Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase

Carla R. Polycarpo\textsuperscript{b,1}, Stephanie Herring\textsuperscript{b}, Amélie Bérube\textsuperscript{a,2}, John L. Wood\textsuperscript{a}, Dieter Söll\textsuperscript{a,b}, Alexandre Ambrogelly\textsuperscript{a,*}

\textsuperscript{a} Department of Chemistry, Yale University, New Haven, CT 06520-8114, USA
\textsuperscript{b} Department of Molecular Biophysics and Biochemistry, Yale University, P.O. Box 208114, 266 Whitney Avenue, New Haven, CT 06520-8114, USA

Received 24 October 2006; accepted 8 November 2006
Available online 20 November 2006

Edited by Lev Kisselev

Abstract In certain methanogenic archaea a new amino acid, pyrrolysine (Pyl), is inserted at in-frame UAG codons in the mRNAs of some methyltransferases. Pyl is directly acylated onto a suppressor tRNA\textsuperscript{Pyl} by pyrrolysyl-tRNA synthetase (PylRS). Due to the lack of a readily available Pyl source, we looked for structural analogues that could be aminoacylated by PylRS on tRNA\textsuperscript{Pyl}. We report here in vitro aminoacylation of tRNA\textsuperscript{Pyl} by PylRS with two Pyl analogues: N-\textalpha{\textprime}{\textprime}-\textalpha{-\textprime}{\textprime}-D-prolyl-\textalpha{-\textprime}{\textprime}-lysine (D-prolyl-lysine) and N-\textalpha{\textprime}{\textprime}-cyclopenterycloxy carbonyl-\textalpha{-\textprime}{\textprime}-lysine (Cyc). Escherichia coli, transformed with the tRNA Pyl and PylRS genes, suppressed a lacZ amber mutant dependent on the presence of N-prolyl-lysine or Cyc in the medium, implying that the E. coli translation machinery is able to use Cyc-tRNA\textsuperscript{Pyl} and N-prolyl-lysine-tRNA\textsuperscript{Pyl} as substrates during protein synthesis. Furthermore, the formation of active \beta-galactosidase showed that a specialized mRNA motif is not essential for stop-codon recoding, unlike for selenocysteine incorporation.

\section*{1. Introduction}

In the \textit{Methanosarcinaceae}, a family of methanogenic archaea, the new amino acid pyrrolysine (Pyl) was shown to be present in the mono-, di- and tri-methylamine methyltransferases (MtmB, MtB and MttB) [1,2]. This new amino acid is co-translationally inserted in response to an in-frame UAG codon located in the corresponding mRNAs [3]. The insertion of this amino acid relies on the presence of a specific suppressor tRNA (tRNA\textsuperscript{Pyl}) [4] and the new class II aminoacyl-tRNA synthetase, pyrrolysyl-tRNA synthetase (PylRS) specific only for its substrates tRNA\textsuperscript{Pyl} and Pyl [5,6]. The mechanism of Pyl-tRNAPyl insertion at UAG is unknown, however, it has been proposed that a specialized pyrrolysine insertion element (PYLIS) is present immediately downstream of the UAG which assists in the recoding event. Although putative secondary structures have been identified, their role in recoding in vivo remains uncertain [7,8].

Determination of the crystal structure of native \textit{Methanosarcina barkeri} MtmB allowed the elucidation of the molecular structure of this new amino acid. Analysis of the electron density showed that pyrrolysine is a dipeptide composed of a lysine modified at its \epsilon-N by a 4-methyl-pyroline-5-carboxylate (Fig. 1) [1]. The electron density of the MtmB crystal structure indicated an anti relationship between the C4-methyl group and the amide group, and it also suggested the presence of a double bond between the nitrogen and C2 in the pyrroline ring [1].

The detailed characterization of the novel synthetase PylRS will rely on the production of practical amounts of Pyl. Based on the first reported structure of Pyl, for which the nature of the C4 group was not ascertained, we initially undertook the synthesis of the Pyl derivative bearing a hydroxyl group at the C4 ring position. As the synthesis of this compound was unsuccessful, we then devised a new synthetic route leading to the more stable Pyl derivative bearing a methyl group at the C4 position (see \textit{supplemental data}). The product obtained (Fig. S1A, compound 8) allowed us to experimentally demonstrate the direct attachment of Pyl to tRNA\textsuperscript{Pyl} by PylRS [6]. Pyl produced by a different synthetic route [9], could be inserted in the \textit{M. barkeri} MtmB protein in an \textit{E. coli} context [5]. However, in our hands this synthesis did not yield the desired product.

