Translation factor accelerating peptide bond formation on the ribosome: EF-P and eIF5A as entropic catalysts and a potential drug targets

Vitalii Mudryi, Frank Peske, Marina Rodnina *

Department of Physical Biochemistry, Max Planck Institute for Multidisciplinary Sciences, 37077 Goettingen, Germany

A R T I C L E   I N F O

Keywords:
Proline
Elongation factor P
Posttranslational modification
Ribosome
EarP
Activation entropy

A B S T R A C T

Elongation factor P (EF-P) and its eukaryotic homolog eIF5A are auxiliary translation factors that facilitate peptide bond formation when several sequential proline (Pro) residues are incorporated into the nascent chain. EF-P and eIF5A bind to the exit (E) site of the ribosome and contribute to favorable entropy of the reaction by stabilizing tRNA binding in the peptidyl transferase center of the ribosome. In most organisms, EF-P and eIF5A carry a postranslational modification that is crucial for catalysis. The chemical nature of the modification varies between different groups of bacteria and between pro- and eukaryotes, making the EF-P-modification enzymes promising targets for antibiotic development. In this review, we summarize our knowledge of the structure and function of EF-P and eIF5A, describe their modification enzymes, and present an approach for potential drug screening aimed at EarP, an enzyme that is essential for EF-P modification in several pathogenic bacteria.

Introduction

Ribosomes synthesize proteins in all living cells by catalyzing peptide bond formation. The addition of each new amino acid into the growing peptide occurs in the peptidyl transferase center, which is composed entirely of ribosomal RNA (rRNA). During peptidyl transfer, the amino group of amino acid attached to the transfer RNA (tRNA) in the A site of the ribosome nucleophilically attacks the carbonyl carbon of the peptidyl-tRNA in the P site. The resulting zwitterionic tetrahedral intermediate is rapidly deprotonated and then decomposes to form the peptide bond. The addition of each new amino acid into the nascent chain is highly exothermic, and this rate enhancement is achieved entirely by lowering the entropy of activation [1]. The entropy effect includes precise positioning of the substrates (i.e., of the two tRNAs in the A and P sites), exclusion of bulk water and ordering water molecules that help proton exclusion of bulk water and ordering water molecules that help proton transfer in the active site, and providing an electrostatic network that stabilizes the reaction intermediates [2,3]. The predominance of the entropic effects explains how the ribosome can catalyze a variety of unnatural reactions if the reaction substrates can be mounted on the ribosome and provide an atomistic understanding of the reaction, although the experimental reaction rates are still difficult to reproduce in the simulations ([3] and references therein). In fact, our current understanding of the mechanism of peptide bond formation is remarkably detailed, in particular taking into account that the reaction involves only a few atoms moving on a 2.5 MDa macromolecular machine [16].

However, some questions remain and this is particularly true for amino acid-specific differences in the rate of peptide bond formation. At each elongation cycle of protein synthesis, an elongation factor called EF-Tu in bacteria or eIF1A in eukaryotes delivers an aminoacyl-tRNA (aa-tRNA) to the A site of the ribosome where the tRNA reads the codon of the messenger RNA (mRNA). If the codon is cognate, the CCA-end of the tRNA that carries the amino acid is accommodated in the A site of the peptidyl transferase center. For example, ribosomes can catalyze reactions using \( \alpha \)-amino acids, \( \beta \)-amino, \( \gamma \)-amino acids, or fluorescent derivatives of amino acids [4-6]. Crystal structures have identified residues of the 23S rRNA and the water molecules that form hydrogen bonds with the substrates and catalyze the reaction [7-9]. Atomic mutagenesis of the rRNA suggested the roles of residues at the active site of the ribosome [10-12]. Biochemical and kinetic experiments that measure kinetic isotope effects and linear free energy relationships identified the rate-limiting transition state with three protons in flight [13-15]. This large body of experimental work (reviewed in [16]) has led to numerous computer studies that model the detailed pathway for peptide bond formation on the ribosome and provide an atomistic understanding of the reaction, although the experimental reaction rates are still difficult to reproduce in the simulations ([3] and references therein). In fact, our current understanding of the mechanism of peptide bond formation is remarkably detailed, in particular taking into account that the reaction involves only a few atoms moving on a 2.5 MDa macromolecular machine [16].

However, some questions remain and this is particularly true for amino acid-specific differences in the rate of peptide bond formation. At each elongation cycle of protein synthesis, an elongation factor called EF-Tu in bacteria or eIF1A in eukaryotes delivers an aminoacyl-tRNA (aa-tRNA) to the A site of the ribosome where the tRNA reads the codon of the messenger RNA (mRNA). If the codon is cognate, the CCA-end of the tRNA that carries the amino acid is accommodated in the A site of the peptidyl transferase center. With 21 natural amino acids that

* Corresponding author.
E-mail address: rodnina@mpinat.mpg.de (M. Rodnina).

https://doi.org/10.1016/j.bbadva.2023.100074

Received 19 December 2022; Received in revised form 6 January 2023; Accepted 7 January 2023

Available online 10 January 2023

2667-1603/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
are specified by the genetic code, there is a large combinatorial variety of reactive amino acid donor-acceptor pairs formed by the C-terminal amino acid at the growing peptide and the incoming amino acid delivered by the aa-tRNA, respectively. For most pairs, the ribosome itself acts as the catalyst to make peptide bonds rapidly and the reaction does not require auxiliary translation factors. However, the actual rates of peptide bond formation may vary considerably between amino acid pairs, with Pro being a particularly difficult substrate both in the A and P site. In bacteria, Pro incorporation in certain contexts is strongly accelerated by a dedicated elongation factor, EF-P [17,18], which enhances peptide bond formation between two consecutive Pro residues by 16-fold [17] (Fig. 1A). Synthesis of longer poly(Pro) sequences is fully dependent on EF-P: In the absence of the factor, ribosomes stop synthesis at PPP and PPG sequences after incorporating two Pro residues, whereas addition of EF-P alleviates stalling and allows the ribosomes to continue translation [17,18]. EF-P is not essential in all bacteria, but strains without EF-P generally show reduced growth and increased susceptibility to stress environments such as low osmolality, detergents, presence of antibiotics and increased temperatures [19–22]. EF-P is essential for Mycobacterium tuberculosis and Neisseria meningitis [23,24]. In eukaryotic cells, an auxiliary factor called eIF5A fulfills the same function [25,26]. In most organisms, both EF-P and eIF5A are posttranslationally modified by unique mechanisms that differ between pro- and eukaryotes and even among different groups of bacteria (Table 1). Modifications are crucial for EF-P function, which makes them potential targets for antimicrobial drug design. In this review, we summarize the recent progress in understanding the function of EF-P/eIF5A, their posttranslational modifications, and illustrate how the enzymes installing the modifications can be used for the search of novel antimicrobials.

