Phenylalanyl-tRNA synthetase

Shruti Chakraborty
Rajat Banerjee
Department of Biotechnology, University of Calcutta, Kolkata, West Bengal, India

Abstract: Phenylalanyl-tRNA synthetase (PheRS), a member of class II aminoacyl-tRNA synthetases, catalyzes the synthesis of phenylalanyl-tRNA\(^{Phe}\) (Phe-tRNA\(^{Phe}\)). Hence, like other aminoacyl-tRNA synthetases, PheRS also plays a crucial role in the cellular translation process. Structural characterization demonstrates remarkable architectural diversity ranging from monomer to hetero-oligomer. Heterotetrameric PheRS contains an editing domain to proofread misincorporation of non-cognate amino acids. However, editing activity is absent in monomeric PheRS. PheRS has also shown some noncanonical functions, such as DNA binding properties, suggesting its involvement in complex regulatory pathways. Engineered mutants with relaxed substrate specificities of PheRS can be a promising tool for chemical and synthetic biology. Because of substantial structural variations among species due to evolutionary divergence, PheRS may be validated as a novel drug target. Human PheRS gene mutations have recently been implicated in several neurological disorders prompting structure-function studies to elucidate the molecular role of PheRS in such pathologies. In this review on PheRS, we will briefly cover all of the aforementioned aspects and our current understanding about the enzyme.

Keywords: PheRS, structural organization, editing, disease, unnatural amino acid, drug targeting

Introduction

Aminoacyl-tRNA synthetases (aaRSs) form a set of housekeeping enzymes that play a crucial role in cellular translation. aaRSs chemically ligate tRNAs with their cognate amino acids producing the aminoacyl-tRNAs (aa-tRNAs) that subsequently bind to the elongation factor, facilitating delivery to the ribosome for mRNA decoding.\(^1\) The charging of tRNA is a two-step process. In the first step, aaRS activates the cognate amino acid (in a tRNA-dependent or a tRNA-independent manner) in the presence of ATP to form aaRS-aminoacyladenylate complex (aaRS-AMP). In the second step, this activated aminoacyl moiety is transferred to the 2' OH or 3' OH group of the terminal adenine of the acceptor end of cognate tRNA.\(^2-4\)

\[
\text{aaRS + amino acid + ATP} \rightarrow \text{aaRS-AMP + PPi} \quad (1)
\]

\[
\text{aaRS-AMP + tRNA} \rightarrow \text{aa-tRNA + AMP} \quad (2)
\]

Twenty-two different amino acids are ligated to their corresponding cognate tRNAs by 22 different aaRSs. However, based on their sequence motifs and catalytic domain topologies, aaRSs can be divided into two classes, class I and class II (Table 1).\(^5-8\)

Every aaRS contains a catalytic domain and an anticodon-binding domain (ABD). In addition to these two domains several aaRSs also possess an editing domain to maintain
Table I Classification of aaRSs

<table>
<thead>
<tr>
<th>Class</th>
<th>Subclass (subunit organization)</th>
<th>Enzymes</th>
<th>Signature motif</th>
<th>Catalytic domain architecture</th>
<th>Site of aminoacylation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td>a (monomeric)</td>
<td>CysRS, IleRS, LeuRS, MetRS, and ValRS</td>
<td>HIGH, and KMSKS</td>
<td>Rossmann fold</td>
<td>2′ OH</td>
</tr>
<tr>
<td></td>
<td>b (monomeric)</td>
<td>ArgRS, GinRS, GluRS, and LysRS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c (homodimeric)</td>
<td>TrpRS, and TyrRS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class II</td>
<td>a (homodimeric)</td>
<td>GlyRS, HisRS, ProRS, SerRS, and ThrRS</td>
<td>Motif 1, motif 2, and motif 3</td>
<td>Antiparallel fold</td>
<td>3′ OH (except PheRS)</td>
</tr>
<tr>
<td></td>
<td>b (homodimeric)</td>
<td>AsnRS, AspRS, and LysRS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c (tetrameric)</td>
<td>AlaRS (heterotetramer), GlyRS (heterotetramer), PheRS (heterotetramer), PyrRS (homodimer), and SepRS (homotetramer)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: *MetRS also exists in homodimeric form under same subclass. *LysRS is present in both the classes. *GlyRS exists in two different forms in two subclasses. *AlaRS also exists in monomeric form under same subclass. *PheRS also exists in monomeric form under same subclass. Adapted from Klipcan L, Finarov i, Moor N, Safro MG. Structural aspects of phenylalanylation and quality control in three major forms of phenylalanyl-tRNA synthetase. J Amino Acids. 2010:2010:983503. Copyright © 2010 Liron Klipcan et al.

Abbreviations: aaRSs, aminoacyl-tRNA synthetases; CysRS, cysteinyl-tRNA synthetase; GinRS, arginyl-tRNA synthetase; HisRS, histidyl-tRNA synthetase; IleRS, isoleucyl-tRNA synthetase; LeuRS, leucyl-tRNA synthetase; MetRS, methionyl-tRNA synthetase; PheRS, phenylalanyl-tRNA synthetase; ProRS, prolyl-tRNA synthetase; SerRS, seryl-tRNA synthetase; ThrRS, threonyl-tRNA synthetase; AsnRS, asparaginyl-tRNA synthetase; AspRS, aspartyl-tRNA synthetase; AlaRS, alaninyl-tRNA synthetase; PylRS, pyrrolysyl-tRNA synthetase; SepRS, O-phosphoseryl-tRNA synthetase.

the translational fidelity. During evolution some aaRSs have also acquired additional domains that are mostly involved in noncanonical functions, like transcriptional regulation, signal transduction, angiogenesis, inflammation, etc.9

In mammalian systems, two separate sets of genes encode two distinct sets of aaRSs, cytosolic, and mitochondrial. Mitochondrial aaRSs (mtaaRSs) are also encoded in the nucleus, translated in the cytosol, and transported to the mitochondrial lumen. In humans, out of 37 aaRSs, 17 aaRSs are exclusively functional in the cytosol, including bifunctional glutamyl-prolyl-tRNA synthetase (GluProRS; responsible for charging both Glu and Pro), 18 are exclusively active in the mitochondria and 2 are dually localized (GlyRS and LysRS).10

