Molecular evolution of protein-RNA mimicry as a mechanism for translational control

Assaf Katz1, Lindsey Solden1, S. Betty Zou2, William Wiley Navarre2 and Michael Ibba1,3,*

1Department of Microbiology, Ohio State University, Columbus, OH 43210, USA, 2Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada and 3Ohio State Biochemistry Program, Center for RNA Biology, Ohio State University, Columbus, OH 43210, USA

Received July 1, 2013; Revised November 19, 2013; Accepted November 21, 2013

ABSTRACT

Elongation factor P (EF-P) is a conserved ribosome-binding protein that structurally mimics tRNA to enable the synthesis of peptides containing motifs that otherwise would induce translational stalling, including polyproline. In many bacteria, EF-P function requires post-translational modification with (R)-β-lysine by the lysyl-tRNA synthetase paralog PoxA. To investigate how recognition of EF-P by PoxA evolved from tRNA recognition by aminoacyl-tRNA synthetases, we compared the roles of EF-P/PoxA polar contacts with analogous interactions in a closely related tRNA/synthetase complex. PoxA was found to recognize EF-P solely via identity elements in the acceptor loop, the domain of the protein that interacts with the ribosome peptidyl transferase center and mimics the 3’-acceptor stem of tRNA. Although the EF-P acceptor loop residues required for PoxA recognition are highly conserved, their conservation was found to be independent of the phylogenetic distribution of PoxA. This suggests EF-P first evolved tRNA mimicry to optimize interactions with the ribosome, with PoxA-catalyzed aminoacylation evolving later as a secondary mechanism to further improve ribosome binding and translation control.

INTRODUCTION

Mimicry is the process by which one species or molecule evolves to be similar to another (1,2). At the molecular level (i.e. molecular mimicry), this is observed in several pathogens that have evolved molecular structures to resemble the host’s own structures to evade or trigger a response from it (2). An example of this is translation initiation factor 2A (eIF2A) that is mimicked by K3L from poxviruses. During viral infections, double-stranded RNA-dependent protein kinase (PKR) phosphorylates eIF2A preventing protein translation, which in turn decreases viral production. This cellular defense mechanism placed a selective pressure on the poxvirus, leading to the evolution of K3L (the mimic), an inhibitor that binds the active site of PKR, thereby preventing eIF2A phosphorylation (3).

In addition to the mimicry of molecules between interacting species, the term ‘molecular mimicry’ is also used to describe the convergent evolution of molecules that coexist within the same species. Some of the most intriguing examples of molecular mimicry are within the transcription and translation apparatus where several proteins mimic nucleic acids to control these processes (4–6). For instance, several translation factors mimic tRNA, which presumably allows these proteins to fit properly within binding sites on the ribosome (6,7). Perhaps one of the most striking examples is EF-P, which not only resembles a complete tRNA molecule in structure (Figure 1), but also emulates the tRNA before its interaction with the ribosome.

EF-P is a non-essential elongation factor that has been shown to facilitate the translation of several proteins involved in pathogenesis and survival under other stressful conditions (8,9). Many EF-P-regulated proteins have a stretch of prolines or other sequence patterns that induce ribosome stalling during translation. EF-P is required for alleviating the stalling of these paused ribosomes, allowing for the translation of these genes to continue (10–12). To function efficiently in translation, EF-P is (R)-β-lysylated at Lys34 (Escherichia coli numbering), a position structurally equivalent to A76 in tRNA (10,13,14). The modification of EF-P is catalyzed by PoxA, a protein paralog of the catalytic domain of lysyl–tRNA synthetase (LysRS). All aminoacyl–tRNA synthetases (aaRS) including LysRS catalyze tRNA aminoacylation in a two-step reaction. First, the substrate...
amino acid is activated with ATP forming an aminoacyl adenylate (aa-AMP) that usually remains stably bound to the enzyme in the absence of tRNA. In a second step, the amino acid is transferred from AMP to the 3’ or 2’ OH of A76 at the 3’-end of tRNA:

\[
\text{aa+ATP} \leftrightarrow \text{PPi+aa-AMP} \\
\text{aa-AMP+tRNA} \leftrightarrow \text{aa-tRNA+AMP}
\]

Similar to all known aaRSs, PoxA also catalyzes the aminoacylation of EF-P in a two-step reaction, which in this case requires ATP and (R)-β-lysine ((R)-β-Lys), an analog of α-lysine that has the amino group in the beta instead of the alpha position (8,14):

\[(R)-\beta\text{-lys+ATP} \leftrightarrow \text{PPi+(R)-\beta\text{-lysyl-AMP}}
\]

\[(R)-\beta\text{-lysyl-AMP+EF-P} \leftrightarrow \text{(R)-\beta\text{-lysyl-EF-P+AMP}}
\]

(R)-β-Lys is transferred to Lys34 of EF-P forming an (R)-β-lysyl-lysine amino acid. This group is further hydroxylated to form (R)-β-lysyl-hydroxylysine in a reaction catalyzed by the product of the yfcM gene (15). Until now it has not been possible to find the physiological relevance of this hydroxylation, as it is not required for alleviating stalled ribosomes, and yfcM deletion strains do not present any obvious phenotype (10,16).

