Strategies for synthetic activation of translation elongation factor P

DISSERTATION

der Fakultät für Biologie

der Ludwig-Maximilians-Universität München

vorgelegt von

Wolfram Johannes Volkwein

München, November 2019

Gutachter:

1. Prof. Dr. Jürgen Lassak

2. Prof. Dr. Andreas Klingl

Datum der Abgabe: 07.11.2019

Datum der mündlichen Prüfung: 10.01.2020

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt wurde. Des Weiteren erkläre ich, dass ich nicht anderweitig ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen. Die folgende Dissertation liegt weder ganz, noch in wesentlichen Teilen einer anderen Prüfungskommission vor.

München, 07.11.2019   Wolfram Volkwein

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources. As well I declare, that I have not submitted a dissertation without success and not passed the oral exam. The present dissertation (neither the entire dissertation nor parts) has not been presented to another examination board.

Munich, 07.11.2019   Wolfram Volkwein

Table of Content

Nomenclature

Amino acids are given in the one-letter code (e.g. K).

The positions of amino acids within a protein are indicated by numbers. The numbering starts with "1" at the first methionine/valine of the wild-type protein.

Amino acid substitutions are termed as follows: The native amino acid is designated in one letter code first, followed by its position in the protein, and the one-letter code for the amino acid substitution introduced. (Example: EF-P K34R, here, lysine at position 34 of EF-P is replaced by an arginine).

Deletions are marked with "Δ".

Abbreviations

AcKRS	acetyl lysyl-tRNA synthetase
A-site	aminoacyl-site
BTH	bacterial two-hybrid
β3Ωβ4	loop between beta-strands β3/β4
CAT	chloramphenicol acetyltransferase
EF-P	elongation factor P
eIF5A	eukaryotic translation initiation factor 5A
E-site	exit-site
fMet-tRNA$_i^{fMet}$	initiator transfer RNA N-formyl-methionyl-tRNA$_i$
GFP	green fluorescent protein
LysRS	lysyl-tRNA synthetase
mRNA	messenger RNA
ncAA	non-canonical amino acid
ORF	open reading frame
PTM	post-translational modification
PylRS	pyrrolysyl-tRNA synthetase
P-site	peptidyl-site
tRNA$_{CUA}$	pyrrolysyl-transfer RNA
tRNA^{Lys}	lysyl-transfer RNA

Publications and Manuscripts Originating from this Thesis

Chapter 2:

<u>Wolfram Volkwein</u>, Sabine Peschek, Benedikt Frederik Camille Graf von Armansperg, Kirsten Jung and Jürgen Lassak (2019). Strategy for the translational activation of EF-P using the amber suppression system. (Manuscript).

Chapter 3:

<u>Wolfram Volkwein</u>, Christopher Maier, Ralph Krafczyk, Kirsten Jung and Jürgen Lassak (2017). A versatile toolbox for the control of protein levels using N^ε-acetyl-L-lysine dependent amber suppression. ACS Synth. Biol. 6, 1892-1902.

Chapter 4:

<u>Wolfram Volkwein</u>*, Ralph Krafczyk*, Pravin Kumar Ankush Jagtap, Marina Parr, Elena Mankina, Jakub Macošek, Zhenghuan Guo, Maximilian Josef Ludwig Johannes Fürst, Miriam Pfab, Dmitrij Frishman, Janosch Hennig, Kirsten Jung and Jürgen Lassak (2019). Switching the post-translational modification of translation elongation factor EF-P. Front. Microbiol. 10.

*Authors contributed equally

Chapter 5:

Jürgen Lassak, Franziska Koller, Ralph Krafczyk and <u>Wolfram Volkwein</u> (2019). Exceptionally versatile – Arginine in bacterial post-translational protein modifications. Biol. Chem. 400, 1397-1427.

(Review Article)

Manuscripts in preparation:

Miriam Pfab, Pavel Kielkowski, Ralph Krafczyk, <u>Wolfram Volkwein</u>, Stephan Sieber, Jürgen Lassak and Kirsten Jung. Synthetic post-translational modifications of elongation factor P using the ligase EpmA.

Ralph Krafczyk, Fei Qi, Alina Sieber, <u>Wolfram Volkwein</u>, Kirsten Jung, Dmitrij Frishman and Jürgen Lassak. Codon choice affects translation efficiency of consecutive prolines in bacteria.

Contributions to Publications and Manuscripts Presented in this Thesis

Chapter 2:

The study was designed by Wolfram Volkwein, Kirsten Jung and Jürgen Lassak. The study was directed by Kirsten Jung and Jürgen Lassak. Wolfram Volkwein, Sabine Peschek and Jürgen Lassak constructed the strains and plasmids. Swimming assays were performed by Wolfram Volkwein, Sabine Peschek and Benedikt Frederik Camille Graf von Armansperg. Wolfram Volkwein and Jürgen Lassak wrote the manuscript.

Chapter 3:

The study was designed by Wolfram Volkwein, Kirsten Jung and Jürgen Lassak and directed as well as coordinated by Kirsten Jung and Jürgen Lassak. Wolfram Volkwein and Christopher Maier constructed all strains and performed the enzyme assays; Ralph Krafczyk performed the parameter optimization for the Lux reporters. All authors contributed to the writing of the manuscript.

Chapter 4:

Dmitrij Frishman, Marina Parr, Elena Mankina and Jürgen Lassak performed bioinformatic analyses. NMR studies were performed by Janosch Hennig, Pravin Kumar Ankush Jagtap and Jakub Macošek. The corresponding proteins were produced and purified by Wolfram Volkwein. Miriam Pfab performed isoelectric focusing experiments. Ralph Krafczyk and Zhenghuan Guo performed *in vitro* rhamnosylation assays. All other biochemical and genetic analyses of β3Ωβ4 substitution variants of *Escherichia coli* and *Pseudomonas putida* were conducted by Wolfram Volkwein and Ralph Krafczyk. Wolfram Volkwein and Maximilian Josef Ludwig Johannes Fürst performed the biochemical analysis with EarP from *Shewanella oneidensis* with contributions from Jürgen Lassak. Jürgen Lassak, Janosch Hennig, Kirsten Jung and Dmitrij Frishman designed the study. The manuscript was written by Wolfram Volkwein, Ralph Krafczyk, Kirsten Jung, Elena Mankina, Pravin Kumar Ankush Jagtap, Janosch Hennig and Jürgen Lassak.

Chapter 5:

Jürgen Lassak, Franziska Koller, Ralph Krafczyk and Wolfram Volkwein designed and discussed the concept of the review. Jürgen Lassak wrote the introduction and the sections about arginylation, advanced glycation end products and arginine citrullination. Franziska Koller wrote the section about ADP-ribosylation. Ralph Krafczyk wrote the sections about arginine glycosylation and arginine methylation. Wolfram Volkwein wrote the section about arginine phosphorylation.

We hereby confirm the above statements:

Wolfram Volkwein   Prof. Dr. Jürgen Lassak

Chapter 4:

In order to study the possibility of unnatural elongation factor P (EF-P) activation, Wolfram Volkwein and Ralph Krafczyk generated, and Ralph Krafczyk also applied a novel reporter strain for the sensitive *in vivo* detection of protein-protein interactions. This analysis showed that *Pseudomonas putida* EarP interacts both with its cognate EF-P and the sequentially unrelated homologue from *Escherichia coli*. Based on this information, Wolfram Volkwein performed extensive *in vivo* functionality analyses of single substitution variants and also produced and purified these variants for NMR studies. These experiments identified several amino acids that are crucial for activity of *P. putida* EF-P and subsequently enabled *in vivo* rhamnosylation - and thus unnatural activation - of *E. coli* EF-P. Complemented by *in vitro* rhamnosylation assays that were established and performed by Ralph Krafczyk, the minimal sequential requirements for modification and activation of elongation factor P by rhamnosylation were determined. Together, Wolfram Volkwein and Ralph Krafczyk managed to demonstrate, that the post-translational modification of *E. coli* EF-P can be switched from lysylation to rhamnosylation and that this unnatural modification can activate the elongation factor.

We hereby confirm the above statements:

Wolfram Volkwein   Dr. Ralph Krafczyk

Summary

During the synthesis of two or more consecutive prolines, translation comes to a halt. To resolve this arrest, bacteria have evolved a specialized translation elongation factor P (EF-P), being orthologous and functionally equivalent to the archaeal/eukaryotic initiation factor a/eIF5A. EF-P binds to the stalled ribosome in order to stabilize and orient the P-site prolyl-tRNA in a way that translation can resume. In this regard, a seven amino acids long loop between beta-strands β3/β4 (β3Ωβ4) is crucial for EF-P function and modified at its tip. In the Gram-negative model organism *Escherichia coli* and around 25% of all bacteria, the ligase EpmA catalyzes formation of an unusual peptide bond between the α-carboxyl group of (*R*)-β-lysine and the ε-amino group of EF-P lysine at position 34 (K34). Another 10% including *Pseudomonads* have evolved a glycosyltransferase termed EarP, which rhamnosylates an arginine at position 32 (R32) in the K34 equivalent position. As the requirements allowing these two chemically distinct modifications to fulfil the same function are unknown, this thesis was intended to shed a light on this so far neglected aspect. To do so, the following **translational** and **post-translational** strategies were developed and applied to modify EF-P synthetically:

With the first strategy, we aimed to incorporate (*R*)-β-lysyl-lysine in a single **translational** step directly, using the pyrrolysyl-tRNA synthetase (PylRS)/tRNA$_{CUA}$ pair dependent amber suppression system from *Methanosarcina mazei*. Therefore, *E. coli* EF-P would be synthetically functionalised in a post-translational independent manner. For this purpose, a screening system was initially developed, allowing to select for PylRS mutants which accept (*R*)-β-lysyl-lysine as a substrate (Chapter 2). In parallel and based on the acquired expertise regarding the amber suppression system, a toolbox to control bacterial protein levels was developed during this thesis (Chapter 3). This toolbox allows to control protein levels translationally, and when combined with transcriptional control, to gradually adjust the protein output in diverse Gram-negative bacteria. Furthermore, the potential of this system to regulate even essential gene products was demonstrated.

In the second strategy, we employed the modification system EarP from *Pseudomonas putida* to synthetically switch the **post-translational** modification of *E. coli* EF-P from (*R*)-β-lysylation to rhamnosylation (Chapter 4). We could show that a single amino acid exchange from lysine to arginine in the post-translational modification site is sufficient to switch the modification. Since the resulting *E. coli* EF-P variant is rhamnosylated but not functional, we furthermore investigated the influence of the amino acid composition of the β3Ωβ4 surrounding the modification site with respect to functionality. Based on these analyses we identified a conserved proline in *E. coli* EF-P located two amino acids upstream of the modification site which, when removed, restores functionality of rhamnosylated *E. coli* EF-P and therefore synthetically activates it.

Zusammenfassung

Bei der Synthese von zwei oder mehr aufeinanderfolgenden Prolinen kommt die Translation zum Stillstand. Um diesen Arrest aufzulösen, haben Bakterien einen speziellen Translationselongationsfaktor P (EF-P) entwickelt, welcher ortholog zu dem archaealen/eukaryotischen Initiationsfaktor a/eIF5A ist. EF-P bindet an das blockierte Ribosom und stabilisiert bzw. orientiert die P-site Prolyl-tRNA so, dass die Translation wiederaufgenommen werden kann. In diesem Zusammenhang ist eine sieben Aminosäuren lange Schleife entscheidend für die EF-P-Funktion. Sie liegt zwischen den Beta-Strängen β3/β4 (β3Ωβ4) und wird an deren Spitze modifiziert. Im Gram-negativen Modellorganismus *Escherichia coli* und rund 25% aller Bakterien katalysiert die Ligase EpmA die Bildung einer ungewöhnlichen Peptidbindung zwischen der α-Carboxylgruppe von (*R*)-β-Lysin und der ε-Aminogruppe von EF-P-Lysin in Position 34 (K34). Weitere 10% der Bakterien, einschließlich *Pseudomonaden*, haben eine Glykosyltransferase namens EarP entwickelt, die ein Arginin in Position 32 (R32) in der äquivalenten Position zu K34 rhamnosyliert. Da die Anforderungen unbekannt sind die es diesen beiden chemisch unterschiedlichen Modifikationen erlauben die gleiche Funktion zu erfüllen, war die vorliegende Arbeit dazu gedacht diesen bisher vernachlässigten Aspekt zu untersuchen. Zu diesem Zweck wurden folgende **translationale** und **posttranslationale** Strategien entwickelt und angewendet um EF-P synthetisch zu modifizieren:

In der ersten Strategie wurde versucht (*R*)-β-Lysyl-Lysin unter Verwendung des Pyrrolysyl-tRNA Synthetase (PylRS)/tRNACUA Paar abhängigen Amber-Suppressions-Systems von *Methanosarcina mazei* **translational** einzubauen. Somit würde *E. coli* EF-P in einem einzigen posttranslational unabhängigen Schritt synthetisch funktionalisiert werden. Um dies zu erreichen wurde initial ein Screening-System entwickelt welches es ermöglicht PylRS-Mutanten zu identifizieren welche (*R*)-β-Lysyl-Lysin als Substrat akzeptieren (Kapitel 2). Parallel hierzu, und auf Grundlage der erworbenen Expertise bezüglich des Amber-Suppressions-Systems, wurde im Rahmen dieser Arbeit ein Werkzeug entwickelt mit dessen Hilfe sich Proteinlevel in diversen Gram-negativen Bakterien steuern lassen (Kapitel 3). Dieses Werkzeug erlaubt, wenn kombiniert mit transkriptionaler Kontrolle, Proteinlevel graduell zu steuern und kann selbst auf essentielle Gene angewendet werden.

In der zweiten Strategie verwendeten wir das Modifikationssystem EarP von *Pseudomonas putida*, um **posttranslational** die Modifikation von *E. coli* EF-P von (*R*)-β-Lysylierung auf Rhamnosylierung umzustellen (Kapitel 4). Wir konnten zeigen, dass ein einzelner Aminosäureaustausch von Lysin zu Arginin an der posttranslationalen Modifikationsstelle ausreicht, um die Modifikation zu wechseln. Da jedoch die resultierende *E. coli* EF P-Variante rhamnosyliert aber nicht funktional war, wurde in einem nächsten Schritt der Einfluss der Aminosäurezusammensetzung der β3Ωβ4 auf die Funktionalität untersucht. Mit Hilfe dieser Analysen konnte ein konserviertes Prolin in *E. coli* EF-P identifiziert werden, welches sich zwei Positionen stromaufwärts bezogen auf die Modifikationsstelle befindet. Abschließend konnten wir

zeigen, dass durch den Austausch dieses Prolins rhamnosyliertes *E. coli* EF-P funktionalisiert und somit synthetisch aktiviert werden kann.

1. Introduction

1.1 The non-uniformity of translation

In all domains of life mRNAs are translated on the ribosomes into polypeptide chains in a highly conserved four step process. This process consists of: initiation, elongation, termination and ribosome recycling. During elongation individual amino acids are connected in the ribosomal peptidyl transferase center via peptide bonds to a polypeptide chain. To achieve this, the ribosome moves along the mRNA during elongation and this was initially seen as a uniform process. However, already in the late 1970s, early 1980s it became evident that this movement takes place rather in a non-uniform manner (Protzel and Morris, 1974; Chaney and Morris, 1978; Lizardi et al., 1979; Abraham and Pihl, 1980; Randall et al., 1980; Varenne et al., 1981; Yanofsky, 1981; Varenne et al., 1982; Varenne et al., 1984). One reason for this resides in the sequence composition (Wohlgemuth et al., 2008). This is not surprising when considering that peptide bond formation has to be achieved between 20 chemically diverse amino acids which allow for 400 different combinations. A peculiar case here represents peptidyl transferase reactions involving proline, due to its unique chemical properties among the proteinogenic amino acids. Its integral pyrrolidine ring gives proline an exceptional rigidity in structure which in turn makes it a poor substrate for peptide bond formation, as acceptor substrate in the ribosomal aminoacyl-site (A-site) (Pavlov et al., 2009) as well as donor substrate in the peptidyl-site (P-site) (Muto and Ito, 2008). In both cases this results in slow peptide bond formation (Wohlgemuth et al., 2008; Johansson et al., 2011; Doerfel et al., 2013; Chevance et al., 2014; Doerfel et al., 2015). This effect becomes particularly apparent when proline acts as acceptor and donor substrate at the same time, resulting in translational arrest of the ribosome during the synthesis of stretches of two or more consecutive prolines. (Tanner et al., 2009; Doerfel et al., 2013; Hersch et al., 2013; Peil et al., 2013; Ude et al., 2013; Woolstenhulme et al., 2013; Elgamal et al., 2014; Starosta et al., 2014; Woolstenhulme et al., 2015). However, despite this translational arrest, polyproline motifs are found extensively in the genomes throughout all domains of life (Mandal et al., 2014; Qi et al., 2018). The reason for this is presumably that the translational drawbacks of consecutive prolines are outweighed by prolines capability to endow proteins with unique structural properties not least exemplified by polyproline helices. Accordingly, life had to find a way to overcome this barrier, and in 2013 the molecular mechanism that made this possible was discovered. In this year, two parallel studies demonstrated independently that in bacteria the specialized translation elongation factor P (EF-P) is responsible for the relief of the translational ribosomal arrest caused by polyproline motifs (Doerfel et al., 2013; Ude et al., 2013). This finding was shortly afterwards also reported for the eukaryotic EF-P homolog, eIF5A (Gutierrez et al., 2013).

1.2 The discovery of EF-P and its molecular function

EF-P itself was already discovered in 1975 by Bernard Glick and Clelia Ganoza (Glick and Ganoza, 1975). In their work they were able to identify a soluble protein fraction containing an extraribosomal factor with the ability to stimulate peptide bond formation

between the initiator transfer RNA *N*-formyl-methionyl-tRNA$_i$ (fMet-tRNA$_i^{fMet}$) and the aminoacyl-tRNA analog, puromycin. Since puromycin represents a special case as aminoacyl acceptor, the team around Glick and Ganoza was curious if EF-P can also stimulate peptide bond formation between fMet-tRNA$_i^{fMet}$ and other aminoacyl acceptors mimicking more the natural aminoacyl tRNAs. To do so, they investigated 1979 the influence of EF-P on peptide bond formation between fMet-tRNA$_i^{fMet}$ and diverse other aminoacyl acceptors (Glick et al., 1979). In this work they could show *in vitro*, that small amino acids like glycine and leucine depend more on EF-P to form the first peptide bond with fMet-tRNA$_i^{fMet}$ than others. Based on these finding they suggested that EF-P stimulates the first peptide bond formation not in general but specifically when aminoacyl acceptors have a smaller and/or less hydrophobic side chain. However, the confirmation of this suggestion *in vivo* did not take place at that time, most likely because *efp* was described as essential (Aoki et al., 1997), which turned out to be incorrect later on. Correct is, that *efp* can be deleted in the majority of the model organisms used in the laboratory, including *Escherichia coli* (Baba et al., 2006), *Pseudomonas aeruginosa* (Balibar et al., 2013) *Salmonella enterica* (Zou et al., 2012), *Shewanella oneidensis* (Lassak et al., 2015), *Agrobacterium tumefaciens* (Peng et al., 2001), *Brucella abortus* (Iannino et al., 2012) and *Bacillus subtilis* (Ohashi et al., 2003) and only in rare cases like *Neisseria meningitides* or *Mycobacterium tuberculosis* deletion of *efp* is lethal (Sassetti et al., 2003; Yanagisawa et al., 2016). Based on the first structural data (Hanawa-Suetsugu et al., 2004), it became evident that the overall structure of EF-P consists of three β-barrel domains which resemble the L-shape of tRNAs (Figure 1).
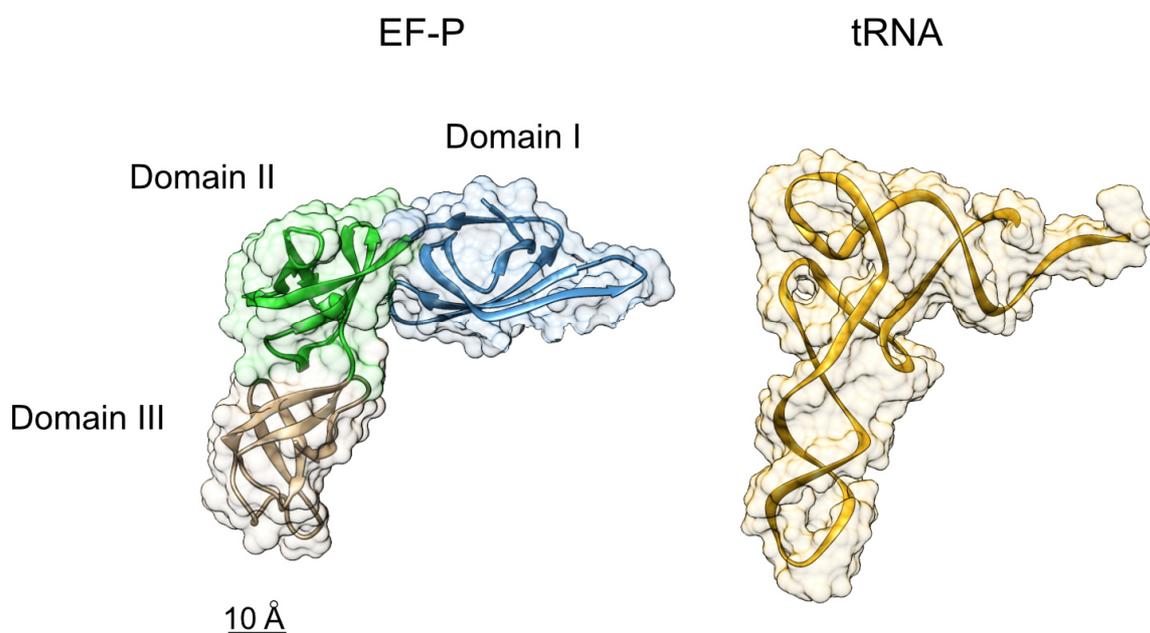
Figure 1: Structure comparison between EF-P and tRNA. (Left) Crystal structure of EF-P from *Thermus thermophiles* (Hanawa-Suetsugu et al., 2004) (PDB: 1UEB). The different colors represent the different β-barrel domains I (blue), II (green) and III (brown) **(Right)** Crystal structure of tRNA^{Phe} from *Saccharomyces cerevisiae* (Jovine et al., 2000) (PDB: 1EVV). For visualization, UCSF Chimera was used (Pettersen et al., 2004).

Then, in 2009, Steitz and colleagues were able to solve the crystal structure of EF-P in complex with the *Thermus thermophilus* 70S ribosome, the ribosomal protein L1, fMet-tRNA$_i^{fMet}$ and the messenger RNA (mRNA) (Blaha et al., 2009). Based on this co-crystal structure they assumed, like Glick and Ganoza 30 years ago, that EF-P is needed in the cell for the correct positioning of the fMet-tRNA$_i^{fMet}$ in the P site, and by that allowing the formation of the first peptide bond during translation elongation.

1.3   EF-P resolves polyproline (PPP) mediated ribosome stalling

In 2013, two independent research groups showed that EF-Ps main task is to stimulate peptide bond formation between prolines (Doerfel et al., 2013; Ude et al., 2013). This discovery arose from *in vivo* experiments in which the polyproline containing pH-sensor CadC of *E. coli* could be identified as a protein whose synthesis is directly dependent on EF-P (Ude et al., 2013). Based on this finding, further *in vivo* analysis of diverse polyproline containing proteins in combination with *in vitro* translation experiments could show that EF-P is required to resolve ribosome stalling caused by three or more consecutive prolines. Independently of this study, the group lead by Marina Rodnina came to the same conclusion (Doerfel et al., 2013). They determined the kinetics of peptide bond formation between the 20 proteinogenic amino acids and puromycin. Here they showed that the addition of EF-P leads to a strong acceleration of the peptide bond formation between proline and puromycin, an affect which could not be observed between the 19 other proteinogenic amino acids and puromycin. Furthermore, *in vitro* translation experiments showed that the synthesis of peptides containing three consecutive prolines was only possible in the presence of EF-P. Shortly afterwards, also a third laboratory supported these results. They showed that eIF5A, the eukaryotic homolog of EF-P, is also required for the translation of polyproline sequences (Gutierrez et al., 2013).

Taken together, these studies prove that EF-P/eIF5A are universal translation factors needed for the effective translation of polyproline sequences on the ribosome. Additionally, and since polyproline motifs are widespread in proteins (Morgan and Rubenstein, 2013; Ude et al., 2013), it provides also a conclusive explanation for properties associated with EF-P, such as its influence on cell growth, motility, viability, virulence, sensitivity to low molarity and susceptibility to certain antibiotics (Kearns et al., 2004; Navarre et al., 2010; Zou et al., 2011; Zou et al., 2012).

1.4   EF-P resolves diprolyl (PP) mediated ribosome stalling

Having shown that EF-P stimulates the peptide bond formation between stretches of three consecutive prolines (PPP), a more differentiated image of this effect was obtained shortly afterwards by our group in a joined work with Daniel Wilson and colleagues (Peil et al., 2013). In this study it could be shown that certain diprolyl motifs can also lead to a ribosomal arrest (Figure 2). The strength of this arrest depends strongly on the sequence context of the diprolyl motif, whereby the amino acid upstream (XPP) and downstream (PPX) of the diprolyl motif has a decisive role here. Based on these data, a hierarchy of arrest motifs was generated, ranging from strong stalling (PPP, DPP and PPN) to weak stalling motifs (CPP, PPR and PPH). These data were confirmed in 2015 by Buskirk and colleagues using ribosome profiling to

characterize pausing at proline motifs in the absence of EF-P (Woolstenhulme et al., 2015), and also by von Heijne and colleagues by using a mechanical pulling force sensor applied to different diprolyl arrest peptides (Cymer et al., 2015).
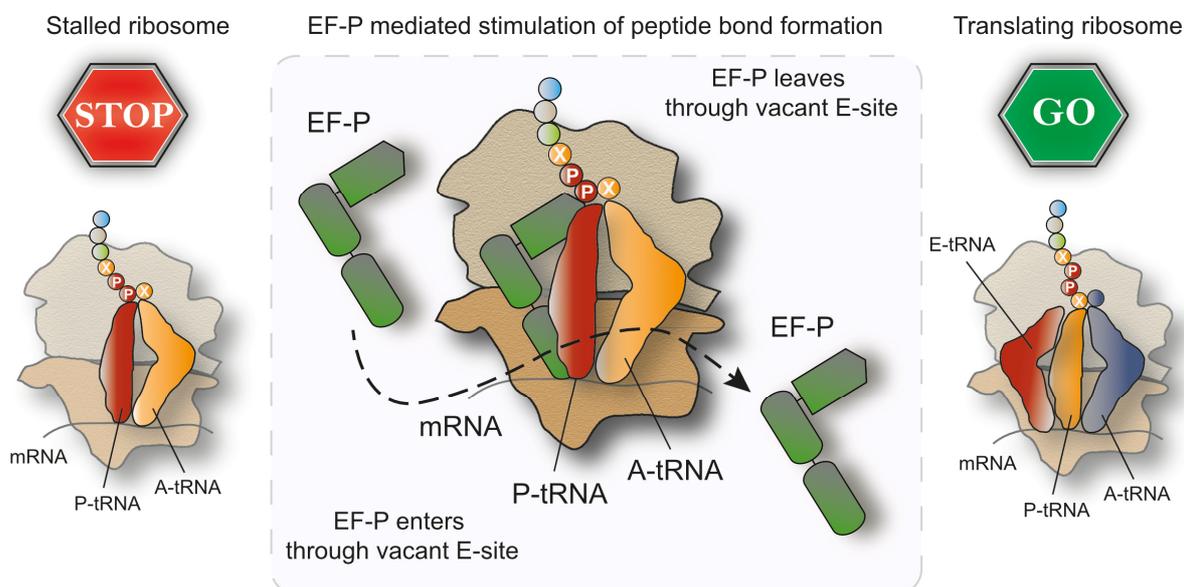
Figure 2: EF-P mediated release of ribosome stalling. (Left) Diprolyl paused ribosome. Red circles with P indicate proline, yellow circles with X indicate any of the 20 proteinogenic amino acids flanking the diprolyl motif and having influence on ribosome stalling. **(Middle)** EF-P enters the stalled ribosome through the vacant E-site and binds to the P-tRNA/ribosome to stimulate peptide bond formation. **(Right)** Afterwards, EF-P dissociates, leaves through the vacant E-site and protein synthesis continues.

1.5 The larger upstream amino acid context of the arrest motif also influences its stalling strength

Interestingly, the absence of EF-P does not lead in all cases to a reduced abundance of polyproline (PPP) containing proteins in the cell, and in some cases leads even to increased levels (Peil et al., 2013; Starosta et al., 2014). The reason for this unexpected fact was found in 2014 by an in-depth analysis of the amino acids preceding the stall site performed by our group, again in cooperation with Daniel Wilson and colleagues (Starosta et al., 2014). This analysis revealed that not only the first amino acid preceding the stall site strongly influence the strength of the arrest motif, but that this effect extends further up to five amino acids upstream of the stall site. Although this effect is strongest for the amino acid directly preceding the arrest motif it is still measurable, although with decreasing strength, till the fifth amino acid upstream. This observation was supported by a parallel study which came to comparable results (Elgamal et al., 2014). Based on these results we assumed that the conformation of the nascent polypeptide chain located within the ribosomal tunnel upstream of the stalling site is responsible for the stalling effect. However, to clarify this question further structural studies are needed.

1.6 Additional factors influencing the translation of proteins containing consecutive prolines

Besides the already mentioned factors (arrest motif strength, context and larger upstream amino acid composition), two other factors strongly affect diprolyl/polyproline

translation. This is on the one hand the translation initiation rate and, on the other hand, the position of the arrest motif within the open reading frame (ORF) (Hersch et al., 2014; Woolstenhulme et al., 2015). The following applies to these two properties:

- High rates of translation initiation result in an increase in EF-P dependence.
- The earlier the arrest motif is embedded in the ORF, the higher is the EF-P dependence.

The theory behind the described dependencies is the following: When having a stalling motif in combination with a high translation initiation rate, a ribosome queue emerges that extends back to the start codon and prevents further binding of new ribosomes. This happens the faster, the nearer the stalling motif is located within the ORF (Hersch et al., 2014; Woolstenhulme et al., 2015).

## 1.7 Post-translational modifications of EF-P

In order to improve its molecular function, EF-P is post-translationally modified. To date, three different strategies have been discovered in bacteria, which are discussed in detail in the following three sections.

1.7.1 β-Lysylation of lysine type EF-P by the EpmABC pathway

Already in 1991, the first suspicions arose that EF-P from *E. coli* owns a post-translational modification (PTM) (Aoki et al., 1991), but experimental proof regarding this aspect was only presented 17 years later by Ganoza and coworkers (Aoki et al., 2008). Their study actually investigated the interaction between *E. coli* EF-P and its ribosome. Within the scope of this work, however, the mass of EF-P was determined by means of mass spectrometry, whereby a mass shift of 144 Da could be detected on lysine at position 34 (K34). Nonetheless, the type of post-translational modification involved remained unclear until 2010 the first hypothesis regarding this was presented (Bailly and de Crecy-Lagard, 2010). Bailly and Crécy-Lagard postulated, on the basis of bioinformatic analyses, that K34 is (R)-β-lysylated. This happens in a two-step process under the participation of EpmA (also known as YjeA, PoxA or GenX) and EpmB (also known as YjeK) as follows: The lysine 2,3-aminomutase EpmB converts (S)-α-lysine into (R)-β-lysine (Figure 3, Step 1) (Behshad et al., 2006), which is then transferred onto K34 of EF-P by EpmA (Figure 3, Step 2). This results in a (R)-β-lysylation of K34. This hypothesis was confirmed in the following years by mutational analyses, *in vitro* experiments and the crystal structure of EF-P in complex with EpmA (Navarre et al., 2010; Sumida et al., 2010; Yanagisawa et al., 2010; Roy et al., 2011).

However, after the PTM pathway responsible for the full functionalisation of EF-P was elucidated, it remained unclear why *in vitro* (R)-β-lysylated EF-P has a 128 Da modification (Navarre et al., 2010; Yanagisawa et al., 2010) whereas endogenous EF-P has a 144 Da modification (Aoki et al., 2008; Park et al., 2012). The mass shift already indicated a hydroxylation (Yanagisawa et al., 2010), but how this hydroxylation is achieved was not clarified until 2012. There, Wilson and colleagues showed that EpmC (also known as YfcM) hydroxylates PTM EF-P in a final step, and that this

hydroxylation takes most likely place at the C5 atom of K34 (Peil et al., 2012) (Figure 3, Step 3).
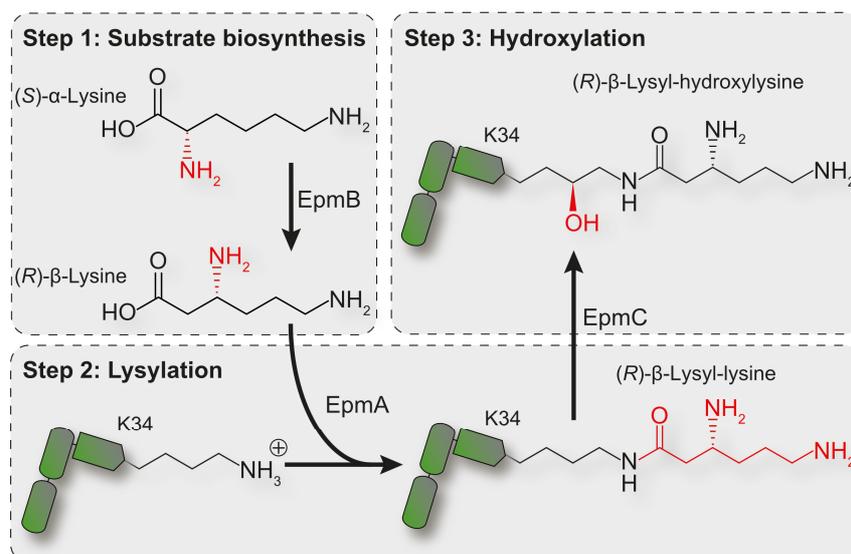
Figure 3: β-Lysylation of EF-P with conserved lysine at position 34 (K34) by EpmABC. (Step 1) The aminomutase EpmB shifts the α amino group of (*S*)-α-lysine in the β position, resulting in (*R*)-β-lysine. **(Step 2)** (*R*)-β-Lysine is attached onto K34 of EF-P by the protein ligase EpmA, resulting in (*R*)-β-lysyl-lysine. **(Step 3)** (*R*)-β-Lysyl-lysine is hydroxylated most likely on the C5 atom by EpmC, resulting in (*R*)-β-lysyl-hydroxylysine.