Structure and function of EF-P and eIF5A

Structures of the factors

EF-P is a small protein that consists of 184–195 amino acids (molecular weight of 20–22 kDa) depending on the organism [27]. Its sequence is highly conserved among bacteria. EF-P comprises three β-barrel domains (I, II and III) decorated with short 310-helices in domain I and domain II (Fig. 1B). Domain I has six β-strands, whereas domain II and domain III have five β-strands each. The Cα-scaffolds of domains I and II are almost perfectly superimposable (with a root mean square deviation (RMSD) of only around 1.7 Å), suggesting that the two domains have evolved through a duplication event [28], possibly driven by the stabilization of EF-P binding to ribosomes through domain II-tRNA interactions (see below). The RMSD between domains III and I is 2 Å and between III and II 2.5 Å; the N-terminal part of domain III shares some similarity to domain I and the C terminal part to domain II.
EF-P has an ξ-shape structure similar to that of tRNAs [28]. The net negative charge of EF-P (isoelectric point 4.90) also resembles that of tRNAs. E. coli EF-P carries a posttranslational modification, β-Lys, at Lys34, which is part of a conserved PKGK motif at the tip of domain I. The place of the modification resembles the CCA end of a tRNA, further emphasizing the molecular mimicry between EF-P and tRNA. The EF-P concentration in E. coli cells is 2–4 μM depending on growth conditions [29].

Some Proteobacteria encode an EF-P-like protein in addition to EF-P (Table 1). The function of EF-P-like proteins is not known. EF-P-like proteins share only about 30% sequence similarity with EF-P, but are almost 90% similar between themselves. AlphaFold predictions suggest that EF-P-like proteins have a three-domain structure like EF-P (EF-P-like proteins has the highest sequence similarity to domain III of EF-P). AlphaFold predictions suggests that EF-P-like proteins have a three-domain structure like EF-P (EF-P-like proteins has the highest sequence similarity to domain III of EF-P). The function of EF-P-like proteins is not known. EF-P-like proteins are predicted not to be modified; 5-AP: 5-aminopentanol; Rha: Rhamnose; 8-Lys: β-Lys-hydroxyl.

**Table 1**

<table>
<thead>
<tr>
<th>Phylum</th>
<th>Order</th>
<th>Species</th>
<th>Modification</th>
<th>Tip motif</th>
<th>EF-P-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tenericutes</td>
<td></td>
<td>Mycoplasma pneumoniae</td>
<td>no</td>
<td>KPGKGP</td>
<td></td>
</tr>
<tr>
<td>Chlamydiae</td>
<td></td>
<td>Chlamydia trachomatis</td>
<td>no</td>
<td>KPGKGP</td>
<td></td>
</tr>
<tr>
<td>Actinobacteria</td>
<td></td>
<td>Mycobacterium tuberculosis</td>
<td>no</td>
<td>KPGKGP</td>
<td></td>
</tr>
<tr>
<td>Firmicutes</td>
<td>Bacilli</td>
<td>Clostridoides difficile</td>
<td>no</td>
<td>KPGKGP</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Listeria monocytogenes</td>
<td>5-AP</td>
<td>KPGKGN</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lactobacillales</td>
<td>Streptococcus agalactiae</td>
<td>5-AP</td>
<td>KPGKGN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Streptococcus pneumoniae</td>
<td>no</td>
<td>KPGKGN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Enterococcus faecalis</td>
<td>no</td>
<td>KPGKGN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Enterococcus faecium</td>
<td>no</td>
<td>KPGKGN</td>
<td></td>
</tr>
<tr>
<td>Proteobacteria</td>
<td>ε</td>
<td>Campylobacter jejuni</td>
<td>no</td>
<td>KPGKGP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β</td>
<td>Neisseria gonorrhoeae</td>
<td>Rha</td>
<td>KGGRSS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>γ</td>
<td>Neisseria meningitidis</td>
<td>Rha</td>
<td>KGGRSS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pseudomonas aeruginosa</td>
<td>Rha*</td>
<td>KGGRNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Legionella pneumophila</td>
<td>KPGKQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Acinetobacter baumanii</td>
<td>KPGKQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aeromonas hydrophila</td>
<td>KSGRNS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enterobacterales</td>
<td>Vibrio cholerae</td>
<td>8-Lys</td>
<td>KPGKQ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Haemophilus influenzae</td>
<td>8-Lys</td>
<td>KPGKQ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Morganella morganii</td>
<td>8-Lys</td>
<td>KPGKQ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteus vulgaris</td>
<td>8-Lys</td>
<td>KPGKQ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Providencia stuartii</td>
<td>8-Lys</td>
<td>KPGKQ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Citrobacter freundii</td>
<td>8-Lys</td>
<td>KPGKQ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Escherichia coli</td>
<td>8-Lys</td>
<td>KPGKQ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Klebsiella pneumoniae</td>
<td>8-Lys</td>
<td>KPGKQ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shigella flexneri</td>
<td>8-Lys</td>
<td>KPGKQ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Enterobacter huxanensis</td>
<td>8-Lys</td>
<td>KPGKQ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Salmonella enterica</td>
<td>8-Lys*</td>
<td>KPGKQ</td>
<td>+</td>
</tr>
</tbody>
</table>

