

Beyond post-translational modifications: Unravelling the functional principles of unmodified translation elongation factor P

Dissertation

Zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

der Fakultät für Biologie
der Ludwig-Maximilians-Universität München
vorgelegt von

Urtė Tomašiūnaitė
aus Vilnius, Litauen

München, Dezember 2024

Gutachter:

1. Prof. Dr. Kirsten Jung
2. Prof. Dr. Christof Osman

Datum der Abgabe: 11.12.2024

Datum der mündlichen Prüfung: 26.02.2025

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist

München, 11.12.2024

Urtė Tomašiūnaitė

Erklärung

Hiermit erkläre ich, *

dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist.

dass ich mich anderweitig einer Doktorprüfung ohne Erfolg **nicht** unterzogen habe.

dass ich mich mit Erfolg der Doktorprüfung im Hauptfach
und in den Nebenfächern
bei der Fakultät für der
unterzogen habe. (Hochschule/Universität)

dass ich ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen.

München, 11.12.2024

Urtė Tomašiūnaitė

*) Nichtzutreffendes streichen

Eigenständigkeitserklärung

Hiermit versichere ich an Eides statt, dass die vorliegende Dissertation mit dem Titel:

Beyond post-translational modifications: Unravelling the functional principles of unmodified translation elongation factor P

von mir selbstständig verfasst wurde und dass keine anderen als die angegebenen Quellen und Hilfsmittel benutzt wurden. Die Stellen der Arbeit, die anderen Werken dem Wortlaut oder dem Sinne nach entnommen sind, wurden in jedem Fall unter Angabe der Quellen (einschließlich des World Wide Web und anderer elektronischer Text- und Datensammlungen) kenntlich gemacht. Weiterhin wurden alle Teile der Arbeit, die mit Hilfe von Werkzeugen der künstlichen Intelligenz de novo generiert wurden, durch Fußnote/Anmerkung an den entsprechenden Stellen kenntlich gemacht und die verwendeten Werkzeuge der künstlichen Intelligenz gelistet. Die genutzten Prompts befinden sich im Anhang. Diese Erklärung gilt für alle in der Arbeit enthaltenen Texte, Graphiken, Zeichnungen, Kartenskizzen und bildliche Darstellungen.

München, 11.12.2024

Urtė Tomašiūnaitė

Affidavit

Herewith I certify under oath that I wrote the accompanying Dissertation myself.

Title: Beyond post-translational modifications: Unravelling the functional principles of unmodified translation elongation factor P

In the thesis no other sources and aids have been used than those indicated. The passages of the thesis that are taken in wording or meaning from other sources have been marked with an indication of the sources (including the World Wide Web and other electronic text and data collections). Furthermore, all parts of the thesis that were de novo generated with the help of artificial intelligence tools were identified by footnotes/annotations at the appropriate places and the artificial intelligence tools used were listed. The prompts used were listed in the appendix. This statement applies to all text, graphics, drawings, sketch maps, and pictorial representations contained in the Work.

München, 11.12.2024

Urtė Tomašiūnaitė

Stand: 26.08.2024

Contents

Eidesstattliche Erklärung	III
Eigenständigkeitserklärung.....	IV
Affidavit	IV
Contents.....	V
Nomenclature.....	VI
Abbreviations.....	VII
Publications Originating from this Thesis	IX
Contributions to Publications Presented in this Thesis.....	X
Summary.....	XI
Zusammenfassung.....	XII
1 Introduction.....	1
1.1 Protein synthesis control in bacteria.....	1
1.1.1 Translation initiation.....	2
1.1.2 Translation elongation.....	3
1.1.3 Translation termination and ribosome recycling	3
1.1.4 Challenges in translation efficiency.....	4
1.2 EF-P discovery and structural elucidation	5
1.2.1 Molecular function of EF-P.....	7
1.2.2 EF-P activation by post-translational modification	8
1.2.3 Fully active unmodified EF-Ps	11
1.2.4 Synthetic activation of EF-P.....	12
1.2.5 The diversity of motifs causing ribosome pausing.....	13
1.3 Scope of the thesis.....	14
2 Decrypting the functional design of unmodified translation elongation factor P	15
2.1 Supplemental Data Tables of Chapter 2	45
3 Versatile Dual Reporter to Identify Ribosome Pausing Motifs Alleviated by Translation Elongation Factor P.....	55
3.1 Supplemental Tables of Chapter 3	78
4 Concluding discussion and outlook.....	83
4.1 <i>R. vannielii</i> EF-P is fully functional in <i>E. coli</i>	83
4.2 Amino acids within the EF-P backbone determine the functionality of unmodified EF-Ps	85
4.3 Preferences in the rescue of distinct polyproline motifs.....	88
4.4 A tuneable plasmid based dual reporter for rapid ribosome pausing measurements	90
4.5 Outlook and future perspectives	92
References for Chapters 1 and 4.....	94
Danksagung	105

Nomenclature

All gene names are written in italics, all protein names are written with a first capital letter. Gene deletions are denoted with “ Δ ”, gene replacements are denoted with “::”.

The positions of amino acids within a protein are represented using a one-letter code, providing information about the location of each amino acid in a primary sequence and amino acid encoded (e. g. P35, proline at the position 35). The first methionine of the wild type protein is designated “1” in the amino acid sequence regardless of whether an N-terminal affinity tag is present. The information about the amino acid substitution by site-directed mutagenesis is added to the native amino acid position in a one-letter code (e.g. V50R, valine was substituted with arginine at the amino acid position 50).

The prefix “anti-” followed by the antigen name was used to denote antibodies (e.g. anti- *R. v.* EF-P antibodies).

Abbreviations

30S IC	30S initiation complex
30S PIC	30S pre-initiation complex
70S IC	70S initiation complex
aa	Amino acid
ABCF	F subfamily of ABC ATPase
A-site	Acceptor-site
aa-tRNA	Aminoacyl tRNA
CAT	Chloramphenicol acetyltransferase
CHL	Chloramphenicol
DNA	Deoxyribonucleic acid
dTDP-L-rhamnose	Deoxythymidine diphospho-L-rhamnose
a/eIF5A	Archaeal/eukaryotic translation initiation factor 5A
EarP	Protein-arginine rhamnosyltransferase
EF-G	Elongation factor G
EF-P	Elongation factor P
EF-Tu	Elongation factor Tu
EpmA	(R)- β -lysine ligase; YjeA
EpmB	Lysine 2,3-aminomutase; YjeK
EpmC	EF-P hydroxylase; YfcM
E-site	Exit-site
GDP	Guanosine diphosphate
GTPase	Guanosine triphosphatase
IF	Initiation factor
mRNA	Messenger RNA
ORF	Open reading frame
PGKGP	Pro-Gly-Lys-Gly-Pro (Proline-Glycine-Lysine-Glycine-Proline)
Pi	Inorganic phosphate
P-site	Peptidyl-site
PTC	Peptidyl transfer center
PTM	Post-translational modification
RBS	Ribosome binding site

Abbreviations

RF	Release factor
RNA	Ribonucleic acid
RNAP	RNA polymerase
RRF	Ribosome recycling factor
SD	Shine Dalgarno
sfGFP	Super folded green fluorescence protein
tRNA	Transfer RNA
tRNA ^{fMet}	Initiator transfer RNA

Publications Originating from this Thesis

Chapter 2:

Tomasiunaite U, Kielkowski P, Krafczyk R, Forné I, Imhof A, Jung K. Decrypting the functional design of unmodified translation elongation factor P. *Cell Rep.* 2024;43(5):114063.

<https://doi.org/10.1016/j.celrep.2024.114063>

Chapter 3:

Tomasiunaite U, Brewer T, Burdack K, Brameyer S, Jung K. Versatile Dual Reporter to Identify Ribosome Pausing Motifs Alleviated by Translation Elongation Factor P. *ACS Synth. Biol.* 2024, 13, 11, 3698–3710

<https://doi.org/10.1021/acssynbio.4c00534>

Contributions to Publications Presented in this Thesis

Chapter 2:

The study was designed by Urte Tomasiunaite and Kirsten Jung. Urte Tomasiunaite constructed all stains and plasmids, and performed enzymatic assays, spot assays and growth curves. Urte Tomasiunaite purified all proteins for mass spectrometry studies. Pavel Kielkowski analysed protein modification status using mass spectrometry. Urte Tomasiunaite, with contributions from Ralph Krafczyk, purified all proteins for antibody generation. Urte Tomasiunaite confirmed antibody sensitivity and conducted all Western Blot analyses. Ignasi Forné, with contributions from Axel Imhof and Urte Tomasiunaite, conducted the proteome analysis using mass spectrometry. All authors analysed the data. All authors contributed to writing the material and methods section. Urte Tomasiunaite and Kirsten Jung wrote the main manuscript. Kirsten Jung supervised the study.

Chapter 3:

The study was designed by Urte Tomasiunaite and Kirsten Jung. Urte Tomasiunaite constructed all plasmids. Urte Tomasiunaite, with contributions from Korinna Burdack, performed all bacterial survival assays, plate-based fluorescence analyses and took microscopy images. Sophie Brameyer, with contributions from Urte Tomasiunaite, analysed the microscopy data. Urte Tomasiunaite and Tess Brewer analysed the flow cytometry data. Tess Brewer performed the bioinformatical analysis of the high throughput screening. All authors contributed to writing the material and methods section. Urte Tomasiunaite and Kirsten Jung wrote the main manuscript. Kirsten Jung supervised the study.

We hereby confirm the above statements:

Urtė Tomašiūnaitė

Kirsten Jung

Summary

The incorporation of distinct amino acids into the elongating polypeptide chain can highly influence the efficiency of protein synthesis. In particular, consecutive prolines impede translation speed due to the delayed formation of the peptide bond between prolines, resulting in ribosome stalling. The rescue of stalled ribosomes at consecutive prolines is facilitated by the translation elongation factor P (EF-P) in bacteria, and by the archaeal/eukaryotic translation initiation factor 5A (a/eIF5A) in archaea and eukaryotes. Although many EF-P orthologs require a specific post-translational modification (PTM) for optimal functionality, which have long been considered to be essential, it was recently shown that actinobacterial EF-Ps are active without modification. The signature motif of these actinobacterial EF-Ps, PGKGP, has also been identified in bacteria of other phyla (PGKGP-subfamily of EF-P). Previous structural studies have indicated that the PTM of EF-P in *Escherichia coli* facilitates peptide bond formation during polyproline synthesis. However, the absence of PTMs in certain bacteria suggests that the molecular mechanism of EF-P during stimulation of the peptide bond formation remains incompletely understood. This work aimed to unravel the functional principles of unmodified EF-Ps during polyproline synthesis. A combination of structural and functional analyses, including the construction of PGKGP-subfamily EF-P variants using site-directed mutagenesis, screening in *E. coli*, and proteome and structural modelling studies, provided insights into the functional principles of EF-P beyond PTMs.

Activity studies with EF-Ps containing the PGKGP motif identified the EF-P from *Rhodomicrobium vannielii* as being capable of rescuing translation of the polyproline-containing CadC in *E. coli*. Mass spectrometry analysis of purified *R. vannielii* EF-P revealed the absence of a PTM, indicating that the protein is functional in *E. coli* in an unmodified state. Additionally, growth and proteome studies demonstrated that the deletion of the native *efp* gene in *E. coli* can be complemented by chromosomal integration of *R. vannielii efp*. Multiple sequence alignments and activity studies with various EF-Ps of the PGKGP-subfamily in *E. coli* identified three functional important amino acids of unmodified EF-Ps. Structural modelling suggests that these amino acids are involved in establishing contacts with the ribosome and the tRNA in *E. coli*.

Given the lack of knowledge about the motif spectra rescued by unmodified EF-Ps, an *in vitro* reporter system was developed to assess ribosome stalling. The conventional experimental set-up for analyzing the ribosome stalling strength of polyproline-containing motifs is time-consuming and hardly allows high-throughput analysis. The coupling of two reporter genes, encoding a fluorophore and an antibiotic resistance, enabled screenings for ribosome stalling based on bacterial fluorescence and survival. The reporter can be precisely tuned by the antibiotic resistance system. As no additional equipment is required, the reporter system can be easily used in any standard laboratory setting and allows high-throughput screening.

In summary, this work provides fundamental insights into the functional principles of unmodified EF-Ps during translation of polyproline motifs. Furthermore, the basis for the development of unmodified EF-Ps with potential applications in a biotechnological context was established.

Zusammenfassung

Die Effizienz der Proteinsynthese kann durch den Einbau bestimmter Aminosäuren in die sich verlängernde Polypeptidkette stark beeinflusst werden. Insbesondere der Einbau aufeinanderfolgender Proline führt zu einer Verlangsamung der Translationsgeschwindigkeit, da die Peptidbindung zwischen Prolinen verzögert gebildet wird, was zu einem Translationsstopp führt. Die Aufhebung dieser Stopps, der durch aufeinanderfolgende Proline zustande kommt, wird in Bakterien durch den Translationselongationsfaktor P (EF-P) und in Archaeen und Eukaryoten durch den archaealen/eukaryotischen Translationsinitiationsfaktor 5A (a/eIF5A) gewährleistet. Obwohl viele EF-P Orthologe für eine optimale Funktionalität eine spezifische post-translationale Modifikation (PTM) erfordern, die lange Zeit als essenziell galten, konnte kürzlich nachgewiesen werden, dass actinobakterielle EF-Ps ohne Modifikation aktiv sind. Das charakteristische PGKGP Motiv actinobakterieller EF-Ps konnte in Bakterien anderer Phyla nachgewiesen werden (EF-P PGKGP-Familie). Frühere Strukturstudien haben gezeigt, dass die PTM von EF-P in *Escherichia coli* die Bildung von Peptidbindungen während der Polyprolinsynthese unterstützt. Da nicht alle Bakterien eine PTM von EF-P aufweisen, scheint der zugrundeliegende molekulare Mechanismus noch nicht vollständig geklärt zu sein. Das Ziel dieser Arbeit war es, die Funktionsprinzipien von unmodifizierten EF-Ps während der Polyprolinsynthese zu entschlüsseln. Eine Kombination aus Struktur- und Funktionsanalysen, einschließlich der Konstruktion von EF-P-Varianten der PGKGP-Familie mittels Mutagenese, Screening in *E. coli* sowie Proteom- und Strukturmodellierungsstudien, ermöglichten Einblicke in die Funktionsprinzipien von EF-P, die über PTMs hinausgehen.

Zunächst wurden verschiedene EF-Ps der PGKGP-Familie auf die Translationsaktivität des Polyprolin-Proteins CadC in *E. coli* gescreent, wobei EF-P aus *Rhodomicrobium vannielii* als besonders aktiv identifiziert wurde. Massenspektrometrische Untersuchungen von gereinigtem *R. vannielii* EF-P zeigten, dass das Protein keine PTM aufweist, was darauf hindeutet, dass es in *E. coli* im unmodifizierten Zustand aktiv ist. Wachstums- und Proteomstudien konnten darüber hinaus zeigen, dass die Deletion des nativen *efp* Gens in *E. coli* durch die chromosomale Integration von *R. vannielii efp* komplementiert werden kann. Sequenzvergleiche und Proteinaktivitätsstudien von EF-P Varianten der PGKGP-Familie in *E. coli* führten zur Identifizierung von drei Aminosäuren, die für die Funktionalität von unmodifizierten EF-Ps kritisch sind. Strukturmodelle deuten darauf hin, dass diese Aminosäuren bei der Herstellung von Kontakten mit dem Ribosom und der tRNA in *E. coli* involviert sind.

Um anhand der Stärke der Ribosomenpause das Motivspektrum von unmodifizierten EF-Ps aufzuklären, wurde ein *in vitro* Reportersystem entwickelt. Gängige Versuchsaufbauten zur Untersuchung der Stärke der Ribosomenpause an polyprolinhaltigen Motiven sind zeitaufwendig und kaum hochdurchsatzfähig. Die Kopplung zweier Reportergene, die für ein Fluorophor und eine Antibiotikaresistenz kodieren, ermöglicht ein Ribosomenpause-Screening auf der Basis bakterieller Fluoreszenz und Vitalität. Durch die Verwendung des Resistenzgens ist eine genaue Feinabstimmung des Systems möglich. Die Verwendung des Systems erfordert keine zusätzliche Ausrüstung, was eine

einfache Implementierung des Systems in jedem Standardlaboratorium und hohen Probendurchsatz ermöglicht.

Diese Arbeit beschreibt neue Einblicke in die funktionalen Prinzipien von unmodifizierten EF-Ps während der Translation von polyprolinhaltigen Motiven. Darüber hinaus wurde die Grundlage für die Entwicklung unmodifizierter EF-Ps mit potenziellen Anwendungen in einem biotechnologischen Kontext geschaffen.

1 Introduction

1.1 Protein synthesis control in bacteria

In 1958, a significant advancement in the field of biology was achieved with the publication of Francis Crick's central dogma of molecular biology (1). Crick outlined the central dogma as the unidirectional genetic information transfer from DNA to RNA via transcription and from RNA to protein via translation (1,2). Regulation of the final stage of protein synthesis is fundamental for the organism to adapt to fluctuating environments and maintain physiological homeostasis (3,4). The mechanisms in protein synthesis are complex and exhibit a high degree of conservation across all domains of life (5,6), though not entirely uniform. While the transcription and translation are spatially and temporally separated in eukaryotes (7), most prokaryotes have no physical barrier to separate these processes (8,9). In 1964, after observing complex formations of DNA and ribosomes during *in vitro* protein synthesis, Marshall Nirenberg's laboratory proposed that transcription and translation are coupled in bacteria (10,11). Later studies demonstrated that the primary enzyme which carries out transcription, RNA polymerase (RNAP), physically interacts with the key player of translation - the ribosome, thereby offering the first insights into co-transcriptional messenger RNA (mRNA) translation (12-14). This spatial proximity between RNAP and the ribosome allows the newly transcribed mRNA to directly bind the ribosome, which initiates the translation immediately. The observation that RNAP pauses if translation is inhibited provided further evidence for the kinetic coupling of transcription and translation in bacteria (14,15). Translation in bacteria involves the action of the ribosome and numerous factors (Figure 1).

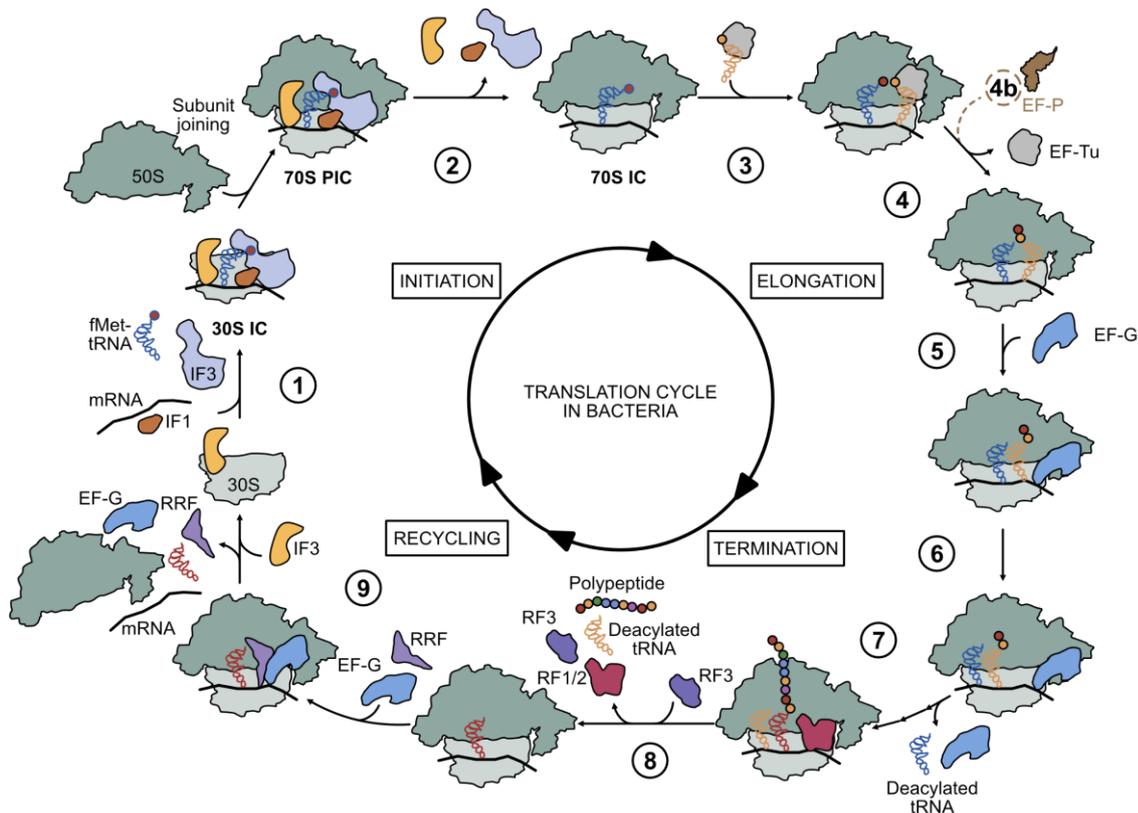


Figure 1. Schematic overview of bacterial translation cycle. (description on the next page)

(Figure 1. description) Depicted are the fundamental steps involved in the process of mRNA translation (initiation, elongation, termination and recycling) with factors involved during this process. 1 – Formation of 30S IC; 2 – Hydrolysis of GFP, dissociation of IFs; 3 – tRNA (acceptor site, A-site) binding; 4 – Hydrolysis of GFP, transfer of the peptidyl; 4b – binding of EF-P (during polyproline synthesis; dashed line); 5 – Formation of the hybrid state, binding of EF-G; 6 – Hydrolysis of the GTP, translocation; 7 – Repeating elongation cycles; 8 – Hydrolysis of the GTP, RF release; 9 – Hydrolysis of the GTP, subunit recycling. 30S IC – 30S initiation complex, 70S PIC – 70S Preinitiation Complex, 70S IC – 70S Initiation Complex, IF – Initiation Factor, EF-G – Elongation factor G, EF-P – Elongation Factor P, EF-Tu – Elongation Factor Tu, mRNA – messenger RNA, RF – Release Factor, RRF – Ribosome Recycling Factor, tRNA^{fMet} – tRNA with formylated methionine. The figure was adapted and extended from (16).

Bacterial ribosomes demonstrate high catalytic rates in protein synthesis, by forming approximately 20 peptide bonds per second (17-19). This underscores the need for a highly regulated and finely tuned translational regulatory mechanism. Thus, translation in bacteria is regulated by diverse translation factors and can be broadly divided into four phases: initiation, elongation, termination and ribosome recycling (Figure 1)(16).

1.1.1 Translation initiation

The initiation of translation represents the initial step of protein synthesis and involves several key players. These include the initiator transfer RNA (tRNA^{fMet}), which has a distinctive N-formylmethionine, differentiating it from other tRNAs (20), the mRNA, 30S and 50S ribosomal subunits and the initiation factors (IF) IF1, IF2 and IF3 (16,21-24)(Figure 1).

In bacteria, translation is canonically initiated in three stages by forming three complexes (16,21-27)(Figure 1). Briefly, the initial stage, the 30S pre-initiation complex (30S PIC) is formed, whereby the 30S ribosomal subunit recruits IF1, IF2, IF3, tRNA^{fMet} and the mRNA. The complex is stabilised by the base pairing of the tRNA^{fMet} anticodon and the start codon on the open reading frame (ORF) of the mRNA, which leads to the 30S initiation complex (30S IC) formation (21-23,26). In the final stage of translation initiation, the 50S ribosomal subunit is recruited and binds with the 30S IC to form the 70S initiation complex (70S IC)(16,23,27)(Figure 1).

Initiation factors play critical roles during translation initiation. IF1 promotes IF2 and IF3 binding to the 30S ribosomal subunit and regulates the selection of tRNA^{fMet} and the mRNA (28-30). Moreover, structural and functional studies have shown IF1 to localise to the ribosomal acceptor site (A-site), where it supports the binding of initiator tRNA to the ribosomal peptidyl-site (P-site)(31,32). IF2 is a guanosine triphosphatase (GTPase). GTP hydrolysis facilitates the binding of tRNA^{fMet} to the ribosomal P-site by interacting with the aminoacyl acceptor stem and the N-formyl-methionine of the tRNA^{fMet} (29,33-37). In addition, IF2 has a strong affinity for the 50S ribosomal subunit, which ultimately promotes docking of the 50S subunit to the 30S IC (29,38-40). IF3 binds to the 30S ribosomal subunit and regulates ribosome assembly, thus preventing premature assembly of 30S and 50S (21,22,25).

1.1.2 Translation elongation

The translation elongation phase represents the central stage of the translation process. During translation, the information encoded on the mRNA is decoded into the nascent chain of amino acids (23,24). Repetitive decoding cycles, peptide bond formation and tRNA translocation, constitute the steps of translation elongation, in which ribosomes play a central role. Ribosomes comprise three tRNA binding sites, designated E- (exit-), P- (peptidyl-) and A- (acceptor-) sites (41). The codons on the mRNA, composed of three consecutive nucleotides, determine the specific amino acid incorporated into the polypeptide chain (42,43). Once the second codon is positioned within the A-site and thus is accessible for the elongator aminoacyl-tRNAs (aa-tRNA), the elongation phase begins. The elongation process is governed by the action of translation elongation factors, namely EF-Tu (elongation factor Tu), EF-G (elongation factor G) and EF-P (elongation factor P)(16,23,24)(Figure 1). SelB is a specialised elongation factor in bacteria, which is required for the delivery of the 21st natural amino acid selenocysteine (44).

A single cycle of translation elongation can be subdivided into three distinct phases. These are the initial attachment of aa-tRNA to the ribosome, peptide bond formation and peptidyl transfer, and subsequent translocation of the aa-tRNAs within the ribosome. In the initial phase, the aa-tRNA is transported by the GTPase EF-Tu to the A-site of the ribosome (16,45-47)(Figure 1). Here, EF-Tu forms a tertiary complex with the aa-tRNA and GTP, which facilitates its binding to the ribosomal bL12 stalk (45,48). In the A-site of the ribosome, the interaction between the aa-tRNA anticodon and the mRNA codon prompts the hydrolysis of GTP (49-54) and the release of inorganic phosphate (Pi) by EF-Tu (55). Once bound to the guanosine diphosphate (GDP), EF-Tu dissociates from the ribosome by releasing the A-site bound aa-tRNA (52,56). In the subsequent phase, a peptide bond is formed between the amino acids of P-site located peptidyl-tRNA and the A-site located aa-tRNA in the peptidyl transfer center (PTC)(52,53). Here, the amino group of the aa-tRNA forms a peptide bond by nucleophilic attacks on the carbonyl carbon of the peptidyl-tRNA, finally resulting in a deacetylated P-site tRNA and an A-site peptidyl-tRNA (57,58). In the third phase, the translation elongation factor EF-G promotes ribosomes translocation after GDP hydrolysis (16,59)(Figure 1). This stimulates the ribosome to move along the mRNA, facilitating the shifting of the P-site located tRNA to the E-site and the A-site located tRNA to the P-site, leaving the A-site empty for new aa-tRNAs (23,24,60).

The action of EF-Tu and EF-G is critical during the synthesis. EF-P is instrumental in synthesising consecutive proline motifs (16,61,62)(Figure 1). Translation elongation continues until the ribosome reaches the stop codon, which initiates translation termination and ribosome recycling.

1.1.3 Translation termination and ribosome recycling

The translation elongation process is terminated when the A-site of the ribosome encounters a stop codon. Stop codons UAA, UAG and UGA are universally conserved and do not code for any amino acid (63,64). Positioning the stop codons in the A-site of the ribosome serves to trigger the recruitment of

release factors (RF), which assist during the termination process (65). Three phases regulate the termination process: identification of the stop codon, cleavage of the ester bond of the peptidyl-tRNA via hydrolysis and detachment of RF1 and RF2 by the action of RF3 (16,23,24)(Figure 1).

The release factors, RF1 and RF2, differ in their ability to bind particular stop codons: UAG and UAA codons are recognised by RF1 via the conserved PVT motif, whereas RF2 recognises UGA and UAA via the conserved SPF motif (66-73). GGQ motif is conserved in both RF1 and RF2, facilitating the peptidyl-tRNA hydrolysis (74), which causes the peptide to release. Peptide release is disturbed when the GGQ motif is mutated (74-77), underlining its functional importance during peptidyl-tRNA hydrolysis. In the final step before ribosome recycling, the GTPase RF3 promotes the RF1 and RF2 dissociation from the ribosome after GTP hydrolysis (16,75,78,79)(Figure 1). Ribosome-bound mRNA and tRNA are released by the ribosome recycling factor (RRF) and EF-G. Accordingly, the reuse of ribosomal subunits for further translation is facilitated (16,23,80-82)(Figure 1).

1.1.4 Challenges in translation efficiency

Bacteria have evolved diverse regulatory mechanisms to fine-tune mRNA translation in response to rapidly changing environmental conditions and stress. These mechanisms include the composition of the nucleotides that lie both adjacent to and within the coding sequence. These sequences can form intramolecular base pairs within the mRNA, resulting in the formation of secondary structures such as hairpins, stem-loops, or pseudoknots (83-85). mRNA structures have been reported to have regulatory effects on translation efficiency, for example during sporulation (86) and toxin production (87,88). They can also modulate the initiation frequency of translation (21). In addition to mRNA structures, the ribosome binding site (RBS) is another element which influences translational regulation. The RBS is located between -20 to +15 nucleotides relative to the translation initiation codon (89). Characteristically, the RBS contains a Shine-Dalgarno (SD) sequence (90,91), whose eukaryotic equivalent is known as the Kozak sequence (92-94). This sequence, just upstream of the start codon, serves as a binding motif for the anti-SD sequence of the 16S rRNA of the ribosome (90,91). The interaction between the SD and anti-SD sequences supports the correct positioning of the start codon within the 30S ribosomal subunit (36,95,96), thus initiating the formation of the translation initiation complex (97-99). However, among 162 analysed prokaryotic genomes, the presence of SD-led genes ranged from ~12% to ~90% (100), suggesting the existence of a high number of non-SD-led mRNAs (23,98,100-102). The translation of mRNAs lacking the SD sequence is as effective as that of SD-led sequences, suggesting that other elements or mRNA features can facilitate the selection of the start codon (103,104). Sequestration of the SD within an mRNA secondary structure may further modulate translational regulation, masking it from the ribosome and ultimately interfering with translation initiation(105-107).

Codon selection within coding sequences represents a further critical element that influences translational efficiency (108-110). The genetic code is degenerate, meaning that multiple synonymous codons can be used to encode each of the 20 proteinogenic amino acids. The utilization of synonymous

codons is, however, not uniform across the genome, resulting in the phenomenon known as "codon usage bias," which has been observed in diverse species (111-113). Codon usage bias correlates with the abundance of cognate tRNAs, thus it has been suggested that synonymous codons exert a different influence on translation efficiencies and translation elongation (111,114,115). Preferred codons (optimal codons) positively influence translation elongation speed, whereas rare codons (non-optimal codons) slow down elongation speed and cause the ribosome to pause (110,116,117). Indeed, the presence of rare codons in genes has been shown to correlate with the suppression of protein expression in *E. coli* (109,118-120). However, these ribosomal pauses can confer advantages, such as facilitating co-translational protein folding (121-126). Thus, ribosomal pauses generate temporal intervals during translation, which enable newly synthesized polypeptides to adopt an optimal conformational folding within an appropriate timeframe. However, it is evident that they have an overall detrimental influence on translational efficiency.

Ribosomal pausing is not solely caused by codon selection. Specific amino acids encoded by the translated mRNA can cause ribosome pausing as well. Among the proteinogenic canonical amino acids, proline has a unique cyclic structure featuring an imino group, thus classifying it as an imino acid. In contrast to other amino acids, in which the nitrogen in the amino group forms a single bond with the carbon, the nitrogen atom in proline's imino group forms a double bond with the carbon atom. This results in the pyrrolidine ring structure of proline being rigid and inflexible, which renders proline an ineffective acceptor and donor during the transpeptidation reaction (127-129). Consequently, during the synthesis of motifs with consecutive prolines, peptide bond formation is slow, which leads to ribosome stalling (61,62,130,131). However, prolines play indispensable roles in amino acid sequences, for example, enabling sharp turns within protein structures (132,133) and regulating protein copy number (61,134). These advantages have compelled organisms across all domains of life to develop strategies to overcome polyproline-caused ribosome stalling. Ultimately, translational factors evolved in bacteria designated as EF-P, which are instrumental in alleviating ribosomal stalls during the translation of polyproline motifs (61,62).

1.2 EF-P discovery and structural elucidation

The discovery of translation elongation factor P can be traced back to 1975, when Bernard Glick and Clelia Ganoza isolated a ribosomal factor shown to improve the binding between N-formyl-Met-tRNA and the aminoacyl-tRNA analogue puromycin (135). The isolated factor was named EF-P, based on the factor's role in peptide bond formation during peptide chain elongation (135). Several years later, in 1979, Glick, Chladek and Ganoza demonstrated that the efficiency of peptide bond formation promoted by EF-P depends on the selected aminoacyl-tRNA (136). In light of this observation, they were able to draw the conclusion that EF-P activity is crucial for peptide bond formation between distinct amino acids, such as leucine or glycine (136). However, despite the fact that EF-P was regarded as a crucial enzyme, with a primary role in the formation of initial peptide bonds, it took more than 30 years for the scientific

community to elucidate the principal mode of action of EF-P, which was found to be in the synthesis of consecutive polypeptides (61,62). The initial assumption that EF-P was implicated in the first peptide bond formation could be traced back to the isolation of its eukaryotic counterpart, the eukaryotic translation initiation factor 5A (eIF5A)(137). At that time, eIF5A was thought to be involved in the initiation of haemoglobin translation (19,138). A few years later, however, several studies demonstrated that eIF5A does not influence the formation of the initiation complex and is not required for the globin biosynthesis, contrary to previous assumptions (139-141). Years later, molecular and biochemical studies showed that eIF5A is involved in translation elongation rather than in translation initiation (142,143).

In 1998, sequence homology between EF-P and eIF5A was identified through multiple sequence alignment, providing preliminary evidence for a potentially analogous mode of action shared by both proteins (144). However, the observed sequence similarity of approximately 20% (144) was insufficient to conclude anything about their similar functional roles or homology. Promising insights into the EF-P site of action were provided by the study of Hanawa-Suetsugu and colleagues in 2004, who published the crystal structure of the EF-P from *Thermus thermophilus* (145). The study demonstrated that the EF-P is composed of three β -barrel domains (domain I, II and III) arranged in an "L" shape, which closely resembles the structural shape of a tRNA. These findings suggested that EF-P has the potential to bind and interact with the ribosome. A comparative structural analysis of the *Thermus thermophilus* EF-P and the *Methanocaldococcus jannaschii* eIF-5A revealed high similarities between two domains, given that eIF-5A is composed of solely two domains (N and C)(145)(Figure 2). Hanawa-Suetsugu and colleagues suggested that the topological similarity between domains II and III of EF-P and domain C of eIF-5A may indicate that domains II and III originated through a duplication event (145)(Figure 2), probably early in the evolution of EF-P. A more detailed understanding about the site and mode of action of EF-Ps was gained from a study conducted by Blaha and colleagues in 2009, who reported the crystal structure of the *Thermus thermophilus* 70S ribosome in complex with EF-P, an initiator tRNA and fragments of an mRNA (146). In the crystal structure, EF-P was localised between the E- and P-sites of the ribosome, suggesting that EF-P acts during translation initiation by promoting the formation of the first peptide bond through optimal positioning of the initiator tRNA (146). Subsequent studies have demonstrated EF-P acts during the elongation phase of translation, rather than the initiation phase (61,62), which is consistent with observations of eIF5A (142,143).

In sum, from the initial discovery of EF-P, many conclusions regarding its function were speculative and often inaccurate. Sequence and structural studies with EF-P supported the research aimed at elucidating the factor's precise function.

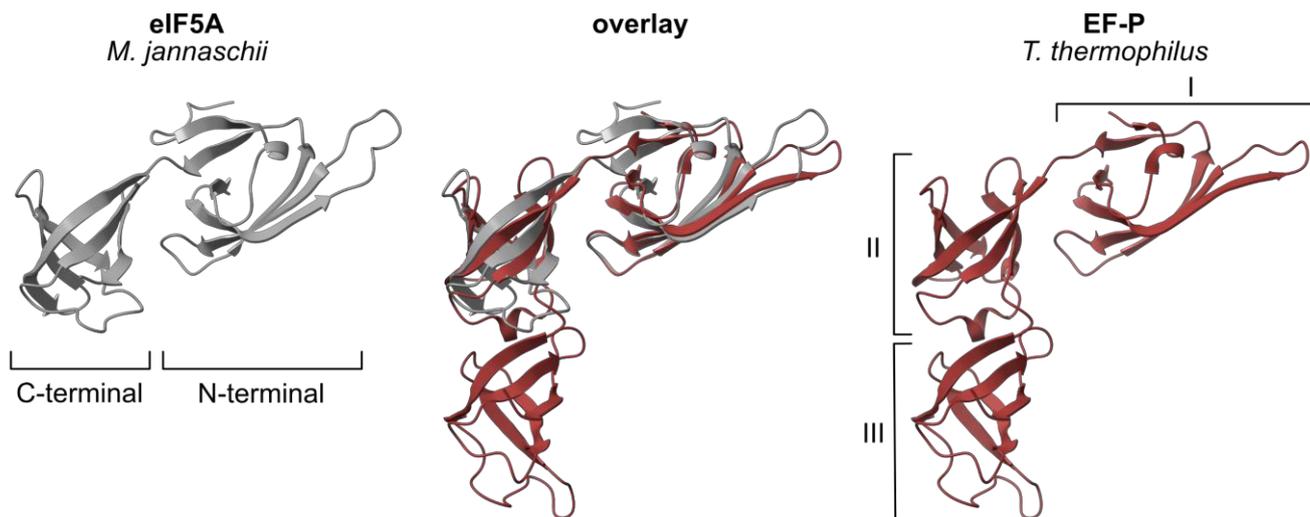


Figure 2. Structural comparison of eIF5A and EF-P. Ribbon diagrams of *Methanococcus jannaschii* eIF5A structure (PDB: 2EIF)(left, coloured in grey) and of *Thermus thermophilus* EF-P (PDB: 1UEB)(right, coloured in red). C-terminal and N-terminal domains of eIF5A, and domains I-III of EF-P are labelled accordingly. The superimposition of eIF5A and EF-P ribbon diagrams is depicted in the middle (overlay). Ribbon colouring and structural superimposition was conducted using UCSF ChimeraX (147,148). EF-P – elongation factor P, eIF5A – eukaryotic initiation factor 5A.

1.2.1 Molecular function of EF-P

The first comprehensive understanding of the molecular function of EF-P was achieved in 2013, when two research groups described the role of EF-P in rescuing stalled ribosomes during the translation of consecutive polyproline motifs (61,62). One of these studies, conducted by Ude and colleagues, investigated CadC-dependent acid stress response in *E. coli* (61). The relationship between EF-P and the CadC was initially observed indirectly, following random deletion of one of the EF-P modifying enzymes, lysine 2,3-aminomutase (EpmB, YjeK), via transposon mutagenesis. Loss of EpmB resulted in reduced CadA activities (61). It was demonstrated that the loss of EF-P and its modifying enzymes could be complemented, rescuing the markedly reduced CadA activities. However, the absence of EF-P and its modifying enzymes did not directly affect the production of CadA, indicating that another protein in the CadC-dependent acid stress response was likely dependent on a functional EF-P (61). This was subsequently demonstrated to be the transcriptional regulator of the *cadBA* operon, CadC. Deletion of *efp*, resulted in a substantial decrease in CadC levels, thereby confirming EF-P's role in CadC production (61). CadC-LacZ translational fusions of varying lengths revealed that translation of a region containing several prolines is dependent on the action of the EF-P (61). Further investigations employing *in vitro* assays confirmed that the translation of polyproline motifs can be rescued by EF-P (61). A study conducted in parallel by Doerfel and colleagues showed the translation of consecutive prolines to be dependent on an active EF-P as well (62). In this study, the formation of the peptide bond between proline and puromycin was found to be reliant on the presence of EF-P (62). Furthermore, the production of the

N-terminal fragment of the release factor glutamine methyltransferase PrmC was significantly influenced by EF-P, particularly when consecutive proline motifs were incorporated into the sequence (62). Functional similarities between EF-P and eIF5A were confirmed half a year later by Gutierrez and colleagues, who proved eIF5A influences the formation of polyproline motifs using *in vivo* and *in vitro* studies (149). In addition to functional studies, structural studies on interactions between the ribosome and EF-P during the translation of polyproline motifs have been instrumental in understanding the mode of action of EF-P. Prior to the identification of the role of EF-P in the synthesis of polyproline motifs (61,62), only structures of EF-P alone or in complex with the ribosome during the synthesis of non-polyproline motifs had been reported (145,146,150). Consequently, in 2017, Huter and colleagues published a cryo-EM structure that revealed the structural interplay between the ribosome, peptidyl tRNA and EF-P during the translation of a polyproline motif (151). The study proposed that the P-site tRNA can become destabilised through the incompatibility of the polyproline-containing nascent chain and the exit tunnel, leading to ribosomal stalling (151). In this scenario EF-P acts as a stabilising factor of the P-site tRNA by interacting with the tRNAs CCA end with its post-translationally modified tip (151), a phenomenon observed with eIF5A as well (152). This stabilisation was hypothesized to promote the peptide bond formation and to enable the formation of a more favourable conformation of the polyproline-containing nascent chain (151). Furthermore, these structural studies indicated that post-translational modifications (PTM) play a critical role in the functionality of EF-P and eIF5A, as they demonstrate a direct involvement in the interaction with the CCA end of the P-site tRNA (146,151,152).

1.2.2 EF-P activation by post-translational modification

Prior to the identification of the molecular function of EF-P, initial studies had already documented that the modification of a positively charged amino acid in the N-terminal domain/domain I on the terminal located activation loop has an impact on EF-P function (146). The depletion of EF-P modifying enzymes results in phenotypes comparable to those observed in *efp* deletion strains, including growth deficiency (153-158), attenuated virulence (155,159,160) and decreased fitness (61,161). These findings highlight the necessity of PTMs for full EF-P functionality. The mechanisms through which PTMs activate initiation factors differ between archaeal and eukaryotic organisms compared to bacteria. In contrast to the single known modification, hypusination, which activates eIF5A in archaea and eukaryotes (162-165), several diverse PTMs exist in bacteria. To date, three PTMs of EF-P have been uncovered: (R)- β -lysylation and hydroxylation of lysine (153,159,166,167), arginine- α -rhamnosylation (155,168) and lysine-5-aminopentanoloylation (169,170)(Figure 3).

(R)- β -lysylation is the most extensively studied EF-P modification, with reports indicating its presence in 25-29% of sequenced bacterial genomes (including *E. coli* and *Salmonella enterica*)(155,171,172). This modification is catalysed by three enzymes: the EF-P (R)- β -lysine ligase (EpmA, also known as YjeA, PoxA or GenX), the lysine 2,3-aminomutase (EpmB, also known as YjeK) and the EF-P hydroxylase (EpmC, also known as YfcM)(Figure 3A). Initial reports described only the role of EpmA and

EpmB during EFP modification, EpmC was discovered later. Accordingly, the first descriptions outlined two distinct modification phases. In the first phase, (S)- α -lysine is converted into (R)- β -lysine by the action of EpmB (173). In the second phase, EpmA ligates the newly generated substrate (R)- β -lysine to the lysine present in the EF-P activation loop (K34 in *E. coli*)(153,159,167). In 2012, Peil and colleagues published a study that completed the pathway required for the lysylation modification, revealing the role of EpmC in the final stage of this particular modification (166). *In vitro* and mass spectrometry studies revealed inconsistencies (166) between the calculated (+128 Da) and the measured (+144 Da) modification mass of the endogenous EF-P (153,159,167,174,175). Peil and colleagues demonstrated that cells with depleted *epmC* exhibited a shift in mass by 16 Da, thereby substantiating the role of EpmC in hydroxylating the modified EF-P tip in the final phase of the lysylation process (166). It should be noted, however, that *epmC* is not present in every bacterial genome in which *epmA* and *epmB* have been identified (155). Moreover, the deletion of *epmC* has been observed to have a negligible impact on EF-P activity (62,131,176,177). Therefore, the precise rationale behind the hydroxylation carried out by EpmC during the (R)- β -lysylation process remains unclear and current research suggests it plays a minor role in EF-P activity.

In contrast to the (R)- β -lysylation modification process, where a lysine situated at the terminus of the EF-P activation loop undergoes modification, the rhamnosylation process involves the modification of an arginine at the equivalent position of EF-P (155,158)(Figure 3B). The subfamily of EF-Ps which contain an arginine activated by rhamnosylation has been identified in 9% of sequenced bacteria (mostly β -proteobacteria and some γ -proteobacteria) and can be found in species like *Neisseria meningitidis*, *Shewanella oneidensis* and *Pseudomonas aeruginosa* (155,156,158). The enzyme which catalyses the rhamnosylation of this family of EF-Ps was initially designated as an uncharacterised protein with the domain DUF 2331 (155). Following the elucidation of its mechanism of function, it was renamed to protein-arginine rhamnosyltransferase (EarP)(155)(Figure 3B). A phylogenetic analysis, coupled with mass spectrometry, *in vivo* and *in vitro* studies, revealed the precise pathway catalysed by EarP (155). EarP uses deoxythymidine diphospho-L-rhamnose (dTDP-L-rhamnose) from the rhamnosylation biosynthesis pathway *rmlABCD* as the substrate for rhamnosylation of EF-P (155,158). While (R)- β -lysine and rhamnose have unique chemical and structural characteristics, both enhance the function of EF-P in an analogous manner (155,158).

5-aminopentanoylation, found in Firmicutes, is the most recently identified PTM of EF-P (169,170,178). Approximately 4.5 % of all sequenced bacteria, including the Gram-positive bacteria *Bacillus subtilis* and *Staphylococcus carnosus*, encode the enzyme YmfI (172) which catalyses the final step of the 5-aminopentanoylation reaction (170,178)(Figure 3C). In contrast to other PTMs of EF-P, the assembly of 5-aminopentanol on lysine situated within the activation loop of EF-P (K32 in *Bacillus subtilis*) involves the formation of several intermediate modifications (178). Mass spectrometry studies uncovered the identity of all the intermediate modifications (hydroxypentanone, pentanone, 5-aminopentanone)(178), as well as the enzyme YmfI which converts the penultimate modification to 5-

aminopentanol (170). However, the identity of the enzymes involved in the generation of the intermediate modifications remains unknown and requires further investigation (178)(Figure 3C).

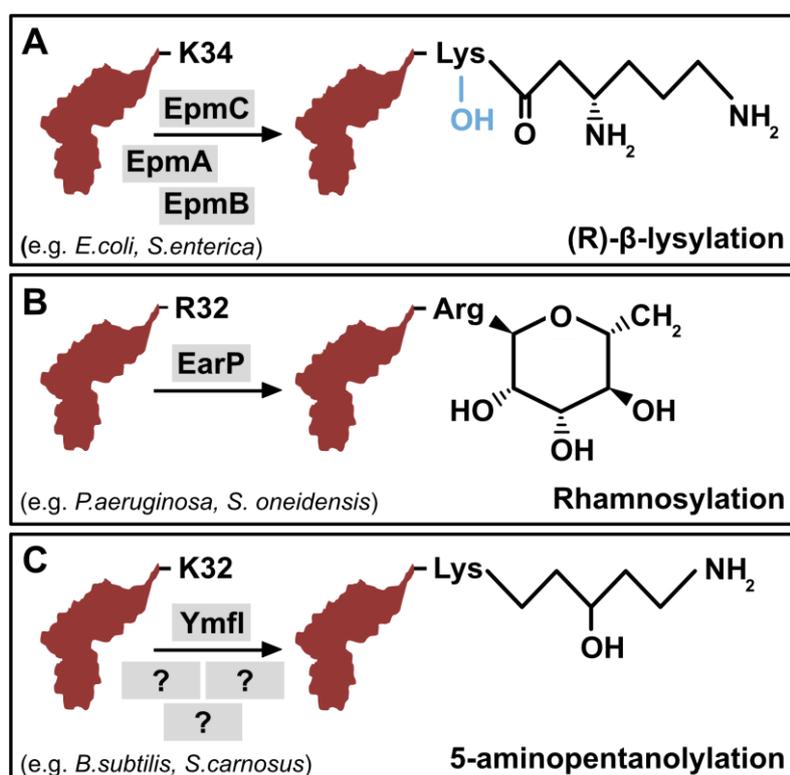


Figure 3. Overview of known EF-P post-translational modification pathways in bacteria. **A** (R)-β-lysylation is catalysed at the conserved lysine at position 34 (K34) of EF-P. Modifying enzymes that regulate this reaction are lysine 2,3-aminomutase (EpmB) and EF-P (R)-beta-lysine ligase (EpmA)(153,159,167,173). In certain species, EF-P hydroxylase EpmC hydroxylates the EF-P on the C5(δ) of K34 (OH depicted in blue)(166). **B** Rhamnosylation is catalysed at the conserved arginine at position 32 (R32) of EF-P. Modifying enzyme that regulates this reaction is the protein-arginine rhamnosyltransferase (EarP)(155,158). **C** 5-aminopentanylation is catalysed at the conserved lysine at position 32 (K32) of EF-P. The last 5-aminopentanylation modification step is catalysed by YmfL (169,170). Modification enzymes catalysing the preliminary steps of the 5-aminopentanylation modification are not yet known and therefore indicated by "?".

Only 50% of sequenced bacteria encode known modification systems. This suggests that the remaining half possess additional modification systems that are essential for EF-P activity, but which have yet to be identified, or that their EF-Ps are active without PTMs. Furthermore, the exact molecular mechanism of modified EF-P in polyproline motif translation remains elusive. Based on the current knowledge, structural studies have indicated that EF-P modifications may exert an influence on the nascent chain stabilisation by interacting with the CCA end of the P-site tRNA (146,151,155). However, it seems that not all modification types extend into the PTC to the same degree due to their diverse chemical structures (19,151,155). Consequently, the precise mechanisms by which PTMs facilitate the translation of polyproline motifs remain a topic for further research.

1.2.3 Fully active unmodified EF-Ps

In 2020, Pinheiro and colleagues reported a notable discovery concerning the activation mechanism of actinobacterial EF-Ps (172). Actinobacteria are a highly relevant group of microorganisms with substantial biotechnological applications, particularly in the production of antibiotics, secondary metabolites, and other biotechnologically relevant compounds (179,180). The study by Pinheiro and colleagues demonstrated that a high proportion of proteins encoded by actinobacteria contain consecutive proline motifs, with 1.17 motifs/protein in *Mycobacterium tuberculosis*, 1.06 motifs/protein in *Nocardia farcinica*, and 1.08 motifs/protein in *Streptomyces coelicolor* (172). In comparison, *E. coli* encodes 0.49 motifs/protein, while *Salmonella enterica* encodes 0.51 motifs/protein (134,172). This suggested that the translation of a high proportion of the proteome of Actinobacteria requires the action of EF-P (172). However, the precise mechanism by which actinobacterial EF-Ps are activated remains elusive, as Actinobacteria encode none of the known EF-P PTM pathways. In 2020, Pinheiro and colleagues were the first to report on a new subclass of EF-Ps, whose activity is not dependent on a PTM (172). Mass spectrometry and isoelectric focusing studies confirmed the absence of a modification in the actinobacterial representatives *Corynebacterium glutamicum*, *Mycobacterium smegmatis*, and *Streptomyces coelicolor* (172). Bioinformatic analyses identified a characteristic palindromic loop sequence associated with these unmodified EF-Ps (PGKGP, Pro-Gly-Lys-Gly-Pro), present in 11% of EF-Ps from all sequenced bacteria (172). Most EF-Ps with this signature motif were detected in the phylum Actinobacteria, including the genera *Bifidobacterium*, *Corynebacterium*, *Gardnerella*, *Mobiluncus*, *Mycobacterium*, *Nocardia*, and *Streptomyces* (172). However, bacteria from other phyla were also shown to encode EF-Ps with the PGKGP signature motif, which do not encode for known EF-P modification enzymes (Figure 4)(172). This indicates about their functionality independent of a modification or dependent on until now unknown modification system. Biochemical studies demonstrated that the presence of both prolines in the PGKGP sequence (P30 and P34 in *C. glutamicum*) is critical for the functionality of this family of EF-Ps, as substitutions of both prolines resulted in the formation of non-functional variants (172). This suggests that the two prolines increase the rigidity of the activation loop of these EF-P, enabling the favourable positioning of the P-site tRNA for efficient polyproline motif translation (172). Active, but unmodified EF-Ps represent a notable divergence from the prevailing understanding of translation factor activation, which typically rely on PTMs in a diverse range of organisms (61,62,149,155,178). The existence of unmodified EF-Ps (172) raises the question about the necessity of the modification of EF-Ps in other bacterial phyla. Further studies should address these questions, such as: is the palindromic PGKGP signature motif strictly necessary for unmodified EF-P to be functional? Are EF-Ps with the palindromic PGKGP signature motif unmodified originating from other bacterial phyla, outside of the Actinobacteria? Is the activation loop the only component that determines the functionality of unmodified EF-Ps? Do other modification systems exist?

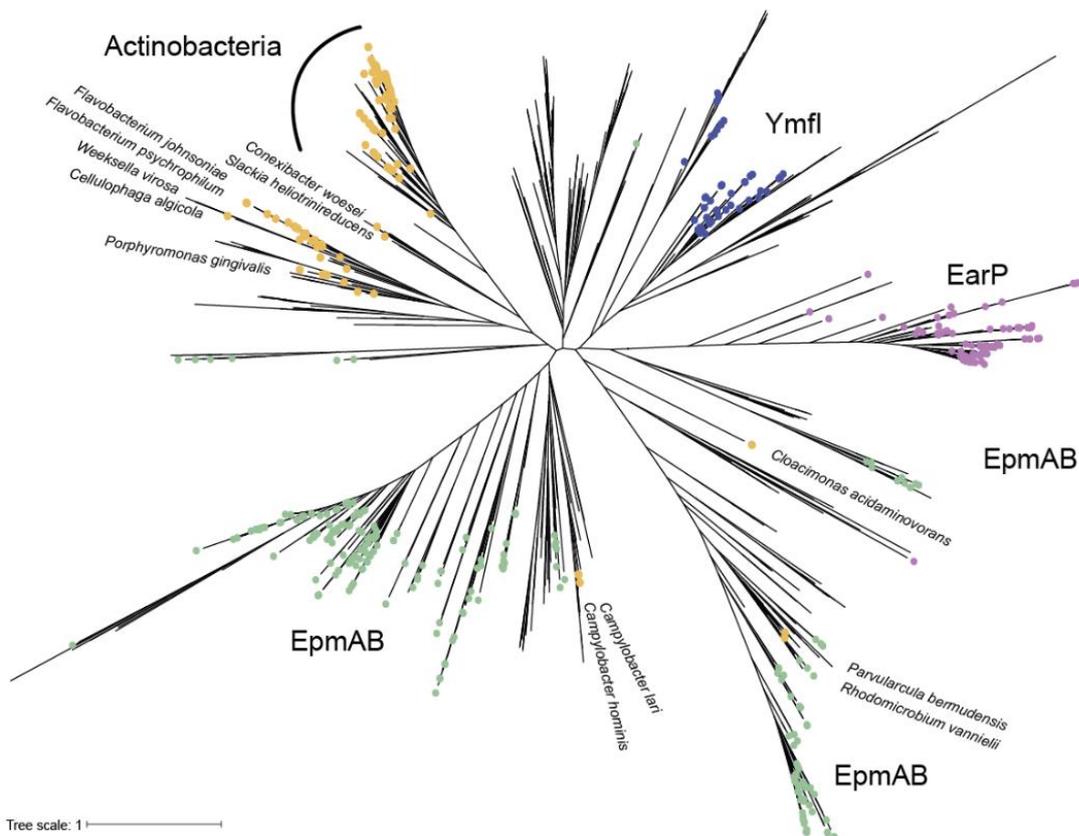


Figure 4. EF-Ps containing a proline at the position 34 are found in bacterial species lacking genes coding for EF-P modification enzyme homologues. Bacterial species with EF-Ps containing a proline at the position 34 are depicted in yellow dots. Bacterial species coding for the modification systems: YmfI (blue), EarP (purple), EpmA and EpmB (green dots). Figure derived from (172).