In the following years, the catalytic mechanism of EpmC could be clarified to a large extent (Kobayashi et al., 2014a; Kobayashi et al., 2014b), but the biological function behind this hydroxylation remains unclear, since deletions of *epmC* seemed to have no effect *in vivo* and *in vitro* on the functionality of EF-P (Peil et al., 2012; Doerfel et al., 2013; Ude et al., 2013). Furthermore, genome analyses show that EpmC is restricted however to some phyla encoding the EpmAB pathway (Lassak et al., 2015) (among them *E. coli* and *Salmonella*). This suggests that EpmC possesses only a minor molecular function, such as stabilization of the PTM by hydroxylation.

1.7.2 5-Amino-pentanolylation of lysine type EF-P by an unknown pathway

The so far described modification pathway from *E. coli* could only be identified in about 25% of all bacteria (Bailly and de Crecy-Lagard, 2010; Lassak et al., 2015), and notably, among them are no Gram-positive. Motivated by this this lack of information, Ibba and co-worker investigated the Gram-positive model organism *B. subtilis* regarding this aspect. There they found that the deletion of *efp* only impairs swarming motility but has no effect on growth or swimming motility (Kearns et al., 2004; Rajkovic et al., 2016). In comparison to that, the deletion of *efp* in Gram-negative bacteria exhibits a severe pleiotropic phenotype. With this knowledge in hand, Ibba and colleagues investigated the importance of EF-P in *B. subtilis* (Rajkovic et al., 2016) and mutated the lysine of the potential PTM site at position 32 (K32) of *B. subtilis* EF-P (equivalent to the PTM site K34 in *E. coli* / *Salmonella*). The result was that swarming motility of these mutants was impaired to the same amount as in the Δ*efp* mutants,

indicating a PTM at this position. To verify this, mass spectrometry experiments were performed and revealed a 5-amino-pentanolylation on K32.

In subsequent studies, Ibba and colleagues have now begun to clarify the modification pathway which leads to this 5-amino-pentanolylation of K32 (Hummels et al., 2017; Witzky et al., 2018). These studies indicate that the PTM is directly assembled on EF-P itself in a four step process (Figure 4, Step 1 to Step 4), but the involved enzymes remain unclear with exception of the enzyme involved in the last step (Witzky et al., 2018). In this step, YmfI (a homolog to the reductase FabG) catalyzes the reduction from 5-aminopentanone to 5-aminopentanol (Figure 4, Step 4) (Hummels et al., 2017). However, the PTM described here is rather an exceptional case and probably limited to Firmicutes (Hummels et al., 2017).
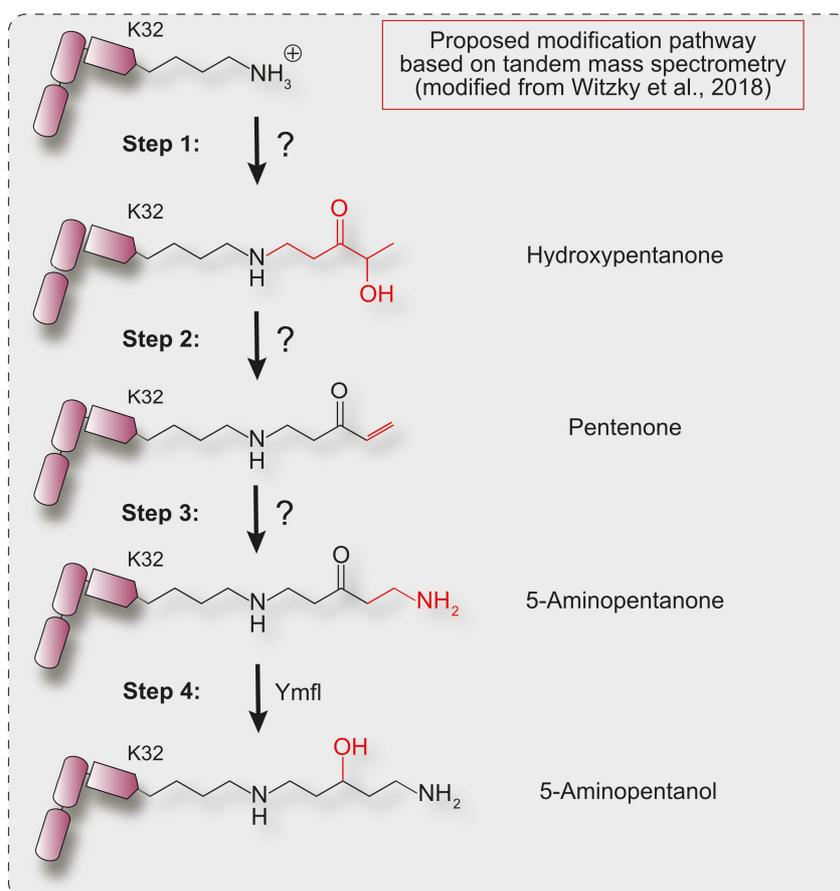
Figure 4: 5-Aminopentanolylation of EF-P with conserved lysine at position 32 (K32). (Step 1 to Step 3) Recent studies suggest, based on tandem mass spectrometry, that 5-aminopentanone is assembled directly on K32 of EF-P from *B. subtilis* (Witzky et al., 2018). The involved modification enzymes remain unknown to date. **(Step 4)** In a final step 5-aminopentanone is converted into 5-aminopentanol by YmfI (Hummels et al., 2017).

1.7.3 Rhamnosylation of arginine type EF-P by the RmlABCD/EarP pathway

So far, the two PTMs mentioned, (*R*)-β-lysylation and 5-amino-pentanolylation, are structurally and chemically reminiscent. In 2015, however, a chemically distinct PTM was discovered which is present in about 10% of all bacteria (Lassak et al., 2015). The basis for this discovery was again provided by bioinformatic analyses, which helped to

identify a subfamily of unusual EF-Ps in genomes lacking the β-lysylation pathway EpmABC. These EF-Ps have a strictly conserved arginine R32, which is the equivalent position to the PTM site K34 in *E. coli*. However, even more interesting, they encode a conserved protein in the direct genomic neighbourhood which has now been referred to as EarP (Lassak et al., 2015). To elucidate whether there is a functional connection between EarP and EF-P, the corresponding genes were individually deleted both in *S. oneidensis* and in *P. aeruginosa*. The deletion of *earP* led to the same mutant phenotype as the deletion of *efp*, which is reflected by a strong growth defect. (Lassak et al., 2015; Rajkovic et al., 2015). Furthermore, reporter assays showed that EF-P can only rescue polyproline stalled ribosomes in the concomitant presence of EarP (Lassak et al., 2015).

Having shown a functional connection between EarP and EF-P, it had to be clarified whether this connection is related to an unknown PTM at R32. Therefore, mass spectroscopy experiments were performed and identified a deoxyhexose moiety on the R32 residue of EF-P (Lassak et al., 2015; Rajkovic et al., 2015). This moiety was not present when *earP* was chromosomally deleted, providing first evidence that EarP modifies EF-P post-translationally. Subsequent *in vitro* experiments identified dTDP-rhamnose as the substrate for EarP, a result which was further supported by *in vitro* glycosylation reactions with purified components (Lassak et al., 2015; Rajkovic et al., 2015). dTDP-β-L-rhamnose is the product of the well characterized RmlABCD pathway which is strictly conserved in bacteria encoding EarP. In an inverting glycosyltransferase reaction (Li et al., 2016; Krafczyk et al., 2017; Sengoku et al., 2018) EarP uses this precursor and attaches rhamnose to EF-P (Lassak et al., 2015; Rajkovic et al., 2015) establishing now the full modification pathway of EF-Ps with conserved arginine at position 32 (Figure 5).
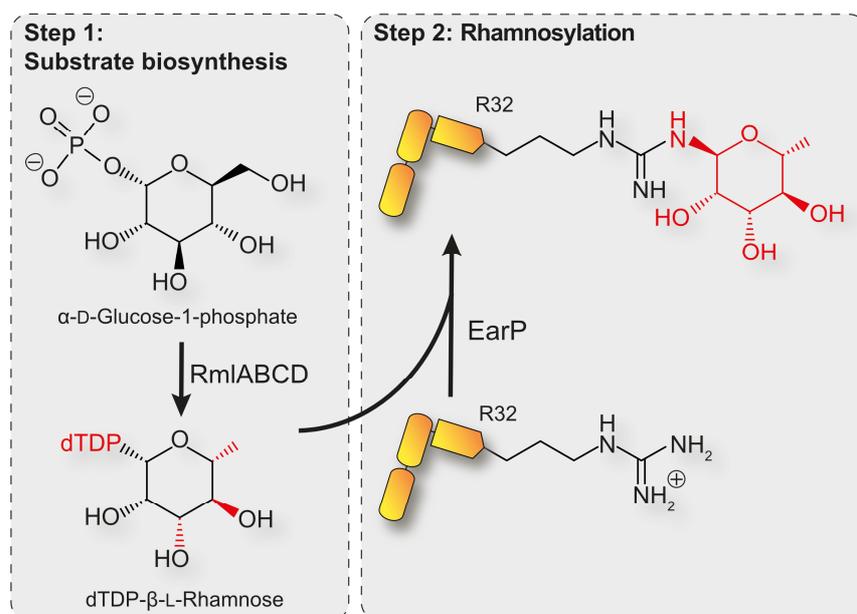

Figure 5: Rhamnosylation of EF-P with conserved arginine at position 32 (R32) by the RmlABCD/EarP pathway. (Step 1) α-D-glucose-1-phosphate is converted in a multistep process to dTDP-β-L-rhamnose by the RmlABCD pathway. **(Step 2)** EarP uses dTDP-β-L-rhamnose as donor substrate to attach rhamnose to arginine at position 32 of EF-P.

1.8 Motivation and strategies for the synthetic activation of EF-P

As the molecular function of EF-P is well understood, the mechanism of how EF-P stimulates the peptide bond formation between prolines on the ribosome and the involvement of the chemically distinct PTMs remains elusive. Although the first cryo-electron microscopy structure of EF-P with (R)-β-lysyl-hydroxylysine modification in complex with the translating ribosome stalled at polyproline stretches sheds some light on the exact role of the PTM (Huter et al., 2017), it remains unclear to date how chemically distinct PTMs can have the same stimulating effect on the ribosome. To understand this so far neglected aspect, this thesis was intended to investigate this by the synthetic functionalisation of *E. coli* EF-P. To achieve this, two different strategies were pursued in parallel, one translational and one post-translational based approach.

1.8.1 Strategy one: Translational activation of *E. coli* EF-P using the amber suppression system

With this strategy we aimed to replace the natural occurring lysine at position 34 (K34) in EF-P from *E. coli* by (R)-β-lysyl-lysine and related derivatives by using the amber suppression system (Wang et al., 2001). With this system non-canonical amino acids (ncAA) can be incorporated site specifically into proteins, after mutating the desired site to an amber stop codon (TAG). The incorporation of the ncAA then takes place through an orthogonal translation system, in our case the pyrrolysyl-tRNA synthetase (PylRS)/tRNA$_{CUA}$ pair from the archaeum *Methanosarcina mazei* (Figure 6) (Crnkovic et al., 2016).
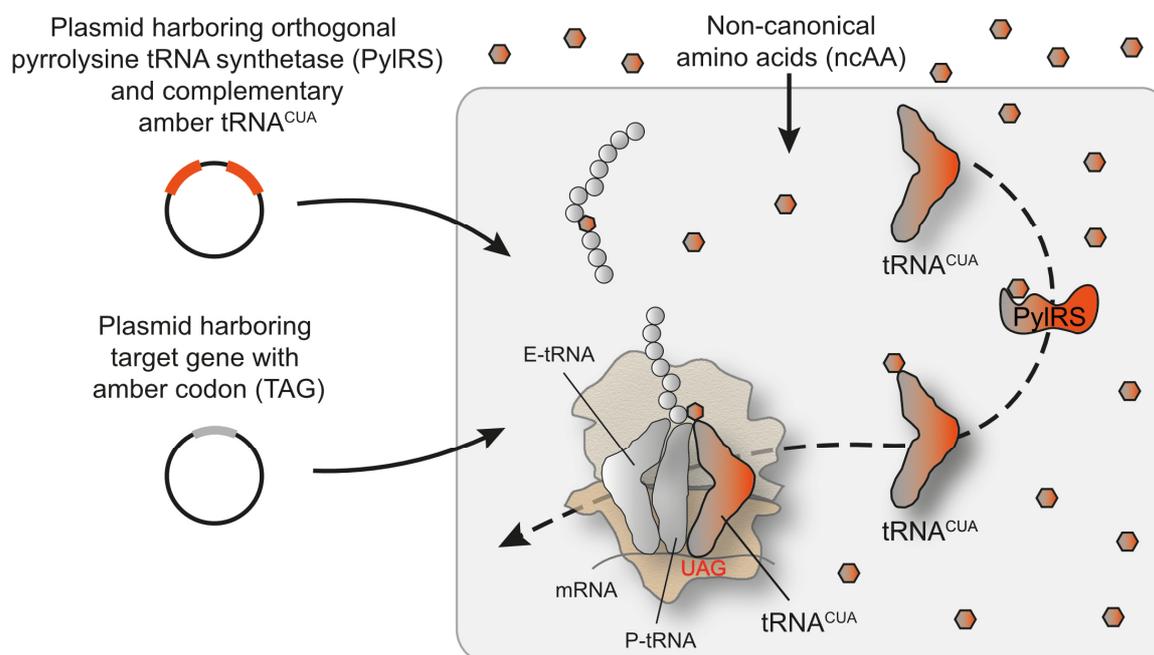
Figure 6: The amber suppression system for the site specific incorporation of ncAAs into proteins. **(Left)** In a first step, cells are transformed with plasmids harboring the orthogonal translation system (PylRS/tRNA$_{CUA}$ pair) and the *efp* gene were at position 34 the natural codon is replaced by an amber stop codon. **(Right)** In the next step, these cells are cultured in a medium that contains the ncAA for incorporation. As a result, the ncAA is incorporated at the amber stop codon position 34 into the growing polypeptide chain by the PylRS/tRNA$_{CUA}$ pair.

Due to the high PylRS substrate promiscuity paired with engineering efforts of the catalytic pocket, it is nowadays possible to incorporate more than 150 different ncAAs with this system (Wan et al., 2014; Dumas et al., 2015). Among them are also numerous lysine derivatives which resemble the naturally occurring post-translational modified lysine variant (*R*)-β-lysyl-lysine of *E. coli*. As part of my master's thesis, several of these lysine derivatives (e.g. N^ε-propionyl-, N^ε-butyryl-, and N^ε-crotonyl-lysine) were already incorporated into *E. coli* EF-P. However, all of these EF-P variants lost their translational activity.

Therefore, we asked whether (*R*)-β-lysyl-lysine can be incorporated directly into EF-P by means of orthogonal translation using an engineered PylRS (Chapter 2). If this would work, subsequent rational design of the catalytic pocket of this PylRS variant might even enable the incorporation of derivatives closely related to (*R*)-β-lysyl-lysine. Ultimately this would allow to draw conclusions regarding the chemical prerequisites of the EF-P modification that stimulates peptide bond formation between prolines.

In addition, the expertise acquired during the master's thesis on the amber suppression system led to the idea of developing a toolbox which allows to control bacterial protein levels of even essential genes *in vivo* based on orthogonal translation. (Chapter 3).

1.8.2 Strategy two: Non-cognate post-translational activation of *E. coli* EF-P

As already mentioned, only two EF-P modification pathways in bacteria are completely understood. These pathways post-translationally modify EF-P, and thus fully functionalise it. This functionalisation is achieved in γ-proteobacteria like *E. coli* by (*R*)-β-lysylation and in β-proteobacteria like *P. aeruginosa* by rhamnosylation. Since the two modification strategies differ significantly from each other but the EF-Ps to be modified are structurally similar (Katz et al., 2014; Krafczyk et al., 2017) (Figure 7), we assumed that it might be possible to switch the EarP acceptor substrate specificity from *P. aeruginosa* EF-P to the non-cognate *E. coli* EF-P. This cross-modification of *E. coli* EF-P would then allow to draw conclusions about the chemical and structural requirements to rescue polyproline mediated ribosome stalling (Chapter 4).

<div align="center">

E. coli EF-P   *P. aeruginosa* EF-P

</div>
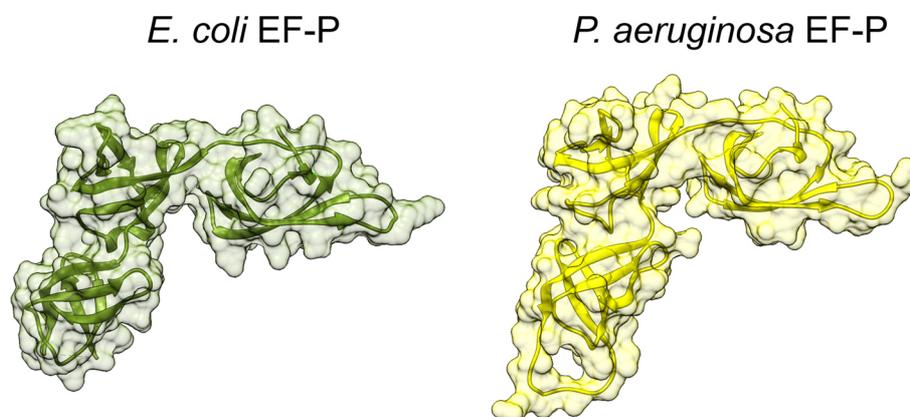
Figure 7: Comparison of EF-P structures from different bacteria. (Left) Crystal structure of *E. coli* EF-P (Yanagisawa et al., 2010) (PDB: 3A5Z). **(Right)** Crystal structure of *P. aeruginosa* EF-P (Choi and Choe, 2011) (PDB: 3OYY). For visualization, UCSF Chimera was used (Pettersen et al., 2004).

2   Strategy for the Translational Activation of EF-P Using the Amber Suppression System

Manuscript

<u>Wolfram Volkwein</u>[1a], Sabine Peschek[1], Benedikt Frederik Camille Graf von Armansperg[1b], Kirsten Jung[1] and Jürgen Lassak[1*] (2019).

[1]Center for Integrated Protein Science Munich, Department of Biology I, Microbiology, Ludwig-Maximilians-Universität München, Munich, Germany.

[a]Present address: Bavarian Health and Food Safety Authority, Oberschleißheim, Germany.

[b]Present address: Max-von-Pettenkofer Institute, Ludwig-Maximilians-Universität München, Munich, Germany.

*To whom correspondence should be addressed:
Jürgen Lassak, Ludwig-Maximilians-Universität München, Munich, Germany.
E-mail: juergen.lassak@lmu.de

Introduction

Ribosome stalling occurs during the synthesis of proteins containing consecutive prolines, and is resolved in bacterial cells by the translation elongation factor EF-P (Doerfel et al., 2013; Ude et al., 2013). EF-P binds between the P- and E-sites of the ribosome and interacts from there with the CAA end of the peptidyl-tRNA. This interaction leads to a conformation of the polyproline-containing nascent peptide chain that promotes the formation of peptide bonds between prolines (Huter et al., 2017). However, in order to fully fulfil its function on the ribosome, EF-P has to be post-translationally modified. In bacteria like *Escherichia coli*, this modification is achieved by the attachment of a (*R*)-β-lysine moiety to a conserved lysine (K34) at the tip of domain I of EF-P, resulting in a (*R*)-β-lysylation of K34. ((Behshad et al., 2006; Bailly and de Crecy-Lagard, 2010; Navarre et al., 2010; Sumida et al., 2010; Yanagisawa et al., 2010)).

Interestingly, the pyrrolysyl-tRNA synthetase (PylRS) allows to charge its cognate amber codon recognizing tRNA$_{CUA}$ with numerous lysine derivatives (Dumas et al., 2015) due to its high substrate promiscuity (Wan et al., 2014) and allowing thereby the incorporation of these derivatives into proteins. Based on this fact the idea arose to use the amber suppression system to incorporate site specifically (*R*)-β-lysyl-lysine into *E. coli* EF-P (EF-P$_{Eco}$) at position 34, and by that, activate EF-P in a post-translational independent manner. For this purpose (*R*)-β-lysyl-lysine was chemical synthesized and subsequently, the wild-type (wt) PylRS was tested for its ability to incorporate (*R*)-β-lysyl-lysine site specifically at position 34 into EF-P$_{Eco}$.

In parallel, a screening system based on the swimming ability of *E. coli* was developed. This system allows to select for PylRS mutants capable of incorporating the lysine derivative of interest in a cheap and simple way.

Results

Incorporation of (*R*)-β-lysyl-lysine into EF-P$_{Eco}$

Since the PylRS is known for its high substrate promiscuity (Yanagisawa et al., 2008b; Wan et al., 2014), we hypothesized that the amber suppression system based on the wt PylRS/tRNA$_{CUA}$ pair might also be capable of incorporating (*R*)-β-lysyl-lysine into EF-P$_{Eco}$. To test this, we replaced the original lysine codon at position 34 of *efp$_{Eco}$* by an amber stop codon leading to the variant EF-P$_{Eco}$ K34$^{(Am)}$. As a result, only wt PylRS/tRNA$_{CUA}$ pair mediated incorporation of (*R*)-β-lysyl-lysine into this variant leads to the production of full-length EF-P$_{Eco}$. Furthermore, and since this synthetic variant harbor a C-terminal His$_6$-tag, (*R*)-β-lysyl-lysine incorporation can be verified indirectly by Western Blot analysis using an Anti His$_6$ antibody afterwards.

In order to determine if the synthetic EF-P$_{Eco}$ K34$^{(Am)}$ variant with incorporated (*R*)-β-lysyl-lysine is functional, experiments were performed in an *E. coli* reporter strain (Figure 8A). In this reporter strain, *efp* is chromosomally deleted (Δ*efp*) and possesses an EF-P dependent *lacZ* activity. As a result, only EF-P in its post-translationally modified form leads to β-galactosidase activity whereas unmodified EF-P does not (for

more details see also (Ude et al., 2013)). Furthermore, not only (*R*)-β-lysyl-lysine was used for wt PylRS/tRNA$_{CUA}$ pair mediated incorporation but also *N*^{ε}-butyryl-lysine (Figure 8B). *N*^{ε}-Butyryl-lysine is a substrate known to be accepted by the wt PylRS/tRNA$_{CUA}$ pair (Gattner et al., 2013) and therefore served as a positive control for successful incorporation.
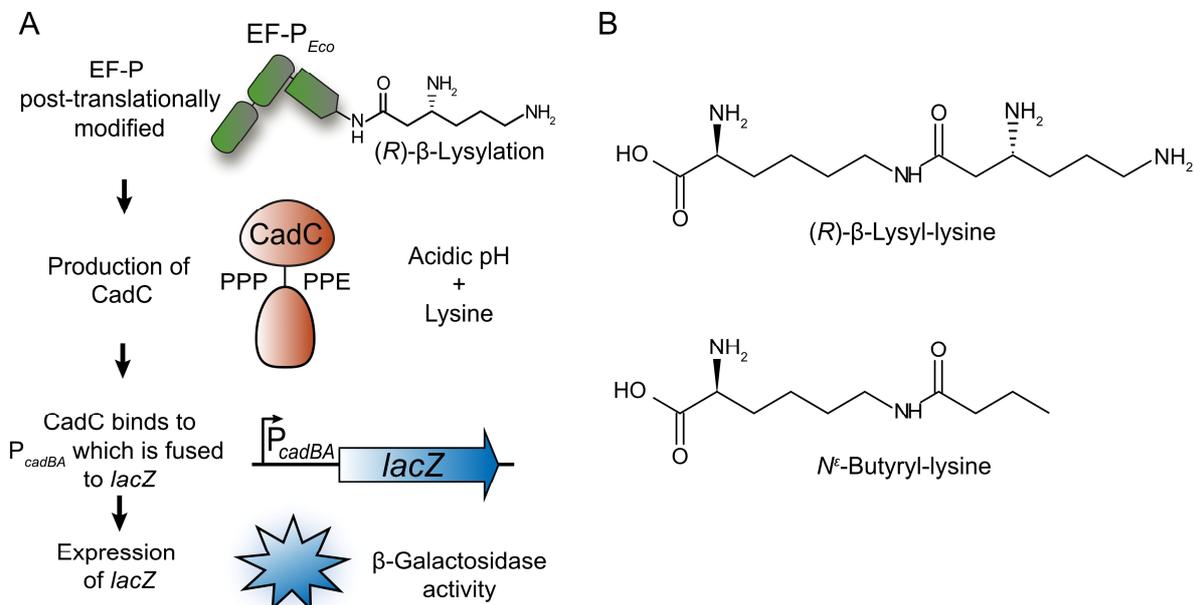
Figure 8: Reporter strain for the detection of EF-P functionality and chemical structures of (*R*)-β-lysyl-lysine and *N*^{ε}-butyryl-lysine. (A) Schematic illustration of the *E. coli* reporter strain used to detect EF-P functionality: The membrane integrated polyproline/diproline containing pH sensor CadC can only be produced in the presence of active (post-translationally modified) EF-P. If CadC is produced - and the cells were grown under mild acidic pH and in the presence of lysine - CadC binds to its promoter *cadBA*, which is fused to *lacZ* in this reporter strain. As a result, the *lacZ* expression and thereby β-galactosidase activity in this strain is dependent on post-translationally modified EF-P. **(B)** Chemical structures of (*R*)-β-lysyl-lysine and *N*^{ε}-butyryl-lysine which are used in this study as substrates for the wt PylRS/tRNA$_{CUA}$ pair.

Our results show that, however, the wt PylRS/tRNA$_{CUA}$ pair is not capable of incorporating (*R*)-β-lysyl-lysine into EF-P$_{Eco}$ K34$^{(Am)}$ (Figure 9A). Furthermore, the functionality of post-translationally modified EF-P$_{Eco}$ (Figure 9A and B, dark grey EF-P$_{Eco}$ bars) was compared to that of other EF-P$_{Eco}$ variants (Figure 9A and B) including unmodified EF-P$_{Eco}$ (Figure 9B, light grey EF-P$_{Eco}$ bar). In line with previous reports (Volkwein et al., 2019) the replacement of lysine by alanine (EF-P$_{Eco}$ K34A) leads to a loss of functionality, while the replacement of lysine by arginine (EF-P$_{Eco}$ K34R) restores functionality to a certain degree (Figure 9A and B). Interestingly, the EF-P$_{Eco}$ K34$^{(Am)}$ *N*^{ε}-butyryl-lysine variant shows also functionality in the Δ*efp* reporter strain (Figure 9A). One possible explanation for this functionality could be that the butyryl moiety is cleaved of by sirtuins like CobB (Ringel et al., 2014), resulting in lysine at position 34 (K34) which can then be lysylated by EpmA again. To exclude this, experiments were additionally performed in the reporter strain background where *efp* and the modification enzyme *epmA* were deleted (Figure 9B, light grey bars). Additionally, this reporter strain allows to produce unmodified EF-P$_{Eco}$ since *epmA* is deleted, and to compare its functionality with the EF-P$_{Eco}$ K34$^{(Am)}$ *N*^{ε}-butyryl-lysine

variant. The result was that the EF-P$_{Eco}$ K34$^{(Am)}$ N^{ε}-butyryl-lysine variant is indeed functional, but that its functionality is below the functionality of unmodified EF-P$_{Eco}$ (Figure 9B, light grey EF-P$_{Eco}$ bar), hence a synthetic translational activation with N^{ε}-butyryl-lysine is not possible. Furthermore, these results show that unmodified EF-P$_{Eco}$, when overproduced, can rescue polyproline mediated ribosome stalling to a certain degree (Figure 9B) which is in line with previous reports (Volkwein et al., 2019).
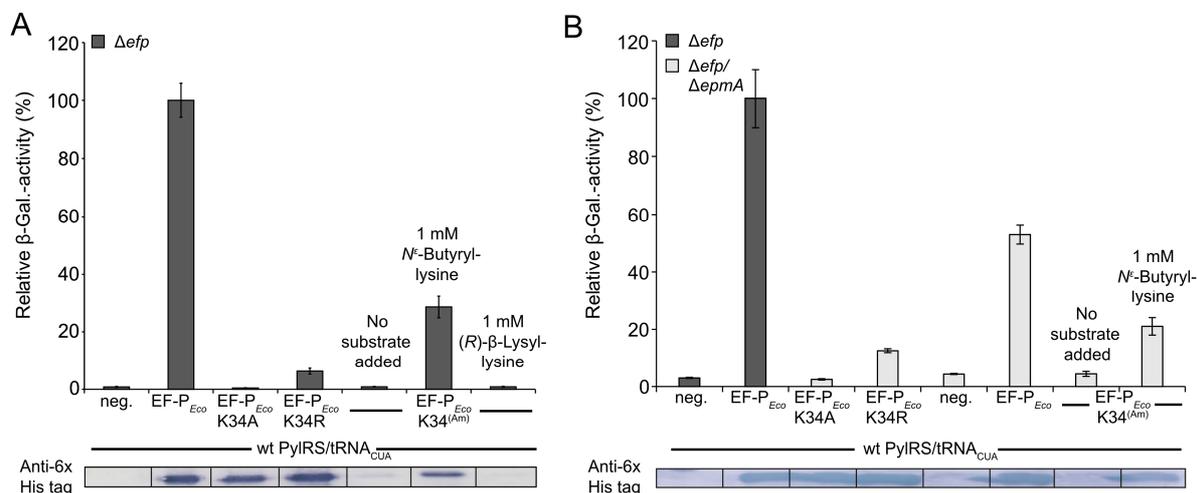
Figure 9: Translational activity of synthetic EF-P$_{Eco}$ variants. (A) *In vivo* expression of different His$_6$-tagged EF-P$_{Eco}$ variants (EF-P$_{Eco}$, EF-P$_{Eco}$ K34A, EF-P$_{Eco}$ K34R, EF-P$_{Eco}$ K34$^{(Am)}$) in combination with the wt PylRS/tRNA$_{CUA}$ pair in the reporter strain where *efp* is deleted (MG-CR-efp-KanS, Δ*efp*). Cells were grown under *cadBA*-inducing conditions (LB, pH 5.8). Where indicated, 1 mM (*R*)-β-lysyl-lysine or 1 mM N^{ε}-butyryl-lysine was added to the medium. EF-P production was verified by Western Blot analysis. β-Galactosidase activities are given in relative Miller units with EF-P$_{Eco}$ in the Δ*efp* background set to 100%. **(B)** *In vivo* expression of different EF-P variants (EF-P$_{Eco}$, EF-P$_{Eco}$ K34A, EF-P$_{Eco}$ K34R, EF-P$_{Eco}$ K34$^{(Am)}$) in combination with wt PylRS/tRNA$_{CUA}$ pair in the described reporter strain. In the reporter strain, either *efp* (MG-CR-efp-KanS, dark grey bars, Δ*efp*) or *efp* in combination with *epmA* (MG-CR-efp-epmA-KanR, light grey bars, Δ*efp* Δ*epmA*) is deleted. Cells were grown under *cadBA*-inducing conditions (LB, pH 5.8). 1 mM N^{ε}-Butyryl-lysine was added to the surrounding medium where indicated. EF-P production was verified by Western Blot analysis. β-Galactosidase activities are given in relative Miller units with EF-P$_{Eco}$ in the Δ*efp* background set to 100%.

Development of a screening system based on *E. coli* FliC

Based on the fact that the wt PylRS is not capable of incorporating (*R*)-β-lysyl-lysine into the growing polypeptide chain, the idea arose to evolve its substrate specificity towards (*R*)-β-lysyl-lysine. That it is indeed possible to evolve the PylRS towards a specific substrate has already been shown repeatedly (Wan et al., 2014). However, none of the described substrates so far resembled the chemical structure of (*R*)-β-lysyl-lysine. For this reason, a mutational screening was considered to be useful, and as a first step a new screening system should be developed and established based on the following idea:

To swim, *E. coli* uses its peritrichous flagella on the cell surface, which are composed of many different proteins including the filament FliC (Berg, 2003). The replacement of a natural codon within the *fliC* gene by an amber stop codon would have the result that the protein production of FliC, and thereby the swimming motility, would depend on the amber suppression activity of the introduced PylRS mutant for a specific substrate. In

our case (*R*)-β-lysyl-lysine. As a consequence, only *E. coli* cells with the desired mutant capable of incorporating (*R*)-β-lysyl-lysine into FliC can produce FliC and are thereby able to swim while the others are not. These cells can then be easily identified and isolated on swimming plates.

In a first step and in order to establish the described screening system in *E. coli*, *fliC* was interrupted in the *E. coli* MG1655 derivative C321.ΔA.exp (Lajoie et al., 2013). We decided to work with this strain because in its chromosome all 321 UAG (amber) stop codons have been replaced by UAA (ochre) stop codons. This permits the deletion of release factor 1, which in turn leads to improved properties for substrate incorporation mediated by the amber suppression system (Lajoie et al., 2013).

Then, in a second step we used the generated strain with interrupted *fliC* (C321.ΔA.exp *fliC*::*npt*) in combination with a PylRS mutant which is capable of incorporating N^{ε}-acetyl-lysine at the amber position into the growing polypeptide chain. This acetyl lysyl-tRNA synthetase (AcKRS) (Umehara et al., 2012) was chosen for the establishment of the screening system due to the following reasons:

- First, it is capable of incorporating the commercially available and cheap substrate N^{ε}-acetyl-lysine (Sigma-Aldrich) instead of the not commercially available substrate N^{ε}-butyryl-lysine.

- Second, the use of the AcKRS allows the production of unaltered FliC if a lysine codon is replaced by an amber stop codon. The reason for this is that after N^{ε}-acetyl-lysine is incorporated, the acetyl group can be subject to cleavage by CobB (AbouElfetouh et al., 2015), resulting in unmodified native protein, which ensures in turn the full functionality of the produced FliC.

To demonstrate the functionality of the screening system, the *E. coli* strains C321.ΔA.exp and C321.ΔA.exp *fliC*::*npt* were transformed with a plasmid encoding for the acetyl lysyl-tRNA synthetase/tRNA$_{CUA}$ pair (AcKRST), FliC or FliC$^{(Am)}$, or a combination of these plasmids (Figure 10A). To determine the influence of these plasmid combinations on the swimming capability in response to N^{ε}-acteyl-lysine, the cells were spotted on 0.3 % (w/v) swimming plates supplemented with (Figure 10B) or without (Figure 10C) N^{ε}-acetyl-lysine. The swimming plates were then incubated for 24 hours and subsequently pictures were taken. These pictures show that the swimming behavior of the stain C321.ΔA.exp is independent of the presence/absence of N^{ε}-acetyl-lysine (Figure 10B and 10C, samples 1 to 6). The plates also show that the swimming capability is abolished when *fliC* is interrupted chromosomally in the strain C321.ΔA.exp *fliC*::*npt* (Figure 10B and 10C, samples 7). The plasmid encoding FliC$^{(Am)}$ alone has no influence on the swimming capability of the cells (Figure 10B and 10C samples 8 and 11). The swimming capability can be restored upon ectopic expression of FliC (Figure 10B and C, samples 9 and 10). Most interestingly, the expression of AcKRST in combination with FliC$^{(Am)}$ results in swimming capability only in the presence of N^{ε}-acetyl-lysine (Figure 10B and C, samples 12 with red frame), proving the applied screening principle as correct.