While PoxA modifies EF-P in a reaction that resembles those catalyzed by aaRSs, the extent of EF-P mimicry to tRNA is unknown. Therefore, we set out to investigate whether EF-P mimics tRNA, not only in structure, but also in its interaction with the LysRS paralog, PoxA. Our results show that PoxA recognizes EF-P mainly through interactions at positions mimicking the acceptor stem of tRNA. Thus, PoxA has retained and even strengthened interactions that were already relevant for recognition of tRNA by aaRSs. Most of the conserved positions in the EF-P sequence correspond to residues contacting the ribosome, with PoxA only recognizing a small subset of these conserved positions. This suggests EF-P first evolved tRNA mimicry to interact with the ribosome, whereas PoxA evolved afterward in a subset of bacterial species to recognize EF-P. The absence of PoxA in many bacteria and the presence of a different pathway for the modification of the EF-P paralog in archaea and eukaryotes (IF5a) further support the hypothesis that tRNA mimicry in this pathway originated to allow for interaction between EF-P and the ribosome, whereas PoxA recognition of the mimic is a more recent addition.

**MATERIALS AND METHODS**

**Strains and vectors, cloning and mutagenesis**

EF-P and PoxA were expressed in BL21 XJB (DE3) strains from plasmids EFP pTYB11 and PoxA...
Complementation assays were performed in *Salmonella typhimurium* 14028s using plasmid pWN403 (pBAD18-poxA) and mutation variants as previously described (8,9). Briefly, hypo-osmolarity was assessed by streaking poxA, (8) and EFP pTYB11 as templates, respectively. All cloning and mutagenesis were confirmed through sequencing at the Plant Microbe Genomics Facility at The Ohio State University and then transformed in *E. coli* BL21 XJB (DE3) cells for overexpression and purification.

**Sensitivity assays**

Complementation assays were performed in *Salmonella typhimurium* 14028s using plasmid pWN403 (pBAD18-poxA) and mutation variants as previously described (8,9). Briefly, hypo-osmolarity was assessed by streaking *Salmonella* on antibiotic medium 2 agar (AB2; Difco, Detroit, MI, USA). Plates were incubated at 37°C for 16 h followed by visual assessment of colony size. For gentamicin and lauryl sulfobetaine sensitivity, overnight cultures of *Salmonella* strains were diluted 1:1000 into fresh Luria-Bertani medium containing either gentamicin and lauryl sulfobetaine. All strains and gentamicin plus lauryl sulfobetaine were cultured at 37°C with shaking. In this condition, all strains grow up to 1–1.15 OD₆₀₀ units in the absence of gentamicin and sulfobetaine.

**Purification of PoxA, EF-P and variant proteins**

All proteins were expressed using autoinduction media and then purified as previously reported (14). Briefly, cells were resuspended in buffer A (50 mM Tris HCl, pH 8.0, 500 mM NaCl and 10% glycerol) plus 0.05% triton X-100 and lysed by sonication. Cell debris was eliminated by centrifugation and the supernatant loaded to a chitin column. Column was washed with five volumes of buffer A plus 0.05% triton X-100, and then with 25 volumes of buffer A. Finally, the column was washed with 2 volumes of buffer A plus 100 mM dithiothreitol and incubated overnight at 4°C. After this, the protein was eluted with buffer A and concentrated using a filter centrifugation apparatus. Then the protein was dialyzed against dialysis buffer (25 mM Tris HCl, pH 8.0, 150 mM NaCl and 4 mM 2-mercaptoethanol) plus 50% glycerol and stored at −20°C. EF-P used for kat/KM determinations was further purified through anion exchange using a Resource Q column. Protein was loaded on the column and washed with 10 volumes of buffer 1 (25 mM Tris HCl, pH 8, 50 mM NaCl and 2 mM 2-mercaptoethanol), then with additional 6 volumes of 10% buffer 2 (25 mM Tris HCl, pH 8, 1 M NaCl and 2 mM 2-mercaptoethanol) in buffer 1. Finally, the protein was eluted in a linear gradient to 14% buffer 2 in buffer 1, concentrated and dialyzed against dialysis buffer plus 20% glycerol. The protein was then stored in aliquots at −80°C. All proteins were quantified in 6 M guanidinium HCl based on absorbance at 280 nm and a molar extinction coefficient of 24750 M⁻¹ cm⁻¹ (EF-P) or 29870 M⁻¹ cm⁻¹ (PoxA) that were calculated using EMBOSS (17).

**Phosphorylation of EF-P-PKA**

EF-P-PKA was phosphorylated for 1 h at 30°C in a reaction mixture that contained 60 μM EF-P-PKA, 50 mM Tris HCl, pH 7.5, 10 mM MgCl₂, 1.4 μM (~9 μCi/μl) 32P-γATP, 200 μM ATP and 30 U/ml PKA (NEB). Then the 32P-EF-P-PKA was purified using a DEAE sepharose column at room temperature. The protein was loaded and the column was cleaned with 50 volumes of buffer 1, then with additional 15 volumes of 5% buffer 2 in buffer 1 and eluted with 10% buffer 2 in buffer 1. The purified protein was dialyzed against dialysis buffer with 50% glycerol and stored in aliquots at −80°C.