1 Ref. [109]; 2Modification has been predicted based on the presents of the modification enzymes in the genome.

**Binding to the ribosome**

EF-P binds to the E site of the ribosome [35,36] in close proximity of the P-site tRNA. Domain III is oriented towards the mRNA, whereas domain I is aligned with the CCA end of the tRNA pointing towards the peptidyl transferase center (Fig. 1B). In E. coli EF-P, the β-Lys modification, which protrudes into the peptide transferase center of the ribosome, contacts the CCA end of the P-site tRNA as well as the conserved A2439 of the 23S rRNA [35]. Its position in the E site implies that EF-P requires an empty E site to bind to the ribosome. During ongoing translation, the E site is mostly occupied by a deacylated tRNA moving out of the ribosome as a result of tRNA–mRNA translocation at the end of every elongation cycle. However, if peptide bond formation is blocked, the E site becomes free for EF-P binding. Thus, one key regulator of EF-P binding to the ribosome is simply the availability of a free E site, which is a hallmark of ribosome stalling during decoding or peptide bond formation.

In addition to the open binding site on the ribosome, there are also more specific recognition elements that affect the catalytic proficiency of EF-P on the ribosome. An important recognition element for EF-P binding is the ω-loop of the P-site tRNA, which in the structure of the ribosome–tRNA复合体, EF-P complex is found within hydrogen bonding distance of EF-P Asp69 [35] (Fig. 1B). The three isoacceptors of tRNAγ and the two isoacceptors of tRNAδ share the same conserved 9-nucleotide ω-loop sequence (AGCCUGGDA) with a stable 4-base pair π-stem. This ω-loop identity is essential for EF-P rescue of poly(Pro)-mediated translation.
stalling, and replacement of the 9-nucleotide β-loop of tRNAPro with that of tRNASer or tRNAAsn abolished EF-P activity [37]. Vice versa, when the β-loops of tRNAAsn or tRNAAla were replaced with that of tRNAPro, EF-P was able to rescue stalling of complexes formed with those hybrid tRNAs charged with Pro [37]. In addition to its interactions with the tRNA β-loop, conserved EF-P residues Tyr183 and Arg186 are close to A42 of the P-site tRNA and G1338 of the 16S rRNA. Mutations of Tyr183 to Ala and Arg186 to Ala render EF-P inactive [35]. The conserved loop in domain III formed by the Gly144-Asp145-Thr146 (GDT) motif is in proximity of the mRNA and may interact with cytosine nucleotides of the CCG Pro codon in the E site [35].

In comparison to EF-P, eIF5A is shifted away from the P-site tRNA and only Arg27 is found within hydrogen bond distance to G4-CS of the tRNA acceptor stem backbone (Fig. 1B). However, hypusine forms numerous interactions by contacting the phosphate group of the tRNA C75-A76 backbone [30,36] as well as the phosphate group between U2807 and A2808 of 25S rRNA, which is similar to the interaction of the ε-Lys modification of EF-P binding to A2439 [30]. eIF5A domain II is inserted between two domains of uL1 and possibly can induce a closed conformation of uL1, thereby favoring a particular conformation of the E site and stabilizing eIF5A binding. The N-terminal extension of domain I protrudes between uL1 and eL2 the same direction as domain II [30].

### Stalling sequences that require EF-P/eIF5A action

Poly(Pro) sequences are more abundant in eukaryotes than in prokaryotes. In *Homo sapiens*, 64% of proteins in the proteome contain at least one PP sequence, 20% have PPP and ~8% have PPP or longer poly(Pro) sequences. Human mitochondria-encoded proteins contain 14 PP sequences and one protein (a subunit of cytochrome c oxidase) contains a PPP sequence. In comparison, 31% of *E. coli* proteins have PP sequences, 2% PPP and 0.3% PPPP (14 proteins) sequences; many proteins have more than one poly(Pro) sequence. However, in some bacteria poly(Pro) sequences are more prevalent than in others. *M. tuberculosis* and *Pseudomonas aeruginosa* have a large number of PP and PPP sequences, with ~2% of *M. tuberculosis* proteins containing a poly(Pro) sequence of four or more consecutive Pro residues. In contrast, *Staphylococcus aureus*, *Clostridium botulinum* and *Bacillus subtilis* have almost no PPP-containing proteins. In *E. coli* poly(Pro) sequences are often found within the first 50 amino acids after the translation start, where they may act as cis-acting element regulating the rate of translation or polysome loading of particular mRNAs. Poly(Pro) sequences are also prevalent in the inter-domain regions of multi-domain proteins and downstream of transmembrane helices, where they can slow down translation to ensure efficient co-translational folding and protein membrane insertion [39]. Given the difficulties of protein synthesis with several consecutive Pro residues, it is remarkable that such sequences were not eliminated during evolution and instead a dedicated factor has evolved to solve the problems. This could be explained by the existence of poly(Pro) sequences that could not be replaced without a functional loss of a universally essential protein. In fact, the active site of Val-tRNA synthetase (ValRS) has a universally conserved, invariant PPP sequence that is important for efficient charging of tRNA synthetase (ValRS) has a universally conserved, invariant PPP sequence that could not be replaced without a functional loss [39].

In addition to PPP, PPPP, and PPG sequences, some XPP/PPX combination (where X denote any amino acid) can result in translation pausing that is rescued by EF-P in vivo in a LacZ reporter assay or in a reconstituted *E. coli* translation system [41]. Most of the XPP/PPX sequences reduce protein synthesis by less than 2-fold, while DPP, PPP, PPG and PPN have stronger (4- to 8-fold) effects. When these sequences were combined in quadruplets, for example to DPPN, the severity of stalling increased, reaching levels comparable to the one measured for the PPP sequence (20-fold) [41]. This broadens the range of proteins that require EF-P for their production, e.g. from about 100 with PPP motifs to approximately 1500 proteins containing different XPP/PPX sequences in *E. coli* [41].