A

C321.ΔA.exp

1: neg.	**2**: + AcKRST	**3**: + FliC
4: + FliC + AcKRST	**5**: + FliC$^{(Am)}$	**6**: + FliC$^{(Am)}$ + AcKRST

C321.ΔA.exp
flic::*npt*

7: neg.	**8**: + AcKRST	**9**: + FliC
10: + FliC + AcKRST	**11**: + FliC$^{(Am)}$	**12**: + FliC$^{(Am)}$ + AcKRST

B   + N^ε-Acetyl-lysine   **C**   - N^ε-Acetyl-lysine

C321.ΔA.exp

C321.ΔA.exp
flic::*npt*

Figure 10: Establishment of a screening system to detect successful substrate incorporation mediated by the amber suppression system. (A) Schematic overview of the different cultures spotted on the swimming plates. The cultures 1-6 correspond to *E. coli* C321.ΔA.exp, and the cultures 7-12 correspond to *E. coli* C321.ΔA.exp *fliC*::*npt* cells, complemented where indicated with a plasmid encoding AcKRST, FliC, FliC$^{(Am)}$ or a combination of these plasmids. **(B)** Swimming capability of the in **(A)** described cultures on swimming plates supplemented with N^ε-acetyl-lysine. **(C)** Swimming capability of the in **(A)** described cultures on swimming plates not supplemented with N^ε-acetyl-lysine. The swimming capability was determined as described in the material and methods section.

Concluding Remarks

The here presented results demonstrate that the FliC-based screening system in principle can be used to identify new PylRS mutants with new substrate specificity. At this point, the next step would be to tested whether other components of the flagella are better suited for this purpose or not. Then, PylRS libraries have to be generated were residues of the binding pocket are randomized to change its substrate specificity towards (*R*)-β-lysyl-lysine. Which residues have to be mutated in order to achieve this remains speculation at this time point, because (*R*)-β-lysyl-lysine has a unique chemical structure compared to all other substrates accepted so far by the PylRS. A further aspect of uncertainty which has to be taken into account is the fact that it is not known if chemically synthesized (*R*)-β-lysyl-lysine is properly taken up into the cells or not. However, the here provided screening system can in principle also be applied to other substrates and provides a cheap and easy screening platform which can be used in nearly every lab.

Material and Methods

All strains, plasmids and primers used in this study are listed in Tables 1 to 3.

Strain construction

The *fliC* gene in the *E. coli* strain C321.ΔA.exp (Lajoie et al., 2013) was deleted using pRed/ET recombination technology in accordance to the Quick & Easy *E. coli* Gene Deletion Kit (Gene Bridges, Germany), resulting in C321.ΔA.exp *fliC::npt*. The strain C321.ΔA.exp itself was a gift from George Church (Addgene plasmid # 49018).

Plasmid construction

The plasmid pBBR1MCS-5-NP-fliC was constructed as follows: The *fliC* gene was amplified from *E. coli* MG1655 using the primer pair P1/P2 and was ligated between the SacI and EcoRI restriction sites of pBBR1MCS-5 (Kovach et al., 1995). In a similar way the plasmid pBBR1MCS-5-NP-fliC K199(Amber) was generated, only with the difference that mismatched primer pairs (P1 to P4) were used for overlap extension PCR to replace the original lysine (K) codon of *fliC* at position 199 by an amber (TAG) codon.

To replace the lysine codon at position 34 in *efp* in the construct pBBR1MCS-4-efp$_{Eco}$, a mismatched primer pair (P9/P10) in combination with primers for overlap-extension PCR (P11/P12) were used. The resulting fragment was then cloned into pBBR1MCS-4 using the XbaI and BamHI restriction sites resulting in pBBR1MCS-4-efp$_{Eco}$ K34(Amber).

β-galactosidase assay

β-Galactosidase activities in the *E. coli* Δ*efp* (MG-CR-efp) and *E. coli* Δ*efp/*Δ*epmA* (MG-CR-efp-epmA-KanR) reporter strain were determined as follows: Cells were grown over night at 37 °C under agitation and microaerobic conditions in 100 mM sodium-phosphate buffered Miller modified LB (pH 5.8) (Bertani, 1951; Miller, 1972). The next day, cells were harvested by centrifugation, and the β-galactosidase activities were determined as described (Tetsch et al., 2008) and are given in relative Miller units (MU) (Miller, 1992).

SDS PAGE and Western Blot

To visualize protein production of EF-P, proteins were separated by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) according to Lämmli (Laemmli, 1970). Afterwards, proteins were transferred to a nitrocellulose membrane by vertical Western blotting. Fully produced EF-P was detected in a first step by a primary Anti-His$_6$ antibody from rabbit (Abcam, Inc.). This primary antibody was then targeted by a secondary Anti-rabbit antibody with conjugated alkaline phosphatase (Rockland). To finally visualize EF-P presence on the nitrocellulose membrane, 0.01% (w/v) *p*-nitro blue tetrazolium chloride (NBT) and 0.045% (w/v) 5-bromo-4-chloro-3-indolyl-phosphate disodium salt (BCIP) was dissolved in 50 mM sodium carbonate buffer (pH 9.5) and added to the membrane.

Swimming assay

Cells were grown in Miller modified LB medium (Bertani, 1951; Miller, 1972) supplemented with 15 µg/mL gentamicin sulfate and 30 µg/mL chloramphenicol till an OD_{600} of 1.2. Then, 2 µL of these cultures were spotted on LB swimming plates containing 0.3% agar (w/v) supplemented with 15 µg/mL gentamicin sulfate, 30 µg/mL chloramphenicol and 5 mM N^ε-acetyl-lysine where indicated. These plates were incubated over night at room temperature. The next day, the incubation temperature was shifted to 37°C, and the plates were incubated for additional 6 hours. Afterwards, pictures were taken.

Synthesis of lysine derivatives

(R)-β-Lysyl-lysine and N^ε-butyryl-lysine were synthesized by the group of Thomas Carell from the Ludwig-Maximilians-Universität München.

Table 1: Strains used in this study

Strain	Genotype	Reference
DH5αλpir	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 ΔlacZYA-argF U169 φ80dlacZΔM15 λpir	(Macinga et al., 1995)
MG1655	F⁻ λ⁻ ilvG rfb50 rph-1	(Blattner et al., 1997)
MG-CR-efp	MG1655 ΔlacZ::tet rpsL150 efp::npt ΔcadBA PcadBA::lacZ	(Lassak et al., 2015)
MG-CR-efp-epmA-KanR	MG1655 ΔlacZ::tet rpsL150 Δefp ΔcadBA epmA::npt PcadBA::lacZ	(Volkwein et al., 2019)
C321.ΔA.exp	E. coli MG1655 Δ(ybhB-bioAB)::zeoR ΔprfA; all 321 UAG codons changed to UAA	(Lajoie et al., 2013)
C321.ΔA.exp fliC::npt	E. coli MG1655 Δ(ybhB-bioAB)::zeoR ΔprfA; all 321 UAG codons changed to UAA, fliC::npt	this study

Table 2: Plasmids used in this study

Plasmid	Feature	Reference
pRed/ET Amp	λ-RED recombinase in pBAD24; Amp^R	GeneBridges, Germany
pBBR1MCS-4	Broad-host-range cloning vector, 5.0 kb, Amp^R	(Kovach et al., 1995)
pBBR1MCS-5	Broad-host-range cloning vector, 4.9 kb, Gm^R	(Kovach et al., 1995)
pACYC-DUET™-1	Standard expression vector, ORI P15A, 4 kb, Cam^R	Novagene
pACYC-DUET-PylRST	Contains the pyrrolysyl-tRNA synthetase (PylRS) / cognate amber suppressor tRNA$_{CUA}$ (pylT) pair from M. mazei, 5.5 kb, Cam^R	(Gattner et al., 2013)
pACYC-DUET-AcKRST	Contains the acetyl lysyl-tRNA-synthetase (AcKRS), a mutated version of the PylRS from M. mazei towards N^ε-acetyl-lysine specificity, according to (Umehara et al., 2012) and the cognate amber suppressor tRNA$_{CUA}$ (pylT), 5.5 kb, Cam^R	(Volkwein et al., 2017)
pBBR1MCS-5-NP-fliC	Contains fliC from E. coli MG1655 fronted by its natural promoter, inserted into the SacI and EcoRI restriction sites of pBBR1MCS-5, Gm^R	this study

pBBR1MCS-5-NP-fliC K199Amber	Derivative from pBBR1MCS5-NP-fliC where additionally the lysine codon (AAA) at position 199 was replaced by an amber stop codon (TAG), Gm^R	this study
pBBR1MCS-4-efp$_{Eco}$	Contains *efp* from *E. coli* fronted by its natural promoter, Amp^R	Jürgen Lassak (unpublished)
pBBR1MCS-4-efp$_{Eco}$ K34A	Derivative from pBBR1 MCS4-*efp$_{Eco}$* where additionally the lysine codon (AAA) at position 34 was replaced by an alanine codon (GCG), Amp^R	Jürgen Lassak (unpublished)
pBBR1MCS-4-efp$_{Eco}$ K34R	Derivative from pBBR1 MCS4-*efp$_{Eco}$* where additionally the lysine codon (AAA) at position 34 was replaced by an arginine codon (AGG), Amp^R	Jürgen Lassak (unpublished)
pBBR1MCS-4-efp$_{Eco}$ K34(Amber)	Derivative from pBBR1 MCS4-*efp$_{Eco}$* where additionally the lysine codon (AAA) at position 34 was replaced by an amber stop codon (TAG), Amp^R	Wolfram Volkwein (unpublished)

Amp^R, Cam^R, and Gm^R are ampicillin, chloramphenicol and gentamycin resistance, respectively.

Table 3: Primers used in this study

Identifier	Oligonucleotide	Sequence (5' - 3')	Restriction site
P1	SacI-NP-fliC	CGC GAG CTC GCG ATT TCC TTT TAT CTT TCG ACA CG	SacI
P2	EcoRI-NP-fliC	CGC TCG AAT TCT TAA CCC TGC AGC AGA GAC AGA AC	EcoRI
P3	fliC-K199Amber-Fw	CCA CAA ACA ATA TTT AGC TTA CTG GAA TTA CC	
P4	fliC-K199Amber-Rev	GGT AAT TCC AGT AAG CTA AAT ATT GTT TGT GG	
P5	fliC-ET-Fw	GGT GGA AAC CCA ATA CGT AAT CAA CGA CTT GCA ATA TAG GAT AAC GAA TCA ATT AAC CCT CAC TAA AGG GCG	
P6	fliC-ET-Rev	ATC AGG CAA TTT GGC GTT GCC GTC AGT CTC AGT TAA TCA GGT TAC AAC GAT AAT ACG ACT CAC TAT AGG CTC	
P7	fliC-chk-Fw	GAT AAC AGG GTT GAC GGC GAT TG	
P8	fliC-chk-Rev	CTT ATC CGG CCT ACA AAA ATG TGC	
P9	efp-K34(Amber)-OL-Fw	GTA AAA CCG GGT TAG GGC CAG GCA TTT	
P10	efp-K34(Amber)-OL-Rev	AAA TGC CTG GCC CTA ACC CGG TTT TAC	
P11	XbaI-NP-efp-Fw	AGT CTA GAA TGC GCA GCA ACA TAC TCA GTG	XbaI
P12	BamHI-efp-GS-His6-Rev	CAG CGG ATC CAC TAG TTC TAG ATT AGT GAT GGT GAT GGT GAT GGC TGC CCT TCA CGC GAG AGA CGT ATT CAC C	BamHI

3 A Versatile Toolbox for the Control of Protein Levels Using $N^{ε}$-Acetyl-L-lysine Dependent Amber Suppression

Reproduced from ACS Synthetic Biology:

Full-text article: https://pubs.acs.org/doi/abs/10.1021/acssynbio.7b00048

Supplementary material: https://pubs.acs.org/doi/suppl/10.1021/acssynbio.7b00048

A Versatile Toolbox for the Control of Protein Levels Using N^ε-Acetyl-L-lysine Dependent Amber Suppression

Wolfram Volkwein, Christopher Maier, Ralph Krafczyk, Kirsten Jung,*[ORCID] and Jürgen Lassak*[ORCID]

Center for integrated Protein Science Munich (CiPSM) at the Department of Biology I, Microbiology, Ludwig-Maximilians-Universität München, Großhaderner Strasse 2-4, 82152 Martinsried, Germany

[S] *Supporting Information*

ABSTRACT: The analysis of the function of essential genes *in vivo* depends on the ability to experimentally modulate levels of their protein products. Current methods to address this are based on transcriptional or post-transcriptional regulation of mRNAs, but approaches based on the exploitation of translation regulation have so far been neglected. Here we describe a toolbox, based on amber suppression in the presence of N^ε-acetyl-L-lysine (AcK), for translational tuning of protein output. We chose the highly sensitive luminescence system LuxCDABE as a reporter and incorporated a UAG stop codon into the gene for the reductase subunit LuxC. The system was used to measure and compare the effects of AcK- and N^ε-(*tert*-butoxycarbonyl)-L-lysine (BocK) dependent amber suppression in *Escherichia coli*. We also demonstrate here that, in combination with transcriptional regulation, the system allows protein production to be either totally repressed or gradually adjusted. To identify sequence motifs that provide improved translational regulation, we varied the sequence context of the amber codon and found that insertion of two preceding prolines drastically decreases luminescence. In addition, using LacZ as a reporter, we demonstrated that a strain encoding a variant with a Pro-Pro amber motif can only grow on lactose when AcK is supplied, thus confirming the tight translational regulation of protein output. In parallel, we constructed an *E. coli* strain that carries an isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible version of the AcK-tRNA synthetase (AcKRS) gene on the chromosome, thus preventing mischarging of noncognate substrates. Subsequently, a diaminopimelic acid auxotrophic mutant ($\Delta dapA$) was generated demonstrating the potential of this strain in regulating essential gene products. Furthermore, we assembled a set of vectors based on the broad-host-range pBBR *ori* that enable the AcK-dependent amber suppression system to control protein output not only in *E. coli*, but also in *Salmonella enterica* and *Vibrio cholerae*.

**KEYWORDS:** *translational regulation, gene silencing, PylRS, Salmonella enterica, Vibrio cholerae, genetic screen*
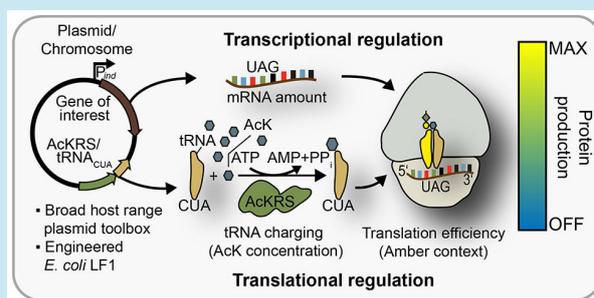
ost gene silencing strategies target gene expression at the level of transcription by exploiting the activity of transcriptional regulators, *e.g.*, TetR,[1] LacI[2] or the CRISPRi system.[3] Alternatively, gene expression can be tightly regulated post-transcriptionally by the use of, *e.g.*, microRNAs or small interfering RNAs.[4] Here, we show that translational regulation can also be employed to sensitively adjust protein levels.

Canonically, 20 different amino acids can be utilized for polypeptide synthesis by the translational apparatus. However, the genetic code can be translationally expanded to direct the incorporation of two natural but noncanonical amino acids (ncAAs), namely selenocysteine (Sec)[8] and pyrrolysine (Pyl).[9] Both Sec-tRNA and Pyl-tRNA recognize codons—UGA and UAG, respectively—that are usually reserved for translation termination. Whereas Sec is found in a variety of proteins in all domains of life,[10,11] Pyl is restricted to a small number of proteins in a few archaeal and bacterial species.[12] In methanogenic archaea such as *Methanosarcina barkeri*, the pyrrolysyl tRNA-synthetase (PylRS) charges its cognate tRNA$_{\text{CUA}}$ (encoded by *pylT*) (Figure 1A) to suppress the

amber stop codon UAG.[13] Due to its substrate promiscuity and its narrow distribution, PylRS has become a powerful means of expanding the genetic code to include synthetic ncAAs.[14] Today, over 150 substrates can be incorporated into proteins by means of amber suppression.[15,16] This in turn has enabled researchers not only to investigate post-translational modifications,[17] but also to introduce fluorescent labels,[18] and ultimately allows one to engineer proteins with modified functions. In 2015 Rovner *et al.* introduced the orthogonal translation system consisting of an aminoacyl-tRNA synthetase and its cognate UAG-reading tRNA from *Methanococcus janaschii*, which is optimized for the incorporation of *p*-acetyl-L-phenylalanine, into an *E. coli* MG1655 derivative.[19] With this strain, they found that three amber mutations, one in each of the open reading frames of the essential genes *murG, dnaA* and *serS*, can be rescued by an external supply of 4-acetyl-L-phenylalanine. Along the same line Yusuke Kato and co-workers demonstrated
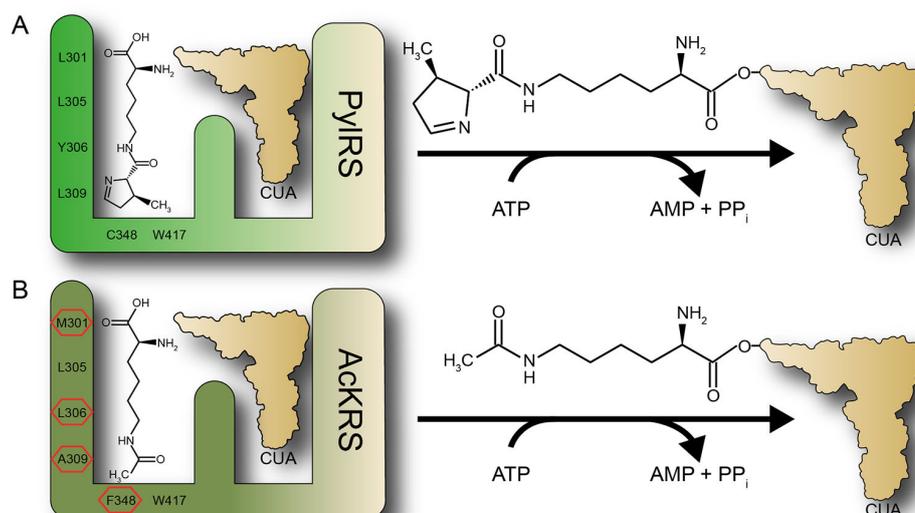
Figure 1. PylRS and AcKRS based amber suppression. (A) The pyrrolysyl-tRNA synthetase (PylRS) of *Methanosarcina mazei* loads its natural substrate pyrrolysine onto its cognate suppressor tRNA$_{CUA}$ (encoded by *pylT*). The six amino acids that form the pyrrole ring binding pocket are indicated.[5,6] (B) Acetyl lysyl-tRNA synthetase (AcKRS), a PylRS variant bearing the highlighted amino acids at positions 301, 306, 309 and 348, which enable it to interact specifically with N^ε-acetyl-L-lysine (AcK).[7]

the tunable translational control using site-specific 3-iodo-L-tyrosine incorporation in *Escherichia coli*.[20,21] Thus, the amber suppression system can be used to efficiently deplete essential gene products at the translational level.

In the present study we describe the assembly of a versatile toolbox for control of protein levels in diverse bacterial species using N^ε-acetyl-L-lysine (AcK)-dependent amber suppression.

RESULTS AND DISCUSSION

Selection of N^ε-Acetyl-L-lysine for Amber Suppression. Our study began with the selection of a suitable substrate to be incorporated by PylRS-mediated amber suppression. We chose N^ε-acetyl-L-lysine (AcK) for this purpose because an evolved variant of PylRS has been reported to charge tRNA$_{CUA}$ with AcK,[7,22] and AcK occurs frequently in bacterial proteins as a product of the post-translational modification of the ε-amino group of lysine residues.[23−26] Lysine acetylation can be catalyzed by protein acetylases such as YfiQ (PatZ) in *E. coli*,[27] but can also result from a nonenzymatic reaction with acetyl phosphate.[24,26] Note that protein acetylation is not a permanent modification, but can be reversed by deacetylases, *e.g.*, CobB, in bacteria.[27] This would allow the production of an unaltered protein when a native lysine codon is replaced by UAG and read by AcK-dependent amber suppression.[22,28] Apart from the fact that AcK is found naturally in proteins, it has no effect on bacterial growth (Supplementary Figure S1), is commercially available and is relatively cheap. For example, supplementing 1 L of culture medium with 1 mM AcK costs around 5€, and is thus about as expensive as IPTG. Taken together, these considerations make AcK an ideal substrate. We used the previously described PylRST expression system from *Methanosarcina mazei*[29] and introduced the mutations reported by Umehara *et al.*[7] to render the enzyme specific for AcK, resulting in the N^ε-acetyl lysyl-tRNA synthetase/tRNA$_{CUA}$ pair (AcKRST) (Figure 1B).
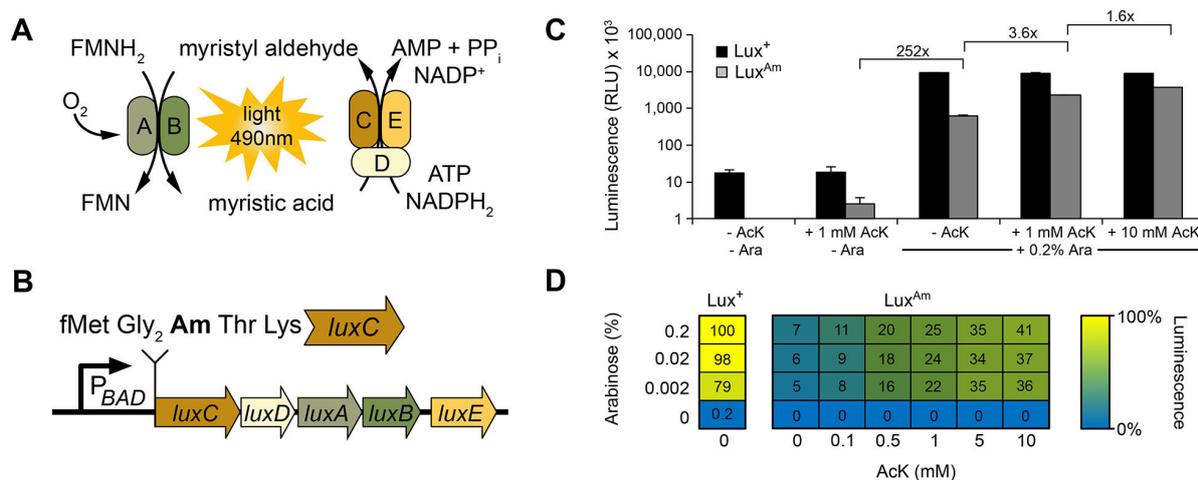
Development of a Luminescence-Based Amber Suppression Reporter. Amber suppression is conventionally quantified using either green fluorescent protein (GFP)[30] or β-galactosidase (LacZ)[7] as the reporter. We decided to use luminescence as the readout, and chose the system based on expression of the *luxCDABE* operon from *Photorhabdus luminescens* (Lux) (Figure 2A, B). In this system *luxA* and *luxB* encode the α- and β-subunits of the luciferase heterodimer (Figure 2A), while the reductase, transferase and synthase specified by LuxC, D and E, respectively, together form the complex that synthesizes and regenerates the myristyl aldehyde substrate required for the bioluminescent reaction (Figure 2A).[31,32] The Lux system is highly sensitive and allows one to quantify amber suppression in living cells in real time.[33]

To make luminescence development not only dependent on amber suppression but also tunable at the transcriptional level we cloned the *luxCDABE* cassette into pBAD/HisA, placing it under the control of the L-arabinose (Ara)-inducible promoter P$_{BAD}$. The stop sequence TAG was inserted at position 3 in the *luxC* gene (Figure 2B)

We note that alterations at the 5′ end of coding sequences might affect translation efficiency due to changes in mRNA folding[34−36] and accordingly one should be careful when inserting TAG here. On the other hand, the region close to the protein start site often lacks structural features important for protein function and therefore tolerates manipulation without perturbation of functionality. In the case of LuxC, secondary structure prediction did not identify any α-helix or β-sheet in the first six amino acids and thus giving first hint for a tolerance toward insertions. Furthermore, there is relative low sequence conservation in this region and some LuxC homologues even have N-terminal extensions, *e.g.*, LuxC of *Photobacterium phosphoreum* altogether indicating this region to be suitable for TAG insertion.

We analyzed Lux-amber reporter (Lux^{Am}) activity in *E. coli* BW25113 which, in addition to pBAD/HisA-Lux(Am), harbors the vector pACYCDuet-AcKRST that codes for AcKRS and its cognate tRNA$_{CUA}$ (Figure 1B, 2C, Table 1). We grew the cells in LB and supplemented the medium with either 0.2% Ara, 1

Figure 2. A luminescence-based amber suppression bioreporter (A) Basic principle of the luminescence system (Lux) from *Photorhabdus luminescens*. The *lux* genes are organized in the *luxCDABE* operon, which encodes the multienzyme complex LuxCDE and the luciferase heterodimer LuxAB. These two orchestrate substrate synthesis/regeneration of myristyl aldehyde and generate luminescence, respectively. (B) The Lux^Am luminescence reporter: The *luxCDABE* operon of *P. luminescens* was placed under the control of the Ara-inducible promoter P_{BAD}, and the third codon of the open reading frame of *luxC* was replaced by an amber stop codon, resulting in the Lux-amber reporter (Lux^Am). (C) Luminescence production of Lux^Am in *E. coli* BW25113. A construct in which the wildtype *luxC* ORF was left unaltered was used as a positive control (Lux^+): Cells harboring the corresponding Lux reporter plasmid in combination with pACYCDuet-AcKRST, which codes for the AcKRS and its cognate tRNA_CUA, were grown in LB overnight and luminescence production was monitored in response to Ara and/or AcK, and is given in relative light units (RLU). Error bars represent the standard deviation of data from three different experiments. (D) Parameter optimization for Lux^+- and Lux^Am-encoding *E. coli* strains. Ara and AcK concentrations were varied and maximal luminescence was determined. The resulting luminescence values ranged from "below detection" (0) RLU (blue) to 9.0×10^6 RLU (yellow).

mM AcK or both compounds together (Figure 2C). The Lux reporter plasmid pBAD/HisA-Lux, in which the wildtype *luxC* ORF was unchanged, served as a positive control (Lux^+). The maximal luminescence was determined and is given in relative light units (RLU). As expected, cells harboring pBAD/HisA-Lux produced light after induction with 0.2% Ara, irrespective of the presence or absence of AcK (Figure 2C). In contrast, in Lux^Am-expressing *E. coli* cells, luminescence was best stimulated by a combination of the two substances. However, under Ara-inducing conditions, we detected significant light emission (1×10^6 RLU) even in the absence of AcK. This result can be explained by translational read-through mediated by binding of near-cognate tRNAs[37] or by misaminoacylation of tRNA_CUA.[38]

Strikingly, while Lux^+-harboring cells showed a certain degree of light production in the absence of Ara, the Lux^Am-containing cells remained completely dark (Figure 2C). Thus, the combination of transcriptional and translational control completely abolishes protein output, in line with an earlier report.[19]

In the next step we systematically varied both the concentration of Ara and AcK and compared the luminescence of Lux^Am with Lux^+ cells (Figure 2D, Supplementary Figure S2). As before, light production by the Lux^+ strain was essentially unaffected by increasing concentrations of AcK (Supplementary Figure S2), and reached a maximum when the growth medium contained ≥0.02% Ara (Figure 2C, Supplementary Figure S2). In contrast, luminescence in the Lux^Am strain responds to the increasing presence of AcK, which allows for adjustment of the light output over a range that extends from "below detection" (blue) to more than 3.8×10^6 RLUs (green), corresponding to 41% of the maximal light intensity that we could measure with Lux^+ (Figure 2C, D, Supplementary Figure S2).

Additionally, we used the Lux^Am luminescence reporter to investigate the potential of N^ε-(*tert*-butoxycarbonyl)-L-lysine (BocK) as alternative substrate for translational protein level control. BocK can be incorporated into the genetic code employing the wildtype PylRST.[39] Like AcK, BocK is commercially available and comparably cheap. In contrast to AcKRS, PylRS is not promiscuous to charge tRNA_CUA with natural amino acids,[7,40] and thus we were curious to compare the two aminoacyl-tRNA-synthetases in our test system with respect to misacylation and amber suppression efficiency. To this end we transformed *E. coli* BW25113 cells harboring pBAD/HisA-Lux(Am) in combination with pACYCDuet-PylRST,[29] pACYCDuet-AcKRST or the empty vector pACYCDuet-1, respectively, in order to discriminate between misacylation and translational read through. Subsequently, light production was analyzed in the resulting strains in the presence of 0.2% Ara with and without the supplement of 10 mM BocK and AcK, respectively (Supplementary Figure S3). Independent of PylRS and AcKRS we measured a base level luminescence of about 2.0×10^5 RLU. This value further increased in cells harboring pACYCDuet-AcKRST to about 5.0×10^5 RLU without the addition of AcK, indicating mischarging of tRNA_CUA by endogenously available noncognate substrates. By contrast, in cells bearing pACYCDuet-PylRST we did not observe such an increase in luminescence clearly showing, that natural amino acids are not recognized by PylRS. Therefore, BocK might be a good alternative to AcKRS in order to control protein output translationally. Note that the BocK supplement resulted in an approximately 5-fold increase in luminescence in cells expressing *pylRS/pylT*. This was slightly (1.0×10^6 RLU) lower compared to the maximal value measured with AcKRST. Thus, and similar to AcK dependent amber suppression, the translational regulation with BocK/PylRST allows a tight

protein output control when combined with inducible gene expression.

Adjustment of the Amber Codon Context Permits Modulation of Protein Output. In the preceding section we showed that the Lux^{Am} strain produces significant levels of luminescence even in the absence of the cognate AcKRS substrate AcK (Figure 2C), which prompted us to ask whether an additional amber codon in the reporter gene might eliminate this background signal. Indeed, the introduction of a second TAG into *luxC* completely abolished light production, despite supplementation with AcK (Table 1). This implies that the second amber codon brings translation of *luxC* to a halt. We therefore undertook further engineering to enable suppression to compete with termination.[41,42]

Table 1. Influence of the 5′ Amber Context on Luminescence Production[a]

motif	AcK (1 mM)	mean × 10³/SD × 10³ (RLU)
Lux$^+$	−	7165 ± 392
Lux$^+$	+	7252 ± 486
Lux^{Am}	−	689 ± 26
Lux^{Am}	+	2406 ± 77
Lux^{2Am}	−	0
Lux^{2Am}	+	0
Lux$^{RA(Am)}$	−	919 ± 68
Lux$^{RA(Am)}$	+	3400 ± 22
Lux$^{DA(Am)}$	−	194 ± 33
Lux$^{DA(Am)}$	+	496 ± 51
Lux$^{HH(Am)}$	−	1434 ± 21
Lux$^{HH(Am)}$	+	3717 ± 633
Lux$^{KDP(Am)}$	−	20 ± 5
Lux$^{KDP(Am)}$	+	35 ± 26
Lux$^{PP(Am)}$	−	112 ± 35
Lux$^{PP(Am)}$	+	691 ± 72
Lux$^{KDPP(Am)}$	−	3 ± 3
Lux$^{KDPP(Am)}$	+	78 ± 7

[a]Cells harboring the corresponding Lux reporter plasmid in combination with pACYCDuet-AcKRST were grown in LB supplemented with 0.2% L-arabinose overnight and luminescence production was monitored in response to AcK and is given in relative light units (RLU).

It is known that the context of an internal stop codon can have a huge impact on translational readthrough.[43,44] This led us to analyze the influence of sequence motifs known to have significant effects on amber suppression. AcK incorporation into the nascent chain competes with translational termination mediated by the release factor RF1. This competition is reflected by the fact that even at saturating AcK concentrations the luminescence of Lux^{Am} cells does not reach the level observed in the Lux$^+$ *E. coli* strain (Figure 2C). However, the efficiency of recognition of the amber stop codon by RF1 is modulated by the base that follows it. While in *E. coli* a 3′ uracil best ensures termination, amber suppression is favored by a neighboring adenine.[45,46] The threonine sense codon ACU that follows the UAG in the Lux^{Am} reporter should therefore promote suppression and was left intact. Accordingly, we focused on the sequence context 5′ to the amber codon. Although the codon directly upstream of the stop site is of importance,[47] in 1994 Mottagui-Tabar and co-workers found that the penultimate codon has a major influence on termination in *E. coli*.[43] In that study, the authors described a

hierarchy of suppression, with arginine codons in this position showing the lowest and aspartate codons stimulating the highest suppression efficiency when UGA was the stop codon. On the basis of these findings, we decided to investigate the influence of R$_{-2}$ A$_{-1}$ and D$_{-2}$ A$_{-1}$ on amber suppression. We also included H$_{-2}$ H$_{-1}$, a context motif that was identified in a screen for efficient incorporation of ncAAs.[44] Moreover, we chose K$_{-3}$ D$_{-2}$ P$_{-1}$ and P$_{-2}$ P$_{-1}$, as well as a combination of the two (K$_{-4}$ D$_{-3}$ P$_{-2}$ P$_1$), to precede the stop site. These sequence contexts are known to cause ribosome stalling and should thus interfere with translation termination.[48−50]

All Lux^{Am} variants were constructed essentially as depicted in Figure 2B, by insertion of the motif after Gly$_2$ in the *luxC* gene. Luminescence was measured under P$_{BAD}$-inducing conditions as described above, and in the presence and absence of 1 mM AcK, respectively (Table 1).