Because these syntheses have a number of drawbacks, chief among them the complexity of the synthesis, low yields and high costs, we investigated the possibility of using simpler and readily available Pyl structural analogues that would still be substrates for PylRS and could be inserted into proteins in \textit{E. coli}.

\section*{2. Materials and methods}

\subsection*{2.1. General}

Oligonucleotide synthesis and DNA sequencing were performed by the Keck Foundation Biotechnology Resource Laboratory at Yale University. Uniformly labeled sodium \textsuperscript{32P} pyrophosphate [1–60 Ci/mmol (1 Ci = 37 GBq)], [\textsuperscript{32P}P] ATP (6000 Ci/mmol) and \textsuperscript{3H} serine (28 Ci/mmol) were from Amersham Biosciences. Mouse raised monoclonal anti-\textit{E. coli} \beta-galactosidase antibody was purchased from US Biological. The pyrrolysine analogues purchased from Sigma were L-lysine (Fig. 1, compound 1), N-\textalpha{-\textprime}-methyl-L-lysine (Fig. 1, compound 2), N-\textalpha{-\textprime}-formyl-L-lysine (Fig. 1, compound 3), N-\textalpha{-\textprime}-acetyl-L-lysine (Fig. 1, compound 4), and N-\textalpha{-\textprime}-cyclopenterycloxy carbonyl-L-lysine (Cyc) (Fig. 1, compound 5). The pyrrolysine analogues N-\epsilon-D-prolyl-L-lysine

Abbreviations: Pyl, pyrrolysine; PylRS, pyrrolysyl-tRNA synthetase

© 2006 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

doi:10.1016/j.febslet.2006.11.028
2.5. ATP-[32P] PPi exchange with pyrrolysine analogues

The reactions were performed at 37 °C, as described in [11], in 100 mM HEPES-NaOH (pH 7.2), 10 mM MgCl₂, 50 mM KCl, 5 mM DTT, 2 mM KF, 2 mM ATP, 2 mM [32P] PPi, (2 cpm/pmol), 200 nM of PylRS, and 500 μM of a pyrrolysine analogue (Fig. 1, compounds 1–8) in a final volume of 0.1 ml. [32P] ATP formation was followed by specific absorption on acid-washed Norit [0.2 ml of a 1% suspension (wt/vol) of Norit in a solution of 0.4 M sodium pyrophosphate solution containing 15% (vol/vol) perchloric acid], filtration on Whatman GF/C filters, and washing with 15 ml of water and 5 ml of ethanol. Cys and N-prolyl-lysine concentration were varied from 300 μM to 3 mM for K₅₀ determination.

2.4. Aminoacylation assays and acid ura gel electrophoresis of aminoaeryl-tRNA

Aminoacylation assays were performed for 90 min at 37 °C in 100 mM HEPES-NaOH (pH 7.2), 30 mM KCl, 15 mM MgCl₂, 5 mM ATP, 5 mM DTT, 2 μM tRNA₅₇₈ transcript, 1.6 μM PylRS, and 500 μM Cys, N-prolyl-lysine or PylE. All reactions were started by addition of the amino acid. The aminoacylation level was then visualized by acid ura gel electrophoresis and Northern blot as previously described [6].