**Active site titration of PoxA**

PoxA active site titration was made in a reaction mix containing 0.1 M glycine, pH 9.0, 30 mM KCl, 10 mM MgCl₂, 30 μM ATP, 0.0065 μM (~0.02 μCi/μl) 32P-γATP, 2.5 mM 2-mercaptoethanol, 4.2 U/ml pyrophosphotase and 4 mM (R)-β-Lys. Reactions were started by adding enzyme and performed at room temperature. At specific time points, aliquots were taken and quenched by mixing with 2.5 volumes of 0.2 M Na acetate, pH 5.0. Aliquots (0.5 μl) of each sample were separated on PEI-F cellulose plates using 0.1 M ammonium acetate plus 5% acetic acid as mobile phase. Plates were analyzed using phosphorimaging, and PoxA active sites were quantified from the plateau of the (R)-β-lysyl-AMP intermediate formation.

**Aminoacylation of EF-P with 14C-α-lysine**

EF-P aminoacylation with 14C-α-Lysine was performed at 37°C in a mixture containing 100 mM glycine, pH 9.0, 30 mM KCl, 10 mM MgCl₂, 8 mM ATP, 3 mM 2-mercaptoethanol, 8.4 U/ml pyrophosphotase and 0.75 μM EF-P. At defined time points, 6 μl of reaction mixture was precipitated on 5% trichloroacetic acid saturated 3MM Whatman paper. The papers were washed three times with 5% trichloroacetic acid, once with ethanol and dried at 85°C. Then the 14C aminoacylated EF-P was quantified using scintillation counting. Additionally, aliquots of the reaction were separated in a sodium dodecyl sulphate-polyacrylamide gel electrophoresis, and 14C-α-Lysyl-EF-P was detected by phosphorimaging to confirm aminoacylation of EF-P.

**Aminoacylation of EF-P with (R)-β-Lys**

Aminoacylation kinetics of EF-P with (R)-β-Lys were assayed at 37°C in a mixture containing 100 mM glycine, pH 9.0, 30 mM KCl, 10 mM MgCl₂, 2.5 mM...
2-mercaptoethanol, 0.1 mg/ml bovine serum albumin, saturating concentrations of ATP (2–8 mM depending on the PoxA mutant), 6–8 mM (R)-β-Lys, 4 mM $^{32}$P-EF-P-PKA as a probe for EF-P aminoacylation and variable concentrations of EF-P (0.5–2.5 μM). Reaction aliquots were taken at defined time points and quenched in an equal volume of stop solution (2% triton X-100, 1% 2-mercaptoethanol, 750 mM Na acetate, pH 5.0). Aminoacylated and non-aminoacylated forms of $^{32}$P-EF-P-PKA were separated through isoelectric focusing using a 4.5–5.4 pH gradient (Pharmalyte 4.5–5.4 GE-Healthcare) and subsequently quantified using phosphorimaging.

**Sequence and structure analysis**

Structure interfaces and polar contacts were determined using PDBe PISA (18). Additional polar contacts previously described were also included (13). Figures were made with The PyMOL Molecular Graphics System (version 1.2.0.3 Schrödinger, LLC). Alignments of 112 sequences of EF-P from organisms of diverse genera were performed with PROMALS3D (19) using EF-P structures from pdb 3a5z-D, 1ueb-A, 3oyy-A to guide the alignment. WebLogo representation of this alignment, as well as from the sequences with or without a Lys residue in the modification position (Supplementary Tables S3 and S4), was performed using the WebLogo online server (20). Variability of each position on the sequence alignments was calculated using the Shannon entropy analysis of the protein variability server (21).

**RESULTS**

**Equivalent contacts between PoxA/EF-P and aspartyl-tRNA synthetase/tRNA are within the acceptor stem**

To investigate how PoxA evolved from LysRS to acquire its substrate specificity for a protein rather than RNA, it would be ideal to first compare the structure of EF-P with that of tRNA$^{\text{lys}}$ bound to LysRS. tRNA$^{\text{Asp}}$ is not well resolved in the corresponding crystal structure (22), so instead we compared the EF-P/PoxA complex structure with that of tRNA$^{\text{Asp}}$ bound to AspRS (pdb 1asy). AspRS is closely related to LysRS in structure and, together with AsnRS, they form the class IIb group of aaRSs (23,24), which share common tRNA identity elements, including the discriminator base (25–27). In addition, the effects of several alanine substitution mutations on the contact surface of the tRNA$^{\text{Asp}}$/AspRS complex have been studied (27), providing a useful reference for comparisons with EF-P/PoxA complex.