Compared to EF-P, which acts on Pro-containing sequences, eIF5A helps to alleviate pauses not only on Pro-containing motifs such as GPP, DPP, PPP, PPG, but also on a range of other sequences, e.g. DDK, PDK, or CFK. In contrast to EF-P, eIF5A facilitates translation termination by increasing the rate of peptide release by eRF1 (by 17-fold in vitro). eIF5A-deficient cells accumulate ribosomes at stop codons, whereas in EF-P-deficient cells no accumulation of ribosomes at stop codons is observed [25]. eIF5A does not form any specific interactions with the β-loop of the P-site tRNA, suggesting that its binding to ribosomes may be independent of the P-site tRNA. Nevertheless, the affinity of eIF5A to the ribosome is very high (Kd = 9 nM), suggesting that at the cellular concentration of the factor (8–15 µM is sufficient to occupy every ribosome (~1 µM)) [42]. Similarly to EF-P, eIF5A stabilizes the CAA end of the P-site tRNA in a productive conformation, helping reactions with slowly reacting amino acids to proceed [30].

It remains unclear how mitochondrial ribosomes solve the problems of poly(Pro) synthesis and whether they employ eIF5A for this purpose. eIF5A lacks a mitochondrial targeting sequence, but considering its small size, the factor may be passively imported into mitochondria. Another possibility is that an alternative start codon in the eIF5A mRNA may drive production of an eIF5A isoform with a 30 amino acid-long N-terminal extension that may contain an mitochondrial import sequence [43]. In fact, mass spectrometry identified eIF5A in mitochondria of human and yeast cells [44,45] and eIF5A seems to play a functional role in mitochondria, as it is upregulated in some cancer cells and the inhibition of eIF5A leads to mitochondria silencing [46]. Vice versa, overexpression of eIF5A in cardiac muscle cell line results in increased ROS production and release of cytochrome c into cytosol [47].

Another interesting observation is that also viral mRNAs, which are translated by the host machinery and are tuned for rapid virus replication, often have high percentage of poly(Pro) motifs, and thus have to hijack eIF5A to ensure rapid synthesis of viral proteins. Especially viruses from the *Herpesviridae* family (human *Herpesvirus* HHV-1, 2, 8) have high amounts of poly(Pro) containing proteins, with 60–90% of proteins containing a PP motif, 25–40% with PPP, and 2–15% with PPPP, suggesting that translation of these viral mRNAs requires eIF5A.

### Mechanism of EF-P action

To understand how EF-P works, it is important to discuss why ribosomes stall during synthesis of poly(Pro) sequences. Pro is unique among the 21 proteinogenic amino acids by having a cyclic side chain, which restricts the possible conformations of Pro itself and of the neighboring amino acids. Peptide bonds between two prolines can adopt isoenergetic cis and trans states, in contrast to all other amino acids, which adopt the energetically favored trans conformation. These unique Pro characteristics appear to be particularly disadvantageous for ribosome-catalyzed peptide bond formation, because the uncatalyzed reaction in solution is as fast with Pro or Pro analogs as with other amino acids [48,49]. Linear free energy relationships of the reaction on the ribosome and in solution using Pro analogs that differed in their ring structure, preferences towards *cis*/*trans* conformers and *exo*/*endo* puckers, as well as their electronic properties suggested that the positioning of Pro-tRNAPro in the peptidyl transferase center is the major determinant for the slow reaction [49]. The unfavorable positioning on the ribosome exaggerates the differences in the intrinsic reactivity, likely by altering the trajectory for the nucleophilic attack. Upon addition of EF-P, the activation energy decreases by the same 2.5 kcal/mol independent of the Pro analog [49]. Comparison of the activation parameters of the reaction with and without EF-P suggests that the acceleration is due to a favorable entropic effect, whereas the enthalpy change is unfavorable. The absence of an enthalpic contribution to catalysis argues against the possibility that EF-P might act by donating functionally active groups, which would predict favorable enthalpic effects [49]. The entropic character of EF-P-assisted catalysis is consistent with the notion that the factor induces a more favorable positioning of the P-site substrate in the peptidyl...
transferase center (Fig. 1A). In fact, cryo-EM structures suggest that the favored conformation of the poly(Pro)-containing nascent chain is incompatible with the peptide exit tunnel of the ribosome and leads to destabilization of the peptidyl-tRNA in the P site [35]. Binding of EF-P stabilizes the P-site tRNA, particularly via interactions between its modification and the CCA end, thereby enforcing an alternative conformation of the poly(Pro)-containing nascent chain, which allows a favorable substrate geometry for peptide bond formation.

In addition to enhancing the formation of peptide bonds, EF-P helps the ribosome to maintain the correct reading frame. Spontaneous

Fig. 2. Posttranslational modifications of EF-P and eIF5A. Modifications are shown in cyan. Left panel, modified residues, Lys or Arg, are shown in black. Lys-residues that are modified with R-β-Lys are also modified with a hydroxyl group (red). Right panel, putative interaction of the modifications with the CCA end of the P-site tRNA. Interactions of β-Lys and hypusine are deduced based on cryo-EM structures [30, 35]. Rhamnose and S-aminopentanol were modeled by replacing β-Lys modification with energy minimization. Putative hydrogen bonds are shown in purple (for H-bond predictions, 2.1 to 3.9 Å distance and a 30° variation was allowed to account for uncertainties due to limited resolution of the cryo-EM structures used for modeling).
frameshifting occurs on mRNA sequences where the same tRNA reads the codons in the 0- and in an alternative frame. These slippery sequences can be made by runs of consecutive Pro codons (CCN), where N is any nucleotide), such as CCC CCU. Another hallmark of frameshifting is ribosome stalling, which opens the time window in which tRNAs can re-equilibrate between the 0- and ±1-frame codons according to their thermodynamic preferences (e.g. if the same codon is found in the 0- and in an alternative frame, the frameshifting efficiency can be as high as 50%) [50]. Slippery CCCCC sequences are particularly prone to ±1 frameshifting when read by the tRNA^Pro isoacceptor with the anticodon GGG, whereas EF-P suppresses frameshifting errors [51]. The molecular mechanism of frameshift control by EF-P is not entirely understood, but in the simplest model EF-P can bind to the ribosome because the E site is free and it reduces the dynamic fluctuations of the tRNA between the 0- and ±1-frame codons by binding to the tRNA^Pro D loop. At the same time, acceleration of peptide bond formation closes the kinetic window for frameshifting, thereby safeguarding ribosomes upon traversing the slippery Pro codons.