As before (Figure 2C), the strain containing the Lux^{Am} reporter construct produced a maximal level of luminescence of about 2.4 × 10⁶ RLU upon supplementation with AcK. This is a 3.5-fold increase relative to cells that were grown without AcK. In comparison, the maximal light intensities emitted by Lux-RA(Am) and Lux-DA(Am) cells grown in the presence of AcK were 3.4 × 10⁶ and 5.0 × 10⁵ RLU, respectively, and differed significantly from each other. This might be explained by altered *luxC* expression, as regulatory elements are located within the first 50 codons.[51] Regardless of these differences, the relative increases in luminescence in the presence of AcK were comparable, at 3.7-fold for Lux$^{RA(Am)}$ and 2.6-fold for Lux$^{DA(Am)}$. Mottagui-Tabar *et al.* showed that, in principle, arginine and aspartate in the second last position to UAG follow the same rules as when UGA is the stop codon, but that the effect with UAG was less pronounced.[43] This latter factor may account for the small differences in light production seen between amber-suppressed Lux$^{RA(Am)}$ and Lux$^{DA(Am)}$.

In agreement with the findings of Pott *et al.*, we observed increased luminescence for Lux$^{HH(Am)}$ compared to the original Lux^{Am} bioreporter.[44] However, this is largely attributable to increased translational read-through, as indicated by the increase in light production seen in the absence of AcK (Table 1). The analysis of amber context motifs that promote ribosome stalling—KDP, PP and KDPP—revealed a strong decrease in luminescence, which was unaffected by the presence of AcK. This decrease is presumably the result of the recruitment of ribosome rescue systems, such as tmRNA, which ultimately lead to abortion of translation.[49] Of the three arrest motifs tested, only the PP motif present in the Lux$^{PP(Am)}$ reporter had a significant effect on amber suppression. In this case, the addition of AcK to the growth medium was correlated with an increase of >6-fold in luminescence—the highest relative change observed with any of the motifs tested. Thus, a diprolyl amber context would appear to be especially beneficial in experiments where coupling with transcriptional regulation must be avoided.

A Diproline Amber Context Allows Tight Regulation of β-Galactosidase Protein Levels. We next tested whether the integration of an amber stop codon into the bacterial chromosome might allow for tighter regulation of protein production in *E. coli*. For these experiments, we chose the β-galactosidase LacZ, which cleaves the disaccharide lactose into D-galactose and D-glucose. This enzyme is essential for growth of *E. coli* cells on lactose as sole carbon source. We constructed two *E. coli* strains, bearing either an Am (TAG) or a PP(Am) (CCG CCG TAG) insertion after the ninth codon in the *lacZ*

coding sequence (Figure 3A). Similar to LuxC, our decision for the insertion position was based on a lack of structural elements as well as comparatively low sequence conservation in the first ten N-terminal amino acids of LacZ.[54] In addition to the
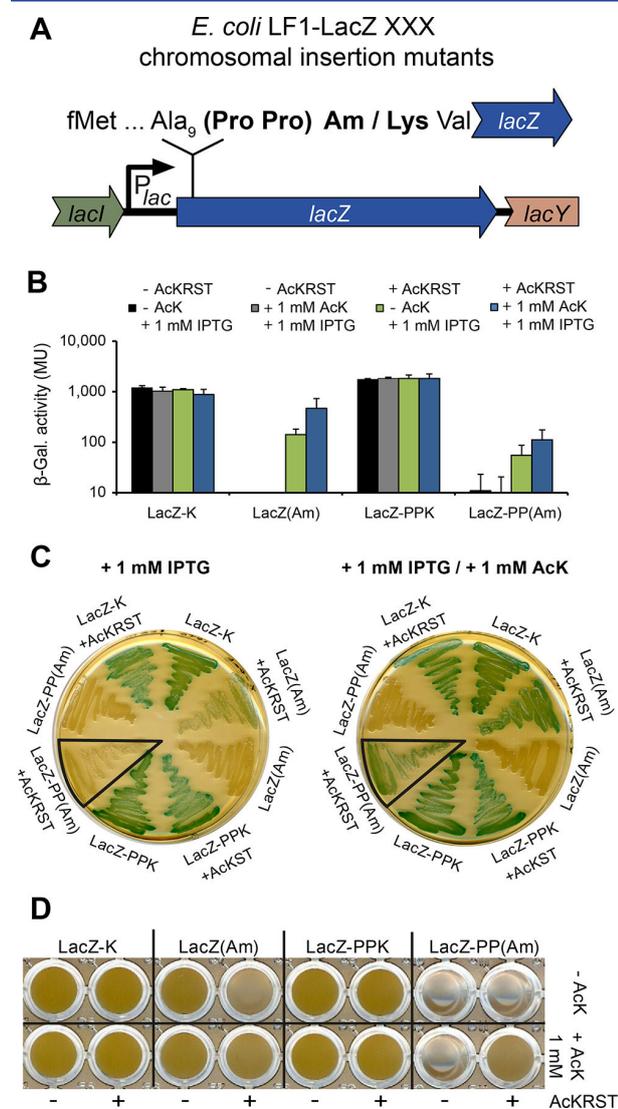

Figure 3. Regulation of LacZ levels in *E. coli* by amber suppression. (A) Genetic organization of the *lacZ* gene in *E. coli* LF1-LacZ strains. *E. coli* LF1 strains were constructed by inserting either an amber (LF1-LacZ(Am)) or PP amber sequence (LF1-LacZ-PP(Am)) after the ninth codon of the *lacZ* gene. Corresponding lysine codon insertions served as positive controls—LF1-LacZ-K and LF1-LacZ-PPK, respectively. (B) Quantitative analysis of amber suppression of β-galactosidase LacZ. To measure the β-galactosidase activity, cells were grown in LB overnight and reporter activity was determined according to Miller,[52] in the presence/absence of the amber suppression machinery (AcKRST). Cells were incubated with IPTG, AcK or a combination of both substances. The activity is given in Miller Units (MU). (C) Qualitative analysis of amber suppression of LacZ. *E. coli* LF1 strains described in (A) were grown on LB agar supplemented with 80 μM X-Gal, 1 mM IPTG and 1 mM AcK as indicated. (D) Growth analysis of *E. coli* LF1 strains with lactose as sole carbon source. LF1 strains with and without AcKRST were grown in M9 minimal medium[53] supplemented with 20 mM lactose in the presence/absence of 1 mM AcK in microtiter plates.

resulting strains LF1-LacZ(Am) and LF1-LacZ-PP(Am) and to further exclude any effects on protein activity we engineered two strains containing a lysine codon instead of TAG for use as positive controls (LF1-LacZ-K and LF1-LacZ-PPK). These two strains were compared with wildtype MG1655 cells and found to be phenotypically indistinguishable with respect to both growth on lactose and β-galactosidase activity (data not shown). By contrast, both LF1-LacZ(Am) and LF1-LacZ-PP(Am) exhibited strongly diminished LacZ production (Figure 3B). Moreover, whereas β-galactosidase activity was still detectable in LF1-LacZ(Am) in the absence of AcK (141 MU; MU = Miller Units), only background levels (55 MU) were measured in strain LF1-LacZ-PP(Am). As expected, addition of 1 mM AcK to the medium increased LacZ production in both cases, to 468 MU (LF1-LacZ(Am)) and 111 MU (LF1-LacZ-PP(Am)), respectively. In agreement with our previous finding for the luminescence reporter Lux^PP(Am), the PP(Am) context resulted in a more pronounced decrease in translational output, confirming that it further reduces leakiness in the system. The virtually complete abrogation of LacZ production is further illustrated by the fact that the color change from white to blue of *E. coli* colonies encoding LacZ-PP(Am) is strictly dependent on AcK (Figure 3C). In contrast, the LF1-LacZ(Am) strain turns blue when grown on LB/X-Gal agar plates containing the *lacZ* inducer IPTG (Figure 3C). Furthermore, LF1-LacZ-PP(Am) containing AcKRST was only able to grow on lactose when AcK was concomitantly supplied, while the residual β-galactosidase activity in LF1-LacZ(Am) cells (Figure 3B) was sufficient to promote growth even in the absence of AcK (Figure 3D). Thus, these data further corroborate our findings with the Lux reporters (Figure 2C, D; Table 1).

The ability to tightly control LacZ synthesis in this way demonstrates the potential of our system for AcK-dependent translational regulation of essential genes with a specific amber context.

Chromosomal Integration of AcKRS Enhances Regulated Amber Suppression. Having shown that amber suppression is a potent molecular tool for the regulation of translational output, we set out to engineer *E. coli* in such a way that it can serve as a host for investigation of the function of essential genes in an AcK-dependent manner. In a previous publication we reported the construction of a strain named LF1, into which we had inserted a kanamycin resistance cassette and a wildtype copy of *rpsL* (conferring streptomycin sensitivity) into the *lac* locus.[55]

The two genes are flanked by FRT sites that mediate Red/ET (Gene Bridges) recombination, which permits scarless insertion of any DNA sequence of interest within 1 day. We used this method to integrate *pylS*/*pylT* into the LF1 strain. While *pylS* was set under the control of P_{lac}, *pylT* (tRNA$_{CUA}$) expression was kept constitutively expressed employing P_{proK}. The genomic context of the resultant insert in the new strain LF1-AcKRST is depicted in Figure 4A. To test amber suppression in *E. coli* LF1-AcKRST, the cells were transformed with pBAD/HisA-Lux(Am), and luminescence was monitored in response to the external supply of IPTG and AcK. As expected, no luminescence was detectable in the absence of both substances. Similarly, supplementation with 1 mM AcK alone did not lead to any measurable light production, while addition of IPTG to the medium on its own resulted in only very weak luminescence of approximately 1.0×10^4 RLU. In stark contrast, the simultaneous presence of AcK and IPTG

A

B
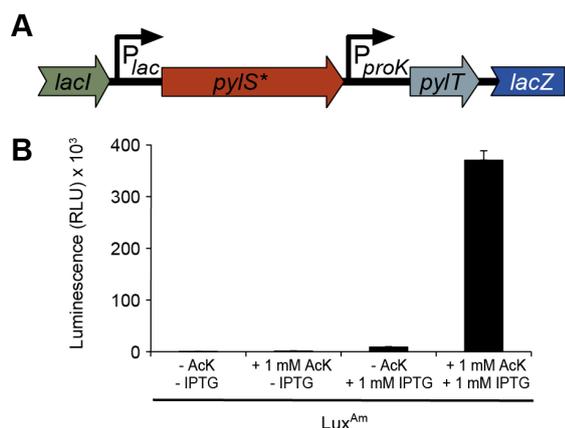
Figure 4. *E. coli* LF1-AcKRST. (A) Genetic organization of the *E. coli* strain LF1-AcKRST. N^ε-acetyl lysyl-tRNA-synthetase (*pylS**) was cloned under the control of the IPTG-inducible *lac* promoter (P_{lac}), while the cognate $tRNA_{CUA}$ (*pylT*) was placed under the control of the constitutive promoter P_{proK}. (B) AcK-dependent amber suppression in *E. coli* LF1-AcKRST. Cells were transformed with the Lux-amber reporter (Lux^{Am}). The resulting strains were then grown overnight in LB, and luminescence development was recorded in response to IPTG and AcK. The maximal luminescence normalized to the OD_{600} from a 16 h time course experiment is depicted. Error bars represent the standard deviation of data from three different experiments.

increased the light intensity to about 3.7×10^5 RLU. Thus, the *E. coli* strain LF1-AcKRST provides an on/off switch for the production of specific proteins. It should be noted that 0.2% Ara was always added to the medium, so that the *luxCDABE* operon was fully induced under all conditions tested. Recall that when *pylS**/*pylT* was provided *in trans* (pACYCDuet-AcKRST) significant luminescence was observed in the absence of AcK (Figure 2C). Thus, the reduction in *pylS**/*pylT* copy

number from about 20−30 in the plasmid-based system[56] to one in the chromosome almost completely abolishes mischarging of $tRNA_{CUA}$. Furthermore, the coupling of chromosomal integration of the amber suppression machinery (AcKRST) with IPTG-inducible expression of *pylS** eliminates the physiological burden that goes along with the constitutive expression of orthogonal tRNAs and aminoacyl tRNA synthetases[57] (Supplementary Figure S4).

To proof functionality of the parental strain *E. coli* LF1-AcKRST in regulating the expression of essential gene products, we aimed to generate a mutant with a defect in cell wall biosynthesis. For this purpose, we chose the gene encoding dihydrodipicolinate synthase DapA, being crucial for the generation of meso diaminopimelic acid (DAP).[58] Accordingly, a *dapA* deletion causes DAP auxotrophy. To this end the *E. coli* strain LF1-AcKRST/*dapA:cam*^{r} was generated using Red/ET (Gene Bridges) recombination. As expected LF1-AcKRST/ *dapA:cam*^{r} (DAP^-) can only grow in LB supplemented with as little as 75 *μ*M DAP (Supplementary Figure S5).

To trace the observed auxotrophy back to the deletion of *dapA* we constructed a plasmid pBAD/HisA-dapA coding for an N-terminal His_6-tagged copy of DapA. Ectopic expression of His_6-DapA in LF1-AcKRST/*dapA:cam*^{r} permitted growth in the absence of DAP after induction with 0.2% Ara (Figure 5A). Beside the fact, that this result clearly links the observed phenotype to the gene deletion, it also demonstrates that the N-terminal extension of the protein sequence does not interfere with the enzymatic activity of DapA. Consequently, a plasmid pBAD/HisA-dapA(Am) was derived from pBAD/HisA-dapA. This daughter plasmid deviates from the parental only by an amber motif as a third codon in the original *dapA* ORF which in turn follows a 39 amino acids (aa) long N-terminal extension encoded in the pBAD/HisA backbone.

We investigated the power of coupled transcriptional and translational regulation by controlling the growth of the DAP^-

A

B
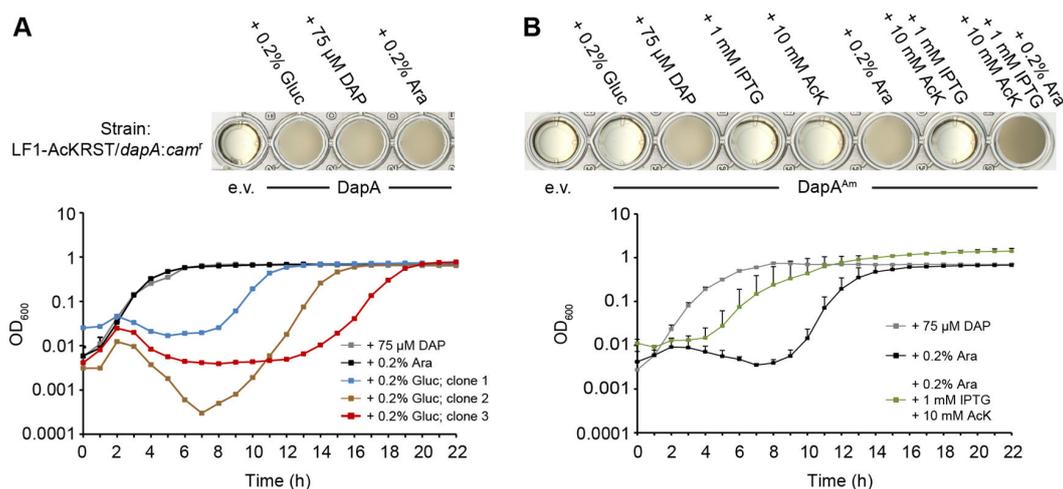
Figure 5. Transcriptional and translational control of DAP auxotrophy. Growth analysis of the LF1-AcKRST/*dapA:cam*^{r} strain harboring the pBAD/HisA-dapA or pBAD/HisA-dapA(Am) plasmid. Cells were grown in a microtiter plate over a time course of 22h in LB with the indicated supplements. Afterwards pictures of the bacterial cultures were taken. (A) Growth curve and picture of the LF1-AcKRST/*dapA:cam*^{r} strain harboring a plasmid borne copy of *dapA* (pBAD/HisA-dapA), or the empty vector e.v. (pBAD/HisA) as negative control, in response to 0.2% (w/v) L-arabinose (Ara), 0.2% (w/v) D-glucose (Gluc) or 75 *μ*M diaminopimelic acid (DAP). Ara and DAP growth curves are depicted as triplicates, Gluc growth curves are depicted as single measurements. (B) Growth curve and picture of the LF1-AcKRST/*dapA:cam*^{r} strain harboring the empty vector (e.v.) as negative control or pBAD/HisA-dapA(Am) ($DapA^{Am}$) containing a plasmid borne copy of DapA with an inserted amber codon, in response to Ara, DAP, Gluc, N^ε-acetyl-L-lysine (AcK) and isopropyl β-D-1-thiogalactopyranoside (IPTG) as indicated. Error bars in the growth curve represent the standard deviation of data from three different experiments.
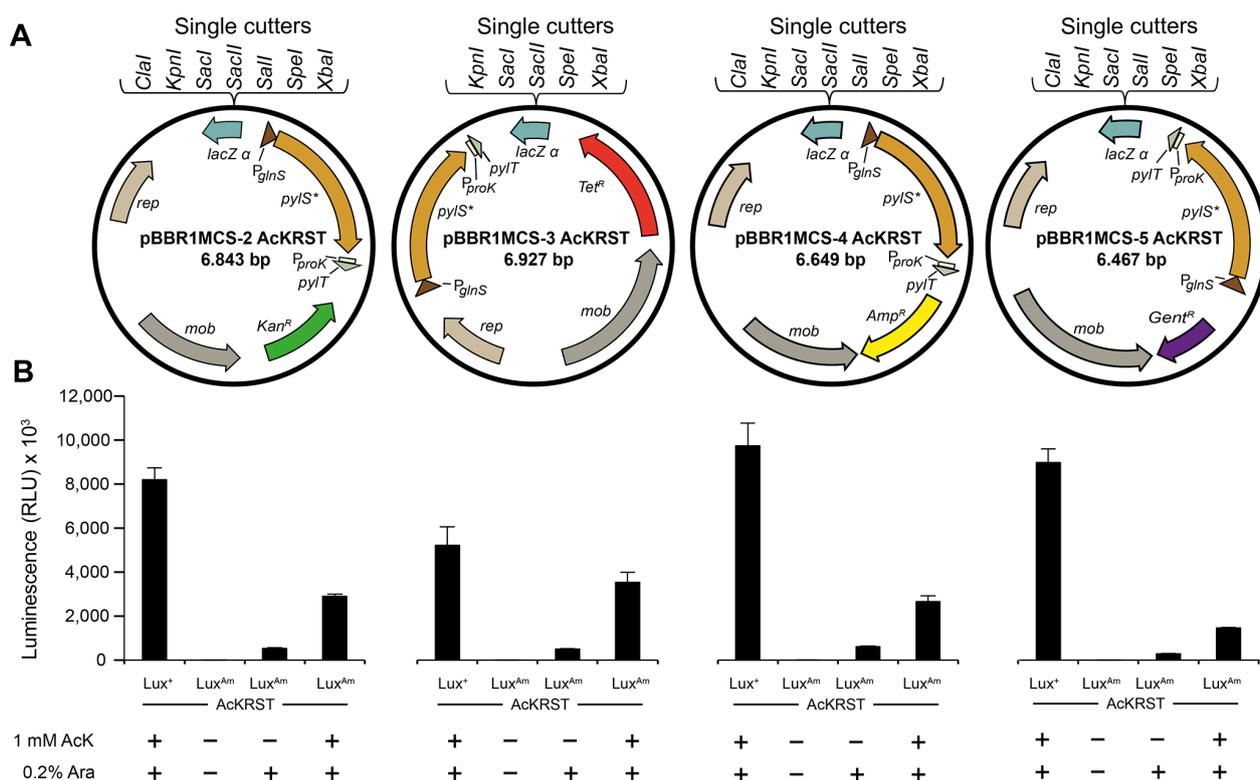
Figure 6. A set of broad-host-range cloning vectors for amber suppression in diverse bacteria. (A) Maps of pBBR1MCS-X AcKRST derivatives. The *pylS*/*pylT* cassette was integrated adjacent to the multiple cloning site in the available broad-host-range plasmids pBBR1MCS-2 to 5. (B) Analysis of amber suppression in cells bearing the pBBR1MCS-X AcKRST derivatives. pBBR1MCS-X AcKRST vectors were transferred into *E. coli* BW25113 together with the Lux-amber reporter (Lux^Am) or Lux reporter (Lux^+), as positive control. The resulting strains were grown overnight in LB and luminescence development was recorded in response to Ara and AcK. The maximal luminescence normalized to the OD_{600} from a 16 h time course experiment is depicted. Error bars represent the standard deviation of data from three different experiments.

strain in the presence of a plasmid borne gene copy. Therefore, the *E. coli* DAP⁻ strain LF1-AcKRST/*dapA:cam*^r was transformed with pBAD/HisA-dapA and pBAD/HisA-dapA(Am), respectively and cells were initially cultivated overnight in LB supplemented with 75 μM DAP. The next day cells were inoculated into LB + 0.2% D-glucose (Gluc) (to catabolically repress P_{BAD} induction), LB + 75 μM DAP and LB + 0.2% Ara (to induce P_{BAD} induction). Growth of the strain harboring the pBAD/HisA-dapA(Am) was additionally monitored in LB + 1 mM IPTG (to induce expression of P_{lac} regulated AckRS), in LB + 1 mM IPTG + 1 mM AcK (to allow Amber suppression) and ultimately in LB containing a combination of all three components (0.2% Ara, 1 mM IPTG and 1 mM AcK) (Figure 5B). As expected, this coupling of AcK dependent amber suppression with Ara induced transcriptional activation permitted growth in the absence of DAP. Notably, growth behavior of cells cultivated with 0.2% L-Ara, 1 mM IPTG and 1 mM AcK was beside a short lag-phase almost identical to what we observed for cells that were supplemented with 75 μM of DAP (Figure 5B). On the other hand, we could not detect an increase in OD_{600} in the cultures lacking either DAP or 0.2% Ara. However, the sole addition of Ara was sufficient to promote overnight growth of LF1-AcKRST/*dapA:cam*^r with pBAD/HisA-dapA(Am). Compared to cultures containing all three supplements (Ara, IPTG and AcK) we observed an extended lag phase and even an initial drop in the starting OD_{600} (Figure 5B). The latter might be explained by cell lysis due to instability of the bacterial cell wall, whereas the first 10 h

without any significant change in OD_{600} indicates the development of a suppressor mutation that promotes growth even under translationally repressing conditions. Similar results were observed when monitoring the growth of cells containing pBAD/HisA-dapA in which P_{BAD} dependent gene expression was catabolically repressed[59] (+0.2% Gluc) (Figure 5A).

Thus, the coupling of transcriptional and translational regulation does not only allow a tunable protein output over several orders of magnitude but also prevents a fast development of suppressor mutation in cases where prolonged incubation is necessary.

Four Broad-Host-Range Cloning Vectors Allow Amber Suppression in Diverse Bacterial Species. The demonstration that amber suppression permits tight regulation of protein output in *E. coli* raised the question of the system's transferability to other bacteria. To answer this question, we made use of the various pBBR1MCS-X derivatives available,[60] bearing resistance cassettes for kanamycin, tetracycline, ampicillin and gentamycin, respectively.[61] The pBBR1 origin of replication present in these plasmids is recognized in a wide range of bacterial hosts, and they are also mobilizable for conjugational transfer.[60]

Moreover, the vectors do not belong to the Inc.P, Inc.Q or Inc.W incompatibility group, and can therefore be coinherited with p15A- or ColE1-dependent replicons.[61] Consequently, we cloned the genes of the amber suppression machinery into pBBR1MCS-2 to 5, keeping the original multiple cloning sites intact, and retaining the potential for blue/white screening

(Figure 6A). The function of the resulting derivatives was validated by transformation of *E. coli* with the Lux[Am] reporter and subsequent analysis of AcK-dependent amber suppression (Figure 6B).

Next we cloned a copy of the NanoLuc luciferase (Nluc) into pBBR1MCS-2 AcKRST (Figure 7A) under the control of the P_{BAD} promoter. The plasmids generated in this way (pBBR1MCS-2 AcKRST Nluc, pBBR1MCS-2 AcKRST Nluc-(Am), pBBR1MCS-2 AcKRST Nluc PP(Am)) were then transferred into *Salmonella enterica* LT2 and *Vibrio cholerae* El Tor C6706 cells, and the wildtype NanoLuc was used as the positive control (Nluc). AcK-dependent amber suppression was analyzed with a NanoLuc variant, containing a UAG as third codon (Nluc[Am]). Again, a position close to the protein start was selected because the enzyme reportedly tolerates N-terminal fusions.[62] In addition, we constructed a variant with PP amber context (Nluc[PP(Am)]) (Figure 7A). The dependence of luminescence production in *S. enterica* and *V. cholerae* cells on supplementation with Ara and/or AcK was then analyzed. As expected, cells encoding a NanoLuc without any amber insertion emit luminescence when grown in the presence of Ara (Figure 7B, C). In contrast, significant luminescence production by cells harboring the Nluc[Am] or Nluc[PP(Am)] reporter required the simultaneous addition of Ara and AcK. As in *E. coli*, the PP(Am) context further diminished the low level of light production seen in the absence of AcK. Thus, our results are 1:1 transferable from *E. coli* to *S. enterica* and *V. cholerae* and with this, amber suppression is applicable to diverse bacteria.

CONCLUDING REMARKS

In the present study we investigated the application of AcK-dependent amber suppression for the regulation of translational outputs. We showed that coupling of transcriptional and translational regulation is an effective way to fully switch off the synthesis of specific proteins. Manipulation of the concentrations of both the transcriptional inducer and AcK permits gradual adjustment of the protein output. Furthermore, and in line with previous reports, we observed that the incorporation of an amber stop codon into a plasmid-borne gene copy resulted in a measurable level of protein synthesis by translational read-through (Figure 2C), which can be modulated by changing the amber context[43,46,49] (Table 1). Such read-through might be beneficial when reduced levels of an essential gene product are needed to maintain viability, but nevertheless cause a mutant phenotype. If necessary, the translational output can be further increased by ectopic expression of the *pylS*/pylT* amber suppression system. It should be noted that in our system mischarging becomes virtually irrelevant when *pylS*/pylT* was integrated in the chromosome and combined with inducible expression. This is exemplified by the fact that *E. coli* LF1-AcKRST cells encoding a Lux[Am] reporter remain completely dark in the absence of IPTG (Figure 4). Only the concomitant presence of IPTG and AcK allowed efficient amber suppression. Thus, chromosomal incorporation of the amber suppression machinery might generally increase specificity for ncAAs.

Altogether, our analyses of AcK-dependent amber suppression as a means of regulating protein production provide the basis for use of the system as a versatile translational control mechanism for investigation of the function of essential genes. Thus, the components available in this molecular toolbox enable fine-tuning of protein synthesis, not only in *E. coli* but
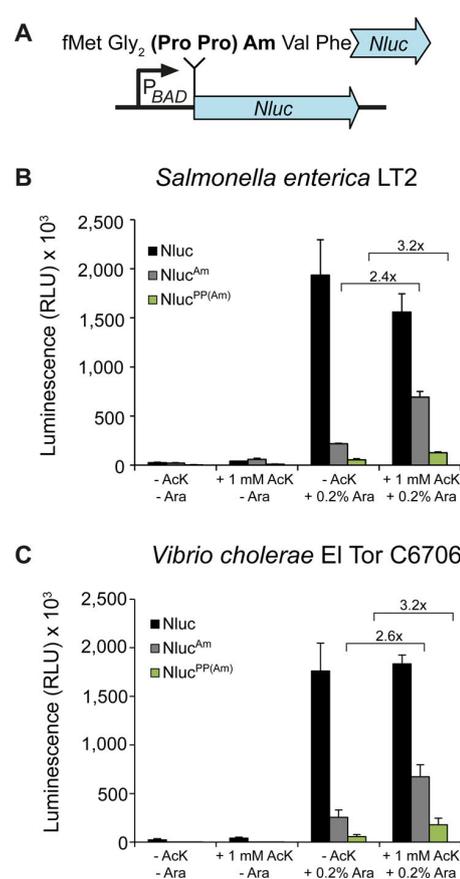
Figure 7. Amber suppression in *Salmonella enterica* and *Vibrio cholerae*. (A) An amber (Nluc[Am]) or PP amber motif (Nluc[PP(Am)]) was inserted after codon 2 of the NanoLuc gene and cloned into the pBBR1MCS-2 AcKRST backbone. A vector bearing the wildtype NanoLuc gene served as the positive control (Nluc). pBBR1MCS-2 AcKRST Nluc, pBBR1MCS-2 AcKRST Nluc(Am) and pBBR1MCS-2 AcKRST Nluc PP(Am) were transferred into *Vibrio cholerae* (B) and *Salmonella enterica* (C). Luminescence was measured from overnight cultures using the Nano-Glo luciferase assay system. Error bars represent the standard deviation of data from three different experiments.

also in other Gram-negative bacteria such as *S. enterica* or *V. cholerae*.

MATERIALS AND METHODS

Bacterial Stains, Oligonucleotides and Plasmids. Strains, primers and plasmids used in this study are listed in Supplementary Tables S1–S3. All oligonucleotides used for PCR amplification and sequencing were synthesized by Sigma-Aldrich. The sequences of all constructed plasmids were verified by Sanger sequencing.

Molecular Biological Methods. *E. coli*, *S. enterica* and *V. cholerae* cells were routinely grown in LB and where indicated in M9 minimal medium.[63] Solidification of the medium was achieved by adding 1.5% (w/v) Agar−Agar (Carl Roth). If necessary, media were supplemented with antibiotics at the following concentrations: ampicillin sodium salt (100 μg/mL), chloramphenicol (30 μg/mL), kanamycin sulfate (50 μg/mL), tetracycline hydrochloride (15 μg/mL), gentamycin sulfate (15 μg/mL), and streptomycin sulfate (50 μg/mL). For blue-white selection, LB-agar plates were additionally supplemented with 80 μM 5-bromo-4-chloro-3-indolyl β-D-galacto-

pyranoside (X-Gal) (Sigma-Aldrich) and 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Sigma-Aldrich) to induce the *lac* promoter P_{lac}. Unless indicated otherwise, 0.2% (w/v) L-arabinose was added to induce the P_{BAD} promoter.

All enzymes and kits were used according to the manufacturers' directions. DNA fragments were purified from agarose gels using a high-yield PCR cleanup and gel extraction kit (Sued-Laborbedarf). Restriction endonucleases were purchased from New England Biolabs. Sequence amplifications were performed by PCR using Phusion high-fidelity DNA polymerase (Finnzymes) or Taq DNA polymerase (New England Biolabs). All mutants were constructed by one- or two-step PCR using mismatched primer pairs.[64]

β-Galactosidase activity assays were performed as essentially described previously.[55]

Plasmid and Strain Construction. pACYCDuet-AcKRST was constructed as follows: A PCR fragment containing the respective mutations in the *pylS* gene was generated as described previously by Umehara *et al.* 2012[7] and cloned into pACYCDuet-PylRST using the internal *Xho*I and *Not*I restriction sites (Supplementary Table S2, Supplementary Table S3).

LuxCDABE-containing pBAD/HisA constructs were generated as follows: The *luxCDABE* operon from *Photorhabdus luminescens* was amplified from pBBR1MCS-5 TT-RBS-lux,[65] and then cloned into pBAD/HisA (Invitrogen) using *Nco*I and *Kpn*I as restriction sites. The resulting vector pBAD/HisA-Lux served as a template for all other *luxCDABE*-bearing plasmids used in this study. To construct *luxC* amber versions we used mismatched primer pairs (Supplementary Table S3) for overlap-extension PCR and cloned the resulting fragments into pBAD/HisA-Lux using *Nco*I and *Bsr*GI, thereby replacing the parental *luxC* fragment. The resulting plasmids, pBAD/HisA-Lux(Am), pBAD/HisA-Lux(2Am), pBAD/HisA-Lux-RA-(Am), pBAD/HisA-Lux-DA(Am), pBAD/HisA-Lux-HH(Am), pBAD/HisA-Lux-KDP(Am), pBAD/HisA-Lux-PP(Am) and pBAD/HisA-Lux-KDPP(Am) are listed and described in Supplementary Table S2.

The constructs pBAD/HisA-dapA and pBAD/HisA-dapA-(Am) were generated as follows: From genomic *E. coli* DNA, *dapA* was amplified and the amber motif was introduced using a mismatched primer (Supplementary Table S3). The resulting fragments were than cloned into the pBAD/HisA backbone using *Xho*I and *Eco*RI as restriction sites. The resulting plasmids are listed in Supplementary Table S2.

To generate pBAD/HisA-Kan and plasmid derivatives (Supplementary Table S2) the ampicillin resistance cassette was replaced by *nptI*, which was amplified from pBBR1MCS-2 (Supplementary Table S3). The PCR fragment was introduced restriction-free into pBAD/HisA as described.[66]

The broad-host-range cloning vectors containing *pylS*/pylT* were constructed as follows: *pylS*/pylT* was amplified from pACYCDuet-AcKRST and the resulting PCR fragment was cloned into pBBR1MCS-2 to 5 plasmids using either the *Nsi*I or the *Age*I restriction site (Supplementary Table S2, Supplementary Table S3).

The Nluc reporter plasmids were generated by amplifying the sequence encoding NanoLuc luciferase from pNL1.1 (Promega). The respective amber motifs were introduced using mismatched primer pairs in a standard PCR (Supplementary Table S3). The fragments obtained were each combined with the arabinose-inducible P_{BAD} promoter fragment (amplified from pBAD/HisA, *via* overlap-extension

PCR), and cloned into pBBR1MCS-2 AcKRST using *Xba*I and *Kpn*I restriction sites. The resulting plasmids, pBBR1MCS-2 AcKRST Nluc, pBBR1MCS-2 AcKRST Nluc(Am), and pBBR1MCS-2 AcKRST Nluc PP(Am) are listed in Supplementary Table S2.

E. coli LF1-LacZ-K, LF1-LacZ(Am), LF1-LacZ-PPK, and LF1-LacZ-PP(Am) strains, as well as LF1-AcKRST, were constructed essentially as described previously by Fried *et al.*[55] and are listed in Supplementary Table S1.

E. coli LF1-AcKRST/*dapA:cam*^{r} was generated by using the pRed/ET recombination technology in accordance to the technical protocol of the Quick & Easy *E. coli* Gene Deletion Kit (Gene Bridges) and is listed in Supplementary Table S1.

Lux-Based Luminescence Activity Assay. Single colonies were transferred into LB supplemented with the appropriate antibiotics, and grown aerobically overnight at 37 °C. On the next day, cells were inoculated at an OD_{600} of 0.01 into fresh LB supplemented with the corresponding antibiotics. Cells were then grown aerobically at 37 °C in a microtiter plate in a Tecan Infinite F500 system (TECAN) and luminescence development in response to AcK and Ara was monitored at 10 min intervals over the course of 16 h. Light units were normalized to OD_{600} and are thus expressed in relative light units (RLU). Subsequently RLUs were background corrected by subtracting the luminescence level detected in cells containing a vector without a *luxCDABE* operon. Each measurement was performed in triplicate.

