figure contains PheRS sequences corresponding to the region around the *E. coli* α subunit residue Ala^{294} (marked by an arrow) which was shown to be involved in Phe binding (Chapters II and III). The PheRS sequences for the two organisms described in this Chapter (*S. typhimurium* and *T. thermophilus*; species names underlined) were taken from Fig. 9 and from Keller (1990*b*), respectively. References to the other sequences, the criteria for establishing the alignment and the used symbols were detailed in the legend to Fig. 8 of Chapter II.

5. References

Ahmad, S., Weisburg, W. G. & Jensen, R. A. (1990). Evolution of Aromatic Amino Acid Biosynthesis and Application to the Fine-Tuned Phylogenetic Positioning of Enteric Bacteria. *J. Bacteriol.* **172**, 1051-1061.

Ankilova, V. N., Reshetnikova, L. S., Chernaya, M. M. & Lavrik, O. I. (1988). Phenylalanyl-tRNA Synthetase from *Thermus thermophilus* HB8. Purification and Properties of the Crystallizing Enzyme. *FEBS Lett.* **227**, 9-13.

Argos, P., Rossmann, M. G., Grau, U. M., Zuber, H., Frank, G. & Tratschin, J. D. (1979). Thermal Stability and Protein Structure. *Biochemistry*, **18**, 5698-5703.

Bowen, D., Littlechild, J. A., Fothergill, J. E., Watson, H. C. & Hall, L. (1988). Nucleotide Sequence of the Phosphoglycerate Kinase Gene from the Extreme Thermophile *Thermus thermophilus*. *Biochem. J.* **254**, 509-517.

Brakhage, A. A., Putzer, H., Shazand, K., Röschenthaler, R. J. & Grunberg-Manago, M. (1989). *Bacillus subtilis* Phenylalanyl-tRNA Synthetase Genes: Cloning and Expression in *Escherichia coli* and *B. subtilis*. *J. Bacteriol.* **171**, 1228-1232.

Brakhage, A. A., Wozny, M. & Putzer, H. (1990). Structure and Nucleotide Sequence of the *Bacillus subtilis* Phenylalanyl-tRNA Synthetase Genes. *Biochimie*, **72**, 725-734. - Erratum (1991): *Biochimie*, **73**, 127.

Brick, P., Bhat, T. N. & Blow, D. M. (1989). Structure of Tyrosyl-tRNA Synthetase Refined at 2.3 Å Resolution. Interaction of the Enzyme with the Tyrosyl Adenylate Intermediate. *J. Mol. Biol.* **208**, 83-98.

Brunie, S., Zelwer, C. & Risler, J.-L. (1990). Crystallographic Study at 2.5 Å Resolution of the Interaction of Methionyl-tRNA Synthetase from *Escherichia coli* with ATP. *J. Mol. Biol.* **216**, 411-424.

Chernaya, M. M., Korolev, S. V., Reshetnikova, L. S. & Safro, M. G. (1987). Preliminary Crystallographic Study of the Phenylalanyl-tRNA Synthetase from *Thermus thermophilus* HB8. *J. Mol. Biol.* **198**, 555-556.

Croft, J. E., Love, D. R. & Bergquist, P. L. (1987). Expression of Leucine Genes from an Extremely Thermophilic Bacterium in *Escherichia coli*. *Mol. Gen. Genet.* **210**, 490-497.

Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N. & Leberman, R. (1990). A Second Class of Synthetase Structure Revealed by X-Ray Analysis of *Escherichia coli* Seryl-tRNA Synthetase at 2.5 Å. *Nature (London)*, **347**, 249-255.

Dessen, P., Ducruix, A., May, R. P. & Blanquet, S. (1990). Low-Resolution Structure of the Tetrameric Phenylalanyl-tRNA Synthetase from *Escherichia coli*. A Neutron Small-Angle Scattering Study of Hybrids Composed of Protonated and Deuterated Protomers. *Biochemistry*, **29**, 3039-3045.

Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990). Partition of tRNA Synthetases into Two Classes Based on Mutually Exclusive Sets of Sequence Motifs. *Nature (London)*, **347**, 203-206.

Fayat, G., Mayaux, J.-F., Sacerdot, C., Fromant, M., Springer, M., Grunberg-Manago, M. & Blanquet, S. (1983). *Escherichia coli* Phenylalanyl-tRNA Synthetase Operon Region. Evidence for an Attenuation Mechanism. Identification of the Gene for the Ribosomal Protein L20. *J. Mol. Biol.* **171**, 239-261.

Feinberg, A. P. & Vogelstein, B. (1984). Addendum to 'A Technique for Radiolabeling DNA Restriction Endonuclease Fragments to High Specific Activity'. *Anal. Biochem.* **137**, 266-267.

Garber, M. B., Yaremchuk, A. D., Tukalo, M. A., Egorova, S. P., Berthet-Colominas, C. & Leberman, R. (1990*a*). Crystals of Seryl-tRNA Synthetase from *Thermus thermophilus*. Preliminary Crystallographic Data. *J. Mol. Biol.* **213**, 631-632.

Garber, M. B., Yaremchuk, A. D., Tukalo, M. A., Egorova, S. P., Fomenkova, N. P. & Nikonov, S. V. (1990*b*). Crystals of Threonyl-tRNA Synthetase from *Thermus thermophilus*. Preliminary Crystallographic Data. *J. Mol. Biol.* **214**, 819-820.

Hartmann, R. K., Wolters, J., Kröger, B., Schultze, S., Specht, T. & Erdmann, V. A. (1989). Does *Thermus* Represent Another Deep Eubacterial Branching? *System. Appl. Microbiol.* **11**, 243-249.

Hennecke, H. (1976). Use of Mutant Enzymes to Demonstrate the Presence of Two Active Sites on Phenylalanyl-tRNA Synthetase from *Escherichia coli*. *FEBS Lett.* **72**, 182-186.

Hennecke, H. & Böck, A. (1975). Altered α Subunits in Phenylalanyl-tRNA Syn-

thetases from *p*-Fluorophenylalanine-Resistant Strains of *Escherichia coli*. *Eur. J. Biochem.* **55**, 431-437.

Hennecke, H., Springer, M. & Böck, A. (1977). A Specialized Transducing λ Phage Carrying the *Escherichia coli* Genes for Phenylalanyl-tRNA Synthetase. *Mol. Gen. Genet.* **152**, 205-210.

Higgins, N. P. & Hillyard, D. (1988). Primary Structure and Mapping of the *hupA* Gene of *Salmonella typhimurium*. *J. Bacteriol.* **170**, 5751-5758.

Kagawa, Y., Nojima, H., Nukiwa, N., Ishizuka, M., Nakajima, T., Yasuhara, T., Tanaka, T. & Oshima, T. (1984). High Guanine plus Cytosine Content in the Third Letter of Codons of an Extreme Thermophile. *J. Biol. Chem.* **259**, 2956-2960.

Kawase, Y., Iwai, S., Inoue, H., Miura, K. & Ohtsuka, E. (1986). Studies on Nucleic Acid Interactions I. Stabilities of Mini-Duplexes ($dG_2A_4XA_4G_2 \cdot dC_2T_4YT_4C_2$) and Self-Complementary d(GGGAAXYTTCCC) Containing Deoxyinosine and Other Mismatched Bases. *Nucl. Acids Res.* **14**, 7727-7736.

Keller, B. (1990*a*). Phenylalanyl-tRNA Synthetase (PheRS) aus *Escherichia coli*, *Salmonella typhimurium* und *Thermus thermophilus*: Identifizierung der PheRS-Gene durch DNA-Hybridisierung. Semesterarbeit ETH Zürich.