1.2.4 Synthetic activation of EF-P

The discovery of diverse EF-P modification systems in bacteria, all of which serve the same function in boosting the translation of polyproline motifs, has left many questions unanswered. The evolutionary rationale behind the emergence of these chemically distinct modification systems to enhance EF-P activity remained unclear. Researchers began to question whether EF-P modification systems could be used interchangeably between bacterial species and whether different EF-P modification systems could be adapted to be compatible with originally differently modified EF-Ps. The feasibility of switching modification systems was demonstrated by Lassak and colleagues (155). Here, replacement of native EF-P in an *E. coli* Δefp strain with *S. odeinensis* EF-P and co-expression of its corresponding modification enzyme EarP, restored pH sensor CadC levels comparable to the wild type (155). Another study demonstrated that the substitution of two amino acids in the *E. coli* EF-P, in conjunction with the co-expression of *S. odeinensis* EarP, were sufficient to facilitate a switch in the modification system from lysylation to rhamnosylation in *E. coli* (181). In this context, the distinct amino acid composition of the EF-P activation loop was shown to be the determining factor in the non-native EF-P adaptation in *E. coli* (181). An alteration of the substrate attached to the EF-P activation loop does not necessarily result in a universal outcome with respect to EF-P activation (182). In this regard, *E. coli* EF-P could be modified

with non-canonical substrates, leading to the formation of novel modifications (182). However, functional assays confirmed that the native substrate, (R)- β -lysine, remained the only substrate capable of fully functionalizing *E. coli* EF-P, outperforming other lysine derivatives tested (182). Thus, the development of diverse modification systems has likely undergone several levels of optimization during evolution, as non-canonical substrates that are structurally similar to the natively used substrate cannot be used to activate EF-P (182). All in all, these results indicate that, throughout the course of evolution, an optimal modification system likely co-evolved with its corresponding activation loop and compatible substrate to ensure the efficient functionality of EF-P (181).

The studies described here have mainly focused on the transferability of different modification systems between bacterial species and the sequence compatibility of the activation loop and the substrate required for EF-P activation. However, there has not been much focus on the influence of individual amino acids present outside the activation loop of EF-P on EF-Ps activity. Despite the high structural similarity of all EF-Ps (172), there are large differences in EF-P sequence identity (145,172). The influence of the sequence composition within the remaining EF-P backbone, which may also contribute to the functionality of EF-P beyond the modification itself, is a subject that has yet to be explored and will be a major part of this thesis (Chapter 2).

1.2.5 The diversity of motifs causing ribosome pausing

Following the discovery of the molecular function of EF-P in facilitating the synthesis of polyproline motifs (61,62), further studies aimed to uncover the full repertoire of motifs that rely on EF-P for their translational rescue (131,183). By the end of 2013, a spectrum of motifs with less than three consecutive prolines was identified (131). With proteome comparisons of *E. coli* wild type with diverse *E. coli* deletion strains ($\Delta yjeK/\Delta epmB$, $\Delta yjeA/\Delta epmA$, $\Delta yfcM/\Delta epmC$), Peil and colleagues showed that ribosomal pausing strength is dependent on the amino acids upstream and downstream of the consecutive prolines in an polyproline motif (X-PP-X)(131). A combination of two strong motifs (XPP and PPX) resulted in stronger ribosomal pauses than either of the motifs considered separately (131). In 2014, Starosta and colleagues reported that the amino acid several amino acids upstream of the consecutive prolines affected ribosomal pausing strength as well, albeit to a lesser extent than the amino acid directly upstream of the prolines (184).

The regulation of translation pausing mediated by polyproline motifs remains a complex area of research. For example, proteomic data studies have demonstrated that the absence of EF-P does not necessarily result in a uniform reduction in proteins containing polyprolines (131,183). Ribosome profiling studies confirmed this phenomenon, as ribosomes were shown to pause at only a subset of XPP motifs in a Δefp background (185,186). A combination of factors, including the motif type, the motif location within the ORF and the rate of translation initiation, have been suggested to influence which polyproline motifs produce ribosomal pausing (19,186-188). In the absence of EF-P, protein production is predicted to be reduced when a strong stalling motif is situated closer to the 5' end of the gene and the translation

initiation rate is high (19,186-188). The codon selection in the polyproline motif adds to the complex regulation of polyproline synthesis (189). As demonstrated by Krafczyk *et al.*, different proline codons (CCA, CCC, CCG or CCU) present in a polyproline motif seemed to show a different impact on protein synthesis (189).

The presence of polyproline motifs in genomes suggests that the functional advantages of these motifs outweigh the intrinsic complexity and metabolic burden for organisms. Nevertheless, further research is required to clarify why certain combinations of encoded amino acids exert a stronger influence on the translation than others. Furthermore, the rescue of only half of the XPPX motifs, along with the discovery of unmodified EF-Ps, raises the question of whether there is a subset of motifs that are preferentially rescued by modified or unmodified EF-Ps. The methods currently used to study the dynamics of translational rescue of polyproline motifs are relatively complex, requiring the integration of *in vivo* reporters into the genome, and costly, requiring specialized equipment and reagents (61,131,182,183,186,189). The second part of this thesis will therefore focus on the development and implementation of a simple plasmid-based dual reporter system that can be easily used in any laboratory to assess ribosome pausing during translation of polyproline motifs (Chapter 3). The ultimate goal of this system is to enable the investigation of the spectrum of polyproline motifs that are preferentially rescued by unmodified translation elongation factors P.

1.3 Scope of the thesis

Over time, several strategies for EF-P activation have been identified. Many studies have concentrated on the role of the modification systems and the activation loop region. Nevertheless, given the considerable discrepancies between the different modification systems and the recent discovery of unmodified EF-Ps, it may prove advantageous to shift the focus towards the protein backbone and investigate whether other EF-P regions are critical for activity beyond the modifications. Thus, the primary objective of the first part of this thesis (Chapter 2) is to investigate the functionality of PGKGP subfamily EF-Ps, which are predicted to be unmodified, in *E. coli*. Additionally, the role of specific amino acids within the PGKGP subfamily EF-P backbone will be analysed to determine their contribution to overall EF-P functionality in the absence of a PTM.

The second part of the thesis (Chapter 3) will focus on the development of a versatile and cost-efficient reporter system to investigate the polyproline motif spectrum, whose translation is dependent on functional EF-Ps. By translationally coupling two reporter genes, coding for a fluorophore and an antibiotic resistance protein, we aim to improve current screening capabilities of polyproline motifs. This reporter system will facilitate future research in ribosome pausing measurements in a cost-effective manner, especially in the context of unmodified EF-Ps, whose spectrum of rescued motifs remains to be elucidated. These findings have the potential to clarify the fundamental understanding behind the evolution of EF-Ps that do not necessitate modification, as their mode of action could be directed preferentially to a subset of polyproline motifs.

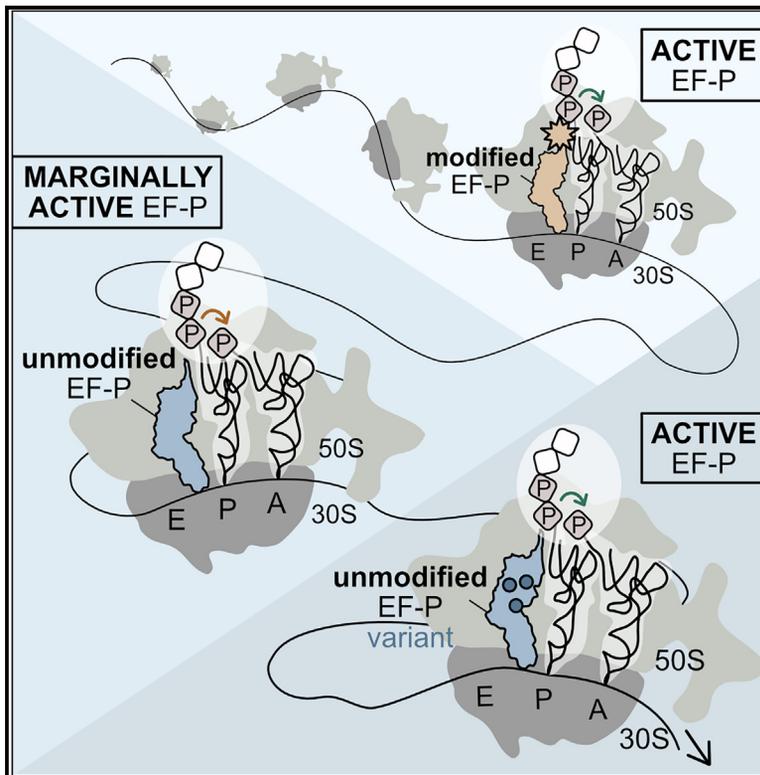
2 Decrypting the functional design of unmodified translation elongation factor P

Tomasiunaite U, Kielkowski P, Krafczyk R, Forné I, Imhof A, Jung K. Decrypting the functional design of unmodified translation elongation factor P. Cell Rep. 2024;43(5):114063.

DOI: <https://doi.org/10.1016/j.celrep.2024.114063>

Decrypting the functional design of unmodified translation elongation factor P

Graphical abstract



Authors

Urte Tomasiunaite, Pavel Kielkowski, Ralph Krafczyk, Ignasi Forné, Axel Imhof, Kirsten Jung

Correspondence

jung@lmu.de

In brief

Tomasiunaite et al. provide insights into the functional design of unmodified translation elongation factor P. They identify critical residues that determine the functionality of unmodified EF-Ps in *E. coli*. The results indicate that the activity of EF-P is not exclusively dependent on the post-translational modification.

Highlights

- The naturally unmodified EF-P of *Rhodospirillum rubrum* is active in *E. coli*
- Three amino acid positions are crucial to maintain the functionality of unmodified EF-P
- Marginally active EF-Ps can be engineered in fully functional unmodified variants in *E. coli*
- A heterologous expression platform for proteins with polyproline motifs is available



Article

Decrypting the functional design of unmodified translation elongation factor P

Urte Tomasiunaite,¹ Pavel Kielkowski,² Ralph Krafczyk,¹ Ignasi Forné,³ Axel Imhof,³ and Kirsten Jung^{1,4,*}¹Faculty of Biology, Microbiology, Ludwig-Maximilians-Universität München, 82152 Martinsried, Germany²Department of Chemistry, Institut für Chemische Epigenetik (ICEM), Ludwig-Maximilians-Universität München, 81375 Munich, Germany³Zentrallabor für Proteinanalytik, Biomedical Center Munich, Ludwig-Maximilians-Universität München, 82152 Martinsried, Germany⁴Lead contact*Correspondence: jung@lmu.de<https://doi.org/10.1016/j.celrep.2024.114063>

SUMMARY

Bacteria overcome ribosome stalling by employing translation elongation factor P (EF-P), which requires post-translational modification (PTM) for its full activity. However, EF-Ps of the PGKGP subfamily are unmodified. The mechanism behind the ability to avoid PTM while retaining active EF-P requires further examination. Here, we investigate the design principles governing the functionality of unmodified EF-Ps in *Escherichia coli*. We screen for naturally unmodified EF-Ps with activity in *E. coli* and discover that the EF-P from *Rhodocrobium vannielii* rescues growth defects of a mutant lacking the modification enzyme EF-P-(R)- β -lysine ligase. We identify amino acids in unmodified EF-P that modulate its activity. Ultimately, we find that substitution of these amino acids in other marginally active EF-Ps of the PGKGP subfamily leads to fully functional variants in *E. coli*. These results provide strategies to improve heterologous expression of proteins with polyproline motifs in *E. coli* and give insights into cellular adaptations to optimize protein synthesis.

INTRODUCTION

The translation of mRNAs to proteins at the ribosome is a highly conserved process in all domains of life. The translational efficiency is dynamic and depends on codon bias, tRNA accessibility, and the chemical nature of amino acids to be incorporated into the polypeptide chain.¹ Among the canonical amino acids, proline is an imino acid. This imino group cannot be optimally oriented for nucleophilic attack onto the carbonyl carbon atom of the P-site substrate during the transpeptidation reaction. In addition, proline is a poor donor during protein synthesis due to the rigidity of the pyrrolidine ring.^{2–5} Consequently, the presence of consecutive proline codons leads to a slowdown in protein synthesis and can cause ribosome stalling.^{6–9} Despite the hindrance, polyproline motifs play essential roles in the catalytic activity of enzymes, contribute to protein-protein interactions, and regulate protein copy numbers,^{7,10–13} rendering them indispensable components of the proteome of all organisms.

To facilitate the synthesis of proteins containing polyproline motifs, organisms from all kingdoms have evolved specialized translation factors. In bacteria, elongation factor P (EF-P) plays a crucial role in this process,^{6,7} while eukaryotes and archaea use initiation factor 5A (e/alf5A).¹⁴ These translation factors bind to the ribosome and enhance peptide bond formation during protein synthesis by stabilizing the P-site tRNA.^{15,16} However, to become fully functional, these translation factors must undergo post-translational modification (PTM).^{6,7,9,14} Notably, in eukaryotes and archaea, e/alf5A undergoes a PTM known as hypusination.¹⁷ On the other hand, bacteria use diverse and unusual

substrates and enzymes to modify their EF-Ps. For instance, in *Escherichia coli* and *Salmonella enterica*, a conserved lysine is β -lysylated by the EF-P-(R)- β -lysine ligase (EpmA) and L-lysine 2,3-aminomutase (EpmB), as well as hydroxylated by the EF-P hydroxylase (EpmC).^{18–22} Other modifications include rhamnosylation of an arginine by the arginine rhamnosyltransferase (EarP) in bacteria like *Pseudomonas putida* and *Shewanella oneidensis*^{23,24} and amino-pentanoylation by the EF-P modification enzyme YmfI in *Bacillus subtilis*^{25,26} (Figure 1A).

Interestingly, there is a group of bacteria, particularly from the Actinobacteria phylum, that possesses functional EF-Ps without the need for any PTM²⁷ (Figure 1A). We have shown that EF-Ps from the genera *Corynebacterium*, *Mycobacterium*, and *Streptomyces* function without PTM.²⁷ These EF-Ps have the otherwise modified lysine at the tip of a β -hairpin, flanked by two proline residues. It is suggested that the presence of this palindromic sequence Pro-Gly-Lys-Gly-Pro (PGKGP) within the hairpin confers rigidity and enables proper positioning of the protruding lysine 32, thereby stabilizing the acceptor arm of the tRNA. The PGKGP sequence has become a signature motif for this subfamily leading to its designation as the PGKGP subfamily of EF-P²⁷ (Figure 1A).

While it has been shown that a differentially modified EF-P variant can complement the deletion of *efp* in *E. coli* when co-expressed with the modification machinery,²³ unmodified EF-P variants were found to be inactive.²⁷ This observation suggests that certain features are responsible for conveying functionality without PTM. The current study was conducted to investigate the underlying design principles that enable the evolution of



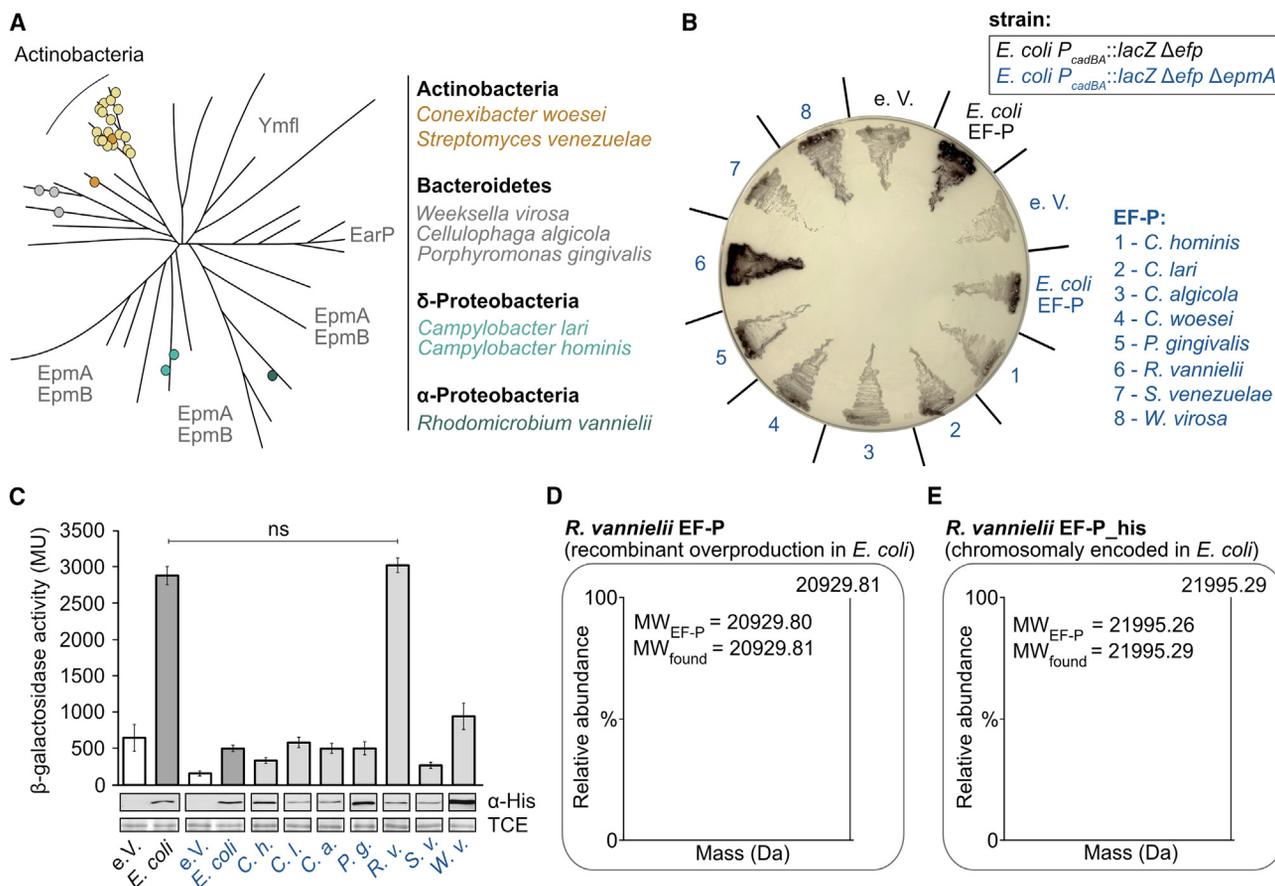


Figure 1. *R. vannielii* EF-P complements Δefp *E. coli* mutants without the necessity of a PTM

(A) Schematic representation of the phylogenetic tree of bacterial EF-Ps, adapted from Pinheiro et al., 2020.²⁷ PGKGP subfamily EF-Ps used in this study are marked with colored dots (Actinobacteria – orange, Bacteroidetes – gray, δ-Proteobacteria – light green, α-Proteobacteria – dark green). Yellow dots represent the remaining PGKGP-type EF-Ps from Actinobacteria.

(B and C) EF-P activity measurements using the P_{cadBA}::lacZ-based reporter assay with S-Gal® or o-nitrophenyl-β-D-galactopyranoside (o-NPG) as substrates for the β-galactosidase activity. Color code in (C) corresponds to the strains used in (B). The β-galactosidase activities are given in Miller units (MU). EF-P production was confirmed by western blot analysis using antibodies against the His-tag. Protein bands corresponding to a 72-kDa protein after staining with 2,2,2-trichloroethanol (TCE) were used as loading controls. The blot is shown split as the samples were originally loaded onto the gel in a different order. The bands were cut and rearranged to achieve the desired order in the graph. Error bars indicate the standard deviation (SD) of three independent biological replicates. Statistics: Student's unpaired two-sided t test (ns, p = 0.2808).

(D) Deconvoluted MS spectra of intact *R. vannielii* EF-P, recombinantly overproduced in *E. coli*.

(E) Deconvoluted MS spectra of intact His-tagged *R. vannielii* EF-P (chromosomally encoded), produced in *E. coli*. See also Table S1; Figures S1–S4.

functional EF-Ps without the need for PTMs. We were able to identify three crucial amino acid positions with pivotal roles in maintaining the functionality of unmodified EF-P. Using this knowledge, we then succeeded in engineering a functional unmodified EF-P in the bacterial workhorse *E. coli*. These findings pave the way for approaches to optimize synthesis of proteins containing polyproline motifs at lower metabolic cost in other laboratory and industrially relevant bacteria.

RESULTS

Screen for naturally unmodified EF-Ps of the PGKGP subfamily that are functional in *E. coli*

Previous study has demonstrated that EF-Ps, belonging to the PGKGP subfamily of various Actinobacteria, are functional

without PTM in their respective hosts, but they are non-functional in *E. coli*.²⁷ Notably, EF-Ps containing the PGKGP loop are also found in other bacterial phyla, which do not encode enzymes required for EF-P modification.²⁷ This observation has led to the hypothesis that these EF-Ps may also be functional in an unmodified state. Here, we screened several EF-Ps of the PGKGP subfamily of species from different phyla for their activity in *E. coli*. To this end, we chose *Campylobacter hominis* and *Campylobacter lari* as representatives of the δ-proteobacteria, *Cellulophaga algicola*, *Porphyromonas gingivalis*, and *Weeksella virosa* as representatives of the Bacteroidetes, *Conexibacter woesei* and *Streptomyces venezuelae* belonging to the Actinobacteria, and *Rhodomicrobium vannielii* as representatives of the α-Proteobacteria (Figure 1A). Despite an overall sequence identity of only 44% (Table S1), the PGKGP activation loop remains conserved.²⁷

To elucidate the activity of these EF-Ps in *E. coli*, a well-established reporter system was used.⁷ This reporter system is based on the finding that the transcriptional regulator CadC of *E. coli* is a polyproline protein that requires modified EF-P for its translation. In Δefp mutants, the copy number of CadC is too low to induce the *cadBA* promoter (tested as $P_{cadBA}::lacZ$). CadC is activated when *E. coli* cells are exposed to low pH in the presence of lysine. To rule out that heterologously produced EF-Ps are modified by the *E. coli*-specific modification system EpmA, a reporter strain with an additional deletion of *epmA* encoding the β -lysine ligase was used. This reporter strain (*E. coli* $P_{cadBA}::lacZ \Delta efp \Delta epmA$) was transformed with plasmids expressing EF-Ps of different representatives of the PGKGP subfamily. *E. coli* was grown at pH 5.8 and β -galactosidase activities were determined. On agar plates containing S-Gal[®] and ferric ions, cells producing β -galactosidase can easily be detected by black precipitates. As a control, *E. coli efp* was expressed both in *E. coli* $P_{cadBA}::lacZ \Delta efp$ and *E. coli* $P_{cadBA}::lacZ \Delta efp \Delta epmA$. Only cells expressing both *efp* and the modification machinery showed black precipitates (Figure 1B). Of the eight selected EF-Ps of the PGKGP subfamily, only EF-P of *R. vannielii* was active. In contrast to *E. coli* EF-P, *R. vannielii* EF-P produced black precipitates in the double deletion $\Delta efp \Delta epmA$ reporter strain (Figure 1B). To obtain quantitative results, β -galactosidase activity was determined using a colorimetric assay. *E. coli* EF-P was able to rescue the production of CadC and induce high β -galactosidase activity in the presence of the PTM machinery (Figure 1C). With the exception of *R. vannielii* EF-P, none of the other EF-Ps was able to complement the *E. coli* Δefp mutant. Remarkably, the activity of the reporter strain producing the *R. vannielii* EF-P was comparable to the activity of the strain with the modified *E. coli* EF-P (Figures 1C and S1).

We then used mass spectrometry-based proteomics (MS) to test whether *R. vannielii* EF-P undergoes PTM in *E. coli*. We first purified recombinantly overproduced *R. vannielii* EF-P using the *E. coli* BL21/pET_SUMO system, which had the advantage that the His-tag could be easily cleaved from the protein after purification. For the detection of the *R. vannielii* EF-P without a His-tag, a polyclonal antibody against this protein was generated (α -625) and used for western blot analysis. The calculated mass of *R. vannielii* EF-P (20,929.80 Da) was consistent with the measured mass of the intact protein (20,929.81 Da), indicating the absence of a modification (Figures 1D and S2A). This observation was confirmed by analysis of peptides after chymotrypsin digestion (Data S1). Overproduction of EF-P in *E. coli* may result in a high level of unmodified EF-P¹⁹ due to an imbalance between EF-P and its modifying enzymes. To avoid any artifact, we inserted *R. vannielii efp* (encoding a C-terminal his-tag) into the *E. coli* genome downstream of the native *efp* promoter and purified this protein (Figures S3A). The calculated mass of the His-tagged *R. vannielii* EF-P (21,995.26 Da) was consistent with the measured mass of the intact protein (21,995.29 Da) (Figures 1E and S2B), confirming that *R. vannielii* EF-P is unmodified in *E. coli*. This observation was further verified by analyzing the peptides after chymotrypsin digestion (Data S1). It has to be noted that the sequence available online (GenBank: NC_014664.1) for *R. vannielii* EF-P differs slightly

from the sequence we determined for the strain used (DSM 162/ATCC 17100) (Figure S4).

Overall, these data demonstrate that an unmodified EF-P variant can support the synthesis of a protein containing polyproline motifs in *E. coli*.

Unmodified *R. vannielii* EF-P is fully functional in *E. coli*

Next, we investigated the functionality of *R. vannielii* EF-P in *E. coli* in more detail. First, we conducted growth tests, including the *E. coli* wild type (WT) and the Δefp , $\Delta epmA$, and $\Delta efp \Delta epmA$ mutants, compared to the complemented $\Delta efp::R. v. efp$ and $\Delta efp::R. v. efp \Delta epmA$ mutants. In the latter two strains, *R. v. efp* is chromosomally encoded under control of the native *E. coli efp* promoter (Figure 2A). After 12 h of growth in complex (LB) medium, the cell density of the mutants Δefp and $\Delta efp \Delta epmA$ was much lower compared to the WT, which is in line with previous observations^{19,28} (Figure 2B). The mutant producing unmodified *E. coli* EF-P ($\Delta epmA$) showed slightly higher cell densities but still exhibited limited growth compared to the WT (Figure 2B). However, expression of *R. v. efp* in both the $\Delta epmA$ and the Δefp mutants suspended the growth defect (Figure 2B). To validate growth differences observed with the spot assay, we monitored the growth of the strains in LB medium at 37°C over time (Figure 2C). The growth defect of the mutant producing the unmodified *E. coli* EF-P ($\Delta epmA$) was overcome by the chromosomal insertion of the *R. v. efp* (Figure 2C). Only marginal differences in doubling times were detected between the WT and the mutants harboring the *R. v. efp* independent of the absence or presence of *epmA* (Figure 2C), confirming our hypothesis of a functional *R. v.* EF-P without the need for a PTM in *E. coli*. No significant differences were found in the amount of EF-P in *E. coli*, regardless of whether *epmA* was present or not (Figures S3A–S3D). These results emphasize that the observed growth defect of the $\Delta epmA$ mutant is due to the absence of the EF-P modification and is not a consequence of possible reduced EF-P levels. Compared to native EF-P, the level of *R. vannielii* EF-P was significantly lower in early exponential growth phase of *E. coli* (Figures S3C and S3D).

The absence of EF-P or its modifying enzymes in *E. coli* is associated with downregulation of proteins containing polyproline motifs and alterations of the proteome.⁹ Therefore, we tested whether unmodified *R. vannielii* EF-P inserted into these mutants could restore the WT proteome. Cells were grown to late exponential growth phase and prepared for proteomic analysis. The changes in protein levels were determined by liquid chromatography-tandem mass spectrometry proteomics using data dependent acquisition (DDA) and label-free quantification (LFQ).²⁹ Identified peptide fragments mapped to a total of 2,378 proteins (Data S2). The differences in proteome comparison between the tested strains can be seen in the principal component analysis (PCA) (Figure 2D). The analyzed proteomes cluster into three populations with large distances between them: (cluster i) WT together with the two mutants producing *R. vannielii* EF-P ($\Delta efp::R. v. efp$; $\Delta efp::R. v. efp \Delta epmA$), (cluster ii) mutants lacking EF-P (Δefp ; $\Delta efp \Delta epmA$), and (cluster iii) the mutant producing unmodified *E. coli* EF-P ($\Delta epmA$)

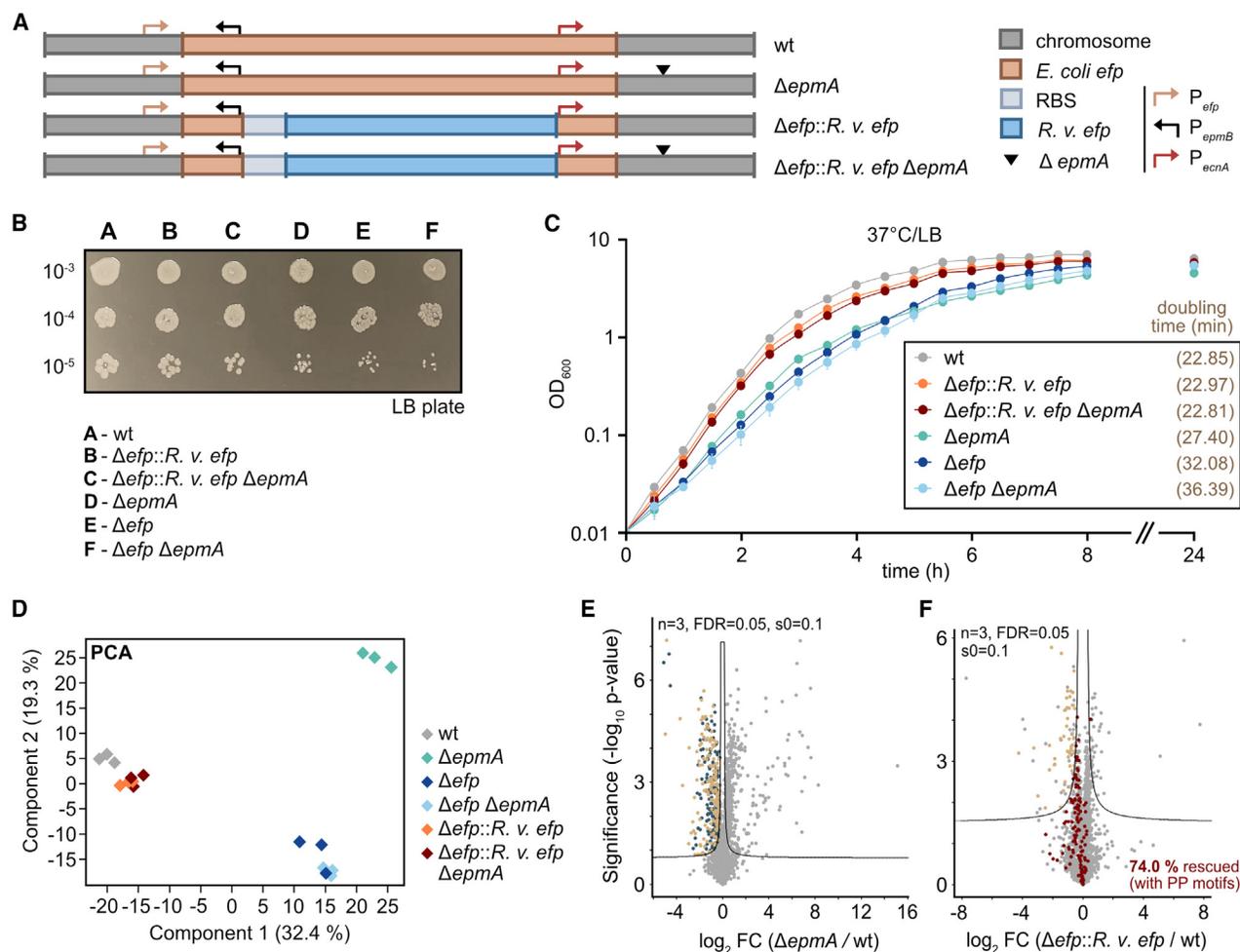


Figure 2. Phenotypic characterization of *E. coli* mutants complemented with *R. vannielii efp*

(A) Schematic overview of *E. coli* BW25113 *efp* deletion mutants with incorporated *R. vannielii efp*. Gene insertions are depicted in colored boxes and promoter locations as colored arrows.

(B) Spot assay of the *E. coli* mutants on an agar LB plate after incubation in LB overnight at 37°C.

(C) Growth of *E. coli* mutants in LB at 37°C. Data are mean values with error bars representing the standard deviation (SD) of four independent biological replicates. The doubling times of the individual strains are given in minutes (min).

(D) Proteome comparison of WT, $\Delta epmA$, Δefp , $\Delta efp \Delta epmA$, $\Delta efp::R. v. efp$, and $\Delta efp::R. v. efp \Delta epmA$ mutants using principal component analysis (PCA). Proteomes of three independent biological replicates were analyzed and are represented in rhombi with the same color.

(E and F) Volcano plot analysis, highlighting the downregulation of proteins with/without polyproline motifs (PP motifs) in *E. coli* mutant $\Delta epmA$ (E) and their rescue in mutant $\Delta efp::R. v. efp$ (F). Downregulated proteins without PP motifs are marked in dark blue dots; with PP motifs in light brown dots; rescued proteins in bordeaux red; and remaining identified proteins in gray. The x axes show the fold change (FC) of the mean value of the \log_2 protein intensity for each protein (LFQ) between two strains. The y axes show the significance level of the observed difference between the two strains ($-\log_{10} p$ value of the t test). See also Figures S3 and S5.

(Figure 2D). These data suggest that the proteome of mutants carrying the *R. v.* EF-P show the highest correlation with the WT *E. coli* strain.

We analyzed the proteomes in more detail with a specific focus on proteins with polyproline motifs. Comparisons between the proteomes of the $\Delta epmA$ mutant and the WT (Figure 2E) as well as between the Δefp and $\Delta efp \Delta epmA$ mutants and the WT (Figures S5A and S5B) revealed a strong protein scattering in the volcano plot indicating strong differences in the proteomes. Complementation of the Δefp mutant with *R. v. efp* (Figure 2F) or the replacement of *E. c. efp* with *R. v. efp* in the $\Delta epmA$

mutant (Figure S5D) reduced the scattering, suggesting high similarity to the WT proteome.

The *E. coli* Δefp and $\Delta efp \Delta epmA$ mutants were characterized by downregulation of proteins with polyproline motifs compared to the WT (Data S2). Replacement of *E. c. efp* with *R. v. efp* allowed translational rescue of 130 (75.6%) and 152 (78.4%) proteins containing polyproline motifs, downregulated in the Δefp and $\Delta efp \Delta epmA$ mutants, respectively (Data S2, Figures S5C and S5D).

It has been previously reported that unmodified *E. coli* EF-P is still able to rescue the translation of polyproline motifs but with

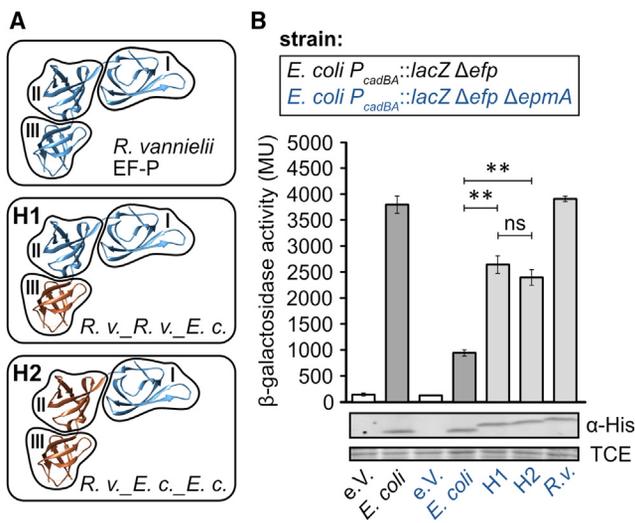


Figure 3. Activity of *E. coli* and *R. vannielii* EF-P hybrids
(A) Schematic overview of the constructed EF-P hybrids H1 (*R. v.*_R. *v.*_E. *c.*) and H2 (*R. v.*_E. *c.*_E. *c.*). Domains of *R. vannielii* EF-P (*R. v.*) are colored in blue and those of *E. coli* EF-P (*E. c.*) in brown.
(B) Activity measurements of the EF-P hybrids in *E. coli* using the $P_{cadBA}::lacZ$ -based reporter assay. The β -galactosidase activities are given in Miller units (MU). EF-P production was confirmed by western blot analysis using antibodies against the His-tag. Protein bands corresponding to a 72-kDa protein after staining with 2,2,2-trichloroethanol (TCE) were used as loading controls. 3D protein structures for simplified EF-P domain representation are taken and modified from PDB: 3A5Z. Error bars indicate the standard deviation (SD) of three independent biological replicates. Statistics: Student's unpaired two-sided t test (**** $p < 0.0001$; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ns $p > 0.05$). $\Delta efp \Delta efp$ (*E. coli* vs. *R. v.*_H1, ** $p = 0.0023$; *E. coli* vs. *R. v.*_H2, ** $p = 0.0021$; *R. v.*_H1 vs. *R. v.*_H2, ns $p = 0.2023$). See also Figure S6.

lower efficiency compared to the modified EF-P.^{6,9,30} Here, we also found that of all the downregulated proteins with polyproline motifs in the Δefp mutant (172 proteins) and the $\Delta efp \Delta efp$ mutant (194 proteins), 47 (27.3%) and 59 (30.4%), respectively, were not downregulated in cells producing unmodified *E. coli* EF-P (Δefp mutant) (Data S2). Importantly, *R. v.* EF-P was able to outperform the rescue potential of unmodified EF-P in *E. coli* by translationally rescuing 128 (74%) proteins with polyproline motifs (Figure 2F; Data S2).

Elucidation of the functional design principles of unmodified EF-P in *E. coli*

Our next aim was to uncover the design principles governing the full functionality of unmodified EF-Ps in *E. coli*. Accordingly, we sought to understand why not all EF-Ps of the PGKGP subfamily are active in *E. coli*. Previous reports showed that the bacterial EF-P consists of three domains, forming a shape similar to a tRNA.^{31,32} To identify the essential domains for EF-P, we constructed *E. coli* and *R. vannielii* EF-P hybrids (Figure 3A) and measured their activities using the $P_{cadBA}::lacZ$ reporter system. To determine the boundaries of each domain, we considered predicted secondary structures and 3D protein structures. Linker regions connecting the three domains were selected as fusion points to minimize effects on the EF-P 3D structure (Figures S6A–S6D; adapted from Uniprot: P0A6N4 and AlphaFold:

AF-P0A6N4-F1^{33–35}). The exchange of domains I and II of *E. c.* EF-P with the respective domains of *R. v.* EF-P (H1: *R. v.*_R. *v.*_E. *c.*) more than doubled the activity of *E. coli* EF-P in the Δefp background (Figure 3B). When only domain I of *E. coli* EF-P was replaced by the corresponding *R. v.* EF-P domain (H2: *R. v.*_E. *c.*_E. *c.*), the same increase in activity was observed. These observations suggest that domain I is of high importance for an active unmodified EF-P in *E. coli*.

To identify these amino acids, we used a synthetic molecular engineering approach, starting with multiple sequence alignments of EF-Ps of the PGKGP subfamily and *E. coli* (Figure 4A), followed by the construction of EF-P variants using site-directed mutagenesis and screening of those with highest activity using the $P_{cadBA}::lacZ$ reporter assay (Figure 4B). The binding of EF-P to the negatively charged ribosomal parts (rRNA) is mediated mainly by charged amino acid side chains.¹⁵ Therefore, we focused on amino acids that differ in charge between functional EF-Ps (*E. c.* EF-P and *R. v.* EF-P) and those that are unfunctional (remaining PGKGP subfamily EF-Ps) in *E. coli*, in particular positions 27, 28, 50, 56, 60, and 65 (Figure 4A).

S. venezuelae EF-P, a representative of the unmodified actinobacterial EF-Ps with low activity in *E. coli*,²⁷ was chosen for molecular engineering (Figure 4C). Out of all substitutions, Q27E, H28F, V50R, N56K, K60S, and T65D (numbering according to the *S. venezuelae* EF-P sequence), only the replacement of valine at position 50 to arginine (V50R) significantly increased the activity of *S. venezuelae* EF-P in *E. coli* (Figure 4C). This result suggested that a single amino acid substitution can already impact the overall functionality of EF-P. In the multiple sequence alignment of EF-Ps from the PGKGP subfamily, we detected a pattern for position 50 (Figure 4A). Strikingly, in almost all EF-P members of the PGKGP subfamily examined, only uncharged and nonpolar amino acids are present at position 50, with the exception of *R. vannielii* EF-P, which contains a positively charged lysine (K). The EF-P of *E. coli* also has a positively charged amino acid (R, arginine) at this position (Figure 4A). We tested the importance of the positively charged amino acid at position 50. Replacement of lysine by a neutral amino acid (K50V) reduced the activity of *R. vannielii* EF-P in *E. coli* to only 15%, whereas substitution by another positively charged amino acid, arginine (K50R), allowed 70% of the original activity (Figure 4D). Based on these results, we investigated the significance of a positively charged amino acid at position 50 in other potentially unmodified EF-Ps from the PGKGP subfamily. Thus, we constructed EF-P variants of *Cellulophaga algicola*, *Campylobacter hominis*, *Campylobacter lari*, *Conexibacter woesei*, *Porphyromonas gingivalis*, and *Weeksella virosa* EF-Ps with substitutions of valine (V) or isoleucine (I) by positively charged arginine (R) and analyzed their activities using the $P_{cadBA}::lacZ$ reporter assay. Indeed, these amino acid substitutions significantly increased the activity of the EF-P variants of *C. hominis*, *P. gingivalis*, and *W. virosa* (Figure 4E), being in line with results for *S. venezuelae* EF-P (Figure 4C) and highlighting the importance of a positively charged amino acid at position 50 for the activity of unmodified EF-P in *E. coli*. Nevertheless, the introduction of a positively charged arginine (R) at position 50 was not sufficient to increase the activity of EF-P of *C. algicola* and *C. lari* although the proteins were synthesized (Figure 4E).

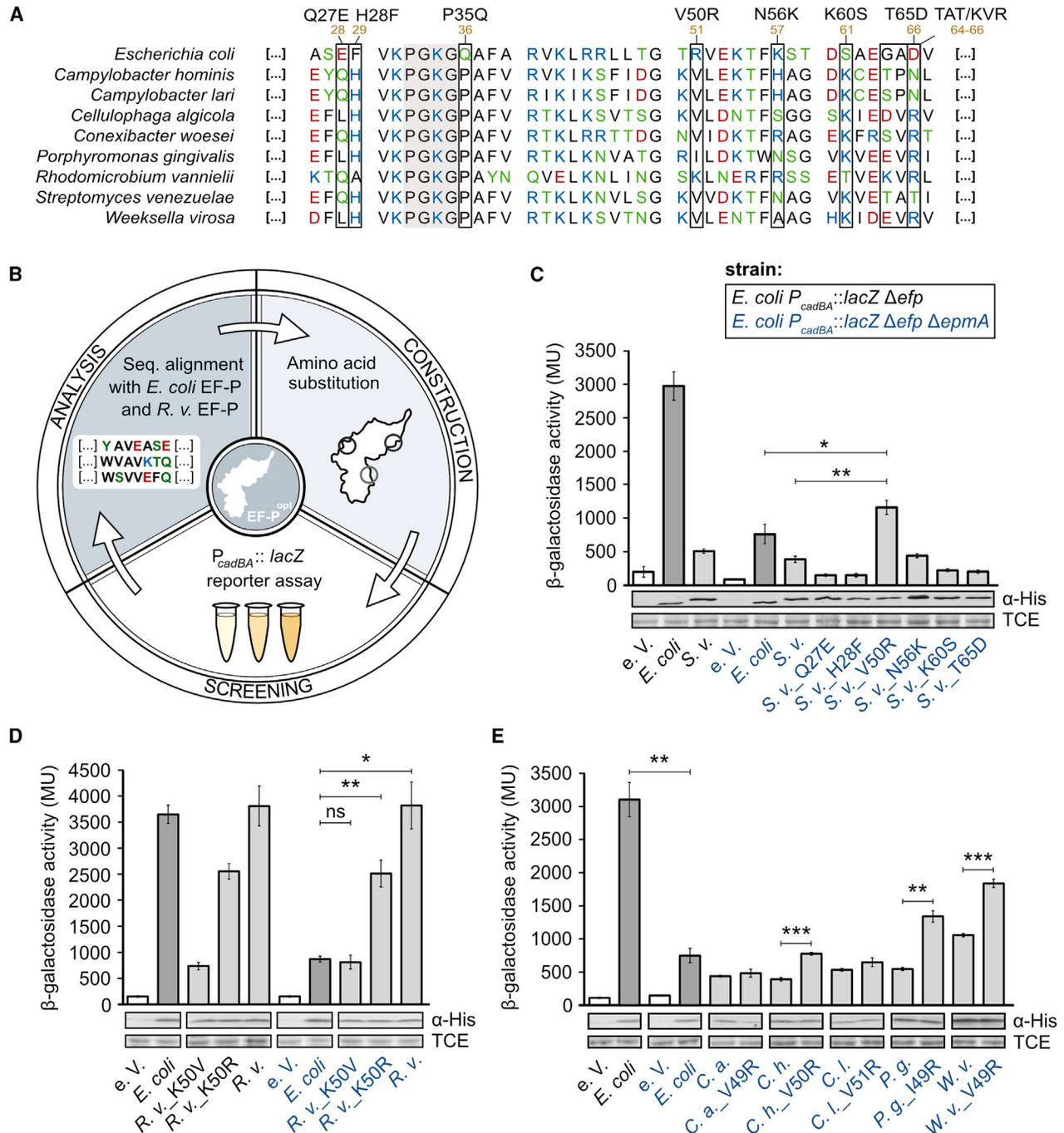


Figure 4. A positively charged amino acid at position 50 has an impact on the activity of PGKGP subfamily EF-Ps in *E. coli*

(A) Multiple sequence alignment of representative EF-Ps of the PGKGP subfamily and *E. coli*. Positions selected for substitution are highlighted and numbered according to *S. venezuelae* EF-P. The brown numbers indicate the positions in *E. coli*.

(B) Workflow for synthetic engineering of unmodified EF-P.

(C–E) Activity measurements of *S. venezuelae* EF-P (C), *R. vannielii* EF-P (D), and PGKGP subfamily EF-P variants (E) in *E. coli* using the $P_{cadBA}::lacZ$ -based reporter assay. Color code in (D) and (E) corresponds to the strains used in (C). The β -galactosidase activities are given in Miller units (MU). EF-P production was confirmed by western blot analysis with antibodies against the His-tag. Protein bands corresponding to a 72-kDa protein after staining with 2,2,2-trichloroethanol (TCE) were used as loading controls. C. a. – *Cellulophaga algicola*, C. h. – *Campylobacter hominis*, C. l. – *Campylobacter lari*, E. c. – *E. coli*, P. g. – *Porphyromonas gingivalis*, S. v. – *Streptomyces venezuelae*, W. v. – *Weeksella virosa*, EF-P^{opt} – optimized EF-P. Error bars indicate the standard deviation (SD) of three

(legend continued on next page)

Overall, these data show that a single amino acid substitution has an effect on the activity of originally non-functional EF-Ps in *E. coli*.

A few amino acid changes are sufficient to generate a fully functional, unmodified EF-P in *E. coli*

The placement of a positively charged arginine at position 50 resulted in increased activity of *S. venezuelae* EF-P (variant A1) in *E. coli*, although not to the level observed for the modified *E. coli* EF-P (Figures 4C and 5A). Thus, we continued with the synthetic molecular engineering approach (Figure 4B) and investigated other amino acid positions. In particular, we focused on the role of proline 34 (P34), which—together with proline 30 (P30)—is part of the β -hairpin of the PGKGP subfamily EF-Ps.²⁷ Remarkably, substitution of P34 by glutamine (Q) enhanced the activity of the *S. venezuelae* EF-P variant (variant A2) (Figure 5A).

Analyzing all positions selected for substitution (Figure 4A), we observed a striking difference between amino acids at position 65 (numbering according to *S. venezuelae* EF-P) in *E. coli* and *R. vannielii*: there is a negative charge in *E. coli* EF-P (aspartic acid, D66) and a positive charge in *R. vannielii* EF-P (arginine, R65). Substitution T65D did not yield active *S. venezuelae* EF-P in *E. coli* (Figure 4C), which prompted us to replace amino acids of the region 63–65_TAT (T63–A64–T65, threonine – alanine – threonine) with those found in *R. vannielii* EF-P, 63–65_KVR (K63–V64–65R, lysine – valine – arginine). This substitution significantly improved the functionality of *S. venezuelae* EF-P in *E. coli* (variant A3) (Figure 5A).

We further investigated whether the combination of these substitutions could have an additive effect to increase the activity of *S. venezuelae* EF-P (Figure 4B). Indeed, the activity of variant A1_A2 (V50R_P35Q) was significantly higher than that of variant A1 (Figure 5A). Similarly, variants A1_A3 (V50R_63–65_KVR) and A2_A3 (P35Q_63–65_KVR) showed increased activities. Finally, the activity of the variant A1_A2_A3 (P35Q_V50R_63–65_KVR) was almost 6-fold higher than that of the original WT protein (*S. v.*) (Figure 5A). Nonetheless, the activity of *S. venezuelae* EF-P_A1_A2_A3 remained significantly lower than that of the modified *E. coli* EF-P (Figure 5A).

We found weaker signals in the western blot for *S. venezuelae* EF-P_A1_A2_A3 compared to *E. coli* EF-P (Figure 5A), which indicates lower production and could explain the lower activity. This observation prompted us to search for other EF-P PGKGP family members that are better expressed in *E. coli*. Since the EF-Ps derived from *P. gingivalis* and *W. virosa* were produced very well in *E. coli* (Figure 4E), we decided to continue synthetic molecular engineering with them. Variants with P34Q or E62K substitutions in both EF-Ps had significantly higher activities than the corresponding native proteins (Figure 5B). It should be noted that in both proteins, only the first amino acid (E62K) had to be substituted to obtain the crucial KVR motif (Figure 4A). Remarkably, the *W. virosa* EF-P variant with all three amino

acid substitutions (*W. virosa* EF-P_S1_S2_S3; EF-P_P34Q_V49R_E62K) reached the activity level of the modified *E. coli* EF-P (Figure 5C). The activity of *W. virosa* EF-P_S1-S3 in *E. coli* was also not affected by the presence of the modifying enzyme EpmA (Figure 5D).

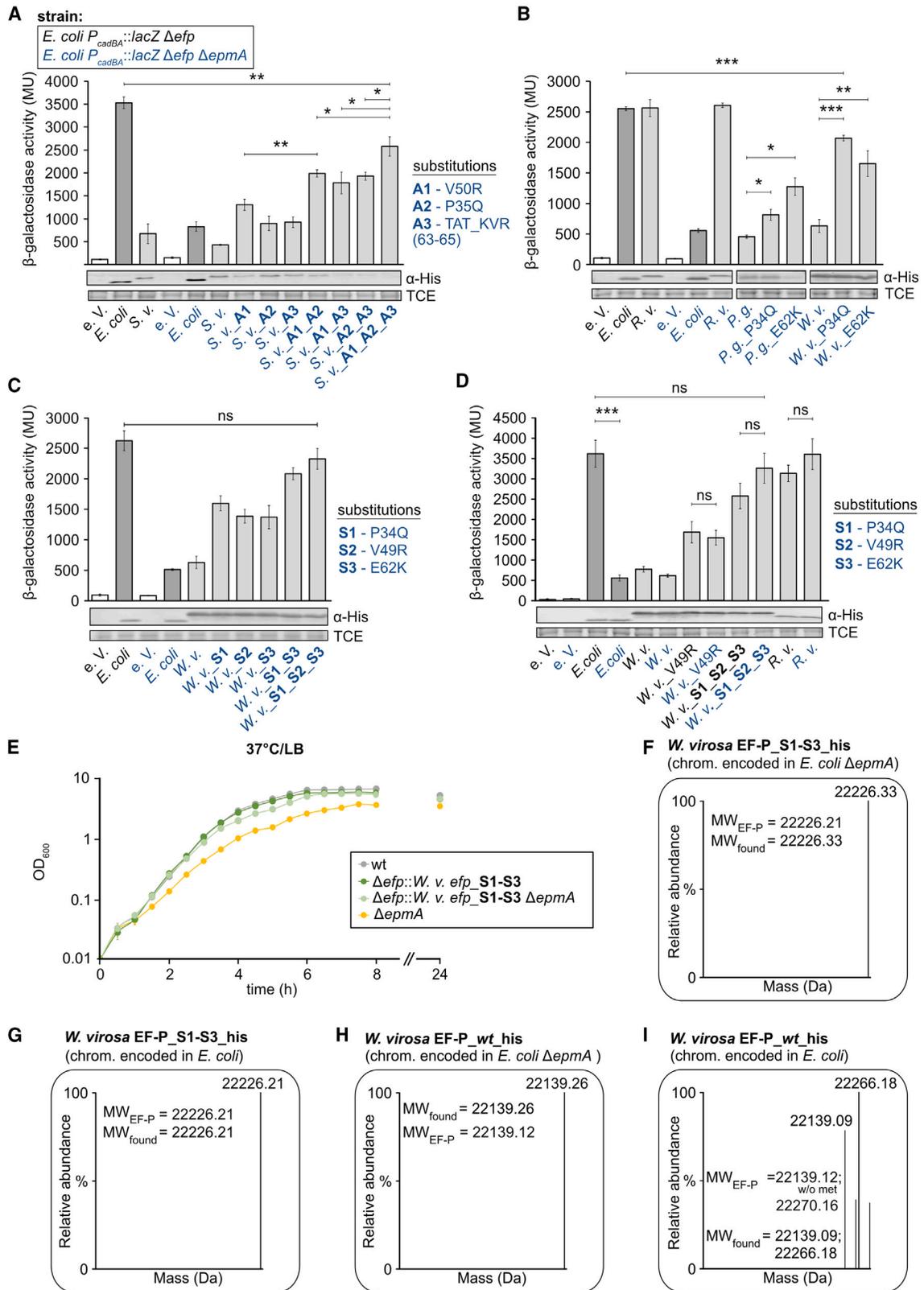
To this end, we characterized *E. coli* strains expressing *W. virosa* EF-P instead of their native modified or unmodified EF-P. For this purpose, *E. coli* mutants with chromosomally inserted *W. virosa* *efp*_S1-S3 were constructed (Figure S3E), and western blot analysis revealed sufficiently high production of this variant (Figures S3F and S3G). *W. virosa* EF-P_S1-S3 was able to rescue the growth phenotypes observed for the Δ *epmA* *E. coli* mutant independent of the presence of *epmA* (Figure 5E). The chromosomally encoded *W. virosa* EF-P_S1-S3 was purified by affinity chromatography. MS analysis of the intact protein confirmed that this EF-P variant was unmodified. The calculated protein mass of the *W. virosa* EF-P_S1-S3 variant was consistent with the measured intact protein mass, regardless of the presence of EpmA (calculated: 22,226.21 Da; measured: Δ *epmA* 22,226.33 Da; *epmA*⁺ 22,226.21 Da) (Figures 5F, 5G, S2C, and S2D). It should be noted that the measured protein masses are given without the first methionine, a phenomenon frequently observed in previous studies.^{36,37} To test whether the amino acid substitutions affected possible modification events, we also analyzed the intact protein mass of purified *W. virosa* WT EF-P, and the calculated protein mass (22,139.12 Da) was consistent with the measured intact protein mass, whether or not the cells expressed EpmA (Δ *epmA*: 22,139.26 Da; *epmA*⁺: 22,139.09 Da) (Figures 5H, 5I, S2E, and S2F). We observed two peaks in the mass spectrometry measurements of *W. virosa* WT EF-P (Figure 5I). These peaks correspond to the mass of *W. virosa* EF-P with (22,266.18 Da) and without (22,139.09 Da) methionine.