In plant systems, a total of 45 aaRSs are encoded for three translationally active chambers – cytosol, mitochondria, and chloroplast. Among these 45 aaRSs, 24 are organellar aaRSs, of which 22 are found to be dually targeted – 17 are found both in the mitochondria and the plastid, 5 are shared between the mitochondria, and the cytosol and 2 are localized exclusively in the chloroplast.11,12

Protozoans also have three translationally active compartments; namely the cytosol, the mitochondrion and the relic plastid or the apicoplast. aaRSs are encoded by nuclear genes and dually targeted to both the cytosol and the apicoplast. However, the mitochondrial fractions of aaRSs are found to be absent in protozoa, suggesting mitochondrial import of aa-tRNAs.13–15 Among the 35 aaRSs found in plasmodia, 16 are restricted to the cytoplasm, 15 for apicoplast, and 4 single copy aaRSs are shared between the cytoplasm and the apicoplast. Recently, it is reported that malarial parasite Plasmodium falciparum encodes a copy of phenylalanyl-tRNA synthetase (PheRS) that is exclusively located in the mitochondria, apart from two other copies, one each for the cytosol and the apicoplast.16

aaRSs have been evaluated as a potential drug target for infectious diseases. Considering that drug resistance happens to be the key nemesis of current antibiotic therapeutics, aaRSs are very attractive as novel drug targets to the researchers, both academia and industries.17,18 Evolutionary divergence of prokaryotic and eukaryotic aaRSs allow orthogonal drug targeting; justifying pathogenic aaRSs as attractive choice for drug development.

In recent years, several disease-causing mutations have been described for both cytosolic aaRSs (ctaRSs) and mtaaRSs encoding genes and mostly found to be associated with disorders of the nervous system.19 Whereas the mutations found in the ctaRSs encoding genes are associated with Charcot–Marie–Tooth disease and related neuropathies, mutations in the mtaaRSs encoding genes are found to be associated with a wider range of diseases.10

In this review, we will discuss PheRS, which is responsible for producing Phe-tRNA\(^*\)e, PheRS has been identified as a member of subclass IIC based on the structural makeup of the catalytic domain and three signature motifs (1, 2, and 3) of α subunit. PheRS is the only exception among the class II aaRSs that attaches the amino acid moiety to the 2′ OH group of the terminal ribose of the acceptor end of the cognate tRNA.20 Here we provide an update on recent studies of PheRS including its role in the pathophysiology, as a protein engineering tool, as a drug target, with a focus on its structural, and functional diversity across the kingdom.
Structural diversity of PheRS

Structurally, PheRS is the most diverse enzyme of the aaRSs family. They show great variations in subunit structure/composition across the three kingdoms of life and exist in three forms – heterotetrameric (αβ2) in eubacteria, heterotetrameric (αβ2) in archaea/eukaryotic cytoplasm, and monomeric in organelles.21 While heterotetrameric (αβ2) form is the most complex among the aaRSs family, monomeric organellar PheRS is the smallest that exhibits aminoacylation activity (Figure 1).22 Multiple sequence alignment shows that archaeal/eukaryotic cytosolic PheRS (crPheRS) have originated from a single ancestor, while eubacterial PheRS is thought to be the predecessor of monomeric organellar PheRS.22

Heterotetrameric PheRS (αβ2): structural insight

The heterotetrameric organization (αβ2) is evolutionarily conserved among eubacterial and archaeal/eukaryotic crPheRS. Whereas α subunit is responsible for catalytic activity, the major function of β subunit is tRNA recognition, binding, and editing of misacylated tRNA.23,24 The eubacterial PheRS has 11 distinct structural domains: A0–A2 in the α subunit and B1–B8 in the β subunit (Table 2).25

Both α and β subunits of archaeal/eukaryotic crPheRS exhibit significant length difference from their bacterial counterparts: 1) the α subunit of archaeal/eukaryotic crPheRS is ~150 residues longer at the N terminus than the bacterial one, 2) β subunit is ~200 amino acids shorter than its bacterial counterpart due to the absence of B2 and B8 domains in the former (Table 3).26 Apart from difference in length, there is also marked divergence in domain organization between the two kingdoms.27,28

Conformational switch: “non-active” to “active” state

α or β subunit on its own is catalytically inactive.29,30 Affinity labeling to localize the substrate binding site in Escherichia coli PheRS (ecPheRS) along with crystallographic data of Thermus thermophilus PheRS (ttPheRS) complexed with phenylalanine (Phe) and phenylalanine-adenylate (PheO-HAMP), the synthetic analog of phenylalanine-adenylate (PheAMP), corroborate that the α subunit of PheRS on its own is unable to carry out the first step of the aminoacylation reaction.31,32 Sequence analysis shows that the amino acid binding loop (residues α257–α263 of ttPheRS; TyrPhePro-PheValGluPro) of α subunit together with the metal-binding loop (residues β450–β465 of ttPheRS; ArgLeuAspLeuArgLeuGluGluAspLeuValGluGluValAspArg) of β subunit are evolutionarily retained among different species and form metal-mediated conserved αβ subunit interface. Conservation of these residues among different PheRS across species implies the importance of the metal-mediated αβ subunit interface in enzymatic activity, especially in the first step of the aminoacylation reaction.32 Crystal structure of ttPheRS/tRNAcomplex suggests that the C-terminal domain of β subunit (B8 domain) plays a significant role in recognizing anticodon triplet of tRNA.24 Comparative

Figure 1 Schematic representation of localization and function of HcpPheRS and HmtPheRS.