2.5. Suppression of E. coli XAC/A24 lacI–lacZ nonsense mutation

E. coli strain XAC/A24 carries an inactivating mutation (Trp 220 to UAG nonsense) in the lacI–lacZ fusion system [12]. The truncated protein resulting from premature termination of protein synthesis at the in-frame stop codon is unable to degrade the chromogenic 2-nitrophenyl β-D-galactopyranoside (ONPG). E. coli strain XAC/A24 cells were co-transformed by electroporation with plasmids carrying the M. barkeri pylS (pCBS) and M. barkeri pylT (pTECH) genes. The same strain carrying an E. coli amber suppressor tRNA₅₇₈ on a pGF1B plasmid [13] was used as a positive control. Transformants were plated on Luria–Bertani (LB) agar with both ampicillin and chloramphenicol and incubated for 16 h at 37 °C. Colonies were transferred to 5 ml of LB liquid medium in the presence of antibiotics and cultured overnight at 37 °C. Fresh LB liquid medium was inoculated with the overnight culture in the presence of antibiotics. Strains carrying the pylS and pylT genes were further supplemented with either 10 mM Cys, 10 mM N-prolyl-lysine, or no amino acid for control. Cells were cultivated at 37 °C until A₆₀₀ reached 0.6. The UAG suppression level was determined by quantitative analysis of the β-galactosidase activity, which was performed according to the standard procedure [14]. Values are the average of triplicate measurements and are reported as the percentage of mutant enzyme activity relative to that of the wild type enzyme produced by the E. coli 1-Z40 strain, which carries the lacI–lacZ fusion with a wild type tryptophan codon in place of the UAG triplet.

2.6. Detection of LacI–LucZ fusion protein by immunoblot analysis

E. coli strain XAC/A24 carrying the plasmid copy of M. barkeri pylS and pylT genes was grown overnight in 5 ml of LB in the presence or absence of 10 mM Cys. The same strain carrying an E. coli amber suppressor tRNA₅₇₈ on a pGF1B plasmid [15] was used as a positive control. Cell free extracts were prepared, filtered on a Microcon 50 (Millipore), concentrated in a speed-vac, and analyzed on a 4–20% polyacrylamide/SDS gel (Biorad). The protein extracts were transferred onto PVDF membrane by using a Bio-Rad blotter. The membrane was incubated with different dilutions of the anti-E. coli β-galactosidase antibodies (1:100–1:1000) for optimal detection. For the immunoblot analysis, the colorimetric Opti–4CN substrate and detection kit (Bio-Rad) was used (horseradish peroxidase conjugate).

2.7. Suppression of E. coli FTP5822 trpA nonsense mutation

E. coli strain FTP5822 carries an inactivating mutation in the trpA gene (Gln 243 to UAG nonsense) [15] which results in tryptophan auxotrophy. Cells were co-transformed by electroporation with plasmids carrying M. barkeri pylS (pCBS) and M. barkeri pylT (pTECH) genes. Transformants were plated on LB agar with both ampicillin and chloramphenicol and incubated for 16 h at 37 °C. Colonies were transferred to 5 ml of LB liquid medium in the presence of antibiotics and cultured for 16 h at 37 °C. Subsequently, cells were washed with M9 minimal medium and used to inoculate 2 ml of M9 minimal medium containing antibiotics and 10 mM Cys or no amino acid as a control. E. coli strain FTP5822 transformed with pGF1B plasmid...
encoding E. coli amber suppressor tRNA \(^{L\text{ys}}\) and empty pTECH vector was used as a positive control. Cells were incubated at 37 °C. Bacterial growth was followed by measuring cell density over time using a Spectronic 20D+ spectrophotometer (Thermo). Values are the average of triplicate measurements.

Measurement of the tryptophan synthetase activity allowed a more quantitative measurement of the suppression efficiency than growth curves. Overnight cultures of E. coli FTP5822 in M9 minimal medium were used to inoculate 5 ml of M9 minimal medium, as before, in the presence or absence of 10 mM Cyc. All cultures were incubated at 37 °C to late log phase, harvested by centrifugation, washed twice with NaCl (0.9%), and resuspended in buffer A (0.05 M potassium phosphate (pH 7.0), 0.1 mg/ml of pyridoxal-5-phosphate, 10 mM 2-mercaptoethanol). Cell extract was prepared [16], micro-dialyzed against buffer A and used immediately for the assay. IGP was freshly prepared and expressed as percentage of mutant enzyme activity relative to that of the wild type TrpA in E. coli strain W3110.