AspRS, in common with all class IIb aaRSs, has two domains: a catalytic domain that interacts with the acceptor stem of tRNA and an anticodon-binding domain that binds to the anticodon stem loop (Figure 1). Most of the binding energy comes from interactions with the anticodon-binding domain (27) that is absent from PoxA, which contains only the catalytic domain. Focusing on the polar contacts (Figure 1 and Table 1), we observe within the catalytic domain of both enzymes some interactions that have been lost from the EF-P/PoxA complex, others that are conserved (Glu102, Glu103, Arg106 and His108 from PoxA) and a group representing novel interactions not present in the AspRS complex (His52, Asp177, Asn180, Glu185, Gln193, Ser218, Arg235, Glu244 and Arg303 from PoxA). Most of the common contacts correspond to AspRS interactions with nucleotides 73 (discriminator base) and 74 of tRNA$^{\text{Asp}}$. The equivalent positions are located in domain 1 (N-terminal) of EF-P along with all of the novel contacts. Non-polar and polar interactions between EF-P and PoxA have similar distributions, with the only notable differences being in a group of contacts formed between EF-P and the second monomer in the PoxA dimer (Pro63, Gly64, His65, Ser66 and Gln67 from PoxA) (Supplementary Figure S1). Overall, this distribution indicates that most of the EF-P surface does not contact PoxA and, furthermore, the protein is tilted ~60° compared with tRNA$^{\text{Asp}}$ binding to AspRS, with domain III of EF-P displaced from where the anticodon binding domain would have been in a class IIb aaRS (Supplementary Figures S1 and S2) (13). To determine the roles of the retained and novel interactions, we constructed alanine replacement variants for all PoxA residues involved in polar interactions with EF-P and further analyzed their effect on EF-P aminoacylation.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Position</th>
<th>Atom</th>
<th>Amino acid</th>
<th>Position</th>
<th>Atom</th>
</tr>
</thead>
<tbody>
<tr>
<td>His</td>
<td>52</td>
<td>O</td>
<td>Lys</td>
<td>31</td>
<td>NZ</td>
</tr>
<tr>
<td>His</td>
<td>52</td>
<td>O</td>
<td>Lys</td>
<td>31</td>
<td>CE</td>
</tr>
<tr>
<td>His</td>
<td>52</td>
<td>ND1</td>
<td>Lys</td>
<td>34</td>
<td>O</td>
</tr>
<tr>
<td>His</td>
<td>52</td>
<td>ND1</td>
<td>Gly</td>
<td>35</td>
<td>O</td>
</tr>
<tr>
<td>His</td>
<td>52</td>
<td>ND1</td>
<td>Gln</td>
<td>36</td>
<td>N</td>
</tr>
<tr>
<td>Glu</td>
<td>102</td>
<td>O</td>
<td>Lys</td>
<td>31</td>
<td>NZ</td>
</tr>
<tr>
<td>Glu</td>
<td>103</td>
<td>OE1</td>
<td>Lys</td>
<td>31</td>
<td>NZ</td>
</tr>
<tr>
<td>Glu</td>
<td>103</td>
<td>OE2</td>
<td>Lys</td>
<td>31</td>
<td>NZ</td>
</tr>
<tr>
<td>Arg</td>
<td>106</td>
<td>NH1</td>
<td>Glu</td>
<td>28</td>
<td>OE1</td>
</tr>
<tr>
<td>Arg</td>
<td>106</td>
<td>NE</td>
<td>Glu</td>
<td>28</td>
<td>OE2</td>
</tr>
<tr>
<td>Arg</td>
<td>106</td>
<td>NH1</td>
<td>Glu</td>
<td>28</td>
<td>OE2</td>
</tr>
<tr>
<td>His</td>
<td>108</td>
<td>NE2</td>
<td>Lys</td>
<td>31</td>
<td>O</td>
</tr>
<tr>
<td>Asp</td>
<td>177</td>
<td>OD1</td>
<td>Lys</td>
<td>54</td>
<td>NZ</td>
</tr>
<tr>
<td>Asp</td>
<td>177</td>
<td>OD2</td>
<td>Lys</td>
<td>54</td>
<td>NZ</td>
</tr>
<tr>
<td>Asn</td>
<td>180</td>
<td>ND2</td>
<td>Thr</td>
<td>55</td>
<td>O</td>
</tr>
<tr>
<td>Glu</td>
<td>185</td>
<td>OE1</td>
<td>Lys</td>
<td>57</td>
<td>NZ</td>
</tr>
<tr>
<td>Gln</td>
<td>193</td>
<td>NE2</td>
<td>Gly</td>
<td>33</td>
<td>O</td>
</tr>
<tr>
<td>Gln</td>
<td>193</td>
<td>NE2</td>
<td>Lys</td>
<td>34</td>
<td>O</td>
</tr>
<tr>
<td>Ser</td>
<td>218</td>
<td>O</td>
<td>Lys</td>
<td>34</td>
<td>NZ</td>
</tr>
<tr>
<td>Arg</td>
<td>235</td>
<td>NH2</td>
<td>Gly</td>
<td>33</td>
<td>O</td>
</tr>
<tr>
<td>Glu</td>
<td>244</td>
<td>OE1</td>
<td>Gly</td>
<td>33</td>
<td>N</td>
</tr>
<tr>
<td>Glu</td>
<td>244</td>
<td>OE2</td>
<td>Gly</td>
<td>33</td>
<td>N</td>
</tr>
<tr>
<td>Arg</td>
<td>303</td>
<td>NH1</td>
<td>Lys</td>
<td>31</td>
<td>O</td>
</tr>
<tr>
<td>Arg</td>
<td>303</td>
<td>NH2</td>
<td>Lys</td>
<td>31</td>
<td>O</td>
</tr>
</tbody>
</table>

$^a$Contacts determined using PDBE Pisa. Additional contacts previously reported are also included (13).
In vitro characterization of PoxA catalyzed amino acid activation and aminoacylation