Abbreviations: FARS, FARS encoding gene; FARS2, HmtPheRS encoding gene; HcpPheRS, human cytosolic phenylalanyl-tRNA synthetase; HmtPheRS, human mitochondrial phenylalanyl-tRNA synthetase; Phe, phenylalanine; Tyr, tyrosine.
studies of the ttPheRS/PheOH-AMP and ttPheRS/tRNA\(^{Phe}\) binary complexes and ttPheRS/PheOH-AMP/tRNA\(^{Phe}\) ternary complex reveal that the structural rearrangements of the enzyme induced by small substrates enable the correct positioning for charging of tRNA\(^{Phe}\). Conformational transition of ttPheRS from “closed” to “open” conformation is driven by the PheOH-AMP and the tRNA\(^{Phe}\) induced concerted movement of the motif 2 and helical loops that promote the second step of aminocyclation.\(^{33}\)

The tRNA\(^{Phe}\) binding pattern is different in two kingdoms, as the B8 domain involved in anticodon-binding is missing in the archaeal and the eukaryotic PheRS.\(^{34,35}\) Crystal structure of the human ttPheRS (HctPheRS) in complex with Phe reveals a novel structural module, containing three domains with helix-turn-helix (HTH) motif (DBD-1, DBD-2, and DBD-3) at the N-terminus of the \(\alpha\) subunit, whose functional role in recognition and binding of tRNA\(^{Phe}\) has been probed by truncation mutagenesis studies. Two mobile loops (motif 2 and helical loops of \(\alpha\) subunit) play a key role in enzymatic function of HctPheRS, similar to ttPheRS (Figure 2). Conformation of the motif 2 loop is found to be similar for both the enzymes. Similarly, the helical loop in HctPheRS containing a highly conserved sequence [xProxxHisProAlaArgAsp(Met/x)(Trp/Gln/His) AspThrPhe] is important for the proper positioning of CCA end of tRNA\(^{Phe}\).\(^{27}\) The tRNA\(^{Phe}\) makes contact with all four subunits of the enzyme. Two different \(\alpha\) subunits of PheRS are involved in the acceptor end and the anticodon stem binding of a tRNA\(^{Phe}\).\(^{27,32}\) Crystal structures clearly depict that the heterotetrameric PheRS is able to bind two molecules of tRNA\(^{Phe}\) simultaneously.\(^{36}\)

### Table 2 Structural organisation of eubacterial PheRS

<table>
<thead>
<tr>
<th>Subunit</th>
<th>Domain</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)</td>
<td>A0–A2</td>
<td>Catalytic site</td>
</tr>
<tr>
<td>(\beta)</td>
<td>B1</td>
<td>Helix-turn-helix (HTH) motif, interacts with DNA</td>
</tr>
<tr>
<td>(\beta)</td>
<td>B2</td>
<td>Secondary tRNA-binding site</td>
</tr>
<tr>
<td>(\beta)</td>
<td>B3–B4</td>
<td>Editing site</td>
</tr>
<tr>
<td>(\beta)</td>
<td>B5</td>
<td>Helix-turn-helix (HTH) motif, interacts with DNA</td>
</tr>
<tr>
<td>(\beta)</td>
<td>B6–B7</td>
<td>Involved in dimerization</td>
</tr>
<tr>
<td>(\beta)</td>
<td>B8</td>
<td>Interacts with the anticodon of tRNA(^{Phe})</td>
</tr>
</tbody>
</table>

Note: Data from Roy and Ibbas.\(^{35}\)

Abbreviation: PheRS, phenylalanyl-tRNA synthetase.

### Table 3 Domain organisation divergence between the kingdoms

<table>
<thead>
<tr>
<th>Kingdom</th>
<th>Subunit</th>
<th>Chain length</th>
<th>Sub-domain</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eubacteria</td>
<td>(\alpha)</td>
<td>350 aa</td>
<td>A0–A2</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>(\beta)</td>
<td>785 aa</td>
<td>B1–B8</td>
<td></td>
</tr>
<tr>
<td>Archaea</td>
<td>(\alpha)</td>
<td>501 aa</td>
<td>A1–A2</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>(\beta)</td>
<td>574 aa</td>
<td>B1, B3–B7</td>
<td></td>
</tr>
<tr>
<td>Eukarya (cytosolic)</td>
<td>(\alpha)</td>
<td>508 aa</td>
<td>A1–A2</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>(\beta)</td>
<td>589 aa</td>
<td>B1, B3–B7</td>
<td></td>
</tr>
</tbody>
</table>

Organelle-specific PheRS: a chimera

Eukaryotic cells contain many sub-cellular compartments designed for performing specialized functions, e.g., mitochondria, chloroplast, endoplasmic reticulum, golgi apparatus etc. Among them, the mitochondria and the chloroplast have their own genomes and encode a few proteins. The proteins responsible for organelle translation including aaRSs are encoded in the nucleus, synthesized in the cytosol and transported to the organelle.\(^{11,37}\) Due to the endosymbiotic origin of the mitochondria and the chloroplast, there is a notion that the prokaryotic aaRSs are the ancestor of the organelle-specific aaRSs.\(^{38,39}\) Whereas a wealth of literature exists on \(mta\)aaRSs, information about chloroplast-specific aaRSs is very limited.

The organelle-specific PheRS lacks the conventional heterotetrameric (\(\alpha_2\beta_2\)) organization. It is monomer with a minimal set of structural domains. It is basically a chimera of the catalytic domain of the \(\alpha\) subunit and B8 domain of the \(\beta\) subunit of bacterial PheRS. Sequence alignment of the mitochondrial PheRS (\(mt\)PheRS) family with prokaryotic PheRS shows that there is an insertion (35–70 amino acid residues) following motif 2. However, the function of this insertion is unknown.\(^{40,41}\)

Like other \(mt\)aaRSs, \(mt\)PheRS is also encoded in the nucleus. It is translated in the cytoplasm as a pre-protein containing mitochondrial targeted sequence. The pre-protein is imported in the mitochondria through translocase of the inner membrane–translocase of the outer membrane complexes.\(^{42,43}\) Once inside the mitochondria, the mitochondrial targeted sequence is processed by the mitochondrial processing peptidase and mitochondrial intermediate peptidase.\(^{44,45}\) Postprocessing, \(mt\)PheRS must refold correctly within the organelle matrix to ensure proper enzymatic function. Defects in refolding can lead to mitochondrial dysfunction, either by triggering an unfolded protein response or by lowering the levels of stable folded active protein.\(^{46}\)