3. Results and discussion

3.1. Pyl analogues are activated and ligated onto tRNA \(^{Pyl}\) by PyIRS in vitro

Pyrrolysine is essentially a modified lysine. The crystal structure of the M. barkeri MtmB enzyme has shown that this novel amino acid is the result of the coupling of a C-4 methyl pyrroline ring to the lysine ε primary amine via an amide bond [1,9]. While pyrrolysine is the natural substrate of PyIRS, lysine is not recognized by the enzyme [6]. In an attempt to obtain a suitable substrate for PyIRS biochemical characterization and gain insight on the contribution of the pyrroline ring to pyrrolyl synthetase recognition by PyIRS, we looked for commercially available or easily synthesizable lysine analogues that would progressively add components of the pyrroline ring (Fig. 1). We then measured the ability of these analogues to promote the PyIRS catalyzed ATP-[\(^{32}\)P] PP \(_e\) exchange reaction. N-ε-methyl-l-lysine, N-ε-formyl-l-lysine and N-ε-L-prolyl-l-lysine (L-prolyl-lysine) were not able to be activated by PyIRS (data not shown), and N-ε-acetyl-l-lysine allowed measurement only slightly above background (Fig. 2A). The N-ε-cyclopentyl-oxy carbonyl-l-lysine (Cyc) and N-ε-d-prolyl-l-lysine (D-prolyl-lysine) compounds were activated by PyIRS, although not as efficiently as the pyrrolysine enamine derivative N-ε-(3-methyl-2,3-dihydro-pyrroloxy carbonyl-l-lysine (PylE) (Fig. 2A). Determination of the \(K_M\) value in ATP-[\(^{32}\)P] PP \(_e\) exchange reaction for Cyc (\(K_M\) 670 \(\mu\)M) and d-prolyl-lysine (\(K_M\) 500 \(\mu\)M) showed a loss in apparent affinity for PyIRS of about 10-fold when compared to the published value for pyrroline (\(K_M\) 53 \(\mu\)M in [5]). In addition to demonstrating amino acid activation, the analogues Cyc, d-prolyl-lysine and PylE were successfully ligated onto tRNA \(^{Pyl}\) by PyIRS in vitro (Fig. 2B).

Interestingly, while d-prolyl-lysine was a substrate for PyIRS the analog L-prolyl-lysine was not, indicating the crucial importance of the pyrroline ring C-5 stereocenter for recognition by the synthetase. The importance of C-5 (R) conformation has been suggested before [6]. Our attempt at synthesizing pyrrolyl synthetise yielded a mixture of diastereoisomers in which the stereocentform of the carbon bearing the amide group was not fixed (see supplemental data). Upon analysis of the amino acid attached by the PyIRS to tRNA \(^{Pyl}\) in vitro, we saw that only one of the diastereoisomers was selected by the synthetase [6]. The pyrroline ring nitrogen does not make a strong contribution to PyIRS binding since its presence in d-prolyl-lysine only resulted in a small improvement in affinity over that of Cyc, but the incorrect positioning of the nitrogen in the PyIRS active site appears to be detrimental to amino acid activation. The molecular basis for the rejection of the L-prolyl-lysine isomer may become apparent when the structural analysis of the amino acid binding site of PyIRS becomes available [18]. The need to reject a Pyl isomer with an incorrect ring stereoconformation is essential as the UAG-coded position for Pyl is in the active site of the methyltransferase enzymes and presumably essential for enzyme activity [1].

One possible source of a pyrrolysine isomer with the incorrect conformation may be from the coupling of a member of the proline biosynthetic pathway to lysine, as proline and Pyl may share an intermediate metabolite with some structural homology. Although nothing is known on Pyl biosynthesis it has been hypothesized that Pyl might result from the coupling of the d-methyl-\(\Delta^1\)-pyrroline-5-carboxylate ring to lysine by an amino acid ligaase that remains to be identified [19]. Cells contain significant concentrations of L-\(\Delta^1\)-pyrroline-5-carboxylate, an intermediate metabolite in proline biosynthesis and close structural analogue to d-methyl-\(\Delta^1\)-pyrroline-5-carboxylate. The erroneous coupling of the proline precursor to lysine

![Fig. 2](image-url)
would then result in a Pyl isomer with an incorrect stereoconformation at the ring level, and consequently interfere with methyltransferase activity.