The catalytic activities of aaRSs are commonly monitored using radiolabeled amino acids. As radioactive (R)-\(\beta\)-Lys is not commercially available, we developed alternative strategies to characterize PoxA activity in vitro. Active site titration was performed using \(\alpha\)-\(32\)P-ATP and non-radioactive (R)-\(\beta\)-Lys, which allowed monitoring of ATP degradation as well as the formation of \(32\)P-AMP and \(32\)P-(R)-\(\beta\)-lysyl adenylate by TLC and phosphorimaging quantification. Some aaRS show half-of-the-sites reactivity, meaning that only one of the monomers in the dimer at a time can perform amino acid activation (28). To determine whether this is true for PoxA, we calculated the concentration of functional PoxA active sites, which were considered to be equivalent to the concentration of \(32\)P-(R)-\(\beta\)-lysyl adenylate at the plateau of the reaction, a method commonly used with aaRSs. Doubling the enzyme concentration resulted in a 2-fold increase of the plateau level, confirming that the substrate concentration was saturating for the reaction. Enzyme concentrations calculated for purified WT PoxA were usually within 90–95% of concentrations calculated from spectrophotometric determinations (data not shown). The fact that active site titration of PoxA correlated reasonably well with spectrophotometric quantifications suggests that both PoxA monomers are active in the dimer at the same time. During active site titration, the concentration of \(32\)P-(R)-\(\beta\)-lysyl adenylate plateaued after the first time point (15 s) for WT and most of the PoxA variants. Some PoxA variants were slower than WT (R106A, R235A and especially R303A), whereas others were almost completely inactive (E244A; Table 2). Owing to the low activity observed during (R)-\(\beta\)-lysyl adenylate formation, PoxA E244A was not analyzed further.

PoxA recognizes EF-P through residues that mimic the tRNA 3'-end

To compare the role of the contacts that are specific to the EF-P/PoxA complex with those that are common to other class IIb aaRSs, we measured the effect of the corresponding alanine substituted variants on PoxA’s ability to aminoacylate EF-P with (R)-\(\beta\)-Lys. Modification of EF-P with (R)-\(\beta\)-Lys adds a positive charge that allows for the separation of the modified protein from the non-aminoacylated form by isoelectric focusing. To allow quantification of both modified and unmodified EF-P, we added a phosphorylation site at the carboxy terminus of EF-P far from the interaction surface with PoxA. Phosphorylation of this site using the catalytic subunit of the cAMP-dependent protein kinase (PKA) and \(\gamma\)-\(32\)P ATP allowed us to quantify the fraction of lysylated EF-P by analyzing phosphor images of the samples separated by isoelectric focusing (Supplementary Figure S3). Based on their aminoacylation activity, the PoxA variants can be divided in three main groups (Table 2). The majority of amino acid replacements have little or no effect on PoxA activity. These include PoxA E102A, R106A, N180A, Q193A, S218A and R303A, which have a \(k_{cat}/K_M\) of at least 70% of the WT protein. A second group composed of H52A, D177A and E185A showed a moderate loss of activity, with a \(k_{cat}/K_M\) between 20 and 50% of WT PoxA. The third group composed of E103A, H108A and R235A retains no detectable activity for aminoacylation of EF-P with (R)-\(\beta\)-Lys.

All inactivating replacements map to PoxA residues that interact with a loop region of EF-P domain 1 where Lys34, the modification site, is located (Figure 2). One of these residues is Glu103 that hydrogen bonds to Lys31 of EF-P (2.9 Å). The equivalent interaction in the tRNA\(^{Asp}\)/AspRS complex corresponds to Ser329 (Glu264 in \(E. coli\) LysRS) that is at hydrogen bonding distance from the discriminator base (G73) (2.5 Å) and is also in close contact with C74 (3.7 Å) of tRNA\(^{Asp}\). Replacement of Ser329 as well as any of the two other amino acids of AspRS that contact the discriminator base (Asn328 and Thr331) had only a mild effect on efficiency of tRNA\(^{Asp}\) aminoacylation (27). Replacement of the PoxA residues that correspond to Asn328 and Thr331 (Glu102 and Arg106) did not affect EF-P aminoacylation. In contrast, replacement of Glu103 in PoxA completely inactivated the enzyme, indicating a much more important role for this position than observed in the corresponding aaRSs.

His108 of PoxA also forms a hydrogen bond with Lys31 through interactions to the backbone oxygen (3.3 Å), and additionally is in close contact with Phe29 (3.4 Å). PoxA His108 corresponds to one of the most conserved residues in structural motif 2 of AspRS, His334 (His270 in \(E. coli\) LysRS), which contacts C74 of tRNA\(^{Asp}\). As with His108 of PoxA, replacement of His334 in AspRS severely reduced substrate binding affinity and aminoacylation (27). In AspRS, substitution of this His with Ala also reduced the rate of amino acid activation (27). In contrast, replacement of His108 in PoxA did not produce any significant effect on the amino acid activation reaction (Table 2), indicating that the role of this His residue has partially diverged between the two enzymes.

The third replacement that inactivates PoxA corresponds to R235A. Arg235 contacts the backbone oxygen of Gly33 (3.6 Å) and Lys34 (4.0 Å) of EF-P. This arginine corresponds to Ser469 of AspRS, which does not contact tRNA\(^{Asp}\). \(E. coli\) LysRS has an arginine (Arg412), which structurally co-localizes with Arg235 of PoxA. Although tRNA\(^{Lys}\) is not well resolved in the LysRS crystal structure, superposing LysRS to AspRS reveals that one of the A76 phosphate oxygens from tRNA\(^{Asp}\) is a short distance from Arg412 of LysRS (5.2 Å), suggesting a possible direct contact with tRNA. Independently of this, we consider Lys34 of EF-P to be functionally analogous to A76 of tRNA, as both are sites of aminoacylation. Alternatively, Lys34 may be viewed as a structural mimic of C75 of tRNA\(^{Asp}\), as both co-localize in the structural alignment between AspRS and PoxA (Figure 2B). Structural alignment of the EF-P/PoxA complex with tRNA\(^{Asp}\)/AspRS shows that Phe29 of EF-P localizes in a position similar to G73 in tRNA\(^{Asp}\) and Lys31 is near C74. These different EF-P residues are grouped together in what we refer to as the acceptor loop, which structurally and functionally mimics the 3'-end of the acceptor stem of tRNA (Figure 2A and B).
Table 2. Effect of alanine substitutions on PoxA activity