### HmtPheRS: conformational plasticity for enzymatic function

Human \(mt\)PheRS (HmtPheRS) is the smallest known enzyme among the aaRSs family and by far the most extensively characterized both structurally and functionally among the \(mt\)PheRS family. Crystal structures of HmtPheRS, complexed with Phe-AMP (Protein Data Bank accession number 3CMQ,
at 2.2 Å resolution and in complex with both tRNA^{Phe} and Phe (Protein Data Bank accession number 3TUP, at a resolution of 3.0 Å) have been solved (Figure 3).\(^47\)\(^,\)\(^48\) The mature *Hmt*PheRS consists of four major parts: the N-terminal region (residues 1–47), the catalytic domain (residues 48–289), the linker region (residues 290–322), and the C-terminal domain (residues 323–415). Crystal structure of *Hmt*PheRS complexed with Phe-AMP reveals a “closed” conformation where the ABD located near the C-terminus overlaps with the acceptor stem of tRNA^{Phe}, if the substrate is fitted in a manner similar to that observed in the bacterial PheRS-tRNA^{Phe} complex. Hence, ABD must undergo a ∼160° hinge-type rotation for correct positioning of tRNA^{Phe}.\(^47\) *Hmt*PheRS therefore exists in two conformations; the “closed” or non-active conformation where it cannot bind tRNA^{Phe} and the “open” or active conformation where it can bind tRNA^{Phe}. This conformational plasticity of *Hmt*PheRS may be attributed to the presence of an extended flexible linker region.\(^48\)\(^,\)\(^49\)

**Plasmodium falciparum mtPheRS: an exception**

A recent study has reported the presence of *mt*PheRS in malaria parasite, *P. falciparum*. Bioinformatic study reveals that the gene for *mt*PheRS is exclusive to malaria parasite among the apicomplexa phyla. Apicoplast PheRS (*aPheRS*) is a ∼67 kDa protein with a signal and transit peptide for apicoplast targeting. *mt*PheRS is the smallest of the three PheRSs present in *Plasmodia* with a molecular mass of ∼53 kDa. Both the organelle specific PheRSs contain a conserved aminocytacylation domain at the N-terminus and an ABD at the C-terminus, however, the editing domain is absent. *mt*PheRS of *P. falciparum* also contains an insertion of ∼45 residues in the aminocytacylation domain, as reported for other *mt*PheRSs. *P. falciparum* contains two copies of tRNA^{Phe}, one is encoded in the nucleus and the other in the apicoplast. On the contrary, *Plasmodia* mitochondrial genome lacks any copy of tRNA genes. Thus, it is presumed that...
like some other organisms, such as \textit{Leishmania}, \textit{Trypanosoma}, and \textit{Toxoplasma} sp., mitochondrial tRNA import is also active in \textit{P. falciparum}. Indeed, it has been shown that tRNA\textsubscript{Phe} is imported into mitochondria from the cytoplasm. Although, the charging status of imported tRNA\textsubscript{Phe} is uncertain, \textit{m}tPheRS is able to charge cytoplasmic tRNA\textsubscript{Phe} in vitro.

**Editing activity: a canonical function of PheRS**

Accurate translation of the genetic code is a high fidelity process. AaRSs play a crucial role in the quality control of mRNA decoding and maintaining the overall error rate of translation $\sim 10^{-4}$.

Two major checkpoints for maintaining the translation fidelity are: 1) aminoacylation of tRNA with the cognate amino acid by aaRS and 2) selection of aa-tRNA by the ribosome. aa-tRNA is delivered to ribosome by the elongation factor (EF) (EF thermo unstable [EF-Tu] in bacteria and EF-1\textalpha{} in archaea and eukaryotes). In case of class II aaRSs, once the misacylated tRNA escapes the editing step, EF-Tu binds it and delivers it to the ribosome and thus the noncognate amino acid is incorporated into the growing polypeptide chain. EF-Tu competes with the editing domain for aa-tRNA binding. Therefore, in class II aaRSs, editing activity is the sole quality control checkpoint that determines translational fidelity. A “double-sieve” model for aaRSs has been proposed to explain the high fidelity of aminoacylation process. According to this hypothesis, the first sieve is the amino acid recognition by the active site of the aaRSs, which acts as a coarse sieve and excludes most of the noncognate amino acids that are larger than the cognate one. The editing site serves as a second fine sieve, which selectively hydrolyzes structurally similar, smaller near-cognate amino acids based on size and chemical character. While the first sieve is functional in all aaRSs, only approximately half of the aaRSs are found to harbor the second sieve. The editing site hydrolyzes the near-cognate amino acid either by hydrolyzing misacylated aminoacyl-adenylate – known as pre-transfer editing or the mischarged aa-tRNA – known as post-transfer editing. Inaccuracies in aminoacylation activity resulting from charging non-cognate tRNAs are less frequent, as selection of cognate tRNA is guided by the presence of tRNA identity elements.

Structural analyses have revealed that the editing domains of class I aaRSs share structural homology. A unique insertion domain, called connective peptide 1 (CP1) domain is responsible for the editing activities in IleRS, ValRS, LeuRS via both pre- and post-transfer editing. In addition to CP1, LeuRS is found to harbor a second insertion domain, called CP2, which is involved in both aminoacylation and editing activities. The editing mechanism of MetRS of class I, however, differs from the other class I aaRS members. MetRS proofreads the non-cognate homocysteine via intramolecular cyclization by the aminoacylation active site. On the contrary, editing domains of class II aaRSs are structurally diverse. For example, ThrRS has an N-terminal N2 domain dedicated for editing activity, whereas, unlike other class II aaRSs, SerRS does not contain any distinct editing domain. The heterotetrameric PheRS from two kingdoms (eubacterial and archaeal/eukaryotic) possess an editing domain, located at the B3/B4 domain of $\beta$ subunit, sharing no structural resemblance with any known editing domains. The monomeric organelle-specific counterpart, however, lacks any editing domain. While Phe is the cognate amino acid for PheRS, it can mischarge tRNA\textsubscript{Phe} with isosteric tyrosine (Tyr) and a non-cognate product, tyrosyl-tRNA\textsubscript{Phe} (Tyr-tRNA\textsubscript{Phe}), is synthesized at low frequency.