Keller, B. (1990*b*). Versuche zur Klonierung von *pheS* Genen aus *Salmonella typhimurium* und *Thermus thermophilus*. Diplomarbeit ETH Zürich.

Koerner, T. J., Myers, A. M., Lee, S. & Tzagoloff, A. (1987). Isolation and Characterization of the Yeast Gene Coding for the α Subunit of Mitochondrial Phenylalanyl-tRNA Synthetase. *J. Biol. Chem.* **262**, 3690-3696.

Koyama, Y. & Furukawa, K. (1990). Cloning and Sequence Analysis of Tryptophan Synthetase Genes of an Extreme Thermophile, *Thermus thermophilus* HB27: Plasmid Transfer from Replica-Plated *Escherichia coli* Recombinant Colonies to Competent *T. thermophilus* Cells. *J. Bacteriol.* **172**, 3490-3495.

Koyama, Y., Arikawa, Y. & Furukawa, K. (1990*a*). A Plasmid Vector for an Extreme Thermophile, *Thermus thermophilus*. *FEMS Microbiol. Lett.* **72**, 97-102.

Koyama, Y., Okamoto, S. & Furukawa, K. (1990*b*). Cloning of α- and β-Galactosidase Genes from an Extreme Thermophile, *Thermus* Strain T2, and Their Expression in *Thermus thermophilus* HB27. *Appl. Environ. Microbiol.*

56, 2251-2254.

Kunai, K., Machida, M., Matsuzawa, H. & Otha, T. (1986). Nucleotide Sequence and Characteristics of the Gene for L-Lactate Dehydrogenase of *Thermus caldophilus* GK24 and the Deduced Amino-Acid Sequence of the Enzyme. *Eur. J. Biochem.* **160**, 433-440.

Kushiro, A., Shimizu, M. & Tomita, K. (1987). Molecular Cloning and Sequence Determination of the *tuf* Gene Coding for the Elongation Factor Tu of *Thermus thermophilus* HB8. *Eur. J. Biochem.* **170**, 93-98.

Li, Z. J., Hillyard, D. & Higgins, P. (1989). Nucleotide Sequence of the *Salmonella typhimurium himA* Gene. *Nucl. Acids Res.* **17**, 8880.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Mayaux, J.-F., Fayat, G., Fromant, M., Springer, M., Grunberg-Manago, M. & Blanquet, S. (1983). Structural and Transcriptional Evidence for Related *thrS* and *infC* Expression. *Proc. Nat. Acad. Sci., U.S.A.* **80**, 6152-6156.

Mechulam, Y., Fayat, G. & Blanquet, S. (1985). Sequence of the *Escherichia coli pheST* Operon and Identification of the *himA* Gene. *J. Bacteriol.* **163**, 787-791.

Menéndez-Arias, L. & Argos, P. (1989). Engineering Protein Thermal Stability. Sequence Statistics Point to Residue Substitutions in α Helices. *J. Mol. Biol.* **206**, 397-406.

Mozhaev, V. V. & Martinek, K. (1984). Structure-Stability Relationships in Proteins: New Approaches to Stabilizing Enzymes. *Enzyme Microb. Technol.* **6**, 50-59.

Nicholls, D. J., Sundaram, T. K., Atkinson, T. & Minton, N. P. (1988). Nucleotide Sequence of the Succinyl-CoA Synthetase Alpha-Subunit from *Thermus aquaticus* B. *Nucl. Acids Res.* **16**, 9858.

Nicholls, D. J., Sundaram, T. K., Atkinson, T. & Minton, N. P. (1990). Cloning and Nucleotide Sequences of the *mdh* and *sucD* Genes from *Thermus aquaticus* B. *FEMS Microbiol. Lett.* **70**, 7-14.