Overall, the finding that the EF-Ps of *S. venezuelae* and *W. virosa* can be significantly activated by a small number of substitutions demonstrates that certain design principles found in *R. vannielii* EF-P can be transferred to other isoforms within the same group. Our data show that three amino acid substitutions are sufficient to convert an inactive EF-P into a fully functional unmodified variant. Moreover, an *E. coli* mutant is now available as a host for heterologous production of proteins containing polypyrroline motifs with lower metabolic and energy costs, since EF-P no longer requires PTM.

DISCUSSION

For optimal synthesis of proteins with polypyrroline stretches, the EF-P of *E. coli* requires a PTM catalyzed by the modifying enzymes EpmA, EpmB, and EpmC.^{18–21} EpmB converts the precursor (S)- α -lysine to (R)- β -lysine, which is then ligated by EpmA in an ATP-dependent condensation reaction to the ϵ -amino group of Lys34 in EF-P. EpmC subsequently catalyzes the hydroxylation of Lys34. Except for hydroxylation, none of

independent biological replicates. Statistics: Student's unpaired two-sided t test (*** $p < 0.0001$; ** $p < 0.001$; * $p < 0.01$; * $p < 0.05$; ns $p > 0.05$). Δ *efp* Δ *epmA* (*E. coli* vs. *S. v.*_V50R, * $p = 0.0396$; *S. v.* vs. *S. v.*_V50R, ** $p = 0.0035$) (C); Δ *efp* Δ *epmA* (*E. coli* vs. *R. v.*_K50V, ns $p = 0.6099$; *E. coli* vs. *R. v.*_K50R, ** $p = 0.0092$; *E. coli* vs. *R. v.*, * $p = 0.0104$) (D); Δ *efp* Δ *epmA* (*C. h.* vs. *C. h.*_V50R, *** $p = 0.0001$; *P. g.* vs. *P. g.*_I49R, ** $p = 0.0037$; *W. v.* vs. *W. v.*_V49R, *** $p = 0.00099$), *E. coli* (in Δ *efp*) vs. *E. coli* (in Δ *efp* Δ *epmA*), ** $p = 0.0022$ (E). See also Figure S1.



(legend on next page)

the modification steps can be bypassed without noticeable effects on translational rescue of proteins with polyproline motifs. Even modification of EF-P with α -lysine instead of β -lysine decreases its activity.³⁸ Notably, the absence of modification in EF-P impairs the synthesis of proteins with polyproline motifs.⁹ This study revealed that only one-third of proteins with polyproline motifs found to be downregulated in the Δefp mutant were rescued by unmodified E-FP ($\Delta efpA$ mutant) (Data S2). Thus, while the lack of PTM does not completely prevent EF-P from its original mode of action, it impacts its efficiency. The PTM of *E. coli* EF-P is associated with metabolic (β -lysine) and energetic costs (ATP) as well as the production of the modification enzymes EpmA, EpmB, and EpmC.

The relatively high abundance of diproline (XPPX) motif-containing proteins in *E. coli* (2,101 motifs, 0.49 motif/protein),¹³ shown to cause ribosome stalling,⁹ underscores the need for an EF-P with maximal functionality to maintain proteome homeostasis, especially at high growth rates.²⁸ Many representatives of Actinobacteria have an even higher number of XPPX motifs within their proteomes.²⁷ In *Streptomyces coelicolor* and *Mycobacterium tuberculosis*, the number of XPPX motifs even surpasses the number of encoded proteins (1.08 motif/protein and 1.17 motif/protein, respectively).²⁷ Accordingly, one would expect PTMs to be crucial in these bacteria. However, the actinobacterial EF-Ps are not modified.²⁷ Both *S. coelicolor* and *M. tuberculosis* grow very slowly.^{27,39–41} In these slowly growing bacteria, an unmodified, and therefore metabolically and energetically less demanding, EF-P could provide a selective advantage when coping with the production of a large number of proteins with polyproline motifs.

However, actinobacterial unmodified EF-Ps could not rescue the translation of PP proteins in *E. coli*.²⁷ To find out why certain EF-Ps exhibit functionality in their native host without PTM, but are non-functional in *E. coli*, we analyzed the activity of EF-Ps belonging to the PGKGP subfamily from different phyla (Figure 1A). We found that of all eight tested PGKGP subfamily EF-Ps, the EF-P from *R. vannielii* was able to rescue the translation of proteins with polyproline motifs in *E. coli* (Figures 1B and 1C). Importantly, the EF-P from

R. vannielii retained its functionality in an unmodified state in *E. coli* (Figures 1D and 1E). *E. coli* strains expressing *R. vannielii efp* rescued growth defects observed in *E. coli* mutants lacking *efp* or *epmA*. Moreover, the proteome profiles of *R. vannielii* EF-P-producing strains closely resemble those of the WT (Figures 2B–2F).

EF-Ps have conserved regions required for ribosomal contact and adopt a tRNA-like structure, hinting at the evolutionary origin of these proteins.^{31,42} The analogy of EF-P to tRNA, which is one of the building blocks of the translational machinery, likely increased the ability of this protein to bind between the ribosomal E- and P-sites⁴² (Figure S7A). Therefore, we hypothesized that inactive EF-Ps might essentially be functional in *E. coli* but hindered from contributing to the translation process by weak binding to the ribosome or insufficient contact to the tRNA. To investigate this, we developed a synthetic molecular engineering workflow involving the substitution of amino acids based on the sequence of the functional *R. vannielii* EF-P (Figure 4B). By replacing single amino acids at positions 35, 50, and 63–65 (numbered according to *S. venezuelae* EF-P, Figure 4A), we successfully converted initially non-functional EF-Ps of *S. venezuelae*, *P. gingivalis*, and *W. virosa* into functional ones in *E. coli* (Figures 4C, 4E, and 5A–5D). Previous studies have reported the importance of certain amino acids at position 35 for EF-P activity. Pinheiro et al. showed that a proline at position 34 (position 35 in *S. venezuelae*) is necessary for EF-P to function with highest activity in *Corynebacterium glutamicum*, as substitutions to alanine, glutamine, glycine, and asparagine decreased activity of the corresponding EF-P variants.²⁷ In *E. coli*, a glutamine (Q) at the corresponding position has been shown to be required for proper function of EF-P.³⁰ These observations are in line with our results, as substitution of this proline to glutamine significantly increased the activity of the corresponding *S. venezuelae*, *P. gingivalis*, and *W. virosa* variants in *E. coli* (Figures 5A–5C). It is suggested that the amino acid at position 35 is essential for the correct positioning of the conserved lysine (either modified or non-modified) in the β -hairpin to get contact with P-site tRNA.^{15,27} The presence of a glutamine at this particular position could be important to

Figure 5. Optimized unmodified EF-P variants are active in *E. coli*

(A–D) Activity measurements of EF-P variants from *S. venezuelae* (A), *P. gingivalis* (B), and *W. virosa* (B, C, and D) in *E. coli* using the $P_{cadBA}::lacZ$ -based reporter assay. Color code in (B), (C), and (D) corresponds to the strains used in (A). The β -galactosidase activities are given in Miller units (MU). EF-P production was confirmed by western blot analysis using antibodies against the His-tag. Protein bands corresponding to a 72-kDa protein after staining with 2,2,2-trichloroethanol (TCE) were used as loading controls.

(E) Growth of *E. coli* mutants expressing *W. virosa*_S1-S3 *efp* in LB at 37°C.

(F and G) Deconvoluted mass spectra of the chromosomally encoded His-tagged *W. virosa* EF-P_S1-S3, heterologously produced in *E. coli* $\Delta efpA$ or *E. coli* WT, respectively. It should be noted that the calculated mass of *W. virosa* corresponds to a protein without the first methionine.

(H and I) Deconvoluted mass spectra of the chromosomally encoded His-tagged *W. virosa* EF-P WT, heterologously produced in *E. coli* $\Delta efpA$ or *E. coli* WT, respectively. Calculated masses correspond to *W. virosa* EF-P_WT_6xhis = 22270.16 Da and *W. virosa* EF-P_WT_6xhis without the first methionine = 22139.12 Da. Error bars in (A)–(D) indicate the standard deviation (SD) of at least three independent biological replicates. Data in (E) are mean values with error bars representing the standard deviation (SD) of four independent biological replicates. Statistics: Student's unpaired two-sided t test (**** $p < 0.0001$; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ns $p > 0.05$). Calculated p values: $\Delta efp \Delta efpA$ (*S. v.*_A1 vs. *S. v.*_A1_A2, ** $p = 0.0045$; *S. v.*_A1_A2 vs. *S. v.*_A1_A2_A3, * $p = 0.0454$; *S. v.*_A1_A3 vs. *S. v.*_A1_A2_A3, * $p = 0.0250$; *S. v.*_A2_A3 vs. *S. v.*_A1_A2_A3, * $p = 0.0349$), *E. coli* (in Δefp) vs. *S. v.*_A1_A2_A3 (in $\Delta efp \Delta efpA$), ** $p = 0.0096$ (a); $\Delta efp \Delta efpA$ (*P. g.* vs. *P. g.*_P34Q, * $p = 0.0244$; *P. g.* vs. *P. g.*_E62K, * $p = 0.0132$; *W. v.* vs. *W. v.*_P34Q, *** $p = 0.0006$; *W. v.* vs. *W. v.*_E62K, ** $p = 0.0094$), *E. coli* (in Δefp) vs. *W. v.*_P34Q (in $\Delta efp \Delta efpA$), *** $p = 0.0005$ (B); *E. coli* (in Δefp) vs. *W. v.*_S1_S2_S3 (in $\Delta efp \Delta efpA$), ns $p = 0.1484$ (C). *E. coli* (in Δefp) vs. *E. coli* (in $\Delta efp \Delta efpA$), *** $p = 0.0003$; *W. v.*_V49R (in Δefp) vs. *W. v.*_V49R (in $\Delta efp \Delta efpA$), ns $p = 0.4876$; *W. v.*_S1_S2_S3 (in Δefp) vs. *W. v.*_S1_S2_S3 (in $\Delta efp \Delta efpA$), ns $p = 0.0506$; *R. v.* (in Δefp) vs. *R. v.* (in $\Delta efp \Delta efpA$), ns $p = 0.1176$; *E. coli* (in Δefp) vs. *W. v.*_S1_S2_S3 (in $\Delta efp \Delta efpA$), ns $p = 0.2580$ (D). See also Figures S1–S3 and S7.

allow contacts with the *E. coli* ribosome (23S rRNA) and consequently stabilization and/or orientation of the β -hairpin with the essential lysine at the tip in the absence of a modification (Figure S7B).

Another interesting finding was that substitution of amino acids at position 50 into an arginine (R) resulted in an increase in activity for all tested variants (Figure 4E). This positively charged amino acid is predicted to be in close contact with the ribosome^{15,16,42} and therefore important for the corresponding EF-P to establish contacts with the ribosome (Figure S7C). We noted that a positively charged amino acid at position 50 is also found in some EF-P orthologs, the eukaryotic/archaeal initiation factor 5A (e/alf5A).³¹ Finally, the substitution of GAD (positions 64_66 in *E. coli*) against KVR had a major impact on the functionality of the unmodified EF-P variants and might be explained by a stabilization of the interaction between EF-P and the P-site tRNA¹⁵ (Figure S7D).

In conclusion, bacteria, archaea, and eukaryotes have evolved different ways to cope with the synthesis of proteins containing polyproline motifs. There are functional EF-Ps that require PTMs, as well as those that do not necessitate modification. According to previous knowledge, it was suggested that the tip of the modified loop of EF-P reaches the carbonyl group of the P-site tRNA substrate in the peptidyl transferase center and stabilizes it for the transpeptidation reaction.^{2,6,43} Our study reveals that naturally occurring unmodified EF-Ps, which have low activity in *E. coli*, can be made functional by altering distinct amino acids. This suggests that the alleviation of the polyproline-dependent ribosome stalling by EF-P may not be exclusively related to the PTM itself but rather to the affinity of EF-P to the ribosome and/or the tRNA-Pro in the P-site.

The modification of EF-P with β -lysine in *E. coli* requires not only a substrate but also energy-intensive enzymes to catalyze this process. Here, we describe an EF-P that is functional in *E. coli* without the need for a PTM, making it attractive for future studies aiming to reallocate cellular energy for other energy-consuming processes in bacteria. An optimally functioning EF-P is important for proteome homeostasis and bacterial virulence.^{44,45} Thus, decrypting the functional design of unmodified EF-P in *E. coli* not only directs protein production in a resource-efficient manner for scientific and industrial applications but also extends the fundamental understanding of the functional principles of this translation factor.

Limitations of the study

Our study provides insights into the functional design of unmodified translation elongation factor P. There are limitations of the study that we would like to address. All of the EF-P variants we generated were tested for production using the western blot technique. As this is a semi-quantitative technique, we could not calculate the specific activity of the variants. In addition, we did not test their individual stability and conformational heterogeneity. We do not exclude other possible substitution combinations that would lead to active unmodified EF-Ps in *E. coli*. To gain insights into the molecular mechanism for the rescue of ribosome stalling by unmodified EF-P, additional data are required, such as determining the affinity of EF-P variants to bind to the ribosome.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
 - Plasmid and bacterial strain construction
 - β -Galactosidase activity assays
 - Protein purification and structural analysis
 - SDS-PAGE and Western Blot analysis
 - Mass spectrometry for identification of modification status
 - Mass spectrometry for identification of modification status
 - Mass spectrometry for proteome analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2024.114063>.

ACKNOWLEDGMENTS

We thank Prof. Bernhard Schink from University of Konstanz for the support in handling *R. vannielii*. We also thank Haoyu Chen and Thomas Krause for excellent technical assistance. This work was financially supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grant JU270/20-1, project number 449926427, RTG2062 (Molecular Principles of Synthetic Biology) to K.J., and SFB 1309, project number 325871075, to K.J., P.K., and A.I.

AUTHOR CONTRIBUTIONS

Conceptualization, U.T. and K.J.; methodology, U.T., P.K., I.F., A.I., and K.J.; investigation, U.T., P.K., R.K., and I.F.; writing – original draft, U.T. and K.J.; funding acquisition and resources, P.K., A.I., and K.J.; supervision, K.J.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 20, 2023

Revised: January 17, 2024

Accepted: March 21, 2024

Published: April 17, 2024

REFERENCES

1. Hanson, G., and Collier, J. (2018). Codon optimality, bias and usage in translation and mRNA decay. *Nat. Rev. Mol. Cell Biol.* 19, 20–30. <https://doi.org/10.1038/nrm.2017.91>
2. Doerfel, L.K., Wohlgemuth, I., Kubyskhin, V., Starosta, A.L., Wilson, D.N., Budisa, N., and Rodnina, M.V. (2015). Entropic Contribution of Elongation Factor P to Proline Positioning at the Catalytic Center of the

- Ribosome. *J. Am. Chem. Soc.* **137**, 12997–13006. <https://doi.org/10.1021/jacs.5b07427>.
3. Muto, H., and Ito, K. (2008). Peptidyl-prolyl-tRNA at the ribosomal P-site reacts poorly with puromycin. *Biochem. Biophys. Res. Commun.* **366**, 1043–1047. <https://doi.org/10.1016/j.bbrc.2007.12.072>.
 4. Pavlov, M.Y., Watts, R.E., Tan, Z., Cornish, V.W., Ehrenberg, M., and Forster, A.C. (2009). Slow peptide bond formation by proline and other N-alkylamino acids in translation. *Proc. Natl. Acad. Sci. USA* **106**, 50–54. <https://doi.org/10.1073/pnas.0809211106>.
 5. Wohlgemuth, I., Brenner, S., Beringer, M., and Rodnina, M.V. (2008). Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. *J. Biol. Chem.* **283**, 32229–32235. <https://doi.org/10.1074/jbc.M805316200>.
 6. Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M.V. (2013). EF-P Is Essential for Rapid Synthesis of Proteins Containing Consecutive Proline Residues. *Science* **339**, 85–88. <https://doi.org/10.1126/science.1229017>.
 7. Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N., and Jung, K. (2013). Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science* **339**, 82–85. <https://doi.org/10.1126/science.1228985>.
 8. Tanner, D.R., Cariello, D.A., Woolstenhulme, C.J., Broadbent, M.A., and Buskirk, A.R. (2009). Genetic Identification of Nascent Peptides That Induce Ribosome Stalling. *J. Biol. Chem.* **284**, 34809–34818. <https://doi.org/10.1074/jbc.M109.039040>.
 9. Peil, L., Starosta, A.L., Lassak, J., Atkinson, G.C., Virumäe, K., Spitzer, M., Tenson, T., Jung, K., Remme, J., and Wilson, D.N. (2013). Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. *Proc. Natl. Acad. Sci. USA* **110**, 15265–15270. <https://doi.org/10.1073/pnas.1310642110>.
 10. Adzhubei, A.A., Sternberg, M.J.E., and Makarov, A.A. (2013). Polyproline-II helix in proteins: structure and function. *J. Mol. Biol.* **425**, 2100–2132. <https://doi.org/10.1016/j.jmb.2013.03.018>.
 11. Starosta, A.L., Lassak, J., Peil, L., Atkinson, G.C., Woolstenhulme, C.J., Virumäe, K., Buskirk, A., Tenson, T., Remme, J., Jung, K., and Wilson, D.N. (2014). A conserved proline triplet in Val-tRNA synthetase and the origin of elongation factor P. *Cell Rep.* **9**, 476–483. <https://doi.org/10.1016/j.celrep.2014.09.008>.
 12. Motz, M., and Jung, K. (2018). The role of polyproline motifs in the histidine kinase EnvZ. *PLoS One* **13**, e0199782. <https://doi.org/10.1371/journal.pone.0199782>.
 13. Qi, F., Motz, M., Jung, K., Lassak, J., and Frishman, D. (2018). Evolutionary analysis of polyproline motifs in *Escherichia coli* reveals their regulatory role in translation. *PLoS Comput. Biol.* **14**, e1005987. <https://doi.org/10.1371/journal.pcbi.1005987>.
 14. Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., Kim, J.R., Saini, P., Buskirk, A.R., and Dever, T.E. (2013). eIF5A promotes translation of polyproline motifs. *Mol. Cell* **51**, 35–45. <https://doi.org/10.1016/j.molcel.2013.04.021>.
 15. Huter, P., Arenz, S., Bock, L.V., Graf, M., Frister, J.O., Heuer, A., Peil, L., Starosta, A.L., Wohlgemuth, I., Peske, F., et al. (2017). Structural Basis for Polyproline-Mediated Ribosome Stalling and Rescue by the Translation Elongation Factor EF-P. *Mol. Cell* **68**, 515–527.e6. <https://doi.org/10.1016/j.molcel.2017.10.014>.
 16. Blaha, G., Stanley, R.E., and Steitz, T.A. (2009). Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. *Science* **325**, 966–970. <https://doi.org/10.1126/science.1175800>.
 17. Park, M.H., Cooper, H.L., and Folk, J.E. (1982). The biosynthesis of protein-bound hypusine (N epsilon -(4-amino-2-hydroxybutyl)lysine). Lysine as the amino acid precursor and the intermediate role of deoxyhypusine (N epsilon -(4-aminobutyl)lysine). *J. Biol. Chem.* **257**, 7217–7222.
 18. Behshad, E., Ruzicka, F.J., Mansoorabadi, S.O., Chen, D., Reed, G.H., and Frey, P.A. (2006). Enantiomeric free radicals and enzymatic control of stereochemistry in a radical mechanism: the case of lysine 2,3-amino-
 19. Yanagisawa, T., Sumida, T., Ishii, R., Takemoto, C., and Yokoyama, S. (2010). A paralog of lysyl-tRNA synthetase aminoacylates a conserved lysine residue in translation elongation factor P. *Nat. Struct. Mol. Biol.* **17**, 1136–1143. <https://doi.org/10.1038/nsmb.1889>.
 20. Navarre, W.W., Zou, S.B., Roy, H., Xie, J.L., Savchenko, A., Singer, A., Edvokimova, E., Prost, L.R., Kumar, R., Ibba, M., and Fang, F.C. (2010). PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. *Mol. Cell* **39**, 209–221. <https://doi.org/10.1016/j.molcel.2010.06.021>.
 21. Peil, L., Starosta, A.L., Virumäe, K., Atkinson, G.C., Tenson, T., Remme, J., and Wilson, D.N. (2012). Lys34 of translation elongation factor EF-P is hydroxylated by YfcM. *Nat. Chem. Biol.* **8**, 695–697. <https://doi.org/10.1038/nchembio.1001>.
 22. Roy, H., Zou, S.B., Bullwinkle, T.J., Wolfe, B.S., Gilreath, M.S., Forsyth, C.J., Navarre, W.W., and Ibba, M. (2011). The tRNA synthetase paralog PoxA modifies elongation factor-P with (R)-beta-lysine. *Nat. Chem. Biol.* **7**, 667–669. <https://doi.org/10.1038/nchembio.632>.
 23. Lassak, J., Keilhauer, E.C., Fürst, M., Wuichet, K., Gödeke, J., Starosta, A.L., Chen, J.M., Søgaard-Andersen, L., Rohr, J., Wilson, D.N., et al. (2015). Arginine-rhamnosylation as new strategy to activate translation elongation factor P. *Nat. Chem. Biol.* **11**, 266–270. <https://doi.org/10.1038/nchembio.1751>.
 24. Rajkovic, A., Erickson, S., Witzky, A., Branson, O.E., Seo, J., Gafken, P.R., Frietas, M.A., Whitelegge, J.P., Faull, K.F., Navarre, W., et al. (2015). Cyclic Rhamnosylated Elongation Factor P Establishes Antibiotic Resistance in *Pseudomonas aeruginosa*. *mBio* **6**, e00823. <https://doi.org/10.1128/mBio.00823-15>.
 25. Rajkovic, A., Hummels, K.R., Witzky, A., Erickson, S., Gafken, P.R., Whitelegge, J.P., Faull, K.F., Kearns, D.B., and Ibba, M. (2016). Translation Control of Swarming Proficiency in *Bacillus subtilis* by 5-Amino-pentanoylated Elongation Factor P. *J. Biol. Chem.* **291**, 10976–10985. <https://doi.org/10.1074/jbc.M115.712091>.
 26. Hummels, K.R., Witzky, A., Rajkovic, A., Tollerson, R., 2nd, Jones, L.A., Ibba, M., and Kearns, D.B. (2017). Carbonyl reduction by YmfI in *Bacillus subtilis* prevents accumulation of an inhibitory EF-P modification state. *Mol. Microbiol.* **106**, 236–251. <https://doi.org/10.1111/mmi.13760>.
 27. Pinheiro, B., Scheidler, C.M., Kielkowski, P., Schmid, M., Forné, I., Ye, S., Reiling, N., Takano, E., Imhof, A., Sieber, S.A., et al. (2020). Structure and Function of an Elongation Factor P Subfamily in Actinobacteria. *Cell Rep.* **30**, 4332–4342.e5. <https://doi.org/10.1016/j.celrep.2020.03.009>.
 28. Tollerson, R., 2nd, Witzky, A., and Ibba, M. (2018). Elongation factor P is required to maintain proteome homeostasis at high growth rate. *Proc. Natl. Acad. Sci. USA* **115**, 11072–11077. <https://doi.org/10.1073/pnas.1812025115>.
 29. Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526. <https://doi.org/10.1074/mcp.M113.031591>.
 30. Volkwein, W., Krafczyk, R., Jagtap, P.K.A., Parr, M., Mankina, E., Macošek, J., Guo, Z., Fürst, M.J.L.J., Pfab, M., Frishman, D., et al. (2019). Switching the Post-translational Modification of Translation Elongation Factor EF-P. *Front. Microbiol.* **10**, 1148. <https://doi.org/10.3389/fmicb.2019.01148>.
 31. Hanawa-Suetsugu, K., Sekine, S.i., Sakai, H., Hori-Takemoto, C., Terada, T., Unzai, S., Tame, J.R.H., Kuramitsu, S., Shirouzu, M., and Yokoyama, S. (2004). Crystal structure of elongation factor P from *Thermus thermophilus* HB8. *Proc. Natl. Acad. Sci. USA* **101**, 9595–9600. <https://doi.org/10.1073/pnas.0308667101>.
 32. Choi, S., and Choe, J. (2011). Crystal structure of elongation factor P from *Pseudomonas aeruginosa* at 1.75 Å resolution. *Proteins* **79**, 1688–1693. <https://doi.org/10.1002/prot.22992>.

33. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
34. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., et al. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 50, D439–D444. <https://doi.org/10.1093/nar/gkab1061>.
35. UniProt Consortium (2023). UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* 51, D523–D531. <https://doi.org/10.1093/nar/gkac1052>.
36. Hayter, J.R., Robertson, D.H.L., Gaskell, S.J., and Beynon, R.J. (2003). Proteome analysis of intact proteins in complex mixtures. *Mol. Cell. Proteomics* 2, 85–95. <https://doi.org/10.1074/mcp.M200078-MCP200>.
37. Lee, S.W., Berger, S.J., Martinović, S., Pasa-Tolić, L., Anderson, G.A., Shen, Y., Zhao, R., and Smith, R.D. (2002). Direct mass spectrometric analysis of intact proteins of the yeast large ribosomal subunit using capillary LC/FTICR. *Proc. Natl. Acad. Sci. USA* 99, 5942–5947. <https://doi.org/10.1073/pnas.082119899>.
38. Pfab, M., Kielkowski, P., Krafczyk, R., Volkwein, W., Sieber, S.A., Lassak, J., and Jung, K. (2021). Synthetic post-translational modifications of elongation factor P using the ligase EpmA. *FEBS J.* 288, 663–677. <https://doi.org/10.1111/febs.15346>.
39. Reiling, N., Homolka, S., Walter, K., Brandenburg, J., Niwinski, L., Ernst, M., Herzmann, C., Lange, C., Diel, R., Ehlers, S., and Niemann, S. (2013). Clade-specific virulence patterns of *Mycobacterium tuberculosis* complex strains in human primary macrophages and aerogenically infected mice. *mBio* 4, e00250-13. <https://doi.org/10.1128/mBio.00250-13>.
40. Bradshaw, E., Saalbach, G., and McArthur, M. (2013). Proteomic survey of the *Streptomyces coelicolor* nucleoid. *J. Proteomics* 83, 37–46. <https://doi.org/10.1016/j.jprot.2013.02.033>.
41. Wards, B.J., and Collins, D.M. (1996). Electroporation at elevated temperatures substantially improves transformation efficiency of slow-growing mycobacteria. *FEMS Microbiol. Lett.* 145, 101–105. <https://doi.org/10.1111/j.1574-6968.1996.tb08563.x>.
42. Katz, A., Solden, L., Zou, S.B., Navarre, W.W., and Ibba, M. (2014). Molecular evolution of protein-RNA mimicry as a mechanism for translational control. *Nucleic Acids Res.* 42, 3261–3271. <https://doi.org/10.1093/nar/gkt1296>.
43. Mudryi, V., Peske, F., and Rodnina, M. (2023). Translation factor accelerating peptide bond formation on the ribosome: EF-P and eIF5A as entropic catalysts and a potential drug targets. *BBA Adv.* 3, 100074. <https://doi.org/10.1016/j.bbadv.2023.100074>.
44. Marman, H.E., Mey, A.R., and Payne, S.M. (2014). Elongation factor P and modifying enzyme PoxA are necessary for virulence of *Shigella flexneri*. *Infect. Immun.* 82, 3612–3621. <https://doi.org/10.1128/IAI.01532-13>.
45. Guo, Q., Cui, B., Wang, M., Li, X., Tan, H., Song, S., Zhou, J., Zhang, L.H., and Deng, Y. (2022). Elongation factor P modulates *Acinetobacter baumannii* physiology and virulence as a cyclic dimeric guanosine monophosphate effector. *Proc. Natl. Acad. Sci. USA* 119, e2209838119. <https://doi.org/10.1073/pnas.2209838119>.
46. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* 13, 731–740. <https://doi.org/10.1038/nmeth.3901>.
47. Tyanova, S., and Cox, J. (2018). Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. *Methods Mol. Biol.* 1711, 133–148. https://doi.org/10.1007/978-1-4939-7493-1_7.
48. Chambers, M.C., Maclean, B., Burke, R., Amodei, D., Ruderman, D.L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egerton, J., et al. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* 30, 918–920. <https://doi.org/10.1038/nbt.2377>.
49. Demichev, V., Messner, C.B., Vernardis, S.I., Lilley, K.S., and Ralser, M. (2020). DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* 17, 41–44. <https://doi.org/10.1038/s41592-019-0638-x>.
50. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* 27, 14–25. <https://doi.org/10.1002/pro.3235>.
51. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. <https://doi.org/10.1002/pro.3943>.
52. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nat. Methods* 19, 679–682. <https://doi.org/10.1038/s41592-022-01488-1>.
53. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. <https://doi.org/10.1038/nmeth.2089>.
54. Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., et al. (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* 50, D543–D552. <https://doi.org/10.1093/nar/gkab1038>.
55. Lassak, J., Henche, A.L., Binnenkade, L., and Thormann, K.M. (2010). ArcS, the cognate sensor kinase in an atypical Arc system of *Shewanella oneidensis* MR-1. *Appl. Environ. Microbiol.* 76, 3263–3274. <https://doi.org/10.1128/AEM.00512-10>.
56. Keseler, I.M., Gama-Castro, S., Mackie, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Kothari, A., Krummenacker, M., Midford, P.E., Muñoz-Rascado, L., et al. (2021). The EcoCyc Database in 2021. *Front. Microbiol.* 12, 711077. <https://doi.org/10.3389/fmicb.2021.711077>.
57. Miller, J.H. (1992). *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria* (Cold Spring Harbor Laboratory Press).
58. Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. <https://doi.org/10.1038/227680a0>.
59. Hughes, C.S., Moggridge, S., Müller, T., Sorensen, P.H., Morin, G.B., and Krijgsveld, J. (2019). Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat. Protoc.* 14, 68–85. <https://doi.org/10.1038/s41596-018-0082-x>.
60. Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* 98, 5116–5121. <https://doi.org/10.1073/pnas.091062498>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti - <i>R. v.</i> EF-P antibodies (α -625)	This study	N/A
6x-His Tag Monoclonal Antibody	Invitrogen	RRID: AB_2536841
Goat anti-Mouse IgG	Abcam	Cat# ab216776; RRID: AB_2933974
Goat anti-Rabbit IgG	Abcam	Cat# ab216773; RRID:AB_2925189
Rabbit polyclonal anti - <i>E. coli</i> EF-P antibody	Pfab et al. ³⁸	N/A
Bacterial and virus strains		
See Data S3 (Sheet 'strain list')	This study	N/A
Chemicals, peptides, and recombinant proteins		
Chymotrypsin	Thermo Fisher Scientific	Cat# 90056
<i>ortho</i> -Nitrophenyl- β -galactoside	Carl Roth GmbH + Co. KG	2492-87-7
Critical commercial assays		
Q5® Site-Directed Mutagenesis Kit	New England Biolabs	Cat# E0554S
iST Kit	Preomics	P.O.00001
Deposited data		
Mass spectrometry proteomics data	ProteomeXchange	ProteomeXchange: PXD044929
Oligonucleotides		
See Data S3 (sheet 'primer list')	This study	N/A
Recombinant DNA		
See Data S3 (sheet 'plasmid list')	This study	N/A
See Data S3 (sheet 'genomic DNA list')	DSMZ	N/A
Software and algorithms		
Freestyle (Xtract Deconvolution algorithm)	Thermo Scientific	https://www.thermofisher.com/order/catalog/product/OPTON-30965?SID=srch-srp-OPTON-30965
Perseus 2.0.9.0	Tyanova et al. ⁴⁶ Tyanova & Cox ⁴⁷	https://maxquant.net/perseus/
The MaxQuant 2.1.0.0	Cox et al. ²⁹	https://www.maxquant.org/
ProteoWizard	Chambers et al. ⁴⁸	http://www.proteowizard.org/download.html
Standalone DIA-NN software 1.8.1	Demichev et al. ⁴⁹	https://github.com/vdemichev/DiaNN
UCSF ChimeraX	Goddard et al., ⁵⁰ Pettersen et al. ⁵¹	https://www.rbvi.ucsf.edu/chimerax
AlphaFold2 ColabFold 1.5.2	Jumper et al., ³³ Mirdita et al. ⁵²	N/A
ImageJ 1.54d	Schneider et al. ⁵³	https://imagej.net/software/imagej/
CLC Main Workbench 8.1.2	QIAGEN, Aarhus, Denmark	https://digitalinsights.qiagen.com/
Other		
Amicon® Ultra-15 Centrifugal Filter Units	Millipore	UFC9003
Carboxylate-coated magnetic beads (hydrophilic)	Cytiva	Cat# 45152105050250
Carboxylate-coated magnetic beads (hydrophobic)	Cytiva	Cat# 65152105050250
Ni-NTA Agarose	Qiagen	Cat# 30230
ZipTip with C4 resin	Millipore	Cat# ZTC04S096
PicoTip™ Emitter, Silica Tip™	New Objectives	FS360-75-8-N-20-C15
PEPMAP100 C18 5UM 0.3 × 5MM	Thermo Fisher Scientific	Cat# 160454
ReproSil-Pur 120 C18-AQ, 1.9 μ m	Dr. Maisch GmbH	Cat# r119.aq.0001
SERVAPOR® dialysis tubing	SERVA	Cat# 44146.04

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Kirsten Jung (jung@lmu.de).

Materials availability

Materials generated in this study are available from the [lead contact](#) upon request with a completed Material Transfer Agreement.

Data and code availability

- Mass spectrometry and proteomics data are deposited to the ProteomeXchange Consortium via the PRIDE⁵⁴ partner repository (ProteomeXchange: PXD044929).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this work paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All bacterial strains used in this study are listed in [Data S3](#). *E. coli* was cultivated in lysogeny broth (LB) supplemented with antibiotics under agitation (750 rpm) at 37°C. Bacterial growth experiments were conducted in 50 mL glass flasks filled with 15 mL LB. For the spot assay, overnight cultures were resuspended in LB to an optical density (600 nm) (OD₆₀₀) of 0.01, spotted in dilutions (10⁻³ to 10⁻⁵) on LB plates and grown for 18 h at 37°C. For CadC production, cells were grown in buffered LB at pH 5.8.⁷ To investigate β-galactosidase activity on solid media, LB agar plates were supplemented with S-Gal® (300 mg/mL) and ferric ammonium citrate (500 mg/L). For EF-P overproduction, growth media were supplemented with L-arabinose [0.2% (w/v)] and chloramphenicol. Antibiotic concentrations used in this study: 34 μg/mL chloramphenicol, 50 μg/mL kanamycin sulfate.

METHOD DETAILS

Plasmid and bacterial strain construction

All primers and plasmids constructed in this study are shown in [Data S3](#). All PCR reactions were conducted using the Q5 polymerase (New England BioLabs) according to manufacturer's instructions. Standard DNA restrictions were performed in rCutSmart buffer and the fragments were ligated into the corresponding vectors using the T4 ligase (New England BioLabs). *efp* genes from various species including a sequence coding for a C-terminal 6xHis tag were cloned in pBAD33. Plasmids were isolated using Hi Yield Plasmid Mini Kit (Sued Laborbedarf), whereas PCR fragments from the agarose gel were purified using the High-Yield PCR Cleanup and Gel Extraction Kit (New England BioLabs). Site-specific mutagenesis of *efp* was carried out with the Q5 Site-Directed Mutagenesis Kit (New England BioLabs) according to manufacturer's instructions.

E. coli mutants expressing *efp* variants were constructed using double homologous recombination with neomycin acetyltransferase and SacB as the selection or counterselection markers, respectively.⁵⁵ Overlapping regions of *E. coli efp* with the promoter of its modification enzyme *epmB* and the lipoprotein entericidin A (*ecnA*) were considered by keeping the 5' and 3' gene regions of *E. coli efp* present in the chromosome⁵⁶ ([Figures 2A, S3A, and S3E](#)). All nucleotide and protein sequences were analyzed using CLC Main Workbench 8.1.2 (Qiagen).

β-Galactosidase activity assays

Reporter strains MG1655 Δ*lacZ* P_{cadBA}:*lacZ* Δ*efp* and MG1655 Δ*lacZ* P_{cadBA}:*lacZ* Δ*efp* Δ*epmA* transformed with plasmids expressing *efp* and its variants were inoculated in 1.8 mL buffered LB pH 5.8 (91.5 mM KH₂PO₄; 8.5 mM K₂HPO₄; 0.2% (w/v) arabinose; chloramphenicol) and microaerobically grown under agitation (Thermomixer Comfort; 700 rpm) at 37°C overnight. The overnight culture was split into three tubes to measure optical density, β-galactosidase activity and detect EF-P in a Western Blot. For the β-galactosidase activity measurements, cells were harvested from 500 μL liquid culture by centrifugation and resuspended in 1 mL cooled (4°C) assay buffer (60 mM Na₂HPO₄; 40 mM NaH₂PO₄; 10 mM KCl; 1 mM MgSO₄; 50 mM β-mercaptoethanol; pH 7.0). Cells were permeabilized by adding 100 μL chloroform and 50 μL 0.1% sodium dodecyl sulfate (SDS). Prior to the measurements, all samples were incubated for 5 min at 37°C. The reaction was started by adding 0.2 mL of the substrate *ortho*-Nitrophenyl-β-galactoside (*o*-NPG, Carl Roth GmbH + Co. KG; 4 mg/mL in assay buffer). Total incubation time with the substrate was recorded when yellow color was developed until the reaction was stopped by adding 0.5 mL of 1 M Na₂CO₃. Samples were centrifuged and 1 mL of the supernatant was used for absorption measurements at 420 nm.^{7,27} The measured β-galactosidase activity is given in Miller Units (MU), calculated according to Miller et al., 1992.⁵⁷

Protein purification and structural analysis

For the recombinant *R. vannielii* EF-P production, *efp* was cloned into the pET expression plasmid and overproduced using the pET expression system in *E. coli* BL21 (DE3) (Invitrogen). Cells were harvested after cultivation in LB supplemented with IPTG (1 mM) at

18°C overnight, and the resulting pellet was frozen at -80°C . Cells with chromosomally incorporated His-tagged *efp* (*R. vannielii* EF-P and *W. virosa* EF-P) were grown in LB until the mid-exponential growth phase, harvested and kept at -80°C until further processing.

For lysis, cells were resuspended in 0.1 M, pH 7.6 sodium phosphate buffer (supplemented with 300 mM NaCl and DNase) and lysed using the high-pressure cell disrupter (Constant Systems), by running the sample twice under 1.9 kbar with subsequent cell fractionation using ultracentrifugation. The His-tagged proteins were purified using Ni-NTA Agarose beads (Qiagen) with 0.1 M, pH 7.6 sodium phosphate buffer supplemented with 300 mM NaCl and 30 mM–200 mM imidazole. Samples were dialyzed to remove imidazole at 4°C for 24 h using SERVAPORE dialysis tubing (MWCO 12000–14000, SERVA) according to manufacturer's instructions. Proteins were concentrated using the Amicon Ultra-15 Centrifugal Filter Units, 3kDa (Millipore).

The *R. v.* EF-P structure was predicted using AlphaFold2 ColabFold (v1.5.2).^{33,52} All structural models were generated using UCSF ChimeraX.^{50,51}

SDS-PAGE and Western Blot analysis

Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)⁵⁸ using 12.5% Bis-tris acrylamide gels. Gels were stained with 2,2,2-trichloroethanol (TCE) to detect protein bands corresponding to a 72 kDa protein which served as loading controls. All proteins were transferred to a nitrocellulose membrane using the wet-blot apparatus (Mini Trans-Blot Cell, Bio-Rad). *E. c.* EF-P was detected by incubating the membrane in Tris-buffered saline (TBS) with primary polyclonal antibodies against *E. coli* EF-P (rabbit, Eurogentec)³⁸ with final concentration of 1:5,000, whereas His-tagged EF-P was detected with primary monoclonal antibodies against the His-Tag (1:10,000; RRID: AB_2536841, Invitrogen). Membranes were incubated with the secondary antibodies conjugated with a fluorophore (1:20,000 in TBS) (anti mouse: Cat# ab216776; anti rabbit: Cat# ab216773) and imaged using the Odyssey CLx (LI-COR Biosciences). For detection of *R. vannielii* EF-P, recombinantly produced EF-P was purified and sent to Eurogentec for polyclonal antibody generation (Speedy 28-day program in rabbits, Eurogentec). The blood serum containing antibodies against the *R. vannielii* EF-P (α 652) was diluted to final concentration of 1:100. Relative protein band intensities in Western blot analysis were calculated using ImageJ1.54d.⁵³

Mass spectrometry for identification of modification status

For top-down EF-P measurements the purified proteins were desalted on the ZipTip with C4 resin (Millipore, ZTC04S096) and eluted with 50% (v/v) acetonitrile 0.1% (v/v) formic acid (FA) buffer resulting in $\sim 10\ \mu\text{M}$ final protein concentration in 200–400 μL total volume. MS measurements were performed on an Orbitrap Eclipse Tribrid Mass Spectrometer (Thermo Fisher Scientific) via direct injection, a HESI-Spray source (Thermo Fisher Scientific) and FAIMS interface (Thermo Fisher Scientific) in a positive, peptide mode. Typically, the FAIMS compensation voltage (CV) was optimized by a continuous scan. The most intense signal was usually obtained at $-5\ \text{CV}$. The MS spectra were acquired with at least 120,000 FWHM, AGC target 100 and 2–5 microscans covering the 800–1100 m/z range. The spectra were deconvoluted in Freestyle (Thermo Scientific) using the Xtract Deconvolution algorithm.

For bottom-up proteomics, samples were prepared in 96-well plate using the optimized SP3 protocol.⁵⁹ The purified EF-P sample (1 μL , 10 μM) was diluted to total volume of 50 μL with 1% NP40, 0.2% (w/v) SDS in 25 mM HEPES, pH 7.5. The protein was loaded onto a mixture of hydrophilic and hydrophobic carboxylate-coated magnetic beads (10 μL each) pre-washed three times with 100 μL of MS-grade H_2O . The magnetic beads with protein sample were mixed at 850 rpm, 1 min at room temperature (RT). To initiate the binding, 60 μL of absolute EtOH was added, and the mixture was incubated at RT for 5 min at 850 rpm. Subsequently, the beads were washed three times with 80% (v/v) EtOH, with incubation at RT for 1 min and 850 rpm between each wash. After the last wash the beads were resuspended in 50 μL of 100 mM ammonium acetate buffer (ABC). The proteins were reduced and alkylated by addition of 5 μL of 100 mM tris(2-carboxyethyl) phosphine (TCEP) and 5 μL of 400 mM chloroacetamide (CAA) and incubation at 95°C for 5 min, 850 rpm. Samples were cooled to RT. The on-beads digestion was performed with chymotrypsin. Chymotrypsin digestion: ABC buffer was supplemented with 10 mM CaCl_2 , 1 μg of chymotrypsin (Thermo Scientific, 90056) and incubated at 25°C overnight. The resulting peptide mixture was eluted from the magnetic beads into a new 1.5 mL tube. The magnetic beads were washed with 50 and 30 μL of 1% (v/v) formic acid and incubated at 40°C , 850 rpm for 5 min. The fractions were added to the first elution fraction. The combined fractions were further purified from remaining magnetic beads. Alternatively, for *R. v.* EF-P overproduced in *E. coli*, the EF-P was directly digested in ABC buffer (50 μL) supplemented with 10 mM CaCl_2 and 1 μg of chymotrypsin (Thermo Scientific, 90056), without previous clean-up on magnetic beads.

MS measurements were performed on an Orbitrap Eclipse Tribrid Mass Spectrometer (Thermo Fisher Scientific) coupled to an UltiMate 3000 Nano-HPLC (Thermo Fisher Scientific) via a nanospray Flex ion source (Thermo Fisher Scientific) equipped with column oven (Sonation) and FAIMS interface (Thermo Fisher Scientific). Peptides were loaded on an Acclaim PepMap 100 μm -precolumn cartridge (5 μm , 100 \AA , 300 μm ID x 5 mm, Thermo Fisher Scientific) and separated at 40°C on a PicoTip emitter (noncoated, 15 cm, 75 μm ID, 8 μm tip, New Objective) that was *in-house* packed with Reprosil-Pur 120 C18-AQ material (1.9 μm , 150 \AA , Dr. A. Maisch GmbH). Buffer composition. Buffer A consists of MS-grade H_2O supplemented with 0.1% FA. Buffer B consists of acetonitrile supplemented with 0.1% FA. The 41- or 126-min LC gradient from 4 to 35.2% buffer B was used. The flow rate was 0.3 $\mu\text{L}/\text{min}$.

Data-independent acquisition

The DIA duty cycle consisted of one MS1 scan followed by 30 MS2 scans with an isolation window of the 4 m/z range, overlapping with an adjacent window at the 2 m/z range. MS1 scan was conducted with Orbitrap at 60000 resolution power and a scan range of 200–1800 m/z with an adjusted RF lens at 30%. MS2 scans were conducted with Orbitrap at 30000 resolution power, RF lens was set

to 30%. The precursor mass window was restricted to a 500–740 m/z range. HCD fragmentation was enabled as an activation method with a fixed collision energy of 35%. FAIMS was performed with one CV at –45V for both MS1 and MS2 scans during the duty cycle.

Data-dependent acquisition

(*R. v. EF-P* overproduced in *E. coli*) For measurements of DDA, the Orbitrap Eclipse Tribrid Mass Spectrometer was operated with the following settings: Polarity: positive; MS1 resolution: 240k; MS1 AGC target: standard; MS1 maximum injection time: 50 ms; MS1 scan range: m/z 375–1500; MS2 ion trap scan rate: rapid; MS2 AGC target: standard; MS2 maximum injection time: 35 ms; MS2 cycle time: 1.7 s; MS2 isolation window: m/z 1.2; HCD stepped normalised collision energy: 30%; intensity threshold: 1.0e4 counts; included charge states: 2–6; dynamic exclusion: 60 s. FAIMS was performed with two alternating CVs, including –50 V and –70 V.

Mass spectrometry for identification of modification status

Computational evaluation of DIA raw files

Raw files were converted in the first step with “MSConvertGUI” as a part of the “ProteoWizard” software⁴⁸ package (<http://www.proteowizard.org/download.html>) to an output mzML format applying the “peakPicking” filter with “vendor msLevel = 1”, and the “Demultiplex” filter with parameters “Overlap Only” and “mass error” set to 10 ppm.

Standalone DIA-NN software under version 1.8.1 was used for protein identification and quantification⁴⁹

First, a spectral library was predicted *in silico* by the software’s deep learning-based spectra, RTs and IMs prediction using Uniprot *E. coli* decoyed FASTA (canonical and isoforms with added *R. vannielii* EF-P sequence). DIA-NN search settings: FASTA digest for library-free search/library generation option was enabled, together with a match between runs (MBR) option and precursor FDR level set at 1%. Library generation was set to smart profiling, Quantification strategy - Robust LC. The mass accuracy and the scan window were set to 0 to allow the software to identify optimal conditions. The precursor m/z range was changed to 500–740 m/z to fit the measuring parameters. Carbamidomethylation was set as a fixed modification, oxidation of methionine and N-term acetylation were set as variable modifications. On the contrary, the small-scale samples of the 96-well plate were calculated without carbamidomethylation as a fixed modification.

Computational evaluation of DDA raw files

MS Raw files were analyzed using MaxQuant software. Searches were performed against the Uniprot database for *E. coli* (*R. vannielii* EF-P sequence). At least two unique peptides were required for protein identification. False discovery rate determination was carried out using a decoy database and thresholds were set to 1% FDR both at peptide-spectrum match and at protein levels.

Mass spectrometry for proteome analysis

Cells were cultivated in LB under constant shaking at 37°C until reaching the exponential growth phase. Cells were harvested ($OD_{600} = 0.5$) and were processed with the iST kit (Preomics) as recommended by the manufacturer. Samples were evaporated to dryness, re-suspended in LC-LOAD buffer to 0.2 μg/μL and injected in an Ultimate 3000 RSLCnano system (Thermo) separated in a 25-cm Aurora column (Ionopticks) with a 100-min gradient from 4 to 40% acetonitrile in 0.1% formic acid. The effluent from the HPLC was directly electrosprayed into an Orbitrap Exploris 480 (Thermo) operated in data dependent mode to automatically switch between full scan MS and MS/MS acquisition. Survey full scan MS spectra (from m/z 350–1200) were acquired with a resolution of $R = 60,000$ at m/z 400 (AGC target of 3×10^6). The 20 most intense peptide ions with charge states between 2 and 6 were sequentially isolated to a target value of 1×10^5 and fragmented at 30% normalized collision energy. Typical mass spectrometric conditions were: spray voltage, 1.5 kV; no sheath and auxiliary gas flow; heated capillary temperature, 275°C; intensity selection threshold, 3×10^5 .

The MaxQuant 2.1.0.0 software was used for protein identification and quantification by label-free quantification (LFQ)²⁹ with the following parameters: Database Uniprot_UP00000625_Ecoli_20220309.fasta including the *Rhodocrobium vannielii* EFP sequence; MS tol, 10 ppm; MS/MS tol, 20 ppm Da; Peptide FDR, 0.1; Protein FDR, 0.01 min; Peptide Length, 7; Variable modifications, Oxidation (M); Fixed modifications, Carbamidomethyl (C); Peptides for protein quantitation, razor and unique; Min. peptides, 1; Min. ratio count, 2. For display and analysis, the Perseus software^{46,47} was used. A list as a reference for proteins with polyproline-motifs in *E. coli* was taken from Qi *et al.*¹³ Data have been uploaded to the PRIDE repository⁵⁴ (ProteomeXchange: PXD044929).

QUANTIFICATION AND STATISTICAL ANALYSIS

All measurements, except intact protein mass measurements and chymotrypsin digestions, are from at least three biological replicates. Statistical analysis was done using Microsoft Office Excel 2019 (student’s unpaired two-sided t test). Error bars in all growth curves and bar graphs represent the standard deviation (SD). Values were considered as significantly different when the calculated p value was below 0.05.

Perseus (2.0.9.0) was used to log₂ transform LFQ intensities, replace missing values from normal distribution and construct the volcano plots. To determine which proteins were differentially expressed between experimental conditions, we applied a t test with a permutation-based FDR calculation with $n = 3$, $FDR = 0.05$ and $s_0 = 0.1$ to the log₂ LFQ protein values, where s_0 controls the relative importance of t test p value and difference between means. At $s_0 = 0$ only the p value matters, while at nonzero s_0 also the difference of means plays a role.⁶⁰

Cell Reports, Volume 43

Supplemental information

**Decrypting the functional design
of unmodified translation elongation factor P**

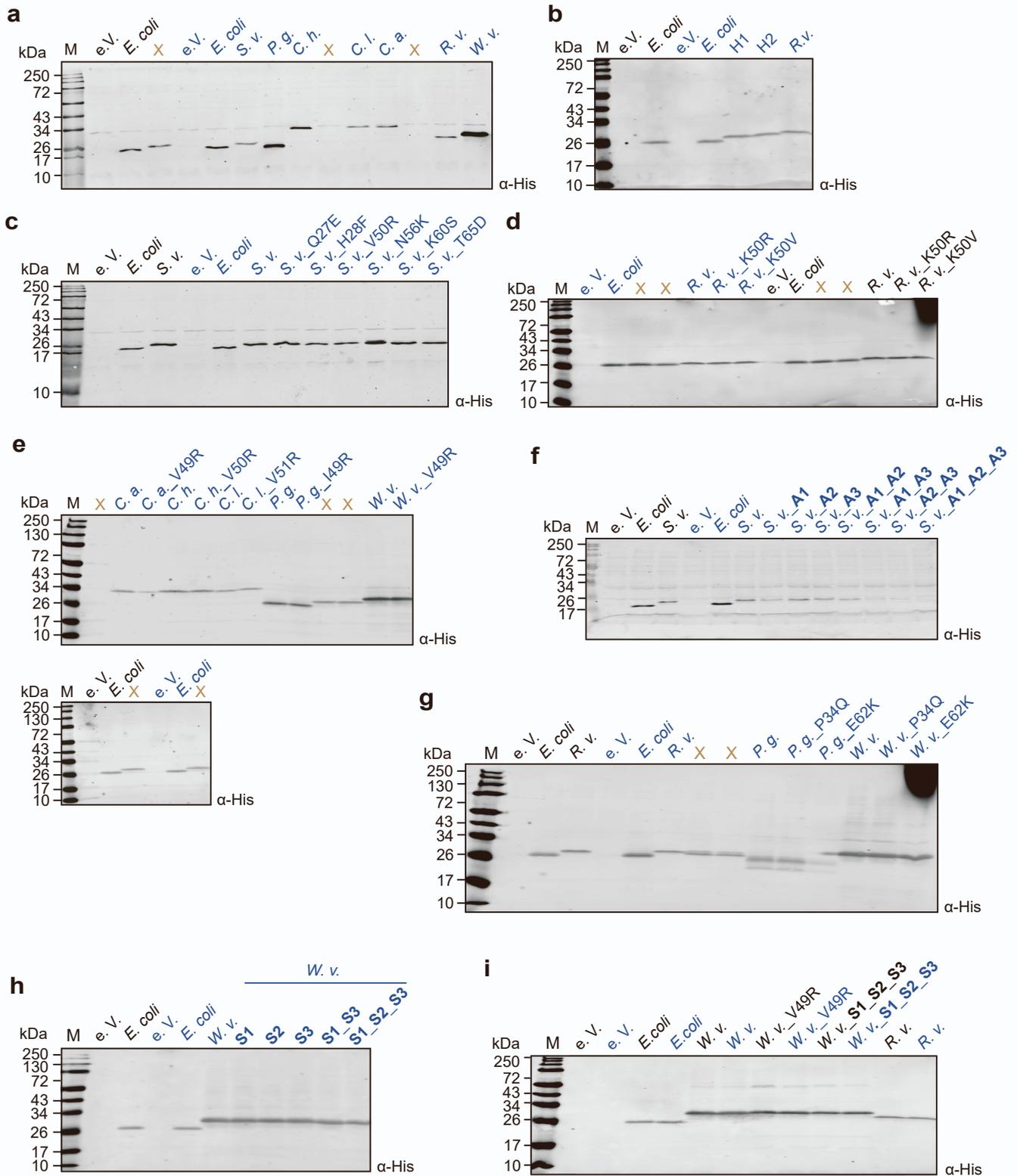
Urte Tomasiunaite, Pavel Kielkowski, Ralph Krafczyk, Ignasi Forné, Axel Imhof, and Kirsten Jung

Supplementary Table S1. Protein sequence identity of *E. coli* EF-P and EF-Ps of the PGKGP-subfamily. Related to Figure 1.