**Critical residues for editing activity of PheRS**

There is an ambiguity over the predominant mode of PheRS editing activity. Both pre-transfer and post-transfer editing have been suggested to be operative in yeast PheRS. However, pre-transfer editing triggered by native tRNA\textsubscript{Phe} has been shown to predominate in yeast ctPheRS. In contrast, eubacterial PheRS hydrolyzes near-cognate Tyr at the post-transfer stage. Conservation mapping has suggested that the editing site is $\sim 40$ Å away from the synthetic site in PheRS and it has been suggested that cis editing is complemented by trans editing for efficient hydrolysis of the mischarged Tyr-tRNA\textsubscript{Phe} that dissociates from the enzyme. $\alpha$Ala294 in motif 3 of $c$ctPheRS, present in amino acid binding pocket of the enzyme, has been found to be the responsible residue for amino acid specificity. The mutant $\alpha$Ala294Gly has been shown to have relaxed substrate specificity for para-substituted Phe analogs. Not only $\alpha$Ala294Gly mutant of $c$ctPheRS, but wild-type $\epsilon$ctPheRS is also able to activate $p$-chloro-$\lambda$-Phe ($p$-Cl-Phe). The editing pocket recognizes near-cognate Tyr, not by size, but based on specific interactions with side chain hydroxyl groups. In eubacterial PheRS editing site, amino acid discrimination is achieved by two highly conserved motifs, “GlyValMetGlyGlyxxSer/Thr” and “GlnProxHisx” and highly conserved amino acid residues, $\beta$Arg244, $\beta$Phe263, $\beta$Gly315, $\beta$Glu334, and $\beta$Ala356. Mutagenesis studies supplemented with biochemical assays establish that all these residues are important for editing.
Phylogenetic analysis of the editing domain shows very low sequence similarity between the two groups, eubacterial and archaeal/eukaryotic. However, there is some resemblance in the architecture of the editing site between these two groups. In case of P. horikoshii PheRS (pHtPheRS), βAsp234 residue, corresponding to βGlu334 of ntPheRS, has been shown to be important for preferentially selecting Phe over Tyr. Replacement of this residue with Ala allows Phe-tRNA^Phe hydrolysis accompanied by a decrease in Tyr-tRNA^Phe hydrolysis activity. Whereas, βHis261 (ntPheRS) residue involved in hydrolysis of Tyr, is found to be replaced by βAsn217 in pHtPheRS. 35 In HctPheRS, βGlu254, and βAsn238 corresponding to βGlu334 and βHis261 in ntPheRS respectively, play a crucial role in recognition and editing of the Tyr moiety.27 The corresponding residues, which are crucial for editing in E. coli are βGlu334 and βHis265 and in yeast are Asn243 and His158.55 The B2 domain has also been shown to play a key role in post-transfer editing. Mutagenesis studies have suggested that B2 domain may play a role in translocating aa-tRNA from synthetic site to editing site.25

**Organelle-specific PheRS lacks editing activity**

Organelle-specific counterparts of PheRS do not contain any domain homologous to B3/B4 domain involved in editing and are devoid of proofreading activity.65 In mammalian mitochondria only 13 proteins are encoded, whereas in plants more than 150 proteins are encoded within the organelles. These editing deficient organelar PheRSs may lead to mistranslation, giving rise to the accumulation of non-native proteins and compromise organelar homeostasis. Some post-translational check-point must operate to maintain the protein populations in functional forms within the organelles. It has been reported that a conserved protein degradation machinery operates as a major contributor of the protein quality control systems in the organelle. The key role of the proteases that are the prime components of the protein degradation machinery is rapid and selective proteolysis of misfolded, non-native proteins to ensure organelles’ viability.59,70

**Error-prone editing of PheRS under oxidative stress**

Reactive oxygen species (ROS) are known to modify the naturally occurring amino acids within the cell and the incorporation of these damaged amino acids into proteins may have pathological consequences in age related disorders, such as Alzheimer’s, Parkinson’s, atherosclerosis, and cataractogenesis.71–73 Exposure of Phe to ROS generates multiple isomers of Tyr: m-Tyr, α-Tyr, and the standard p-Tyr.74 On the other hand, the exposure of Tyr to ROS or catalytic conversion of Tyr by tyrosinase enzyme generates 3,4-dihydroxy-L-Phe, known as levodopa or L-DOPA.75 L-DOPA is the immediate biosynthetic precursor of the neurotransmitter dopamine and is used to treat Parkinson’s disease and dopamine-responsive dystonia.76,77 Incorporation of oxidized amino acids in proteins may lead to protein aggregation. Proteins bearing oxidized residues evade recognition by protein degradation machinery leading to accumulation of noncanonical proteins inside the cell.73 It has been shown that both HctPheRS and HmtPheRS catalyze attachment of m-Tyr to tRNA^Phe. The affinity of HctPheRS toward m-Tyr is one magnitude lower than that of HmtPheRS and it discriminates more efficiently between the cognate Phe and non-cognate m-Tyr.74 Both the enzymes, however, cannot discriminate L-DOPA as efficiently as Tyr and thereby mischarge tRNA^Phe. Unlike HmtPheRS, which lacks any editing activity, HctPheRS is able to proofread these ROS modified analogs of Phe, but in much lower level than the standard substrate Tyr. Kinetic experimental data show that specificity constants for L-DOPA over Phe for both HctPheRS and HmtPheRS are much higher than the estimated translational error rate (<10⁻⁸).79,80 The overall error rate of aa-tRNA synthesis is determined by two factors; aaRSs’ intrinsic proofreading activity and the ratio of cognate/non-cognate amino acid pool inside the cell.81 L-DOPA is the most widely used medicine for Parkinson’s disease. Increased availability of L-DOPA may lead to misincorporation of L-DOPA into proteins. L-DOPA bound proteins are majorly found in age-related disorders, such as atherosclerosis and cataract. L-DOPA can be further modified to dopaquinone, which crosslinks proteins forming covalent links with other biomolecules.75 These modified proteins form stable aggregates that are resistant to proteolysis and extremely cytotoxic causing extensive damage to other biological molecules.84 Thus, protein-bound L-DOPA may play a critical role in age-related disease progression by contributing to lysosomal membrane permeabilization-initiated apoptosis.85 Indeed, it has been found that the levels of protein-incorporated L-DOPA have increased in cells of L-DOPA treated patients.86 Moor et al have demonstrated that PheRS has the potential to misacylate tRNA^Phe with L-DOPA and incorporate it into proteins instead of Phe.79 Increased levels of L-DOPA bound proteins in the cells of L-DOPA treated patients may be attributed to the inability of proofreading activity of HmtPheRS.80 Thus, long-term usage of L-DOPA may contribute to the toxic effects associated with L-DOPA treatment, further complicating the disease scenario.84,86
Noncanonical function of PheRS