The luminescence assay for the *E. coli* LF1-AcKRST strain was performed as described above but with the difference that the cells on the second day were inoculated in LB, containing besides the antibiotic additionally 0.2% Ara (w/v). Subsequently, luminescence development in response to 1 mM AcK and 1 mM IPTG was detected as described before.

NanoLuc Luciferase Assay in *S. enterica* and *V. cholerae*. Single colonies of *S. enterica* and *V. cholerae* were transferred into LB supplemented with kanamycin sulfate (50 μg/mL) and grown overnight aerobically at 37 °C. On the next day, these cultures were used to inoculate the main cultures to an OD_{600} of 0.01. To determine AcK- and Ara-dependent luminescence development, all main cultures were grown aerobically overnight at 37 °C in LB supplemented with kanamycin sulfate (50 μg/mL) and containing either 1 mM AcK, 0.2% Ara (w/v), or a combination of both compounds. A culture grown in the absence of both served as negative control. OD_{600} was determined from overnight cultures and the Nano-Glo luciferase assay reagent (Promega) was prepared according to manufacturer's instructions. Then, the experimental cultures were mixed 1:1 with Nano-Glo luciferase assay reagent, transferred into a 96-well microtiter plate, and luminescence was determined 5 min later with a Tecan Infinite F500 system (TECAN). Light units were normalized to OD_{600} and expressed in relative light units (RLU). Each measurement was performed in triplicate.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.7b00048.

　　Table S1−S4; Figures S1−S5 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors
*E-mail: jung@lmu.de.
*E-mail: juergen.lassak@lmu.de.

ORCID
Kirsten Jung: 0000-0003-0779-6841
Jürgen Lassak: 0000-0003-3936-3389

Author Contributions
The study was designed by W.V., K.J. and J.L. and directed as well as coordinated by K.J. and J.L.; W.V. and C.M. constructed all strains and performed the enzyme assays; R.K. performed the parameter optimization for the Lux reporters. All authors contributed to the writing of the manuscript

Notes
The authors declare no competing financial interest.

■ ABBREVIATIONS

AcK, N^ε-acetyl-L-lysine; AcKRS, N^ε-acetyl lysyl-tRNA-synthetase; ncAA, noncanonical amino acid; Sec, selenocysteine; Pyl, pyrrolysine; PylRS, pyrrolysyl-tRNA-synthetase; PylRST, pyrrolysyl-tRNA-synthase/tRNA$_{CUA}$ pair; AcKRST, N^ε-acetyl lysyl-tRNA-synthetase/tRNA$_{CUA}$ pair; Nluc, NanoLuc luciferase

■ REFERENCES

(1) Guo, X. V., Monteleone, M., Klotzsche, M., Kamionka, A., Hillen, W., Braunstein, M., Ehrt, S., and Schnappinger, D. (2007) Silencing *Mycobacterium smegmatis* by using tetracycline repressors. *J. Bacteriol.* 189, 4614–4623.
(2) Higuchi, M., Tsutsumi, R., Higashi, H., and Hatakeyama, M. (2004) Conditional gene silencing utilizing the *lac* repressor reveals a role of SHP-2 in *cagA*-positive *Helicobacter pylori* pathogenicity. *Cancer Sci.* 95, 442–447.
(3) Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., and Lim, W. A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183.
(4) Filipowicz, W., Jaskiewicz, L., Kolb, F. A., and Pillai, R. S. (2005) Post-transcriptional gene silencing by siRNAs and miRNAs. *Curr. Opin. Struct. Biol.* 15, 331–341.
(5) Kavran, J. M., Gundllapalli, S., O'Donoghue, P., Englert, M., Soll, D., and Steitz, T. A. (2007) Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proc. Natl. Acad. Sci. U. S. A.* 104, 11268–11273.
(6) Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., and Yokoyama, S. (2008) Crystallographic studies on multiple conformational states of active-site loops in pyrrolysyl-tRNA synthetase. *J. Mol. Biol.* 378, 634–652.
(7) Umehara, T., Kim, J., Lee, S., Guo, L. T., Soll, D., and Park, H. S. (2012) N-acetyl lysyl-tRNA synthetases evolved by a CcdB-based selection possess N-acetyl lysine specificity *in vitro* and *in vivo*. *FEBS Lett.* 586, 729–733.

(8) Yoshizawa, S., and Bock, A. (2009) The many levels of control on bacterial selenoprotein synthesis. *Biochim. Biophys. Acta, Gen. Subj.* 1790, 1404–1414.
(9) Krzycki, J. A. (2013) The path of lysine to pyrrolysine. *Curr. Opin. Chem. Biol.* 17, 619–625.
(10) Bulteau, A. L., and Chavatte, L. (2015) Update on selenoprotein biosynthesis. *Antioxid. Redox Signaling* 23, 775–794.
(11) Stock, T., and Rother, M. (2009) Selenoproteins in Archaea and Gram-positive bacteria. *Biochim. Biophys. Acta, Gen. Subj.* 1790, 1520–1532.
(12) Gaston, M. A., Jiang, R., and Krzycki, J. A. (2011) Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Curr. Opin. Microbiol.* 14, 342–349.
(13) James, C. M., Ferguson, T. K., Leykam, J. F., and Krzycki, J. A. (2001) The amber codon in the gene encoding the monomethylamine methyltransferase isolated from *Methanosarcina barkeri* is translated as a sense codon. *J. Biol. Chem.* 276, 34252–34258.
(14) Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* 35, 225–249.
(15) Wan, W., Tharp, J. M., and Liu, W. R. (2014) Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim. Biophys. Acta, Proteins Proteomics* 1844, 1059–1070.
(16) Dumas, A., Lercher, L., Spicer, C. D., and Davis, B. G. (2015) Designing logical codon reassignment - Expanding the chemistry in biology. *Chem. Sci.* 6, 50–69.
(17) Chatterjee, C., and Muir, T. W. (2010) Chemical approaches for studying histone modifications. *J. Biol. Chem.* 285, 11045–11050.
(18) Lang, K., Davis, L., and Chin, J. W. (2015) Genetic encoding of unnatural amino acids for labeling proteins. *Methods Mol. Biol.* 1266, 217–228.
(19) Rovner, A. J., Haimovich, A. D., Katz, S. R., Li, Z., Grome, M. W., Gassaway, B. M., Amiram, M., Patel, J. R., Gallagher, R. R., Rinehart, J., and Isaacs, F. J. (2015) Recoded organisms engineered to depend on synthetic amino acids. *Nature* 518, 89–93.
(20) Minaba, M., and Kato, Y. (2014) High-yield, zero-leakage expression system with a translational switch using site-specific unnatural amino acid incorporation. *Appl. Environ. Microbiol.* 80, 1718–1725.
(21) Kato, Y. (2015) Tunable translational control using site-specific unnatural amino acid incorporation in *Escherichia coli*. *PeerJ* 3, e904.
(22) Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2008) Genetically encoding N^ε-acetyllysine in recombinant proteins. *Nat. Chem. Biol.* 4, 232–234.
(23) Kim, D., Yu, B. J., Kim, J. A., Lee, Y. J., Choi, S. G., Kang, S., and Pan, J. G. (2013) The acetylproteome of Gram-positive model bacterium *Bacillus subtilis*. *Proteomics* 13, 1726–1736.
(24) Weinert, B. T., Iesmantavicius, V., Wagner, S. A., Scholz, C., Gummesson, B., Beli, P., Nystrom, T., and Choudhary, C. (2013) Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. *Mol. Cell* 51, 265–272.
(25) Zhang, K., Zheng, S., Yang, J. S., Chen, Y., and Cheng, Z. (2013) Comprehensive profiling of protein lysine acetylation in *Escherichia coli*. *J. Proteome Res.* 12, 844–851.
(26) Kuhn, M. L., Zemaitaitis, B., Hu, L. I., Sahu, A., Sorensen, D., Minasov, G., Lima, B. P., Scholle, M., Mrksich, M., Anderson, W. F., Gibson, B. W., Schilling, B., and Wolfe, A. J. (2014) Structural, kinetic and proteomic characterization of acetyl phosphate-dependent bacterial protein acetylation. *PLoS One* 9, e94816.
(27) Lima, B. P., Antelmann, H., Gronau, K., Chi, B. K., Becher, D., Brinsmade, S. R., and Wolfe, A. J. (2011) Involvement of protein acetylation in glucose-induced transcription of a stress-responsive promoter. *Mol. Microbiol.* 81, 1190–1204.
(28) AbouElfetouh, A., Kuhn, M. L., Hu, L. I., Scholle, M. D., Sorensen, D. J., Sahu, A. K., Becher, D., Antelmann, H., Mrksich, M., Anderson, W. F., Gibson, B. W., Schilling, B., and Wolfe, A. J. (2015) The *E. coli* sirtuin CobB shows no preference for enzymatic and nonenzymatic lysine acetylation substrate sites. *MicrobiologyOpen* 4, 66–83.

(29) Gattner, M. J., Vrabel, M., and Carell, T. (2013) Synthesis of ε-N-propionyl-, ε-N-butyryl-, and ε-N-crotonyl-lysine containing histone H3 using the pyrrolysine system. *Chem. Commun. 49*, 379−381.

(30) Kuhn, S. M., Rubini, M., Fuhrmann, M., Theobald, I., and Skerra, A. (2010) Engineering of an Orthogonal Aminoacyl-tRNA Synthetase for Efficient Incorporation of the Non-natural Amino Acid O-Methyl-L-tyrosine using Fluorescence-based Bacterial Cell Sorting. *J. Mol. Biol. 404*, 70−87.

(31) Belas, R., Mileham, A., Cohn, D., Hilman, M., Simon, M., and Silverman, M. (1982) Bacterial bioluminescence: isolation and expression of the luciferase genes from *Vibrio harveyi*. *Science 218*, 791−793.

(32) Close, D., Xu, T., Smartt, A., Rogers, A., Crossley, R., Price, S., Ripp, S., and Sayler, G. (2012) The evolution of the bacterial luciferase gene cassette (*lux*) as a real-time bioreporter. *Sensors 12*, 732−752.

(33) Schultz, D. W., and Yarus, M. (1990) A simple and sensitive *in vivo* luciferase assay for tRNA-mediated nonsense suppression. *J. Bacteriol. 172*, 595−602.

(34) 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.

(35) Bentele, K., Saffert, P., Rauscher, R., Ignatova, Z., and Bluthgen, N. (2013) Efficient translation initiation dictates codon usage at gene start. *Mol. Syst. Biol. 9*, 675.

(36) Goodman, D. B., Church, G. M., and Kosuri, S. (2013) Causes and effects of N-terminal codon bias in bacterial genes. *Science 342*, 475−479.

(37) O'Donoghue, P., Prat, L., Heinemann, I. U., Ling, J., Odoi, K., Liu, W. R., and Soll, D. (2012) Near-cognate suppression of amber, opal and quadruplet codons competes with aminoacyl-tRNA[Pyl] for genetic code expansion. *FEBS Lett. 586*, 3931−3937.

(38) Gundllapalli, S., Ambrogelly, A., Umehara, T., Li, D., Polycarpo, C., and Soll, D. (2008) Misacylation of pyrrolysine tRNA *in vitro* and *in vivo*. *FEBS Lett. 582*, 3353−3358.

(39) Ambrogelly, A., Palioura, S., and Soll, D. (2007) Natural expansion of the genetic code. *Nat. Chem. Biol. 3*, 29−35.

(40) Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., and Yokoyama, S. (2008) Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode $N^{ε}$-(o-azidobenzyloxycarbonyl) lysine for site-specific protein modification. *Chem. Biol. 15*, 1187−1197.

(41) Johnson, D. B., Xu, J., Shen, Z., Takimoto, J. K., Schultz, M. D., Schmitz, R. J., Xiang, Z., Ecker, J. R., Briggs, S. P., and Wang, L. (2011) RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat. Chem. Biol. 7*, 779−786.

(42) Johnson, D. B. F., Wang, C., Xu, J. F., Schultz, M. D., Schmitz, R. J., Ecker, J. R., and Wang, L. (2012) Release Factor One Is Nonessential in *Escherichia coli*. *ACS Chem. Biol. 7*, 1337−1344.

(43) Mottagui-Tabar, S., Bjornsson, A., and Isaksson, L. A. (1994) The second to last amino acid in the nascent peptide as a codon context determinant. *EMBO J. 13*, 249−257.

(44) Pott, M., Schmidt, M. J., and Summerer, D. (2014) Evolved sequence contexts for highly efficient amber suppression with noncanonical amino acids. *ACS Chem. Biol. 9*, 2815−2822.

(45) Phillips-Jones, M. K., Watson, F. J., and Martin, R. (1993) The 3′ codon context effect on UAG suppressor tRNA is different in *Escherichia coli* and human cells. *J. Mol. Biol. 233*, 1−6.

(46) Xu, H., Wang, Y., Lu, J., Zhang, B., Zhang, Z., Si, L., Wu, L., Yao, T., Zhang, C., Xiao, S., Zhang, L., Xia, Q., and Zhou, D. (2016) Re-exploration of the Codon Context Effect on Amber Codon-Guided Incorporation of Noncanonical Amino Acids in *Escherichia coli* by the Blue-White Screening Assay. *ChemBioChem 17*, 1250−1256.

(47) Bjornsson, A., Mottagui-Tabar, S., and Isaksson, L. A. (1996) Structure of the C-terminal end of the nascent peptide influences translation termination. *EMBO J. 15*, 1696−1704.

(48) Hayes, C. S., Bose, B., and Sauer, R. T. (2002) Proline residues at the C terminus of nascent chains induce SsrA tagging during translation termination. *J. Biol. Chem. 277*, 33825−33832.

(49) 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.

(50) 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.

(51) Karlin, S., Mrazek, J., and Campbell, A. M. (1998) Codon usages in different gene classes of the *Escherichia coli* genome. *Mol. Microbiol. 29*, 1341−1355.

(52) Miller, J. H. (1992) *A Short Course in Bacterial Genetics: A Laboratory Manual for Escherichia coli and Related Bacteria*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 456 pp.

(53) Epstein, W., and Kim, B. S. (1971) Potassium transport loci in *Escherichia coli* K-12. *J. Bacteriol. 108*, 639−644.

(54) Juers, D. H., Heightman, T. D., Vasella, A., McCarter, J. D., Mackenzie, L., Withers, S. G., and Matthews, B. W. (2001) A structural view of the action of *Escherichia coli* (*lacZ*) β-Galactosidase. *Biochemistry 40*, 14781−14794.

(55) Fried, L., Lassak, J., and Jung, K. (2012) A comprehensive toolbox for the rapid construction of *lacZ* fusion reporters. *J. Microbiol. Methods 91*, 537−543.

(56) Lutz, R., and Bujard, H. (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I_1-I_2 regulatory elements. *Nucleic Acids Res. 25*, 1203−1210.

(57) Hong, S. H., Kwon, Y. C., and Jewett, M. C. (2014) Non-standard amino acid incorporation into proteins using *Escherichia coli* cell-free protein synthesis. *Front. Chem.*, DOI: 10.3389/fchem.2014.00034.

(58) Laber, B., Gomis-Ruth, F. X., Romao, M. J., and Huber, R. (1992) *Escherichia coli* dihydrodipicolinate synthase. Identification of the active site and crystallization. *Biochem. J. 288* (Pt 2), 691−695.

(59) Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J. Bacteriol. 177*, 4121−4130.

(60) Antoine, R., and Locht, C. (1992) Isolation and molecular characterization of a novel broad-host-range plasmid from *Bordetella bronchiseptica* with sequence similarities to plasmids from gram-positive organisms. *Mol. Microbiol. 6*, 1785−1799.

(61) Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., 2nd, and Peterson, K. M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene 166*, 175−176.

(62) Boute, N., Lowe, P., Berger, S., Malissard, M., Robert, A., and Tesar, M. (2016) NanoLuc Luciferase - A Multifunctional Tool for High Throughput Antibody Screening. *Front. Pharmacol. 7*, 27.

(63) Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor, NY.

(64) Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene 77*, 51−59.

(65) Gödeke, J., Heun, M., Bubendorfer, S., Paul, K., and Thormann, K. M. (2011) Roles of two *Shewanella oneidensis* MR-1 extracellular endonucleases. *Appl. Environ. Microbiol. 77*, 5342−5351.

(66) van den Ent, F., and Lowe, J. (2006) RF cloning: a restriction-free method for inserting target genes into plasmids. *J. Biochem. Biophys. Methods 67*, 67−74.

3.1 Supplementary Material

Supplementary material: https://pubs.acs.org/doi/suppl/10.1021/acssynbio.7b00048

Supporting Information

A Versatile Toolbox for the Control of Protein Levels using N^{ε}-acetyl-L-lysine Dependent Amber Suppression

Wolfram Volkwein[1], Christopher Maier[1], Ralph Krafczyk[1], Kirsten Jung[1]* & Jürgen Lassak[1]*

[1]Center for integrated Protein Science Munich (CiPSM) at the Department of Biology I, Microbiology,

Ludwig-Maximilians-Universität München, Großhaderner Strasse 2-4, 82152 Martinsried, Germany.

*To whom correspondence should be addressed:

E-mail: jung@lmu.de

E-mail: juergen.lassak@lmu.de

Supplementary Tables

Supplementary Table S1. Strains used in this study.

Stain	Relevant genotype or description	Reference or source
E. coli DH5αλpir	F⁻ 80*lacZ*ΦM15 (*lacZYA-argF*)U196 *recA1 hsdR17 deoR thi-1 supE44 gyrA96 relA1*/pir	[1]
E. coli BW25113	F⁻ λ⁻ Δ(*araD-araB*)567 Δ*lacZ4787*(::rrnB-3) *rph-1* Δ(*rhaD-rhaB*)568 *hsdR514*	[2]
E. coli MG1655	wild-type; F- lambda- *ilvG rfb50 rph-1*	[3]
Salmonella enterica	Serovar Typhimurium LT2	ATCC 19585
Vibrio cholera C6706	O1 El Tor isolate from Peru	Centers for Disease Control and Prevention (CDC)
E. coli LF1	F⁻ λ⁻ *ilvG rfb50 rph-1 rpsL150* P$_{lac}$::*rpsL-neo-kan::lacZ*$^{\Delta 1-100bp}$, *Kan^r, Strp^s*	[4]
E. coli LF1-LacZ-K	Derivative from *E. coli* LF1, contains an insertion of a lysine codon (AAA) after position 9 of the *lacZ* gene, *Kan^s Str^r*	this study
E. coli LF1-LacZ(Am)	Derivative from *E. coli* LF-1, contains an insertion of an amber codon (TAG) after position 9 of the *lacZ* gene, *Kan^s, Str^r*	this study
E. coli LF1-LacZ-PPK	Derivative from *E. coli* LF1, contains an insertion of a proline-proline-lysine motif (CCG CCG AAA) after position 9 of the *lacZ* gene, *Kan^s, Str^r*	this study
E. coli LF1-LacZ-PP(Am)	Derivative from *E. coli* LF1, contains an insertion of a proline-proline-amber motif (CCG CCG TAG) after position 9 of the *lacZ* gene, *Kan^s, Str^r*	this study
E. coli LF1-AcKRST	Derivative from *E. coli* LF1, contains the acetyl lysyl-tRNA-synthetase (AcKRS), a mutated version of the PylRS from *Methanosarcina mazei* towards N$^\epsilon$-acetyl lysine specificity, according to Umehara *et al*.[5] and the cognate amber suppressor tRNA$_{CUA}$ (*pylT*), *Kan^s, Str^r*	this study
E. coli LF1-AcKRST/*dapA:cam^r*	Derivative from *E. coli* LF1-AcKRST were the *dapA* gene was replaced by a chloramphenicol resistance cassette using λ-RED recombinase, *Cam^r*	this study

Supplementary Table S2. Plasmids used in this study.

Plasmid	Features	References
pBBR1MCS-2	Broad-host-range cloning vector, 5.1 kb, ND[a]; *Kan*[r]	[6]
pBBR1MCS-3	Broad-host-range cloning vector, 5.2 kb, ND[a]; *Tet*[r]	[6]
pBBR1MCS-4	Broad-host-range cloning vector, 5.0 kb, ND[a]; *Amp*[r]	[6]
pBBR1MCS-5	Broad-host-range cloning vector, 4.9 kb, ND[a]; *Gm*[r]	[6]
pBBR1MCS-2 AcKRST	Broad-host-range cloning vector, Contains the acetyl lysyl-tRNA synthetase (*pylS**) under control of P$_{glnS}$ and the suppressor tRNA$_{CUA}$ (*pylT*) under the control of P$_{proK}$, inserted into the *Nsi*I restriction site, 6.8 kb, ND[a]; *Kan*[r]	this study
pBBR1MCS-3 AcKRST	Broad-host-range cloning vector. Contains the acetyl lysyl-tRNA synthetase (*pylS**) under control of P$_{glnS}$ and the suppressor tRNA$_{CUA}$ (*pylT*) under the control of P$_{proK}$, inserted into the *Age*I restriction site, 6.9 kb, ND[a]; *Tet*[r]	this study
pBBR1MCS-4 AcKRST	Broad-host-range cloning vector. Contains the acetyl lysyl-tRNA synthetase (*pylS**) under control of P$_{glnS}$ and the suppressor tRNA$_{CUA}$ (*pylT*) under the control of P$_{proK}$, inserted into the *Nsi*I restriction site, 6.6 kb, ND[a]; *Amp*[r]	this study
pBBR1MCS-5 AcKRST	Broad-host-range cloning vector. Contains the acetyl lysyl-tRNA synthetase (*pylS**) under control of P$_{glnS}$ and the suppressor tRNA$_{CUA}$ (*pylT*) under the control of P$_{proK}$, inserted into the *Nsi*I restriction site, 6.5 kb, ND[a]; *Gm*[r]	this study
pBBR1MCS-5 TT-RBS-lux	PCR template for the amplification of the *lux* operon *luxCDABE* from *Photorhabdus luminescens*, ND[a]; *Gm*[r]	[7]
pBAD/HisA	pBR322-derived expression vector, contains the promotor P$_{BAD}$ of the arabinose operon *araBAD* from *E. coli* and its regulatory gene *araC*, 4.1 kb, *Amp*[r]	Invitrogen
pBAD/HisA-Lux	Contains the *lux* operon inserted into the *Nco*I / *Kpn*I restriction sites, 9.8 kb, *Amp*[r]	this study
pBAD/HisA-Lux(Am)	Contains the *lux* operon with an amber codon (TAG) in *luxC* at position 3, inserted into the *Nco*I / *Bsr*GI restriction sites of pBAD/HisA-Lux, 9.8 kb, *Amp*[r]	this study
pBAD/HisA-Lux(2Am)	Contains the *lux* operon with two amber codons (TAG) in *luxC* at position 3, inserted into the *Nco*I / *Bsr*GI restriction sites of pBAD/HisA-Lux, 9.8 kb, *Amp*[r]	this study
pBAD/HisA-Lux-RA(Am)	Contains the *lux* operon with an arginine alanine amber motif (CGG GCT TAG) in *luxC* at position 3, inserted into the *Nco*I / *Bsr*GI restriction sites of pBAD/HisA-Lux, 9.8 kb, *Amp*[r]	this study
pBAD/HisA-Lux-DA(Am)	Contains the *lux* operon with an aspartic-acid alanine amber motif (GAT GCT TAG) in *luxC* at position 3, inserted into the *Nco*I / *Bsr*GI restriction sites of pBAD/HisA-Lux, 9.8 kb, *Amp*[r]	this study
pBAD/HisA-Lux-HH(Am)	Contains the *lux* operon with an histidine histidine amber motif (CAC CAC TAG) in *luxC* at position 3, inserted into the *Nco*I / *Bsr*GI restriction sites of pBAD/HisA-Lux, 9.8 kb, *Amp*[r]	this study
pBAD/HisA-Lux-KDP(Am)	Contains the *lux* operon with an lysine aspartic-acid proline amber motif (AAA GAT CCG TAG) in *luxC* at position 3, inserted into the *Nco*I / *Bsr*GI restriction sites of pBAD/HisA-Lux, 9.8 kb, *Amp*[r]	this study
pBAD/HisA-Lux-PP(Am)	Contains the *lux* operon with an proline amber motif (CCG TAG) in *luxC* at position 3, inserted into the *Nco*I / *Bsr*GI restriction sites of pBAD/HisA-Lux, 9.8 kb, *Amp*[r]	this study
pBAD/HisA-Lux-KDPP(Am)	Contains the *lux* operon with an lysine aspartic acid proline proline amber motif (AAA GAT CCG CCG TAG) in *luxC* at position 3, inserted into the *Nco*I / *Bsr*GI restriction sites of pBAD/HisA-Lux, 9.8 kb, *Amp*[r]	this study
pNL1.1	PCR template for amplification of NanoLuc® (*Nluc*), 3.1 kb, *Amp*[r]	Promega
pBBR1MCS-2 AcKRST Nluc	Derived from pBBR1MCS-2 AcKRST, contains *Nluc* under control of P$_{BAD}$ inserted into the *Xba*I / *Kpn*I restriction sites, 8.7 kb, *Kan*[r]	this study
pBBR1MCS-2 AcKRST Nluc(Am)	Derived from pBBR1MCS-2 AcKRST, contains *Nluc* under control of P$_{BAD}$ inserted into the *Xba*I / *Kpn*I restriction sites, insertion of an amber codon in *Nluc* at position 3, 8.7 kb, *Kan*[r]	this study
pBBR1MCS-2 AcKRST Nluc PP(Am)	Derived from pBBR1MCS-2 AcKRST, contains *Nluc* under control of P$_{BAD}$ inserted into the *Xba*I / *Kpn*I restriction sites, insertion of an proline proline amber motif in *Nluc* at position 3, 8.7 kb, *Kan*[r]	this study
pBAD/HisA-Kan	pBAD/HisA (Invitrogen) derived expression vector, the ampicillin resistance cassette (*bla*) was replaced with a	this study

	kanamycin resistance cassette from pBBR1MCS-2, 4.2 kb, *Kan^r*	
pBAD/HisA-Kan Lux	Derived from pBAD/HisA (Kan), contains the *lux* operon inserted restriction free, 9.8 kb, *Kan^r*	this study
pBAD/HisA-Kan-Lux(Am)	Derived from pBAD/HisA (Kan), contains the *lux* operon inserted restriction free, insertion of an amber codon (TAG) in *luxC* at position 3, 9.8 kb, *Kan^r*	this study
pBAD/HisA-dapA	Contains *dapA* inserted between the *Xho*I / *Eco*RI restriction sites, 5.0 kb, *Amp^r*	this study
pBAD/HisA-dapA(Am)	Derived from pBAD/HisA-dapA, contains *dapA* with an amber codon on position 3 of the *dapA* open reading frame, inserted between the *Xho*I / *Eco*RI restriction sites, 5.0 kb, *Amp^r*	this study
pACYCDuet™-1	Standard expression vector, ORI P15A, 4 kb, *Cam^r*	Novagen
pACYCDuet-PylRST	Contains the pyrrolysyl-tRNA-synthetase (PylRS) / cognate amber suppressor tRNA$_{CUA}$ (*pylT*) pair from *Methanosarcina mazei*, 5.5 kb, *Cam^r*	[8]
pACYCDuet-AcKRST	Contains the acetyl lysyl-tRNA-synthetase (AcKRS), a mutated version of the PylRS from *Methanosarcina mazei* towards N^ϵ-acetyl lysine specificity, according to Umehara *et al*.[5] and the cognate amber suppressor tRNA$_{CUA}$ (*pylT*), 5.5 kb, *Cam^r*	this study
pRed/ET	λ-RED recombinase in pBAD24; *Amp^r*	Gene Bridges

[a] ND, the incompatibility group of pBBR1 MCS plasmids has not been defined [9]; compatible with IncP, IncQ and IncW group plasmids, as well as ColE1- and P15a-based replicons [6]. *Amp^r*, *Cam^r*, *Kan^r*, *Tet^r*, *Gm^r* and *Str^r* are ampicillin, chloramphenicol, kanamycin, tetracycline, gentamycin, and streptomycin resistance, respectively.

Supplementary Table S3. Primers used in this study.

Primer name	Sequence	Restriction site	Reference
Sequencing and control primers			
AcKRS chk fw	AAT AAG TTC CTC ACA AAG GCA AAC GAA GAC		this study
pBBR1MCS245 NsiI chk rev	CAA GGC GAC AAG GTG CTG ATG		this study
pBBR1MCS3 AgeI chk rev	TGC GAT GAG TGG CAG GGC GGG GC		this study
Seq33 fw-100	GGC GTC ACA CTT TGC TAT GC		[10]
AraC chk fw	CCT GAC CGC GAA TGG TGA GAT TGA GA		this study
Ara Prom chk fw	CTG ACG CTT TTT ATC GCA ACT CTC TAC TG		this study
dapA chk fw	CAT GAA GCT CCG CAA GCG GT		this study
dapA chk rev	CAC CAG ATA ATG TTG CGA TGA CAGT		this study
Primers to generate a mutated version of the PylRS from _Methanosarcina mazei_ towards _Nε_-acetyl lysine specificity (AcKRS). Described primers a directly derived from the library primers described by Umehara _et al._[5]			
AcKRS L301M Y306L L309A OL fw	AGA ACT TCT GCC TGA GAC CCA TGA TGG CTC CAA ACC TTC TGA ACT ACG CGC GCA AGC TTG ACA GGG CCC		this study
AcKRS L301M Y306L L309A OL rev	GGG CCC TGT CAA GCT TGC GCG CGT AGT TCA GAA GGT TTG GAG CCA TCA TGG GTC TCA GGC AGA AGT TCT		this study
AcKRS C348F OL fw	CAT GCT GAA CTT CTT TCA GAT GGG ATC G		this study
AcKRS C348F OL rev	CGA TCC CAT CTG AAA GAA GTT CAG CAT G		this study
MPYSf	TTT CCC TGA ATT CCG GCA AGC		[5]
AcKRS rev	ATC GGC GAG AAA GTC AGC AGG CCG CGC GGC CGC TTA CAG GTT GGT AGA AAT CCC GTT ATA GTA AGA CTC		this study
Primers for Lux reporter constructs			
KpnI LuxE rev	GCT TCG AAT TCC CAT ATG GTA CCT TAT CAA ACG CTT CGG TTA AGC TCA A	_Kpn_I	this study
NcoI LuxC-wt fw	CTA ACA GGA GGA ATT AAC CAT GGG CAC TAA AAA AAT TTC ATT CAT TAT TAA CGG CCA G	_Nco_I	this study
NcoI LuxC:Am fw	CTA ACA GGA GGA ATT AAC CAT GGG CTA GAC TAA AAA AAT TTC ATT CAT TAT TAA CGG CCA G	_Nco_I	this study
NcoI LuxC:2Am fw	CTA ACA GGA GGA ATT AAC CAT GGG CTA GTA GAC TAA AAA AAT TTC ATT CAT TAT TAA CGG CCA G	_Nco_I	this study
NcoI LuxC:RA Am fw	CTA ACA GGA GGA ATT AAC CAT GGG CCG GGC TTA GAC TAA AAA AAT TTC ATT CAT TAT TAA CGG CCA G	_Nco_I	this study
NcoI LuxC:DA Am fw	CTA ACA GGA GGA ATT AAC CAT GGG CGA TGC TTA GAC TAA AAA AAT TTC ATT CAT TAT TAA CGG CCA G	_Nco_I	this study
NcoI LuxC:HH Am fw	CTA ACA GGA GGA ATT AAC CAT GGG CCA CCA CTA GAC TAA AAA AAT TTC ATT CAT TAT TAA CGG CCA G	_Nco_I	this study
NcoI LuxC:KDP Am fw	CTA ACA GGA GGA ATT AAC CAT GGG CAA AGA TCC GTA GAC TAA AAA AAT TTC ATT CAT TAT TAA CGG CCA G	_Nco_I	this study
NcoI LuxC:PP Am fw	CTA ACA GGA GGA ATT AAC CAT GGG CCC GCC GTA GAC TAA AAA AAT TTC ATT CAT TAT TAA CGG CCA G	_Nco_I	this study
NcoI LuxC:KDPP Am fw	CTA ACA GGA GGA ATT AAC CAT GGG CAA AGA TCC GCC GTA GAC TAA AAA AAT TTC ATT CAT TAT TAA CGG CCA G	_Nco_I	this study
LuxC rev	TCA AAA TCT TTT TTG GCA TTC GGT		this study
Primers for _E. coli_ LF1 LacZ XXX strain construction and verification			
lacI 583bp sense	GTC TGC GTC TGG CTG GCT GGC ATA		[4]
lacZ 500bp anti	CGA CTG GGA AAA CCC TGG CGT TAC CCA ACC		[4]
lacZ 220bp anti	AGC TTT CCG GCA CCG CTT CTG		[4]
LacZ-K-OL-fw	TCA CTG GCC AAA GTC GTT TTA C		this study
LacZ-K-OL-rev	GTA AAA CGA CTT TGG CCA GTG A		this study
LacZ-PPK-OL-fw	TCA CTG GCC CCG CCG AAA GTC GTT TTA C		this study
LacZ-PPK-OL-rev	GTA AAA CGA CTT TCG GCG GGG CCA GTG A		this study
LacZ-Am-OL-fw	TCA CTG GCC TAG GTC GTT TTA C		this study
LacZ-Am-OL-rev	GTA AAA CGA CCT AGG CCA GTG A		this study
LacZ-PPAm-OL-fw	TCA CTG GCC CCG CCG TAG GTC GTT TTA C		this study
LacZ-PPAm-OL-rev	GTA AAA CGA CCT ACG GCG GGG CCA GTG A		this study
Primers for _E. coli_ LF1-AcKRST strain construction			
Plac-AcKRST-OL-fw	GGA AAC AGC TAT GGA TAA AAA ACC ACT AAA		this study

	CAC TCT GAT A		
Plac-AcKRST-OL-rev	GGT TTT TTA TCC ATA GCT GTT TCC TGT GTG AAA TTG TTA TCC		this study
Plac-OL-sRBS-lacZ-fw	ACA AAT ACA ATA AAT AAA GTA AGG AGG TAC ATT ATG ACC ATG ATT ACG GAT TCA CTG GCC G		this study
Plac-AcKRST-OL-sRBS-lacZ-rev	AAT GTA CCT CCT TAC TTT ATT TAT TGT ATT TGT GCA AAA AAG CCT GCT CGT TGA GC		this study
Primer for *E. coli* LF1-AcKRST/*dapA*:*cam*[r] strain construction			
FRT dapA fw	CCA GGC GCG ACT TTT GAA CAG AGT AAG CCA TCA AAT CTC CCT AAA CTT TAA ATT AAC CCT CAC TAA AGG GCG		this study
FRT dapA rev	CAT ACC AAA CGT ACC ATT GAG ACA CTT GTT TGC ACA GAG GAT GGC CCA TGT AAT ACG ACT CAC TAT AGG GCT C		this study
Primers for *dapA* constructs			
dapA EcoRI fw	CCG GAA TTC TTA CAG CAA ACC GGC ATG CTT AAG	*Eco*RI	this study
dapA XhoI rev	CCG CTC GAG ATG TTC ACG GGA AGT ATT GTC GCG	*Xho*I	this study
Amber dapA XhoI rev	CCG CTC GAG ATG TTC TAG GGA AGT ATT GTC GCG	*Xho*I	this study
Primers for broad-host-range cloning vectors			
AcKRST NsiI fw	TGC ATG CAT TCA TCA ATC ATC CCC ATA ATC CTTG	*Nsi*I	this study
AcKRST NsiI rev	TGC ATG CAT GCA AAA AAG CCT GCT CGT TGA GCA G	*Nsi*I	this study
AcKRST BspEI fw	GGC ATT CTC CGG ATC ATC AAT CAT CCC CAT AAT CCT TG	*Bsp*EI	this study
AcKRST BspEI rev	GGC ATT CTC CGG AGC AAA AAA GCC TGC TCG TTG AGC AG	*Bsp*EI	this study
Primers for resistance cassette exchange in pBAD/HisA			
pBAD/HisA amp kan ex fw	TAT ATG AGT AAA CTT GGT CTG ACA GTC AGA AGA ACT CGT CAA GAA GGC GA		this study
pBAD/HisA amp kan ex rev	CTT TTT GTT TAT TTT TCT AAA TAC ATA GCT TGC AGT GGG CTT ACA TGG CG		this study
Primers for NanoLuc® reporter constructs			
Ara XbaI fw	TGC TCT AGA TTA TGA CAA CTT GAC GGC TAC	*Xba*I	this study
Ara Nluc OL rev	CAT GGT TAA TTC CTC CTG TTA GC		this study
Ara Nluc OL fw	GCT AAC AGG AGG AAT TAA CCA TGG TCT TCA CAC TC		this study
Ara Nluc am OL fw	GCT AAC AGG AGG AAT TAA CCA TGG GCT AGG TCT TCA CAC TCG AAG AT		this study
Ara Nluc am PP OL fw	GCT AAC AGG AGG AAT TAA CCA TGG CCG CGC CGT AGG TCT TCA CAC TCG AAG AT		this study
Nluc KpnI rev	CGG GGT ACC TTA CGC CAG AAT GCG TTC	*Kpn*I	this study

Supplementary Table S4. Amber suppression in *Salmonella enterica* LT2 (A) and *Vibrio cholerae* El Tor C6706 (B). Cells were grown aerobically overnight at 37 °C in LB and on the next day, luminescence production in response to *N*ε-acetyl-L-lysine (AcK) and L-arabinose (Ara) was measured for the reporter pBBR1MCS-2 AcKRST Nluc, pBBR1MCS-2 AcKRST Nluc^Am and pBBR1MCS-2 AcKRST Nluc^PP(Am), using the Nano-Glo® luciferase assay system.