Multiple sequence alignment and percent identity matrix were calculated using the Multiple Sequence Alignment Tool from Clustal Omega (Clustal 2.1) and is given in percentage (%).

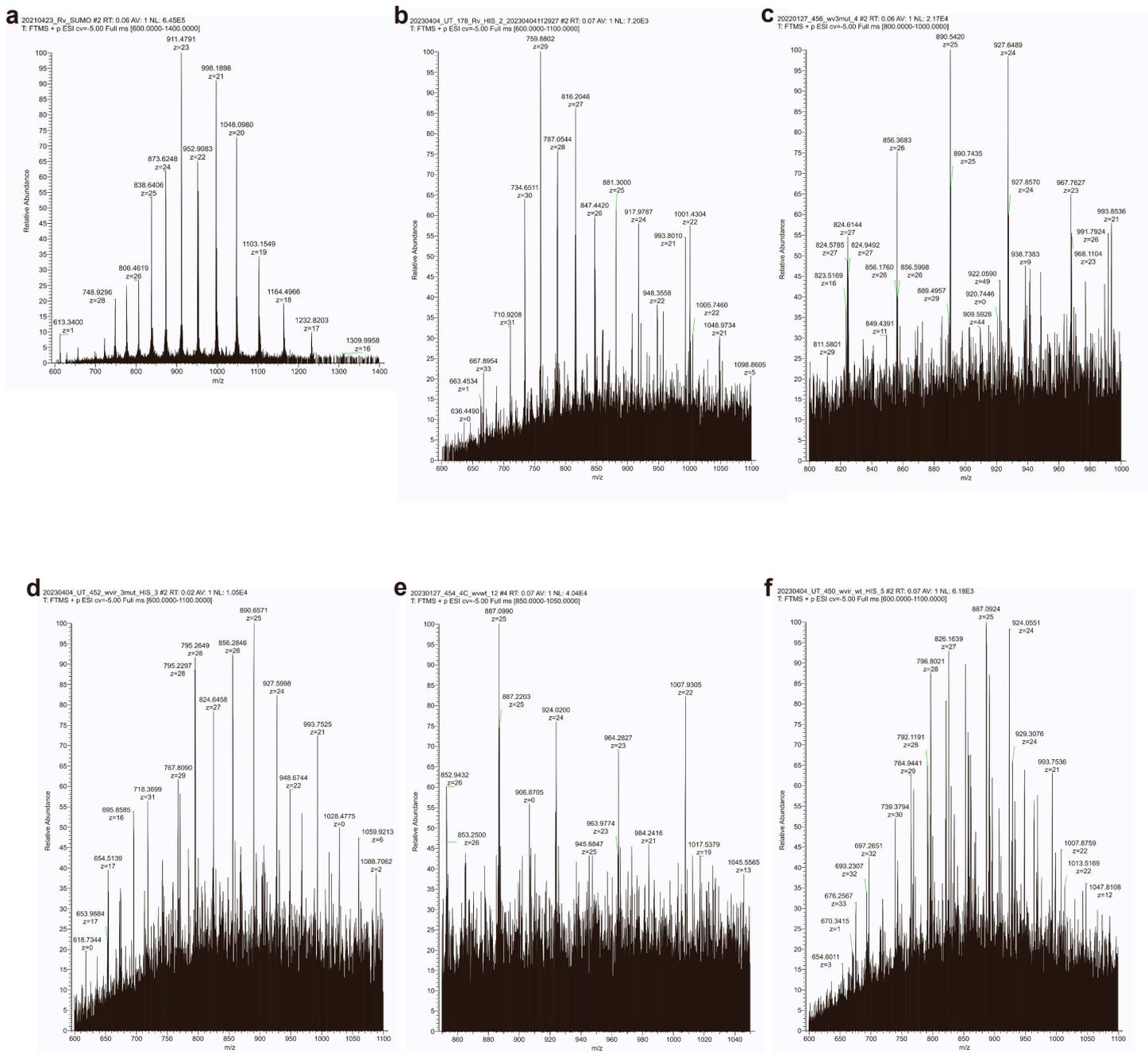
	<i>E. c.</i>	<i>C. h.</i>	<i>C. l.</i>	<i>C. a.</i>	<i>C. w.</i>	<i>P. g.</i>	<i>R. v.</i>	<i>S. v.</i>	<i>W. v.</i>
<i>E. c.</i>		44.92	44.15	39.13	40.54	41.08	32.09	41.94	37.84
<i>C. h.</i>	44.92		88.83	40.86	44.86	43.01	40.64	43.32	40.86
<i>C. l.</i>	44.15	88.83		40.86	45.95	40.86	37.97	42.25	40.32
<i>C. a.</i>	39.13	40.86	40.86		44.86	53.19	32.97	43.01	66.49
<i>C. w.</i>	40.54	44.86	45.95	44.86		45.65	35.14	46.49	50.54
<i>P. g.</i>	41.08	43.01	40.86	53.19	45.65		35.68	49.46	57.98
<i>R. v.</i>	32.09	40.64	37.97	32.97	35.14	35.68		36.56	31.89
<i>S. v.</i>	41.94	43.32	42.25	43.01	46.49	49.46	36.56		41.94
<i>W. v.</i>	37.84	40.86	40.32	66.49	50.54	57.98	31.89	41.94	

<p><i>E. c.</i> - <i>Escherichia coli</i> <i>C. h.</i> - <i>Campylobacter hominis</i> <i>C. l.</i> - <i>Campylobacter lari</i> <i>C. a.</i> - <i>Cellulophaga algicola</i> <i>C. w.</i> - <i>Conexibacter woesei</i> <i>P. g.</i> - <i>Porphyromonas gingivalis</i> <i>R. v.</i> - <i>Rhodomicrobium vannielii</i> <i>S. v.</i> - <i>Streptomyces venezuelae</i> <i>W. v.</i> - <i>Weeksella virosa</i></p>



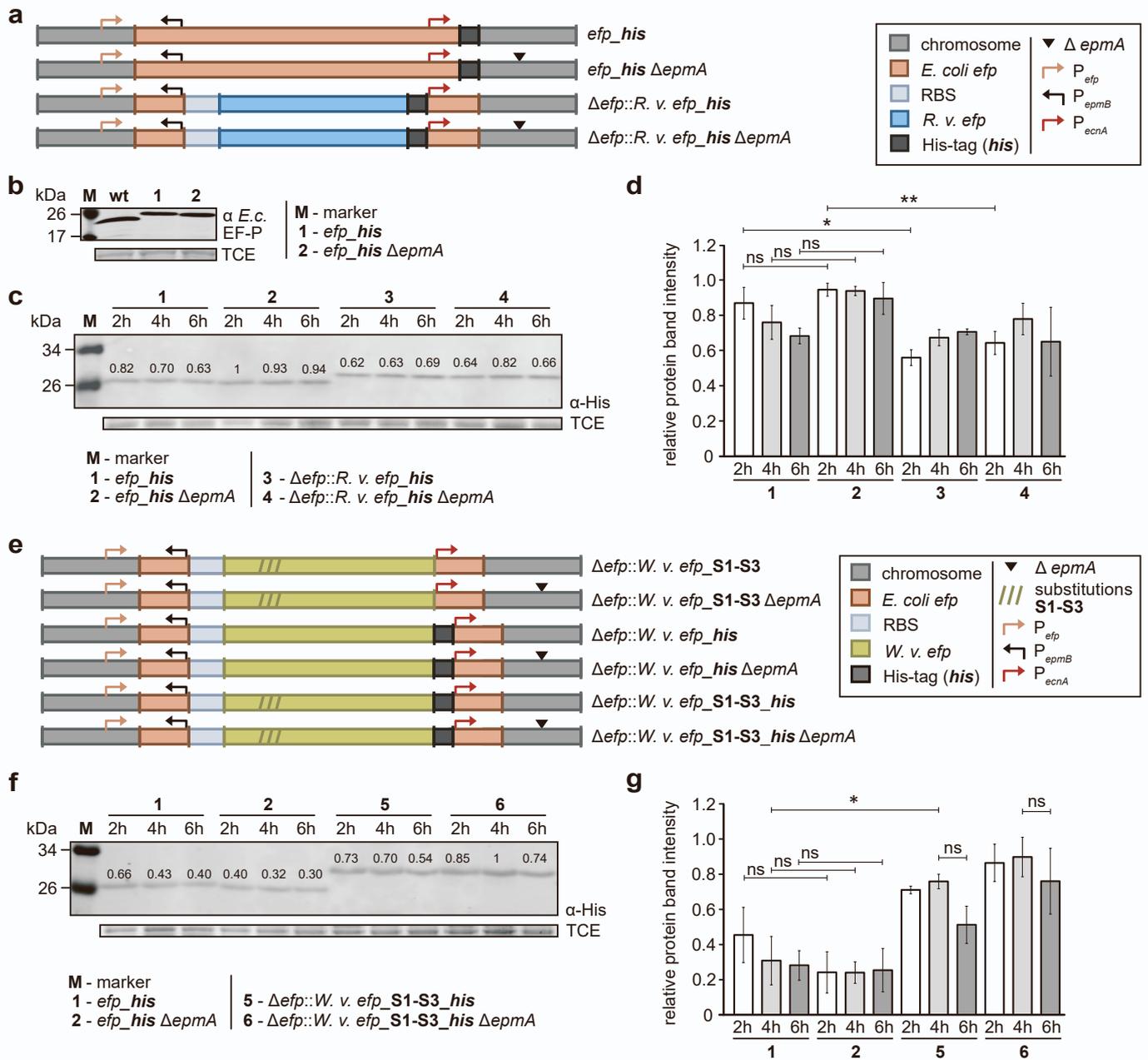
Supplementary Figure S1. Western blot analysis of His-tagged EF-P variants in *E. coli*. Related to Figures 1, 3, 4 and 5.

Uncropped Western blot membranes related to Fig. 1c (a), Fig. 3b (b), Fig. 4c (c), Fig. 4d (d), Fig. 4e (e), Fig. 5a (f), Fig. 5b (g), Fig. 5c (h) and Fig. 5d (i). EF-P production was confirmed by Western Blot analysis using antibodies against the His-tag. Lanes labelled with "X" correspond to samples not shown in the cropped images.



Supplementary Figure S2. Mass spectrometry analysis of EF-Ps produced in *E. coli*. Related to Figures 1 and 5.

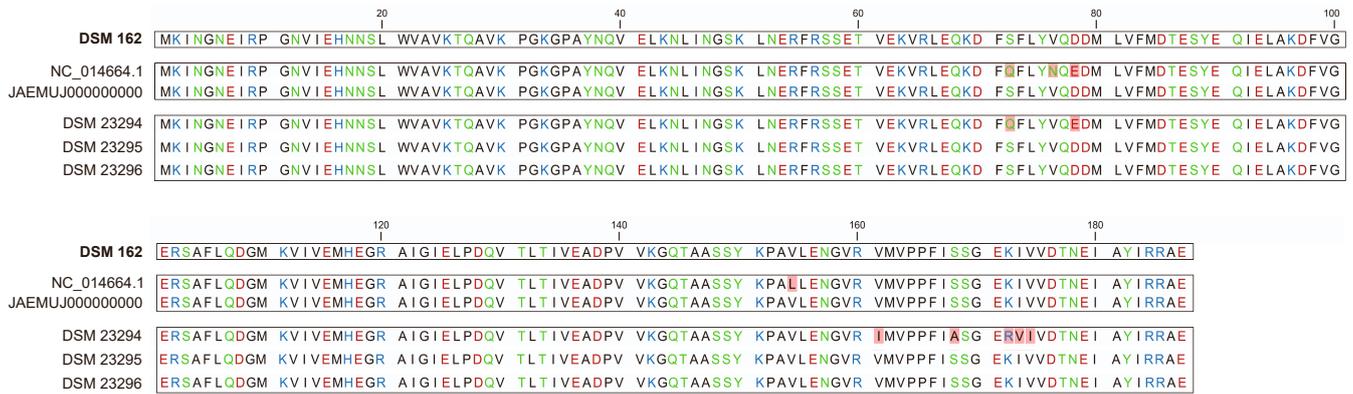
a Intact mass spectra of recombinantly overproduced *R. vannielii* EF-P, purified from *E. coli* wt. **b** Intact mass spectra of His-tagged *R. vannielii* EF-P (chromosomally encoded), purified from *E. coli* wt. **c** Intact mass spectra of His-tagged *W. virosa* EF-P variant with substitutions S1-S3 (chromosomally encoded), purified from *E. coli* Δ *epmA*. **d** Intact mass spectra of His-tagged *W. virosa* EF-P variant with substitutions S1-S3 (chromosomally encoded), purified from *E. coli* wt. **e** Intact mass spectra of His-tagged *W. virosa* EF-P wt (chromosomally encoded), purified from *E. coli* Δ *epmA*. **f** Intact mass spectra of His-tagged *W. virosa* EF-P wt (chromosomally encoded), purified from *E. coli* wt. S – substitution (S1 – P34Q; S2 – V49R; S3 – E62K).



Supplementary Figure S3. Protein production of His-tagged EF-P variants in *E. coli*. Related to Figures 1, 2 and 5.

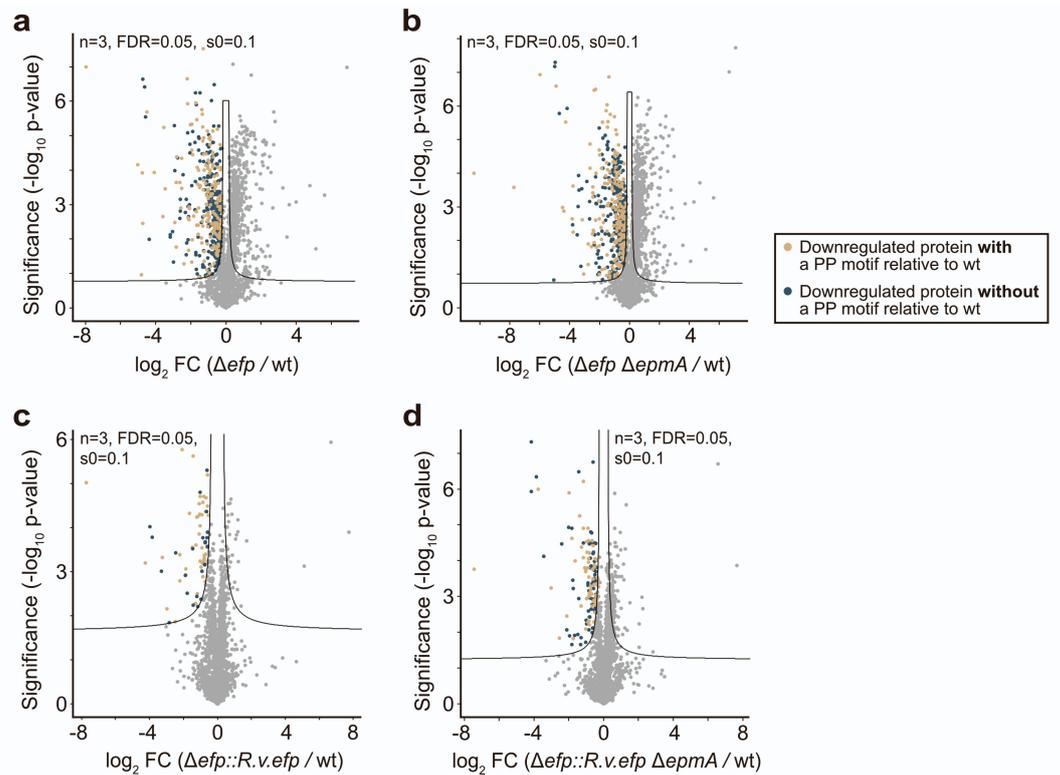
a Schematic overview of the constructed *E. coli* mutants expressing his-tagged *efp* or *R. v. efp*. **b** Western blot analysis of EF-P-His production in *E. coli* mutants after cells were grown in LB medium to mid-exponential phase using antibodies against *E. coli* EF-P. **c** Western blot analysis of *R. v.* EF-P production in *E. coli* mutants after cells were grown in LB medium to early-exponential (2h), mid-exponential (4h) and early-stationary (6h) phase using antibodies against the His-tag. **d** Quantification of His-tagged *E. coli* and *R.vanniellii* EF-P production in cells of three independent biological replicates. The numbers in bold correspond to the strains used in **c**. **e** Schematic overview of the constructed *E. coli* mutants expressing his-tagged *efp* variants of *W. virosa*. **f** Western blot analysis of *W. virosa* EF-P variant production in *E. coli* mutants after cells were grown in LB medium to early-exponential (2h), mid-exponential (4h) and early-stationary (6h) phase using antibodies against the His-tag. **g** Quantification of His-tagged *E. coli* and *W. virosa* EF-P variant production in cells of three independent biological replicates. The numbers in bold correspond to the strains used in **f**. Relative band intensities calculated using ImageJ are displayed above the corresponding protein bands. Protein bands corresponding to a 72 kDa protein after staining with 2,2,2-trichloroethanol (TCE) were used as loading controls. Gene insertions in **a** and **e** are depicted in coloured

boxes, and promoter locations as coloured arrows. S1-S3 - amino acid substitutions (S1 – P34Q; S2 – V49R; S3 – E62K). Statistics: error bars represent the standard deviation (SD) of at least three independent biological replicates; student's unpaired two-sided t-test (****p < 0,0001; ***p < 0,001; **p < 0,01; *p < 0,05; ns p > 0,05). 2h (Δefp_his vs. $\Delta efp_his \Delta epmA$, ns p = 0.3557; Δefp_his vs. $\Delta efp::R. v. efp_his$, *p = 0.0232; $\Delta efp_his \Delta epmA$ vs. $\Delta efp::R. v. efp_his \Delta epmA$, **p = 0.0092), 4h (Δefp_his vs. $\Delta efp_his \Delta epmA$, ns p = 0.1104), 6h (Δefp_his vs. $\Delta efp_his \Delta epmA$, ns p = 0.0608) (d). 2h (Δefp_his vs. $\Delta efp_his \Delta epmA$, ns p = 0.2041), 4h (Δefp_his vs. $\Delta efp_his \Delta epmA$, ns p = 0.5752; Δefp_his vs. $\Delta efp::W. virosa_efp_S1-S3_his$, *p = 0.0350, 6h (Δefp_his vs. $\Delta efp_his \Delta epmA$, ns p = 0.7399), 4h vs. 6h ($\Delta efp::W. virosa_efp_S1-S3_his$, ns p = 0.0657; $\Delta efp::W. virosa_efp_S1-S3_his \Delta epmA$, ns p = 0.4331) (g).



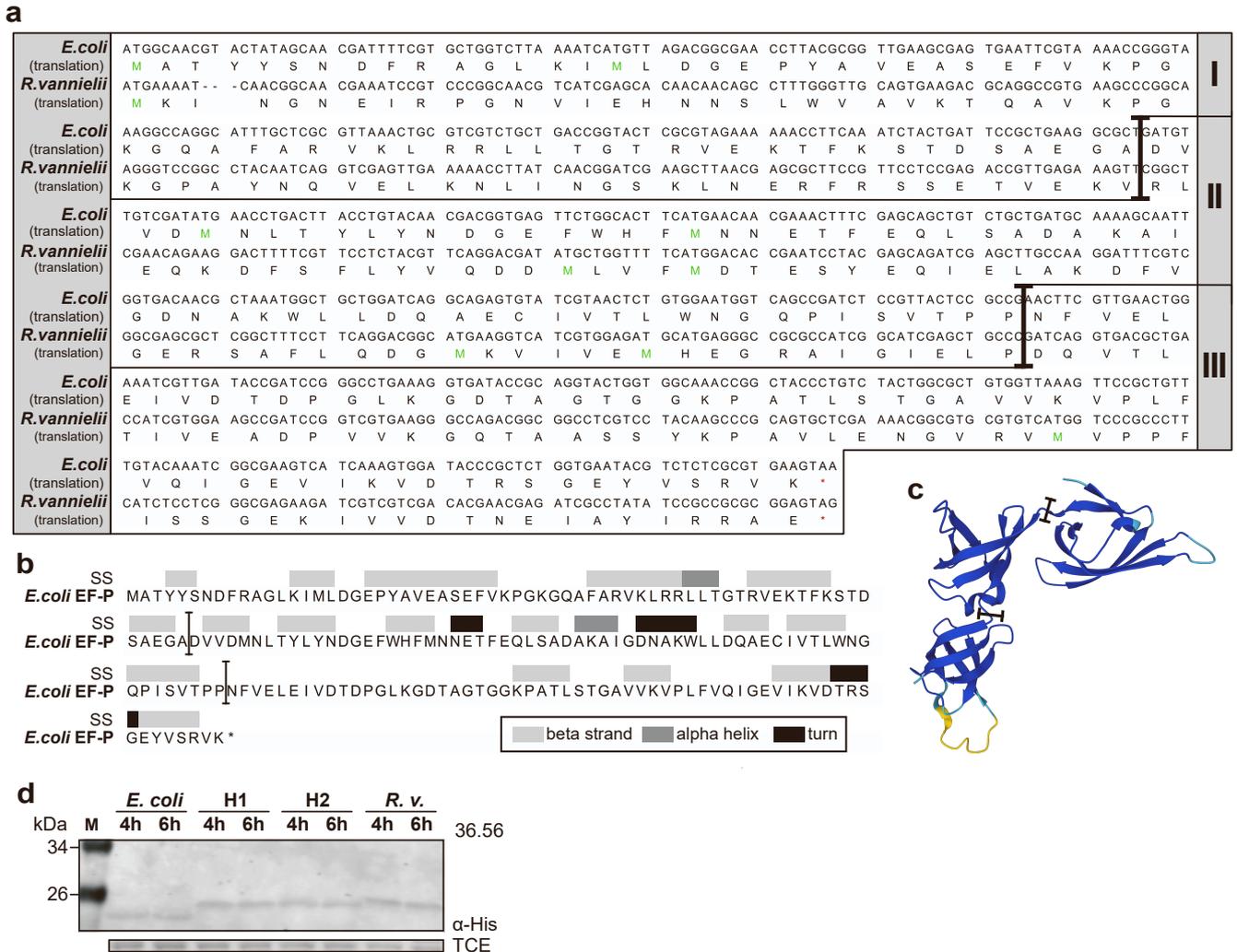
Supplementary Figure S4. Comparison of the sequence of EF-Ps from different *R. vannielii* strains and the publicly available sequences. Related to Figure 1.

Different colors represent the polarity of the amino acids (black- hydrophobic, light green-hydrophilic, red-acidic, blue-basic). Differences in amino acids between strains are highlighted in pink. Discrepancies between the EF-P sequences of the type strain DSM 162 used in this study and the publicly available sequence of *R. vannielii* ATCC171000 (GenBank accession number NC_014664.1) were ruled out by comparison with the revised EF-P sequence of *R. vannielii* ATCC171000 (GenBank accession number JAEMUJ000000000.1) [S1], and the sequencing results of EF-P from evolutionarily related strains (DSM 23294, DSM 23295, and DSM 23296).



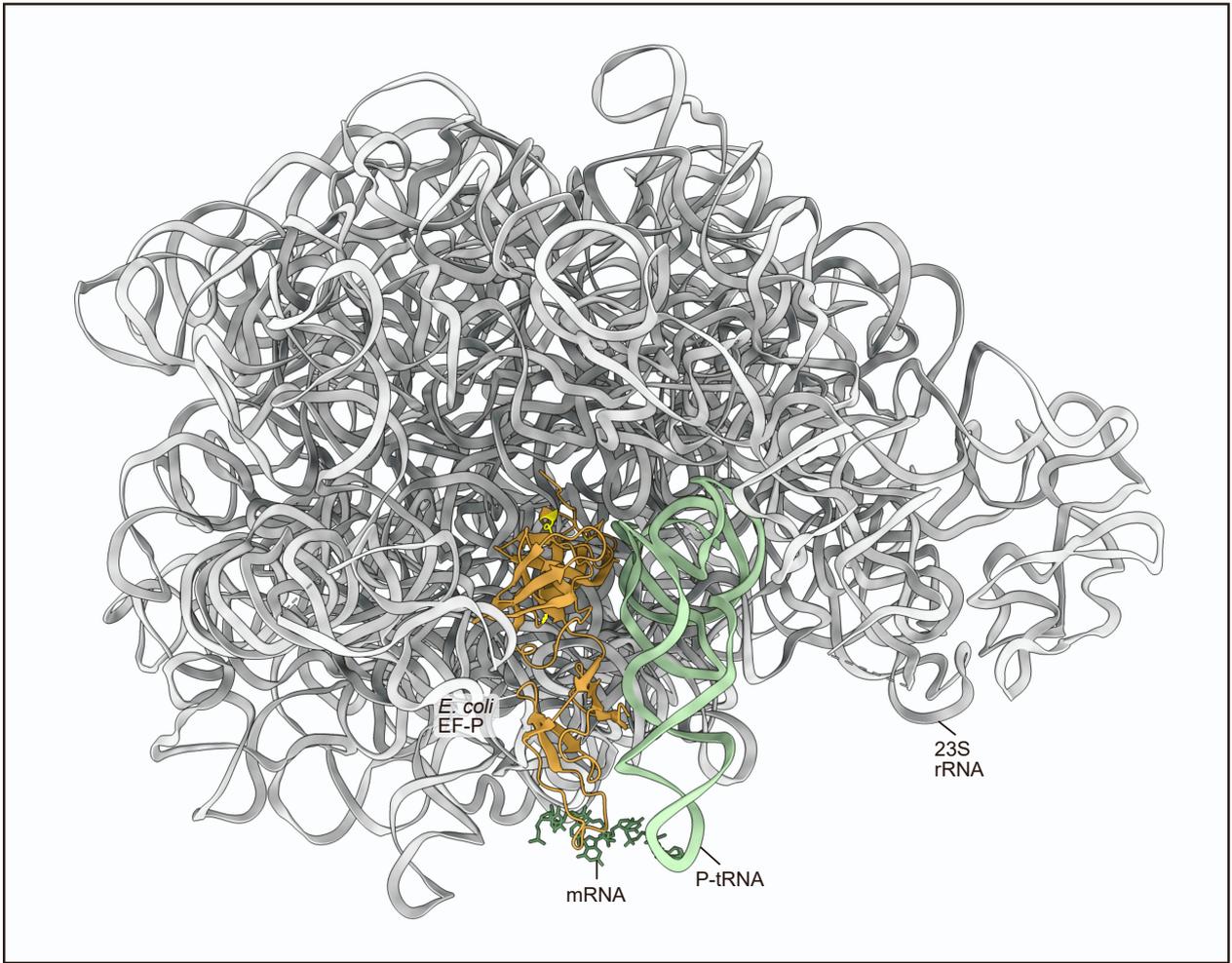
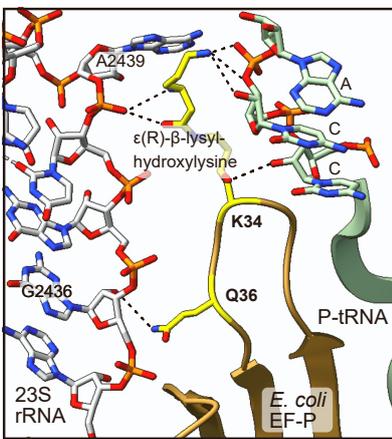
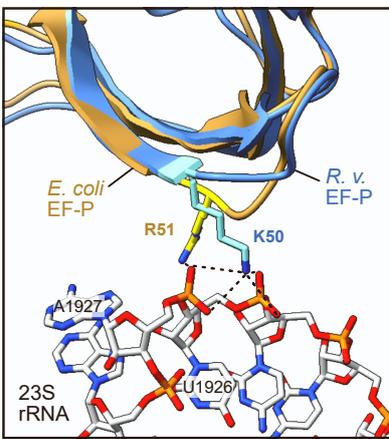
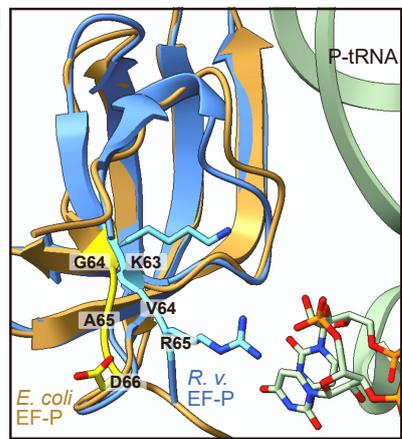
Supplementary Figure S5. Proteome-wide analysis of polyproline-containing proteins in the *E. coli* mutants Δefp , $\Delta efp \Delta epmA$ after complementation with *R. v. efp*. Related to Figure 2.

a-d Volcano plot analysis highlight the downregulation of proteins with/without polyproline motifs (PP-motifs) in *E. coli* mutants Δefp (a), $\Delta efp \Delta epmA$ (b), $\Delta efp::R. v. efp$ (c) and $\Delta efp::R. v. efp \Delta epmA$ (d) compared to wild type (wt). The x-axes show for each protein the fold change (FC) of the mean value of the \log_2 protein intensity (LFQ) between two strains. The y-axes show the significance level of the observed difference between the two strains ($-\log_{10}$ p-value of the t-test). The test was adjusted for multiple comparisons (permutation-based FDR with $n=3$, $FDR=0.05$ and $s_0=0.1$).



Supplementary Figure S6. Construction principles of *R. vannielii* EF-P and *E. coli* EF-P hybrids. Related to Figure 3.

a Nucleotide and protein sequence alignment of *E. coli* EF-P and *R. vannielii* EF-P with selected boundaries (black bulk arrows) for each EF-P domain (I-III). **b** Secondary structure and **c** 3D structural model of *E. coli* EF-P with selected boundaries between all domains (I-III) for hybrid construction (adapted from UniProt P0A6N4, and AlphaFold AF-P0A6N4-F1). **d** Western blot analysis of EF-P hybrid production in cells grown in LB medium to mid-exponential (4h) and early-stationary (6h) phase using antibodies against the His-tag. Protein bands corresponding to a 72 kDa protein after staining with 2,2,2-trichloroethanol (TCE) were used as loading controls. Structure prediction and the confidence score (pLDDT) were calculated by AlphaFold. 3D model confidence: very high – in dark blue (pLDDT > 90); confident – in light blue (90 > pLDDT > 70); low – yellow (70 > pLDDT > 50); very low – red (pLDDT < 50).

a**b****c****d**

Supplementary Figure S7. Model of potential interactions between *E. coli* / *R. vannielii* EF-P and 23S rRNA / P-site-tRNA in context of amino acid substitutions P35Q, V50R, and 63-65_KVR in *S. venezuelae* EF-P. Related to Figure 5.

The Cryo-EM structure of polyproline-stalled ribosome in the presence of *E. coli* EF-P was adopted from Huter *et al.*, 2017 [S2] (PDB accession number: 6ENU). Structural models were generated using UCSF ChimeraX [S3,S4]. The *R. v.* EF-P structure was predicted using AlphaFold2 ColabFold (v1.5.2) [S5,S6]. **a** Schematic presentation of *E. coli* EF-P (brown), localized relative to the ribosome (23S rRNA) (grey), mRNA (dark green) and P-site-bound tRNA^{Pro} (light green). **b** Potential interaction between Q36 of *E. coli* EF-P and 23S rRNA (G2436) is indicated with a dashed line. As described before [S2] the interactions of ϵ (R)- β -lysyl-hydroxylysine (after post translational modification of K34) with the CCA end of the P-site tRNA (P-tRNA) and the 23S rRNA (A2439) are shown in dashed lines. **c** Potential interactions between R51 of *E. coli* EF-P or K50 of *R. v.* EF-P with the phosphate backbone of the 23S rRNA (U1926, A1927), indicated in dashed lines. Domain I of the predicted *R. v.* EF-P was superimposed with domain I of the cryo-EM structure of *E. coli* EF-P. **d** Potential interactions between the motif KVR (positions 63-65 in *R. v.* EF-P) and the P-tRNA. For comparison, the position of the corresponding GAD motif (amino acids 64-66) in *E. coli* EF-P is marked. Amino acids K34, Q36, R51, G64, A65 and D66 in *E. coli* EF-P correspond to K33, P35, V50, T63, A64 and T65, respectively, in *S. venezuelae* EF-P (**Fig. 4a**).

References

- S1. Conners, E.M., Davenport, E.J., and Bose, A. (2021). Revised Draft Genome Sequences of *Rhodomicrobium vannielii* ATCC 17100 and *Rhodomicrobium udaipurensis* JA643. *Microbiol Resour Announc* 10. 10.1128/MRA.00022-21.
- S2. Huter, P., Arenz, S., Bock, L.V., Graf, M., Frister, J.O., Heuer, A., Peil, L., Starosta, A.L., Wohlgemuth, I., Peske, F., et al. (2017). Structural Basis for Polyproline-Mediated Ribosome Stalling and Rescue by the Translation Elongation Factor EF-P. *Molecular Cell* 68, 515-+. 10.1016/j.molcel.2017.10.014.
- S3. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* 27, 14-25. 10.1002/pro.3235.
- S4. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* 30, 70-82. 10.1002/pro.3943.
- S5. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-+. 10.1038/s41586-021-03819-2.
- S6. Mirdita, M., Schutze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nat Methods* 19, 679-682. 10.1038/s41592-022-01488-1.

2.1 Supplemental Data Tables of Chapter 2

The tables presented originate from Chapter 2 (<https://doi.org/10.1016/j.celrep.2024.114063>) and have been reformatted to facilitate their transfer from Microsoft Excel to Microsoft Word format.

Data S1. Analysis of the PTM status. Chymotrypsin digestion of *R. v.* EF-P, related to Figure 1

Supplemental online material

Data S2. Summary of the proteomic data. Comparison of detected proteins with and without PP motifs between WT and mutants, related to Figure 2

Supplemental online material

Data S3. Plasmid primer strain genomic DNA, related to STAR Methods**Plasmid List**

general			
vector name	insert	insert source	reference
709-FLPe	Plasmid for removal of FRT flanked resistance cassette (Cat. No. A104; Amp ^R)		GeneBridges, Germany
FRT-PGK-gb2-neo-FRT	PCR-template (plasmid DNA) for generation of a FRT-flanked PGK-gb2-neo cassette, Kan ^R		GeneBridges, Germany
pBAD33 (e.V.)	Cm ^R -cassette, p15A origin, araC coding sequence, ara operator		[1]
pET SUMO	pBR322 origin, <i>lacI</i> , T7 <i>lac</i> promoter, N-terminal <i>6xhis</i> , SUMO coding sequence, Kan ^R		Invitrogen
pNPTS138-R6KT	<i>mobRP4⁺ori-R6K sacB</i> ; suicide plasmid for gene deletions; Km ^R		[2]
<i>E. coli efp</i>			
vector name	insert	insert source	reference
pBAD33_ <i>E. coli efp_wt_6xhis</i>	<i>E. coli efp</i> fused to C-terminal <i>6xhis</i>	MG1655	This study
pNTPS138_500_ <i>E. coli efp_6xhis_500</i>	<i>E. coli efp</i> (fused to C-terminal <i>6xhis</i>) flanked by homologous regions for chromosomal gene substitution	BW25113; DSM 162 (genomic DNA)	This study
PGKGP subfamily <i>efps</i>			
vector name	insert	insert source	reference
pBAD33_ <i>C. algicola efp_6xhis</i>	<i>Cellulophaga algicola efp</i> fused to C-terminal <i>6xhis</i>	DSM-14237 (genomic DNA)	This study
pBAD33_ <i>C. algicola efp_V49R_6xhis</i>	<i>Cellulophaga algicola efp</i> with a substitution of valine to arginine at the position 49; fused to C-terminal <i>6xhis</i>	DSM-14237 (genomic DNA)	This study
pBAD33_ <i>C. hominis efp_6xhis</i>	<i>Campylobacter hominis efp</i> fused to C-terminal <i>6xhis</i>	DSM 21671 (genomic DNA)	This study
pBAD33_ <i>C. hominis efp_V50R_6xhis</i>	<i>Campylobacter hominis efp</i> with a substitution of valine to arginine at the position 50; fused to C-terminal <i>6xhis</i>	DSM 21671 (genomic DNA)	This study
pBAD33_ <i>C. lari efp_6xhis</i>	<i>Campylobacter lari efp</i> fused to C-terminal <i>6xhis</i>	DSM-11375 (genomic DNA)	This study
pBAD33_ <i>C. lari efp_V51R_6xhis</i>	<i>Campylobacter lari efp</i> with a substitution of valine to arginine at the position 51; fused to C-terminal <i>6xhis</i>	DSM-11375 (genomic DNA)	This study
pBAD33_ <i>C. woesei efp_6xhis</i>	<i>Conexibacter woesei efp</i> fused to C-terminal <i>6xhis</i>	DSM 14684 (genomic DNA)	This study
pBAD33_ <i>P. gingivalis efp_6xhis</i>	<i>Porphyromonas gingivalis efp</i> fused to C-terminal <i>6xhis</i>	DSM 20709 (genomic DNA)	This study
pBAD33_ <i>P. gingivalis efp_P34Q_6xhis</i>	<i>Porphyromonas gingivalis efp</i> with a substitution of proline to glutamine at the position 34; fused to C-terminal <i>6xhis</i>	DSM 20709 (genomic DNA)	This study
pBAD33_ <i>P. gingivalis efp_I49R_6xhis</i>	<i>Porphyromonas gingivalis efp</i> with a substitution of isoleucine to arginine at the position 49; fused to C-terminal <i>6xhis</i>	DSM 20709 (genomic DNA)	This study

pBAD33_ <i>P. gingivalis</i> <i>efp</i> _E62K_6xhis	<i>Porphyromonas gingivalis efp</i> with a substitution of glutamic acid to lysine at the position 62; fused to C-terminal 6xhis	DSM 20709 (genomic DNA)	This study
--	--	-------------------------	------------

R. vannielii efp

vector name	insert	insert source	reference
H1_pBAD33_I R. v._II R. v._III E. c. <i>efp</i> _6xhis	<i>efp</i> hybrid with the <i>efp</i> domains I - II from <i>R. vannielii</i> and <i>efp</i> domain III from <i>E. coli</i> ; fused to C-terminal 6xhis	DSM 162 (genomic DNA); MG1655	This study
H2_pBAD33_I R. v._II E. c._III E. c. <i>efp</i> _6xhis	<i>efp</i> hybrid with the <i>efp</i> domain I from <i>R. vannielii</i> and <i>efp</i> domains II - III from <i>E. coli</i> ; fused to C-terminal 6xhis	DSM 162 (genomic DNA); MG1655	This study
pBAD33_ <i>R. vannielii efp</i> _6xhis	<i>Rhodocrobium vannielii efp</i> fused to C-terminal 6xhis	DSM 162 (genomic DNA)	This study
pBAD33_ <i>R. vannielii efp</i> _K50R_6xhis	<i>Rhodocrobium vannielii efp</i> with a substitution of lysine to arginine at the position 50; fused to C-terminal 6xhis	DSM 162 (genomic DNA)	This study
pBAD33_ <i>R. vannielii efp</i> _K50V_6xhis	<i>Rhodocrobium vannielii efp</i> with a substitution of lysine to valine at the position 50; fused to C-terminal 6xhis	DSM 162 (genomic DNA)	This study
pET_SUMO_ <i>R. vannielii efp</i>	<i>Rhodocrobium vannielii efp</i> fused to N-terminal 6xhis-SUMO	DSM 162 (genomic DNA)	This study
pNTPS138_500_ <i>E. coli efp</i> ₁₋₁₇ _RBS_ <i>R. vannielii efp</i> _ <i>E. coli efp</i> ₁₈₂₋₁₈₉ _500	<i>Rhodocrobium vannielii efp</i> flanked by homologous regions, including <i>E. coli efp</i> ₁₋₅₁ and <i>E. coli efp</i> ₅₄₇₋₅₆₄ , for chromosomal gene substitution	BW25113; DSM 162 (genomic DNA)	This study
pNTPS138_500_ <i>E. coli efp</i> ₁₋₁₇ _RBS_ <i>R. vannielii efp</i> _his_ <i>E. coli efp</i> ₁₈₂₋₁₈₉ _500	<i>Rhodocrobium vannielii efp</i> (fused to C-terminal 6xhis) flanked by homologous regions, including <i>E. coli efp</i> ₁₋₅₁ and <i>E. coli efp</i> ₅₄₇₋₅₆₄ , for chromosomal gene substitution	BW25113; DSM 162 (genomic DNA)	This study

S. venezuelae efp

vector name	insert	insert source	reference
pBAD33_ <i>S. venezuelae efp</i> _6xhis	<i>Streptomyces venezuelae efp</i> fused to C-terminal 6xhis		[3]
pBAD33_ <i>S. venezuelae efp</i> _Q27E_6xhis	<i>Streptomyces venezuelae efp</i> with a substitution of glutamine to glutamic acid at the position 27; fused to C-terminal 6xhis	[3]	This study
pBAD33_ <i>S. venezuelae efp</i> _H28F_6xhis	<i>Streptomyces venezuelae efp</i> with a substitution of histidine to phenylalanine at the position 28; fused to C-terminal 6xhis	[3]	This study
pBAD33_ <i>S. venezuelae efp</i> _N56K_6xhis	<i>Streptomyces venezuelae efp</i> with a substitution of asparagine to lysine at the position 56; fused to C-terminal 6xhis	[3]	This study
pBAD33_ <i>S. venezuelae efp</i> _K60S_6xhis	<i>Streptomyces venezuelae efp</i> with a substitution of lysine to serine at the position 60; fused to C-terminal 6xhis	[3]	This study
pBAD33_ <i>S. venezuelae efp</i> _T65D_6xhis	<i>Streptomyces venezuelae efp</i> with a substitution of threonine to aspartic acid at the position 65; fused to C-terminal 6xhis	[3]	This study
pBAD33_ A1_ <i>S. venezuelae efp</i> _V50R_6xhis	<i>Streptomyces venezuelae efp</i> with a substitution of valine to arginine at the position 50; fused to C-terminal 6xhis	[3]	This study
pBAD33_ A2_ <i>S. venezuelae efp</i> _P35Q_6xhis	<i>Streptomyces venezuelae efp</i> with a substitution of proline to glutamine at the position 35; fused to C-terminal 6xhis	[3]	This study
pBAD33_ A3_ <i>S. venezuelae efp</i> _TAT-KVR_63-65_6xhis	<i>Streptomyces venezuelae efp</i> with a substitution of amino acids TAT* to KVR** at the positions 63-65; fused to C-terminal 6xhis	[3]	This study

pBAD33_A1_A2_S. venezuelae <i>efp_V50R_P35Q_6xhis</i>	<i>Streptomyces venezuelae efp</i> with substitutions: valine to arginine at the position 50, proline to glutamine at the position 35; fused to C-terminal <i>6xhis</i>	[3]	This study
pBAD33_A1_A3_S. venezuelae <i>efp_V50R_TAT-KVR_63-65_6xhis</i>	<i>Streptomyces venezuelae efp</i> with substitutions: valine to arginine at the position 50, TAT* to KVR** at the positions 63-65; fused to C-terminal <i>6xhis</i>	[3]	This study
pBAD33_A2_A3_S. venezuelae <i>efp_P35Q_TAT-KVR_63-65_6xhis</i>	<i>Streptomyces venezuelae efp</i> with substitutions: proline to glutamine at the position 35, TAT* to KVR** at the positions 63-65; fused to C-terminal <i>6xhis</i>	[3]	This study
pBAD33_A1_A2_A3_S. venezuelae <i>efp_V50R_P35Q_TAT-KVR_63-65_6xhis</i>	<i>Streptomyces venezuelae efp</i> with substitutions: valine to arginine at the position 50, proline to glutamine at the position 35, TAT* to KVR** at the positions 63-65; fused to C-terminal <i>6xhis</i>	[3]	This study

<i>W. virosa efp</i>			
vector name	insert	insert source	reference
pBAD33_W. virosa <i>efp_6xhis</i>	<i>Weeksella virosa efp</i> fused to C-terminal <i>6xhis</i>	DSM 16922 (genomic DNA)	This study
pBAD33_S1_W. virosa <i>efp_P34Q_6xhis</i>	<i>Weeksella virosa efp</i> with a substitution of proline to glutamine at the position 34; fused to C-terminal <i>6xhis</i>	DSM 16922 (genomic DNA)	This study
pBAD33_S2_W. virosa <i>efp_V49R_6xhis</i>	<i>Weeksella virosa efp</i> with a substitution of valine to arginine at the position 49; fused to C-terminal <i>6xhis</i>	DSM 16922 (genomic DNA)	This study
pBAD33_S3_W. virosa <i>efp_E62K_6xhis</i>	<i>Weeksella virosa efp</i> with a substitution of glutamic acid to lysine at the position 62; fused to C-terminal <i>6xhis</i>	DSM 16922 (genomic DNA)	This study
pBAD33_S1_S3_W. virosa <i>efp_P34Q_E62K_6xhis</i>	<i>Weeksella virosa efp</i> with substitutions: proline to glutamine at the position 34; glutamic acid to lysine at the position 62; fused to C-terminal <i>6xhis</i>	DSM 16922 (genomic DNA)	This study
pBAD33_S1_S2_S3_W. virosa <i>efp_P34Q_V49R_E62K_6xhis</i>	<i>Weeksella virosa efp</i> with substitutions: proline to glutamine at the position 34; valine to arginine at the position 49; glutamic acid to lysine at the position 62; fused to C-terminal <i>6xhis</i>	DSM 16922 (genomic DNA)	This study
pNTPS138_E. coli <i>efp</i> ₁₋₅₁ _RBS_W. <i>virosa efp_wt_6xhis_E. coli efp</i> ₅₄₇₋₅₆₄	<i>Weeksella virosa efp</i> (fused to C-terminal <i>6xhis</i>) flanked by homologous regions, including <i>E. coli efp</i> ₁₋₅₁ and <i>E. coli efp</i> ₅₄₇₋₅₆₄ , for chromosomal gene substitution	BW25113; DSM 16922 (genomic DNA)	This study
pNTPS138_E. coli <i>efp</i> ₁₋₅₁ _RBS_W. <i>virosa efp_S1-S3_E. coli efp</i> ₅₄₇₋₅₆₄	<i>Weeksella virosa efp_S1-S3</i> flanked by homologous regions, including <i>E. coli efp</i> ₁₋₅₁ and <i>E. coli efp</i> ₅₄₇₋₅₆₄ , for chromosomal gene substitution	BW25113; DSM 16922 (genomic DNA)	This study
pNTPS138_E. coli <i>efp</i> ₁₋₅₁ _RBS_W. <i>virosa efp_S1-S3_6xhis_E. coli efp</i> ₅₄₇₋₅₆₄	<i>Weeksella virosa efp_S1-S3</i> (fused to C-terminal <i>6xhis</i>) flanked by homologous regions, including <i>E. coli efp</i> ₁₋₅₁ and <i>E. coli efp</i> ₅₄₇₋₅₆₄ , for chromosomal gene substitution	BW25113; DSM 16922 (genomic DNA)	This study

* threonine-alanine-threonine; ** lysine-valine-arginine; Amp^R, Kan^R/Km^R – ampicillin, kanamycin resistance

Primer List

<i>epmA</i> deletion			
primer name	5'-3' sequence	restriction site	
P1	pRED_ <i>epmA</i> _del_F	TCG CGT TGC GAG TAG ACT TCG TGC CCT TGT CAA AAA CTG GAG ATT TAA CTA ATT AAC CCT CAC TAA AGG GCG	
P2	pRED_ <i>epmA</i> _del_R	ATG GCG CTT ATC ACG CCA TTC TTC GCT GTT AAT TCA GTA ATT TTT CAG AAT AAT ACG ACT CAC TAT AGG GCT C	
<i>E. coli efp</i>			
primer name	5'-3' sequence	restriction site	
P3	pB33_ <i>e. coli efp</i> _wt_F	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GGC AAC GTA CTA TAG CAA CGA TTT TC	XbaI
P4	pB33_ <i>e. coli efp</i> _wt_R	AGC CAA GCT TTT AGT GAT GGT GAT GGT GAT GAG ATC TCT TCA CGC GAG AGA CGT ATT CAC CA	HindIII
P5	pNPTS_ <i>e. coli efp</i> _6xhis_F	CAT CAC CAT CAC TAA TGC GGT TGT GGT GCG	
P6	pNPTS_ <i>e. coli efp</i> _6xhis_R	GTG ATG AGA TCT CTT CAC GCG AGA GAC GTA TTC	
PGKGP subfamily <i>efps</i>			
primer name	5'-3' sequence	restriction site	
P7	pB33_ <i>c. algicola efp</i> _wt_F	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GGC ATC AAC ATC AGA TAT TAG AAA AG	XbaI
P8	pB33_ <i>c. algicola efp</i> _wt_R	AGC CAA GCT TTT AGT GAT GGT GAT GGT GAT GAG ATC TTT CTT TAA CAC GTT CCA TGT AAG AA	HindIII
P9	pB33_ <i>c. algicola efp</i> _V49R_F	TTC AGG TAA GAG GTT AGA TAA TAC ATT TTC TGG	
P10	pB33_ <i>c. algicola efp</i> _V49R_R	GTT ACA CTT TTC AAC TTT GTA C	
P11	pB33_ <i>c. hominis efp</i> _wt_F	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GTC TTA CTC AAT GGG AGA TCT TAA AA	XbaI
P12	pB33_ <i>c. hominis efp</i> _wt_R	AGC CAA GCT TTT AGT GAT GGT GAT GGT GAT GAG ATC TTT TAT TTG CTT TTT CTA TGT ATT CA	HindIII
P13	pB33_ <i>c. hominis efp</i> _V50R_F	AGA CGG TAA AAG GTT GGA AAA AAC TTT CCA C	
P14	pB33_ <i>c. hominis efp</i> _V50R_R	ATA AAA GAT TTT ATT TTT ACG CG	
P15	pB33_ <i>c. lari efp</i> _wt_F	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GGC TTC TTA TGG AAT GGG AGA TTT AA	XbaI
P16	pB33_ <i>c. lari efp</i> _wt_R	AGC CAA GCT TTT AGT GAT GGT GAT GGT GAT GAG ATC TTT TAT TTG CTC TTT CTA TGT ATT CC	HindIII
P17	pB33_ <i>c. lari efp</i> _V51R_F	CGA TGG TAA GAG GTT AGA AAA AAC TTT CCA TG	
P18	pB33_ <i>c. lari efp</i> _V51R_R	ATA AAA GAT TTA ATT TTA ATA CGA ACA AAA G	
P19	pB33_ <i>c. woeseii efp</i> _wt_F	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GAT TTC TAC GAA CCA GCT GAA GAA CG	XbaI
P20	pB33_ <i>c. woeseii efp</i> _wt_R	AGC CAA GCT TCT AGT GAT GGT GAT GGT GAT GAG ATC TCG CGC GGG ACA CGT AGT CGC CGG AG	HindIII
P21	pB33_ <i>p. gingivalis efp</i> _wt_F	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GGC TAC AAC GGC AGA CTT TCG CAA CG	XbaI
P22	pB33_ <i>p. gingivalis efp</i> _wt_R	AGC CAA GCT TTT AGT GAT GGT GAT GGT GAT GAG ATC TTT CCT TTA CTC TAC CGA TGT AGG AA	HindIII
P23	pB33_ <i>p. gingivalis efp</i> _P34Q_F	CGG CAA AGG CCA GGC TTT CGT CC	
P24	pB33_ <i>p. gingivalis efp</i> _P34Q_R	GGC TTG ACG TGA AGG AAT TCG AC	
P25	pB33_ <i>p. gingivalis efp</i> _I49R_F	ACA GGA CGT AGG CTC GAT AAG ACC	
P26	pB33_ <i>p. gingivalis efp</i> _I49R_R	AGC CAC GTT CTT GAG TTT G	

P27 pB33_p. *gingivalis_efp_E62K_F* CAA GGT CGA AAA AGT CAG AAT AGA ACG ACG
 P28 pB33_p. *gingivalis_efp_E62K_R* ACG CCG CTG TTC CAG GTC