Nishiyama, M., Matsubara, N., Yamamoto, K., Iijima, S., Uozumi, T. & Beppu, T. (1986). Nucleotide Sequence of the Malate Dehydrogenase Gene of *Thermus flavus* and Its Mutation Directing an Increase in Enzyme Activity. *J. Biol.*

Chem. **261**, 14178-14183.

Norrander, J., Kempe, T. & Messing, J. (1983). Construction of Improved M13 Vectors Using Oligodeoxynucleotide-Directed Mutagenesis. *Gene*, **26**, 101-106.

Nureki, O., Muramatsu, T., Suzuki, K., Kohda, D., Matsuzawa, H., Ohta, T., Miyazawa, T. & Yokoyama, S. (1991). Methionyl-tRNA Synthetase Gene from an Extreme Thermophile, *Thermus thermophilus* HB8. Molecular Cloning, Primary-Structure Analysis, Expression in *Escherichia coli*, and Site-Directed Mutagenesis. *J. Biol. Chem.* **266**, 3268-3277.

Ørskov, F. (1984). Genus I. *Escherichia.* In *Bergey's Manual of Systematic Bacteriology.* (Krieg, N. R. & Holt, J. G., eds), vol. 1, pp. 420-423, Williams & Wilkins, Baltimore, MD.

Oshima, T. & Imahori, K. (1974). Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a Nonsporulating Thermophilic Bacterium from a Japanese Thermal Spa. *Int. J. Syst. Bacteriol.* **24**, 102-112.

Rould, M. A., Perona, J. J., Söll, D. & Steitz, T. A. (1989). Structure of *E. coli* Glutaminyl-tRNA Synthetase Complexed with $tRNA^{Gln}$ and ATP at 2.8 Å Resolution. *Science*, **246**, 1135-1142.

Ruff, M., Cavarelli, J., Mikol, V., Lorber, B., Mitschler, A., Giege, R., Thierry, J. C. & Moras, D. (1988). A High Resolution Diffracting Crystal Form of the Complex between Yeast $tRNA^{Asp}$ and Aspartyl-tRNA Synthetase. *J. Mol. Biol.* **201**, 235-236.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual* (2^{nd} ed), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sanni, A., Mirande, M., Ebel, J.-P., Boulanger, Y., Waller, J.-P. & Fasiolo, F. (1988). Structure and Expression of the Genes Encoding the α and β Subunits of Yeast Phenylalanyl-tRNA Synthetase. *J. Biol. Chem.* **263**, 15407-15415.

Sato, S., Nakada, Y., Kanaya, S. & Tanaka, T. (1988). Molecular Cloning and Nucleotide Sequence of *Thermus thermophilus* HB8 *trpE* and *trpG*. *Biochim. Biophys. Acta*, **950**, 303-312.

Seidler, L., Peter, M., Meissner, F. & Sprinzl, M. (1987). Sequence and Identification of the Nucleotide Binding Site for the Elongation Factor Tu from *Thermus*

thermophilus HB8. *Nucl. Acids Res.* **15**, 9263-9277.

Shine, J. & Dalgarno, L. (1974). The 3'-Terminal Sequence of *Escherichia coli* 16S Ribosomal RNA: Complementarity to Nonsense Triplets and Ribosome Binding Sites. *Proc. Nat. Acad. Sci., U.S.A.* **71**, 1342-1346.

Smith, C. M., Koch, W. H., Franklin, S. B., Foster, P. L., Cebula, T. A. & Eisenstadt, E. (1990). Sequence Analysis and Mapping of the *Salmonella typhimurium* LT2 *umuDC* Operon. *J. Bacteriol.* **172**, 4964-4978.

Southern, E. M. (1975). Detection of Specific Sequences Among DNA Fragments Separated by Gel Electrophoresis. *J. Mol. Biol.* **98**, 503-517.

Springer, M., Mayaux, J.-F., Fayat, G., Plumbridge, J. A., Graffe, M., Blanquet, S. & Grunberg-Manago, M. (1985). Attenuation Control of the *Escherichia coli* Phenylalanyl-tRNA Synthetase Operon. *J. Mol. Biol.* **181**, 467-478.