A

Salmonella enterica LT2	- AcK - Ara	+AcK (1mM)	+ Ara (0.2%)	+AcK (1mM) + Ara (0.2%)
■ nLuc	23,925	38,640	1,935,681	1,559,454
☐ Amber nLuc	19,516	56,584	215,665	692,714
☐ PP Amber nLuc	2,332	8,646	52,880	124,702

B

Vibrio cholerae El Tor C6706	- AcK - Ara	+AcK (1mM)	+ Ara (0.2%)	+AcK (1mM) + Ara (0.2%)
■ nLuc	24,780	41,781	1,760,650	1,834,598
☐ Amber nLuc	686	2,163	256,137	672,902
☐ PP Amber nLuc	33	203	56,303	177,643

Supplementary Figures
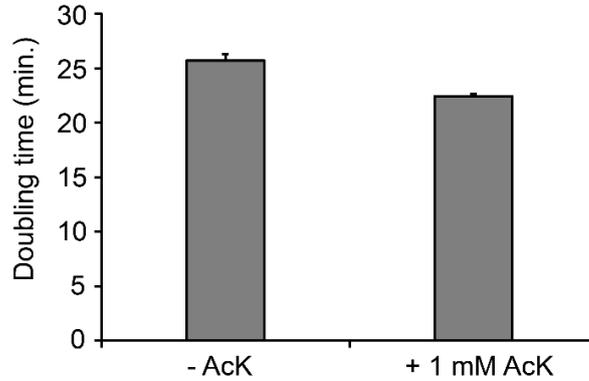
Supplementary Figure S1. Influence of $N\varepsilon$-acetyl-L-lysine (AcK) on the doubling time of *E. coli* BW25113 during exponential growth. Cells were grown in LB in the absence (- AcK) and in the presence of 1 mM AcK (+ 1 mM AcK). Error bars represent the standard deviation of data from three different experiments.
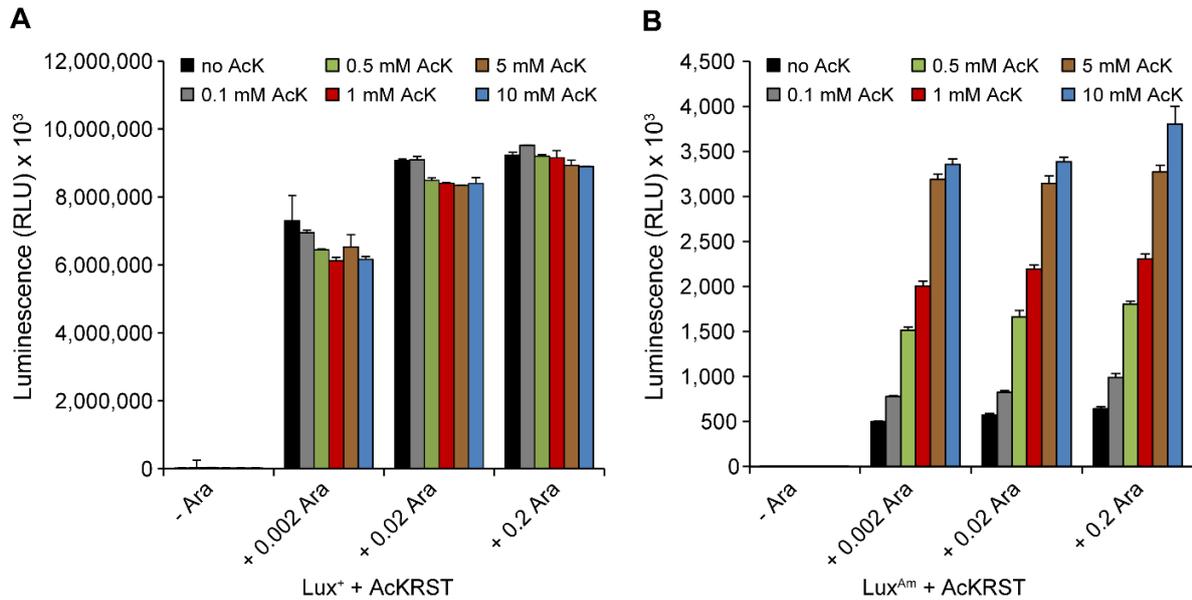
Supplementary Figure S2. Luminescence production of Lux[+] (A) and Lux[Am] (B) in *E. coli* BW25113 in response to L-arabinose (Ara) and of $N\varepsilon$-acetyl-L-lysine (AcK). The maximal luminescence normalized to the OD_{600} from a 16 h time course experiment is depicted in response to Ara and AcK for Lux[+] (A) and Lux[Am] (B). Error bars represent the standard deviation of data from three different experiments.

	+ 0.2% Ara	+ 0.2% Ara + 10 mM AcK	+ 0.2% Ara + 10 mM BocK
Lux^{Am}	158,400	209,028	174,064
Lux^{Am} + AcKRST	445,435	3,426,005	533,867
Lux^{Am} + PylRST	123,738	168,880	2,362,942
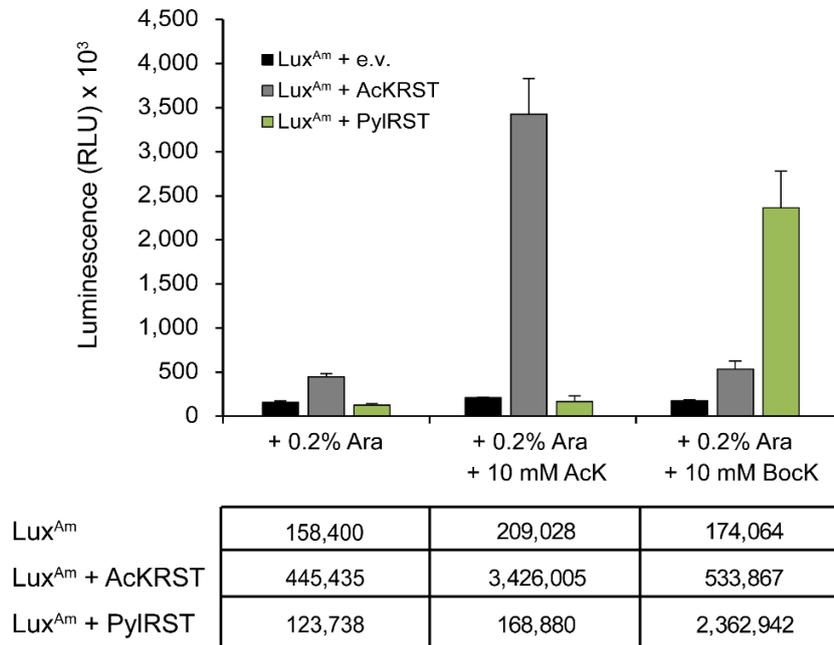
Supplementary Figure S3. Comparison of PylRS and AcKRS based amber suppression. *E. coli* BW25113 cells were transformed with the Lux-amber reporter (Lux^{Am}) in combination with either pACYCDuet-AcKRST (encoding the acetyl lysyl-tRNA synthetase/tRNA$_{CUA}$ pair, AcKRST), pACYCDuet-PylRST (encoding the pyrrolysyl-tRNA synthase/tRNA$_{CUA}$ pair, PylRST) or the empty vector (e.v.) as negative control. These cells were grown aerobically in LB supplemented with 0.2% (w/v) L-arabinose (Ara) at 37°C in a microtiter plate within a Tecan Infinite® F500 system (TECAN). The maximal luminescence production in response to $N\varepsilon$-acetyl-L-lysine (AcK) and $N\varepsilon$-*tert*-butoxycarbonyl-L-lysine (BocK) normalized to the OD$_{600}$ from a 16 h time course experiment is depicted. Error bars represent the standard deviation of data from three different experiments.
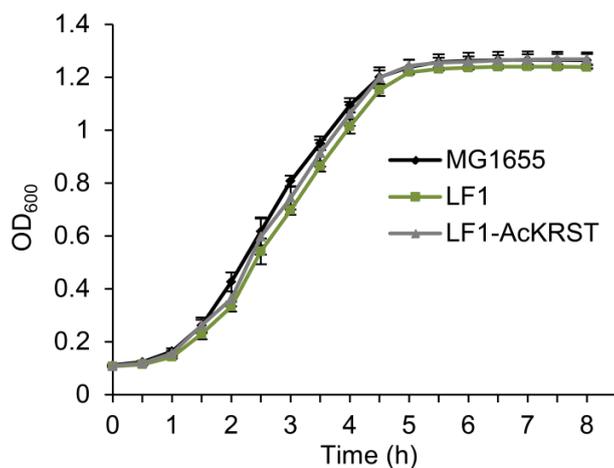
Supplementary Figure S4. Influence of the chromosomal integration of the acetyl lysyl-tRNA synthetase/tRNA$_{CUA}$ pair (AcKRST) on cell growth. *E. coli* MG1655, LF1 and LF1-AcKRST were grown aerobically at 37°C in a microtiter plate within a Tecan Infinite® F500 system (TECAN) in LB. Error bars represent the standard deviation of data from three different experiments.

Supplementary Figure S5. Determination of the minimal diaminopimelic acid (DAP) concentration required for the growth of the *dapA* deficient LF1-AcKRST/*dapA*:*cam*^{r} strain. Depicted are cells grown for 16h at 37 °C aerobically in a microtiter plate in LB, supplemented with varying DAP concentrations ranging from 600 µM to 0 µM.

REFERENCES

 (1) Macinga, D. R., Parojcic, M. M. and Rather, P. N. (1995) Identification and analysis of *aarP*, a transcriptional activator of the 2'-N-acetyltransferase in *Providencia stuartii*. *J. Bacteriol. 177*, 3407-3413.

 (2) Datsenko, K. A. and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA 97*, 6640-6645.

 (3) Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. and Shao, Y. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science 277*, 1453-1462.

 (4) Fried, L., Lassak, J. and Jung, K. (2012) A comprehensive toolbox for the rapid construction of *lacZ* fusion reporters. *J. Microbiol. Methods 91*, 537-543.

 (5) Umehara, T., Kim, J., Lee, S., Guo, L. T., Soll, D. and Park, H. S. (2012) *N*-acetyl lysyl-tRNA synthetases evolved by a CcdB-based selection possess *N*-acetyl lysine specificity in vitro and in vivo. *FEBS Lett. 586*, 729-733.

 (6) Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., 2nd and Peterson, K. M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene 166*, 175-176.

 (7) Gödeke, J., Heun, M., Bubendorfer, S., Paul, K. and Thormann, K. M. (2011) Roles of two *Shewanella oneidensis* MR-1 extracellular endonucleases. *Appl. Environ. Microbiol. 77*, 5342-5351.

 (8) Gattner, M. J., Vrabel, M. and Carell, T. (2013) Synthesis of ε-*N*-propionyl-, ε-*N*-butyryl-, and ε-*N*-crotonyl-lysine containing histone H3 using the pyrrolysine system. *Chem. Commun. 49*, 379-381.

 (9) Antoine, R. and Locht, C. (1992) Isolation and molecular characterization of a novel broad-host-range plasmid from *Bordetella bronchiseptica* with sequence similarities to plasmids from gram-positive organisms. *Mol. Microbiol. 6*, 1785-1799.

 (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.

4 Switching the Post-translational Modification of Elongation Factor EF-P

Reproduced from Frontiers in Microbiology:

Wolfram Volkwein*, Ralph Krafczyk*, Pravin Kumar Ankush Jagtap, Marina Parr, Elena Mankina, Jakub Macošek, Zhenghuan Guo, Maximilian Josef Ludwig Johannes Fürst, Miriam Pfab, Dmitrij Frishman, Janosch Hennig, Kirsten Jung and Jürgen Lassak (2019). Switching the post-translational modification of translation elongation factor EF-P. Front. Microbiol. 10.

*Authors contributed equally

Full-text article: https://doi.org/10.3389/fmicb.2019.01148

Supplementary material:

https://www.frontiersin.org/articles/10.3389/fmicb.2019.01148/full#supplementary-material

Switching the Post-translational Modification of Translation Elongation Factor EF-P

Wolfram Volkwein[1†], Ralph Krafczyk[1†‡], Pravin Kumar Ankush Jagtap[2§], Marina Parr[3,4], Elena Mankina[3], Jakub Macošek[2,5], Zhenghuan Guo[1], Maximilian Josef Ludwig Johannes Fürst[1,6||], Miriam Pfab[1¶], Dmitrij Frishman[3,4], Janosch Hennig[2#], Kirsten Jung[1$] and Jürgen Lassak[1*]

[1] Center for Integrated Protein Science Munich, Department of Biology I, Microbiology, Ludwig-Maximilians-Universität München, Munich, Germany, [2] Structural and Computational Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany, [3] Department of Bioinformatics, Wissenschaftszentrum Weihenstephan, Technische Universität München, Freising, Germany, [4] St. Petersburg State Polytechnic University, Saint Petersburg, Russia, [5] Faculty of Biosciences, Collaboration for Joint PhD Degree Between EMBL and Heidelberg University, Heidelberg, Germany, [6] Molecular Enzymology Group, University of Groningen, Groningen, Netherlands

Tripeptides with two consecutive prolines are the shortest and most frequent sequences causing ribosome stalling. The bacterial translation elongation factor P (EF-P) relieves this arrest, allowing protein biosynthesis to continue. A seven amino acids long loop between beta-strands β3/β4 is crucial for EF-P function and modified at its tip by lysylation of lysine or rhamnosylation of arginine. Phylogenetic analyses unveiled an invariant proline in the -2 position of the modification site in EF-Ps that utilize lysine modifications such as *Escherichia coli*. Bacteria with the arginine modification like *Pseudomonas putida* on the contrary have selected against it. Focusing on the EF-Ps from these two model organisms we demonstrate the importance of the β3/β4 loop composition for functionalization by chemically distinct modifications. Ultimately, we show that only two amino acid changes in *E. coli* EF-P are needed for switching the modification strategy from lysylation to rhamnosylation.

Keywords: IF5A, EarP, EpmA, bacterial two-hybrid, glycosylation, TDP-rhamnose, *Pseudomonas aeruginosa*, NleB

INTRODUCTION

Protein biosynthesis is a universally conserved three-step process that occurs on ribosomes and provides a platform for tRNA mediated amino acid delivery. During translation elongation aminoacyl-tRNAs bind to the ribosomal A-site and peptide bond formation is mediated by a peptidyl-tRNA located in the P-site. Relocation of the P-site tRNA to the E-site enables its exiting from the ribosome. The speed of incorporating amino acids into the growing polypeptide chain varies and strongly depends on their chemical nature (Pavlov et al., 2009). Due to its rigid structure, proline in particular delays the peptidyl transfer reaction, being both a poor A-site donor and P-site acceptor substrate (Muto and Ito, 2008; Wohlgemuth et al., 2008; Pavlov et al., 2009; Johansson et al., 2011; Doerfel et al., 2013, 2015). When translating stretches of two or more prolines, ribosomes become arrested (Tanner et al., 2009; Doerfel et al., 2013; Gutierrez et al., 2013; Hersch et al., 2013; Peil et al., 2013; Ude et al., 2013; Woolstenhulme et al., 2013, 2015;

Elgamal et al., 2014; Starosta et al., 2014a). Thus consecutive prolines are disfavored in evolution (Qi et al., 2018). However, the structural benefits of polyproline sequences in proteins (Adzhubei et al., 2013; Starosta et al., 2014b) seem to outweigh the translational drawback and favored the evolution of a specialized universal elongation factor termed e/aIF5A in eukaryotes/archaea and EF-P in bacteria (Doerfel et al., 2013; Gutierrez et al., 2013; Ude et al., 2013). Upon polyproline mediated stalling e/aIF5A and EF-P are recruited to the ribosome to a location between the P- and E-tRNA binding sites (Blaha et al., 2009; Saini et al., 2009; Melnikov et al., 2016a,b; Schmidt et al., 2016; Huter et al., 2017).

With its three domains (**Figures 1A,B**), EF-P spans both ribosomal subunits and forms an L-shaped, tRNA mimicking structure (Hanawa-Suetsugu et al., 2004; Katz et al., 2014). Whereas the two OB-Folding domains (Oligonucleotide Binding) II and III are likely to be involved in P-site tRNA^{Pro} (Katoh et al., 2016) and E-site codon (Huter et al., 2017) recognition, the EF-P KOW-like N-domain I is crucial for the catalytic activity. Specifically, a seven amino acid long apical loop region between beta-strands three and four (β3Ωβ4) protrudes toward the peptidyl transferase center (Blaha et al., 2009; Huter et al., 2017). A conserved positively charged residue at the loop tip mediates stabilization and positioning of the CCA-end of the P-site tRNA^{Pro} in a way favorable for peptide bond formation (Doerfel et al., 2013, 2015; Lassak et al., 2015). EF-P activity is further enhanced by post-translational extensions of this specific tip residue (Doerfel et al., 2013; Lassak et al., 2015). Interestingly the underlying bacterial modifications appear to be chemically diverse (Lassak et al., 2016; **Figure 1A**). In a subset of bacteria including the Gram-negative model organism *Escherichia coli*, a lysine residue K34 is β-lysylated (Bailly and de Crecy-Lagard, 2010; Navarre et al., 2010; Yanagisawa et al., 2010) with (R)-β-lysine (Behshad et al., 2006) at the ε-amino group, employing the catalytic activity of the EF-P specific ligase EpmA (YjeA, PoxA, GenX) (Roy et al., 2011). Subsequent hydroxylation by EpmC (YfcM) (Peil et al., 2012; Kobayashi et al., 2014) presumably at the (R)-β-lysyl-lysine C5 atom (Huter et al., 2017) completes the modification, but is negligible for function (Bullwinkle et al., 2013). A chemically related amino acid – 5-amino-pentanolyl-lysine – was found on *Bacillus subtilis* EF-P (Rajkovic et al., 2016). By contrast, activity of a distinct EF-P subset encoded in the β-proteobacterial subdivision and certain γ-proteobacteria such as *Pseudomonas putida* and *Shewanella oneidensis* depends on α-rhamnosylation of arginine at the equivalent position (Lassak et al., 2015; Rajkovic et al., 2015; Yanagisawa et al., 2016). This glycosylation is mediated by the GT-B folding glycosyltransferase EarP (Krafczyk et al., 2017; Sengoku et al., 2018) belonging to the enzyme family GT104 according to the CAZy database (Coutinho et al., 2003).

Despite their distinct chemical nature both lysine as well as arginine modifications of EF-P promote proline-proline peptide bond formation at the ribosome. We therefore asked whether there is a specific conservation pattern around the modified residue of diverse EF-Ps and if so how such a specific context contributes to modification efficiency and ribosome rescue. Using bioinformatics and site directed mutagenesis, we were able to show that EarP mediated modification of

E. coli EF-P requires only the substitution of the protruding lysine by arginine. However, this protein derivative remained translationally inactive. Notably, we recognized a selective pressure on the amino acid located at the second position N-terminal of the modification site. While bacteria encoding EF-P with protruding lysine contain an invariant proline, those with arginine instead strictly select against it. Strikingly, the additional substitution of this residue in this context in *E. coli* EF-P led to a variant that even promotes peptide bond formation in polyproline arrested ribosome upon arginine rhamnosylation. We therefore reason that the presence or absence of this specific proline orients β3Ωβ4 in a way that results in translationally active EF-Ps with modifications similar to either (R)-β-lysylation or α-rhamnosylation.

MATERIALS AND METHODS

Plasmid and Strain Construction

All strains, plasmids and oligonucleotides used in this study are listed and described in the **Supplementary Tables S1–S3**. All kits and enzymes were used according to manufacturer's instructions. Plasmid DNA was isolated using the Hi Yield® Plasmid Mini Kit from Süd Laborbedarf. DNA fragments were purified from agarose gels using the Hi Yield® Gel/PCR DNA fragment extraction kit from Süd Laborbedarf. All restriction enzymes, DNA modifying enzymes and the Q5® high fidelity DNA polymerase for PCR amplification were purchased from New England BioLabs.

Escherichia coli strain KV1 for bacterial two-hybrid analysis was constructed as follows: The *luxCDABE* operon from *Photorhabdus luminescens* was amplified from pBAD/HisA-Lux (Volkwein et al., 2017) and integrated into *E. coli* LF1 as essentially described previously by Fried et al. (2012). To keep the ability of blue/white screening, a synthetic ribosomal binding site predicted by RBS calculator (Salis et al., 2009; Espah Borujeni et al., 2014) was introduced upstream of the *lacZ* start site. Afterward *cyaA* was deleted using Red®/ET® recombination technology and the kanamycin cassette was removed using the 709-FLPe/amp expression vector in accordance to the Quick and Easy *E. coli* Gene Deletion Kit (Gene Bridges, Germany). In the same way, *epmA*$_{Eco}$ was deleted in the *E. coli* Δ*efp* reporter strain MG-CR-efp-KanS, resulting in the Δ*efp*$_{Eco}$/Δ*epmA*$_{Eco}$ reporter strain MG-CR-efp-epmA-KanR. The Δ*efp*$_{Eco}$ reporter strain MG-CR-efp-KanS itself was generated by removing the kanamycin resistance cassette from MG-CR-efp (Lassak et al., 2015) using also the Quick and Easy *E. coli* Gene Deletion Kit of Gene Bridges according to the manufacturer's instructions.

Growth Conditions

Escherichia coli cells were routinely grown in Miller modified Lysogeny Broth (LB) (Bertani, 1951; Miller, 1972; Bertani, 2004) at 37°C aerobically under agitation, if not indicated otherwise. When required, media were solidified by using 1.5% (w/v) agar. The medium was supplemented with antibiotics at the following

concentrations when indicated: 100 μg/ml ampicillin sodium salt, 50 μg/ml kanamycin sulfate, 30 μg/ml chloramphenicol, or 15 μg/ml tetracycline hydrochloride. Plasmids carrying the P$_{BAD}$ promoter (Guzman et al., 1995) were induced with L-arabinose at a final concentration of 0.2% (w/v).

SDS-PAGE and Western Blotting

For protein analyses cells were subjected to 12% (w/v) sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) as described by Laemmli (1970). To visualize and confirm protein separation, 2,2,2-trichloroethanol was incorporated into the polyacrylamide gels (Ladner et al., 2004) and detected within a Gel Doc^{TM} EZ gel documentation system (Bio-Rad). Afterward the proteins were transferred onto nitrocellulose membranes (Whatman) which were then subjected to immunoblotting. In a first step the membranes were incubated either with 0.1 μg/mL Anti-6×His® antibody (Abcam, Inc.) to detect EF-P, or with 0.25 μg/ml Anti-Arg^{Rha} antibody (Li et al., 2016; Krafczyk et al., 2017) to visualize rhamnosylation. These primary antibodies (rabbit) were then targeted with 0.2 μg/ml Anti-rabbit alkaline phosphatase-conjugated secondary antibody (Rockland). Localization was visualized by adding development solution [50 mM sodium carbonate buffer, pH 9.5, 0.01% (w/v) p-nitro blue tetrazolium chloride (NBT) and 0.045% (w/v) 5-bromo-4-chloro-3-indolyl-phosphate (BCIP)].

β-Galactosidase Activity Assay

Escherichia coli Δ*efp* (MG-CR-efp-KanS) or Δ*efp*/Δ*epmA* (MG-CR-efp-epmA-KanR) reporter strain cells, in which *lacZ* expression is controlled by the *cadBA* promoter, were grown overnight (o/n) in 100 mM sodium-phosphate buffered Miller modified LB (pH 5.8) under microaerobic conditions and with agitation at 37°C. On the next day, cells were harvested by centrifugation, and the β-galactosidase activities were determined as described (Tetsch et al., 2008) and are given in relative Miller units (MU) (Miller, 1992).

Whenever the plasmid based reporter system pBBR1MCS-3 XPPX *lacZ* (Peil et al., 2013) was used, cells were grown o/n in 100 mM sodium-phosphate buffered Miller modified LB (pH 5.8), microaerobically under agitation at 37°C. Whenever the *E. coli* Δ*epmA* reporter strain (MG-CL-12-yjeA) (Ude, 2013) was used, cells were grown in potassium buffered KE minimal medium (Epstein and Kim, 1971) pH 5.8, supplemented with 10 mM lysine, 0.2% glycerol and antibiotics in the appropriate concentrations. Whenever *efp*$_{Ppu}$ and *earP*$_{Ppu}$ were co-expressed from pBBR1MCS2 (Kovach et al., 1995) and pBAD33, respectively, cells were grown in 100 mM sodium-phosphate buffered Miller modified LB (pH 5.8), microaerobically under agitation at 30°C. Whenever *efp*$_{Ppu}$ and *earP*$_{Ppu}$ were co-overexpressed from pBAD24 and pBAD33, respectively, cells were grown in 100 mM sodium-phosphate buffered Miller modified LB (pH 5.8) and 20 mM arabinose, aerobically under agitation at 30°C. In all cases, the cells were harvested by centrifugation on the next day, and the β-galactosidase activities were determined

as described (Tetsch et al., 2008) and are given in MU (Miller, 1992).

NMR Experiments

To obtain labeled proteins for NMR studies, bacterial overproductions were performed in M9 glucose minimal medium (Miller, 1972) containing either ^{15}N-labeled ammonium chloride alone (pET-SUMO-*efp*$_{Eco}$K34R, pET-SUMO-*efp*$_{Eco}$ P32S, pET-SUMO-*efp*$_{Eco}$ P32S K34R, pET-SUMO-*efp*$_{Ppu}$), or ^{15}N-labeled ammonium chloride in combination with ^{13}C labeled glucose (pET-SUMO-*efp*$_{Eco}$, pET-SUMO-*efp*$_{Eco}$ loop$_{Ppu}$). Overproduction of these N-terminally His$_6$-SUMO tagged hybrid EF-P variants was induced in *E. coli* BL21 (DE3) by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; Sigma Aldrich) during exponential growth. Until the induction point the cells were grown at 37°C, after IPTG induction the temperature was shifted to 18°C and the cells were grown o/n. On the next day, the cells were harvested by centrifugation. The resulting pellet was resuspended on ice in dialysis buffer 1 (100 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, pH 6.5, 1 mM DTT). Cells were lysed using a continuous-flow cabinet from Constant Systems Ltd., at 1.35 kbar, in combination with sonication. The resulting lysate was centrifuged for 40 min at 4°C at 39,810 × *g*. The His$_6$-SUMO tagged proteins were purified in a first step with nickel-nitrilotriacetic acid (Ni-NTA; Qiagen) according to the manufacturer's instructions, using 20 mM imidazole for washing and 250 mM imidazole for elution. Subsequently, imidazole was removed by dialysis o/n at 4°C in dialysis buffer 1. Afterward, the His$_6$-SUMO tag was cleaved off using His$_6$-Ulp1 protease (Starosta et al., 2014b), followed by a second Ni-NTA purification step to remove the His$_6$-SUMO tag itself as well as the His$_6$ tagged Ulp1 protease. As a final step, the purified protein was dialyzed again o/n at 4°C in dialysis buffer 1.

C-terminally His$_6$-tagged EarP$_{Ppu}$ for NMR interaction studies was overproduced in *E. coli* LMG194 cells harboring a pBAD33-*earP*$_{Ppu}$ plasmid in Miller modified LB at 37°C. During exponential growth, 0.2% (w/v) L-arabinose was added. After induction, the temperature was shifted to 18°C, and the cells were grown o/n. On the next day, the cells were harvested by centrifugation. The resulting pellet was resuspended on ice in dialysis buffer 2 (100 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, pH 7.5, 50 mM NaCl, 5 mM DTT). Cell lysis, centrifugation of the lysate and the first Ni-NTA purification step was performed as described above. In a final step, the purified protein was dialyzed o/n in dialysis buffer 2 to remove imidazole from the purification step.

All ^{15}N NMR relaxation experiments for EF-P and its variants were performed in 100 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, pH 6.5 and 1 mM DTT. NMR data were recorded at 298 K for ~ 0.15–0.18 mM of EF-P$_{Eco}$ and its variants except for EF-P$_{Eco}$ P32S for which the data were recorded at 0.09 mM due to low yields of expression. Pulse experiments were performed on an 800 MHz Bruker Avance III NMR spectrometer equipped with a TXI cryogenic probehead. Amide ^{15}N relaxation data of R_1, R_2, and steady-state heteronuclear {^{1}H}–^{15}N-NOE experiments were performed as described before (Farrow et al., 1994;

Korzhnev et al., 2002). T_1 data were measured with 11 different relaxation delays: 20, 50, 100, 150, 250, 400, 500, 650, 800, 1000, and 1300 ms, whereby 150 ms was used as duplicate. T_2 data were determined by using eight different relaxation delays: 16, 32, 48, 64, 80. 96, 112, and 128 ms using 16 ms as duplicate. Duplicate time points were used for error estimation. The correlation time (τ_c) of the protein molecule was estimated using the ratio of averaged T_2/T_1 values (Farrow et al., 1994). Steady-state heteronuclear {^{1}H}–^{15}N-NOE experiments were recorded with and without 3 s of ^{1}H saturation. All relaxation experiments were acquired as pseudo-3D experiments. The spectra were processed with NMRPipe (Delaglio et al., 1995) and peak integration and relaxation parameter calculation was performed using PINT (Niklasson et al., 2017).

For the titration of EF-P$_{Eco}$ and its variants with 2× EarP$_{Ppu}$, both the proteins were dialyzed against 100 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, pH 7.5, 50 mM NaCl, and 1 mM DTT. Experiments were recorded on an 800 MHz Bruker NMR spectrometer equipped with a TXI cryogenic probehead at 298 K. Protein backbone assignments for EF-P$_{Eco}$ and EF-P$_{Eco}$ loop$_{Ppu}$ were obtained from HNCACB, CBCA(CO)NH, and HNCA experiments (Sattler et al., 1999). Data analysis was performed in CcpNmr Analysis software (Vranken et al., 2005). Resonance assignments of EF-P variants have been deposited at the BMRB with the following accession codes: 27811.

In vitro Rhamnosylation Studies

To obtain EF-P variants for *in vitro* rhamnosylation studies, protein overproductions were performed in *E. coli* LMG194 cells, grown in Miller modified LB, harboring the following C-terminally His$_6$-tagged EF-P constructs (see also **Supplementary Table S2**):

- EF-P$_{Ppu}$: pBAD24-*efp*$_{Ppu}$, pBAD24-*efp*$_{Ppu}$ S30P, pBAD24-*efp*$_{Ppu}$ G31A, pBAD24-*efp*$_{Ppu}$ R32K, pBAD24-*efp*$_{Ppu}$ N33G, pBAD24-*efp*$_{Ppu}$ S34Q
- EF-P$_{Eco}$: pBAD24-*efp*$_{Eco}$ K34R, pBAD24-*efp*$_{Eco}$ P32S K34R, pBAD24-*efp*$_{Eco}$ K34R G35N, pBAD24-*efp*$_{Eco}$ K34R Q36S, pBAD24-*efp*$_{Eco}$ P32S K34R G35N, pBAD24-*efp*$_{Eco}$ P32S K34R Q36S, pBAD24-*efp*$_{Eco}$ K34R G35N Q36S, pBAD24-EF-P$_{Eco}$ loop$_{Ppu}$, pBAD24-EF-P$_{Eco}$ domainI$_{Ppu}$

Furthermore, C-terminally His$_6$ tagged EarP$_{Ppu}$ was overproduced in *E. coli* LMG194 harboring pBAD33-*earP*$_{Ppu}$.