**EF-P modifications**

To date, three types of posttranslational EF-P modifications have been identified in bacteria, all installed at the same functional end of the molecule carrying a conserved motif at the tip of domain 1, but different among bacterial phila (Table 1). EF-P is modified by β-Lys in *E. coli* and *Salmonella enterica* [52–54], with rhamnose in *P. aeruginosa* and *N. meningitidis* [22,24,55], and with 5-aminopentanol in *B. subtilis* and *S. aureus* [56,57] (Table 1 and Fig. 2). In the phylum Actinobacteria, which includes *Corynebacterium*, *Mycobacterium* and *Streptomyces*, EF-P does not have a posttranslational modification [58]. The modification state of EF-P in many other bacteria is not known and it is thus possible that other unique modifications will be discovered in the future. Alignment of EF-P with different modifications suggest that despite their chemical diversity, the modifications serve the same function, i.e., stabilization of the CCA end of the P-site tRNA (Fig. 2). This was also demonstrated by the rescue of an *E. coli* strain, which lacks the β-Lys modification enzymes, with rhamnosylated EF-P [24]. In the following we describe how different types of the modifications and installed and how these enzymes can be used as drug design targets.

### β-Lys-hydroxyl modification

The β-Lys-hydroxyl modification occurs in three steps. In the first step, S-α-Lys is isomerized to R-β-Lys catalyzed by EpnB [59]. Than EpnA modifies EF-P at the ε-amino group of Lys34 (E. coli numbering) with R-β-Lys utilizing one molecule of ATP [54]. In the final step, EpnC hydroxylates Lys34 [60]. Hydroxylation is not essential for EF-P activity [17] and its prevalence in vivo is not clear. The structure of EpnA resembles that of Lys-tRNA synthetase (LysRS), the enzyme that charges tRNA^Lys with Lys at the cost of ATP hydrolysis, except that EpnA lacks the tRNA anticodon-binding domain of LysRS. *E. coli* EpnA and LysRS are superimposable with an RMSD of only 1.4–1.7 Å on the Cα backbone. The active sites of EpnA and LysRS are highly homologous and the binding mode of EF-P to EpnA is very similar to that of tRNA to LysRS. EF-P Lys34 binding to EpnA mimics the 3'-terminal A of the tRNA’s CAA end on LysRS. The K_M of EF-P binding to EpnA (3.2 μM) is similar to that of tRNA^Lys to LysRS (1.9 μM) [61,62]. Nevertheless, EpnA transfers β-Lys specifically to EF-P and cannot ligate β-Lys to a tRNA [61]. EpnA discriminates against (R)-α-Lys and (S)-α-Lys due to their much lower affinity compared to R-β-Lys (K_M ~20 mM and 5 mM vs 0.2 mM, respectively). An Ala298 to Gly mutation in EpnA reverses the substrate selectivity in favor of (S)-α-Lys [63]. While the affinities of EpnA and LysRS to their amino acid substrates differ only 5-fold (0.2 mM and 0.04 mM, respectively), EpnA is a very inefficient enzyme with a k_cat/K_M = 3 × 10^8 M^-1 s^-1 compared to LysRS (k_cat/K_M = 10^10 M^-1 s^-1) [64].

The functional importance of the β-Lys modification has been demonstrated in vivo and in vitro. In *E. coli* deletion of EpnA decreases bacterial growth to the same extent as the deletion of EF-P [54], which is essential for growth at elevated temperatures [65], affects the growth rate at physiological temperature (37°C), but is not essential at 25°C [21]. In the pathogenic *S. enterica*, the knockout mutant of EpnA has attenuated virulence and shows reduced growth rate, stress resistance, and pyruvate oxidase activity. An EpnA-deficient strain is 10,000 times attenuated in a mouse model when inoculated orally and 1000 times when inoculated interperitoneally [52]. EpnA mutants also had increased susceptibility to a wide range of antibiotics, including antimicrobial peptides, lipophilic chelators, detergents and inhibitors of cell wall and protein synthesis [52]. Quantitative analysis of the catalytic activity of non-modified EF-P in rMPPG synthesis in vitro shows a decrease in the catalytic rate k_cat (6-fold) and the catalytic proficiency k_cat/K_M (100-fold) compared to fully modified EF-P, indicating that the modification is important both for EF-P binding to the ribosome and for catalysis [17]. In the cryo-EM structures of ribosomes stalled during poly (Pro) synthesis, the CCA end of the peptidyl-tRNA and the peptidyl moiety are better resolved with EF-P than without the factor and the local resolution indicated less flexibility of the P-site tRNA in the presence of EF-P [35] (Fig. 2). Molecular dynamic simulations suggested that in the absence of EF-P or when EF-P was not modified, the CCA end the P-site peptidyl-Pro-tRNA adopts a conformation that shifts Pro by 2 Å away from the A-site tRNA and the CCA end showed altogether increased conformational motion, explaining why Pro incorporation is more efficient with than without EF-P [35].

### Rhamnose modification

In *P. aeruginosa* EF-P is modified by an enzyme called EarP. EarP uses deoxythymidine diphosphate-ε-rhamnose (dTDP-ε-Rha) as a substrate for EF-P modification. dTDP-ε-Rha synthesis is part of the pathway to incorporate ε-Rha into bacterial cell wall. dTDP-ε-Rha is synthesized from glucose-1-phosphate and deoxythymidine triphosphate (dTTP), involving four enzymes: glucose-1-phosphate thymidyltransferase (RlmA), dTDP-ε-glucose 4,6-dehydratase (RlmB), dTDP-deoxy-ε-xylo-4-hexulose 3,5-epimerase (RmCl) and dTDP-6-deoxy-ε-lyxo-4-hexulose reductase (RmID) [66–69]. In cell wall biogenesis, dTDP-ε-Rha is a component of a critical linker, ε-N-acetylglucosamine-ε-Rha disaccharide, which connects the galactan moiety of arabinogalactan to peptidoglycan via a phosphodiester linkage [70]. In *Streptococcus pyogenes* Rha-containing cell wall polysaccharides comprise half of cell wall mass [71].