Springer, M., Trudel, M., Graffe, M., Plumbridge, J., Fayat, G., Mayaux, J. F., Sacerdot, C., Blanquet, S. & Grunberg-Manago, M. (1983). *Escherichia coli* Phenylalanyl-tRNA Synthetase Operon Is Controlled by Attenuation *in Vivo*. *J. Mol. Biol.* **171**, 263-279.

Yakhnin, A. V., Vorozheykina, D. P. & Matvienko, N. I. (1990). Nucleotide Sequence of the *Thermus thermophilus* HB8 *rps12* and *rps7* Genes Coding for the Ribosomal Proteins S12 and S7. *Nucl. Acids Res.* **18**, 3659.

Yanofsky, C. & vanCleemput M. (1982). Nucleotide Sequence of *trpE* of *Salmonella typhimurium* and Its Homology with the Corresponding Sequence of *Escherichia coli*. *J. Mol. Biol.* **154**, 235-246.

Abstract

The enzyme phenyalanyl-tRNA synthetase (PheRS) catalyses the covalent coupling of the amino acid phenylalanine to the corresponding $tRNA^{Phe}$. Neither the tertiary structure nor the topology of the substrate binding sites are known for PheRS. In this work, it was attempted, by molecular-genetic means, to characterize the phenylalanine binding site of PheRS from the bacterium *Escherichia coli*.

To localize amino acid residues in PheRS which interact with phenylalanine, two types of PheRS mutants with altered substrate binding properties were analyzed. Since both mutations mapped to *pheS* (the gene encoding the α subunit of PheRS), the mutated *pheS* genes were cloned and sequenced. One of the mutations, causing a lowered affinity for phenylalanine in PheRS, resulted in an amino acid exchange in motif 2, a generally conserved sequence of class II aminoacyl-tRNA synthetases to which PheRS belongs. The second mutation affected motif 3, another class II-specific sequence. The resulting amino acid exchange (alanine-to-serine at position 294 of the PheRS α subunit) was responsible for the resistance of the *E. coli* mutant strain against the substrate analogue *para*-fluoro-phenylalanine (*p*-F-Phe). The exclusion of *p*-F-Phe from the enzymatic reaction by the mutant PheRS (in contrast to the wild-type enzyme) may now be explained by alterations of steric interactions.

The sequence around position 294 (Gly^{292}-Phe^{293}-Ala^{294}-Phe^{295}-Gly^{296}) is also well conserved in *Bacillus subtilis* and yeast cytoplasmic and mitochondrial PheRSs. To enlarge the number of sequences for comparisons, the *pheS* genes from *Salmonella typhimurium* and *Thermus thermophilus* were cloned and partially sequenced. The sequence data confirmed that amino acid residue 294 was located in a PheRS-specific site. The choice of *T. thermophilus* was influenced by the fact that its PheRS had been crystallized; the availability of amino acid sequence information will therefore greatly help elucidate the tertiary structure.

Replacement of the phenylalanines at positions 293 and 295 by selected other amino acids revealed that these residues do not directly interact with the amino acid substrate, but seem to affect PheRS stability. Replacements at position 294 showed that this residue contacts the *para*-position of the substrate's aromatic ring. The most interesting exchange, alanine-to-glycine, generated an enzyme with relaxed substrate specificity for *para*-substituted phenylalanine analogues. This mutant can possibly be exploited in an *in vivo* system for the incorporation of non-proteinogenic amino acids into proteins of pharmaceutical relevance.

Kurzfassung

Das Enzym Phenylalanyl-tRNA Synthetase (PheRS) katalysiert die kovalente Kopplung der Aminosäure Phenylalanin an die zugehörige $tRNA^{Phe}$. Von der PheRS sind weder Tertiärstruktur noch Topologie der Substratbindungsstellen bekannt. In dieser Arbeit sollte versucht werden, mit molekulargenetischen Methoden die Bindungsstelle für Phenylalanin in der PheRS des Bakteriums *Escherichia coli* zu charakterisieren.