<table>
<thead>
<tr>
<th>PoxA variant (AspRS equivalent)</th>
<th>Area of tRNA contacted in AspRS(^a)</th>
<th>Effect on amino acid activation</th>
<th>(k_{\text{cat}}/K_M) ((\mu\text{M}^{-1} \text{s}^{-1}))(^b)</th>
<th>Relative (k_{\text{cat}}/K_M) ((R))</th>
<th>(\Delta \Delta G^\ddagger) (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td></td>
<td></td>
<td>0.05 ± 0.005</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>H52A (gap)</td>
<td>Gap</td>
<td>Active</td>
<td>0.018 ± 0.007</td>
<td>0.37</td>
<td>0.62</td>
</tr>
<tr>
<td>E102A (N328)</td>
<td>Acceptor arm (G73)</td>
<td>Active</td>
<td>0.049 ± 0.022</td>
<td>0.98</td>
<td>0.011</td>
</tr>
<tr>
<td>R106A (T331)</td>
<td>Acceptor arm (G73 + C74)</td>
<td>Reduced</td>
<td>0.055 ± 0.005</td>
<td>1.1</td>
<td>-0.057</td>
</tr>
<tr>
<td>H108A (H334)</td>
<td>Acceptor arm (C74)</td>
<td>Active</td>
<td>0.037 ± 0.008</td>
<td>0.73</td>
<td>0.19</td>
</tr>
<tr>
<td>D177A (gap)</td>
<td>Gap</td>
<td>Active</td>
<td>0.027 ± 0.003</td>
<td>0.53</td>
<td>0.39</td>
</tr>
<tr>
<td>N180A (gap)</td>
<td>Gap</td>
<td>Active</td>
<td>0.013 ± 0.002</td>
<td>0.27</td>
<td>0.81</td>
</tr>
<tr>
<td>E185A (gap)</td>
<td>Gap</td>
<td>Active</td>
<td>0.055 ± 0.009</td>
<td>1.1</td>
<td>-0.057</td>
</tr>
<tr>
<td>Q193A (D421)</td>
<td>No contact</td>
<td>Active</td>
<td>0.047 ± 0.009</td>
<td>0.93</td>
<td>0.044</td>
</tr>
<tr>
<td>S218A (E451)</td>
<td>No contact</td>
<td>Reduced</td>
<td>0.044 ± 0.004</td>
<td>0.87</td>
<td>0.086</td>
</tr>
</tbody>
</table>

\(^a\)Data from Eriani and Gangloff.\(^{27}\)
\(^b\)\(k_{\text{cat}}/K_M\) was estimated using sub-saturating EF-P concentrations from the slope of the equation, \(V = k_{\text{cat}} [E][S]/K_M\).
\(^\ddagger\)Compared with wild-type.
\(^\ddagger\)\(\Delta \Delta G^\ddagger\) represents differences in the transition state binding \(\Delta G\) and was estimated to be RT\(\ln[k_{\text{cat}}/K_M]_{\text{mut}}/(k_{\text{cat}}/K_M)_{\text{WT}}\) (29).

Replacement of most amino acids involved in the polar contacts that are unique to the PoxA/EF-P complex (i.e. absent from the tRNA\(^{\text{Asp}}\)/AspRS complex) have a minimal effect on aminoacylation activity. The only exceptions were R235A (discussed earlier) along with three other variants with modest effects on activity, H52A, D177A and E185A (Figure 2C). His52 contacts the acceptor loop of EF-P (with contacts to Lys31, Lys34, Gly35 and Gln36), but the other two amino acids from this group interact with amino acids outside the acceptor loop in a region that is not used by AspRS to contact tRNA\(^{\text{Asp}}\).

**EF-P identity elements**

Replacement of most of the amino acids in PoxA involved in polar contacts with EF-P did not lead to any significant effect on the aminoacylation kinetics, with the exception of three positions that contact Phe29, Lys31, Gly33 and Lys34 in the acceptor loop. To confirm the relevance of these contacts, we replaced each of these amino acids with Ala in the EF-P acceptor loop and analyzed their effect on aminoacylation with \(^{14}\)C-\(\alpha\)-lysine (Figure 3). Previous experiments have shown that substituting Lys34 with either Arg or Ala prevents aminoacylation of EF-P (13,14). Replacements of Phe29 and Lys31 with Ala also completely abolished PoxA aminoacylation, confirming their role as identity elements (Figure 3). Substitution of Gly33 with Ala also greatly diminished PoxA’s ability to modify EF-P, although some aminoacylation activity was retained. Taken together, our data indicate that residues Phe29, Lys31, Gly33 and Lys34 are the identity elements of EF-P for PoxA aminoacylation and that these amino acids are clustered in the acceptor loop. These contacts mimic those of the tRNA acceptor stem with the aaRS, while novel contacts elsewhere have at most a secondary role in this interaction. Although EF-P and tRNAs are composed of different monomers and bind their respective enzymes with different geometries (Supplementary Figure S2), our data indicate that the EF-P/PoxA complex evolved a molecular recognition strategy that is analogous to the one used by class IIb aaRSs, the evolutionary ancestors of PoxA.