There are several structures of EarP from *P. aeruginosa* [72], an apo structure, as well as structures of EF-P with TDP, TDP-Rha, or EF-P (PDB ID: 6J7J, 6J7L, 6J7K, 6J7M). Also available are apo and ligand-bound structures of EarP from *P. putida* [73] (PDB ID: 5NV8) and *N. meningitidis* [74] (PDB ID: 5WXJ, 5WXK, 5WXX). *P. aeruginosa* EarP has 52% similarity with *N. meningitidis* and 68% with *P. putida* protein. EarP is a member of the glycosyltransferase superfAMILY B and consists of two tandem Rossmann-fold domains separated by a deep cleft. The N-terminal domain (NTD) consists of four parallel β-sheets and two β-hairpins formed on each side and surrounded by four α-helices and a 310-helix. The C-terminal domain (CTD) contains five parallel β-sheets with an antiparallel strand, surrounded by ten α-helices and one 310-helix [72].

Glycosyltransferases transfer sugars from donors to a variety of molecules including proteins, metabolites, lipids and nucleic acids. Unlike other glycosyltransferases, EarP is substrate-specific and glycosylates Arg32 of EF-P only. EarP recognizes domain I of EF-P and Arg32 through multiple side-chain interactions along both the sheets of the two proteins. The affinity of EarP to EF-P is in the sub-micromolar range [72,74] and the high specificity is achieved by recognition of both the shape of EF-P and the amino acid sequence. TDP-Rha binds to EarP at the cleft between the NTD and CTD, and the affinity to TDP-Rha (6.3 μM) is five times higher than to TDP (31.4 μM) [72]. Thymine, deoxyribose and pyrophosphate interact with the CTD, while Rha interacts with both the
EF-P induces a conformational change of the Rha ring. EarP Asp16 acts as a specific accommodation of other bases [72,74]. The Rha binding pocket is tight and specific for the chair conformation of the Rha ring, which forms van der Waals contacts with surrounding amino acids [72-74]. Binding of EF-P induces a conformational change of the Rha ring. EarP Asp16 acts as a general base that activates Arg32. The $\gamma$-nitrogen of activated Arg32 nucleophilically attacks the C1 atom of the Rha ring accomplishing the inverting rhamnosylation, whereas the dTDP moiety leaves [74].

The deletion of EarP in *P. aeruginosa* or *N. meningitis* strongly inhibits growth comparable to the deletion of EF-P itself. The mutants show a decrease in virulence and motility and an increased susceptibility to antibiotics targeting cell wall synthesis [22,24,55]. This is important because *P. aeruginosa* and *N. meningitis* are pathogenic bacteria that often have advanced antibiotic resistance mechanisms, causing hospital-acquired infections such as ventilator-associated pneumonia, various sepsis syndromes and bacterial meningitis [75].

In some cases, installing the modification may regulate the life cycle of a bacterium. For example, *Micavibrio aeruginosavorus*, which is an epibiotic predator bacteria that preys on *P. aeruginosa* [76], encodes EarP, the modification enzyme installing the rhamnose modification, but not the enzymes required for rhamnose synthesis [77,78]. One potential scenario is that when attacking its prey, *M. aeruginosavorus* disrupts the *P. aeruginosavorus* membrane, which frees intracellular rhamnose from the host cells, enabling the predator to install the EF-P modification. This results in synthesis of a subset of 128 poly(Pro)-containing proteins that are required for the *M. aeruginosavorus* growth phase [77,78]. This is the only well-documented example where EF-P modification plays a regulatory role responding to environmental cues.

### 5-aminopentanol modification

In *B. subtilis* and *S. aureus* EF-P is modified with 5-aminopentanol at Lys32 (B. subtilis numbering) [56,57]. The modification pathway is less well studied compared to $\gamma$-Lys, Rha and hypusine modifications. The 5-aminopentanol modification proceeds through several intermediates: hydroxypentenone, pentenone, 5-aminopentenone and 5-aminopentanal. The exact mechanisms and the enzymes responsible for each step are not known [56], but six genes – *yrbB*, *gsAB*, *ymlf*, *yaAO*, *yfKA* and *ywGL* – were identified as potential modification enzymes by forward genetic screen analysis and mass spectrometry. Of these genes, Ymlf is important for the last step, i.e. the reduction of EF-P-attached 5-aminopentanone to 5-aminopentanol. Ymlf is homologous to FabG, which catalyzes the NADPH-dependent reduction of $\beta$-ketocetyl-ACP substrates to $\beta$-hydroxyacyl-ACP products, the first reductive step in the elongation cycle of fatty acid biosynthesis [79]. Interestingly, in the absence of Ymlf accumulation of EF-P modified with 5-aminopentanone has inhibiting activity on *B. subtilis* swarming motility. The mutation of the conserved Lys32 in *B. subtilis* EF-P, which prevents post-translational modification, restores the swarming motility in the absence of Ymlf, showing that 5-aminopentanone modified EF-P might be toxic for cells [80].

### Non-modified EF-P

*C. glutamicum* EF-P does not carry a post-translational modification. Its Lys at the tip of domain I is important for EF-P function, but it can be exchanged to Arg (which cannot be modified) with only a slight loss in activity. In *Actinobacteria* the Lys residue is a part of a PGKGP motif, which is unique for the phylum. The PG sequences in the palindromic enhance the stability of the loop, possibly compensating for the absence of a modification [81]. Based on the available structures, the tip cannot reach the CCA end of the P-site tRNA, raising the question of how unmodified EF-P works and whether other interactions compensate for the lack of modification [82]. Notably, EF-P from *Mycobacterium tuberculosis* also contains the same PGKGP motif and is predicted to be non-modified; given the medical relevance of this pathogen, it would be important to check this prediction, e.g. by mass spectrometry.