To overproduce proteins, cells harboring corresponding plasmids were grown at 37°C, and during exponential growth, 0.2% (w/v) L-arabinose was added to induce protein production. After induction, the temperature was shifted to 18°C, and the cells were grown o/n. On the next day, the cells were harvested by centrifugation and the resulting pellet was resuspended in buffer 3 (100 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, pH 7.5, 50 mM NaCl). Cells were then lysed by sonication and the resulting cell lysate was clarified by centrifugation for 40 min at 4°C at 39,810 × *g*. The His$_6$ tagged proteins were then purified using Ni-NTA beads (Qiagen) according to the manufacturer's instructions, whereby 20 mM imidazole was used for washing and 250 mM imidazole for elution of the His$_6$ tagged proteins. Subsequently, imidazole

was removed by dialysis o/n at 4°C in buffer 3, followed by a second dialysis step at the next morning for 5 h, again in buffer 3. The resulting proteins were then used for *in vitro* rhamnosylation assays.

Kinetic parameters were determined by varying TDP-β-L-rhamnose (TDP-Rha) concentrations while keeping concentrations of EarP$_{Ppu}$ (0.1 μM) and unmodified EF-P$_{Ppu}$ (2.5 μM) constant. A mixture of EarP$_{Ppu}$ and unmodified EF-P$_{Ppu}$ was equilibrated to 30°C in 100 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, pH 7.6. The reaction was started by the addition of TDP-Rha and was stopped after 20 s of incubation at 30°C by the addition of one volume twofold Laemmli buffer (Laemmli, 1970) and incubation at 95°C for 5 min. Samples were subjected to SDS-PAGE and rhamnosylated EF-P$_{Ppu}$ was detected using an Anti-Arg^{Rha} antibody (Krafczyk et al., 2017). A secondary FITC coupled Anti-rabbit antibody (Abcam, United Kingdom) was used to visualize rhamnosylation in a LI-COR Odyssey CLx. Band intensities were quantified using ImageJ (Schneider et al., 2012). K_m values were determined by fitting relative band intensities to the Michaelis–Menten equation using SigmaPlot. The K_m of 5 μM TDP-Rha was determined using commercially available substrate (Carbosynth, United Kingdom). Previously, we determined a K_m of 50 μM using biochemically synthesized TDP-Rha (Krafczyk et al., 2017). After rigorous assessment of this discrepancy we found that contaminations with ammonium acetate were responsible for a miscalculation of the TDP-Rha concentration in stock solutions.

In vitro rhamnosylation of EF-P$_{Eco}$ and EF-P$_{Ppu}$ variants was conducted in 100 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, pH 7.5 containing 50 mM NaCl. A master mix containing 25 μM of the corresponding EF-P variant and 100 μM TDP-Rha was prepared in a reaction tube and divided into 10 μl aliquots. The reaction was started by addition of 10 μl of 0.5 μM EarP solution and stopped after distinct time intervals by addition of 20 μl twofold Laemmli buffer and immediate heating to 95°C in an Eppendorf ThermoMixer for 2 min. All samples were diluted by a factor of 10 in onefold Laemmli buffer and 20 μl (corresponding to 0.5 μg of EF-P) were subjected to SDS-PAGE and Western blotting. Rhamnosylated EF-P was detected and visualized using a polyclonal rabbit Anti-Arg^{Rha} and a fluorescence labeled Anti-rabbit antibody, respectively. Band intensities were determined using ImageJ (Schneider et al., 2012). Relative rhamnosylation rates were calculated by plotting the normalized linear range (intensity t$_x$/intensity$_{max}$) of the time course and determining the slope of the resulting graphs.

Isoelectric Focusing

To investigate lysylation of *E. coli* EF-P C-terminally His$_6$-tagged EF-P$_{Eco}$ was overproduced in *E. coli* BW25113 and *E. coli* BW25113 Δ*epmA* cells, harboring the pBAD33-*efp*-His$_6$ plasmid, and were grown in Miller modified LB at 37°C. Furthermore, *E. coli* BW25113 was transformed with pBAD33-*efp*-His$_6$-*epmAB* to produce post-translationally modified EF-P. During exponential growth, 0.2% (w/v) L-arabinose was added to induce protein production and cells were grown o/n at 18°C. On the next day, cells were harvested by centrifugation. The resulting

pellet was resuspended on ice in HEPES buffer (50 mM HEPES, 100 mM NaCl, 50 mM KCl, 10 mM MgCl$_2$, 5% (w/v) glycerol, pH 7.0). Cells were lysed using a continuous-flow cabinet from Constant Systems Ltd., at 1.35 kb. The resulting lysates were clarified by centrifugation for 1.5 h at 4°C at 234,998 × g. The His$_6$-tagged proteins were purified with Ni-NTA according to the manufacturer's instructions, using 20 mM imidazole for washing and 400 mM imidazole for elution. In a final step, the purified protein was dialyzed o/n in HEPES buffer to remove imidazole from the purification step.

For isoelectric focusing 0.5 µg of protein per lane was loaded on a native vertical isoelectric focusing gel with a pH gradient range of 4–7 (SERVAGel^{TM}) containing approximately 3% (v/v) SERVALYT^{TM}. Prior to loading, samples were mixed with twofold IEF sample buffer according to the manufacturer's instructions and wells were rinsed with SERVA IEF Cathode buffer. Focusing was conducted for 1 h at 50 V, 1 h at 300 V and finally bands were sharpened for 30 min at 500 V. Western blotting was conducted as described above using 0.1 µg/ml Anti-EF-P$_{Eco}$ (Eurogentec).

Bacterial Two-Hybrid Analysis

Protein-protein interactions were detected using the bacterial adenylate cyclase two-hybrid system kit (Euromedex) according to the manufacturer's instructions. This system is based on functional reconstitution of split *Bordetella pertussis* adenylate cyclase CyaA, which in turn activates the *lac* promoter P$_{lac}$ being dependent on the cAMP receptor protein CAP (**Supplementary Figure S1A**). The *E. coli* KV1 strain used in this study was generated by start to stop deletion of the *cyaA* gene from *E. coli* LF1 (Fried et al., 2012) and subsequent incorporation of the *lux* operon at the *lac* locus by using described methods. Applicability of this strain was tested by assessing the self-interaction of the GCN4 leucine zipper in *E. coli* KV1 and the commercially available bacterial two-hybrid strains *E. coli* BTH101 (Euromedex) and *E. coli* DHM1 (Euromedex) on X-Gal containing screening plates. For this purpose, the reporter strains were co-transformed with the plasmids pKT25-zip (Euromedex) and pUT18C-zip (Euromedex) that encode for protein hybrids of the leucine zipper and the corresponding CyaA fragment. While all reporter strains respond with comparable β-galactosidase mediated color formation, KV1 exhibits an additional light output (**Supplementary Figure S1B**). Liquid cultures of transformants containing 50 µg/ml kanamycin sulfate, 100 µg/ml carbenicillin and 0.5 mM IPTG were inoculated from single colonies and incubated at 30°C for 8 h. 2 µl of liquid culture were spotted on LB plates containing 40 µg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and 0.5 mM IPTG as well as 50 µg/ml kanamycin sulfate and 100 µg/ml carbenicillin. Pictures were taken after 32 h of incubation at 30°C.

For measuring interaction strength between EarP$_{Ppu}$ and EF-P$_{Ppu}$ or EF-P$_{Eco}$, chemically competent *E. coli* KV1 cells were co-transformed with pKT25-EarP$_{Ppu}$ and either pUT18C-EF-P$_{Ppu}$ or pUT18C-EF-P$_{Eco}$. Transformants carrying leucine-zipper-reporter hybrids (pUT18-zip and pKT25-zip) were used as positive controls, whereas transformants containing pUT18C

and pKT25 vector backbones served as negative controls. Single colonies were picked and used to inoculate 96-well plates, with each well containing 200 µl of LB medium supplemented with 0.5 mM IPTG as well as 50 µg/ml kanamycin sulfate and 100 µg/ml carbenicillin. Plates were incubated at 30°C and under moderate agitation (550 rpm in Eppendorf ThermoMixer) for 16 h. On the next morning, Costar 96 Well White plates containing 200 µl of LB medium (0.5 mM IPTG, 50 µg/ml kanamycin sulfate, 100 µg/ml carbenicillin) were inoculated with 2 µl of o/n culture and luminescence output was monitored in 10-min intervals for 40 h in a Tecan Spark with 240 rpm at 30°C.

Bioinformatic Analyses
Fold Recognition and Comparison of EF-P Structures

The EF-P$_{Ppu}$ fold recognition model was generated using the online user interface of the I-TASSER server (Zhang, 2008; Roy et al., 2010; Yang et al., 2015). Chain B of the *Pseudomonas aeruginosa* EF-P crystal structure (3OYY:B) (Choi and Choe, 2011) was used as "template without alignment." The resulting model exhibits a C-score (confidence score for estimating the quality of fold recognition models predicted by I-TASSER. Range: (−5; 2); high values signify a model with a high confidence and vice-versa) of 1.27 and an estimated TM-score [measure for the similarity between the predicted model and the native structures. Range: (0,1), with 1 indicating a perfect match] of 0.89 ± 0.07, indicating high confidence and correct topology. The structural alignment of the N-terminal KOW-like EF-P N-domains of the EF-P$_{Ppu}$ fold recognition model and *E. coli* EF-P (3A5Z:H) (Katz et al., 2014) was generated using the UCSF Chimera (Pettersen et al., 2004) MatchMaker function (Chain pairing: Best-aligning pair of chains; Alignment algorithm: Needleman–Wunsch; Matrix: Blosum-62; Gap extension penalty: 1; Matching to 2.0 angstroms) and resulted in an RMSD of 1.005 angstroms between 51 pruned atom pairs. Only amino acids 1–63 of each EF-P structure were used for the alignment.

Sequence Data and Domain Analysis

Using the search query "Bacteria"[Organism] AND ["reference genome"(refseq category) OR "representative genome"(refseq category)] AND "complete genome"[filter] we obtained from the RefSeq database (O'Leary et al., 2016) a collection of 1644 proteomes corresponding to complete and/or representative genomes. HMMER searches (Finn et al., 2015) against the locally installed Pfam database (Finn et al., 2016) was used to identify domains in proteins using the *hmmscan e*-value cut-off of 0.001.

We created an initial dataset of EF-P proteins by considering only those gene products that exclusively contain the three domains of interest – "EFP_N" (KOW-like domain), "EFP" (OB domain), "Elong-fact-P_C" (C-terminal). Subsequently we excluded from the dataset the proteins annotated in the UniProt database (UniProt Consortium, 2018) as YeiP, which are EF-P's paralogs and possess the same domain architecture. Additional sequence comparisons did not yield any misannotated YeiP proteins.

Sequences of predicted KOW-like domains, in which the β3Ωβ4 loop is located, were aligned using the "E-INS-i" algorithm from the MAFFT software suite

(Katoh and Standley, 2013). According to the trimAl tool (Capella-Gutierrez et al., 2009) the MSA of KOW-like domain sequences does not contain any sequences particularly prone to introduce poorly aligned regions. Eleven sequences were manually deleted from the set as they introduce gaps in the multiple alignment of the β3Ωβ4 region (alignment positions 31–37) and the alignment was re-computed. MSA file is available in the **Supplementary Materials**.

The final EF-P set contains 1166 sequences from 1137 genomes, including 29 genomes containing two EF-P paralogs. Sequence logos were built using ggseqlogo R package (Wagih, 2017).

The EF-P-containing genomes were scanned for the EpmA, EpmC, EarP, DHS, and YmfI proteins. We identified 358, 143, 100, and 128 EpmA, EpmC, EarP, and DHS proteins based, which are single-domain proteins containing the "tRNA-synt_2," "EpmC," "EarP" (Lassak et al., 2015) and "DS" (Brochier et al., 2004) domains, respectively. Orthologs of the YmfI protein (*Uniprot ID: O31767*) from *Bacillus subtilis* (Hummels et al., 2017; Rajkovic and Ibba, 2017) were obtained from the orthologous group 508579 of the OMA database (Altenhoff et al., 2018).

Phylogenetic Analysis

IQ-Tree 1.6.10 (Nguyen et al., 2015) was used to infer a phylogenetic tree of KOW-like domains by the maximum likelihood method, with the LG substitution matrix and the number of standard non-parametric bootstrap replicates set to 100. The tree file in PDF format and its visualization including bootstrap support values are available as **Supplementary Dataset S2**. Using the *ete3* python package (Huerta-Cepas et al., 2016) the tree was rooted to the midpoint outgroup and converted to ultrametric. The evolutionary reconstruction of ancestral states was performed using the *ace* function from the *phytools* R package (Revell, 2012), which implements the maximum likelihood estimation. We used the *ggtree* R package (Yu et al., 2017) to visualize the evolutionary reconstruction of ancestral states on the tree of KOW-like domains and annotate it with the amino acid located at the 34th position, the presence or absence of a certain modification enzyme, and the taxonomy for Proteobacteria, Firmicutes and Actinobacteria. A tree with a more detailed taxonomic annotation is available in **Supplementary Materials**.

RESULTS AND DISCUSSION

K34R Substitution of *E. coli* EF-P Is Sufficient for Non-cognate Rhamnosylation by EarP

Canonically, N-linked protein glycosylation occurs at a consensus sequence motif (Helenius and Aebi, 2004). By contrast, the glycosyltransferase EarP seems to recognize rather the overall shape of domain I of its target EF-P (Krafczyk et al., 2017; Sengoku et al., 2018; **Figure 1B**). Notably, and despite large sequence diversity all EF-Ps are structurally similar (Yanagisawa

et al., 2010; Choi and Choe, 2011; Lassak et al., 2016). In this regard the EF-Ps of the two model organisms *E. coli* (EF-P$_{Eco}$) and *P. putida* (EF-P$_{Ppu}$) are highly superimposable (**Figure 1B**) although they share a sequence identity below 30%. We were therefore curious whether cross-interaction between the non-cognate partners EF-P$_{Eco}$ and EarP$_{Ppu}$ is possible. To this end we constructed a highly sensitive bacterial two-hybrid *E. coli* reporter strain KV1 to combine it with the plasmid system which was reported previously (Karimova et al., 1998) (**Supplementary Figure S1A**). Distinct from the original strains we used bioluminescence as a readout (see section "Materials and Methods") and generated C-terminal fusions of the two complementary fragments T25 and T18 of the *Bordetella pertussis* adenylate cyclase with EarP$_{Ppu}$ and the EF-Ps of *E. coli* and *P. putida* (T25-EarP$_{Ppu}$, T18-EF-P$_{Ppu}$, T18-EF-P$_{Eco}$). Interaction strength of protein pairs was assessed by determining the maximal light output in a 40 h time course experiment (**Supplementary Figure S1B**). Cells co-expressing the cognate interaction partners EarP$_{Ppu}$ and EF-P$_{Ppu}$ emitted a maximum of 7,000 RLU (**Figure 1C**). When EarP$_{Ppu}$ and the non-cognate EF-P$_{Eco}$ were co-produced, a maximal light emission of 255 RLU was observed. Although this is substantially lower, the measured value is significantly above background levels (maximal RLU of <100) and thus clearly demonstrates cross-interaction of EarP$_{Ppu}$ and EF-P$_{Eco}$.

Knowing that EF-P$_{Eco}$ and EarP$_{Ppu}$ do cross-interact, we next assessed whether cross-rhamnosylation also occurs. Therefore we took advantage of a previously introduced rhamnosylarginine specific antibody (Li et al., 2016; Krafczyk et al., 2017) and used it to test *in vitro* rhamnosylation over time. The rhamnosylation efficiency of EarP$_{Ppu}$ with its cognate partner EF-P$_{Ppu}$ was determined with 50 μM TDP-β-L-rhamnose (TDP-Rha) (=10-fold K_m; **Supplementary Figure S2A**) and set to 100%. Next we tested an EF-P$_{Eco}$ variant in which solely the modification site K34 was changed to arginine (K34R$_{Eco}$). However, we did not observe any rhamnosylation *in vitro* (**Figure 1D** and **Supplementary Figure S2B**) presumably as a result of suboptimal contacts between EF-P$_{Eco}$ and EarP$_{Ppu}$ (Sengoku et al., 2018; **Figure 1C**). As important interaction sites between EF-P and its corresponding modification system are predominantly located within the first 65 amino acids (Navarre et al., 2010; Yanagisawa et al., 2010; Krafczyk et al., 2017; Sengoku et al., 2018) we now swapped the N-domain of EF-P$_{Eco}$ with the one from EF-P$_{Ppu}$ (EF-P$_{Eco}$ domainI$_{Ppu}$). In line with our expectations EF-P$_{Eco}$ domainI$_{Ppu}$ was readily modified (**Figure 1D**). For this reason, in a subsequent step we tested an EF-P$_{Eco}$ variant with swapped β3Ωβ4 – EF-P$_{Eco}$ loop$_{Ppu}$, containing in total four amino acid substitutions P32S, K34R, G35N, and Q36S. Although the efficiency was strongly reduced (1% compared to EF-P$_{Ppu}$) this variant could still be rhamnosylated (**Figure 1D**).

We previously observed that impairments of EarP variants with largely reduced *in vitro* rhamnosylation rates could be compensated *in vivo* (EF-P modification and functionality) when EF-P and the variants were co-overproduced (Krafczyk et al., 2017; **Supplementary Figure S3**). This could be explained by an increase of the local protein concentrations within the cells.
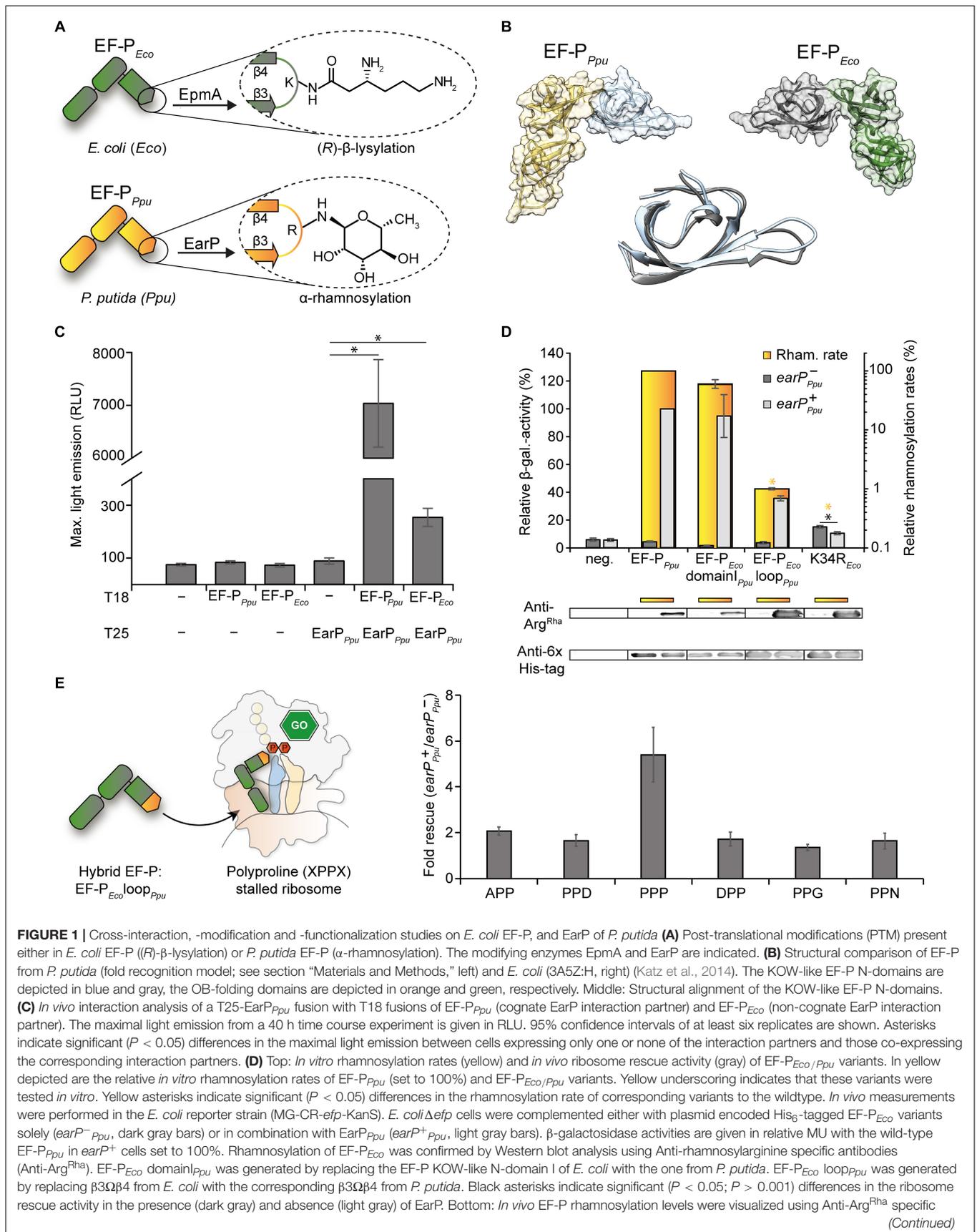
FIGURE 1 | Cross-interaction, -modification and -functionalization studies on *E. coli* EF-P, and EarP of *P. putida* **(A)** Post-translational modifications (PTM) present either in *E. coli* EF-P ((R)-β-lysylation) or *P. putida* EF-P (α-rhamnosylation). The modifying enzymes EpmA and EarP are indicated. **(B)** Structural comparison of EF-P from *P. putida* (fold recognition model; see section "Materials and Methods," left) and *E. coli* (3A5Z:H, right) (Katz et al., 2014). The KOW-like EF-P N-domains are depicted in blue and gray, the OB-folding domains are depicted in orange and green, respectively. Middle: Structural alignment of the KOW-like EF-P N-domains. **(C)** *In vivo* interaction analysis of a T25-EarP$_{Ppu}$ fusion with T18 fusions of EF-P$_{Ppu}$ (cognate EarP interaction partner) and EF-P$_{Eco}$ (non-cognate EarP interaction partner). The maximal light emission from a 40 h time course experiment is given in RLU. 95% confidence intervals of at least six replicates are shown. Asterisks indicate significant ($P < 0.05$) differences in the maximal light emission between cells expressing only one or none of the interaction partners and those co-expressing the corresponding interaction partners. **(D)** Top: *In vitro* rhamnosylation rates (yellow) and *in vivo* ribosome rescue activity (gray) of EF-P$_{Eco/Ppu}$ variants. In yellow depicted are the relative *in vitro* rhamnosylation rates of EF-P$_{Ppu}$ (set to 100%) and EF-P$_{Eco/Ppu}$ variants. Yellow underscoring indicates that these variants were tested *in vitro*. Yellow asterisks indicate significant ($P < 0.05$) differences in the rhamnosylation rate of corresponding variants to the wildtype. *In vivo* measurements were performed in the *E. coli* reporter strain (MG-CR-*efp*-KanS). *E. coli* Δ*efp* cells were complemented either with plasmid encoded His$_6$-tagged EF-P$_{Eco}$ variants solely (*earP*$^-_{Ppu}$, dark gray bars) or in combination with EarP$_{Ppu}$ (*earP*$^+_{Ppu}$, light gray bars). β-galactosidase activities are given in relative MU with the wild-type EF-P$_{Ppu}$ in *earP*$^+$ cells set to 100%. Rhamnosylation of EF-P$_{Eco}$ was confirmed by Western blot analysis using Anti-rhamnosylarginine specific antibodies (Anti-Arg^{Rha}). EF-P$_{Eco}$ domainI$_{Ppu}$ was generated by replacing the EF-P KOW-like N-domain I of *E. coli* with the one from *P. putida*. EF-P$_{Eco}$ loop$_{Ppu}$ was generated by replacing β3Ωβ4 from *E. coli* with the corresponding β3Ωβ4 from *P. putida*. Black asterisks indicate significant ($P < 0.05$; $P > 0.001$) differences in the ribosome rescue activity in the presence (dark gray) and absence (light gray) of EarP. Bottom: *In vivo* EF-P rhamnosylation levels were visualized using Anti-Arg^{Rha} specific

(Continued)

Consequently, we reinvestigated rhamnosylation of K34R$_{Eco}$, EF-P$_{Eco}$ loop$_{Ppu}$, and EF-P$_{Eco}$ domainI$_{Ppu}$ by EarP$_{Ppu}$ in *E. coli*. Following this approach, even the single substituted EF-P$_{Eco}$ variant reached modification levels comparable to wildtype EF-P$_{Ppu}$ (**Figure 1D**, bottom). Thus, we were ultimately able to test these EF-P$_{Eco}$ variants on their ability to rescue ribosome stalling in an EarP dependent manner. EF-P functionality was measured using a previously established β-galactosidase dependent reporter system (Ude et al., 2013; **Figure 1D**, top). The assay is based on the effective translation of the polyproline motif containing acid stress responsive transcriptional regulator CadC (Buchner et al., 2015; Schlundt et al., 2017) and activation of its cognate promoter P$_{cadBA}$ fused to *lacZ* (Ude et al., 2013): β-galactosidase activity is low in *E. coli* cells lacking *efp* but becomes elevated when complementing with both a copy of *earP*$_{Ppu}$ and *efp*$_{Ppu}$ provided in *trans* (**Figure 1D**). Similarly, EF-P$_{Eco}$ domainI$_{Ppu}$ rescues ribosome stalling upon rhamnosylation, indicating that binding of the two distinct EF-Ps from *E. coli* and *P. putida* to the ribosome occurs presumably at the same position. Hence the structural determinants for proper orientation of the respective protruding residue (lysine or arginine) and accordingly the corresponding modification may predominantly lay in the β3Ωβ4 composition. In line with this assumption rhamnosylated EF-P$_{Eco}$ loop$_{Ppu}$ also alleviates the translational arrest occurring at the CadC nascent chain. We note that this rescue activity is not restricted to three consecutive prolines, but encompasses also other diprolyl arrest motifs as demonstrated for APP, DPP, PPD, PPG, and PPN (**Figure 1E**). In contrast to EF-P$_{Eco}$ loop$_{Ppu}$ we measured an unexpected increase in relative β-galactosidase level to about 20% of wildtype activity after introducing K34R$_{Eco}$ into Δ*efp*$_{Eco}$ cells pointing toward a partial complementation of the mutant phenotype. However, this activity was lowered in the concomitant presence of *earP*$_{Ppu}$, which suggests an inhibitory effect of the modification. Presumably, the otherwise preserved EF-P$_{Eco}$ loop composition in K34R$_{Eco}$ precludes proper alignment of rhamnosylarginine with the CCA-end of the P-site tRNA^{Pro}.
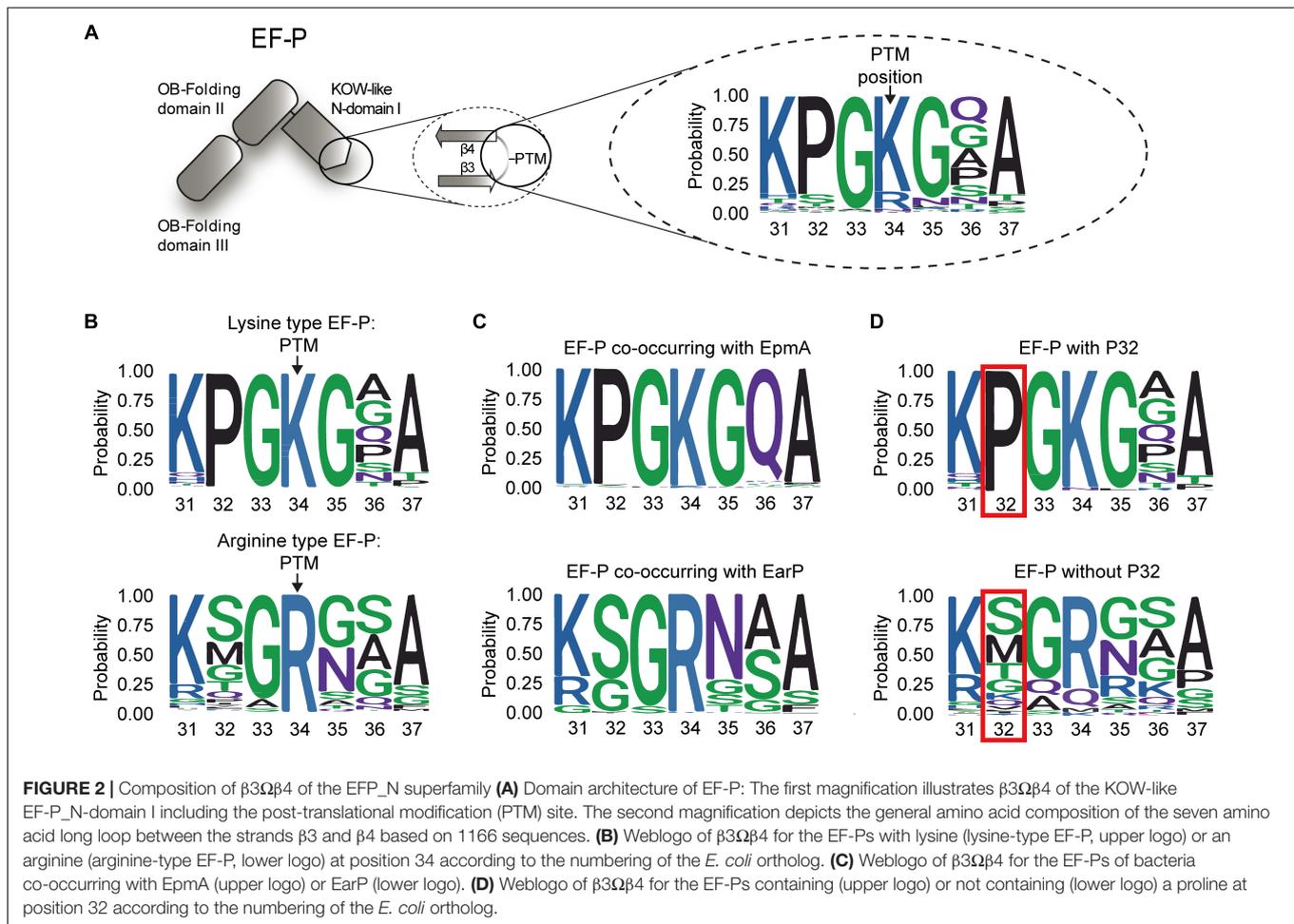
Collectively, these data demonstrate that EarP-mediated rhamnosylation can tolerate substantial changes in the primary sequence of the target protein EF-P. The capability of EF-P to alleviate polyproline dependent translational stalling is, however, strongly affected by changes in β3Ωβ4 sequence. This in turn suggests that both of the modifications rely on a certain acceptor loop architecture that orients the protruding residues in a way favorable for promoting ribosome rescue.

The Sequence Composition of the EF-P β3Ωβ4 Determines Ribosome Rescue With Distinct Modifications

Having shown that the EF-P$_{Eco}$ β3Ωβ4 composition is crucial for rhamnosylation dependent rescue of polyproline arrested ribosomes, we next examined the role of the specific loop amino acids on protein function. Therefore, we initially constructed a phylogenetic tree based on 1166 EF-P sequences. To define modification-specific protein subsets, EpmA and EarP orthologs were collected as described previously (Lassak et al., 2015; **Supplementary Dataset S1**). The EF-P modification system present in *B. subtilis* was excluded in this study, as the full pathway is still poorly understood (Rajkovic et al., 2016; Witzky et al., 2018). A first sequence logo of β3Ωβ4 numbered according to the *E. coli* protein (amino acids 31 to 37) was generated based on the complete EF-P dataset (**Figure 2A** and **Supplementary Dataset S1**). In line with earlier reports (Bailly and de Crecy-Lagard, 2010), the vast majority (81.11%) of EF-Ps have a lysine at the β3Ωβ4 tip (K34), whereas arginine is the second most frequent amino acid occurring in 14.75% of the proteins (**Supplementary Dataset S1**). The remaining 4.11% contain A (0.51%), M (0.77%), N (2.23%), and Q (0.6%) in this position. We next extracted two subsets of proteins with either a protruding lysine (lysine-type) or arginine (arginine-type) (**Figure 2B**). This analysis revealed a highly conserved proline in the second position N-terminal of the modification site (P32) in the lysine type subset being almost absent in the arginine-type EF-Ps. Consistently, bacteria with EpmA pathway have this proline in the EF-P sequence whereas those with EarP do not (**Figure 2C**). With few exceptions the two modification systems thus appear to be mutually exclusive (Lassak et al., 2015; **Supplementary Dataset S1**). Based on these observations, EF-P sequences were grouped according to the presence or absence of P32 (**Figure 2D**). Besides lysine (98.73%), we also found that alanine (100%) and asparagine in the protruding position strongly co-occur with proline (100%), whereas other types of amino acids co-occur with P32 extremely rarely or not at all: arginine (2.33%), methionine (0%), and glutamine (0%) (**Supplementary Dataset S1**).

Subsequent to and based on our bioinformatic analysis we mutated β3Ωβ4 of the EF-Ps of both *E. coli* and *P. putida*. EF-P functionality was measured *in vivo*, again using the CadC dependent β-galactosidase reporter system (Ude et al., 2013; **Figure 3**). The partial P$_{cadBA}$ activation with plasmid-based K34R$_{Eco}$ (**Figure 1D**) intrigued us to first investigate the effect of overproduced unmodified wildtype EF-P$_{Eco}$. Therefore, we ectopically expressed *efp*$_{Eco}$ in a reporter strain lacking the *E. coli* lysyl ligase EpmA and measured the β-galactosidase outcome (**Supplementary Figure S4**). Intriguingly, P$_{cadBA}$ was 50% reactivated compared to a *trans* complementation with *epmA*$_{Eco}$. Presumably, the lysine K34 side chain forms important stabilizing contacts with the CCA-end of the P-site tRNA^{Pro} (Huter et al., 2017), which can in part compensate for a lack of modification. We were therefore next curious whether unmodified amino acids other than lysine and arginine can promote EF-P$_{Eco}$ functionality without modification. Hence, we substituted K34 by any other amino acid (A, M, N, Q) to

FIGURE 2 | Composition of β3Ωβ4 of the EFP_N superfamily **(A)** Domain architecture of EF-P: The first magnification illustrates β3Ωβ4 of the KOW-like EF-P_N-domain I including the post-translational modification (PTM) site. The second magnification depicts the general amino acid composition of the seven amino acid long loop between the strands β3 and β4 based on 1166 sequences. **(B)** Weblogo of β3Ωβ4 for the EF-Ps with lysine (lysine-type EF-P, upper logo) or an arginine (arginine-type EF-P, lower logo) at position 34 according to the numbering of the *E. coli* ortholog. **(C)** Weblogo of β3Ωβ4 for the EF-Ps of bacteria co-occurring with EpmA (upper logo) or EarP (lower logo). **(D)** Weblogo of β3Ωβ4 for the EF-Ps containing (upper logo) or not containing (lower logo) a proline at position 32 according to the numbering of the *E. coli* ortholog.

be found in the protruding position of β3Ωβ4 of the various EF-Ps (**Figure 3A** and **Supplementary Dataset S1**). However, none of the resultant protein variants was translationally active, indicating on the one hand that side chain similarities only between arginine and lysine seems to be high enough to preserve certain of the above-mentioned interactions. On the other hand, the significantly lower activity with K34R$_{Eco}$ (20%) compared to unmodified EF-P$_{Eco}$ (50%) points toward a non-stimulating or even negative effect, possibly caused by the guanidino group.