R. vannielii efp

primer name	5'-3' sequence	restriction site
P29 pB33_r. <i>vannielii_efp_wt_F</i>	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GAA AAT CAA CGG CAA CGA AAT CCG TC	XbaI
P30 pB33_r. <i>vannielii_efp_wt_R</i>	AGC CAA GCT TCT AGT GAT GGT GAT GGT GAT GAG ATC TCT CCG CGC GGC GGA TAT AGG CGA TC	HindIII
P31 ov_r. v.I_e. c.II and III_F	CCG AGA CCG TTG AGA AAG TTG ATG TTG TCG ATA TGA ACC T	
P32 ov_r. v.I_e. c.II and III_R	AGG TTC ATA TCG ACA ACA TCA ACT TTC TCA ACG GTC TCG G	
P33 ov_r. v.II_e. c.III_F	CCA TCG GCA TCG AGC TGC CCA ACT TCG TTG AAC TGG AAA T	
P34 ov_r. v.II_e. c.III_R	ATT TCC AGT TCA ACG AAG TTG GGC AGC TCG ATG CCG ATG G	
P35 pB33_r. <i>vannielii_efp_K50R_F</i>	AAC GGA TCG AGG CTT AAC GAG	
P36 pB33_r. <i>vannielii_efp_K50R_R</i>	GAT AAG GTT TTT CAA CTC GAC	
P37 pB33_r. <i>vannielii_efp_K50V_F</i>	CAA CGG ATC GGT GCT TAA CGA G	
P38 pB33_r. <i>vannielii_efp_K50V_R</i>	ATA AGG TTT TTC AAC TCG AC	
P39 pET_r. <i>vannielii_efp_SUMO_F</i>	ATG AAA ATC AAC GGC AAC GAA ATC CGT CC	
P40 pET_r. <i>vannielii_efp_SUMO_R</i>	CTA CTC CGC GCG GCG GAT ATA G	
P41 pNPTS_pre_r. <i>vannielii_efp_500_F</i>	GAT CGG ATC CCG CAG CAA CAT ACT CAA GTG CAG	BamHI
P42 pNPTS_ov_pre_r. <i>vannielii_efp_500_R</i>	AGC TGT CCT CCT CTT TAA TAT CTA TAA CAT GAT TTT AAG ACC AGC ACG AAA ATC G	
P43 pNPTS_ov_ec_efp 1-17_r. <i>vannielii_efp_F</i>	TTA AAG AGG AGG ACA GCT ATG AAA ATC AAC GGC AAC GAA ATC CGT CC	
P44 pNPTS_ov_r. <i>vannielii_efp_ec_efp182-189_R</i>	AGA GAC GTA TCT ACT CCG CGC GGC GGA TAT AG	
P45 pNPTS_ov_post_r. <i>vannielii_efp_500_F</i>	CGG AGT AGA TAC GTC TCT CGC GTG AAG TAA TGC	
P46 pNPTS_post_r. <i>vannielii_efp_500_R</i>	GAT CGC TAG CCA GAA AAA ACA AAC GGC ACG ACA CAG	NheI
P47 pNPTS_r. <i>vannielii_efp_6xhis_F</i>	CAT CAC CAT CAC TAG ATA CGT CTC TCG CGT GAA GTA ATG	
P48 pNPTS_r. <i>vannielii_efp_6xhis_R</i>	GTG ATG AGA TCT CTC CGC GCG GCG GAT ATA	

S. venezuelae efp

primer name	5'-3' sequence	restriction site
P49 pB33_s. <i>venezuelae_efp_Q27E_F</i>	CGT CGA GTT CGA GCA CGT CAA GCC	
P50 pB33_s. <i>venezuelae_efp_Q27E_R</i>	ACG GAC CAG AGC TGG CCG	
P51 pB33_s. <i>venezuelae_efp_H28F_F</i>	CGA GTT CCA GTT CGT CAA GCC CGG	
P52 pB33_s. <i>venezuelae_efp_H28F_R</i>	ACG ACG GAC CAG AGC TGG	
P53 pB33_s. <i>venezuelae_efp_P35Q_F</i>	CGG CAA GGG CCA GGC CTT CGT GC	
P54 pB33_s. <i>venezuelae_efp_P35Q_R</i>	GGC TTG ACG TGC TGG AAC TCG ACG	
P55 pB33_s. <i>venezuelae_efp_V50R_F</i>	GTC CGG CAA GAG GGT CGA CAA GAC	
P56 pB33_s. <i>venezuelae_efp_V50R_R</i>	AGC ACG TTC TTG AGC TTG	
P57 pB33_s. <i>venezuelae_efp_N56K_F</i>	AGA CCT TCA AAG CCG GTG TGA AG	

P58	pB33_s. <i>venezuelae_efp</i> _N56K_R	TGT CGA CGA CCT TGC CGG
P59	pB33_s. <i>venezuelae_efp</i> _K60S_F	GCC GGT GTG AGC GTC GAG ACG G
P60	pB33_s. <i>venezuelae_efp</i> _K60S_R	GTT GAA GGT CTT GTC GAC
P61	pB33_s. <i>venezuelae_efp</i> _T63K_F	GAA GGT CGA GAA GGC CAC CAT CG
P62	pB33_s. <i>venezuelae_efp</i> _T63K_R	ACA CCG GCG TTG AAG GTC
P63	pB33_s. <i>venezuelae_efp</i> _A64V_F	GGT CGA GAC GGT CAC CAT CGA CC
P64	pB33_s. <i>venezuelae_efp</i> _A64V_R	TTC ACA CCG GCG TTG AAG G
P65	pB33_s. <i>venezuelae_efp</i> _T65D_F	CGA GAC GGC CGA CAT CGA CCG C
P66	pB33_s. <i>venezuelae_efp</i> _T65D_R	ACC TTC ACA CCG GCG TTG
P67	pB33_s. <i>venezuelae_efp</i> _T65R_F	CGA GAC GGC CAG AAT CGA CCG CC
P68	pB33_s. <i>venezuelae_efp</i> _T65R_R	ACC TTC ACA CCG GCG TTG
P69	pB33_s. <i>venezuelae_efp</i> _TAT/KVR_F	CAG AAT CGA CCG CCG CGA GAT G
P70	pB33_s. <i>venezuelae_efp</i> _TAT/KVR_R	ACC TTC TCG ACC TTC ACA CCG GC
P71	pB33_s. <i>venezuelae_efp</i> _R145T_F	CCA GGG TGA CAC CTC CAC CGG TGG CAC C
P72	pB33_s. <i>venezuelae_efp</i> _R145T_R	ACG CCC GGG TCG GTG TGC
P73	pB33_s. <i>venezuelae_efp</i> _141-146_F	GAC ACC GCC ACC GGT GGC ACC AAG CCC
P74	pB33_s. <i>venezuelae_efp</i> _141-146_R	ACC CTT CAA GCC CGG GTC GGT GTG CTC

W. virosa efp

	primer name	5'-3' sequence	restriction site
P75	pB33_w. <i>virosa_efp</i> _wt_F	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GGC AAC AAC TTC AGA TAT CAG AAA AG	XbaI
P76	pB33_w. <i>virosa_efp</i> _wt_R	AGC CAA GCT TTT AGT GAT GGT GAT GGT GAT GAG ATC TTT CTT TCA CGC GTT CTA CAT ACG AA	HindIII
P77	pB33_w. <i>virosa_efp</i> _P34Q_ov1_F	CCA GGG AAA GGA CAG GCA TTT GTT CGT	
P78	pB33_w. <i>virosa_efp</i> _P34Q_ov1_R	ACG AAC AAA TGC CTG TCC TTT CCC TGG	
P79	pB33_w. <i>virosa_efp</i> _V49R_F	CAA CGG AAA AAG GTT AGA AAA CAC TTT CGC	
P80	pB33_w. <i>virosa_efp</i> _V49R_R	GTT ACT GAT TTT AAT TTT GTA CG	
P81	pB33_w. <i>virosa_efp</i> _E62K_F	TAA AAT CGA TAA AGT GCG TGT AGA AAC C	
P82	pB33_w. <i>virosa_efp</i> _E62K_R	TGA CCC GCT GCG AAA GTG	
P83	pNPTS_ov_w. <i>virosa_efp</i> _gib_F	TAT TAA AGA GGA GGA CAG CTA TGG CAA CAA CTT CAG ATA T	
P84	pNPTS_ov_w. <i>virosa_efp</i> _gib_R	ACT TCA CGC GAG AGA CGT ATT TAT TCT TTC ACG CGT TCT A	
P85	pNPTS_ov_vec_w. <i>virosa_efp</i> _gib_F	TAG AAC GCG TGA AAG AAT AAA TAC GTC TCT CGC GTG AAG T	
P86	pNPTS_ov_vec_w. <i>virosa_efp</i> _gib_R	ATA TCT GAA GTT GTT GCC ATA GCT GTC CTC CTC TTT AAT A	
P87	pNPTS_ov_w. <i>virosa_efp</i> _6xhis_gib_F	ACG CGT GAA AGA AAG ATC TCA TCA CCA TCA CCA TCA CTA AAT ACG TCT CTC GCG TGA AGT A	
P88	pNPTS_ov_w. <i>virosa_efp</i> _6xhis_gib_R	TAC TTC ACG CGA GAG ACG TAT TTA GTG ATG GTG ATG GTG ATG AGA TCT TTC TTT CAC GCG T	

check primers

	primer name	5'-3' sequence	restriction site
P89	check_r. <i>vannielii_efp</i> _im chrom_F	AGA TCG CCC GTC GCC GGA TTG TAA ATG AGG CCC GCC ACG AG	
P90	check_r. <i>vannielii_efp</i> _im chrom_R	GAA CCT TGT GCG CCC GAA GCC ATC CCG GTG CTC GGC CGC GA	
P91	check_epmA_125_F	AGT CTG AAA TTT GAC GAG ATC GAT T	

P92	check_epmA_125_R	AGC GCG TGT ATA ACA CGT GCC CAC T
P93	check_pBAD33_F	GGC GTC ACA CTT TGC TAT GC
P94	check_pBAD33_R	CAGTTCCTACTCTCGCATG
P95	check_pET_T7 prom_F	CCT ATA GTG AGT CGT ATT A
P96	check_pET_T7 term_R	TAT GCT AGT TAT TGC TCA G
P97	check_pNPTS_F	GAC CAT GAT TAC GCC AAG CTA CGT
P98	check_pNPTS_R	TGT GCT GCA AGG CGA TTA AGT TGG

Strain List

general		
strain	description/ features	reference
BL21 (DE)	F ⁻ ompT gal dcm lon hsdSB(r _B ⁻ m _B ⁻) λ(DE3)	[4]
BW25113	Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB)568, hsdR514	[5]
DH5α λpir	endA1 hsdR17 glnV44 (= supE44) thi-1 recA1 gyrA96 relA1 φ80'lacΔ(lacZ)M15 Δ(lacZYA-argF)U169 zdg-232::Tn10 uidA::pir +	[6]
WM3064	thrB1004 pro thi rpsL hsdS lacZΔM15 RP4-1360 Δ(araBAD)567 ΔdapA1341::[erm pir]	[7]
<i>E. coli</i> EF-P modification		
strain	description/ features	reference
Δefp (JW4107)	deletion of <i>efp</i> in BW25113	[8]
ΔepmA	deletion of <i>epmA</i> in BW25113	This study
Δefp ΔepmA	deletion of <i>efp</i> and <i>epmA</i> in BW25113	[9]
<i>efp</i> _his	native <i>efp</i> C-terminally fused to 6xhis in BW25113	This study
<i>efp</i> _his ΔepmA	native <i>efp</i> C-terminally fused to 6xhis; deletion of <i>epmA</i> in BW25113	This study
<i>R. vannielii</i> efp		
strain	description/ features	reference
Δefp:: <i>R. v.</i> efp	replacement of native <i>efp</i> with <i>Rhodocrobium vannielii</i> <i>efp</i> in BW25113	This study
Δefp:: <i>R. v.</i> efp ΔepmA	replacement of native <i>efp</i> with <i>Rhodocrobium vannielii</i> <i>efp</i> ; deletion of <i>epmA</i> in BW25113	This study
Δefp:: <i>R. v.</i> efp_his	replacement of native <i>efp</i> with <i>Rhodocrobium vannielii</i> <i>efp</i> (C-terminally fused to 6xhis) in BW25113	This study
Δefp:: <i>R. v.</i> efp_his ΔepmA	replacement of native <i>efp</i> with <i>Rhodocrobium vannielii</i> <i>efp</i> (C-terminally fused to 6xhis); deletion of <i>epmA</i> in BW25113	This study
<i>W. virosa</i> efp		
strain	description/ features	reference
Δefp:: <i>W. v.</i> efp_his	replacement of native <i>efp</i> with <i>Weeksella virosa</i> <i>efp</i> in BW25113	This study
Δefp:: <i>W. v.</i> efp_his_ΔepmA	replacement of native <i>efp</i> with <i>Weeksella virosa</i> <i>efp</i> ; deletion of <i>epmA</i> in BW25113	This study
Δefp:: <i>W. v.</i> efp_S1-S3	replacement of native <i>efp</i> with <i>Weeksella virosa</i> <i>efp</i> containing substitutions S1, S2 and S3 (P34Q, V49R and E62K) in BW25113	This study
Δefp:: <i>W. v.</i> efp_S1-S3 ΔepmA	replacement of native <i>efp</i> with <i>Weeksella virosa</i> <i>efp</i> containing substitutions S1, S2 and S3 (P34Q, V49R and E62K); deletion of <i>epmA</i> in BW25113	This study

$\Delta efp::W. v. efp_S1-S3_his$	replacement of native <i>efp</i> with <i>Weeksellia virosa efp</i> containing substitutions S1, S2 and S3 (P34Q, V49R and E62K), C-terminally fused to <i>6xhis</i> in BW25113	This study
$\Delta efp::W. v. efp_S1-S3_his$ $\Delta epmA$	replacement of native <i>efp</i> with <i>Weeksellia virosa efp</i> containing substitutions S1, S2 and S3 (P34Q, V49R and E62K), C-terminally fused to <i>6xhis</i> ; deletion of <i>epmA</i> in BW25113	This study

Reporter for EF-P activity measurements

strain	description/ features	reference
MG1655 <i>P_{cadBA}::lacZ</i> $\Delta cadB$ Δefp	deletion of <i>cadB</i> and <i>cadA</i> by substitution with <i>lacZ</i> ; deletion of <i>efp</i> in MG1655	[10]
MG1655 <i>P_{cadBA}::lacZ cadBA</i> Δefp $\Delta epmA$	deletion of <i>cadB</i> and <i>cadA</i> by substitution with <i>lacZ</i> ; deletion of <i>efp</i> and <i>epmA</i> in MG1655	[10]

Genomic DNA List

strain	description	source
DSM 11375	genomic DNA isolated from <i>Campylobacter lari</i> (DSM No.: 11375), type strain	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH
DSM 14237	genomic DNA isolated from <i>Cellulophaga algicola</i> (DSM No.: 14237), type strain	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH
DSM 14684	genomic DNA isolated from <i>Conexibacter woesei</i> (DSM No.: 14684), type strain	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH
DSM 16922	genomic DNA isolated from <i>Weeksellia virosa</i> (DSM No.: 16922), type strain	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH
DSM 20709	genomic DNA isolated from <i>Porphyromonas gingivalis</i> (DSM No.: 20709), type strain	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH
DSM 21671	genomic DNA isolated from <i>Campylobacter hominis</i> (DSM No.: 21671), type strain	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH
DSM 162	genomic DNA isolated from <i>Rhodomicrobium vannielii</i> (DSM No.: 162), type strain	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH
DSM 23294	genomic DNA isolated from <i>Rhodomicrobium vannielii</i> (DSM No.: 23294)	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH
DSM 23295	genomic DNA isolated from <i>Rhodomicrobium vannielii</i> (DSM No.: 23295)	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH
DSM 23296	genomic DNA isolated from <i>Rhodomicrobium vannielii</i> (DSM No.: 23296)	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH

References for Supplemental Data Tables of Chapter 2

1. Guzman LM, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J Bacteriol* 177, 4121-4130 (1995).
2. Lassak J, Henche AL, Binnenkade L, Thormann KM. ArcS, the cognate sensor kinase in an atypical Arc system of *Shewanella oneidensis* MR-1. *Appl Environ Microbiol* 76, 3263-3274 (2010).
3. Pinheiro B, et al. Structure and Function of an Elongation Factor P Subfamily in Actinobacteria. *Cell Rep* 30, 4332-4342 e4335 (2020).
4. Studier, F.W. and B.A. Moffatt, Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*, 1986. 189(1): p. 113-30.
5. Datsenko, K.A. and B.L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*, 2000. 97(12): p. 6640-5.
6. Macinga, D.R., M.M. Parojcic, and P.N. Rather, Identification and analysis of *aarP*, a transcriptional activator of the 2'-*N*-acetyltransferase in *Providencia stuartii*. *J Bacteriol*, 1995. 177(12): p. 3407-13.
7. W. Metcalf, University of Illinois, Urbana
8. Baba, T., et al., Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*, 2006. 2: p. 2006 0008.
9. Lassak J, et al. Arginine-rhamnosylation as new strategy to activate translation elongation factor P. *Nat Chem Biol* 11, 266-270 (2015).
10. Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N. and Jung, K. (2013) Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science*, 339, 82-85.

3 Versatile Dual Reporter to Identify Ribosome Pausing Motifs Alleviated by Translation Elongation Factor P

Tomasiunaite U, Brewer T, Burdack K, Brameyer S, Jung K. Versatile Dual Reporter to Identify Ribosome Pausing Motifs Alleviated by Translation Elongation Factor P. ACS Synth. Biol. 2024, 13, 11, 3698–3710 <https://doi.org/10.1021/acssynbio.4c00534>

Reprinted with permission from DOI: <https://doi.org/10.1021/acssynbio.4c00534>. Copyright 2024 American Chemical Society.

Versatile Dual Reporter to Identify Ribosome Pausing Motifs Alleviated by Translation Elongation Factor P

Urte Tomasiunaite, Tess Brewer, Korinna Burdack, Sophie Brameyer, and Kirsten Jung*

Cite This: *ACS Synth. Biol.* 2024, 13, 3698–3710

Read Online

ACCESS |



Metrics & More



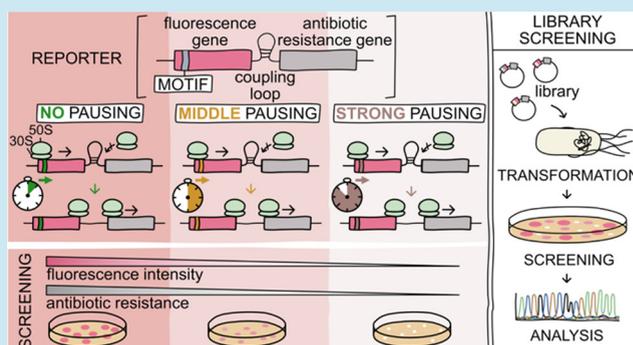
Article Recommendations



Supporting Information

ABSTRACT: Protein synthesis is influenced by the chemical and structural properties of the amino acids incorporated into the polypeptide chain. Motifs containing consecutive prolines can slow the translation speed and cause ribosome stalling. Translation elongation factor P (EF-P) facilitates peptide bond formation in these motifs, thereby alleviating stalled ribosomes and restoring the regular translational speed. Ribosome pausing at various polyproline motifs has been intensively studied using a range of sophisticated techniques, including ribosome profiling, proteomics, and in vivo screening, with reporters incorporated into the chromosome. However, the full spectrum of motifs that cause translational pausing in *Escherichia coli* has not yet been identified. Here, we describe a plasmid-based dual reporter for rapid assessment of pausing motifs. This reporter contains two coupled genes encoding mScarlet-I and chloramphenicol acetyltransferase to screen motif libraries based on both bacterial fluorescence and survival. In combination with a diprolyl motif library, we used this reporter to reveal motifs of different pausing strengths in an *E. coli* strain lacking *efp*. Subsequently, we used the reporter for a high-throughput screen of four motif libraries, with and without prolines at different positions, sorted by fluorescence-associated cell sorting (FACS) and identify new motifs that influence the translational efficiency of the fluorophore. Our study provides an in vivo platform for rapid screening of amino acid motifs that affect translational efficiencies.

KEYWORDS: stalling motifs, ribosome stalling, screening for ribosome pausing strength, translational efficiency, dual reporter, proteins with polyproline motifs, elongation factor P



INTRODUCTION

The regulation of protein synthesis is critical for the maintenance of physiological balance and the survival of organisms. A number of factors can modulate translational efficiency, including charge-specific interactions between the ribosome peptide exit tunnel and the nascent chain,^{1–3} mRNA secondary structures,^{4–7} tRNA abundance and modification,^{8–10} and codon selection.^{11,12} It has been widely demonstrated that certain combinations of amino acids, in particular stretches of polyprolines, have the ability to negatively impact translational efficiency and, in certain instances, cause the ribosome to stall.^{13–16} The profound negative effect of proline on the translation rate can be attributed to its unique structural nature, which includes a pyrrolidine ring that renders it conformationally rigid. This ultimately results in proline being a poor donor and poor acceptor during the transpeptidation reaction.^{17–20} Despite the fact that proline is a challenging factor during translation, diprolyl motifs (XPPX) are very common in bacterial proteomes. In *Escherichia coli*, there are 2184 of these motifs (0.51 motif/protein), in *Salmonella enterica* 2324 motifs (0.51 motif/protein), in *Pseudomonas aeruginosa* 4074 motifs (0.73

motif/protein), and *Streptomyces coelicolor* even contains 8719 motifs (1.08 motif/protein).^{21,22} The prevalence of these motifs is not limited to bacterial proteomes—they are also common among the proteomes of eukaryotes.²³ The ubiquity of these motifs in the proteomes of organisms underlines their functional and regulatory importance to shape protein structure,^{24,25} adjust copy number,²¹ and incorporate transmembrane proteins.²⁶

The biosynthesis of proteins containing polyproline motifs is mediated by distinct translation factors that have evolved in organisms from all domains of life. While eukaryotes and archaea encode initiation factor 5A (e/aIF5A),²⁷ bacteria have evolved elongation factor P (EF-P), which reduces ribosomal stalling by assisting during the translation of polyproline motifs.^{14,15} The EF-Ps of numerous bacterial species are

Received: August 6, 2024
Revised: October 5, 2024
Accepted: October 9, 2024
Published: October 19, 2024



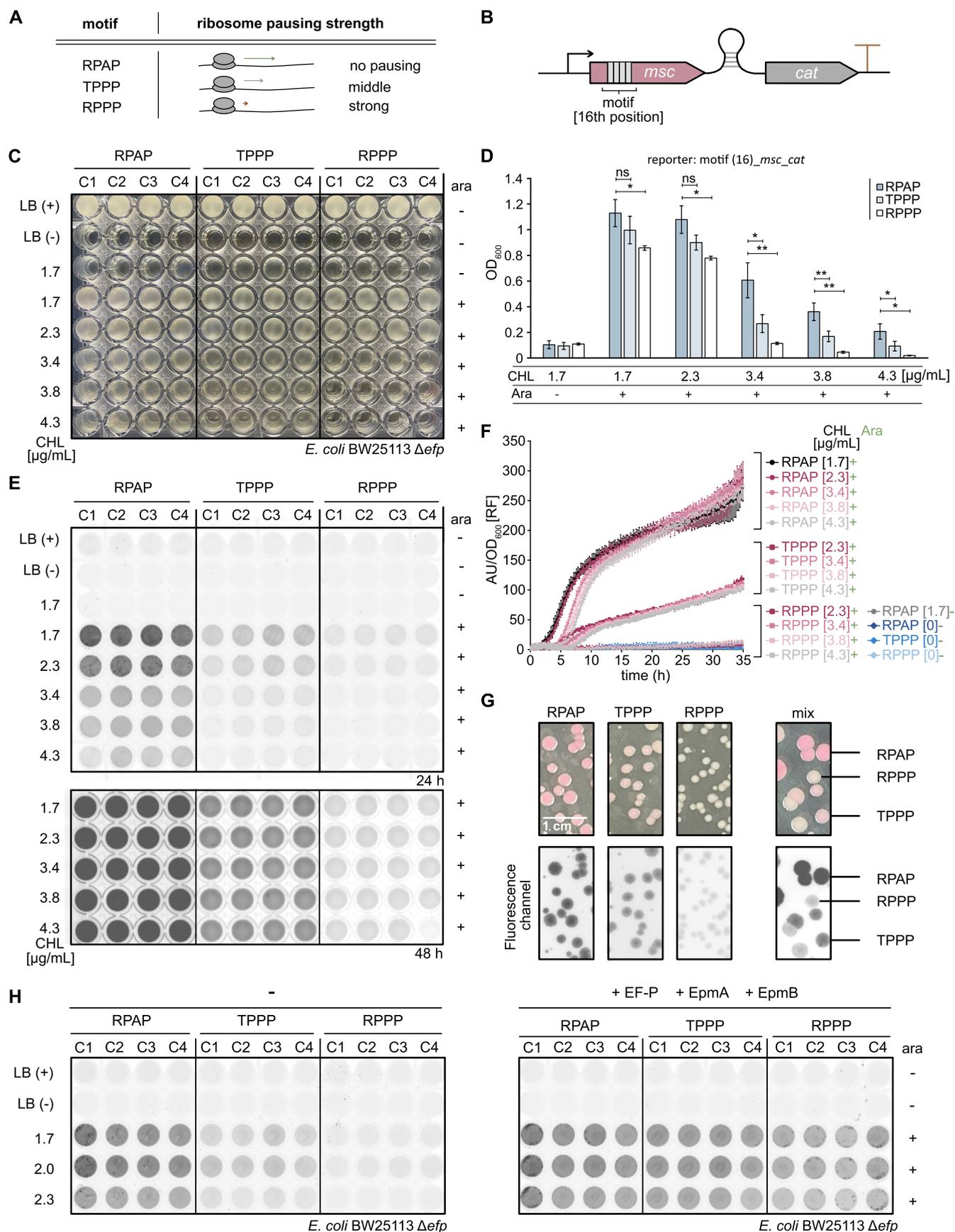


Figure 1. Development of a dual-reporter system to screen for ribosome pausing motifs. (A) Schematic of known ribosome pausing strengths at motifs RPAP (Arg-Pro-Ala-Pro), TPPP (Thr-Pro-Pro-Pro), and RPPP (Arg-Pro-Pro-Pro) alleviated by elongation factor P.^{14,38,41} (B) Construction of a dual reporter for translational efficiency measurements. The reporter consists of a gene coding for fluorophore (mScarlet-I) and a resistance

Figure 1. continued

gene cassette (*chloramphenicol acetyltransferase*, *cat*), both linked by a weak translational coupling device.⁴³ Nucleotides encoding the RPAP, TPPP, and RPPP motifs were incorporated into the fluorophore gene sequence starting from the 16th nucleotide position (16th). (C, D) 96-well assay plate of the growth measurements (C) and growth quantification (D) of *E. coli* BW25113 Δ *efp*, transformed with the dual-reporter plasmid pBAD24_motif_mscarlet-I_cat. When indicated, the medium was supplemented with 0.2% (w/v) arabinose (Ara). (E) Fluorescence measurements of the 96-well assay plate 24 h (upper panel) or 48 h (bottom panel) after bacteria inoculation. *E. coli* BW25113 Δ *efp* was transformed with the dual-reporter plasmid pBAD24_motif_mscarlet-I_cat. (F) Relative fluorescence (RF) measurements over time of *E. coli* BW25113 Δ *efp* with dual-reporter plasmid pBAD24_motif_mscarlet-I_cat. The M9 medium was supplemented with different concentrations of chloramphenicol (CHL). mScarlet-I fluorescence signal, given in arbitrary units (AU), was normalized to the measured optical density (OD₆₀₀) and shown as RF. (G) Colony morphology and fluorescence analysis of *E. coli* BW25113 Δ *efp* with dual-reporter plasmid pBAD24_motif_mscarlet-I_cat. Bacteria were plated on LB agar assay plates supplemented with 0.2% (w/v) Ara and 3.4 μ g/mL CHL. (H) EF-P complementation assay. Fluorescence measurements of the 96-well assay plate 30 h after bacteria inoculation. *E. coli* BW25113 Δ *efp* was transformed with the dual-reporter plasmid pBAD24_motif_mscarlet-I_cat (left panel) or with the dual-reporter plasmid pBAD24_motif_mscarlet-I_cat and pBAD_E. coli *efp-epmA-epmB* (right panel). Error bars indicate the standard deviation (SD) of three independent biological replicates. Statistics: student's unpaired two-sided *t* test (*****p* < 0.0001; ****p* < 0.001; ***p* < 0.01; **p* < 0.05; ns *p* > 0.05). 1.7 μ g/mL CHL and 0.2% (w/v) Ara (RPAP vs TPPP: ns *p* = 0.178 [D]; RPAP vs RPPP: **p* = 0.016 [D]); 2.3 μ g/mL CHL and 0.2% (w/v) Ara (RPAP vs TPPP: ns *p* = 0.055 [D]; RPAP vs RPPP: **p* = 0.013 [D]); 3.4 μ g/mL CHL and 0.2% (w/v) Ara (RPAP vs TPPP: **p* = 0.015 [D]; RPAP vs RPPP: ***p* = 0.008 [D]); 3.8 μ g/mL CHL and 0.2% (w/v) Ara (RPAP vs TPPP: ***p* = 0.009 [D]; RPAP vs RPPP: ***p* = 0.003 [D]); 4.3 μ g/mL CHL and 0.2% (w/v) Ara (RPAP vs TPPP: **p* = 0.038 [D]; RPAP vs RPPP: **p* = 0.012 [D]). C1–C4 is the clone number.

dependent on unique post-translational modifications (PTM), which are necessary to improve their activity.^{28–35} In contrast, certain EF-Ps, particularly from the subfamily with the activation loop PGKGP (Pro-Gly-Lys-Gly-Pro), exhibit full functionality in the absence of PTMs, although with reduced levels of activity in *E. coli*.^{22,36} Intriguingly, our recent findings even demonstrate that some unmodified EF-Ps, which are barely active in *E. coli*, can be engineered to be functional in *E. coli*.³⁶ This suggests that full EF-P functionality is not exclusively dependent on PTM, as previously assumed. Notably, we observed variations in the efficacy of distinct polyproline motif rescue among modified and unmodified EF-Ps, suggesting that the underlying general mechanisms of protein synthesis are not yet fully understood and require further investigation.

A number of studies have employed a range of techniques, including ribosome profiling, mass spectrometry, and in vitro and in vivo reporter systems, to identify motifs associated with different degrees of ribosome pausing.^{14,16,37–41} The majority of these studies focus on motifs that contain at least two prolines, as EF-P is well known to support their synthesis during the transpeptidation reaction.^{16,38} The literature on this topic is limited with regards to motifs that contain a single proline or even no proline, whose rescue is dependent on a functional EF-P.^{37,42} The existing knowledge gaps highlight the need for simple, accessible tools that can be utilized by every laboratory at a low cost with a minimal input of materials to investigate translational dynamics.

In this study, we present a user-friendly, plasmid-based reporter system for investigating translational efficiencies of amino acid motifs. Translational coupling⁴³ of the genes encoding the fluorophore mScarlet-I and chloramphenicol acetyltransferase enhances screening capacity and allows dual output measurements of cell fluorescence and viability. High-throughput screening demonstrates that this system effectively identifies unusual sequence motifs, particularly those lacking proline, which cause reduced translation rates.

RESULTS

Plasmid-Based In Vivo Dual-Reporter System Enables Ribosome Pausing Measurements in a Δ *efp* Mutant. To facilitate the rapid evaluation of which amino acid motifs can influence translational efficiency, we established a plasmid-

based reporter system. Motifs that are known to cause various levels of ribosome pausing were chosen for initial experiments: RPAP (Arg-Pro-Ala-Pro) represented no pausing strength, TPPP (Thr-Pro-Pro-Pro) represented middle pausing strength, and RPPP (Arg-Pro-Pro-Pro) represented strong pausing strength^{14,38,41} (Figure 1A). To identify differences in translational efficiency resulting from distinct sequence motifs, an *E. coli* BW25113 mutant lacking *elongation factor P* (Δ *efp*) was transformed with these reporter plasmids. The initial reporter design consisted of an antibiotic resistance cassette (*chloramphenicol acetyltransferase*, *cat*) with motifs integrated at the N-terminus (Figure S1A, left panel). This approach allowed estimation of the translational efficiencies of *cat* influenced by the motif sequence based on bacterial survival in the presence of chloramphenicol. A 24 h incubation in media containing serial dilutions of chloramphenicol demonstrated differences in bacterial survival only between clones carrying the reporters RPAP_cat and RPPP_cat, particularly at chloramphenicol concentrations 3.4 μ g/mL, 3.8 μ g/mL, and 4.3 μ g/mL (Figure S1B,C). However, no significant differences between the clones carrying the reporters RPAP_cat and TPPP_cat could be detected at any of the chloramphenicol concentrations tested (Figure S1C).

We hypothesized that reducing *cat* expression would enhance the screening sensitivity by making the translational efficiency of RPAP and TPPP motifs more distinguishable. As previously described,^{43,44} weak translational coupling of a fluorophore gene upstream of a resistance gene cassette can support attenuation of antibiotic resistance and enable fine-tuning of screenings. The translational coupling sequence has been reported to form mRNA secondary structures that affect downstream gene translation.⁴³ As the ribosome reaches the termination codon of the first gene, it disrupts these structures, facilitating translation of the second gene. Key factors, including the properties of the Shine–Dalgarno sequence and the choice of the start codon, determine the efficiency of downstream translation.⁴³ Accordingly, we constructed a dual-reporter plasmid encoding for super folded green fluorescent protein (sfGFP) and CAT with a weak translational coupling device⁴³ connecting the genes (Figure S1A, middle and right panels). We selected the weak coupling 1 device⁴³ because it minimizes the use of antibiotics during selection⁴³ and has been shown to be effective in previous studies.⁴⁴ Of all

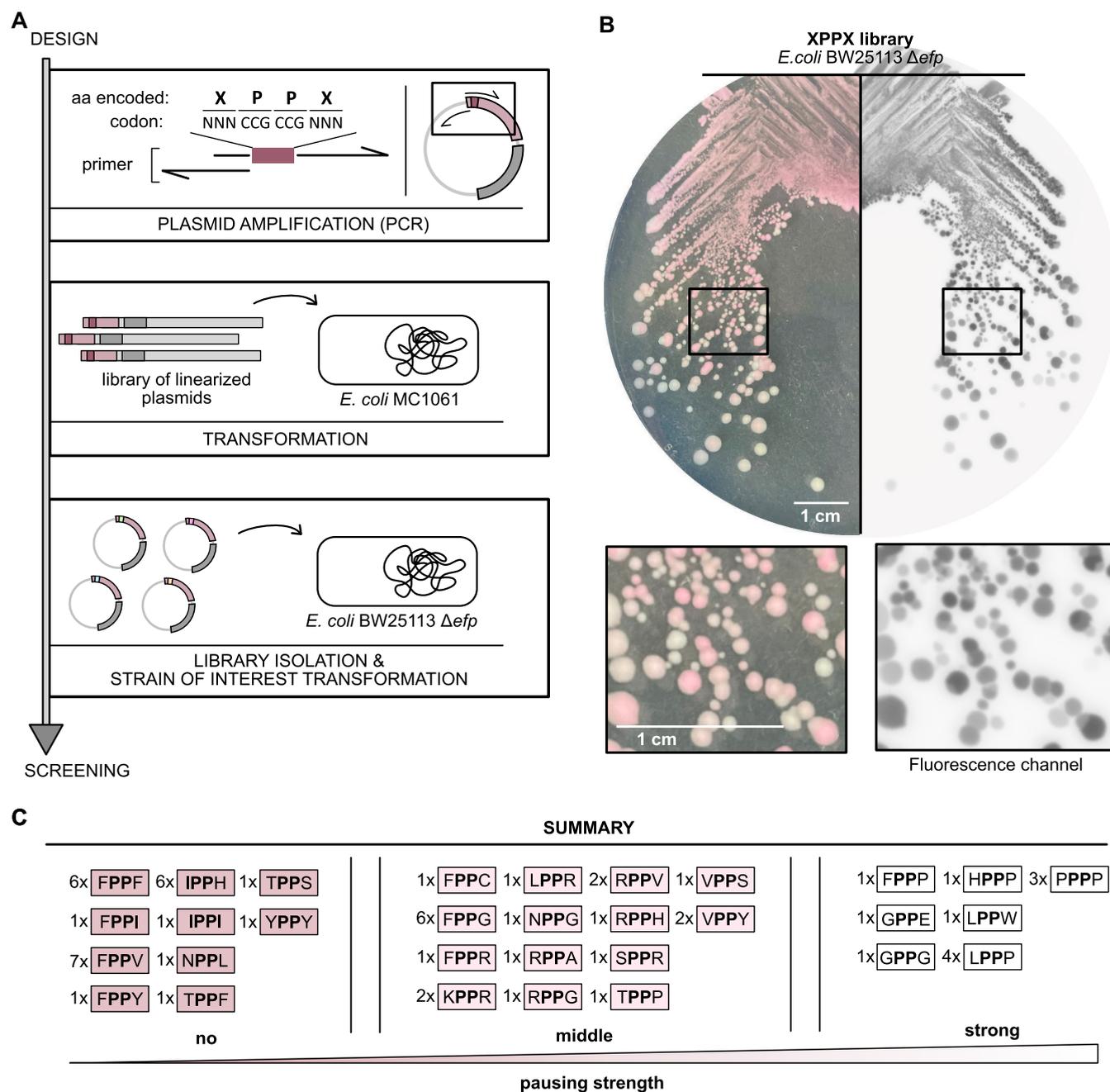


Figure 2. XPPX library screening on agar assay plates distinguish different pausing strengths. (A) Graphical workflow to generate the XPPX (X-Pro-Pro-X) library (X represents all possible amino acids; N, all possible nucleotides). (B) Colony morphology and fluorescence analysis of *E. coli* BW25113 Δ *efp*, transformed with dual-reporter plasmid pBAD24_XPPX_*mScarlet-1_cat* encoding the XPPX library. The images in the rectangles are enlarged under the agar plate. (C) Summary of the sequencing results of 60 clones that were picked according to their fluorescence intensity and viability. Motifs identified in pink colonies with high fluorescence intensities and large colony diameters were classified as those causing no pausing strength. Motifs observed in light pink colonies with intermediate fluorescence intensities and intermediate colony diameter were categorized as those causing middle pausing strength. Motifs identified in white-yellow colonies with low fluorescence intensities and small colony diameters were grouped as those causing strong pausing strength. The numbers indicate the count of plasmids found, which contain the indicated motif. aa—amino acid.

antibiotic resistance markers described and analyzed by Renning et al. in their dual-reporter system,⁴³ we retained *cat* as part of our reporter system, as chloramphenicol resistance emerged as the most promising option based on our preliminary tests (data not shown). Motifs were incorporated into the gene encoding the fluorophore (Figure S1A, middle and right panels). End point optical density quantification revealed that chloramphenicol influenced the

survival of all clones tested with the dual-reporter motif *sfGFP_cat*, regardless of the specific motif present (Figure S1D). Significant differences between all clones in media supplemented with 3.4 or 3.8 μ g/mL chloramphenicol were only detectable when motifs were at the 16th nucleotide position of the fluorophore sequence (Figure S1E). This observation confirms the essentiality of the first 15 nucleotides,

which are critical for the initial synthesis of the nascent peptide chain.⁴⁵

To increase visibility on agar plates, we replaced *sfgfp* with *mScarlet-I*, since the fluorophore mScarlet-I changes the colony color on agar plates to bright pink, as shown in earlier studies⁴⁶ (Figure 1B and Figure S2). End point quantification of optical densities demonstrated that all Δ *efp* cells containing reporters encoding RPAP, TPPP, and RPPP motifs exhibited differences in survival at 3.4, 3.8, and 4.3 μ g/mL chloramphenicol concentrations (Figure 1C,D). In particular, cells carrying reporter RPPP_ *mScarlet-I_cat* struggled to survive at chloramphenicol concentrations of 3.4 μ g/mL or higher (Figure 1C,D). These findings imply that the translational efficiency of different motifs can be measured with the dual-reporter system based on bacterial survival. Fluorescence intensity measurements after 24 h (Figure 1E, upper panel) were in line with the observations from the bacterial survival assay (Figure 1C,D). We also extended the incubation period to 48 h to ensure an equal optical density of bacterial culture in all wells, thereby demonstrating that the low fluorescence intensities observed with the RPPP_ *mScarlet-I_cat* (Figure 1E, upper panel) reporter were not caused by differential cell abundance (Figure 1E, bottom panel).

We monitored relative fluorescence intensities for 35 h in M9 minimal medium to investigate translational efficiencies over time (Figure 1F). The Δ *efp* strain with the reporter RPAP_ *mScarlet-I_cat* exhibited the greatest relative fluorescence (RF) intensities, reaching a range of 200–300 RF between 25 and 35 h (Figure 1F). The initial 10 h of monitoring revealed the impact of the supplemented chloramphenicol concentrations on the relative fluorescence (Figure 1F). The Δ *efp* strain containing the reporter TPPP_ *mScarlet-I_cat* exhibited intermediate fluorescence intensities, reaching a range of 80–110 RFU between 25 and 35 h (Figure 1F). Here, the impact of the supplemented chloramphenicol concentrations on the relative fluorescence curves was detected at the time interval of 4–14 h (Figure 1F). The low relative fluorescence intensities observed with strains containing the reporter plasmid RPPP_ *mScarlet-I_cat* were consistent with the end point fluorescence intensity measurements in Figure 1E. Regardless of the motif present, fluorescence was consistently low in the absence of arabinose and chloramphenicol, indicating low levels of leaky expression (Figure 1F).

To assess the reporter's suitability for plate-based screens, we examined the growth of the Δ *efp* strain carrying reporters with RPAP, TPPP, or RPPP motifs on LB agar plates (Figure 1G). Preliminary analysis showed that LB agar plates with 3.4 μ g/mL chloramphenicol exhibited the most significant differences in colony morphology and fluorescence between cells containing reporters with either the RPAP or RPPP motifs (Figure S3A–C).

After a 120 h incubation period at room temperature, the colonies of Δ *efp* with RPAP_ *mScarlet-I_cat* exhibited pink color and the largest colony diameter (Figure 1G). In contrast, Δ *efp* colonies expressing RPPP_ *mScarlet-I_cat* exhibited a yellow-white color and smaller colony diameters. In line with our measurements, Δ *efp* with TPPP_ *mScarlet-I_cat* exhibited a light pink color and an intermediate colony diameter (Figure 1G). When EF-P and its modification machinery were overproduced in the Δ *efp* strain, translation of the fluorophore was rescued for all of the motifs (Figure 1H). This finding suggests that the observed reduction in fluorescence in the

RPPP and TPPP reporters in Δ *efp* can be attributed to ribosome pausing events caused by the motifs in the absence of EF-P.

Taken together, our results demonstrate that the strength of ribosome pausing at amino acid motifs can be effectively measured using a dual-reporter system. Overall, these data show that mScarlet-I is an appropriate choice for the dual-reporter system, enabling sensitive colony color-contrast- and fluorescence-based screenings on plates to assess ribosome pausing strengths at specific motifs.

Screening of Libraries with Diprolyl Motifs Demonstrates Variation in Translational Efficiencies. We used plate-based screening of a diprolyl library (XPPX, X-Pro-Pro-X) in Δ *efp* to assess the efficacy of our dual reporter for large-scale translational efficiency measurements. We used a three-step protocol to generate the XPPX library, as illustrated in detail in Figure 2A. First, the dual-reporter plasmid was amplified by using primers with complementary overhangs. The forward primer was designed to be degenerate, incorporating all possible nucleotide (N) combinations before and after the two CCG codons coding for prolines (NNN-CCG-CCG-NNN) in the motif (Figure 2A). This strategy enables the incorporation of any amino acid in the positions upstream and downstream of the two prolines in the motif XPPX. Second, the *E. coli* MC1061 strain was transformed with the linearized XPPX library plasmids, allowing the library to circularize.^{43,44,47} Third, the circularized XPPX library was isolated, and the Δ *efp* mutant was transformed with the library (Figure 2A). The bacterial library populations exhibited diverse colony morphologies after plating on the assay plates (Figure 2B). Qualitatively, we found that the colonies with the largest diameter had a deep pink color, while the smaller colonies tended to be lighter pink or white-yellow (Figure 2B). The fluorescence analysis confirmed our previous observation that increasing levels of pink pigment resulted in a higher fluorescence (Figures 1G and 2B). These observations confirm that the amino acids which surround the two prolines in a diprolyl motif determine its translational efficiency, in line with published studies.^{16,21,39}

As a first pass, we manually screened 60 clones by grouping them into three categories based on colony color (pink, light pink, and white-yellow colonies) and determined their motifs by Sanger sequencing (Figure 2C). We found some motifs overlapped in adjacent categories (for example, between pink and light pink colonies) but found no overlaps between the pink and white-yellow color categories. Overlapping motifs were recategorized into the category associated with weaker effects on translational efficiency. Altogether, we identified 26 sequences that had a marginal impact on the translational efficiency of the fluorophore (colonies with pink color), 22 sequences that caused moderate decreases in translational efficiency (colonies with light pink color), and 12 sequences that caused a strong decrease in translational efficiency (colonies with white-yellow color) (Figure 2C). Interestingly, motifs encoding FPPX were found across all categories of translational efficiency (X indicates any amino acid). When encoded in the FPPX motif, phenylalanine, isoleucine, valine, and tyrosine (FPPF, FPPI, FPPV, and FPPY) had a marginal influence on translational efficiency of the fluorophore. Cysteine, glycine, and arginine (FPPC, FPPG, and FPPR) caused a moderate decrease in translational efficiency, and proline (FPPP) had the strongest effect. These results indicate that the amino acid downstream of the two prolines in a

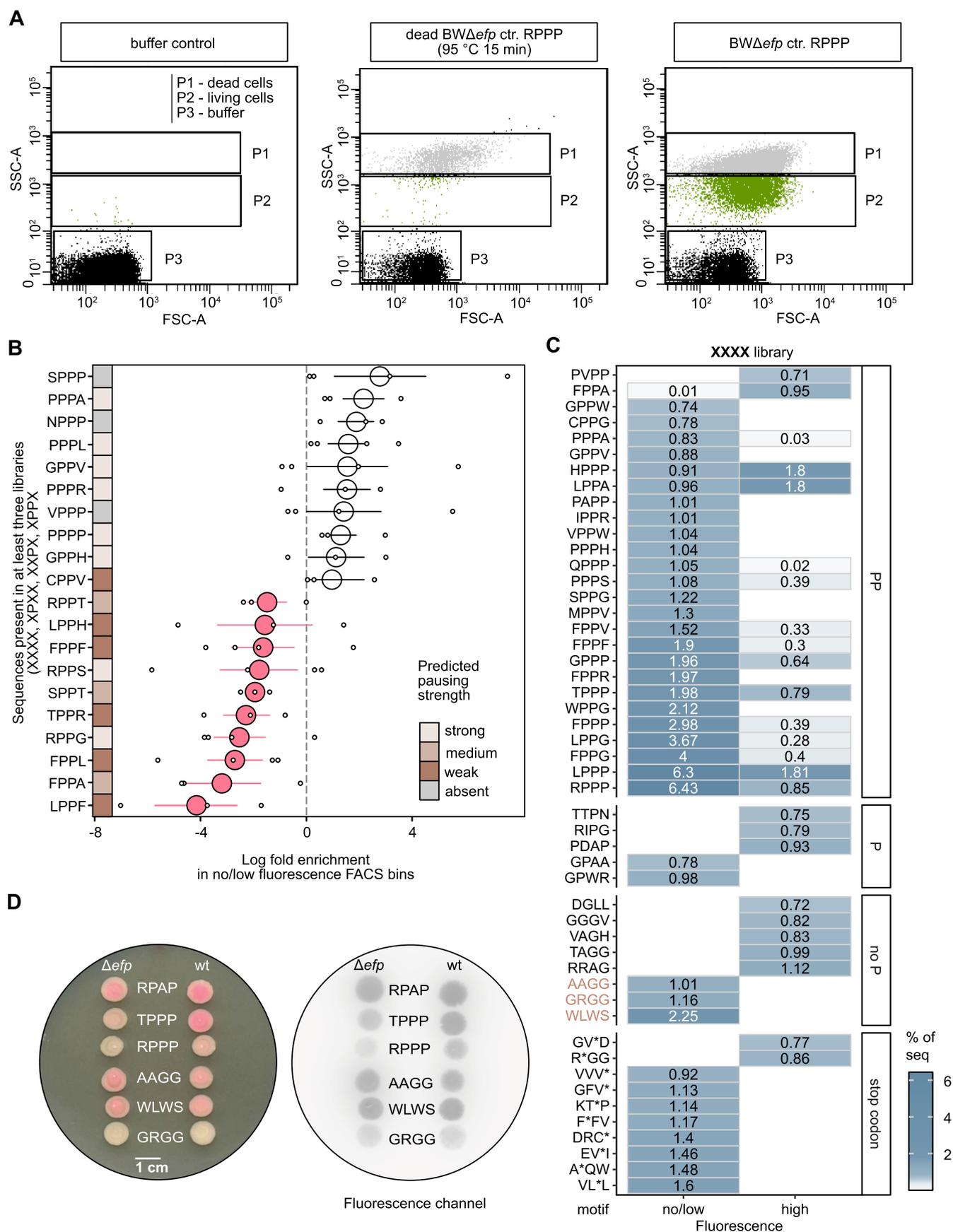


Figure 3. High-throughput screening of four motif libraries (XPXX, XPPX, XPPP, and XXXX). (A) Gating strategy of living cells was assessed by flow cytometry. Living cells (P2) were sorted from the recorded buffer events (P3) and dead cells (P1), considering the forward scatter area (FSC-A) and side scatter area (SSC-A). (B) Summary of the motifs found in at least three out of four libraries (X-X-X-X; X-Pro-X-X; X-X-Pro-X; X-Pro-

Figure 3. continued

Pro-X) with the top ten highest and top ten lowest log fold enrichment. Libraries were expressed in *E. coli* BW25113 Δ *efp*. The log fold enrichment was calculated as the natural logarithm of the sequence count found in the no/low fluorescence FACS bin divided by the sequence count found in the high fluorescence FACS bin. Small circles—values in each library, big circles—mean values, lines—standard error of the mean (SEM). Big circles were colored according to the calculated log fold enrichment value: <0, in pink; >0, in white. Pausing strength predictions were taken from Qi et al., 2018.²¹ (C) The 50 most frequent amino acid motifs found in the XXXX motif library with the largest count differences between no/low fluorescence and high fluorescence FACS bins. The motifs were divided into categories of motifs containing at least two prolines (PP), one proline (P), no proline (no P), and a stop codon (stop codon, *). X represents all possible amino acids. (D) Translational efficiency measurements in Δ *efp* and wild-type (wt) cells carrying dual-reporter plasmids with various motifs, analyzed using a spot assay. Motifs analyzed: RPAP (Arg-Pro-Ala-Pro), TPPP (Thr-Pro-Pro-Pro), RPPP (Arg-Pro-Pro-Pro), AAGG (Ala-Ala-Gly-Gly), WLWS (Trp-Leu-Trp-Ser), and GRGG (Gly-Arg-Gly-Gly). The right panel shows the fluorescence analysis of the spot assay plate.

diproyl motif critically determines translational efficiency, consistent with previous findings.²¹

Together, these results demonstrate that our dual reporter can differentiate the pausing strength of the diproyl motifs in a Δ *efp* strain. We show that the strength of ribosome pausing can be assessed on agar plates supplemented with chloramphenicol by analyzing the viability and the fluorescence of the colonies.

High-Throughput Screening Uncovers New Motifs That Reduce the Translational Efficiencies. To assess the reporter's capability for high-throughput translational efficiency measurements, we expanded our investigations using flow cytometry. To enhance the diversity of motifs that we could analyze, we designed additional libraries with either a fixed CCG codon coding for a single proline (XPXX, NNN-CCG-NNN-NNN; XXPX, NNN-NNN-CCG-NNN) or completely randomized codons (XXXX, NNN-NNN-NNN-NNN). We sorted cells containing these plasmid libraries (about one million cells per library) with fluorescence-associated cell sorting (FACS) and analyzed the sorted cells with Illumina high-throughput sequencing. In a preanalysis, samples were treated with high temperatures (15 min at 95 °C, Figure 3A) to distinguish viable cells (P2) from nonviable cells (P1) with flow cytometry. As a first pass at identifying sequences that influence translational efficiencies, we calculated the log fold enrichment of particular motifs in the no/low fluorescence FACS bin versus the high fluorescence bin across multiple independent libraries (Figure 3B). We found that while generally the same motifs had similar values across different libraries, there was some variation (e.g., while RPPG was depleted in the no/low fluorescence bin in the XXXX, XPXX, and XXPX libraries, it was slightly enriched in the XPPX library, Figure 3B). Furthermore, we also detected variation in mScarlet-I expression by microscopy among cells carrying the same reporter plasmid (Figure S4A–C). Thus, the appearance of motifs in both FACS bins could reflect heterogeneity in plasmid-based fluorophore expression, a phenomenon also reported in earlier studies^{48,49} (Figure S4D). Ultimately, analyzing the relationship between both bins provides the most precise measurement of translational efficiency. To validate our results, we compared the top 20 motifs with the lowest and the highest translational efficiencies with predicted ribosome pausing strengths associated with distinct motifs from earlier studies²¹ (Figure 3B). We found good agreement between predicted pausing strengths and our results, indicating that our system is generally suitable for assessing translational efficiencies, despite variability in plasmid copy numbers within the bacterial population (Figure S4D). Furthermore, while our analysis primarily focused on translational efficiency at the amino acid level, it is likely that regulatory mechanisms at the

codon level also play a role, as described in previous literature.⁴¹ For example, several unique codon combinations were present for the amino acid motif RPPP in one of our libraries (Figure S5). Differences in translational efficiency of different codon combinations that code for the same amino acid motif could also cause broad distributions of certain motifs across two distinct bins (Figure 3B).

A number of studies have reported on ribosomal pausing, which occurs at motifs containing consecutive prolines that can be rescued by EF-P.^{14–16,21,41} So far, only few studies report about motifs containing no prolines, which depend on EF-P for efficient translation.^{37,42} Consequently, we wanted to see whether we could detect any nonproline motifs within the XXXX library that reduced the translational efficiency of the fluorophore. We extracted the top 50 sequences with the largest difference in sequence counts between FACS bins and found four groups of motifs: those containing at least two prolines (PP), those containing one proline (P), those containing no proline (no P), and those containing a stop codon (Figure 3C). In accordance with our expectations, motifs containing a stop codon were highly abundant in the no/low fluorescence FACS bin (Figure 3C). In our system, stop codons cause the rapid termination of fluorophore translation, leading to the production of a nonfunctional protein lacking fluorescence. We found that certain motifs within the PP section appeared in both FACS bins but with differing abundances, which may be due to the heterogeneity phenomenon (Figure S4D) and/or codon effects (Figure S5). However, the majority of motifs in the PP section appeared in the no/low fluorescence FACS bins, consistent with previous studies indicating that the presence of at least two prolines hinders translation.^{16,21}

Interestingly, our experimental system identified two motifs containing a single proline (GPAA, GPWR) that decreased the efficiency of fluorophore translation. Additionally, three motifs without a proline (AAGG, GRGG, WLWS) were identified (Figure 3C). To verify these results, we transformed the Δ *efp* and wild-type strain with newly generated dual-reporter plasmids containing the motifs AAGG, GRGG, or WLWS and spotted cells on agar assay plates. We found that these motifs negatively impact the translational efficiency in the Δ *efp* mutant (Figure 3D), which is consistent with the observations shown in Figure 3C. The spot assay confirmed that the AAGG and WLWS motifs cause an intermediate reduction in the translational efficiency of the fluorophore, while the GRGG motif causes a strong reduction in translation efficiency in the Δ *efp* mutant (Figure 3D). When EF-P and its modification machinery were present (wt), translation of the fluorophore could be rescued for control motifs TPPP and RPPP (Figure 3D). However, no rescue of fluorophore translation could be

detected for the motifs AAGG, GRGG, and WLWS (Figure 3D), suggesting that additional factors influence the translation efficiency at these motifs.