It is postulated that aaRSs have acquired a number of non-canonical functions during evolution, while preserving their canonical aminocacylation activity. Functional expansion may occur either by the introduction of new structural domains to the existing enzyme or through transformation of idle domains into ones with new functions without disrupting the enzymatic activity. \(t\)PheRS has been demonstrated to bind double-stranded DNA. A detailed work has shown that the B5 domain of \(\beta\) subunit containing HTH motif is involved in DNA binding. Specifically, amino acid residues Glu438-Gly439-Pro440-Thr441 and Pro447 are found to be responsible for DNA binding as demonstrated by mutagenesis studies. Although, the exact function is not understood, PheRS may contribute to some other biological functions, apart from protein translation, via DNA binding. \(t\)PheRS also shares significant structural homology (a HTH motif and a Src-homology 3 fold) with \(E.\ coli\) biotin synthetase/repressor (BirA) protein, which acts as a transcriptional regulator.\(^{99,90}\) However, no studies have been carried out till now to elucidate the functional significance of these conserved motifs in PheRS.

PheRS as a tool in protein engineering

Selective incorporation of unnatural amino acids into proteins has become a useful tool for systematic engineering of proteins. This can be achieved either by site-specific method or by residue-specific method.\(^{91,92}\) The site-specific method exploits an orthogonal tRNA-aaRS pair that operates independently of the endogenous pair of synthetase-tRNA of host cell for introducing a particular unnatural amino acid at a specific site of a protein by decoding a reassigned nonsense codon in the gene of interest.\(^{93}\) The \textit{amber} codon (UAG) is most frequently exploited among the three stop codons (\textit{amber}–UAG, \textit{ochre}–UAA, and \textit{opal}–UGA), because it is least utilized in \(E.\ coli\) and most suppressor tRNA can efficiently recognize it.\(^{94}\) Residue-specific method is comparatively less complicated in a way that it requires only engineering of aaRSs to expand the specificity of their amino acid recognition site. In this way, several analogs of cognate amino acid can be incorporated into a protein. Both the methods have been manipulated to introduce various analogs of Phe into protein. For in vitro site-specific incorporation of unnatural amino acid, \(E.\ coli\) cell-free translational system has been used containing a gene with a nonsense codon mutation and precylated suppressor tRNA. Preacylation of suppressor tRNA with \(p\)-iodo-L-Phe (\(p\)-I-Phe) is carried out with \(\alpha\)Ala294Gly mutant of ecPheRS that has relaxed substrate specificity.\(^{95}\) In vivo incorporation of \(p\)-bromo-L-Phe (\(p\)-Br-Phe) has been carried out using yeast suppressor tRNA (tRNA\(_{\text{Phe}}^{\text{ala}}\)) and \(\alpha\)Thr415Gly mutant of yeast PheRS in \(E.\ coli\). \(\alpha\)Ala294 mutant of ecPheRS has already been shown to have relaxed substrate specificity and is able to incorporate \(p\)-Cl-Phe into recombinant protein.\(^{68}\) A computer simulation study corroborates the finding that \(\alpha\)Thr251 and \(\alpha\)Ala294 of \(E.\ coli\) (corresponding to \(\alpha\)Val261 and \(\alpha\)Ala314 in \(t\)PheRS) are crucial for amino acid specificity and replacement of both the amino acids with Gly enables efficient incorporation of \(p\)-acetyl-L-Phe in vivo replacing Phe.\(^{96}\) Sequence alignment shows \(\alpha\)Thr251 in \(ec\)PheRS is equivalent to \(\alpha\)Thr415 in yeast PheRS. A high-throughput screening method has been developed to identify yeast PheRS variant from a library with high substrate-specificity for a Phe analog. The approach is to obtain yeast PheRS variants from the yeast PheRS library, which leads to loss of fluorescence of green fluorescence protein upon efficient incorporation of \(2\)-L-naphthylalanine, a Phe analog, into multiple sites of green fluorescence protein by yeast PheRS, exploiting the degeneracy of the Phe codons. A yeast PheRS variant (\(y\)PheRS\(_{\text{naph}}\): containing Asn412Gly, Thr415Gly, Ser418Cys, and Ser437Phe mutations) has been identified, which shows 60-fold increased selectivity toward \(2\)-L-naphthylalanine vs tryptophan compared to \(\alpha\)Thr415Gly variant of yeast PheRS.\(^{97}\)

In contrast to site-specific method that requires multiple gene manipulation, residue-specific method is a much simpler way for unnatural amino acid incorporation into protein. This can be achieved by enlarging the amino acid binding pocket of aaRSs by point mutations, which have relaxed substrate specificity toward structurally similar amino acid analogs.\(^{98}\) This strategy has been used in eukaryotic organism yeast as well as in mammalian cells.\(^{100}\) This method also has been utilized to incorporate Phe analogs, replacing Phe, in domesticated silkworm, \textit{Bombyx mori}. Phe is an essential amino acid for \textit{B. mori}\(^{101}\) and is incorporated in the major protein component of \textit{B. mori} silk in a periodic manner, making it an appropriate target for unnatural amino acid incorporation. Silk proteins produced by \textit{B. mori} have a wide range of industrial applications\(^{105-105}\) and this strategy will improve its application further. \(\alpha\)Ala450Gly mutant of \(B.\ mori\) PheRS (\(Bm\)PheRS) has been generated and incorporation of Phe analogs, \(p\)-Cl- and \(p\)-Br-Phe, into proteins by this mutant are checked both in vitro and in vivo (in \(BmN\), a \(B.\ mori\) cell line). Although, in vitro assay shows aminocacylation of \(B.\ mori\) tRNA\(_{\text{Phe}}\) with both \(p\)-Cl- and \(p\)-Br-Phe, incorporation of only \(p\)-Cl-Phe is observed in vivo.\(^{106,107}\) Another point mutation, replacing \(\alpha\)Thr407 (corresponding to \(\alpha\)Thr415 residue of yeast
PhRS) with Ala or Gly, has been introduced into BmPhRS to relax the amino acid specificity further and to increase the incorporation efficiency of unnatural amino acid into protein. Aminocacylation capabilities of αAla450Gly, αThr407Ala and αThr407Gly BmPhRS mutants are checked both in vitro and in vivo in BmN using six para-substituted Phe analogs. The data show that αThr407Ala mutant is able to incorporate p-Br-Phe in proteins in vivo. Further, p-Br-Phe, p-cyano-Phe, p-azido-Phe, and p-I-Phe can be incorporated in proteins in vivo by lowering the concentration of competing amino acid Phe in culture media.¹⁰⁸ This strategy has been extended to generate transgenic B. mori expressing the αAla450Gly mutant in silk glands. p-Cl-, p-Br-, and p-azido-substituted analogs are efficiently incorporated into silk fibroin by adding them to the diet and decreasing the Phe content.¹⁰⁹