3.2. Insertion of Pyl analogues into proteins in E. coli

Since Cyc and d-prolyl-lysine were suitable substrates for PyrRS in vitro, we then tested whether they would be substrates in an in vivo context and whether they would be incorporated into proteins. Previous experiments have shown that when Pyl is added to the growth media of an E. coli strain containing plasmids bearing pylS, pylT, and mtmB genes from M. barkeri, Pyl can be incorporated into MtmB proteins at UAG, indicating that the bacterial translation machinery can use Pyl-tRNA<sub>Pyl</sub> (Fig. 3A). The E. coli type copy of the <i>lacI</i> gene was expressed from the plasmid inserted in the lac<sup>-</sup>lac<sup>+</sup> fusion gene or from the lac<sup>-</sup>lac<sup>+</sup> fusion gene have been inserted in the lac<sup>-</sup> part of a lac<sup>-</sup> lac<sup>+</sup> fusion gene have been extensively used to assess sequence-function relationships of a wide range of tRNA molecules by quantifying the β-galactosidase activity resulting from suppression of the amber stop codon [20–22]. We transformed the E. coli XAC/A24 strain with plasmids carrying pylS and pylT genes. The β-galactosidase activity could only be detected when the transformant cells were grown in LB containing either Cyc or d-prolyl-lysine, indicating that the PyS/tRNA<sub>Pyl</sub> pair is truly orthogonal in E. coli and that the Pyl analogues are inserted in the LacI–LacZ fusion protein (Fig. 3A). The β-galactosidase activity in cells grown with the Pyl analogues was compared to that from either an E. coli strain containing in its genome a wild type copy of the lac<sup>-</sup> lac<sup>+</sup> fusion gene or from the E. coli XAC/A24 strain transformed with a plasmid carrying an E. coli amber suppressor tRNA<sub>Lys</sub>. The amber suppressor tRNA<sub>Lys</sub> allowed ~38% read-through efficiency as previously reported [13], and the suppression efficiency in cells grown on Cyc and d-prolyl-lysine were ~25% and ~18%, respectively (Fig. 3A). Interestingly, d-prolyl-lysine (K<sub>M</sub> 500 μM) was a moderately better substrate than Cyc (K<sub>M</sub> 670 μM) for amino acid activation by PyrRS in vitro, yet the addition of d-prolyl-lysine in the growth medium resulted in slightly lower suppression efficiency relative to growth on Cyc. This suggests either a more limited level of cellular uptake of d-prolyl-lysine or cellular degradation.

In addition to quantifying LacI–LacZ activity, we also directly established that UAG read-through in lac<sup>-</sup> lac<sup>+</sup> mRNA was dependent on Cyc using an immunoblot analysis. This showed the presence of a full length 150 kDa LacI–LacZ protein only in cells grown in the presence of Cyc, implying the incorporation of the Pyl analogue at the suppression site within lacI (Fig. 3B). Finally, we analyzed the in vivo charging level of tRNA<sub>Pyl</sub> using the acid gel electrophoresis method. While tRNA<sub>Pyl</sub> extracted from cells grown in the presence of Cyc was significantly aminoacylated, it was only present in a decylated form when Cyc was absent from the culture media (Fig. 3C). Combined, this evidence indicates PyrRS is specifically acylating tRNA<sub>Pyl</sub> with the Pyl analogues in vivo, and the E. coli translation machinery is able to use this charged tRNA in protein synthesis at levels comparable to other amber suppressor aminoacyl-tRNAs.

In order to see whether Cyc insertion is being influenced by the particular context in which the UAG codon is placed in lac<sup>-</sup> lac<sup>+</sup>, we investigated whether the Pyl analogue could be incorporated at a similar level in another E. coli protein. Because extensive analysis of E. coli trpA mutant strains have been conducted previously [15], and used as a powerful tool to screen libraries of mutant tRNAs [23], we chose to use this system as a reporter for Cyc incorporation. The E. coli strain FTP5822 carries an inactivating mutation in the trpA gene (Gln 243 to UAG nonsense), which results in a tryptophan auxotrophy phenotype due to the organism’s inability to synthesize this amino acid. This phenotype can be rescued either when the cells are grown on an exogenic source of tryptophan or when provided with a suitable suppressor tRNA. We transformed the E. coli FTP5822 strain with plasmids carrying the.