Overall, we find that translational regulation appears to be more complex than previously thought. Consequently, our system offers a suitable platform to gain more insights into this phenomenon in any laboratory.

DISCUSSION

Translational coupling is a naturally occurring event in which the movement of ribosomes translating upstream genes also directs the translation of downstream genes. This mechanism ensures the coordinated and sequential synthesis of proteins involved in shared biochemical pathways or protein complexes, thereby ensuring efficient cellular function and effective resource allocation.^{50–54} Translational coupling has been used in experimental studies to identify optimal translation initiation regions (TIRs) and enhance the synthesis of distinct proteins.^{43,44,55} In this study, we used translational coupling in the design of a dual reporter to investigate translational regulation. This system can be effectively adapted for use on agar assay plates, providing a dual output by selecting for fluorescence and viability and allowing uncomplicated screening.

As a component of the dual-reporter system, mScarlet-I offers several advantages in plate-based translational efficiency screenings. Translational coupling of mScarlet-I to the resistance gene cassette (*cat*) resulted in the attenuation of *cat*, which improved reporter sensitivity (Figure 1). The presence of *cat* alone allowed only coarse distinctions in translational efficiencies—only motifs that are known to cause strong ribosome pausing and motifs that cause no pausing at all could be distinguished (Figure S1A–C). This suggests that when present alone in the reporter, CAT “neutralizes” chloramphenicol so effectively that small differences in ribosome pausing become indistinguishable. In contrast, the incorporation of mScarlet-I into the reporter allowed for finer discrimination between pausing strengths caused by different motifs (Figure 1). In addition, the bright fluorescence, good color contrast with the plate background (Figures 1G and 2B), and ability to emit light in the visible red spectrum (FPbase ID: 6VVTk)^{56,57} make mScarlet-I a perfect component of the reporter so that measurements could be performed without additional equipment. Overall, the generation of a dual reporter composed of a fluorescent gene (*mScarlet-I*) and a resistance gene cassette (*cat*) offers the advantage to measure two independent outputs, fluorescence intensity and cell viability (Figures 1 and 2). This configuration also provides tunability, as researchers can adjust antibiotic concentrations, thereby increasing the stringency of the fluorescence analysis (Figures 1D, and S1C–E).

Previous studies have demonstrated that in addition to consecutive prolines, sequences upstream and downstream of prolines can also influence translational efficiency in *E. coli*.^{16,38} We were able to confirm and reproduce these observations, particularly in the context of downstream sequences (FPPX) in Δefp (Figure 2). Motifs with no or intermediate influence on translational efficiency (FPPF, FPPI, FPPV, and FPPY) and high influence on translational efficiency (FPPP) (Figure 2C) were consistent with previously published proteome and ribosome profiling data.^{16,39} However, we do not exclude the possibility that sequences further away from the motif, as well as the codon context of the motif, might also influence

translational efficiency. This should be investigated in future studies.

Our investigations of motifs within the XXXX libraries, conducted using flow cytometry and high-throughput screening, identified new motifs containing one (GPAA, GPWR) or no prolines (AAGG, GRGG, and WLWS) that affect translational efficiency (Figure 3C,D). Two of the motifs we identified were found to contain more than one glycine. Previous studies have demonstrated that sequences rich in glycine have the potential to cause ribosome pausing.^{37,58} mRNA sequences that are similar to the Shine–Dalgarno (SD) sequence can trigger a SD–aSD (Shine–Dalgarno–anti-Shine–Dalgarno) interaction, which in turn causes ribosome pausing.^{5,59,60} Nevertheless, it is still unclear how glycine codons affect translational pausing,⁶¹ as some studies suggest that internal SD sequences have no/low effect on translational speed.^{62,63} The present study provides a platform that can support future research aimed at clarifying the influence of glycine on translational pausing in different sequence contexts.

In addition to EF-P, the ATP-Binding Cassette family-F (ABCF) ATPase proteins, like YfmR and Uup, as well as the RNA binding protein YebC, have recently been shown to be involved in the translational rescue of polyproline motifs.^{64–67} The reporter described here could easily be adapted to study the full spectrum of motifs rescued by these systems in any laboratory. Additionally, while the spectrum of motifs that cause EF-P-dependent ribosome pausing has been studied extensively, these experiments have primarily focused on *E. coli* and its β -lysylated post-translationally modified EF-P.^{16,21,39} Our recent study showed differences in motif rescue preferences between wild-type *E. coli* and a mutant strain in which the native EF-P was replaced with an unmodified PGKGP subfamily EF-P.³⁶ This suggests that there is a subset of motifs that are preferentially rescued by different EF-P types. If so, different EF-P types could drive the evolution of bacterial genomes toward motifs that they preferentially rescue, particularly in highly expressed proteins where maximizing the translation rate is essential.^{21,68} Such motif bias spectra could be studied with our reporter system.

MATERIALS AND METHODS

Bacterial Growth. All bacterial strains used in this study were cultivated in lysogenic broth (LB) under agitation at 37 °C. If required, growth media was solidified by the addition of 1.5% (w/v) agar. The following antibiotics were used in this study: carbenicillin sodium salt (100 μ g/mL) and chloramphenicol (default concentration 34 μ g/mL; reporter assays: 1.7, 2.3, 3.4, and 3.8 μ g/mL).

Molecular Cloning. All strains, plasmids, and DNA oligonucleotides used are listed in Data S1. Kits, enzymes, and polymerases were used according to the manufacturer's instructions. Q5 High-Fidelity DNA Polymerase (New England BioLab) and OneTaq DNA Polymerase were used for PCR amplification. DNA oligonucleotides were ordered from Merck. Amplified genes coding for the fluorophores (*sfGFP* and *mScarlet-I*) and the antibiotic resistance cassette (*chloramphenicol acetyltransferase*, CAT) were linked together with a hairpin loop (weak coupling)⁴³ by overlap PCR. DNA was isolated and purified with a High-Yield PCR Cleanup and Gel Extraction Kit (Sued Laborbedarf Gauting). All DNA restriction digests were carried out using restriction enzymes in rCutSmart buffer (New England BioLab). All ligations were performed with the T4 DNA ligase (New England BioLab) at

25 °C for 2 h. Plasmids were isolated with the HiYield Plasmid Mini Kit (Sued Laborbedarf Gauting) and stored at −20 °C. Plasmids were verified by colony PCR and Sanger sequencing. All nucleotide sequences were analyzed with CLC Main Workbench version 8.1.2 (Qiagen).

Growth and Fluorescence Measurements in 96-Well Plates. To determine the optimal concentration of chloramphenicol for the pausing strength measurements with the dual reporter, serial dilutions were tested. Plasmids pBAD24₋motif_{-sfgfp_cat} and pBAD24₋motif_{-mscarlet-I_cat} (motif: RPAP, TPPP, or RPPP) were independently transformed into *E. coli* BW25113 Δ efp and incubated overnight at 37 °C on LB agar plates supplemented with 100 μ g/mL carbenicillin sodium salt. Single colonies were inoculated into fresh LB supplemented with 100 μ g/mL carbenicillin sodium salt and incubated overnight at 37 °C under constant shaking (180 rpm). The overnight cultures were reinoculated into fresh LB supplemented with 100 μ g/mL carbenicillin sodium salt and grown until reaching an absorption of 0.3 at 600 nm (OD₆₀₀). Each culture was transferred to a final OD₆₀₀ of 0.01 into a 96-well plate containing 0.2% (w/v) arabinose and different concentrations of chloramphenicol (1.7, 2.3, 3.4, 3.8 μ g/mL) with a final volume of 200 μ L/well. Plates were incubated overnight at 37 °C with constant shaking (150 rpm). End point OD₆₀₀ was determined with Tecan Infinite 200 Pro (number of flashes: 25). Fluorescence was measured with a GE Typhoon Trio Imager (laser: green, 532 nm; emission filter: 580 BP 30 Cy3; photomultiplier tube, PMT: 450).

For all time-course fluorescence and growth measurements, the cells were grown for 20 h in M9 minimal media⁶⁹ (composition: 33.7 mM Na₂HPO₄, 22 mM KH₂PO₄, 8.55 mM NaCl, 9.35 mM NH₄Cl, 1 mM MgSO₄, 0.3 mM CaCl₂, 1 μ g biotin, 1 μ g thiamin, trace elements, 0.4% [w/v] glucose) prior to the start of the measurements in the plate reader. Growth and fluorescence were monitored in 10 min intervals for 35 h at 37 °C with Tecan Infinite 200 Pro (excitation wavelength: 565 nm; emission wavelength: 594 nm; gain: 50; number of flashes: 25; shaking between measurements: 180 rpm, orbital).

EF-P complementation assay plates were incubated for 24 to 34 h until fluorescence measurements.

Generation of XPPX, PXXX, XPXX, XPPX, XXXP, and XXXX Libraries. The forward primers were designed to contain 6–12 degenerated nucleotides that code for all possible codons. The fixed prolines in the motif were coded by the CCG codon. Degenerated codons were designed to appear 15 nucleotides downstream of the start codon of the fluorophore gene. Libraries were amplified using the degenerated forward primer and reverse primer with complementary regions to the forward primers to enable plasmid circularization afterward (Figure 2A and Figure S6). The PCR program used for library amplification was as follows: initial denaturation, 98 °C for 30 s (sec); amplification (35 cycles), 98 °C for 10 s, followed by 57 °C for 30 s and 72 °C for 180 s; final extension, 72 °C for 120 s. 40 μ L of the PCR product was treated with DpnI. Chemically competent *E. coli* MC1061 was transformed with the digested and purified PCR products to enable plasmid circularization by homologous recombination.^{43,44,47} The transformants were transferred to a flask with 15 mL of LB, supplemented with 100 μ g/mL carbenicillin sodium salt, and incubated overnight at 37 °C under constant shaking (180 rpm). The libraries were isolated with the HiYield Plasmid Mini Kit (Sued Laborbedarf Gauting).

Screening of Libraries on Assay Plates. The final libraries were prepared according to the protocol described by Rennig et al.⁴³ and Shilling et al.,⁴⁴ with certain modifications. Chemically competent *E. coli* BW25113 Δ efp was transformed with 500 ng of the library with a 1 h recovery phase in fresh LB without antibiotics. The transformants were inoculated into 3 mL of fresh LB containing 100 μ g/mL carbenicillin sodium salt and incubated overnight at 37 °C under constant shaking (180 rpm). 30 μ L of this overnight culture was reinoculated into 3 mL of fresh LB (containing 100 μ g/mL carbenicillin sodium salt) and incubated for another round overnight at 37 °C under constant shaking (180 rpm). A 100 μ L portion of this overnight culture was inoculated into fresh 5 mL of LB supplemented with 100 μ g/mL carbenicillin sodium salt and grown until reaching an OD₆₀₀ of 0.3. 10³–10⁵ cells were plated on assay agar plates, supplemented with 3.4 μ g/mL chloramphenicol and 0.2% (w/v) arabinose. Plates were incubated for 16 h at 30 °C and kept afterward for ~120 h at 25 °C. Fluorescence was measured with a GE Typhoon Trio Imager (laser: green, 532 nm; emission filter: 580 BP 30 Cy3; pmt: 450).

Microscopy. To analyze the effect of different motifs on the production of mScarlet-I, chemically competent *E. coli* BW25113 Δ efp was transformed with the plasmids pBAD24₋motif_{-mscarlet-I_cat} (motif: RPAP, TPPP or RPPP). Cells carrying the plasmids were cultivated overnight at 37 °C under constant shaking (180 rpm) in LB supplemented with 100 μ g/mL carbenicillin sodium salt. The overnight cultures were inoculated into fresh LB supplemented with 100 μ g/mL carbenicillin sodium salt and grown at 37 °C under constant shaking (180 rpm) until reaching an OD₆₀₀ of 0.15. 0.2% (w/v) arabinose was then added to the cultures, and incubation was continued for 24 h. All cells were washed in phosphate-buffered saline (PBS; 93.6 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄·2H₂O, 2 mM KH₂PO₄), and 2 μ L of a culture with an OD₆₀₀ of 0.5 was spotted on a 1% (w/v) agarose pad (in PBS) and covered with a coverslip. Microscopic pictures were taken using a Leica DMI8 inverted microscope equipped with a Leica DFC365 FX camera (Wetzlar, Germany). An excitation wavelength of 546 nm and a 605 nm emission filter with a 75 nm bandwidth were used for mScarlet-I fluorescence with an exposure of 20 ms, gain 1, and 100% intensity. To quantify the relative fluorescent intensities (RF) of single cells, phase contrast and fluorescent images were analyzed using the ImageJ⁷⁰ plugin MicrobeJ.⁷¹ Default settings of MicrobeJ were used for cell segmentation (fit shape, rod-shaped bacteria) apart from the following settings: area: 0.1-max μ m²; length: 1.2–5 μ m; width: 0.1–1 μ m; curvature 0.0–0.15 and angularity 0.0–0.25 for *E. coli* cells. In total, 516 cells were quantified per strain. The background of the agarose pad was subtracted from each cell per field of view.

Flow Cytometry and FACS. The influence of different motifs on the translational efficiency of the fluorophore mScarlet-I was analyzed by using high-throughput screening with flow cytometry and fluorescence-associated cell sorting (FACS). Cells containing the dual-reporter plasmid were grown in LB supplemented with 100 μ g/mL carbenicillin sodium salt and 0.2% (w/v) arabinose for 16 h at 37 °C under constant shaking (180 rpm). The overnight incubation was necessary to provide enough time for mScarlet-I to mature before FACS analysis. Cells were washed in PBS and adjusted to an OD of 0.1 (10⁸ cells) in 10 mL PBS (4 °C) prior to the measurements. For the dead cell control, the culture was boiled

at 95 °C for 15 min. The gating strategy for living cells is depicted in Figure 3A. Flow cytometry measurements and FACS were performed using FACSAriaFusionII Cell Sorter (BD Biosciences) with the following settings: 438 V(FSC), 277 V (SSC), and 490 V (PE-Texas Red). Red fluorescence intensities were quantified using PE-Texas Red and displayed on a standard logarithmic scale. Flow cytometry data was processed with FACSDiva Software (BD Biosciences). Motifs found in cells with no/low fluorescence (intensity below $10^{2.1}$ on the logarithmic scale) were classified as those capable of slowing down the translational efficiency of the fluorophore. In contrast, motifs found in cells with high fluorescence (intensity $>1 \times 10^{3.1}$ on the logarithmic scale) were classified as those having only a marginal influence on the translational efficiency of the fluorophore.

Library Preparation for Illumina High-Throughput Sequencing. Plasmids from each FACS bin were isolated with the HiYield Plasmid Mini Kit (Sued Laborbedarf) and served as template DNA for Illumina-specific amplification required for sequencing.⁷² Custom-made DNA oligos were designed (Merck), consisting of 21 bp Illumina primer sequence, distinct barcoding tag,⁷² and a sequence complementary to the dual-reporter plasmid (forward primer: 81 bp upstream of the library motif, reverse primer: 167 bp downstream of the library motif; Figure S6). Purified PCR products were loaded on a 1% (w/v) agarose gel to estimate the amount of each amplicon and pooled in equimolar concentrations.⁷² The library was finalized for sequencing according to the manufacturer's protocols (http://supportres.illumina.com/documents/documentation/system_documentation/miseq/preparing-libraries-for-sequencing-on-miseq-15039740-d.pdf). Paired-end sequencing of 2×300 bp (base pairs) with two additional 8 bp index reads was performed on an Illumina MiSeq platform (Illumina Inc.) with v3 chemistry. All Sequencing data generated in this study are deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB77738.

Bioinformatic Analysis of Illumina Library Sequencing. Sequences were first quality filtered using FASTQ version 0.23.4.⁷³ Sequences were trimmed using a dynamic approach: if the mean quality of bases within a sliding window of 4 nucleotides dropped below a cutoff of 20, sequences were trimmed starting at the first base within the window (--cut_right). Paired reads were corrected using the overlapping region (--correction). Next, PANDAseq version 2.11 was used with default parameters to merge paired-end sequences.⁷⁴ SABRE version 1.0 was used to demultiplex sequences,⁷⁵ and custom Python scripts were used to extract the sequence region containing the library. The R package Vegan version 2.6 was used to rarefy sequences to the level of the sample with the lowest sequence count (this level varied depending on the samples being compared but was never lower than 98335 sequences).⁷⁶ All statistical analyses on the sequence data were done in R version 4.3.2,⁷⁷ and all related figures were created with ggplot2 version 3.4.4.⁷⁸

Statistical Analysis and Reproducibility. All measurements are from at least three biological replicates except the plate-based library screening (Figure 2C, the count of found motifs is indicated in the figure) and FACS. Data comparisons were done using the student's unpaired two-sided *t* test (Microsoft Office Excel 2021) and are described in each figure legend. Error bars in bar graphs represent the standard

deviation (SD). Values with a calculated *p* value below 0.05 were considered as significantly different.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acssynbio.4c00534>.

Bacterial survival-based determination of translational efficiencies by using a heterologous dual-reporter system; annotated sequence of the pBAD24_motif_mscarlet-1_weak_coupling1_cat expression cassette; the influence of the chloramphenicol concentration and the motif on the survival and fluorescence of the reporter strain; heterogeneity in plasmid-based fluorophore expression; distribution of codons encoding the motif RPPP in the library XXXX; primer design for Illumina library sequencing (PDF)

strains, plasmids; and DNA oligonucleotides (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Kirsten Jung – Faculty of Biology, Microbiology, Ludwig-Maximilians-Universität München, 82152 Planegg-Martinsried, Germany; orcid.org/0000-0003-0779-6841; Email: jung@lmu.de

Authors

Urte Tomasiunaite – Faculty of Biology, Microbiology, Ludwig-Maximilians-Universität München, 82152 Planegg-Martinsried, Germany; orcid.org/0000-0001-6729-9279

Tess Brewer – Faculty of Biology, Microbiology, Ludwig-Maximilians-Universität München, 82152 Planegg-Martinsried, Germany

Korinna Burdack – Faculty of Biology, Microbiology, Ludwig-Maximilians-Universität München, 82152 Planegg-Martinsried, Germany

Sophie Brameyer – Faculty of Biology, Microbiology, Ludwig-Maximilians-Universität München, 82152 Planegg-Martinsried, Germany; orcid.org/0000-0002-6779-2343

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acssynbio.4c00534>

Author Contributions

U.T. and K.J. designed the research; U.T. and K.B. performed the research; U.T., T.B., and S.B. analyzed the data; U.T. and K.J. wrote the manuscript; K.J. funding acquisition.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Dr. Andreas Brachmann and the Genomics Service Unit at the LMU for fruitful discussions and excellent support in the sequencing of the Illumina libraries. The authors thank the Core Facility Flow Cytometry at Biomedical Center Munich for the great technical support. The authors thank Lukas Voshagen for great technical assistance. This work was financially supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grant JU270/20-1, project number 449926427, and RTG2062 (Molecular Principles of Synthetic Biology) to K.J.

REFERENCES

- (1) Lu, J.; Kobertz, W. R.; Deutsch, C. Mapping the electrostatic potential within the ribosomal exit tunnel. *J. Mol. Biol.* **2007**, *371*, 1378–1391.
- (2) Lu, J.; Deutsch, C. Electrostatics in the ribosomal tunnel modulate chain elongation rates. *J. Mol. Biol.* **2008**, *384*, 73–86.
- (3) Kramer, G.; Boehringer, D.; Ban, N.; Bukau, B. The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat. Struct. Mol. Biol.* **2009**, *16*, 589–597.
- (4) de Smit, M. H.; van Duin, J. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 7668–7672.
- (5) Wen, J. D.; Lancaster, L.; Hodges, C.; Zeri, A. C.; Yoshimura, S. H.; Noller, H. F.; Bustamante, C.; Tinoco, I. Following translation by single ribosomes one codon at a time. *Nature* **2008**, *452*, 598–603.
- (6) Kudla, G.; Murray, A. W.; Tollervey, D.; Plotkin, J. B. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **2009**, *324*, 255–258.
- (7) Chen, C.; Zhang, H.; Broitman, S. L.; Reiche, M.; Farrell, I.; Cooperman, B. S.; Goldman, Y. E. Dynamics of translation by single ribosomes through mRNA secondary structures. *Nat. Struct. Mol. Biol.* **2013**, *20*, 582–588.
- (8) Varenne, S.; Buc, J.; Lloubes, R.; Lazdunski, C. Translation is a non-uniform process. Effect of tRNA availability on the rate of elongation of nascent polypeptide chains. *J. Mol. Biol.* **1984**, *180*, 549–576.
- (9) Nedialkova, D. D.; Leidel, S. A. Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. *Cell* **2015**, *161*, 1606–1618.
- (10) Liu, F.; Clark, W.; Luo, G.; Wang, X.; Fu, Y.; Wei, J.; Wang, X.; Hao, Z.; Dai, Q.; Zheng, G.; et al. ALKBH1-Mediated tRNA Demethylation Regulates Translation. *Cell* **2016**, *167*, 816–828.
- (11) Ikemura, T. Codon usage and tRNA content in unicellular and multicellular organisms. *Mol. Biol. Evol.* **1985**, *2*, 13–34.
- (12) Komar, A. A. The Yin and Yang of codon usage. *Hum. Mol. Genet.* **2016**, *25*, R77–R85.
- (13) Tanner, D. R.; Cariello, D. A.; Woolstenhulme, C. J.; Broadbent, M. A.; Buskirk, A. R. Genetic identification of nascent peptides that induce ribosome stalling. *J. Biol. Chem.* **2009**, *284*, 34809–34818.
- (14) Ude, S.; Lassak, J.; Starosta, A. L.; Kraxenberger, T.; Wilson, D. N.; Jung, K. Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science* **2013**, *339*, 82–85.
- (15) Doerfel, L. K.; Wohlgemuth, I.; Kothe, C.; Peske, F.; Urlaub, H.; Rodnina, M. V. EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science* **2013**, *339*, 85–88.
- (16) Peil, L.; Starosta, A. L.; Lassak, J.; Atkinson, G. C.; Virumae, K.; Spitzer, M.; Tenson, T.; Jung, K.; Remme, J.; Wilson, D. N. Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 15265–15270.
- (17) Muto, H.; Ito, K. Peptidyl-prolyl-tRNA at the ribosomal P-site reacts poorly with puromycin. *Biochem. Biophys. Res. Commun.* **2008**, *366*, 1043–1047.
- (18) Wohlgemuth, I.; Brenner, S.; Beringer, M.; Rodnina, M. V. Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. *J. Biol. Chem.* **2008**, *283*, 32229–32235.
- (19) Pavlov, M. Y.; Watts, R. E.; Tan, Z.; Cornish, V. W.; Ehrenberg, M.; Forster, A. C. Slow peptide bond formation by proline and other N-alkylamino acids in translation. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 50–54.
- (20) Doerfel, L. K.; Wohlgemuth, I.; Kubyshkin, V.; Starosta, A. L.; Wilson, D. N.; Budisa, N.; Rodnina, M. V. Entropic Contribution of Elongation Factor P to Proline Positioning at the Catalytic Center of the Ribosome. *J. Am. Chem. Soc.* **2015**, *137*, 12997–13006.
- (21) Qi, F.; Motz, M.; Jung, K.; Lassak, J.; Frishman, D. Evolutionary analysis of polyproline motifs in *Escherichia coli* reveals their regulatory role in translation. *PLoS Comput. Biol.* **2018**, *14*, No. e1005987.
- (22) Pinheiro, B.; Scheidler, C. M.; Kielkowski, P.; Schmid, M.; Forne, I.; Ye, S.; Reiling, N.; Takano, E.; Imhof, A.; Sieber, S. A.; et al. Structure and Function of an Elongation Factor P Subfamily in Actinobacteria. *Cell Rep.* **2020**, *30*, 4332–4342 e4335.
- (23) Morgan, A. A.; Rubenstein, E. Proline: the distribution, frequency, positioning, and common functional roles of proline and polyproline sequences in the human proteome. *PLoS One* **2013**, *8*, No. e53785.
- (24) Rath, A.; Davidson, A. R.; Deber, C. M. The structure of “unstructured” regions in peptides and proteins: role of the polyproline II helix in protein folding and recognition. *Biopolymers* **2005**, *80*, 179–185.
- (25) Adzhubei, A. A.; Sternberg, M. J.; Makarov, A. A. Polyproline-II helix in proteins: structure and function. *J. Mol. Biol.* **2013**, *425*, 2100–2132.
- (26) Motz, M.; Jung, K. The role of polyproline motifs in the histidine kinase EnvZ. *PLoS One* **2018**, *13*, No. e0199782.
- (27) Gutierrez, E.; Shin, B. S.; Woolstenhulme, C. J.; Kim, J. R.; Saini, P.; Buskirk, A. R.; Dever, T. E. eIF5A promotes translation of polyproline motifs. *Mol. Cell* **2013**, *51*, 35–45.
- (28) Behshad, E.; Ruzicka, F. J.; Mansoorabadi, S. O.; Chen, D.; Reed, G. H.; Frey, P. A. Enantiomeric free radicals and enzymatic control of stereochemistry in a radical mechanism: the case of lysine 2,3-aminomutases. *Biochemistry* **2006**, *45*, 12639–12646.
- (29) Yanagisawa, T.; Sumida, T.; Ishii, R.; Takemoto, C.; Yokoyama, S. A paralog of lysyl-tRNA synthetase aminoacylates a conserved lysine residue in translation elongation factor P. *Nat. Struct. Mol. Biol.* **2010**, *17*, 1136–1143.
- (30) Navarre, W. W.; Zou, S. B.; Roy, H.; Xie, J. L.; Savchenko, A.; Singer, A.; Edvokimova, E.; Prost, L. R.; Kumar, R.; Ibba, M.; Fang, F. C. PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. *Mol. Cell* **2010**, *39*, 209–221.
- (31) Peil, L.; Starosta, A. L.; Virumae, K.; Atkinson, G. C.; Tenson, T.; Remme, J.; Wilson, D. N. Lys34 of translation elongation factor EF-P is hydroxylated by YfcM. *Nat. Chem. Biol.* **2012**, *8*, 695–697.
- (32) Lassak, J.; Keilhauer, E. C.; Furst, M.; Wuichet, K.; Godeke, J.; Starosta, A. L.; Chen, J. M.; Sogaard-Andersen, L.; Rohr, J.; Wilson, D. N.; et al. Arginine-rhamnosylation as new strategy to activate translation elongation factor P. *Nat. Chem. Biol.* **2015**, *11*, 266–270.
- (33) Rajkovic, A.; Erickson, S.; Witzky, A.; Branson, O. E.; Seo, J.; Gafken, P. R.; Frietas, M. A.; Whitelegge, J. P.; Faull, K. F.; Navarre, W.; et al. Cyclic Rhamnosylated Elongation Factor P Establishes Antibiotic Resistance in *Pseudomonas aeruginosa*. *mBio* **2015**, *6*, No. e00823.
- (34) Rajkovic, A.; Hummels, K. R.; Witzky, A.; Erickson, S.; Gafken, P. R.; Whitelegge, J. P.; Faull, K. F.; Kearns, D. B.; Ibba, M. Translation Control of Swarming Proficiency in *Bacillus subtilis* by 5-Amino-pentanolylated Elongation Factor P. *J. Biol. Chem.* **2016**, *291*, 10976–10985.
- (35) Hummels, K. R.; Witzky, A.; Rajkovic, A.; Tollerson, R.; Jones, L. A.; Ibba, M.; Kearns, D. B. Carbonyl reduction by YmfI in *Bacillus subtilis* prevents accumulation of an inhibitory EF-P modification state. *Mol. Microbiol.* **2017**, *106*, 236–251.
- (36) Tomasiunaite, U.; Kielkowski, P.; Krafczyk, R.; Forne, I.; Imhof, A.; Jung, K. Decrypting the functional design of unmodified translation elongation factor P. *Cell Rep.* **2024**, *43*, No. 114063.
- (37) Hersch, S. J.; Wang, M.; Zou, S. B.; Moon, K. M.; Foster, L. J.; Ibba, M.; Navarre, W. W. Divergent protein motifs direct elongation factor P-mediated translational regulation in *Salmonella enterica* and *Escherichia coli*. *mBio* **2013**, *4*, No. e00180-13.
- (38) Starosta, A. L.; Lassak, J.; Peil, L.; Atkinson, G. C.; Virumae, K.; Tenson, T.; Remme, J.; Jung, K.; Wilson, D. N. Translational stalling at polyproline stretches is modulated by the sequence context upstream of the stall site. *Nucleic Acids Res.* **2014**, *42*, 10711–10719.
- (39) Woolstenhulme, C. J.; Guydosh, N. R.; Green, R.; Buskirk, A. R. High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. *Cell Rep.* **2015**, *11*, 13–21.

- (40) Pfab, M.; Kielkowski, P.; Krafczyk, R.; Volkwein, W.; Sieber, S. A.; Lassak, J.; Jung, K. Synthetic post-translational modifications of elongation factor P using the ligase EpmA. *FEBS J.* **2021**, *288*, 663–677.
- (41) Krafczyk, R.; Qi, F.; Sieber, A.; Mehler, J.; Jung, K.; Frishman, D.; Lassak, J. Proline codon pair selection determines ribosome pausing strength and translation efficiency in bacteria. *Commun. Biol.* **2021**, *4*, No. 589.
- (42) Sieber, A.; Parr, M.; von Ehr, J.; Dhamotharan, K.; Kielkowski, P.; Brewer, T.; Schäpers, A.; Krafczyk, R.; Qi, F.; Schlundt, A. et al. EF-P and its Paralog EfpL (YeiP) Differentially Control Translation of Proline-Containing Sequences *bioRxiv* 2024 DOI: 10.1101/2024.04.15.589488.
- (43) Rennig, M.; Martinez, V.; Mirzadeh, K.; Dunas, F.; Rojsater, B.; Daley, D. O.; Norholm, M. H. H. TARSyn: Tunable Antibiotic Resistance Devices Enabling Bacterial Synthetic Evolution and Protein Production. *ACS Synth. Biol.* **2018**, *7*, 432–442.
- (44) Shilling, P. J.; Mirzadeh, K.; Cumming, A. J.; Widesheim, M.; Kock, Z.; Daley, D. O. Improved designs for pET expression plasmids increase protein production yield in *Escherichia coli*. *Commun. Biol.* **2020**, *3*, No. 214.
- (45) Verma, M.; Choi, J.; Cottrell, K. A.; Lavagnino, Z.; Thomas, E. N.; Pavlovic-Djuranovic, S.; Szczesny, P.; Piston, D. W.; Zaher, H. S.; Puglisi, J. D.; Djuranovic, S. A short translational ramp determines the efficiency of protein synthesis. *Nat. Commun.* **2019**, *10*, No. 5774.
- (46) Kolbe, K.; Bell, A. C.; Prosser, G. A.; Assmann, M.; Yang, H. J.; Forbes, H. E.; Gallucci, S.; Mayer-Barber, K. D.; Boshoff, H. I.; Barry, C. E., III Development and Optimization of Chromosomally-Integrated Fluorescent *Mycobacterium tuberculosis* Reporter Constructs. *Front. Microbiol.* **2020**, *11*, No. 591866.
- (47) Mirzadeh, K.; Martinez, V.; Toddo, S.; Guntur, S.; Herrgard, M. J.; Elofsson, A.; Norholm, M. H.; Daley, D. O. Enhanced Protein Production in *Escherichia coli* by Optimization of Cloning Scars at the Vector-Coding Sequence Junction. *ACS Synth. Biol.* **2015**, *4*, 959–965.
- (48) Jahn, M.; Vorpahl, C.; Hubschmann, T.; Harms, H.; Müller, S. Copy number variability of expression plasmids determined by cell sorting and Droplet Digital PCR. *Microb. Cell Fact.* **2016**, *15*, No. 211.
- (49) Boy, C.; Lesage, J.; Alfenore, S.; Gorret, N.; Guillouet, S. E. Plasmid expression level heterogeneity monitoring via heterologous eGFP production at the single-cell level in *Cupriavidus necator*. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 5899–5914.
- (50) Aksoy, S.; Squires, C. L.; Squires, C. Translational coupling of the *trpB* and *trpA* genes in the *Escherichia coli* tryptophan operon. *J. Bacteriol.* **1984**, *157*, 363–367.
- (51) de Smit, M. H.; van Duin, J. Control of Prokaryotic Translational Initiation by mRNA Secondary Structure. In *Progress in Nucleic Acid Research and Molecular Biology*; Elsevier, 1990; Vol. 38, pp 1–35.
- (52) Lesage, P.; Chiaruttini, C.; Graffe, M.; Dondon, J.; Milet, M.; Springer, M. Messenger RNA secondary structure and translational coupling in the *Escherichia coli* operon encoding translation initiation factor IF3 and the ribosomal proteins, L35 and L20. *J. Mol. Biol.* **1992**, *228*, 366–386.
- (53) Rex, G.; Surin, B.; Besse, G.; Schneppe, B.; McCarthy, J. E. The mechanism of translational coupling in *Escherichia coli*. Higher order structure in the *atpHA* mRNA acts as a conformational switch regulating the access of de novo initiating ribosomes. *J. Biol. Chem.* **1994**, *269*, 18118–18127.
- (54) Chiaruttini, C.; Milet, M.; Springer, M. Translational coupling by modulation of feedback repression in the IF3 operon of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 9208–9213.
- (55) Ferro, R.; Rennig, M.; Hernandez-Rollan, C.; Daley, D. O.; Norholm, M. H. H. A synbio approach for selection of highly expressed gene variants in Gram-positive bacteria. *Microb. Cell Fact.* **2018**, *17*, No. 37.
- (56) Lambert, T. J. FPbase: a community-editable fluorescent protein database. *Nat. Methods* **2019**, *16*, 277–278.
- (57) Bindels, D. S.; Haarbosch, L.; van Weeren, L.; Postma, M.; Wiese, K. E.; Mastop, M.; Aumonier, S.; Gotthard, G.; Royant, A.; Hink, M. A.; Gadella, T. W., Jr. mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nat. Methods* **2017**, *14*, 53–56.
- (58) Woolstenhulme, C. J.; Parajuli, S.; Healey, D. W.; Valverde, D. P.; Petersen, E. N.; Starosta, A. L.; Guydosh, N. R.; Johnson, W. E.; Wilson, D. N.; Buskirk, A. R. Nascent peptides that block protein synthesis in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E878–887.
- (59) Li, G. W.; Oh, E.; Weissman, J. S. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* **2012**, *484*, 538–541.
- (60) Chevance, F. F. V.; Guyon, S. L.; Hughes, K. T. The effects of codon context on *in vivo* translation speed. *PLoS Genet.* **2014**, *10*, No. e1004392.
- (61) Rodnina, M. V. The ribosome in action: Tuning of translational efficiency and protein folding. *Protein Sci.* **2016**, *25*, 1390–1406.
- (62) Mohammad, F.; Woolstenhulme, C. J.; Green, R.; Buskirk, A. R. Clarifying the Translational Pausing Landscape in Bacteria by Ribosome Profiling. *Cell Rep.* **2016**, *14*, 686–694.
- (63) Borg, A.; Ehrenberg, M. Determinants of the rate of mRNA translocation in bacterial protein synthesis. *J. Mol. Biol.* **2015**, *427*, 1835–1847.
- (64) Hong, H. R.; Prince, C. R.; Tetreault, D. D.; Wu, L.; Feaga, H. A. YfmR is a translation factor that prevents ribosome stalling and cell death in the absence of EF-P. *Proc. Natl. Acad. Sci. U.S.A.* **2024**, *121*, No. e2314437121.
- (65) Takada, H.; Fujiwara, K.; Atkinson, G. C.; Chiba, S.; Haurlyuk, V. Resolution of ribosomal stalling by EF-P and ABCF ATPases YfmR and YkpA/YbiT. *Nucleic Acids Res.* **2024**, *52*, No. 9854.
- (66) Ignatov, D.; Shanmuganathan, V.; Ahmed-Begrich, R.; Alagesan, K.; Frese, C. K.; Wang, C.; Krause, K.; Charpentier, E. Novel RNA-Binding Protein YebC Enhances Translation of Proline-Rich Amino Acid Stretches in Bacteria *bioRxiv* 2024 DOI: 10.1101/2024.08.26.607280.
- (67) Chadani, Y.; Yamanouchi, S.; Uemura, E.; Yamasaki, K.; Niwa, T.; Ikeda, T.; Kurihara, M.; Iwasaki, W.; Taguchi, H. The ABCF proteins in *Escherichia coli* individually cope with 'hard-to-translate' nascent peptide sequences. *Nucleic Acids Res.* **2024**, *52*, 5825–5840.
- (68) Brewer, T. E.; Wagner, A. Translation stalling proline motifs are enriched in slow-growing, thermophilic, and multicellular bacteria. *ISME J.* **2022**, *16*, 1065–1073.
- (69) Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor, N.Y.); Cold Spring Harbor, NY, 1972.
- (70) Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671–675.
- (71) Ducret, A.; Quardokus, E. M.; Brun, Y. V. MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nat. Microbiol.* **2016**, *1*, No. 16077.
- (72) Unterseher, M.; Siddique, A. B.; Brachmann, A.; Persoh, D. Diversity and Composition of the Leaf Microbiome of Beech (*Fagus sylvatica*) Are Affected by Local Habitat Conditions and Leaf Biochemistry. *PLoS One* **2016**, *11*, No. e0152878.
- (73) Chen, S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* **2023**, *2*, No. e107.
- (74) Masella, A. P.; Bartram, A. K.; Truszkowski, J. M.; Brown, D. G.; Neufeld, J. D. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinf.* **2012**, *13*, No. 31.
- (75) Joshi, N. Sabre - A Barcode Demultiplexing and Trimming Tool for FastQ Files 2011 <https://github.com/najoshi/sabre>.
- (76) Oksanen, J.; Blanchet, F. G.; Kindt, R.; Legendre, P.; Minchin, P.; O'Hara, R.; Solymos, P.; Stevens, M.; Szoecs, E.; Wagner, H.; Barbour, M.; Bedward, M.; Bolker, B.; Borcard, D.; Carvalho, G.; Chirico, M.; De Caceres, M.; Durand, S.; Evangelista, H.; FitzJohn, R.; Friendly, M.; Furneaux, B.; Hannigan, G.; Hill, M.; Lahti, L.; McGlenn, D.; Ouellette, M.; Cunha, E. R.; Smith, T.; Stier, A.; Braak, C. J. F.; Weedon, J. *Community Ecology Package. Vegan: Community*

Ecology Package; R package Version 2.6-4, 2022. <https://CRAN.R-project.org/package=vegan>.

(77) R Core Team. *R: A Language and Environment for Statistical Computing*, 2021. <https://www.R-project.org/>.

(78) Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag New York, 2016. <https://CRAN.R-project.org/package=ggplot2>.

Supporting Information

A versatile dual reporter to identify ribosome pausing motifs alleviated by translation elongation factor P

Urte Tomasiunaite¹, Tess Brewer¹, Korinna Burdack¹, Sophie Brameyer¹, Kirsten Jung^{1*}

¹ Faculty of Biology, Microbiology, Ludwig-Maximilians-Universität München, Martinsried, Germany

*Correspondence: jung@lmu.de

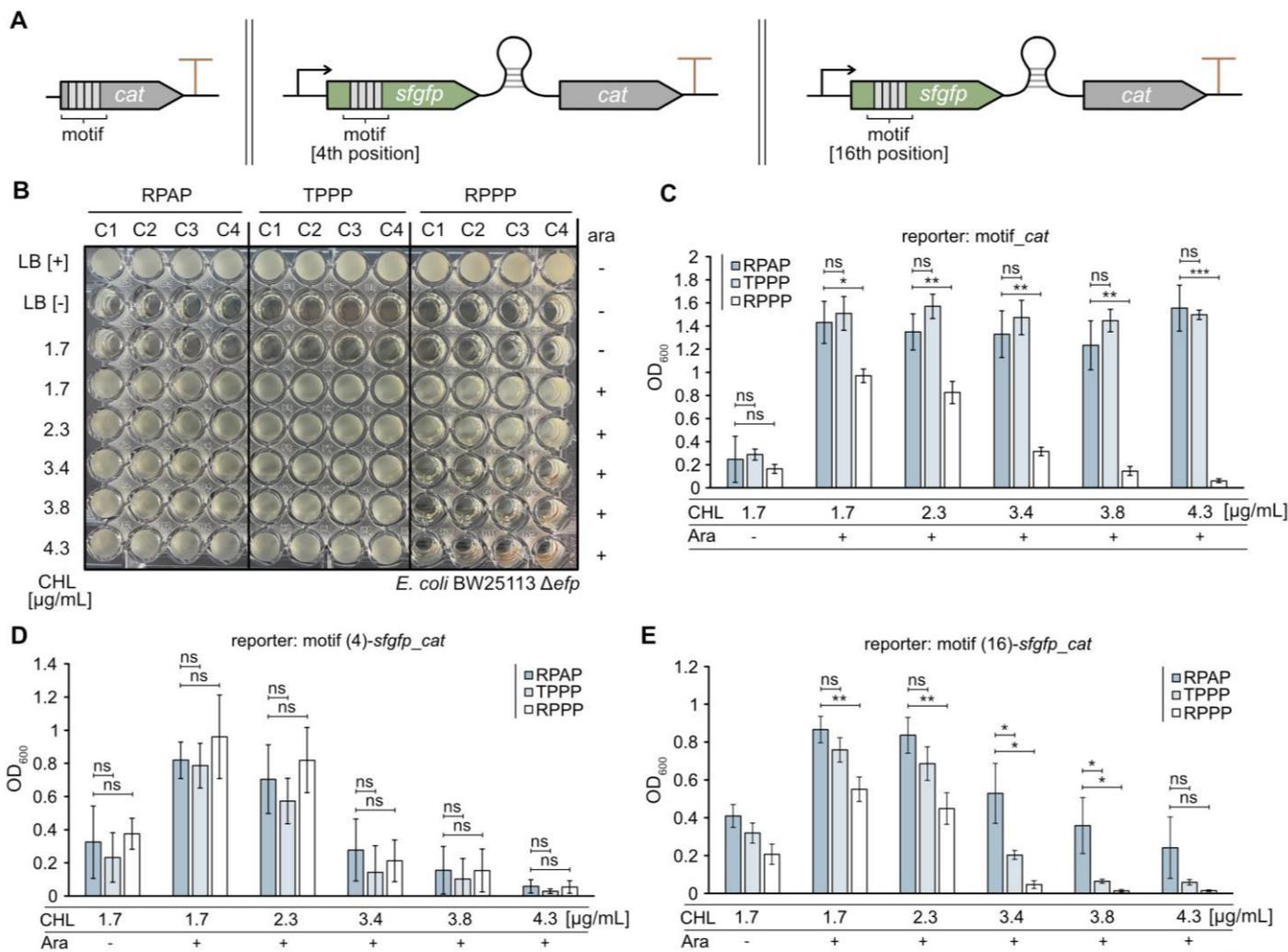


Figure S1. Bacterial survival-based determination of translational efficiencies by using a heterologous dual reporter system. **A** Construction of reporters for translational efficiency measurements. The reporters were designed to consist of either one reporter gene encoding an antibiotic resistance protein (left panel; chloramphenicol acetyltransferase, CAT) or two reporter genes encoding a fluorophore sGFP and CAT, linked with a weak translational coupling device¹. The motifs were located either from the fourth (4th) or from the sixteenth (16th) nucleotide position of the gene coding for the fluorescent protein. The arrows illustrate promoter locations. **B**, **C** 96-well assay plate of the growth measurements (**B**) and growth quantification (**C**) of *E. coli* BW25113 Δ *efp*. The strain contains the reporter plasmid pBAD24_motif_cat. **D**, **E** Growth quantification of *E. coli* BW25113 Δ *efp* containing the dual reporter plasmid pBAD24_motif_sfgfp_cat with the motif at the 4th (**D**) or the 16th (**E**) nucleotide position of the gene sequence coding for fluorescent protein. RPAP, TPPP and RPPP motifs were used to assess no stalling/ stalling of the ribosome (Figure 1A). Error bars indicate the standard deviation (SD) of three independent biological replicates. Statistics: student's unpaired two-sided t test (**** $p < 0.0001$; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ns $p > 0.05$). 1.7 μ g/mL CHL and no Ara (RPAP vs TPPP: ns $p = 0.750$ [C], ns $p = 0.570$ [D]; RPAP vs RPPP: ns $p = 0.524$ [C], ns $p = 0.732$ [D]); 1.7 μ g/mL CHL and 0.2 % (w/v) Ara (RPAP vs TPPP: ns $p = 0.589$ [C], ns $p = 0.752$ [D], ns $p = 0.099$ [E]; RPAP vs RPPP: * $p = 0.017$ [D], ns $p = 0.423$ [D], ** $p = 0.001$ [E]); 2.3 μ g/mL CHL and 0.2 % (w/v) Ara (RPAP vs TPPP: ns $p = 0.092$ [C], ns $p = 0.399$ [D], ns $p = 0.093$ [E]; RPAP vs RPPP: ** $p = 0.004$ [C], ns $p = 0.513$ [D], ** $p = 0.002$ [E]); 3.4 μ g/mL CHL and 0.2 % (w/v) Ara (RPAP vs TPPP: ns $p = 0.359$ [C], ns $p = 0.381$ [D], * $p = 0.036$ [E]; RPAP vs RPPP: ** $p = 0.003$ [C], ns $p = 0.634$ [D], * $p = 0.013$ [E]), 3.8 μ g/mL CHL and 0.2 % (w/v) Ara (RPAP vs TPPP: ns $p = 0.184$ [C], ns $p = 0.646$ [D], * $p = 0.041$ [E]; RPAP vs RPPP: ** $p = 0.002$ [C], ns $p = 0.991$ [D], * $p = 0.027$ [E]); 4.3 μ g/mL CHL and 0.2 % (w/v) Ara (RPAP vs TPPP: ns $p = 0.649$ [C], ns $p = 0.311$ [D], ns $p = 0.147$ [E]; RPAP vs RPPP: *** $p = 0.001$ [C], ns $p = 0.925$ [D], ns $p = 0.095$ [E]). Ara – arabinose, C1-C4 – clone number, CHL – chloramphenicol, *msc* - *mscarlet-1*, *sfgfp* - *super folded green fluorescent protein*.



Figure S2. Sequence of the pBAD24_motif_mscarlet-I_weak coupling1_cat expression cassette. The reporter cassette is composed of two reporter genes: *mscarlet-I* (pink arrow) and *chloramphenicol acetyl transferase* (*cat*, grey arrow). Both reporter genes are connected by a translational coupling device sequence (weak coupling 1¹, blue arrow). The motif responsible for the ribosome pausing (yellow arrow) is incorporated at the 5'-end of the *mscarlet-I* gene sequence. The restriction sites *NheI* (green box) and *PstI* (light pink box) were used to clone the construct into the pBAD24 vector. N - nucleotide.

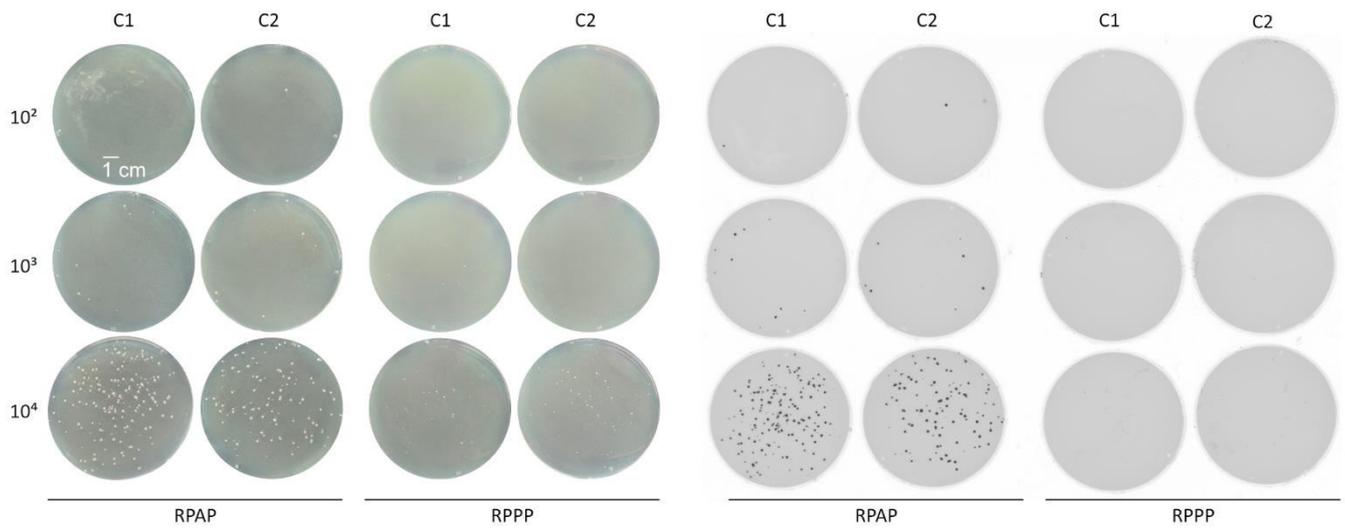
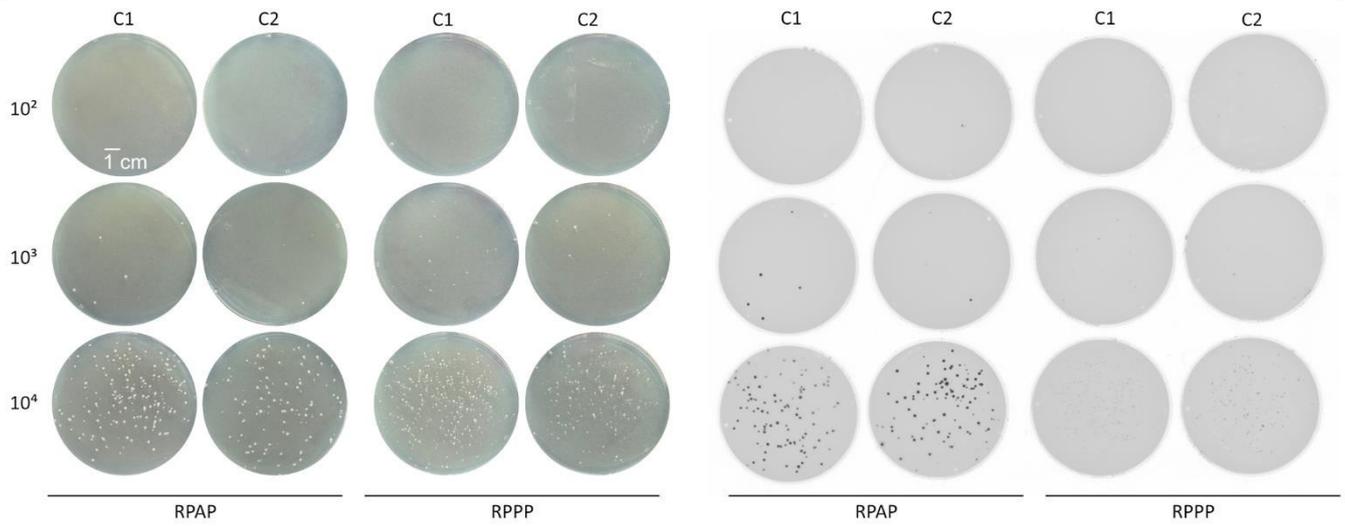
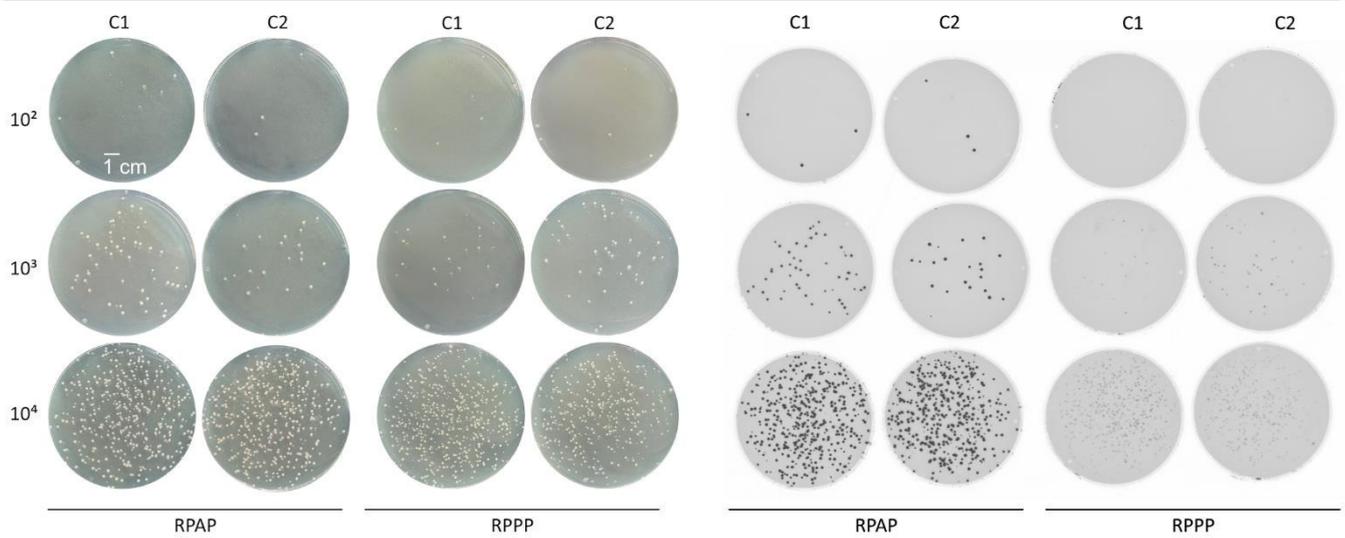
A**4.3 μg chloramphenicol****B****3.8 μg chloramphenicol****C****3.4 μg chloramphenicol**

Figure S3. Influence of the chloramphenicol concentration and the motif on the survival and fluorescence of the reporter strain. A-C Growth (left panel) and colony fluorescence (right panel) analysis of *E. coli* BW25113 Δefp on LB agar assay plates, supplemented with 0.2 % (w/v) arabinose and 4.3 μg (**A**), 3.8 μg (**B**) or 3.4 μg (**C**) chloramphenicol. $10^2 - 10^4$ of cells were plated. The strain was transformed with the dual reporter plasmid pBAD24_motif_mscarlet-I_cat, containing either RPAP (Arg-Pro-Ala-Pro) or RPPP (Arg-Pro-Pro-Pro) amino acid motif. C1 – C2 – clone number.