**PhRS as a therapeutic target**

The dramatic increase of multidrug resistance in pathogens against existing drugs drives researchers to search for novel targets that can be exploited to develop novel therapeutics.¹⁷,¹¹⁰ Several reasons have made aaRS an attractive target for the development of antimicrobial agents, i) they are indispensable for cellular function, ii) abundant genetic and protein structural information available, iii) due to their conserved nature, unlikely to undergo mutations and develop resistance upon being targeted, iv) sequence and structural divergence among the kingdoms facilitate easy targeting of pathogenic aaRSs with high selectivity and hence reduced toxicity.¹¹¹,¹¹² Although, several antimicrobial agents against aaRSs have been reported and patented so far, mupirocin (Bactroban) is the only commercially available antibacterial agent, used against methicillin-resistant *Staphylococcus aureus*. It selectively inactivates eubacterial IleRS.¹¹³ The structural differences between eukaryotic and prokaryotic PhRS lead to search for novel PhRS inhibitors with broad spectrum efficacy and high selectivity. A series of compounds have been isolated, namely, spirocyclic furans and pyrroli, structural differences between eukaryotic and prokaryotic aaRSs have been reported so far, mupirocin (Bactroban) is the only commercially available antibacterial agent, used against methicillin-resistant *Staphylococcus aureus* (Bactroban) is the only commercially available antibacterial agent, used against methicillin-resistant *Staphylococcus aureus*. It selectively inactivates eubacterial IleRS.¹¹³ The structural differences between eukaryotic and prokaryotic PhRS lead to search for novel PhRS inhibitors with broad spectrum efficacy and high selectivity. A series of compounds have been isolated, namely, spirocyclic furans and pyrrolidines, ethanalamines, and benzyl phenyl ethers, but with nominal efficacy.¹¹⁴–¹¹⁷

A novel class of PhRS inhibitor, phenyl-thiazolyurea-sulfonamides, has been reported to be a potent antibiotic against both gram-positive and gram-negative bacteria, with IC₅₀ in the nanomolar range. These compounds competitively inhibit the aminocacylation of tRNA³⁵th with respect to the cognate substrate Phe. The advantage of these compounds is that they do not show any cross-reactivity to the mammalian enzymes, both cytoplasmic and mitochondrial.¹¹⁸ However, the compounds also have some drawbacks, eg, low aqueous solubility, high level of plasma protein binding, and Phe complementation reduces the efficacy of the compounds significantly. To overcome these weaknesses, a high throughput screening has been performed to isolate novel chemical scaffolds with improved efficacy. Finally, three novel chemical scaffolds (pyridinyl anilines, trifluoromethyl pyrazoles, and thiazole amides) have been identified and crystal structures of *Pseudomonas aeruginosa* PhRS have been determined in complex with phenyl-thiazolyurea-sulfonamides as well as these chemical scaffolds to identify the binding mode of the inhibitor. In addition to Phe binding pocket, an auxiliary hydrophobic pocket, just below the substrate binding pocket has been identified, where the inhibitors can bind but Phe cannot. Subsequent studies have identified three phenyl-thiazolyurea-sulfonamide-resistant mutations in *E. coli*, which suggest that the auxiliary pocket should be avoided for drug designing against PhRS in the future.¹¹⁹

**PhRS in pathophysiology**

**Human disorders associated with PhRS mutations**

**Human disorders associated with PhRS mutations**

Human cells harbor two distinct sets of aaRSs encoded by two specific sets of genes; cytoplasmic and mitochondrial.¹²⁰ Recent studies have revealed the association of aaRSs mutations with human pathologies.¹⁰,¹⁹,¹¹² Many disease-associated mutations have been reported for *ctaaRSs*, but none involving *HetPhRS* to date. On the contrary, so far five disease-causing mutations have been identified within FARS2 gene encoding *HetPhRS* (Figure 4 and Table 4).¹²²–¹²⁵ FARS2 gene is located in chromosome 6; containing 7 exons, of which 2–7 are protein coding.¹²² The effects of the disease-associated mutations on aminoacylation activity and enzyme stability have been analyzed in *HmtPhRS*. Structural modeling has shown that the mutation I329T is located near the ATP binding site of the catalytic domain, impairing the ATP binding capacity of the mutant protein. The D391V mutation has been found to be located at the interface of the catalytic domain and ABD, impairing Phe binding affinity. Y144C mutation is also located at the interface and found to impair tRNA³⁵th binding.¹²² D325Y mutation is located near the ATP binding site of catalytic domain and found
to impair ATP binding and consequently aminoacylation activity of the enzyme. A previous report shows that mutating amino acid residues (S57C and N280S), located far from the catalytic or tRNA binding site, do not alter aminoacylation activity but greatly affect refolding ability and protein stability. This suggests that improper folding and reduced stability may also contribute to diseases through mutations; even if they do not directly compromise enzymatic activity.

Autoantibody against PheRS: antisynthetase syndrome

Antisynthetase syndrome is a rare, chronic autoimmune disease characterized by myositis, arthralgia, Raynaud phenomenon, skin changes called mechanic’s hands and interstitial lung disease. The hallmark of the disease is the presence of serum autoantibodies against the aaRSs. To date, a total of eight antisynthetase autoantibodies have been reported, including autoantibody against PheRS (Table 5). A combination of immunoprecipitation and mass spectrometric data have identified a novel pattern with two bands corresponding to 60 and 70 kDa, matching to PheRS α and β chain, respectively.