Für die Lokalisierung von PheRS-Aminosäuren, die mit Phenylalanin interagieren, wurden zwei Typen von PheRS-Mutanten mit beeinträchtigter Substratbindung untersucht. Da die Mutationen in *pheS*, dem Gen für die PheRS α-Untereinheit kartierten, wurden die mutierten *pheS*-Gene kloniert und sequenziert. Eine Mutation, die zu einer erniedrigten Affinität des Enzyms für Phenylalanin führte, ergab einen Aminosäureaustausch in Motiv 2, einer generell konservierten Sequenz der sogenannten Klasse II-Aminoacyl-tRNA Synthetasen, zu denen PheRS gehört. Die zweite Mutation betraf das ebenfalls für Klasse II spezifische Sequenz-Motiv 3. Der resultierende Aminosäure-Austausch (Alanin zu Serin an Position 294 der PheRS α-Untereinheit) war verantwortlich für die Resistenz eines entsprechenden *E. coli* Mutanten-Stamms gegen das Substrat-Analog *p*-Fluor-Phenylalanin (*p*-F-Phe). Der Ausschluss von *p*-F-Phe von der Reaktion der Mutanten-PheRS (im Gegensatz zum Wildtyp-Enzym) konnte mit veränderten sterischen Interaktionen erklärt werden.

Die Region unmittelbar um Position 294 (Gly^{292}-Phe^{293}-Ala^{294}-Phe^{295}-Gly^{296}) ist gut konserviert in PheRS von *Bacillus subtilis* und Hefe (aus Cytoplasma und Mitochondrien). Um weitere Sequenzvergleiche anstellen zu können, wurden die *pheS*-Gene aus *Salmonella typhimurium* und *Thermus thermophilus* kloniert und teilweise sequenziert. Die Sequenzierungsresultate bestätigten, dass Aminosäure 294 in einer für PheRS spezifischen Region liegt. Die Wahl von *T. thermophilus* erfolgte auch deshalb, weil *T. thermophilus* PheRS-Kristalle vorhanden sind, die es später erlauben sollten, mit Hilfe der Aminosäuresequenz die dreidimensionale Struktur zu ermitteln.

Ein gezielter Ersatz der Phenylalanine 293 und 295 durch ausgewählte andere Aminosäuren zeigte, dass diese Positionen nicht direkt mit dem Aminosäure-Substrat interagieren, jedoch für die PheRS-Stabilität wichtig sind. Austausche an Position 294 bestätigten hingegen, dass hier Interaktionen mit der *para*-Position

des aromatischen Rings des Substrats stattfinden. Der interessanteste Austausch, Alanin zu Glycin, ergab ein Enzym mit relaxierter Spezifität für *para*-substituierte Phenylalanin-Analoga. Dieses Mutanten-Enzym lässt sich möglicherweise in einem *in vivo* Produktionssystem für den Einbau von nicht-proteinogenen Aminosäuren in pharmazeutisch relevante Proteine einsetzen.

Seite Leer /
Blank leaf

Curriculum Vitae

Born on June 12, 1961	in Wetzikon (ZH), Switzerland
1968 - 1976	Primary and secondary education in Wetzikon (ZH)
1976 - 1980	Gymnasium, Kantonsschule Zürcher Oberland in Wetzikon; final examination: Matura type C
1981 - 1986	Studies in Biotechnology at the Swiss Federal Institute of Technology (ETH) in Zürich; Diploma in Natural Sciences
1986	Diploma Thesis at the Institute of Molecular Biology I, University of Zürich.
1986 - 1991	Assistant researcher at the Institute of Microbiology ETH Zürich; Ph. D. Thesis