![Fig. 3. Incorporation of Cyc and d-prolyl-lysine into β-galactosidase. (A) β-Galactosidase activity catalyzed by the wild type enzyme produced by the E. coli FTP5822 strain, which carries the lacI–lacZ fusion with a wild type tryptophan codon in place of the UAG triplet (L1). E. coli XAC/A24 carrying lacI–lacZ with Trp220 to UAG mutation transformed with E. coli suppressor tRNA<sub>Lys</sub> (L2). E. coli XAC/A24 strain transformed with M. barkeri pylS and pylT grown in presence of either 10 mM Cyc (L3) of 10 mM d-prolyl-lysine (L4), no amino acid (L5). E. coli XAC/A24 strain transformed with pylT and empty pCBS2 vector (L6). (B) In vivo aminoacylation of tRNA<sub>Pyl</sub> measured by acid gel electrophoresis. Total tRNA extracted from E. coli XAC/A24 strain transformed with M. barkeri pylS and pylT grown in presence (L3) or absence (L5) of 10 mM Cyc. (C) Immunoblot of full-length LacI–LacZ fusion protein. Total protein extract from E. coli XAC/A24 strain transformed E. coli suppressor tRNA<sub>Lys</sub> (L2) and from E. coli XAC/A24 strain transformed with M. barkeri pylS and pylT grown in presence (L3) or absence (L5) of 10 mM Cyc.](image-url)
pylS and pylT genes and grew the cells in M9 minimal media in the presence or absence of Cyc. The presence of the Pyl analogue in the culture media restored bacterial growth, implying that a functional full-length tryptophan synthase had been made (Fig. 4A). This was confirmed by assaying tryptophan synthase activity in cell free extracts of the E. coli grown in presence of Cyc (2.6 U/mg, 16% of wild type activity). To compare suppression activity, E. coli FTP5822 strain was also transformed with the amber suppressor tRNA Lys and tryptophan synthase activity was measured (5.5 U/mg, 33% of wild type activity). As with the lac–lacZ read-through system, tRNA Pyl suppression was reduced relative to the suppressor tRNA Lys (Fig. 4A).

Fig. 4. Incorporation of Cyc and 6-prolyl-lysine into tryptophan synthase. (A) Growth curve in M9 minimal media of E. coli FTP5822 strain transformed with E. coli suppressor tRNA<sub>Lys</sub> (●), E. coli FTP5822 strain transformed with M. barkeri pylS and pylT grown in presence (○) or absence (□) of 10 mM Cyc. E. coli FTP5822 strain transformed with either M. barkeri pylS (□) or M. barkeri pylT (△) and grown in presence of 10 mM Cyc. (B) In vivo aminoacylation of tRNA<sub>Pyl</sub> as judged by acid gel electrophoresis. Total tRNA extracted from E. coli FTP5822 strain transformed with M. barkeri pylS and pylT grown in presence (L1) or absence (L2) of 10 mM Cyc.

3.3. Conclusions and perspectives

We identified two Pyl analogues, the commercially available N-6-cyclopentyloxycarbonyl-L-lysine (Cyc) and N-6-6-prolyl-L-lysine (6-prolyl-lysine) which is easily synthesized, that can circumvent the lack of a reliable and cheap source of synthetic Pyl for both in vitro and in vivo study of PylRS. By doing so, we also established the essentiality of the pyrroline ring stereochemistry, thus gaining some preliminary insight on the structure–activity relationships for Pyl recognition by PylRS. Determination of the chemical pathway to Pyl and the genes responsible for Pyl biosynthesis will allow the future production of the Pyl required for active site characterization of PylRS. Furthermore, we establish that the PLYS element is not essential for pyrrolysine insertion at UAG codons, but raise the possibility that PLYS may enhance read-through levels.

Acknowledgements: We are grateful to Frances Pagel, Emanuel Murgo, Hiroaki Nakano and William McClain for the gift of plasmids and E. coli strains. This work was supported NIH grant GM22854 (to D.S.), and by support from Bristol-Myers Squibb, Amgen, Yamanouchi, Pflizer, and Merck (to J.L.W.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2006.11.028.

References