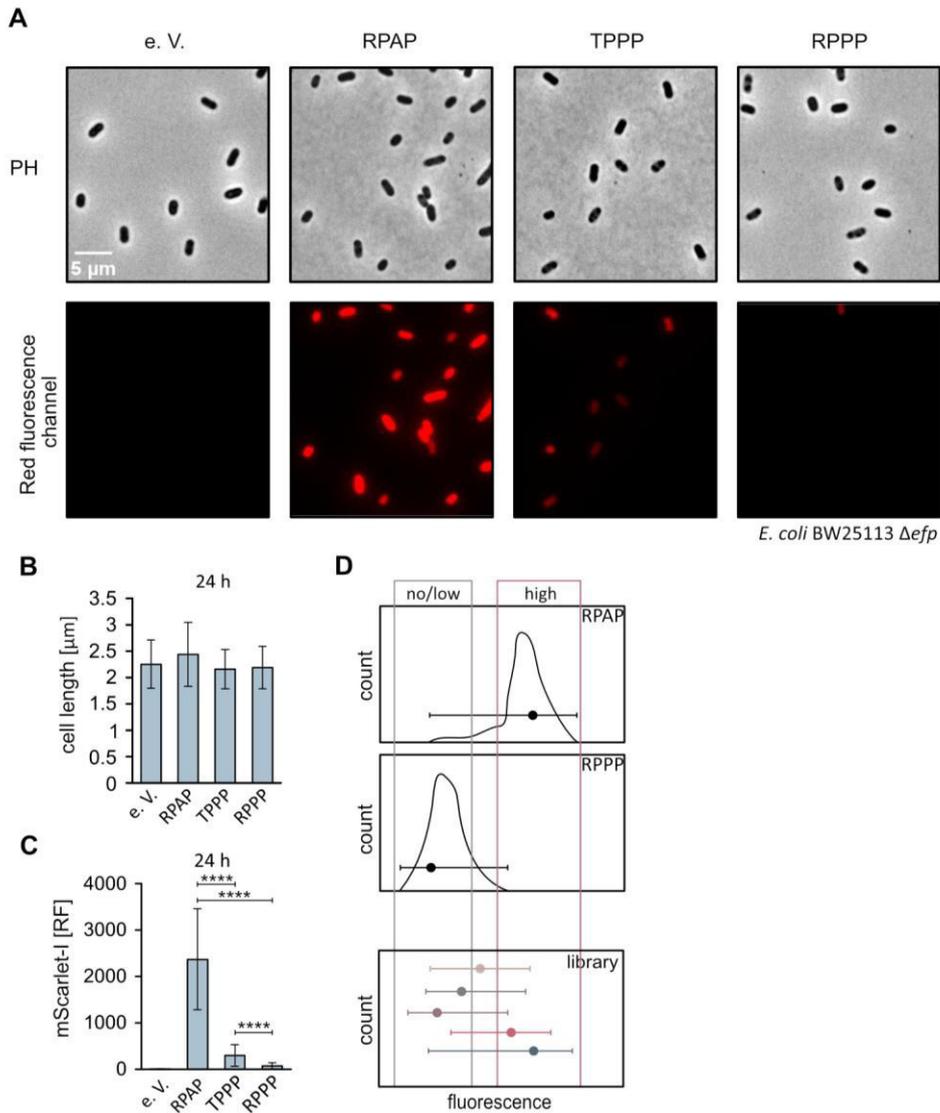


Figure S4. Heterogeneity in plasmid-based fluorophore expression. **A** Fluorescence microscopic images of BW25113 Δ *efp* transformed with dual reporter plasmid pBAD24_motif_*mScarlet-1_cat*. The top panel shows phase contrast (PH) images, the bottom panel - images from the red fluorescent channel. RPAP (Arg-Pro-Ala-Pro), TPPP (Thr-Pro-Pro-Pro) and RPPP (Arg-Pro-Pro-Pro) amino acid motifs were used to assess no stalling/ stalling of the ribosome. **B** Bar graph showing the average bacterial cell length from a total of 516 cells containing the reporter with the indicated amino acid motif. **C** Quantification of relative cell fluorescence from microscopic images. Microscopic images were taken 24 h after induction with 0.2 % (w/v) arabinose. **D** Schematic representation of the heterogeneity phenomenon in flow cytometry. e. V. – empty vector, h – hour, RF – relative fluorescence. Error bars indicate the standard deviation (SD) with following statistics: student's unpaired two-sided t test (**** $p < 0,0001$; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ns $p > 0.05$). RPAP vs TPPP: **** $p < 0,0001$; RPAP vs RPPP: **** $p < 0,0001$; TPPP vs RPPP: **** $p < 0,0001$ [C].

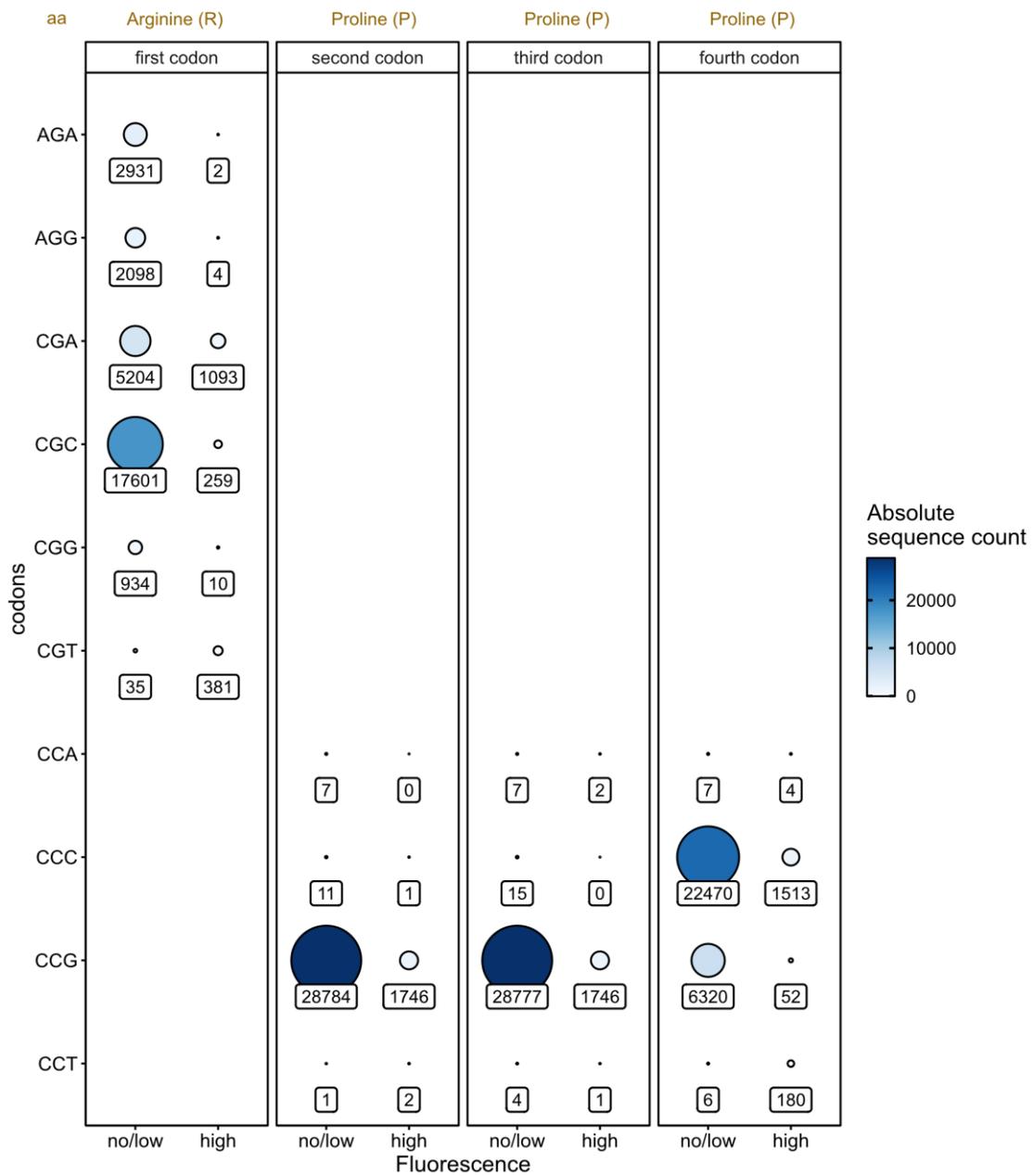


Figure S5. Distribution of codons encoding the motif RPPP in the library XXXX. Absolute sequence counts are depicted in numbers. The size of the circles corresponds to the count magnitude. aa – amino acid.

	illumina primer sequence	barcoding tag	complementary sequence to target DNA
illumina_seq_F1	TACACGACGCTCTCCGATCT	TCAT	TTAGCGGATCCTACCTGACG
illumina_seq_F2	TACACGACGCTCTCCGATCT	AAGTGA	TTAGCGGATCCTACCTGACG
illumina_seq_F3	TACACGACGCTCTCCGATCT	TGCGAGA	TTAGCGGATCCTACCTGACG
illumina_seq_F4	TACACGACGCTCTCCGATCT	GACATCCA	TTAGCGGATCCTACCTGACG
illumina_seq_R2	CAGACGTGTGCTCTCCGATCT	CGCTCA	ACATGAACTGAGGGGACAGG
illumina_seq_R3	CAGACGTGTGCTCTCCGATCT	GCTAACA	ACATGAACTGAGGGGACAGG
illumina_seq_R4	CAGACGTGTGCTCTCCGATCT	TTGACCAG	ACATGAACTGAGGGGACAGG

Figure S6. Primer design for Illumina library sequencing. Forward primers (F1-F4) and reverse primers (R2-R4) were designed to contain a sequence required for Illumina sequencing, a barcoding tag² for sample distinction, and a sequence complementary to the target DNA (library plasmids).

References

1. Rennig, M., Martinez, V., Mirzadeh, K., Dunas, F., Rojsater, B., Daley, D.O., and Norholm, M.H.H. (2018). TARSyn: Tunable Antibiotic Resistance Devices Enabling Bacterial Synthetic Evolution and Protein Production. *ACS Synth Biol* 7, 432-442. [10.1021/acssynbio.7b00200](https://doi.org/10.1021/acssynbio.7b00200).
2. Unterseher, M., Siddique, A.B., Brachmann, A., and Persoh, D. (2016). Diversity and Composition of the Leaf Mycobiome of Beech (*Fagus sylvatica*) Are Affected by Local Habitat Conditions and Leaf Biochemistry. *PLoS One* 11, e0152878. [10.1371/journal.pone.0152878](https://doi.org/10.1371/journal.pone.0152878).

3.1 Supplemental Tables of Chapter 3

The tables presented were sourced from Chapter 3 (<https://doi.org/10.1021/acssynbio.4c00534>) and have been reformatted to facilitate their transfer from Microsoft Excel to Microsoft Word format.

Data S1: Strains, plasmids and DNA oligonucleotides

Strain and Plasmid List

strain	genotype or description	Reference
<i>E. coli</i> BW25113	$\Delta(\text{araD-araB})567$, $\Delta\text{lacZ4787}>::\text{rrnB-3}$, λ -, <i>rph-1</i> , $\Delta(\text{rhaD-rhaB})568$, <i>hsdR514</i>	[1]
<i>E. coli</i> BW25113 Δefp	<i>F</i> -, $\Delta(\text{araD-araB})567$, $\Delta\text{lacZ4787}>::\text{rrnB-3}$, λ -, <i>rph-1</i> , $\Delta(\text{rhaD-rhaB})568$, Δefp , <i>hsdR514</i>	[2]
<i>E. coli</i> DH5 α λpir	<i>endA1 hsdR17 glnV44</i> (= <i>supE44</i>) <i>thi-1 recA1 gyrA96 relA1</i> $\phi 80^{\lambda}\text{lac}\Delta(\text{lacZ})\text{M15}$ $\Delta(\text{lacZYA-argF})\text{U169 zdg-232}::\text{Tn10}$ <i>uidA::pir</i> +	[3]
JW4107	$\Delta(\text{araD-araB})567$, $\Delta\text{lacZ4787}>::\text{rrnB-3}$, λ -, <i>rph-1</i> , $\Delta(\text{rhaD-rhaB})568$, $\Delta\text{efp-772}::\text{kan}$, <i>hsdR514</i>	[4]
plasmid	genotype or description	Reference
pBAD24 (e. V.)	pBR322 origin, arabinose inducible P _{BAD} promoter, Amp ^R -cassette	[5]
pBAD33 (e. V.)	p15A origin, arabinose inducible P _{BAD} promoter, CHL ^R -cassette	[5]
pBAD24_RPAP_cat	<i>chloramphenicol acetyl transferase</i> (<i>cat</i>); N-terminal motif coding for RPAP (arg-pro-ala-pro)	this study
pBAD24_RPPP_cat	<i>cat</i> ; N-terminal motif coding for RPPP (arg-pro-pro-pro)	this study
pBAD24_TPPP_cat	<i>cat</i> ; N-terminal motif coding for TPPP (thr-pro-pro-pro)	this study
pBAD24_pos2_RPAP_sfgfp_whp1_cat	<i>sfgfp</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for RPAP (arg-pro-ala-pro), nucleotide position of the motif in <i>sfgfp</i> : 4-15	this study
pBAD24_pos2_RPPP_sfgfp_whp1_cat	<i>sfgfp</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for RPPP (arg-pro-pro-pro), nucleotide position of the motif in <i>sfgfp</i> : 4-15	this study
pBAD24_pos2_TPPP_sfgfp_whp1_cat	<i>sfgfp</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for TPPP (thr-pro-pro-pro), nucleotide position of the motif in <i>sfgfp</i> : 4-15	this study
pBAD24_pos6_RPAP_sfgfp_whp1_cat	<i>sfgfp</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for RPAP (arg-pro-ala-pro), nucleotide position of the motif in <i>sfgfp</i> : 16- 27	this study
pBAD24_pos6_RPPP_sfgfp_whp1_cat	<i>sfgfp</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for RPPP (arg-pro-pro-pro), nucleotide position of the motif in <i>sfgfp</i> : 16- 27	this study
pBAD24_pos6_TPPP_sfgfp_whp1_cat	<i>sfgfp</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for TPPP (thr-pro-pro-pro), nucleotide position of the motif in <i>sfgfp</i> : 16- 27	this study
pBAD24_pos6_RPAP_mscarlet-I_whp1_cat	<i>mscarlet-I</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1);	this study

Versatile Dual Reporter to Identify Ribosome Pausing Motifs Alleviated by Translation Elongation Factor P

	motif coding for RPAP (arg-pro-ala-pro), nucleotide position of the motif in <i>mscarlet-I</i> : 16- 27	
pBAD24_pos6_RPPP_ <i>mscarlet-I</i> _whp1_cat	<i>mscarlet-I</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for RPPP (arg-pro-pro-pro), nucleotide position of the motif in <i>mscarlet-I</i> : 16- 27	this study
pBAD24_pos6_TPPP_ <i>mscarlet-I</i> _whp1_cat	<i>mscarlet-I</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for TPPP (thr-pro-pro-pro), nucleotide position of the motif in <i>mscarlet-I</i> : 16- 27	this study
pBAD24_pos6_AAGG_ <i>mscarlet-I</i> _whp1_cat	<i>mscarlet-I</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for AAGG (ala-ala-gly-gly), nucleotide position of the motif in <i>mscarlet-I</i> : 16- 27	this study
pBAD24_pos6_WLWS_ <i>mscarlet-I</i> _whp1_cat	<i>mscarlet-I</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for WLWS (trp-leu-trp-ser), nucleotide position of the motif in <i>mscarlet-I</i> : 16- 27	this study
pBAD24_pos6_GRGG_ <i>mscarlet-I</i> _whp1_cat	<i>mscarlet-I</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for GRGG (gly-arg-gly-gly), nucleotide position of the motif in <i>mscarlet-I</i> : 16- 27	this study
pBAD33b_ <i>E. c. efp</i> _ <i>epmA</i> _ <i>epmB</i> _ <i>nptI</i>	C-terminally His-tagged <i>efp</i> , <i>epmA</i> and <i>epmB</i> in pBAD33, <i>cat</i> substituted with <i>neomycin phosphotransferase I</i> (<i>nptI</i>) gene	this study
pBAD33- <i>efp</i> -His ₆ - <i>epmAB</i> _cat	C-terminal His-tagged <i>efp</i> and <i>epmAB</i> from <i>E. coli</i>	[6]
plasmid libraries		
lib_pBAD24_pos6_XPPX_ <i>mscarlet-I</i> _whp1_CAT	library; <i>mscarlet-I</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for XPPX (X-pro-pro-X), X codes for any amino acid, nucleotide position of the motif in <i>mscarlet-I</i> : 16- 27	this study
lib_pBAD24_pos6_XPXX_ <i>mscarlet-I</i> _whp1_CAT	library; <i>mscarlet-I</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for XPXX (X-pro-X-X), X codes for any amino acid, nucleotide position of the motif in <i>mscarlet-I</i> : 16- 27	this study
lib_pBAD24_pos6_XXPX_ <i>mscarlet-I</i> _whp1_CAT	library; <i>mscarlet-I</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for XXPX (X-pro-X-X), X codes for any amino acid, nucleotide position of the motif in <i>mscarlet-I</i> : 16- 27	this study
lib_pBAD24_pos6_XXXX_ <i>mscarlet-I</i> _whp1_CAT	library; <i>mscarlet-I</i> fused with <i>cat</i> , both genes connected with short sequence (weak hairpin loop 1); motif coding for XXXX (X-X-X-X), X codes for any amino acid, nucleotide position of the motif in <i>mscarlet-I</i> : 16- 27	this study

Versatile Dual Reporter to Identify Ribosome Pausing Motifs Alleviated by Translation Elongation Factor P

DNA and Oligonucleotides

check primers			
	primer name	5'-3' sequence	restriction site
P1	check_pBAD_F	GGC GTC ACA CTT TGC TAT GC	
P2	check_pBAD_R	CAG TTC CCT ACT CTC GCA TG	
P3	check_nptI_F	TGA GTA GGA CAA ATC CGC CGG G	
P4	check_nptI_R	CCG ATT TAG AGC TTG ACG G	
illumina sequencing			
	primer name	5'-3' sequence	restriction site
P5	illumina_seq_F1	TAC ACG ACG CTC TTC CGA TCT TCA TTT AGC GGA TCC TAC CTG ACG	
P6	illumina_seq_F2	TAC ACG ACG CTC TTC CGA TCT AAG TGA TTA GCG GAT CCT ACC TGA CG	
P7	illumina_seq_F3	TAC ACG ACG CTC TTC CGA TCT TGC GAG ATT AGC GGA TCC TAC CTG ACG	
P8	illumina_seq_F4	TAC ACG ACG CTC TTC CGA TCT GAC ATC CAT TAG CGG ATC CTA CCT GAC G	
P9	illumina_seq_R2	CAG ACG TGT GCT CTT CCG ATC TCG CTC AAC ATG AAC TGA GGG GAC AGG	
P10	illumina_seq_R3	CAG ACG TGT GCT CTT CCG ATC TGC TAA CAA CAT GAA CTG AGG GGA CAG G	
P11	illumina_seq_R4	CAG ACG TGT GCT CTT CCG ATC TTT GAC CAG ACA TGA ACT GAG GGG ACA GG	
reporter cloning			
	primer name	5'-3' sequence	restriction site
P12	pB24_RPAP_cat_F	ATC CGC TAG CAG GAG GAA TTC ACC ATT GAG AAA AAA ATC CGC CCG GCG CCG ACT GGA TAT ACC A	NheI
P13	pB24_RPPP_cat_F	ATC CGC TAG CAG GAG GAA TTC ACC ATT GAG AAA AAA ATC CGC CCG CCG CCG ACT GGA TAT ACC A	NheI
P14	pB24_TPPP_cat_F	ATC CGC TAG CAG GAG GAA TTC ACC ATT GAG AAA AAA ATC ACC CCG CCG CCG ACT GGA TAT ACC A	NheI
P15	pB24_2RPAP_sfgfp_F	ATC CGC TAG CAG GAG GAA TTC ACC ATG CGC CCG GCG CCG GGA AGC AAA GGA GAA GAA CTT TTC A	NheI
P16	pB24_2RPPP_sfgfp_F	ATC CGC TAG CAG GAG GAA TTC ACC ATG CGC CCG CCG CCG GGA AGC AAA GGA GAA GAA CTT TTC A	NheI
P17	pB24_2TPPP_sfgfp_F	ATC CGC TAG CAG GAG GAA TTC ACC ATG ACC CCG CCG CCG GGA AGC AAA GGA GAA GAA CTT TTC A	NheI
P18	pB24_6RPAP_sfgfp_F	ATC CGC TAG CAG GAG GAA TTC ACC ATG AGC AAA GGA GAA CGC CCG GCG CCG GAA CTT TTC ACT GGA GTT GT	NheI
P19	pB24_6RPPP_sfgfp_F	ATC CGC TAG CAG GAG GAA TTC ACC ATG AGC AAA GGA GAA CGC CCG CCG CCG GAA CTT TTC ACT GGA GTT GT	NheI
P20	pB24_6TPPP_sfgfp_F	ATC CGC TAG CAG GAG GAA TTC ACC ATG AGC AAA GGA GAA ACC CCG CCG CCG GAA CTT TTC ACT GGA GTT GT	NheI
P21	pB24_6RPAP_mscl_F	ATC CGC TAG CAG GAG GAA TTC ACC ATG GTG AGC AAG GGC CGC CCG GCG CCG GAG GCA GTG ATC AAG GAG TT	NheI
P22	pB24_6RPPP_mscl_F	ATC CGC TAG CAG GAG GAA TTC ACC ATG GTG AGC AAG GGC CGC CCG CCG CCG GAG GCA GTG ATC AAG GAG TT	NheI
P23	pB24_6TPPP_mscl_F	ATC CGC TAG CAG GAG GAA TTC ACC ATG GTG AGC AAG GGC ACC CCG CCG CCG GAG GCA GTG ATC AAG GAG TT	NheI
P24	ov_sfgfp_whp1_R	AAT AGG AGG ACC TCC TAT TTC ATT TGT AGA GCT CAT CCA TGC CA	

Versatile Dual Reporter to Identify Ribosome Pausing Motifs Alleviated by Translation Elongation Factor P

P25	ov_whp1_cat_F	AAT AGG AGG TCC TCC TAT TGA GAA AAA AAT CAC TGG ATA TAC CA	
P26	ov_mscl_whp1_R	AAT AGG AGG ACC TCC TAT TTT ACT TGT ACA GCT CGT CCA TGC CG	
P27	pB24_cat_PstI_R	AGC CCT GCA GGT CGA CTC TAG ATT ACG CCC CGC CCT GCC ACT CAT CG	PstI
P28	pB24_AAGG_mscl_F	GGG GGA GAG GCA GTG ATC AAG G	
P29	pB24_AAGG_mscl_R	GGC AGC GCC CTT GCT CAC CAT G	
P30	pB24_WLWS_mscl_F	TGG TCC GAG GCA GTG ATC AAG G	
P31	pB24_WLWS_mscl_R	CAG CCA GCC CTT GCT CAC CAT G	
P32	pB24_GRGG_mscl_F	GGT GGG GAG GCA GTG ATC AAG G	
P33	pB24_GRGG_mscl_R	ACG GCC GCC CTT GCT CAC CAT G	

resistance cassette cloning

	primer name	5'-3' sequence	restriction site
P34	ov_nptI_ins_F	CTT TTG TTT ATT TTT CTA AAC TTG GCC GGG TTA CAT TGC ACA AGA TA	
P35	ov_nptI_ins_R	GCC GCT ACA GGG CGC GTA AAT CAA TCT TTG GCC GGC GCC GTC CCG TC	
P36	ov_nptI_backb_F	AGA TTG ATT TAC GCG CCC TGT AGC GGC GCA TTA AGC GCG GCG GGT GT	
P37	ov_nptI_backb_R	TTT AGA AAA ATA AAC AAA AGA GTT TGT AGA AAC GCA AAA AGG CCA TC	

library cloning

	primer name	5'-3' sequence	restriction site
P38	lib_mscl_XPPX_F	C ATG GTG AGC AAG GGC NNN CCG CCG NNN GAG GCA GTG ATC AAG GAG T	
P39	lib_mscl_XXXX_F	C ATG GTG AGC AAG GGC NNN NNN NNN NNN GAG GCA GTG ATC AAG GAG T	
P40	lib_mscl_XPXX_F	C ATG GTG AGC AAG GGC NNN CCG NNN NNN GAG GCA GTG ATC AAG GAG T	
P41	lib_mscl_XXPX_F	C ATG GTG AGC AAG GGC NNN NNN CCG NNN GAG GCA GTG ATC AAG GAG T	
P42	lib_mscl_R	GCC CTT GCT CAC CAT GGT GAA TTC CTC CTG CTA GC	

References for Chapter 3 Supplemental Tables

1. Datsenko, K.A. and B.L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*, 2000. 97(12): p. 6640-5.
2. Sieber, A., Parr, M., von Ehr, J., Dhamotharan, K., Kielkowski, P., Brewer, T., Schäpers, A., Krafczyk, R., Qi, F., Schlundt, A., et al. (2024). EF-P and its paralog EfpL (YeiP) differentially control translation of proline-containing sequences. *bioRxiv*, 2024.2004.2015.589488. 10.1101/2024.04.15.589488.
3. Macinga, D.R., M.M. Parojcic, and P.N. Rather, Identification and analysis of *aarP*, a transcriptional activator of the 2'-*N*-acetyltransferase in *Providencia stuartii*. *J Bacteriol*, 1995. 177(12): p. 3407-13.
4. Baba, T., et al., Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*, 2006. 2: p. 2006 0008.
5. Guzman LM, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177, 4121-4130 (1995).
6. Volkwein, W., Krafczyk, R., Jagtap, P.K.A., Parr, M., Mankina, E., Macosek, J., Guo, Z., Furst, M., Pfab, M., Frishman, D., et al. (2019). Switching the Post-translational Modification of Translation Elongation Factor EF-P. *Front Microbiol* 10, 1148. 10.3389/fmicb.2019.01148.

4 Concluding discussion and outlook

The synthesis of consecutive prolines highly depends on a fully functional EF-P (61,62). In addition to its critical roles in cellular fitness and environmental cell adaptation during stress (61,154,161), its relevance for bacterial virulence renders EF-P a suitable target for antibiotic research and drug production (155,190-192). The discovery of functional EF-Ps that do not require PTMs (172) makes EF-P an intriguing protein for energy-efficient protein production in biotechnological applications. However, the functional characteristics of EF-P at the molecular level remain incomplete, and its potential applications in biotechnology necessitate further investigations. The existence of both unmodified and modified EF-Ps prompts questions on the evolution and necessity of such chemically diverse modification systems for the translation of polyprolines. This thesis uncovers the fundamental principles that govern the functionality of unmodified EF-Ps, with a particular focus on the correlation between the sequence identity of these EF-Ps and their functionality in *Escherichia coli* (Chapter 2). The novel *in vivo* reporter system for ribosomal pausing measurements (Chapter 3) can serve as a tool for future studies to investigate the spectra of polyproline motifs rescued by unmodified EF-Ps and by other factors recently identified to be involved in the polyproline synthesis (YfmR, Uup, YebC)(193-196).

4.1 *R. vannielii* EF-P is fully functional in *E. coli*

In many bacterial species, PTMs boost the activity of EF-Ps (153,155,158,159,169,170,173). Although bacteria have evolved species-specific modification pathways to activate their EF-Ps, studies demonstrated that EF-Ps with particular modifications are also interchangeable across species. For example, the co-expression of *Shewanella oneidensis* EF-P with its corresponding modification enzyme EarP in *E. coli* can complement the deletion of the native *efp* (155). Minor substitutions within the activation loop of *E. coli* EF-P were sufficient for *E. coli* to switch the modification from lysylation to rhamnosylation upon co-expression of the modification enzyme EarP (181). However, the rhamnosylated *E. coli* EF-P could not fully rescue the translation of pH sensor CadC, which contains polyprolines (181), suggesting that additional factors might hinder the rhamnosylated EF-P from reaching full activities. These observations indicate EF-Ps and their cognate modification systems have co-evolved, working together to maximize activity in the native host. The discovery of unmodified EF-Ps from Actinobacteria with full activity (172) has prompted a debate regarding the necessity of the modifications developed in other species. It is reasonable to expect that unmodified EF-Ps, which have been evolutionarily optimised over a considerable period of time, could be easily transferable to other species and functionally replace other EF-Ps that are dependent on a PTM. However, heterologous expression of unmodified EF-Ps from Actinobacteria have only been shown to restore native EF-P function in other Actinobacteria, such as *Corynebacterium glutamicum* (172). When expressed in *E. coli*, whose native EF-P requires a PTM (153,159,166,167), the actinobacterial EF-Ps are not functional (172). Accordingly, this thesis examined the activities of further EF-Ps that were isolated from strains that do not exclusively belong to the Actinobacteria phylum but were members of the Bacteroidetes, α -Proteobacteria and δ -Proteobacteria

(Chapter 2)(172). The latter two phyla are more closely related to γ -Proteobacteria phyla, from which *E. coli* originates, than to Actinobacteria. Here, it could be hypothesised that bacteria species closely related to *E. coli* might develop EF-Ps more compatible with *E. coli*'s translational apparatus, such as the ribosome or tRNA. Among the PGKGP subfamily EF-Ps evaluated in this thesis, the EF-P from *Rhodomicrobium vannielii* effectively rescued translation of the polyproline-containing pH sensor CadC, achieving levels similar to those observed with the native *E. coli* EF-P (Chapter 2). Mass spectrometry studies with purified *R. vannielii* EF-P confirmed that it is unmodified in *E. coli* (Chapter 2). These findings support the hypothesis that PGKGP serves as a signature motif for unmodified EF-Ps beyond the Actinobacteria phyla. However, it remains unclear whether this EF-P was horizontally transferred to other bacterial phyla or is an example of convergent evolution (197).

R. vannielii is a photoheterotrophic anaerobic bacterium named in honour of the Dutch-American microbiologist Cornelis Bernardus van Niel (198). Isolated from the Gaden Hot Spring in Malaysia (50-58 °C), *R. vannielii* can grow at elevated temperatures (199). Here, EF-P could play a critical role, since earlier studies have reported that EF-P performance seems to be more crucial at higher temperatures (161). Therefore, the absence of an EF-P modification could be advantageous for *R. vannielii* in terms of energy relocation that is not needed for modifying enzyme production towards other processes, helping the bacterium to cope with challenging conditions. Similarly to certain actinobacterial representatives (*Streptomyces coelicolor*: 1.08 XPPX motif/protein; *Mycobacterium tuberculosis*: 1.17 XPPX motif/protein)(172), *R. vannielii* exhibits relatively high abundance of polyproline-containing motifs within its proteome (0.88 XPPX motif/protein, unpublished data), whereas *E. coli* has 0.49 XPPX motif/protein (134). It thus appears plausible that bacteria adapted to survive in extreme conditions, and which are forced to cope with polyproline motif translation, did evolve an EF-P with high-performance capabilities despite the absence of a PTM. However, studies of *R. vannielii* EF-P in the native host were not within the scope of this thesis and represent a valuable topic for further research.

In contrast to the *R. vannielii* EF-P, all other PGKGP subfamily EF-Ps tested could not rescue CadC synthesis (in Chapter 2 Figure 1B-C). High diversity in amino acid sequence composition indicates the potential for incompatibility between tested PGKGP subfamily EF-Ps and *E. coli*'s translational machinery (172)(Chapter 2). Despite the conservation observed in the protein 3D structures of the PGKGP subfamily EF-Ps together with EF-Ps, which require a PTM for optimal activity (172), it can be suggested that sequence variations might influence the functional characteristics of these proteins. The 44 % sequence identity of PGKGP subfamily EF-Ps and *E. coli* EF-P underscores the high diversity in the amino acid sequence among compared EF-Ps (in Chapter 2 Table S1). It is, therefore, unsurprising that all tested EF-Ps performed differently and that only one EF-P candidate was found to rescue translational stalling (Chapter 2).

Taken together, it has been demonstrated that an unmodified EF-P can substitute for the native modified EF-P in synthesising the polyproline-containing pH sensor CadC in *E. coli*. Given the structural homology and sequence diversity of EF-Ps within the PGKGP subfamily, it seems possible that the amino acid type within the EF-P backbone may influence protein function in *E. coli*.

4.2 Amino acids within the EF-P backbone determine the functionality of unmodified EF-Ps

A molecular mechanism of EF-P acting during polyproline synthesis in *E. coli* was proposed based on structural investigations (151). It has been suggested that the incompatibility between the polyproline nascent chain and the exit tunnel destabilises the P-site tRNA. In this scenario, the modified EF-P loop makes contact with the CCA end of the P-site tRNA, stabilizing it and thereby allowing the nascent chain to adopt a more favourable conformation, improving the peptide bond formation (151). This model highlights the necessity of a modification system of EF-P for the translation of polyproline sequences. Therefore, it remains unclear how unmodified EF-Ps, which lack this extension of the activation loop, can reach the CCA end. As the *R. vannielii* EF-P and *E. coli* EF-P have comparable levels of activity (Chapter 2) it may be possible that amino acids situated within the EF-P backbone, apart from the activation loop, may have functional roles. Functional studies with constructed EF-P hybrids narrowed the search for amino acids with the greatest impact on unmodified EF-P functionality in *E. coli* (Chapter 2). These investigations revealed that *R. vannielii* EF-P domain I had the greatest effect on EF-P activity in *E. coli* (in Chapter 2 Figure 3). This observation would not be a surprise in the context of modified EF-Ps, given that all activation loops which require a PTM are localized within domain I (150,152,155,156,166,169,174). However, in the case of *R. vannielii* EF-P, mass spectrometry studies verified no modification at this position (Chapter 2), indicating that, in addition to the activation loop, other amino acids present in domain I, may also be involved in regulating the activity of EF-P.

Multiple sequence alignments revealed the presence of amino acids in the PGKGP subfamily of EF-Ps that exhibited common characteristics distinct from those observed in *E. coli* and *R. vannielii* EF-Ps (Chapter 2). These observations suggested about their critical roles in EF-P functionality in *E. coli*. Three amino acid positions within the *S. venezuelae* EF-P backbone were identified with considerable impact on its activity in *E. coli* using an established synthetic molecular engineering approach (in Chapter 2 Figure 4B). Particularly, substituting amino acids with neutral side chains with amino acids with charged side chains at positions 35, 50 and 63-65 (positions numbered according to *S. venezuelae* EF-P) resulted in elevated EF-P activity in *E. coli* (Chapter 2). Amino acid replacements at equivalent positions in *P. gingivalis* EF-P and *W. virosa* EF-P improved their activities in *E. coli* as well, demonstrating that certain amino acids boost the functionality of unmodified EF-Ps (Chapter 2). Notably, in *W. virosa* EF-P, the substitution of three positions was sufficient to render it a fully functional EF-P, with activities comparable to the modified *E. coli* EF-P (Chapter 2). These results highlight the applicability of the developed synthetic engineering system to improve the activity of other unmodified EF-P candidates in *E. coli*.

Structural alignments of the PGKGP subfamily EF-P wild types and their corresponding variants revealed no notable alterations in protein structure (data not shown). This indicated that the reduced activity of unmodified EF-Ps (Chapter 2) is more likely due to the incompatibility between these EF-Ps and the surrounding translational apparatus of *E. coli* than overarching structural differences. Structural

modelling of the *R. vannielii* EF-P with *E. coli*'s translational machinery (151) indicates that amino acids with charged side chains may interact with either the P-site tRNA (*S. venezuelae* EF-P substitutions: P35Q, V50R) or with the ribosome (23S rRNA; *S. venezuelae* substitutions: 63-65)(in Chapter 2 Figure S7). Accordingly, these distinct amino acids might be critical for EF-P compatibility with the translational machinery, ultimately determining how effective those unmodified EF-P can contribute during the polyproline motif translation in *E. coli* (Figure 5). Recent studies have proposed a model for the kinetics of EF-P binding to the translational complex in *E. coli* (200,201). Based on these studies, EF-P constantly scans the ribosome by rapidly binding to the translation complex. A proper binding to the translational machinery was achieved after *E. coli* EF-P was modified (200). The unmodified PGKGP subfamily EF-P variants generated in this study may contain a more compatible protein structure to the *E. coli* ribosome, resulting in improved binding to the translation machinery compared to their wild type counterparts (Figure 5). This could allow improved positioning during the rapid scanning process. However, additional structural analyses are needed for further insight into the precise molecular mechanism of EF-P during polyproline synthesis, such as using Cryo-EM with PGKGP subfamily EF-Ps in complex with translational machinery.

Taken together, this thesis developed a platform for optimizing and engineering unmodified EF-Ps intending to improve their functionality in *E. coli*. Structural modelling provided initial insights into the role of critical amino acids in making contact with the surrounding translational machinery. This thesis provides a framework for further research to understand the molecular principles behind the functionality of unmodified EF-Ps. By demonstrating that the native modified *E. coli* EF-P can be functionally replaced by unmodified EF-Ps to maintain translation efficiency indicates that these EF-Ps are promising candidates in biotechnological applications. This could be particularly advantageous in a biotechnological context, as strains can redirect the energy originally dedicated to the production of modification enzymes towards the synthesis of the desired biotechnological product.

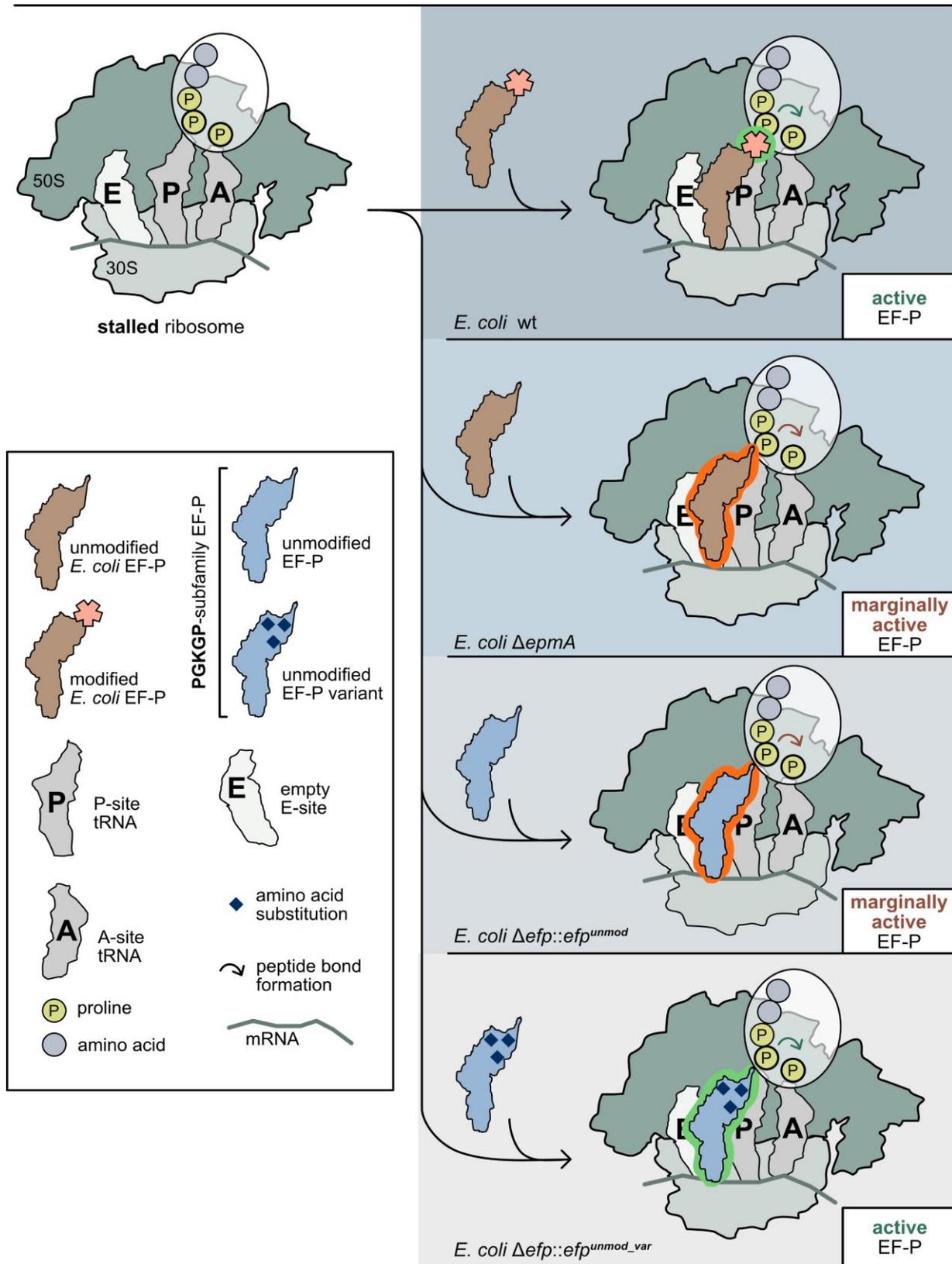


Figure 5. Schematic illustration of stalled ribosome rescue at polyproline motifs by activated EF-Ps. *E. coli* EF-P requires a post-translational modification (PTM) to alleviate stalled ribosomes and facilitate fast translation of polyproline motifs. Amino acid substitutions render unmodified EF-Ps from the PGKGP subfamily to be active during the translation of polyproline motifs. Marginally active EF-Ps are indicated by an orange background, active EF-Ps are indicated by a green background. EF-P – elongation factor P.

4.3 Preferences in the rescue of distinct polyproline motifs

Deletion of *efp* can have severe consequences for bacteria, leading to detrimental phenotypes (155,159,161,169,202) or cell death (156,191). Notably, in certain species, the absence of the modifying enzyme leads to growth defects similar to those caused by the deletion of the *efp* gene alone (154,155,157). This further emphasizes the significance of modifications in those species, as these EF-Ps in an unmodified state can only partially compensate for the absence of fully functional modified EF-Ps. In line with previous studies, this thesis reveals comparable phenotypes observed with *E. coli* mutants that either lacked one of the genes responsible for the modification ($\Delta epmA$) or the gene encoding EF-P (Δefp) (Chapter 2). Both mutants had growth defects compared to wild type (Chapter 2). Chromosomal integration of the *R. vannielii* EF-P was sufficient to rescue growth defects observed with *E. coli* mutants lacking genes encoding EF-P or EpmA (see Chapter 2 Figures 2A-C). Based on these results, it appears that the *R. vannielii* EF-P is capable of fulfilling the role of the native modified EF-P in *E. coli*, not only during the rescue of CadC synthesis (in Chapter 2 Figure 1) but potentially also during the synthesis of other proteins (in Chapter 2 Figure 2). Consistent with these observations, the overall proteome profiles of *E. coli* wild type and the *E. coli* mutant encoding *R. vannielii* EF-P clustered together the closest in the PCA analysis (Chapter 2). Thus, in this study, an *E. coli* mutant strain encoding an unmodified EF-P was constructed, which maintains proteome homeostasis comparable to *E. coli* wild type. This further highlights the interchangeability of both EF-Ps.

A more comprehensive investigation of the rescue of proteins with diprolyl motifs (XPPX) was done by comparison of XPPX-containing protein abundances between the wild type and all EF-P mutant strains (in Chapter 2 Figure 2E and Figure S5). Here, diverse scatter profiles could be detected, indicating different preferences of these mutants regarding XPPX motif-containing protein synthesis. The mutant strain encoding *R. vannielii* EF-P could rescue 74 % of downregulated proteins containing XPPX motifs (in Chapter 2 Figure 2F). At first sight, this seemed unexpected, as the two strains clustered together most closely in the PCA analysis based on their proteomes (Chapter 2). However, previous studies have shown that not all proteins with polyproline motifs are equally rescued by wild type and mutants lacking the *efp* or lacking genes encoding modification enzymes (131,172,184) (Figure 6). In particular, some proteins with consecutive prolines were found to be more abundant in *E. coli* mutants lacking the modification enzymes or EF-P compared to the wild type strain (131,184). These data further highlight that, in addition to the fully functional EF-P, other factors may influence the strength of ribosome stalling. These factors could include the sequences upstream and downstream of the motif (Figure 6), the context of the codons, the location of the motif within the ORF or the composition of the motif itself (131,184,186,189). Such sequence characteristics may help the cell to precisely control protein abundance even in the presence of a fully functional EF-P. This might explain why the *E. coli* mutant expressing *R. vannielii* EF-P could not rescue all proteins with XPPX motifs, as shown in Chapter 2.

4.4 A tuneable plasmid based dual reporter for rapid ribosome pausing measurements

EF-Ps from Actinobacteria have been shown to function effectively within their native hosts without requiring any PTMs (172). However, the complete range of motifs these EF-Ps can rescue is still not completely understood. The discovery of motifs comprising fewer than three prolines, which were shown to highly impact protein synthesis in *Escherichia coli* (131)(Figure 6), broadens the spectrum of motifs that necessitate further investigation with unmodified EF-Ps. Interestingly, the type of amino acid present upstream and downstream of the two prolines within the polyproline motif shows a notable degree of variability in its impact on the strength of ribosome pausing (131)(Figure 6). Thus far, no ultimate rule has been established to explain how these amino acid variations influence ribosome stalling thus necessitating further investigations. The design and discovery of unmodified EF-Ps that demonstrate functionality in *E. coli* (as discussed in Chapter 2) raised questions regarding which specific motifs these EF-Ps might be able to rescue. Common methods for assessing ribosome pausing include ribosome profiling, proteomics, and *in vitro* techniques (131,182-184,186). Implementing these methodologies requires specialized equipment and reagents, rendering them impractical for many research laboratories. The development of *in vivo* systems for investigating ribosome pausing in living organisms could help overcome these barriers. Examples of these reporters are a chromosomally integrated fluorophore with stalling motifs (172), the P_{cadBA} *lacZ*-based reporter system (61), the HisL reporter system (189) or the L-serine biosynthesis-based reporter system (203). One advantage of these systems is that they permit stable control over reporter expression, providing constant ribosome pausing measurements. However, the constraints associated with these systems are either that the measured motifs remain the same due to measurements of the same protein with polyprolines (61) or that the exchange of motifs is technically challenging due to the necessity of chromosomal incorporation of the motif (172,189,203). Additionally, chromosomal reporters complicate the generation of libraries, whereas plasmid-based library generation is relatively straightforward. To enhance the efficacy of screenings, thereby rendering them suitable for high-throughput library screenings, a plasmid-based dual reporter was developed (Chapter 3).

Plasmid-based reporter systems, which use fluorophores to measure ribosome pausing have been described previously (158,161,178). However, these have mainly used the fluorescent protein GFP as a reporter, which requires additional equipment such as microscopes or spectrometers for measurements. This dependency on specialized tools could be a drawback when aiming for more accessible or high-throughput methods. To overcome this limitation, the fluorophore mScarlet-I was chosen to be part of the reporter, as it is directly detectable from the colony through a colour change, eliminating the need for a microscope or spectrometer (Chapter 3). Moreover, the fusion of the gene encoding mScarlet-I with a resistance gene cassette encoding chloramphenicol acetyltransferase (CAT, chloramphenicol acetyltransferase) enabled the assessment of ribosome pausing based on two distinct outcomes: cell fluorescence and bacterial survival (Chapter 3). The combination of a fluorophore and an antibiotic resistance protein enhances the tunability of the system (Figure 7). In particular, higher antibiotic concentrations facilitate the preferential identification of weaker pausing motifs, as cells carrying

plasmids with strong pausing motifs are unable to express the antibiotic resistance gene in sufficient quantities to survive (in Chapter 3 Figure 1C-D and S1C-E)(Figure 7). This underscores the tunability of the system that allows directed studies of motifs of interest (Figure 7). Future investigations could employ this system to pre-determine the antibiotic concentrations required to exclude motifs that are not central to the study (e.g. motifs causing strong ribosome pauses), thus facilitating a more targeted analysis.

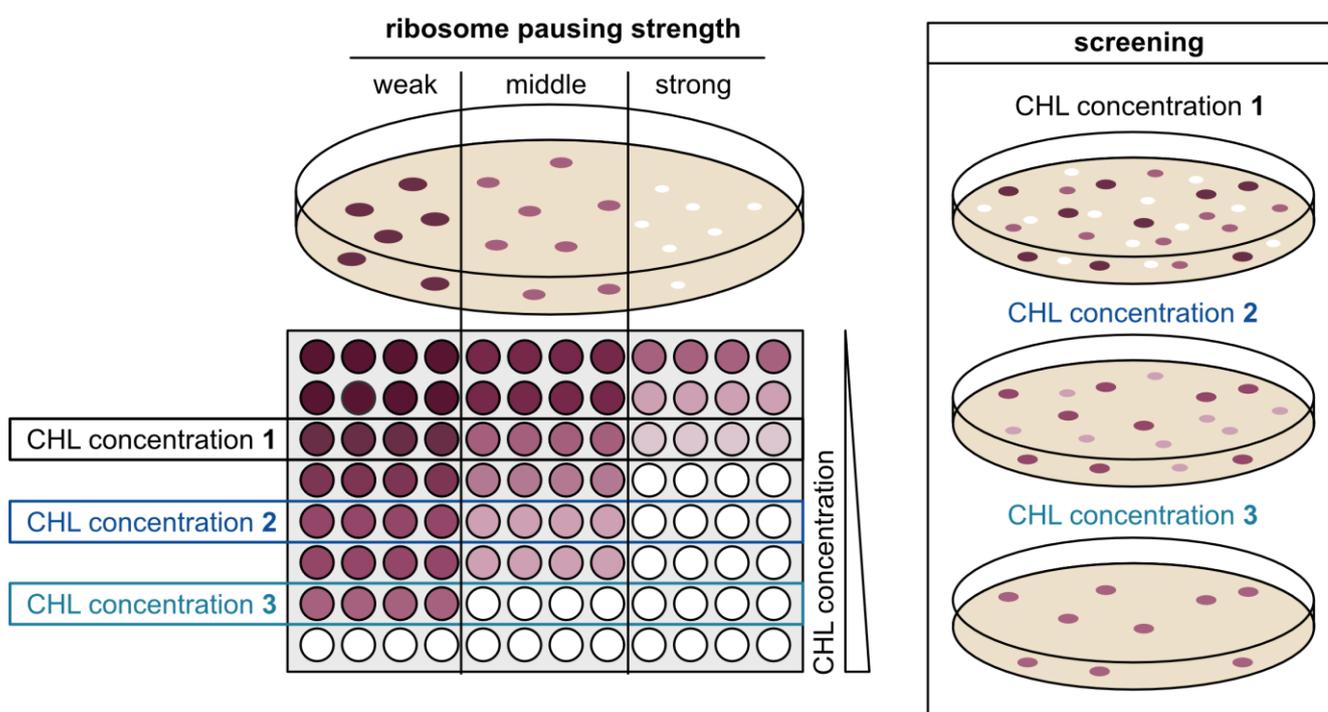


Figure 7. Tuneability of the dual reporter system during measurements of the ribosome pausing. The illustration depicts a 96-well plate (left bottom) inoculated with clones containing a dual-reporter plasmid with motifs causing varying ribosome pausing strengths (columns 1 to 4 – weak pausing, columns 5 to 8 – middle pausing, columns 9 to 12 – strong pausing). The chloramphenicol (CHL) concentration in the agar assay plates dictates the selective growth of clones carrying the dual reporter plasmid with stalling motifs (right, screening), highlighting the system's tunability.

Consistent with activity assays, proteomics and ribosome profiling data from previous studies (131,186)(Figure 7), the dual reporter system allowed the detection of motifs with different effects on translation (in Chapter 3 Figure 2C). High throughput screening with motif libraries identified the strength of additional motifs (strong pausing: PPPA, PPPL, GPPV, PPPR, PPPP, GPPH; weak pauses: LPPH, FPPF, TPPR, FPPL, LPPF)(in Chapter3 Figure 3B), which were in line with previous studies as well (134) (Figure 6)(131). These observations suggest that the system generated in this study is suitable for measuring ribosome pausing, as it successfully reproduces data from studies using alternative systems for measurements. While screening the XXXX library, motifs lacking prolines (AAGG, GRGG, WLWS) were also identified (in Chapter 3 Figure 3C). Notably, the translation of these motifs could not be rescued by a modified EF-P in *E. coli* (in Chapter 3 Figure 3D), which is not unexpected given that the XXXX motif

screening was carried out only in strains lacking EF-P (Δefp). The fact that EF-P could not rescue the translation of these motifs suggests that other mechanisms or factors could be involved during the translation of AAGG, GRGG and WLWS. These findings highlight the reporter system's broader applicability; therefore, future research could employ the dual reporter system to examine motifs that rely not only on EF-P to be rescued but also on alternative mechanisms.

Recently, members of the F subfamily of ABC ATPases (ABCF), namely YfmR/UuP, have been shown to be involved in translating polyproline motifs like EF-P (193,194,196). Severe growth defects were detected with a strain lacking *yfmR* and *efp* in *B. subtilis* (193,194). Additionally, *yfmR* and *efp* were shown to have partially overlapping functions since the deletion of *yfmR* in *B. subtilis* Δefp mutants results in a more pronounced ribosomal arrest during polyproline synthesis compared to the deletion of *efp* alone (193). Besides YfmR/UuP, another RNA binding translation factor, YebC, was also recently described to be involved in the translation of polyproline motifs (195). Ribosome profiling data, together with *in vivo* and *in vitro* studies, suggest that ribosomal pausing at polyproline stretches increases in the absence of YebC (195). In conclusion, it can be argued that EF-P is not the sole factor responsible for alleviating stalled ribosomes during polyproline synthesis. However, the spectrum of motifs these newly discovered systems can rescue is currently limited. The dual reporter system described in this thesis represents a versatile platform for future research in this area.

4.5 Outlook and future perspectives

This thesis provides fundamental insights into the functional principles of unmodified EF-Ps using a combination of functional studies, proteomic studies and structural modelling. The expression of unmodified EF-P variants, designed in this study, in other bacterial species would further improve the understanding of the roles the identified critical amino acids play during polyproline motif translation (Figure 8A). Structural studies (e. g. Cryo-EM) with unmodified EF-P variants and translational machinery (e.g. ribosome, tRNA) could provide further evidence regarding the spatial positioning of the EF-P during the peptide bond formation (Figure 8B). Determination of the binding kinetics would give insights about the substrate recognition elements necessary for unmodified EF-P to bind to the ribosome and tRNA during translation in time (Figure 8C). Moreover, it could provide evidence about the frequency with which these EF-Ps can bind to their respective binding partners. Strains with chromosomally integrated unmodified EF-Ps in *E. coli* would provide a platform for heterologous protein production in synthetic biology and biotechnological applications. The production of biotechnological and chemical compounds in these strains should be examined with respect to quantities produced, product quality, and potential energy savings (Figure 8D).