Outlook

The crystal structures of PheRS, both with and without ligand, have been solved from all three kingdoms of life. This is an unprecedented advantage for several applications as observed in some recent studies. All these following observations have proved that PheRSs indeed snugly fit into future emerging field of research, particularly in the area of synthetic biology, proteomics, genetics, a model system to understand the mechanism for aaRS related diseases, to name a few. All of these applications can be attributed to our knowledge of PheRS structure. The approach of unnatural amino acid incorporation during protein synthesis has emerged as a probe for studying protein structure and function as well as enabled designing new proteins with novel applications. Recently, one group has demonstrated that yeast PheRS can be manipulated to incorporate Phe analog Nal at several specific sites of a protein very efficiently. Caenorhabditis elegans PheRS has been engineered to achieve cell-selective bioorthogonal noncanonical amino acid tagging (cell-selective BONCAT) of low-abundance proteins. Coupling this method with stable-isotope labeling of amino acids in cell culture enables identification of expression patterns of proteins in targeted subsets of cells. This study provides first-hand information that proteomic analysis of rare cells is possible that otherwise pose difficulty for enriching, identifying and quantifying proteins in complex biological systems. Two recently published reports state that PheRS can be used as a counter selection marker in genome recombineering, genome mutagenesis, and plasmid curing experiments. Mayazaki develops two highly efficient variants (Thr251Ala/Ala294Gly and Thr251Ser/Ala294Gly) of E. coli PheRS that show improved incorporation efficiency of a toxic analog 4-chloro-phenylalanine over the single mutant (Ala294Gly) leading to the high lethality in common bacterial growth media. These two variants can be used as counter selection marker with high selection efficiency.

Another group has developed a genetic model of T. thermophilus that can potentially be used for genome engineering of this microorganism. Interestingly, this model has been generated by chromosomal manipulation of PheRS gene that is sensitive to 4-chloro-phenylalanine. This conditionally lethal mutant has been used as a counter selection marker for introducing genomic point mutations or deletions in T. thermophilus. To date, the mechanism of aaRS mutation

### Table 4 Disease-related mutations of HmtPheRS

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Exon</th>
<th>Amino acid change</th>
<th>Domain</th>
<th>Disease</th>
<th>Ethnicity</th>
<th>Amino acid conservation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.431A&gt;G</td>
<td>3</td>
<td>Y144C</td>
<td>CD</td>
<td>Developmental delay, myoclonus seizures</td>
<td>Saudi Arabian</td>
<td>Conserved</td>
<td>122,123</td>
</tr>
<tr>
<td>c.986T&gt;C</td>
<td>5</td>
<td>I329T</td>
<td>CD</td>
<td>Alpers encephalopathy</td>
<td>Finnish</td>
<td>Conserved</td>
<td>122</td>
</tr>
<tr>
<td>c.1172A&gt;T</td>
<td>6</td>
<td>D391V</td>
<td>ABD</td>
<td>Alpers encephalopathy</td>
<td>Finnish</td>
<td>Conserved</td>
<td>122</td>
</tr>
<tr>
<td>c.973G&gt;T</td>
<td>5</td>
<td>D325Y</td>
<td>CD</td>
<td>Infantile-onset epilepsy</td>
<td>British</td>
<td>Conserved</td>
<td>124</td>
</tr>
<tr>
<td>c.1255C&gt;T</td>
<td>7</td>
<td>R419C</td>
<td>ABD</td>
<td>Global developmental delay, dysarthria and tremor</td>
<td>Mixed European</td>
<td>Not conserved</td>
<td>125</td>
</tr>
</tbody>
</table>

**Abbreviations:** ABD, anticodon binding domain; CD, catalytic domain; HmtPheRS, human mitochondrial phenylalanyl-tRNA synthetase.

### Table 5 Antisynthetase antibodies

<table>
<thead>
<tr>
<th>aaRS (antigen)</th>
<th>Class</th>
<th>Autoantibody</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenylalanyl-tRNA synthetase</td>
<td>II</td>
<td>Zo</td>
<td>127</td>
</tr>
<tr>
<td>Histidyl-tRNA synthetase</td>
<td>II</td>
<td>Jo-1</td>
<td>128</td>
</tr>
<tr>
<td>Threonyl-tRNA synthetase</td>
<td>II</td>
<td>PL-7</td>
<td>129</td>
</tr>
<tr>
<td>Alanyl-tRNA synthetase</td>
<td>II</td>
<td>PL-12</td>
<td>130</td>
</tr>
<tr>
<td>Isoleucyl-tRNA synthetase</td>
<td>I</td>
<td>Oj</td>
<td>131</td>
</tr>
<tr>
<td>Glycyl-tRNA synthetase</td>
<td>II</td>
<td>Ej</td>
<td>131</td>
</tr>
<tr>
<td>Asparaginyl-tRNA synthetase</td>
<td>II</td>
<td>KS</td>
<td>132</td>
</tr>
<tr>
<td>Tyrosyl-tRNA synthetase</td>
<td>I</td>
<td>Ha</td>
<td>133</td>
</tr>
</tbody>
</table>
related genetic disorders is poorly understood. Lu et al have recently developed a double-sieving-defective model of PheRS in *Drosophila melanogaster* and demonstrated its usefulness in understanding the underlying mechanism of aaRS related genetic disorders.137

**Acknowledgments**

SC is supported by a fellowship from the University Grant Commission (UGC), Government of India, and is currently working as a senior research fellow. This work was supported by funds from the Department of Science and Technology (India) to RB (SR/SO/BB- 0018/2012).

**Disclosure**

The authors report no conflicts of interest in this work.

**References**

5. Roy H, Ibba M. Phenylalanyl-tRNA synthetase contains a dispensable RNA-binding domain that contributes to the editing of noncognate aminoacyl-tRNA. *Biochemistry*. 2006;45(30):9156–9162.
25. Roy H, Ibba M. Phenylalanyl-tRNA synthetase contains a dispensable RNA-binding domain that contributes to the editing of noncognate aminoacyl-tRNA. *Biochemistry*. 2006;45(30):9156–9162.
83. Chakraborty and Banerjee