**Conservation of EF-P binding sites to PoxA and the ribosome**

It is well established that positions in a protein that interact with other molecules tend to be more conserved, as their variability during evolution is constrained by the need to recognize their molecular partners, which imposes a strong selective pressure (21,30). Thus, one would expect in organisms where PoxA is present that the contacts made to EF-P would be well conserved, especially in the case of the four identity elements that are essential for EF-P aminoacylation. Similarly, it would be predicted that this would not be the case for EF-P from organisms that lack the poxA gene, as these positions are not under the same selective constrains. The acceptor loop and other PoxA contacting positions are well conserved in EF-P sequences from organisms that encode a poxA gene (Table 3 and Figures 4 and 5 and Supplementary Figures S4 and S5). The most conserved residues include the modification site (Lys34 following E. coli EF-P numbering) and some surrounding positions that are completely invariant (32, 33 and 35). Most contacts outside of the acceptor loop are also well conserved, with some minor differences observed for organisms that lack the poxA gene but retain a Lys at position 34. Position 29, one of the four EF-P identity elements, is usually Phe or Tyr in organisms that have the poxA gene, but it is frequently His in organisms that lack poxA. Other positions interacting with PoxA such as residues 28 and 36 of the acceptor loop or 57 and 58 distal from it show a similar pattern of conservation. Overall, most positions in EF-P contacting PoxA are strictly conserved, regardless of whether or not the corresponding organism has a poxA gene. This includes all the other identity elements (EF-P residues 31, 33 and 34), as well as several other contacting positions (e.g. residues 8, 30, etc.)
The strong conservation of these residues in EF-P, regardless of the presence of the poxA gene, suggests that there are other constraints on their variability beyond the need to bind PoxA. The most likely selective pressure that constrains variability of EF-P’s surface amino acids is the need to maintain functional interactions with the ribosome to prevent translational stalling. Most of the EF-P residues that contact PoxA, including all of the identity elements, also correspond to positions that interact with the ribosome or tRNA.
Table 3. Conservation of the amino acid sequence of the acceptor loop from EF-P

<table>
<thead>
<tr>
<th>Position</th>
<th>Organisms with PoxA</th>
<th>Organisms without PoxA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Variability^c</td>
<td>Amino acid identity^d</td>
</tr>
<tr>
<td>Glu28</td>
<td>1.198</td>
<td>Usually acid (71%)</td>
</tr>
<tr>
<td>Phe29</td>
<td>1.611</td>
<td>Usually Phe or Tyr (79%)</td>
</tr>
<tr>
<td>Val30</td>
<td>0.735</td>
<td>Usually Val (86%)</td>
</tr>
<tr>
<td>Lys31</td>
<td>1.292</td>
<td>Usually Lys (71%)</td>
</tr>
<tr>
<td>Pro32</td>
<td>0.000</td>
<td>Pro (100%)</td>
</tr>
<tr>
<td>Gly33</td>
<td>0.000</td>
<td>Gly (100%)</td>
</tr>
<tr>
<td>Lys34</td>
<td>0.000</td>
<td>Lys (100%)</td>
</tr>
<tr>
<td>Gly35</td>
<td>0.000</td>
<td>Gly (100%)</td>
</tr>
<tr>
<td>Gln36</td>
<td>1.484</td>
<td>Usually Gln (71%)</td>
</tr>
<tr>
<td>Ala37</td>
<td>0.000</td>
<td>Ala (100%)</td>
</tr>
<tr>
<td>Phe38</td>
<td>0.371</td>
<td>Usually Phe (93%)</td>
</tr>
<tr>
<td>Ala39</td>
<td>2.414</td>
<td>Variable</td>
</tr>
<tr>
<td>Arg40</td>
<td>0.735</td>
<td>Usually Arg (86%)</td>
</tr>
</tbody>
</table>

^aPositions are indicated according to E. coli EF-P numbering. Positions that contact PoxA are highlighted in bold letters.

^bIdentity elements for PoxA aminoacylation.

^cShannon entropy was calculated as an index of variability using Protein Variability Server. A value of 0 indicates no variability at all, whereas 4.322 indicates maximum variability (equal probability to find any of the 20 amino acids) (21).

^dPercentage of occurrence of the amino acids in each position is in parentheses.

Figure 4. Conservation of EF-P residues involved in PoxA contacts. WebLogo representation of a fragment of the alignment of diverse EF-Ps that contain a lysine in the equivalent to the modification position. The top WebLogo (A) corresponds to an alignment of 14 EF-P sequences from organisms that have a poxA gene encoded in their genome. The bottom WebLogo (B) corresponds to an alignment of 17 EF-P sequences from organisms that do not have a poxA gene. The acceptor loop is marked with a black line, with the modification position highlighted with a triangle. Contacts to PoxA (based on pdb 3a52) are marked with red triangles and contacts to the ribosome (pdb 3huw and 3hux) are indicated with cyan triangles. Numbering of the WebLogo positions corresponds to the full alignment, and corresponding positions on EF-P from E. coli are indicated below the triangles (for full alignments see Supplementary Figure S4).
that emergence of PoxA aminoacylation activity in bacterial evolution could have gradually improved fitness. These results are in line with recent reports indicating that unmodified EF-P can stimulate translation of proteins with polyproline stretches, albeit with a much lower efficiency than the modified protein (10), further suggesting that EF-P may have originally functioned without post-translational modification.