This thesis describes the development and the characterization of a plasmid-based dual reporter system for ribosome pausing measurements. Future studies using this platform can investigate the translational rescue of polyproline motif containing protein libraries (Chapter 3) in strains expressing unmodified EF-Ps/EF-P variants (generated in Chapter 2)(Figure 8E). These studies would reveal the

motif spectra's these EF-Ps are capable to rescue. Moreover, motif libraries could be investigated in strains lacking newly described translational regulators YfmR, UuP or YebC, either individually or in combination with *efp* deletion (Figure 8F). This would provide evidence regarding the potential for these translational regulators to compensate for the absence of one another in the coverage of the motif spectra rescued. Ultimately, the knowledge about the motif spectra rescued by these translational factors will enable their targeted application in protein production in synthetic and biotechnological contexts.

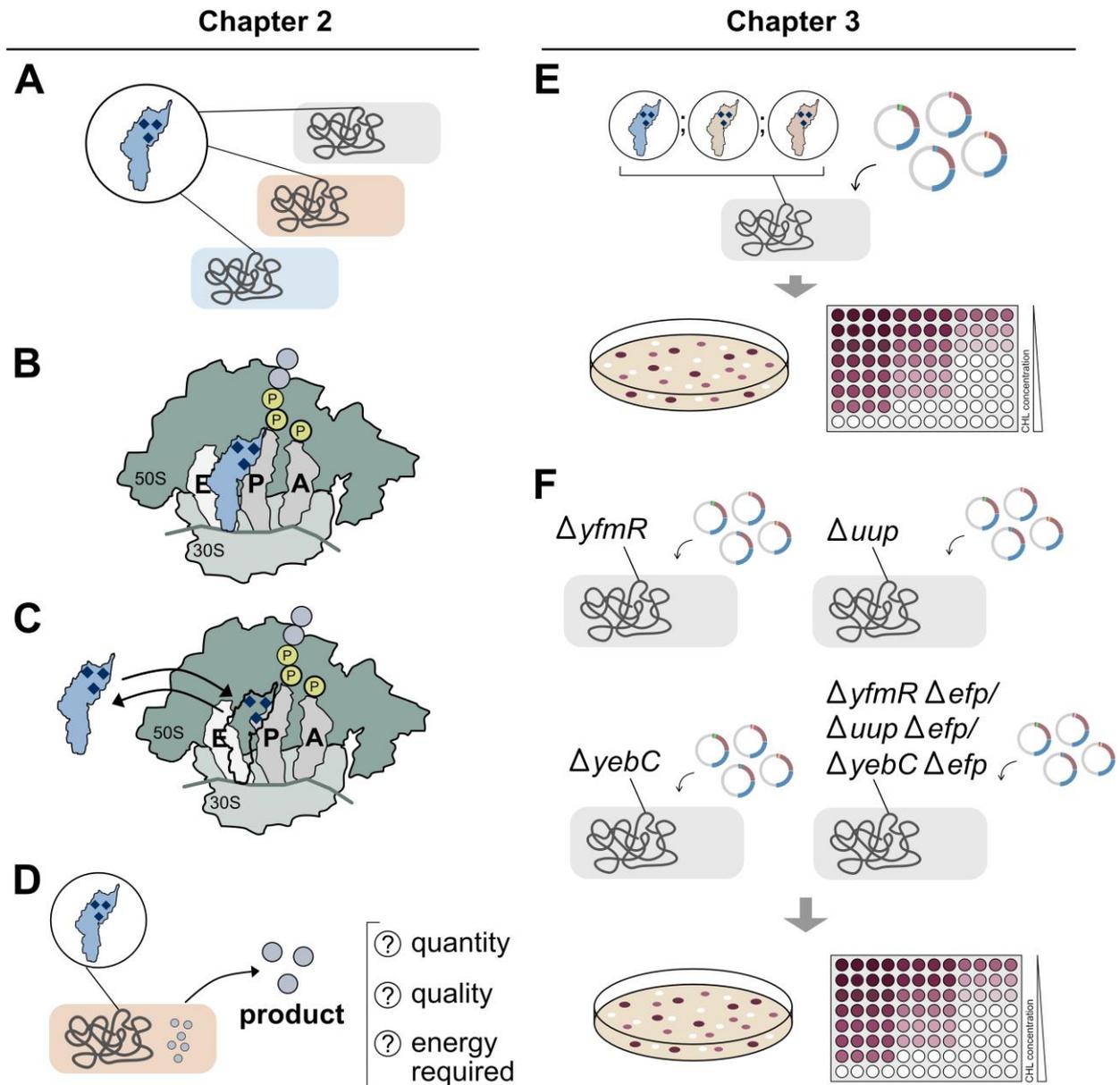


Figure 8. Schematic illustration of future research opportunities. The illustration depicts following future experiments based on data generated in Chapter 2 and Chapter 3: (A) unmodified EF-P variant expression in other bacterial species, (B) structural studies (e.g. Cryo-EM) with unmodified EF-P variants, (C) binding kinetic studies with unmodified EF-P variants, (D) studies on biotechnological compound production in bacterial strains with unmodified EF-P variants, (E) studies on translational rescue of polyproline-containing protein libraries by using the dual reporter system in strains encoding for unmodified EF-P variants, (F) studies on translational rescue of polyproline-containing protein libraries in strains lacking the genes coding for YfmR, UuP, YebC and EF-P.

References for Chapters 1 and 4

1. Crick, F.H. (1958) On protein synthesis. *Symp Soc Exp Biol*, **12**, 138-163.
2. Cobb, M. (2017) 60 years ago, Francis Crick changed the logic of biology. *PLoS Biol*, **15**, e2003243.
3. Tollerson, R., 2nd and Ibba, M. (2020) Translational regulation of environmental adaptation in bacteria. *J Biol Chem*, **295**, 10434-10445.
4. Advani, V.M. and Ivanov, P. (2019) Translational Control under Stress: Reshaping the Translatome. *Bioessays*, **41**, e1900009.
5. Ganoza, M.C., Kiel, M.C. and Aoki, H. (2002) Evolutionary conservation of reactions in translation. *Microbiol Mol Biol Rev*, **66**, 460-485, table of contents.
6. Ouzounis, C. and Kyrpides, N. (1996) The emergence of major cellular processes in evolution. *FEBS Lett*, **390**, 119-123.
7. Slobodin, B. and Dikstein, R. (2020) So close, no matter how far: multiple paths connecting transcription to mRNA translation in eukaryotes. *EMBO Rep*, **21**, e50799.
8. Stent, G.S. (1964) The Operon: On Its Third Anniversary. Modulation of Transfer Rna Species Can Provide a Workable Model of an Operator-Less Operon. *Science*, **144**, 816-820.
9. Irastortza-Olaziregi, M. and Amster-Choder, O. (2020) Coupled Transcription-Translation in Prokaryotes: An Old Couple With New Surprises. *Front Microbiol*, **11**, 624830.
10. Byrne, R., Levin, J.G., Bladen, H.A. and Nirenberg, M.W. (1964) The in Vitro Formation of a DNA-Ribosome Complex. *Proc Natl Acad Sci U S A*, **52**, 140-148.
11. Bladen, H.A., Byrne, R., Levin, J.G. and Nirenberg, M.W. (1965) An Electron Microscopic Study of a DNA-Ribosome Complex Formed *in Vitro*. *J Mol Biol*, **11**, 78-83.
12. Kohler, R., Mooney, R.A., Mills, D.J., Landick, R. and Cramer, P. (2017) Architecture of a transcribing-translating expressome. *Science*, **356**, 194-197.
13. Miller, O.L., Jr., Hamkalo, B.A. and Thomas, C.A., Jr. (1970) Visualization of bacterial genes in action. *Science*, **169**, 392-395.
14. Proshkin, S., Rahmouni, A.R., Mironov, A. and Nudler, E. (2010) Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science*, **328**, 504-508.
15. Landick, R., Carey, J. and Yanofsky, C. (1985) Translation activates the paused transcription complex and restores transcription of the trp operon leader region. *Proc Natl Acad Sci U S A*, **82**, 4663-4667.
16. Shetty, S. and Varshney, U. (2021) Regulation of translation by one-carbon metabolism in bacteria and eukaryotic organelles. *J Biol Chem*, **296**, 100088.
17. Pedersen, S. (1984) *Escherichia coli* ribosomes translate *in vivo* with variable rate. *EMBO J*, **3**, 2895-2898.
18. Liang, S.T., Xu, Y.C., Dennis, P. and Bremer, H. (2000) mRNA composition and control of bacterial gene expression. *J Bacteriol*, **182**, 3037-3044.
19. Hummels, K.R. and Kearns, D.B. (2020) Translation elongation factor P (EF-P). *FEMS Microbiol Rev*, **44**, 208-218.
20. Schmitt, E., Panvert, M., Blanquet, S. and Mechulam, Y. (1998) Crystal structure of methionyl-tRNA^{Met} transformylase complexed with the initiator formyl-methionyl-tRNA^{Met}. *EMBO J*, **17**, 6819-6826.
21. Gualerzi, C.O. and Pon, C.L. (2015) Initiation of mRNA translation in bacteria: structural and dynamic aspects. *Cell Mol Life Sci*, **72**, 4341-4367.

22. Milon, P. and Rodnina, M.V. (2012) Kinetic control of translation initiation in bacteria. *Crit Rev Biochem Mol Biol*, **47**, 334-348.
23. Rodnina, M.V. (2018) Translation in Prokaryotes. *Cold Spring Harb Perspect Biol*, **10**.
24. Schmeing, T.M. and Ramakrishnan, V. (2009) What recent ribosome structures have revealed about the mechanism of translation. *Nature*, **461**, 1234-1242.
25. Duval, M., Simonetti, A., Caldelari, I. and Marzi, S. (2015) Multiple ways to regulate translation initiation in bacteria: Mechanisms, regulatory circuits, dynamics. *Biochimie*, **114**, 18-29.
26. Milon, P., Maracci, C., Filonava, L., Gualerzi, C.O. and Rodnina, M.V. (2012) Real-time assembly landscape of bacterial 30S translation initiation complex. *Nat Struct Mol Biol*, **19**, 609-615.
27. Goyal, A., Belardinelli, R., Maracci, C., Milon, P. and Rodnina, M.V. (2015) Directional transition from initiation to elongation in bacterial translation. *Nucleic Acids Res*, **43**, 10700-10712.
28. Gualerzi, C., Risuleo, G. and Pon, C.L. (1977) Initial rate kinetic analysis of the mechanism of initiation complex formation and the role of initiation factor IF-3. *Biochemistry*, **16**, 1684-1689.
29. Milon, P., Konevega, A.L., Gualerzi, C.O. and Rodnina, M.V. (2008) Kinetic checkpoint at a late step in translation initiation. *Mol Cell*, **30**, 712-720.
30. Milon, P., Carotti, M., Konevega, A.L., Wintermeyer, W., Rodnina, M.V. and Gualerzi, C.O. (2010) The ribosome-bound initiation factor 2 recruits initiator tRNA to the 30S initiation complex. *EMBO Rep*, **11**, 312-316.
31. Carter, A.P., Clemons, W.M., Jr., Brodersen, D.E., Morgan-Warren, R.J., Hartsch, T., Wimberly, B.T. and Ramakrishnan, V. (2001) Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science*, **291**, 498-501.
32. Dahlquist, K.D. and Puglisi, J.D. (2000) Interaction of translation initiation factor IF1 with the *E. coli* ribosomal A site. *J Mol Biol*, **299**, 1-15.
33. Guenneugues, M., Caserta, E., Brandi, L., Spurio, R., Meunier, S., Pon, C.L., Boelens, R. and Gualerzi, C.O. (2000) Mapping the fMet-tRNA(f)(Met) binding site of initiation factor IF2. *EMBO J*, **19**, 5233-5240.
34. Simonetti, A., Marzi, S., Myasnikov, A.G., Fabbretti, A., Yusupov, M., Gualerzi, C.O. and Klaholz, B.P. (2008) Structure of the 30S translation initiation complex. *Nature*, **455**, 416-420.
35. Julian, P., Milon, P., Agirrezabala, X., Lasso, G., Gil, D., Rodnina, M.V. and Valle, M. (2011) The Cryo-EM structure of a complete 30S translation initiation complex from *Escherichia coli*. *PLoS Biol*, **9**, e1001095.
36. Hussain, T., Llacer, J.L., Wimberly, B.T., Kieft, J.S. and Ramakrishnan, V. (2016) Large-Scale Movements of IF3 and tRNA during Bacterial Translation Initiation. *Cell*, **167**, 133-144 e113.
37. Canonaco, M.A., Calogero, R.A. and Gualerzi, C.O. (1986) Mechanism of translational initiation in prokaryotes. Evidence for a direct effect of IF2 on the activity of the 30 S ribosomal subunit. *FEBS Lett*, **207**, 198-204.
38. Antoun, A., Pavlov, M.Y., Lovmar, M. and Ehrenberg, M. (2006) How initiation factors tune the rate of initiation of protein synthesis in bacteria. *EMBO J*, **25**, 2539-2550.
39. Antoun, A., Pavlov, M.Y., Lovmar, M. and Ehrenberg, M. (2006) How initiation factors maximize the accuracy of tRNA selection in initiation of bacterial protein synthesis. *Mol Cell*, **23**, 183-193.
40. Grigoriadou, C., Marzi, S., Kirillov, S., Gualerzi, C.O. and Cooperman, B.S. (2007) A quantitative kinetic scheme for 70 S translation initiation complex formation. *J Mol Biol*, **373**, 562-572.
41. Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H. and Noller, H.F. (2001) Crystal structure of the ribosome at 5.5 Å resolution. *Science*, **292**, 883-896.
42. Nirenberg, M.W. and Matthaei, J.H. (1961) The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc Natl Acad Sci U S A*, **47**, 1588-1602.

43. Nirenberg, M., Leder, P., Bernfield, M., Brimacombe, R., Trupin, J., Rottman, F. and O'Neal, C. (1965) RNA codewords and protein synthesis, VII. On the general nature of the RNA code. *Proc Natl Acad Sci U S A*, **53**, 1161-1168.
44. Forchhammer, K., Leinfelder, W. and Bock, A. (1989) Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. *Nature*, **342**, 453-456.
45. Kothe, U., Wieden, H.J., Mohr, D. and Rodnina, M.V. (2004) Interaction of helix D of elongation factor Tu with helices 4 and 5 of protein L7/12 on the ribosome. *J Mol Biol*, **336**, 1011-1021.
46. Rodnina, M.V., Pape, T., Fricke, R., Kuhn, L. and Wintermeyer, W. (1996) Initial binding of the elongation factor Tu.GTP.aminoacyl-tRNA complex preceding codon recognition on the ribosome. *J Biol Chem*, **271**, 646-652.
47. Rodnina, M.V. and Wintermeyer, W. (2001) Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. *Annu Rev Biochem*, **70**, 415-435.
48. Diaconu, M., Kothe, U., Schlunzen, F., Fischer, N., Harms, J.M., Tonevitsky, A.G., Stark, H., Rodnina, M.V. and Wahl, M.C. (2005) Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation. *Cell*, **121**, 991-1004.
49. Gromadski, K.B. and Rodnina, M.V. (2004) Kinetic determinants of high-fidelity tRNA discrimination on the ribosome. *Mol Cell*, **13**, 191-200.
50. Gromadski, K.B., Daviter, T. and Rodnina, M.V. (2006) A uniform response to mismatches in codon-anticodon complexes ensures ribosomal fidelity. *Mol Cell*, **21**, 369-377.
51. Rodnina, M.V., Fricke, R., Kuhn, L. and Wintermeyer, W. (1995) Codon-dependent conformational change of elongation factor Tu preceding GTP hydrolysis on the ribosome. *EMBO J*, **14**, 2613-2619.
52. Pape, T., Wintermeyer, W. and Rodnina, M.V. (1998) Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the *E. coli* ribosome. *EMBO J*, **17**, 7490-7497.
53. Pape, T., Wintermeyer, W. and Rodnina, M. (1999) Induced fit in initial selection and proofreading of aminoacyl-tRNA on the ribosome. *EMBO J*, **18**, 3800-3807.
54. Blanchard, S.C., Gonzalez, R.L., Kim, H.D., Chu, S. and Puglisi, J.D. (2004) tRNA selection and kinetic proofreading in translation. *Nat Struct Mol Biol*, **11**, 1008-1014.
55. Kothe, U. and Rodnina, M.V. (2006) Delayed release of inorganic phosphate from elongation factor Tu following GTP hydrolysis on the ribosome. *Biochemistry*, **45**, 12767-12774.
56. Liu, W., Chen, C., Kavaliauskas, D., Knudsen, C.R., Goldman, Y.E. and Cooperman, B.S. (2015) EF-Tu dynamics during pre-translocation complex formation: EF-Tu.GDP exits the ribosome via two different pathways. *Nucleic Acids Res*, **43**, 9519-9528.
57. Beringer, M. and Rodnina, M.V. (2007) The ribosomal peptidyl transferase. *Mol Cell*, **26**, 311-321.
58. Xu, B., Liu, L. and Song, G. (2021) Functions and Regulation of Translation Elongation Factors. *Front Mol Biosci*, **8**, 816398.
59. Rodnina, M.V., Savelsbergh, A., Katunin, V.I. and Wintermeyer, W. (1997) Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature*, **385**, 37-41.
60. Rodnina, M.V., Peske, F., Peng, B.Z., Belardinelli, R. and Wintermeyer, W. (2019) Converting GTP hydrolysis into motion: versatile translational elongation factor G. *Biol Chem*, **401**, 131-142.
61. Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N. and Jung, K. (2013) Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science*, **339**, 82-85.
62. Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H. and Rodnina, M.V. (2013) EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science*, **339**, 85-88.

63. Brenner, S., Stretton, A.O. and Kaplan, S. (1965) Genetic code: the 'nonsense' triplets for chain termination and their suppression. *Nature*, **206**, 994-998.
64. Weigert, M.G. and Garen, A. (1965) Base composition of nonsense codons in *E. coli*. Evidence from amino-acid substitutions at a tryptophan site in alkaline phosphatase. *Nature*, **206**, 992-994.
65. Capecchi, M.R. (1967) Polypeptide chain termination in vitro: isolation of a release factor. *Proc Natl Acad Sci U S A*, **58**, 1144-1151.
66. Scolnick, E., Tompkins, R., Caskey, T. and Nirenberg, M. (1968) Release factors differing in specificity for terminator codons. *Proc Natl Acad Sci U S A*, **61**, 768-774.
67. Milman, G., Goldstein, J., Scolnick, E. and Caskey, T. (1969) Peptide chain termination. 3. Stimulation of *in vitro* termination. *Proc Natl Acad Sci U S A*, **63**, 183-190.
68. Scolnick, E.M. and Caskey, C.T. (1969) Peptide chain termination. V. The role of release factors in mRNA terminator codon recognition. *Proc Natl Acad Sci U S A*, **64**, 1235-1241.
69. Korostelev, A., Asahara, H., Lancaster, L., Laurberg, M., Hirschi, A., Zhu, J., Trakhanov, S., Scott, W.G. and Noller, H.F. (2008) Crystal structure of a translation termination complex formed with release factor RF2. *Proc Natl Acad Sci U S A*, **105**, 19684-19689.
70. Korostelev, A., Zhu, J., Asahara, H. and Noller, H.F. (2010) Recognition of the amber UAG stop codon by release factor RF1. *EMBO J*, **29**, 2577-2585.
71. Laurberg, M., Asahara, H., Korostelev, A., Zhu, J., Trakhanov, S. and Noller, H.F. (2008) Structural basis for translation termination on the 70S ribosome. *Nature*, **454**, 852-857.
72. Weixlbaumer, A., Jin, H., Neubauer, C., Voorhees, R.M., Petry, S., Kelley, A.C. and Ramakrishnan, V. (2008) Insights into translational termination from the structure of RF2 bound to the ribosome. *Science*, **322**, 953-956.
73. Santos, N., Zhu, J., Donohue, J.P., Korostelev, A.A. and Noller, H.F. (2013) Crystal structure of the 70S ribosome bound with the Q253P mutant form of release factor RF2. *Structure*, **21**, 1258-1263.
74. Frolova, L.Y., Tsivkovskii, R.Y., Sivolobova, G.F., Oparina, N.Y., Serpinsky, O.I., Blinov, V.M., Tatkov, S.I. and Kisselev, L.L. (1999) Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. *RNA*, **5**, 1014-1020.
75. Zavialov, A.V., Mora, L., Buckingham, R.H. and Ehrenberg, M. (2002) Release of peptide promoted by the GGQ motif of class 1 release factors regulates the GTPase activity of RF3. *Mol Cell*, **10**, 789-798.
76. Mora, L., Heurgue-Hamard, V., Champ, S., Ehrenberg, M., Kisselev, L.L. and Buckingham, R.H. (2003) The essential role of the invariant GGQ motif in the function and stability *in vivo* of bacterial release factors RF1 and RF2. *Mol Microbiol*, **47**, 267-275.
77. Shaw, J.J. and Green, R. (2007) Two distinct components of release factor function uncovered by nucleophile partitioning analysis. *Mol Cell*, **28**, 458-467.
78. Freistroffer, D.V., Pavlov, M.Y., MacDougall, J., Buckingham, R.H. and Ehrenberg, M. (1997) Release factor RF3 in *E.coli* accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner. *EMBO J*, **16**, 4126-4133.
79. Koutmou, K.S., McDonald, M.E., Brunelle, J.L. and Green, R. (2014) RF3:GTP promotes rapid dissociation of the class 1 termination factor. *RNA*, **20**, 609-620.
80. Janosi, L., Hara, H., Zhang, S. and Kaji, A. (1996) Ribosome recycling by ribosome recycling factor (RRF)--an important but overlooked step of protein biosynthesis. *Adv Biophys*, **32**, 121-201.
81. Karimi, R., Pavlov, M.Y., Buckingham, R.H. and Ehrenberg, M. (1999) Novel roles for classical factors at the interface between translation termination and initiation. *Mol Cell*, **3**, 601-609.

82. Gao, N., Zavialov, A.V., Li, W., Sengupta, J., Valle, M., Gursky, R.P., Ehrenberg, M. and Frank, J. (2005) Mechanism for the disassembly of the posttermination complex inferred from cryo-EM studies. *Mol Cell*, **18**, 663-674.
83. Brion, P. and Westhof, E. (1997) Hierarchy and dynamics of RNA folding. *Annu Rev Biophys Biomol Struct*, **26**, 113-137.
84. Batey, R.T., Rambo, R.P. and Doudna, J.A. (1999) Tertiary Motifs in RNA Structure and Folding. *Angew Chem Int Ed Engl*, **38**, 2326-2343.
85. Pleij, C.W., Rietveld, K. and Bosch, L. (1985) A new principle of RNA folding based on pseudoknotting. *Nucleic Acids Res*, **13**, 1717-1731.
86. Mearls, E.B., Jackter, J., Colquhoun, J.M., Farmer, V., Matthews, A.J., Murphy, L.S., Fenton, C. and Camp, A.H. (2018) Transcription and translation of the sigG gene is tuned for proper execution of the switch from early to late gene expression in the developing *Bacillus subtilis* spore. *PLoS Genet*, **14**, e1007350.
87. Masachis, S. and Darfeuille, F. (2018) Type I Toxin-Antitoxin Systems: Regulating Toxin Expression via Shine-Dalgarno Sequence Sequestration and Small RNA Binding. *Microbiol Spectr*, **6**.
88. Masachis, S., Tourasse, N.J., Lays, C., Faucher, M., Chabas, S., Iost, I. and Darfeuille, F. (2019) A genetic selection reveals functional metastable structures embedded in a toxin-encoding mRNA. *Elife*, **8**.
89. Dreyfus, M. (1988) What constitutes the signal for the initiation of protein synthesis on *Escherichia coli* mRNAs? *J Mol Biol*, **204**, 79-94.
90. Shine, J. and Dalgarno, L. (1974) The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci U S A*, **71**, 1342-1346.
91. Steitz, J.A. and Jakes, K. (1975) How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proc Natl Acad Sci U S A*, **72**, 4734-4738.
92. Kozak, M. (1981) Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res*, **9**, 5233-5252.
93. Kozak, M. (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res*, **15**, 8125-8148.
94. Kozak, M. (2002) Pushing the limits of the scanning mechanism for initiation of translation. *Gene*, **299**, 1-34.
95. Vellanoweth, R.L. and Rabinowitz, J.C. (1992) The influence of ribosome-binding-site elements on translational efficiency in *Bacillus subtilis* and *Escherichia coli* *in vivo*. *Mol Microbiol*, **6**, 1105-1114.
96. Studer, S.M. and Joseph, S. (2006) Unfolding of mRNA secondary structure by the bacterial translation initiation complex. *Mol Cell*, **22**, 105-115.
97. Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene*, **234**, 187-208.
98. Ma, J., Campbell, A. and Karlin, S. (2002) Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures. *J Bacteriol*, **184**, 5733-5745.
99. Nakagawa, S., Niimura, Y., Miura, K. and Gojobori, T. (2010) Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proc Natl Acad Sci U S A*, **107**, 6382-6387.
100. Chang, B., Halgamuge, S. and Tang, S.L. (2006) Analysis of SD sequences in completed microbial genomes: non-SD-led genes are as common as SD-led genes. *Gene*, **373**, 90-99.

101. Tolstrup, N., Sensen, C.W., Garrett, R.A. and Clausen, I.G. (2000) Two different and highly organized mechanisms of translation initiation in the archaeon *Sulfolobus solfataricus*. *Extremophiles*, **4**, 175-179.
102. Weiner, J., 3rd, Herrmann, R. and Browning, G.F. (2000) Transcription in *Mycoplasma pneumoniae*. *Nucleic Acids Res*, **28**, 4488-4496.
103. Espah Borujeni, A., Channarasappa, A.S. and Salis, H.M. (2014) Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res*, **42**, 2646-2659.
104. Hockenberry, A.J., Pah, A.R., Jewett, M.C. and Amaral, L.A. (2017) Leveraging genome-wide datasets to quantify the functional role of the anti-Shine-Dalgarno sequence in regulating translation efficiency. *Open Biol*, **7**.
105. de Smit, M.H. and van Duin, J. (1994) Translational initiation on structured messengers. Another role for the Shine-Dalgarno interaction. *J Mol Biol*, **235**, 173-184.
106. Salis, H.M., Mirsky, E.A. and Voigt, C.A. (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol*, **27**, 946-950.
107. Espah Borujeni, A., Cetnar, D., Farasat, I., Smith, A., Lundgren, N. and Salis, H.M. (2017) Precise quantification of translation inhibition by mRNA structures that overlap with the ribosomal footprint in N-terminal coding sequences. *Nucleic Acids Res*, **45**, 5437-5448.
108. Frumkin, I., Lajoie, M.J., Gregg, C.J., Hornung, G., Church, G.M. and Pilpel, Y. (2018) Codon usage of highly expressed genes affects proteome-wide translation efficiency. *Proc Natl Acad Sci U S A*, **115**, E4940-E4949.
109. Kudla, G., Murray, A.W., Tollervey, D. and Plotkin, J.B. (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. *Science*, **324**, 255-258.
110. Liu, Y., Yang, Q. and Zhao, F. (2021) Synonymous but Not Silent: The Codon Usage Code for Gene Expression and Protein Folding. *Annu Rev Biochem*, **90**, 375-401.
111. Ikemura, T. (1985) Codon usage and tRNA content in unicellular and multicellular organisms. *Mol Biol Evol*, **2**, 13-34.
112. Plotkin, J.B. and Kudla, G. (2011) Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet*, **12**, 32-42.
113. Emery, L.R. and Sharp, P.M. (2011) Impact of translational selection on codon usage bias in the archaeon *Methanococcus maripaludis*. *Biol Lett*, **7**, 131-135.
114. Sabi, R. and Tuller, T. (2014) Modelling the efficiency of codon-tRNA interactions based on codon usage bias. *DNA Res*, **21**, 511-526.
115. Spencer, P.S., Siller, E., Anderson, J.F. and Barral, J.M. (2012) Silent substitutions predictably alter translation elongation rates and protein folding efficiencies. *J Mol Biol*, **422**, 328-335.
116. Mohammad, F., Woolstenhulme, C.J., Green, R. and Buskirk, A.R. (2016) Clarifying the Translational Pausing Landscape in Bacteria by Ribosome Profiling. *Cell Rep*, **14**, 686-694.
117. Samatova, E., Daberger, J., Liutkute, M. and Rodnina, M.V. (2020) Translational Control by Ribosome Pausing in Bacteria: How a Non-uniform Pace of Translation Affects Protein Production and Folding. *Front Microbiol*, **11**, 619430.
118. Chen, G.T. and Inouye, M. (1994) Role of the AGA/AGG codons, the rarest codons in global gene expression in *Escherichia coli*. *Genes Dev*, **8**, 2641-2652.
119. Lakey, D.L., Voladri, R.K., Edwards, K.M., Hager, C., Samten, B., Wallis, R.S., Barnes, P.F. and Kernodle, D.S. (2000) Enhanced production of recombinant *Mycobacterium tuberculosis* antigens in *Escherichia coli* by replacement of low-usage codons. *Infect Immun*, **68**, 233-238.
120. Stenstrom, C.M., Jin, H., Major, L.L., Tate, W.P. and Isaksson, L.A. (2001) Codon bias at the 3'-side of the initiation codon is correlated with translation initiation efficiency in *Escherichia coli*. *Gene*, **263**, 273-284.

121. Liu, Y. (2020) A code within the genetic code: codon usage regulates co-translational protein folding. *Cell Commun Signal*, **18**, 145.
122. Pechmann, S. and Frydman, J. (2013) Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat Struct Mol Biol*, **20**, 237-243.
123. Zhang, G. and Ignatova, Z. (2011) Folding at the birth of the nascent chain: coordinating translation with co-translational folding. *Curr Opin Struct Biol*, **21**, 25-31.
124. Komar, A.A. and Jaenicke, R. (1995) Kinetics of translation of gamma B crystallin and its circularly permuted variant in an in vitro cell-free system: possible relations to codon distribution and protein folding. *FEBS Lett*, **376**, 195-198.
125. O'Brien, E.P., Vendruscolo, M. and Dobson, C.M. (2012) Prediction of variable translation rate effects on cotranslational protein folding. *Nat Commun*, **3**, 868.
126. O'Brien, E.P., Vendruscolo, M. and Dobson, C.M. (2014) Kinetic modelling indicates that fast-translating codons can coordinate cotranslational protein folding by avoiding misfolded intermediates. *Nat Commun*, **5**, 2988.
127. Pavlov, M.Y., Watts, R.E., Tan, Z., Cornish, V.W., Ehrenberg, M. and Forster, A.C. (2009) Slow peptide bond formation by proline and other N-alkylamino acids in translation. *Proc Natl Acad Sci U S A*, **106**, 50-54.
128. Muto, H. and Ito, K. (2008) Peptidyl-prolyl-tRNA at the ribosomal P-site reacts poorly with puromycin. *Biochem Biophys Res Commun*, **366**, 1043-1047.
129. Wohlgemuth, I., Brenner, S., Beringer, M. and Rodnina, M.V. (2008) Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. *J Biol Chem*, **283**, 32229-32235.
130. Tanner, D.R., Cariello, D.A., Woolstenhulme, C.J., Broadbent, M.A. and Buskirk, A.R. (2009) Genetic identification of nascent peptides that induce ribosome stalling. *J Biol Chem*, **284**, 34809-34818.
131. Peil, L., Starosta, A.L., Lassak, J., Atkinson, G.C., Virumae, K., Spitzer, M., Tenson, T., Jung, K., Remme, J. and Wilson, D.N. (2013) Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. *Proc Natl Acad Sci U S A*, **110**, 15265-15270.
132. Rath, A., Davidson, A.R. and Deber, C.M. (2005) The structure of "unstructured" regions in peptides and proteins: role of the polyproline II helix in protein folding and recognition. *Biopolymers*, **80**, 179-185.
133. Adzhubei, A.A., Sternberg, M.J. and Makarov, A.A. (2013) Polyproline-II helix in proteins: structure and function. *J Mol Biol*, **425**, 2100-2132.
134. Qi, F., Motz, M., Jung, K., Lassak, J. and Frishman, D. (2018) Evolutionary analysis of polyproline motifs in *Escherichia coli* reveals their regulatory role in translation. *PLoS Comput Biol*, **14**, e1005987.
135. Glick, B.R. and Ganoza, M.C. (1975) Identification of a soluble protein that stimulates peptide bond synthesis. *Proc Natl Acad Sci U S A*, **72**, 4257-4260.
136. Glick, B.R., Chladek, S. and Ganoza, M.C. (1979) Peptide bond formation stimulated by protein synthesis factor EF-P depends on the aminoacyl moiety of the acceptor. *Eur J Biochem*, **97**, 23-28.
137. Anderson, W.F., Bosch, L., Cohn, W.E., Lodish, H., Merrick, W.C., Weissbach, H., Wittmann, H.G. and Wool, I.G. (1977) International symposium on protein synthesis. Summary of Fogarty Center-NIH Workshop held in Bethesda, Maryland on 18-20 October, 1976. *FEBS Lett*, **76**, 1-10.
138. Prichard, P.M., Gilbert, J.M., Shafritz, D.A. and Anderson, W.F. (1970) Factors for the initiation of haemoglobin synthesis by rabbit reticulocyte ribosomes. *Nature*, **226**, 511-514.

139. Kemper, W.M., Berry, K.W. and Merrick, W.C. (1976) Purification and properties of rabbit reticulocyte protein synthesis initiation factors M2Balpha and M2Bbeta. *J Biol Chem*, **251**, 5551-5557.
140. Benne, R. and Hershey, J.W. (1978) The mechanism of action of protein synthesis initiation factors from rabbit reticulocytes. *J Biol Chem*, **253**, 3078-3087.
141. Benne, R., Brown-Luedi, M.L. and Hershey, J.W. (1978) Purification and characterization of protein synthesis initiation factors eIF-1, eIF-4C, eIF-4D, and eIF-5 from rabbit reticulocytes. *J Biol Chem*, **253**, 3070-3077.
142. Zanelli, C.F., Maragno, A.L., Gregio, A.P., Komili, S., Pandolfi, J.R., Mestriner, C.A., Lustri, W.R. and Valentini, S.R. (2006) eIF5A binds to translational machinery components and affects translation in yeast. *Biochem Biophys Res Commun*, **348**, 1358-1366.
143. Saini, P., Eyler, D.E., Green, R. and Dever, T.E. (2009) Hypusine-containing protein eIF5A promotes translation elongation. *Nature*, **459**, 118-121.
144. Kyrpides, N.C. and Woese, C.R. (1998) Universally conserved translation initiation factors. *Proc Natl Acad Sci U S A*, **95**, 224-228.
145. Hanawa-Suetsugu, K., Sekine, S., Sakai, H., Hori-Takemoto, C., Terada, T., Unzai, S., Tame, J.R., Kuramitsu, S., Shirouzu, M. and Yokoyama, S. (2004) Crystal structure of elongation factor P from *Thermus thermophilus* HB8. *Proc Natl Acad Sci U S A*, **101**, 9595-9600.
146. Blaha, G., Stanley, R.E. and Steitz, T.A. (2009) Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. *Science*, **325**, 966-970.
147. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H. and Ferrin, T.E. (2018) UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci*, **27**, 14-25.
148. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H. and Ferrin, T.E. (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci*, **30**, 70-82.
149. Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., Kim, J.R., Saini, P., Buskirk, A.R. and Dever, T.E. (2013) eIF5A promotes translation of polyproline motifs. *Mol Cell*, **51**, 35-45.
150. Choi, S. and Choe, J. (2011) Crystal structure of elongation factor P from *Pseudomonas aeruginosa* at 1.75 Å resolution. *Proteins*, **79**, 1688-1693.
151. Huter, P., Arenz, S., Bock, L.V., Graf, M., Frister, J.O., Heuer, A., Peil, L., Starosta, A.L., Wohlgemuth, I., Peske, F. *et al.* (2017) Structural Basis for Polyproline-Mediated Ribosome Stalling and Rescue by the Translation Elongation Factor EF-P. *Mol Cell*, **68**, 515-527 e516.
152. Schmidt, C., Becker, T., Heuer, A., Braunger, K., Shanmuganathan, V., Pech, M., Berninghausen, O., Wilson, D.N. and Beckmann, R. (2016) Structure of the hypusylated eukaryotic translation factor eIF-5A bound to the ribosome. *Nucleic Acids Res*, **44**, 1944-1951.
153. Yanagisawa, T., Sumida, T., Ishii, R., Takemoto, C. and Yokoyama, S. (2010) A paralog of lysyl-tRNA synthetase aminoacylates a conserved lysine residue in translation elongation factor P. *Nat Struct Mol Biol*, **17**, 1136-1143.
154. Zou, S.B., Hersch, S.J., Roy, H., Wiggers, J.B., Leung, A.S., Buranyi, S., Xie, J.L., Dare, K., Ibba, M. and Navarre, W.W. (2012) Loss of elongation factor P disrupts bacterial outer membrane integrity. *J Bacteriol*, **194**, 413-425.
155. Lassak, J., Keilhauer, E.C., Furst, M., Wuichet, K., Godeke, J., Starosta, A.L., Chen, J.M., Sogaard-Andersen, L., Rohr, J., Wilson, D.N. *et al.* (2015) Arginine-rhamnosylation as new strategy to activate translation elongation factor P. *Nat Chem Biol*, **11**, 266-270.
156. Yanagisawa, T., Takahashi, H., Suzuki, T., Masuda, A., Dohmae, N. and Yokoyama, S. (2016) *Neisseria meningitidis* Translation Elongation Factor P and Its Active-Site Arginine Residue Are Essential for Cell Viability. *PLoS One*, **11**, e0147907.

157. Klee, S.M., Sinn, J.P., Holmes, A.C., Lehman, B.L., Krawczyk, T., Peter, K.A. and McNellis, T.W. (2019) Extragenic Suppression of Elongation Factor P Gene Mutant Phenotypes in *Erwinia amylovora*. *J Bacteriol*, **201**.
158. Rajkovic, A., Erickson, S., Witzky, A., Branson, O.E., Seo, J., Gafken, P.R., Frietas, M.A., Whitelegge, J.P., Faull, K.F., Navarre, W. *et al.* (2015) Cyclic Rhamnosylated Elongation Factor P Establishes Antibiotic Resistance in *Pseudomonas aeruginosa*. *mBio*, **6**, e00823.
159. Navarre, W.W., Zou, S.B., Roy, H., Xie, J.L., Savchenko, A., Singer, A., Edvokimova, E., Prost, L.R., Kumar, R., Ibba, M. *et al.* (2010) PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. *Mol Cell*, **39**, 209-221.
160. Klee, S.M., Mostafa, I., Chen, S., Dufresne, C., Lehman, B.L., Sinn, J.P., Peter, K.A. and McNellis, T.W. (2018) An *Erwinia amylovora* yjeK mutant exhibits reduced virulence, increased chemical sensitivity and numerous environmentally dependent proteomic alterations. *Mol Plant Pathol*, **19**, 1667-1678.
161. Tollerson, R., 2nd, Witzky, A. and Ibba, M. (2018) Elongation factor P is required to maintain proteome homeostasis at high growth rate. *Proc Natl Acad Sci U S A*, **115**, 11072-11077.
162. Park, M.H., Cooper, H.L. and Folk, J.E. (1981) Identification of hypusine, an unusual amino acid, in a protein from human lymphocytes and of spermidine as its biosynthetic precursor. *Proc Natl Acad Sci U S A*, **78**, 2869-2873.
163. Park, M.H., Cooper, H.L. and Folk, J.E. (1982) The biosynthesis of protein-bound hypusine (N epsilon -(4-amino-2-hydroxybutyl)lysine). Lysine as the amino acid precursor and the intermediate role of deoxyhypusine (N epsilon -(4-aminobutyl)lysine). *J Biol Chem*, **257**, 7217-7222.
164. Cooper, H.L., Park, M.H., Folk, J.E., Safer, B. and Braverman, R. (1983) Identification of the hypusine-containing protein hy+ as translation initiation factor eIF-4D. *Proc Natl Acad Sci U S A*, **80**, 1854-1857.
165. Prunetti, L., Graf, M., Blaby, I.K., Peil, L., Makkay, A.M., Starosta, A.L., Papke, R.T., Oshima, T., Wilson, D.N. and de Crecy-Lagard, V. (2016) Deciphering the Translation Initiation Factor 5A Modification Pathway in Halophilic Archaea. *Archaea*, **2016**, 7316725.
166. Peil, L., Starosta, A.L., Virumae, K., Atkinson, G.C., Tenson, T., Remme, J. and Wilson, D.N. (2012) Lys34 of translation elongation factor EF-P is hydroxylated by YfcM. *Nat Chem Biol*, **8**, 695-697.
167. Roy, H., Zou, S.B., Bullwinkle, T.J., Wolfe, B.S., Gilreath, M.S., Forsyth, C.J., Navarre, W.W. and Ibba, M. (2011) The tRNA synthetase paralog PoxA modifies elongation factor-P with (R)-beta-lysine. *Nat Chem Biol*, **7**, 667-669.
168. Krawczyk, R., Macosek, J., Jagtap, P.K.A., Gast, D., Wunder, S., Mitra, P., Jha, A.K., Rohr, J., Hoffmann-Roder, A., Jung, K. *et al.* (2017) Structural Basis for EarP-Mediated Arginine Glycosylation of Translation Elongation Factor EF-P. *mBio*, **8**.
169. Rajkovic, A., Hummels, K.R., Witzky, A., Erickson, S., Gafken, P.R., Whitelegge, J.P., Faull, K.F., Kearns, D.B. and Ibba, M. (2016) Translation Control of Swarming Proficiency in *Bacillus subtilis* by 5-Amino-pentanolylated Elongation Factor P. *J Biol Chem*, **291**, 10976-10985.
170. Hummels, K.R., Witzky, A., Rajkovic, A., Tollerson, R., 2nd, Jones, L.A., Ibba, M. and Kearns, D.B. (2017) Carbonyl reduction by Ymfl in *Bacillus subtilis* prevents accumulation of an inhibitory EF-P modification state. *Mol Microbiol*, **106**, 236-251.
171. Bailly, M. and de Crecy-Lagard, V. (2010) Predicting the pathway involved in post-translational modification of elongation factor P in a subset of bacterial species. *Biol Direct*, **5**, 3.
172. Pinheiro, B., Scheidler, C.M., Kielkowski, P., Schmid, M., Forne, I., Ye, S., Reiling, N., Takano, E., Imhof, A., Sieber, S.A. *et al.* (2020) Structure and Function of an Elongation Factor P Subfamily in Actinobacteria. *Cell Rep*, **30**, 4332-4342 e4335.

173. Behshad, E., Ruzicka, F.J., Mansoorabadi, S.O., Chen, D., Reed, G.H. and Frey, P.A. (2006) Enantiomeric free radicals and enzymatic control of stereochemistry in a radical mechanism: the case of lysine 2,3-aminomutases. *Biochemistry*, **45**, 12639-12646.
174. Aoki, H., Xu, J., Emili, A., Chosay, J.G., Golshani, A. and Ganoza, M.C. (2008) Interactions of elongation factor EF-P with the *Escherichia coli* ribosome. *FEBS J*, **275**, 671-681.
175. Park, J.H., Johansson, H.E., Aoki, H., Huang, B.X., Kim, H.Y., Ganoza, M.C. and Park, M.H. (2012) Post-translational modification by beta-lysylation is required for activity of *Escherichia coli* elongation factor P (EF-P). *J Biol Chem*, **287**, 2579-2590.
176. Starosta, A.L., Lassak, J., Peil, L., Atkinson, G.C., Woolstenhulme, C.J., Virumae, K., Buskirk, A., Tenson, T., Remme, J., Jung, K. *et al.* (2014) A conserved proline triplet in Val-tRNA synthetase and the origin of elongation factor P. *Cell Rep*, **9**, 476-483.
177. Bullwinkle, T.J., Zou, S.B., Rajkovic, A., Hersch, S.J., Elgamal, S., Robinson, N., Smil, D., Bolshan, Y., Navarre, W.W. and Ibba, M. (2013) (R)-beta-lysine-modified elongation factor P functions in translation elongation. *J Biol Chem*, **288**, 4416-4423.
178. Witzky, A., Hummels, K.R., Tollerson, R., 2nd, Rajkovic, A., Jones, L.A., Kearns, D.B. and Ibba, M. (2018) EF-P Posttranslational Modification Has Variable Impact on Polyproline Translation in *Bacillus subtilis*. *mBio*, **9**.
179. Lewin, G.R., Carlos, C., Chevrette, M.G., Horn, H.A., McDonald, B.R., Stankey, R.J., Fox, B.G. and Currie, C.R. (2016) Evolution and Ecology of Actinobacteria and Their Bioenergy Applications. *Annu Rev Microbiol*, **70**, 235-254.
180. Selim, M.S.M., Abdelhamid, S.A. and Mohamed, S.S. (2021) Secondary metabolites and biodiversity of actinomycetes. *J Genet Eng Biotechnol*, **19**, 72.
181. Volkwein, W., Krafczyk, R., Jagtap, P.K.A., Parr, M., Mankina, E., Macosek, J., Guo, Z., Furst, M., Pfab, M., Frishman, D. *et al.* (2019) Switching the Post-translational Modification of Translation Elongation Factor EF-P. *Front Microbiol*, **10**, 1148.
182. Pfab, M., Kielkowski, P., Krafczyk, R., Volkwein, W., Sieber, S.A., Lassak, J. and Jung, K. (2021) Synthetic post-translational modifications of elongation factor P using the ligase EpmA. *FEBS J*, **288**, 663-677.
183. Hersch, S.J., Wang, M., Zou, S.B., Moon, K.M., Foster, L.J., Ibba, M. and Navarre, W.W. (2013) Divergent protein motifs direct elongation factor P-mediated translational regulation in *Salmonella enterica* and *Escherichia coli*. *mBio*, **4**, e00180-00113.
184. Starosta, A.L., Lassak, J., Peil, L., Atkinson, G.C., Virumae, K., Tenson, T., Remme, J., Jung, K. and Wilson, D.N. (2014) Translational stalling at polyproline stretches is modulated by the sequence context upstream of the stall site. *Nucleic Acids Res*, **42**, 10711-10719.
185. Elgamal, S., Katz, A., Hersch, S.J., Newsom, D., White, P., Navarre, W.W. and Ibba, M. (2014) EF-P dependent pauses integrate proximal and distal signals during translation. *PLoS Genet*, **10**, e1004553.
186. Woolstenhulme, C.J., Guydosh, N.R., Green, R. and Buskirk, A.R. (2015) High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. *Cell Rep*, **11**, 13-21.
187. Hersch, S.J., Elgamal, S., Katz, A., Ibba, M. and Navarre, W.W. (2014) Translation initiation rate determines the impact of ribosome stalling on bacterial protein synthesis. *J Biol Chem*, **289**, 28160-28171.
188. Lassak, J., Wilson, D.N. and Jung, K. (2016) Stall no more at polyproline stretches with the translation elongation factors EF-P and IF-5A. *Mol Microbiol*, **99**, 219-235.
189. Krafczyk, R., Qi, F., Sieber, A., Mehler, J., Jung, K., Frishman, D. and Lassak, J. (2021) Proline codon pair selection determines ribosome pausing strength and translation efficiency in bacteria. *Commun Biol*, **4**, 589.

190. Marman, H.E., Mey, A.R. and Payne, S.M. (2014) Elongation factor P and modifying enzyme PoxA are necessary for virulence of *Shigella flexneri*. *Infect Immun*, **82**, 3612-3621.
191. Guo, Q., Cui, B., Wang, M., Li, X., Tan, H., Song, S., Zhou, J., Zhang, L.H. and Deng, Y. (2022) Elongation factor P modulates *Acinetobacter baumannii* physiology and virulence as a cyclic dimeric guanosine monophosphate effector. *Proc Natl Acad Sci U S A*, **119**, e2209838119.
192. Choi, E., Choi, S., Nam, D., Park, S., Han, Y., Lee, J.S. and Lee, E.J. (2017) Elongation factor P restricts *Salmonella's* growth by controlling translation of a Mg(2+) transporter gene during infection. *Sci Rep*, **7**, 42098.
193. Hong, H.R., Prince, C.R., Tetreault, D.D., Wu, L. and Feaga, H.A. (2024) YfmR is a translation factor that prevents ribosome stalling and cell death in the absence of EF-P. *Proc Natl Acad Sci U S A*, **121**, e2314437121.
194. Takada, H., Fujiwara, K., Atkinson, G.C., Chiba, S. and Hauryliuk, V. (2024) Resolution of ribosomal stalling by EF-P and ABCF ATPases YfmR and YkpA/YbiT. *Nucleic Acids Res*, **52**, 9854-9866.
195. Ignatov, D., Shanmuganathan, V., Ahmed-Begrich, R., Alagesan, K., Frese, C.K., Wang, C., Krause, K. and Charpentier, E. (2024) Novel RNA-binding protein YebC enhances translation of proline-rich amino acid stretches in bacteria. *bioRxiv*, 2024.2008.2026.607280.
196. Chadani, Y., Yamanouchi, S., Uemura, E., Yamasaki, K., Niwa, T., Ikeda, T., Kurihara, M., Iwasaki, W. and Taguchi, H. (2024) The ABCF proteins in *Escherichia coli* individually cope with 'hard-to-translate' nascent peptide sequences. *Nucleic Acids Res*, **52**, 5825-5840.
197. Brewer, T.E. and Wagner, A. (2024) Horizontal Gene Transfer of a key Translation Factor and its Role in Polyproline Proteome Evolution. *Mol Biol Evol*, **41**.
198. Duchow, E. and Douglas, H.C. (1949) *Rhodomicrobium Vanniellii*, a New Photoheterotrophic Bacterium. *J Bacteriol*, **58**, 409-416.
199. Ainon, H., Tan, C.J. and Vikineswary, S. (2006) Biological Characterization of *Rhodomicrobium vanniellii* Isolated from a Hot Spring at Gadek, Malacca, Malaysia. *Malaysian Journal of Microbiology*.
200. Mudryi, V., Frister, J.O., Peng, B.Z., Wohlgemuth, I., Peske, F. and Rodnina, M.V. (2024) Kinetic mechanism and determinants of EF-P recruitment to translating ribosomes. *Nucleic Acids Res*.
201. Mohapatra, S., Choi, H., Ge, X., Sanyal, S. and Weisshaar, J.C. (2017) Spatial Distribution and Ribosome-Binding Dynamics of EF-P in Live *Escherichia coli*. *mBio*, **8**.
202. Peng, W.T., Banta, L.M., Charles, T.C. and Nester, E.W. (2001) The chvH locus of *Agrobacterium* encodes a homologue of an elongation factor involved in protein synthesis. *J Bacteriol*, **183**, 36-45.
203. Trachtmann, N., Bikmullin, A., Validov, S. and Sprenger, G.A. (2024) *Escherichia coli* Reporter Strains Allow for the In Vivo Evaluation of Recombinant Elongation Factor Protein (EF-P). *Applied Microbiology*, **4**, 1335-1347.

Grammarly (<https://www.grammarly.com>) was used for final grammatical corrections

Danksagung

Zunächst möchte ich allen, die mich auf dem Weg zu dieser Dissertation unterstützt haben, meinen aufrichtigen Dank aussprechen.

Mein besonderer Dank gilt Prof. Dr. Kirsten Jung, die mir die Möglichkeit gegeben hat, an einem so spannenden Projekt in ihrer Arbeitsgruppe zu forschen und mitzuwirken. Ihr Vertrauen, Ihre Offenheit für innovative Ansätze und Ihre kontinuierliche Unterstützung waren maßgeblich für die Entstehung dieser Arbeit. Ich möchte mich sehr herzlich für die exzellente Zusammenarbeit bedanken, die man sich nicht anders hätte wünschen können.

Anschließend möchte ich der Prüfungskommission meinen herzlichen Dank aussprechen, insbesondere Herrn Prof. Dr. Christof Osman, der sich bereit erklärt hat, das Zweitgutachten zu übernehmen.

Mein besonderer Dank gilt der Graduiertenschule (GRK2062) sowie der Deutschen Forschungsgemeinschaft (DFG) für die gebotenen Ressourcen, die entscheidend für die Verwirklichung dieser Arbeit waren. Den Kollaborationspartnern, Pavel Kielkovski, Ralph Krafczyk, Ignasi Forne, Prof. Dr. Axel Imhof, Tess Brewer, Korinna Burdack, Sophie Brameyer bin ich für die tolle wissenschaftliche Zusammenarbeit sehr dankbar. Meinen Studierenden, Thomas Krause und Lukas Voshagen, danke ich für ihre Begeisterung und ihren grandiosen Einsatz, der einen frischen Wind in die Projekte gebracht hat.

Durch ihre Kollegialität und kontinuierliche Unterstützung haben die Mitarbeitenden der AG K. Jung, AG H. Jung und AG Lassak ein Umfeld geschaffen, das von Produktivität und einem angenehmen Miteinander geprägt war, in dem ich mich sehr wohlfühlt habe. Ein besonderer Dank geht an die Organisationskünstlerin des Lehrstuhls, Grazyna; für ihre unermüdliche Unterstützung in der *Universitätsbürokratie* und die erstklassige Steuerberatung. Mein herzliches Dankeschön gilt auch Ingrid, Tatjana, Slavica, Jürgen und Frank, da sie eine unverzichtbare Bereicherung für die Gruppe waren/sind und das Labor sowie die Seminare "am Laufen gehalten haben". Besonders hervorheben möchte ich gerne meine *Laborbuddies* Korinna, Dania, Luti, die mir im Labor ein zweites Zuhause geschaffen haben und die zu wahren Freundinnen geworden sind. Elena (vielleicht nächstes Mal lieber *Tosca* statt *Electra*) und Nicola G. danke ich ganz besonders für die tolle Freundschaft, auch außerhalb des Labors. Für die lustigen Abende danke ich selbstverständlich meiner Substanzcrew - Korinna, Dania, Sophie und Sabine. Außerdem gilt mein besonderer Dank für die angenehme Zusammenarbeit meinen Büro- und Laborkollegen Steffi, Dimi, Miriam, Leo, Sebastian, Bibakhya und Agata. Meinen *EF-P-Freunden* danke ich für die wertvollen wissenschaftlichen Austausche und persönlichen Gespräche, vor allem Rebecca, Tamara, Erika, Giovanni, Alina und Judith. Ich möchte auch Herrn Professor Dr. Jung und den verbleibenden Mitgliedern der Arbeitsgruppen, besonders Julia, Ana, Jin, Kilian, Sophia, Nicola S., Larissa und Michelle, für die tolle wissenschaftliche Arbeitsatmosphäre danken.

Mein größtes Dankeschön für den außergewöhnlichen Rückhalt gilt natürlich meiner Familie, meinen Freunden und HS. Ohne euch wäre diese Arbeit nicht realisierbar gewesen. Ich danke euch für euer unermüdliche Motivation und eure enorme Unterstützung.