### Hypusine modification of eIF5A

The hypusine modification is unique for eIF5A and is formed in a two-step reaction. In the first step, 4-aminobutyl is transferred from the polyamine spermidine to the terminal amino group of Lys by NAD-dependent deoxyhypusynine synthase (DHS), producing deoxyhypusine. In the second step, deoxyhypusine hydroxylase (DOHH) adds a hydroxy group, producing hypusine. DHS forms a functional tetramer and binds spermidine with a $K_d$ of 4 $\mu$M [83]. The DHS-catalyzed reaction can be divided into four steps: (1) spermidine is oxidized resulting in NADH and dehydrosperrmidine formation; (2) dehydrosperrmidine is cleaved producing dianimo propane (DAP) and the butyloamine moiety; (3) butyloamine is attached to the catalytic Lys residue of DHS through an imine bond; and (4) butyloamine is transferred onto the eIF5A Lys [83]. DOHH is a non-heme diiron enzyme featuring highly conserved His-Glu motifs that offer two potential iron coordination sites. DOHH$_{peroxo}$ displays extreme longevity in comparison to other peroxo-diiron species, providing the possibility of a structural study of the active enzyme [84]. In yeast, eIF5A and DHS are essential, while the deletion of the gene for DOHH only reduces growth. All three genes are essential in higher eukaryotes [85].

### EF-P/eIF5A in bioengineering and medicine

#### EF-P in protein production

The stabilization of the P-site peptidyl tRNA in a productive conformation makes EF-P an interesting tool for a wide range of applications in bioengineering. Co-expression of EF-P increases expression yields of proteins that contain poly(Pro) sequences, which can be utilized when upscaling production of industrially relevant enzymes [82].

Furthermore, EF-P can be a key factor to produce proteins with novel physicochemical properties and biological functions. One of the approaches to design new proteins is through incorporation of unnatural amino acids. Due to the progress in genetic reprogramming, it is now possible to mischarge tRNAs with unnatural amino acids and deliver them to the ribosome. However, once on the ribosome, such unnatural substrates are often poor substrates in the peptidyl transfer reaction, which creates a major bottleneck for efficient protein production. A hallmark of inefficient synthesis is ribosome stalling prior to the peptidyl bond formation step, which is the state of the ribosome that allows recruitment of EF-P. In fact, EF-P can enhance the incorporation of a range of unnatural or non-proteogenic amino acids into polypeptide chains such as $\alpha$-amino acids [86,87], $\beta$-amino acids [87], or 2-aminobenzoic acids [88]. In some cases, EF-P increased the production by 2–8-fold, while in other cases, in particular when several sequential unnatural amino acids were incorporated, EF-P is essential for synthesis of these synthetic proteins [86,87].

#### Screening for potential inhibitors of EarP

Inhibition of EF-P appears to be an attractive anti-bacterial strategy, because the virulence of some clinically relevant pathogens, such as *S. enterica*, *P. aeruginosa*, and *N. meningitis*, depends on proteins that require EF-P for their synthesis [52,54,55]. Unfortunately, EF-P itself is not a promising target for inhibition with small compounds, because its three-dimensional structure is built of convex surfaces, making it difficult to find or design compounds that bind to EF-P and inhibit its interaction with the ribosome. Furthermore, the similarity of EF-P and eIF5A structures and the conservation of the EF-P/eIF5A binding sites between bacterial and mammalian ribosomes, make it unlikely to find a specific enough inhibitor for bacterial but not for eukaryotic cells. However, EF-P-modifying enzymes are bacteria- and substrate-specific, which offers interesting perspectives for drug design.
In the following we illustrate a potential approach for searching inhibitors of EarP, which was chosen because EarP is present in many bacterial pathogens and its mechanism of action is well studied [22,24,55]. This is even more supported by the fact that drug-resistant P. aeruginosa and N. meningitis are listed as critical and high priority pathogens for new antibiotic discovery on the WHO priority list. Further advantages of EarP are that it does not have homologs in eukaryotes and that Rha is uniquely used by prokaryotes. EarP structures from P. aeruginosa and N. meningitis are available. In comparison, the pathway for the 5-aminopentanol modification is not known in detail and no structures of the modifying enzymes are available [56], which makes the screen development less straightforward. Targeting EpmA appears to be less promising because of potential off-target effects due to its similarity to eukaryotic LysRS [54].

EarP from P. aeruginosa uses TDP-rhamnose to modify EF-P with rhamnose, thereby releasing TDP (Fig. 3). Introduced thymidylate kinase (ThK) from E. coli transfers a phosphate from TDP onto ADP, producing ATP. ATP is then utilized by a luciferase to produce chemiluminescence, which is used as a readout. All proteins for this assay can be purified and at optimized concentrations the luminescence signal is stable for performing an automated high throughput screening in deep-well format (Fig 3). Such assays provide a straightforward opportunity to develop new antibiotics against some of the most challenging bacterial pathogens.

eIF5A disease association

Eukaryotes have two isoforms of eIF5A, eIF5A1 and eIF5A2, both of which are hypusinated. Both forms are upregulated in tumor cells such as lung and colorectal adenocarcinoma, glioblastoma, pancreatic adenocarcinoma, cervical cancer and chronic myeloid leukemia. As the upregulation of eIF5A2 is particularly striking, it was proposed to be an adenocarcinoma, cervical cancer and chronic myeloid leukemia. As the bacterial pathogens and its mechanism of action is well studied [22,24,55], this is even more supported by the fact that drug-resistant P. aeruginosa and N. meningitis are available. In comparison, the pathway for the 5-aminopentanol modification is not known in detail and no structures of the modifying enzymes are available [56], which makes the screen development less straightforward. Targeting EpmA appears to be less promising because of potential off-target effects due to its similarity to eukaryotic LysRS [54].