**DISCUSSION**

Our results demonstrate that EF-P is recognized mainly through contacts with residues that mimic the 3'-end of the acceptor stem of tRNA. These amino acids are all localized in a loop of domain 1 at the amino terminus of EF-P that we have termed the acceptor loop, as it resembles the acceptor stem of tRNA in both structure and function. It is surprising that although EF-P closely resembles almost the complete structure of a tRNA, PoxA only relies on a small region of EF-P for its recognition (Figure 5 and Supplementary Figure S4). This is in contrast to the recognition of tRNA by a canonical aaRS where the contacts are spread throughout the entire tRNA (Figure 1) and identity elements are usually located at both the acceptor and anticodon stems (31). The only known exception is GluQRS, a paralog of the...
catalytic domain of GluRS that aminoacylates the queuosine-modified wobble base of tRNA\(^{\text{Asp}}\) with glutamate rather than the acceptor stem of tRNA\(^{\text{Glu}}\) (32,33). In this case, the tRNA\(^{\text{Asp}}\) anticodon stem and loop mimic the acceptor stem of tRNA\(^{\text{Glu}}\) and are recognized by GluQRS (34). Molecular docking experiments showed that GluQRS does not contact most of the tRNAs (34) and indicate that tRNA\(^{\text{Asp}}\) nevertheless maintains a complete canonical tRNA shape due to the evolutionary constraints of its need to interact with AspRS, the translation machinery and several nucleotide-modifying enzymes. In a similar manner, the structure of EF-P is also constrained by the need to maintain its interactions not only with PoxA, but also with YfcM and the ribosome. Ribosome contacts are well conserved and distributed throughout the structure of EF-P (Figures 4 and 5 and Supplementary Figures S4 and S5), in contrast to contacts with PoxA that are confined to a subset of these residues (Figure 5 and Supplementary Figures S4 and S5), supporting the idea that functionality at the ribosome is the strongest determinant underlying the structural and functional mimicry of tRNA by EF-P. Furthermore, PoxA is mostly present in proteobacteria and scarcely found in other bacteria (35), suggesting that post-translation modification with (R)-\(\beta\)-Lys is a feature that appeared only recently in evolution to improve the affinity of EF-P for the ribosome. Deletion of \(\text{poxA}\) generates phenotypes less severe than an \(\text{efp}\) deletion, indicating that unmodified EF-P retains part of its activity on the ribosome. Additionally, recent reports show that the role of EF-P modification is mainly to enhance binding to the ribosome and that unmodified EF-P can still stimulate synthesis of proteins with polyproline stretches, although with a much lower efficiency than its aminoacylated version (10). The alternative hypothesis, where EF-P originally interacted with PoxA and interaction with the ribosome appeared later in evolution, is highly unlikely and unsupported by our data. Such a scenario would imply that independent of its binding to the ribosome, EF-P had an alternative unknown function that would have allowed for its retention during evolution. Additionally, the alternative hypothesis implies that all species except the proteobacteria lost \(\text{poxA}\) or that there was widespread lateral transfer of \(\text{efp}\) to all domains of life. Therefore, we favor the hypothesis that EF-P originally interacted with the ribosome, as it is the most parsimonious.

The modification of EF-P by PoxA is analogous to the modification of IF5a (the homolog of EF-P) with a hydroxylated polypeptide by two enzymes, deoxyhypusine synthase and deoxyhypusine hydroxylase (36,37). The presence of two alternative pathways that produce a similar outcome highlights the relevance of EF-P modification, but at the same time raises the question of how EF-P functions in organisms that lack PoxA or deoxyhypusine synthase. In some cases, the lack of PoxA may be indicative of a lack of EF-P modification, which, in turn, may result from alternative mechanisms of ribosome binding or optimization of polyproline translation. In other instances, organisms may have instead evolved an alternative pathway for modifying EF-P to enhance binding to the ribosome, similar to the previously mentioned IF5a modification in archaea and eukaryotes.

Our data suggest that tRNA mimicry originally allowed EF-P to bind the ribosome, thereby increasing the fitness of certain organisms by facilitating synthesis of polyproline containing proteins. PoxA evolution from LysRS likely occurred as a secondary process facilitated by the fact that EF-P already had several key features similar to that of a tRNA. The emergence of the PoxA modification pathway would then have provided additional advantages, as modified EF-P is far more efficient in stimulating translation of polyproline stretches than the unmodified form (10). It is critical that PoxA concurrently lost the ability to recognize the canonical substrates of LysRS to prevent synthesis of (R)-\(\beta\)-Lys-tRNA, which would not be a substrate for translation (38,39). Taken together, these and previous data illustrate the precise mechanism by which molecular mimicry can be used to enhance protein synthesis, and thus, the fitness of the cell.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

The authors thank Medha Raina and Andrei Rajkovic for helpful discussions.

**FUNDING**

National Institutes of Health (NIH) [GM065183 to M.I.]; Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN) [386286-10 to W.W.N.]; Vanier Canada Graduate Scholarship from NSERC (to S.B.Z.). Funding for open access charge: NIH.

**Conflict of interest statement.** None declared.

**REFERENCES**