EarP from P. aeruginosa uses TDP-rhamnose to modify EF-P with rhamnose, thereby releasing TDP (Fig. 3). Introduced thymidylate kinase (ThK) from E. coli transfers a phosphate from TDP onto ADP, producing ATP. ATP is then utilized by a luciferase to produce chemiluminescence, which is used as a readout. All proteins for this assay can be purified and at optimized concentrations the luminescence signal is stable for performing an automated high throughput screening in deep-well format (Fig 3). Such assays provide a straightforward opportunity to develop new antibiotics against some of the most challenging bacterial pathogens.

eIF5A disease association

Eukaryotes have two isoforms of eIF5A, eIF5A1 and eIF5A2, both of which are hypusinated. Both forms are upregulated in tumor cells such as lung and colorectal adenocarcinoma, glioblastoma, pancreatic adenocarcinoma, cervical cancer and chronic myeloid leukemia. As the upregulation of eIF5A2 is particularly striking, it was proposed to be an oncogene and a tumor marker [89], eIF5A depletion and knockdown of its modifying enzymes slows growth of cancer cells and reduces metastasis for a range of cancer types [89–91]. In colorectal cancer, hypusinated eIF5A promotes growth directly by increasing expression of the Myc oncogene, which has five poly(Pro) motifs (PPA, PPL, PPT, PPH, and APP) [90]. Interestingly, Myc activates transcription of ornithine decarboxylase, the first enzyme in the polyamine biosynthesis pathway enabling production of more putrescine for hypusine modification [92]. Another protein associated with tumor formation, p53, has also five poly(Pro) motifs (PPL, APP, PQQ, PPPQ, and PPE); inhibition of eIF5A hypusination suppresses p53 translation [93]. Other effects of eIF5A on p53 translation may be indirect, for example, eIF5A increases expression of HIF-1α, overexpression of which allows cancerogenic cells to survive in a hypoxic environment. Because HIF-1α does not have poly(Pro) motifs, eIF5A most probably regulates its abundance indirectly through proteins required for the proteasomal degradation of HIF-1α [94]. Overexpression of eIF5A in pancreatic cancer cells is also linked to abnormal activation of the sonic hedgehog (shH) protein pathway. Inhibition of eIF5A expression – and hence of the shH signaling pathway – suppresses pancreatic cell proliferation and tumor growth [91]. There is also an interesting link between eIF5A and diabetes, as a tissue-specific Dhs deletion in mice shows a significantly worse glucose tolerance under high fat diet compared to control animals and an impairment in insulin secretion [95].

Inhibition of the eIF5A modification has a potential for cancer treatment applications [89]. So far, only one compound, N3-guanyl-1,7-diaminoheptane (GC7) that inhibits DHS was shown to be effective in treatment of hepatocellular carcinoma, colorectal cancer and leukemia [96–99]. GC7 mimics spermidine and competes with it for the binding site on DHS (PDB ID: 6XJ1, 1RQD); one potential disadvantage of this inhibitor is its propensity for off-target interactions with other polyamine-binding molecules in the cell [99,100]. GC7 also has a potential for treatment of diabetes and inflammatory pain [99,101,102]. GC7 as well as 5-lipoxygenase inhibitor Zileuton that inhibits DOHH in vitro might act against malaria [103]. eIF5A might be also considered a promising target for antivirus therapies, as many viral proteins contain a large number of poly(Pro) sequences. For example, inhibition of eIF5A modification decreases HHV-8 replication and proliferation of the infected cell [104]. HHV-8 requires eIF5A for the synthesis of poly(Pro)-containing latency-associated nuclear antigen (LANA), which has 12 PP and 2 PPP motifs and functions as a molecular tether that ensures segregation of episomes to the progeny nuclei [104]. eIF5A is also important for the replication of the Ebola (EBOV) and the Marburg virus (MARV). Depletion of hypusinated eIF5A followed by infection with EBOV or MARV resulted in 3 log lower virus titers compared to cells with functional eIF5A [105]. A possible mechanism for such reduction is the decreased synthesis of the VP30 protein required for viral transcription, which has a stalling PPA motif in EBOV and a strong stalling PPPPPP motif in MARV [105]. The majority of all EBOV and MARV proteins contain PP motifs. In HIV, eIF5A is required for the production of functional Rev, a protein that mediates transport of viral mRNAs from the nucleus to the cytoplasm. Mutations eIF5A block virus replication [106,107]. In this case, eIF5A has a function outside of translation, as eIF5A binds Rev and mediates its transport out of the nucleus. Interestingly, nuclear eIF5A carries an acetyl group instead of hypusine, indicating a specialization of nuclear vs. cytoplasmic forms of eIF5A [108].

Perspectives

The protein synthesis machinery is universally conserved and has evolved to accept different proteinogenic amino acids in a uniform way regardless of their chemical nature. However, some protein motifs appear to be challenging for the ribosome and require an auxiliary factor to bring the synthesis to speed. The discovery of EF-P/eIF5A as specialized translation factors underscores the importance of specialized
mechanisms to rescue translation in certain contexts. While it is clear that EF-P/eIF5A are essential for synthesis of a subset of cellular proteins, it remains unclear whether the factors also take part in the regulation of translation in response to environmental cues. In biology, postranslational modifications are often used for signaling, which is mediated by enzymes acting as writers and erasers of a given modification. The writers for various EF-P modifications are known, but erasers have not been discovered so far; it would be interesting to systematically explore whether such erasers exist in nature. On the other hand, the establishment of EF-P modifications may depend on the availability of the modification substrates, which, in turn, may be regulated by metabolic pathways responding to the environment. While EF-P appears to be modified in most bacteria, it is not clear why evolution proceeded towards such diversity of modifications and whether the list of all possible modifications is already complete. Moreover, it is not clear how EF-P functions in those organisms where a modification is apparently not installed. Other unexplored territory are EF-P-like proteins, the function for which is completely unknown. The difference in bacterial and eukaryotic EF-P/eIF5A modification pathways might allow designing new specific antimicrobials. By establishing the assay for testing inhibitors of the Rha modification of EF-P, we have made first steps towards targeting specific groups of pathogens. Recent research also shows successful examples of using EF-P to incorporate unnatural amino acid analogs into proteins. Also here the effect of different modifications can be tested – and maybe more efficient chemical modifications designed – to artificially evolve EF-P molecules that are specialized in helping the ribosome to synthesize proteins unknown to nature.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data cited in this review are publicly available

Funding

This work was supported by the Max Planck Society and the Deutsche Forschungsgemeinschaft (Leibniz Prize, SFB1565, and EXC 2067/1-390729940).

References