Having demonstrated that substitution of K34 in *E. coli* EF-P is hardly tolerated, we went on to analyze the impact of its context residues. Coherent with its high degree of conservation in lysine-type EF-Ps (**Figures 2B,D**) an exchange of P32 (P32S$_{Eco}$, P32G$_{Eco}$) analogous to the arginine-type EF-P sequence logo (**Figure 2C**) abolishes β-galactosidase activity (**Figure 3A**). Similarly, a substitution of G33 (G33A$_{Eco}$, G33S$_{Eco}$) is not tolerated and leads to a loss of function of EF-P$_{Eco}$. In comparison, when mutating G35 (G35N$_{Eco}$) and Q36 (Q36S$_{Eco}$) a residual rescue activity of 3.8 and 21.1%, respectively, is retained. Altogether our analysis of EF-P$_{Eco}$ β3Ωβ4 unveils important determinants for protein function and thus explains their high degree of conservation.

In our complementary analysis with the EF-P of *P. putida* KT2440 we generated the substitution variants K29R$_{Ppu}$, S30P$_{Ppu}$, R32K$_{Ppu}$, N33G$_{Ppu}$, and S34Q$_{Ppu}$ according to amino acids predominantly found in the lysine-type sequence logo (**Figure 2B**). We also constructed K29A$_{Ppu}$, S30A$_{Ppu}$, S30G$_{Ppu}$, G31A$_{Ppu}$, G31S$_{Ppu}$, N33D$_{Ppu}$, S34A$_{Ppu}$, and A35S$_{Ppu}$ to further study the impact of the corresponding positions on EF-P activity and rhamnosylation efficiency. An *in vitro* time course analysis was performed (**Figure 3B**) with wild-type EF-P$_{Ppu}$ as well as its variants S30P$_{Ppu}$, G31A$_{Ppu}$, R32K$_{Ppu}$, N33G$_{Ppu}$, and S34Q$_{Ppu}$. This revealed relative rhamnosylation rates with S30P$_{Ppu}$ and G31A$_{Ppu}$ (<1% of wild-type activity) being slowest, while N33G$_{Ppu}$ and S34Q$_{Ppu}$ reach 62 and 12% compared to wild-type EF-P$_{Ppu}$, respectively (**Figure 3B**). Corresponding to the *E. coli* EF-P variants we also assessed the capability of EF-P$_{Ppu}$ in alleviating the translational arrest on consecutive prolines *in vivo*. As for the cross-modified K34R$_{Eco}$ (**Figure 1D**), we found that overproduction of EarP compensates for reduced rhamnosylation efficiency and accordingly all EF-P$_{Ppu}$ substitutions – except changes of R32 – were fully modified *in vivo* (**Figure 3B**). Regardless, S30P$_{Ppu}$ reaches only 4% of wild-type β-galactosidase activity and hence remains almost inactive even upon rhamnosylation (**Figure 3B**). This result is notably

FIGURE 3 | Activity of *E. coli* and *P. putida* EF-P variants with single amino acid substitutions in β3ΩβΩβ4 **(A)** Top: *In vivo* activity measurements of *E. coli* EF-P β3Ωβ4 variants using the described *E. coli* reporter strain. In this strain, either *efp* (*E. coli* MG-CR-efp-KanS, $epmA^+_{Eco}$, dark gray bars) or both, *efp* as well as *epmA* were deleted (*E. coli* MG-CR-efp-epmA-KanR, $epmA^-_{Eco}$, light gray bars) and complemented with a plasmid encoded *E. coli* His₆-tagged EF-P variant. β-galactosidase activities are given in relative MU with the wild-type EF-P$_{Eco}$ in the $epmA^+$ background set to 100%. Asterisks indicate significant ($P < 0.05$) differences in the ribosome rescue activity of a given EF-P$_{Eco}$ variant in the presence of EpmA compared to the same variant in the absence of EpmA. Bottom: Production of EF-P$_{Eco}$ variants was confirmed by Western blot analysis using Anti-6×His® **(B)** Top: *In vitro* rhamnosylation rates (yellow) and ribosome rescue activity (gray) of EF-P$_{Eco/Ppu}$ variants. Shown in yellow are the relative *in vitro* rhamnosylation rates of EF-P$_{Ppu}$ wild type (set to 100%) and EF-P$_{Ppu}$ amino acid substitution variants. Yellow underscoring depicts that these variants were tested *in vitro*. Yellow asterisks indicate significant ($P < 0.05$) differences in the rhamnosylation rate of corresponding variants to the wildtype. *In vivo* activity measurements of *P. putida* EF-P β3Ωβ4 using the *E. coli* reporter strain (*E. coli* MG-CR-efp-KanS). *E. coli* Δ*efp* cells were complemented either with a plasmid-encoded *P. putida* His₆-tagged EF-P variant solely (*earP⁻$_{Ppu}$*, dark gray bars) or in combination with EarP$_{Ppu}$ (*earP⁺$_{Ppu}$*, light gray bars). β-galactosidase activities are given in relative MU with the wild-type EF-P$_{Ppu}$ in *earP⁺* cells set to 100%. Black asterisks indicate significant ($P < 0.05$; $P > 0.0001$) differences in the ribosome rescue activity of a given EF-P$_{Ppu}$ variant in the presence of EarP compared to the same variant in the absence of EarP. Bottom: Rhamnosylation levels of EF-P$_{Ppu}$ variants were detected by Western blot analysis using Anti-Arg^{Rha}.

reminiscent of what we saw with the corresponding *E. coli* EF-P converse exchange P32S. Contrary to S30P$_{Ppu}$, the alanine and glycine substitutions S30A$_{Ppu}$ and S30G$_{Ppu}$ reached 39 and 96% of wild-type β-galactosidase activity, respectively. These data support our observation of a strong selection against proline in the arginine-type EF-Ps, but at the same time allowing for a certain degree of freedom in the -2 position of the modification site. Substitutions of EF-P$_{Ppu}$ in positions N33, S34, and A35 as well as K29 are also tolerated without significant activity

loss (**Figure 3B**). Similar to G33 in EF-P$_{Eco}$ (Yanagisawa et al., 2010; **Figure 3A**), the position equivalent G31 in EF-P$_{Ppu}$ is crucial for both modification efficiency and protein function (**Figure 3B**), which might be explained by sterically-hindering interactions with either the ribosome or EarP caused by longer side chains. Interestingly and in contrast to K34R$_{Eco}$, R32K$_{Ppu}$ is not only inactive but the β-galactosidase activity measured with this variant is even below the level of an Δ*efp$_{Eco}$* deletion strain. This drastic phenotype indicates an inhibitory effect on

FIGURE 4 | *In vivo* and *in vitro* characterization of EF-P hybrids **(A)** Top: *In vivo* rhamnosylation and functionality analysis of EF-P hybrids co-expressed with EarP from *P. putida* or *S. oneidensis*. Measurements were performed in the *E. coli* reporter strain (MG-CR-efp-KanS). *E. coli* Δefp cells were complemented either with plasmid encoded His$_6$-tagged EF-P$_{Eco/Ppu}$ variants solely (earP$^-$, dark gray bars) or in combination with either EarP from *P. putida* (earP$^+_{Ppu}$, light gray bars) or EarP from *S. oneidensis* (earP$^+_{So}$, blue bars). β-galactosidase activities are given in relative MU with the wild-type EF-P$_{Ppu}$ in earP$^+_{Ppu}$ cells set to 100%. Asterisks indicate significant ($P < 0.05$) differences in the ribosome rescue activity of a given EF-P$_{Eco}$ variant in the presence of EarP$_{Ppu}$ or EarP$_{So}$ compared to the same variant in the absence of EarP. Bottom: *In vivo* rhamnosylation levels were detected and visualized using Anti-Arg^{Rha}. Production of EF-P$_{Ppu}$ and EF-P$_{Eco}$ variants was detected and visualized using Anti-6×His®. **(B)** NMR-single titration experiment: Average change in the intensity ratio of EF-P$_{Eco}$ and its variants upon titration with twofold EarP$_{Ppu}$. Error bars represent standard deviation of the intensity ratios over all signals within each titration.
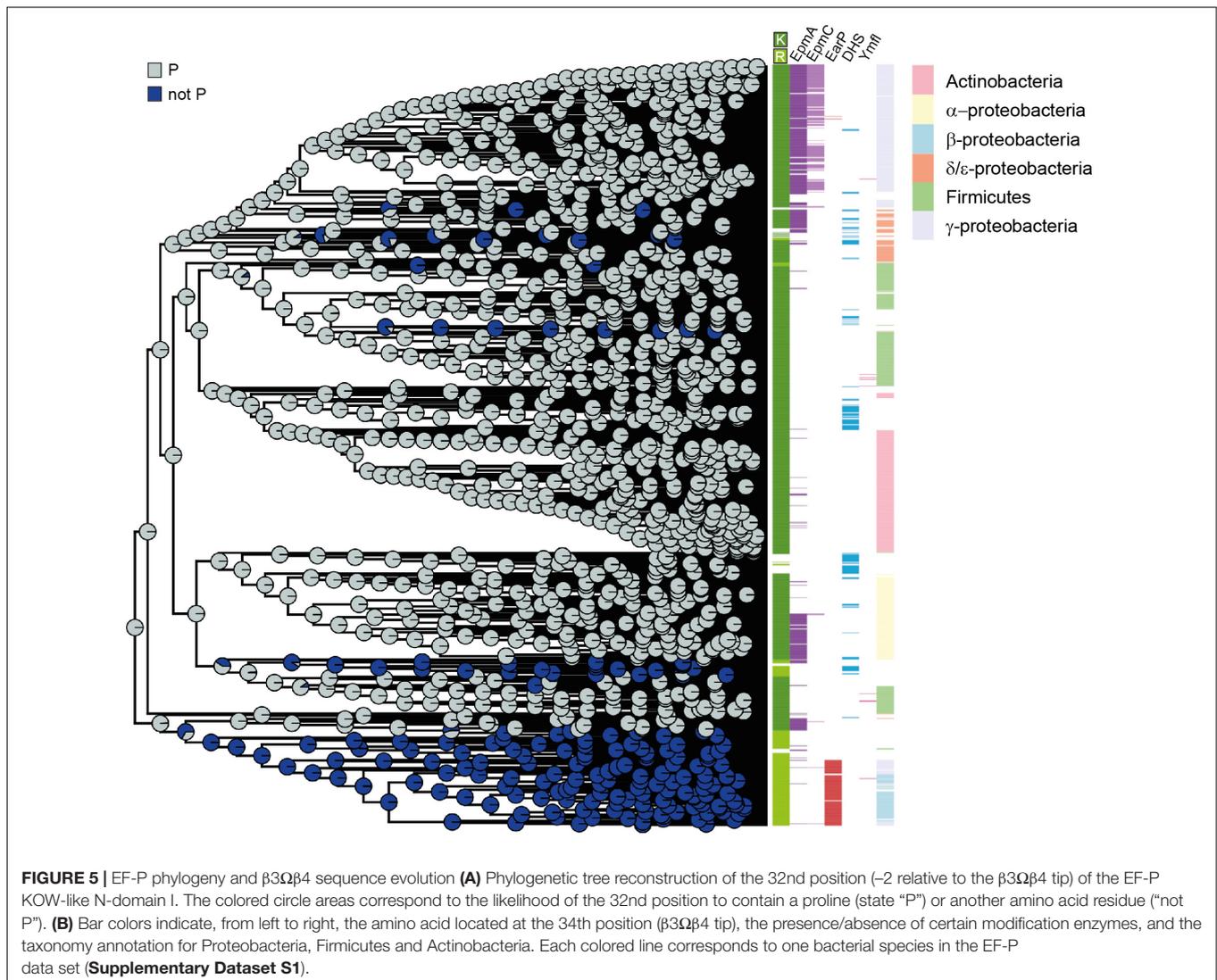
polyproline translation. Notably, we saw the same when testing unmodified EF-P of both *P. putida* (**Figure 1D**) or *S. oneidensis* (Lassak et al., 2015). A similar phenomenon was also observed by others when analyzing the growth of *P. aeruginosa* harboring the EF-P R32K variant (Rajkovic et al., 2015). The most plausible explanation is a distinct orientation of the protruding residue that depends on the β3Ωβ4 composition.

The Essential Proline P32 in *E. coli* EF-P Prevents Activation With Rhamnosylarginine

Our mutational analysis clearly shows that the presence of P32 is crucial for the activity of EF-P$_{Eco}$ on the one hand and prevents ribosome rescue by EF-P$_{Ppu}$ on the other hand. Accordingly, we were curious whether a double substitution P32S/K34R$_{Eco}$ is sufficient to become translationally active upon modification by EarP (**Figure 4A**). Unfortunately, this EF-P variant was unable to promote CadC translation as seen by the low β-galactosidase activities in the Δefp$_{Eco}$ P$_{cadBA}$::*lacZ* reporter strain. However, at the same time we noticed a reduced *in vivo* rhamnosylation efficiency, which might mask a potential rescue effect. Therefore, we tested whether the EarP ortholog from *S. oneidensis* MR–1 might be a more efficient modifier. Indeed, upon co-expression of *earP$_{So}$*, rhamnosylation of P32S/K34R$_{Eco}$ reached significantly higher levels, which concomitantly resulted in a partial ribosome rescue (**Figure 4A**). To understand the effect of substituting *E. coli* EF-P P32 especially on rhamnosylation efficiency by EarP$_{Ppu}$ we determined its interactions with wildtype EF-P$_{Eco}$ and variants at a molecular level by performing NMR titration experiments (**Figure 4B** and

Supplementary Figures S5A,B). EF-P$_{Eco}$ interacts with EarP$_{Ppu}$ as shown by the decrease in total amount of peak intensities in the EF-P$_{Eco}$ ^{15}N-HSQC spectrum upon EarP$_{Ppu}$ titration. Physical interaction leads to an increased molecular tumbling time, which in turn decreases transverse relaxation times and peak intensities. The interaction was substantially enhanced in the K34R$_{Eco}$ variant, resulting in even lower peak intensities. This was expected, as R34 makes important contacts with EarP and its cognate EF-P in *Neisseria meningitidis* (Sengoku et al., 2018). In contrast to K34R$_{Eco}$, we observed reduced interaction strength in the P32S variant as peak intensities were stronger than for EF-P$_{Eco}$ wild type. This result might be counterintuitive, however, only if one ignores that EF-P must not only be efficiently rhamnosylated by EarP, but at the same time has to interact optimally with the P-site tRNA on the ribosome. In this light, the substitution of proline might be regarded as an evolutionary consequence to maintain functionality at the expense of rhamnosylation efficiency. In line with the findings for K34R$_{Eco}$ and P32S$_{Eco}$, the K34R/P32S$_{Eco}$ double substitution variant showed intermediate interaction with EarP$_{Ppu}$ compared to K34R$_{Eco}$ and increased further with the EF-P$_{Eco}$ loop$_{Ppu}$ variant. In addition to K34R$_{Eco}$ and P32S$_{Eco}$, the EF-P$_{Eco}$ loop$_{Ppu}$ construct bears two additional substitutions at positions 35 and 36, which seem to be important for EF-P/EarP interaction. Thus, we can interpret our finding as an adjustment to compensate for the negative interaction effect that we saw with P32S$_{Eco}$.

It is possible that substitution of proline P32 causes substantial changes in the loop dynamics due to its rigid nature. To test this, we performed ^{15}N R_1, R_2, and steady-state heteronuclear {^{1}H}–^{15}N-NOE relaxation experiments on EF-P$_{Eco}$ and its variants and

FIGURE 5 | EF-P phylogeny and β3Ωβ4 sequence evolution **(A)** Phylogenetic tree reconstruction of the 32nd position (–2 relative to the β3Ωβ4 tip) of the EF-P KOW-like N-domain I. The colored circle areas correspond to the likelihood of the 32nd position to contain a proline (state "P") or another amino acid residue ("not P"). **(B)** Bar colors indicate, from left to right, the amino acid located at the 34th position (β3Ωβ4 tip), the presence/absence of certain modification enzymes, and the taxonomy annotation for Proteobacteria, Firmicutes and Actinobacteria. Each colored line corresponds to one bacterial species in the EF-P data set (**Supplementary Dataset S1**).

compared it with EF-P$_{Ppu}$. Our analysis suggests that substitution of single EF-P$_{Eco}$ β3Ωβ4-loop residues with residues from EF-P$_{Ppu}$ or even with the complete β3Ωβ4 does not significantly alter the NMR relaxation properties of β3Ωβ4 and hence its dynamics (**Supplementary Figures S5C–F**). Thus, differences observed in the interaction of EF-P$_{Eco}$ and its variants with EarP$_{Ppu}$ can be attributed to the molecular nature of resulting interactions rather than changes in the loop dynamics.

Our observation that substitution of P32 weakens the interaction strength between EF-P$_{Eco}$ and EarP$_{Ppu}$ explains the differences in cross-complementing the Δefp P$_{cadBA}$::*lacZ* mutant phenotype with K34R/P32S$_{Eco}$ in combination with the rhamnosyltransferase ortholog either from *S. oneidensis* or *P. putida* (**Figure 4A**). It is also indicative that further sequence determinants in the β3Ωβ4 are contributing to EF-P recognition by EarP and accordingly being in line with our *in vitro* rhamnosylation studies on EF-P$_{Ppu}$ substitution variants (**Figure 3B**). Consequently, we additionally substituted G35 and Q36 for asparagine and serine, respectively. Both resultant

EF-P$_{Eco}$ variants P32S/K34R/G35N$_{Eco}$ and P32S/K34R/Q36S$_{Eco}$ alleviated CadC translation when co-producing EarP$_{Ppu}$, exhibited by a twofold and threefold increase in β-galactosidase activities, respectively (**Figure 4A**). However, neither the double substitution K34R/G35N$_{Eco}$, K34R/Q36S$_{Eco}$ nor the triple exchange K34R/G35N/Q36$_{Eco}$ had an alleviating effect on the translational arrest (**Figure 4A**). In summary, our analysis clearly shows that cross-activation of EF-P$_{Eco}$ by EarP$_{Ppu}$ is strictly prohibited in the presence of P32, whereas on the contrary cross-modification solely depends on the protruding residue to be arginine. Combined with our *in vitro* interaction analysis we conclude that specifically the selection against that proline is an adaptation to rescue polyproline stalled ribosomes with α-rhamnosylarginine rather than for efficient modification (**Supplementary Figure S2B**). On the other hand, our data also implies that additional adjustments in the β3Ωβ4 sequence composition have been made to compensate for the negative effect of rhamnosylation (**Figure 1D**).

The Phylogeny of the β3Ωβ4 Composition Unveils the Functional Relationship Between P32 and the Modification Site

Our findings prompted us to investigate the evolutionary order of events resulting in the observed co-occurrence patterns between the residues occupying either the modification site (position 34 according to the numbering of *E. coli* EF-P) or the respective position two amino acids upstream. To this end, we performed a phylogenetic tree reconstruction using the maximum likelihood method from the phytools R package (Revell, 2012). We note that, according to the given bootstrap values, especially the deep branches of the tree might be rather randomly placed and thus the obtained results should be handled with caution (**Supplementary Datasets S2, S3**). Nevertheless, in combination with our biochemical and biophysical data these bioinformatics analyses might provide a plausible rationale for the β3Ωβ4 composition and EF-P modification strategies.

As the lysine at the β3Ωβ4 tip was found in more than three quarters of all EF-P sequences and is also conserved in the eukaryotic and archaeal orthologs e/aIF5A (Dever et al., 2014), we hypothesized that this amino acid is evolutionary ancient. Indeed, we found EF-P with a protruding lysine to be most likely at the root of our tree with subsequent emergence of the first arginine, followed by asparagine, glutamine, and methionine (**Supplementary Figure S6A**).

When reconstructing evolutionary scenarios for position 32 (**Figure 5A**), proline is the most likely amino acid in an EF-P progenitor. The subsequent selection against it in certain EF-P subpopulations strongly correlates with arginine in the protruding position (**Supplementary Figure S6A**) and in turn strengthens our observations that this specific residue is crucial for optimal orientation of β3Ωβ4 (**Figures 3**, **4**). This scenario is further corroborated by the fact that the structural restrictions caused by proline in position 32 have favored the evolution of lysine β-lysylation and lysine 5-aminopentanolylation. Whereas P32 seems to be incompatible with rhamnosylation. Our data further implies that the modification system for the latter emerged subsequent to the phylogenetic recruitment of R34/noP32 (**Figure 5B** and **Supplementary Figure S6B**).

CONCLUSION

In this study, we provide a comprehensive analysis of EF-P β3Ωβ4 and how its sequence composition allows functionalization by chemically and structurally distinct modifications. It might also help to predict the type of novel, yet undiscovered EF-P post-translational functionalization strategies in the >50% of bacteria which do not encode any known modification enzyme. Our assumption is supported by the recent identification of lysine 5-amino-pentanolylation which takes place in *B. subtilis* and presumably a few other firmicutes (Hummels et al., 2017). This EF-P activation strategy chemically resembles β-lysylation and also occurs on a β3/β4 loop with an invariant proline two amino

acids upstream of the modification site. Further, in certain prokaryotes that have a β3Ωβ4 similar to EF-P$_{Eco}$ with lysine at the loop tip or alternatively an asparagine (**Figure 2D**), one can identify a deoxyhypusine synthase (DHS) like protein (**Figure 5B**; Brochier et al., 2004). In eukaryotes and archaea, DHS elongates a lysine in the EF-P ortholog IF5A by an amino-butyryl moiety (Wolff et al., 1990; Prunetti et al., 2016). Accordingly, it is plausible that the bacterial ortholog might attach an analogous modification onto the respective EF-Ps although the experimental connection remains elusive.

The evolutionary flexibility in modification systems and β3Ωβ4 sequence composition is not fully understood yet. However, one could speculate that besides the universally conserved role in alleviating ribosome stalling at polyproline stretches (Gutierrez et al., 2013; Ude et al., 2013) diverse EF-Ps might have extended functionality. In this regard it was reported (Pelechano and Alepuz, 2017; Schuller et al., 2017) that IF5A also acts on non-polyproline arrest motifs and even facilitates termination. Although EF-P activity seems to be restricted to the alleviation of translational arrest situations at consecutive prolines (Ude et al., 2013; Woolstenhulme et al., 2015) it should be noted that all global analyses thus far were performed solely in *E. coli* (Peil et al., 2013; Woolstenhulme et al., 2015) and *Salmonella enterica* (Hersch et al., 2013), both of which depend on (*R*)-β-lysylation of lysine. One might therefore speculate whether other EF-Ps with distinct modifications and β3Ωβ4 sequence composition might have expanded functions similar to eIF5A. EarP-dependent EF-Ps might therefore be of particular interest. First evolved in β-proteobacteria, this EF-P type seems to have spread into certain γ-proteobacterial orders and other phyla (Lassak et al., 2015). Conversely, however, horizontal gene transfer events of EpmABC-dependent EF-Ps into the β-proteobacterial subdivision hardly occur. This in turn could indicate a selection in favor of EF-P arginine rhamnosylation caused either by an expanded target spectrum or improved functionality.

The results of this study also demonstrate the possibility of switching the EarP acceptor substrate specificity. The interaction of EarP with its cognate EF-P has been shown to be both sequence- and structure dependent (Krafczyk et al., 2017; Sengoku et al., 2018). Our data show that a substitution of lysine to arginine in the EF-P of *E. coli* K34R$_{Eco}$ is already sufficient to allow for rhamnosylation in an EarP dependent manner. As the EF-Ps of *E. coli* and *P. putida* share only 30% identity in the EarP-interacting EF-P_N domain, sequence-specific contacts between EF-P and EarP (Krafczyk et al., 2017; Sengoku et al., 2018) might only enhance interaction strength between the two proteins. This is further supported by our corresponding bacterial two-hybrid and *in vitro* NMR analyses. The recognition motif for the AIDA-associated heptosyltransferase Aah has been described as a "short β-strand–short acceptor loop–short β-strand" (Charbonneau et al., 2012). Analogously the two beta-strands bracketing β3Ωβ4 might constitute a structural recognition motif for EarP dependent rhamnosylation. Determining the minimal recognition motif is of particular interest as this information allows for targeted rhamnosylation even for proteins other than

EF-P. Thus, our study also lays the foundation to evolve EarP into a glycosynthase that can ultimately be used in heterologous production of eukaryotic glycoproteins.

AUTHOR CONTRIBUTIONS

DF, MPa, EM, and JL performed the bioinformatic analyses. JH, PJ, and JM performed the NMR studies. WV produced and purified the corresponding proteins. MPf performed isoelectric focusing experiments. RK and ZG performed *in vitro* rhamnosylation assays. WV and RK conducted all other biochemical and genetic analyses of β3Ωβ4 substitution variants of *E. coli* and *P. putida*. WV and MF performed the biochemical analysis with EarP from *S. oneidensis* with contributions from JL. JL, JH, KJ, and DF designed the study. WV, RK, KJ, EM, PJ, JH, and JL wrote the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.01148/full#supplementary-material

REFERENCES

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. doi: 10.1016/j.jmb.2013.03.018

Altenhoff, A. M., Glover, N. M., Train, C. M., Kaleb, K., Warwick Vesztrocy, A., Dylus, D., et al. (2018). The OMA orthology database in 2018: retrieving evolutionary relationships among all domains of life through richer web and programmatic interfaces. *Nucleic Acids Res.* 46, D477–D485. doi: 10.1093/nar/gkx1019

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. doi: 10.1186/1745-6150-5-3

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. doi: 10.1021/bi061328t

Bertani, G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* 62, 293–300.

Bertani, G. (2004). Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *J. Bacteriol.* 186, 595–600. doi: 10.1128/jb.186.3.595-600.2004

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. doi: 10.1126/science.1175800

Brochier, C., Lopez-Garcia, P., and Moreira, D. (2004). Horizontal gene transfer and archaeal origin of deoxyhypusine synthase homologous genes in bacteria. *Gene* 330, 169–176. doi: 10.1016/j.gene.2004.01.018

Buchner, S., Schlundt, A., Lassak, J., Sattler, M., and Jung, K. (2015). Structural and functional analysis of the signal-transducing linker in the pH-responsive one-component system CadC of *Escherichia coli*. *J. Mol. Biol.* 427, 2548–2561. doi: 10.1016/j.jmb.2015.05.001

Bullwinkle, T. J., Zou, S. B., Rajkovic, A., Hersch, S. J., Elgamal, S., Robinson, N., et al. (2013). (*R*)-beta-lysine-modified elongation factor P functions in translation elongation. *J. Biol. Chem.* 288, 4416–4423. doi: 10.1074/jbc.M112.438879

Capella-Gutierrez, S., Silla-Martinez, J. M., and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348

Charbonneau, M. E., Cote, J. P., Haurat, M. F., Reiz, B., Crepin, S., Berthiaume, F., et al. (2012). A structural motif is the recognition site for a new family of bacterial protein *O*-glycosyltransferases. *Mol. Microbiol.* 83, 894–907. doi: 10.1111/j.1365-2958.2012.07973.x

Choi, S., and Choe, J. (2011). Crystal structure of elongation factor P from *Pseudomonas aeruginosa* at 1.75 A resolution. *Proteins* 79, 1688–1693. doi: 10.1002/prot.22992

Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B. (2003). An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* 328, 307–317. doi: 10.1016/s0022-2836(03)00307-3

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293.

Dever, T. E., Gutierrez, E., and Shin, B. S. (2014). The hypusine-containing translation factor eIF5A. *Crit. Rev. Biochem. Mol. Biol.* 49, 413–425. doi: 10.3109/10409238.2014.939608

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. doi: 10.1126/science.1229017

Doerfel, L. K., Wohlgemuth, I., Kubyshkin, V., Starosta, A. L., Wilson, D. N., Budisa, N., et al. (2015). Entropic contribution of elongation factor P to proline positioning at the catalytic center of the ribosome. *J. Am. Chem. Soc.* 137, 12997–13006. doi: 10.1021/jacs.5b07427

Elgamal, S., Katz, A., Hersch, S. J., Newsom, D., White, P., Navarre, W. W., et al. (2014). EF-P dependent pauses integrate proximal and distal signals during translation. *PLoS Genet.* 10:e1004553. doi: 10.1371/journal.pgen.1004553

Epstein, W., and Kim, B. S. (1971). Potassium transport loci in *Escherichia coli* K-12. *J. Bacteriol.* 108, 639–644.

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. doi: 10.1093/nar/gkt1139

Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., et al. (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* 33, 5984–6003. doi: 10.1021/bi00185a040

Finn, R. D., Clements, J., Arndt, W., Miller, B. L., Wheeler, T. J., Schreiber, F., et al. (2015). HMMER web server: 2015 update. *Nucleic Acids Res.* 43, W30–W38. doi: 10.1093/nar/gkv397

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285. doi: 10.1093/nar/gkv1344

Fried, L., Lassak, J., and Jung, K. (2012). A comprehensive toolbox for the rapid construction of *lacZ* fusion reporters. *J. Microbiol. Methods* 91, 537–543. doi: 10.1016/j.mimet.2012.09.023

Gutierrez, E., Shin, B. S., Woolstenhulme, C. J., Kim, J. R., Saini, P., Buskirk, A. R., et al. (2013). eIF5A promotes translation of polyproline motifs. *Mol. Cell* 51, 35–45. doi: 10.1016/j.molcel.2013.04.021

Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose P$_{BAD}$ promoter. *J. Bacteriol.* 177, 4121–4130. doi: 10.1128/jb.177.14.4121-4130.1995

Hanawa-Suetsugu, K., Sekine, S., Sakai, H., Hori-Takemoto, C., Terada, T., Unzai, S., et al. (2004). Crystal structure of elongation factor P from *Thermus thermophilus* HB8. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9595–9600. doi: 10.1073/pnas.0308667101

Helenius, A., and Aebi, M. (2004). Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* 73, 1019–1049. doi: 10.1146/annurev.biochem.73.011303.073752

Hersch, S. J., Wang, M., Zou, S. B., Moon, K. M., Foster, L. J., Ibba, M., et al. (2013). Divergent protein motifs direct elongation factor P-mediated translational regulation in *Salmonella enterica* and *Escherichia coli*. *mBio* 4:e00180-13. doi: 10.1128/mBio.00180-13

Huerta-Cepas, J., Serra, F., and Bork, P. (2016). ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* 33, 1635–1638. doi: 10.1093/molbev/msw046

Hummels, K. R., Witzky, A., Rajkovic, A., Tollerson, R. II, Jones, L. A., Ibba, M., et al. (2017). Carbonyl reduction by YmfI in *Bacillus subtilis* prevents accumulation of an inhibitory EF-P modification state. *Mol. Microbiol.* 106, 236–251. doi: 10.1111/mmi.13760

Huter, P., Arenz, S., Bock, L. V., Graf, M., Frister, J. O., Heuer, A., et al. (2017). Structural basis for polyproline-mediated ribosome stalling and rescue by the translation elongation factor EF-P. *Mol. Cell* 68:e516. doi: 10.1016/j.molcel.2017.10.014

Johansson, M., Ieong, K. W., Trobro, S., Strazewski, P., Aqvist, J., Pavlov, M. Y., et al. (2011). pH-sensitivity of the ribosomal peptidyl transfer reaction dependent on the identity of the A-site aminoacyl-tRNA. *Proc. Natl. Acad. Sci. U.S.A.* 108, 79–84. doi: 10.1073/pnas.1012612107

Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc. Natl. Acad. Sci. U.S.A.* 95, 5752–5756. doi: 10.1073/pnas.95.10.5752

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Katoh, T., Wohlgemuth, I., Nagano, M., Rodnina, M. V., and Suga, H. (2016). Essential structural elements in tRNA^{Pro} for EF-P-mediated alleviation of translation stalling. *Nat. Commun.* 7:11657. doi: 10.1038/ncomms11657

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. doi: 10.1093/nar/gkt1296

Kobayashi, K., Katz, A., Rajkovic, A., Ishii, R., Branson, O. E., Freitas, M. A., et al. (2014). The non-canonical hydroxylase structure of YfcM reveals a metal ion-coordination motif required for EF-P hydroxylation. *Nucleic Acids Res.* 42, 12295–12305. doi: 10.1093/nar/gku898

Korzhnev, D. M., Skrynnikov, N. R., Millet, O., Torchia, D. A., and Kay, L. E. (2002). An NMR experiment for the accurate measurement of heteronuclear spin-lock relaxation rates. *J. Am. Chem. Soc.* 124, 10743–10753. doi: 10.1021/ja0204776

Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., II, et al. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166, 175–176. doi: 10.1016/0378-1119(95)00584-1

Krafczyk, R., Macosek, J., Jagtap, P. K. A., Gast, D., Wunder, S., Mitra, P., et al. (2017). Structural basis for EarP-mediated arginine glycosylation of translation elongation factor EF-P. *mBio* 8:e01412-17. doi: 10.1128/mBio.01412-17

Ladner, C. L., Yang, J., Turner, R. J., and Edwards, R. A. (2004). Visible fluorescent detection of proteins in polyacrylamide gels without staining. *Anal. Biochem.* 326, 13–20. doi: 10.1016/j.ab.2003.10.047

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0

Lassak, J., Keilhauer, E. C., Fürst, M., Wuichet, K., Gödeke, J., Starosta, A. L., et al. (2015). Arginine-rhamnosylation as new strategy to activate translation elongation factor P. *Nat. Chem. Biol.* 11, 266–270. doi: 10.1038/nchembio.1751

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. doi: 10.1111/mmi.13233

Li, X., Krafczyk, R., Macosek, J., Li, Y. L., Zou, Y., Simon, B., et al. (2016). Resolving the α-glycosidic linkage of arginine-rhamnosylated translation elongation factor P triggers generation of the first Arg^{Rha} specific antibody. *Chem. Sci.* 7, 6995–7001. doi: 10.1039/c6sc02889f

Melnikov, S., Mailliot, J., Rigger, L., Neuner, S., Shin, B. S., Yusupova, G., et al. (2016a). Molecular insights into protein synthesis with proline residues. *EMBO Rep.* 17, 1776–1784. doi: 10.15252/embr.201642943

Melnikov, S., Mailliot, J., Shin, B. S., Rigger, L., Yusupova, G., Micura, R., et al. (2016b). Crystal structure of hypusine-containing translation factor eIF5A bound to a rotated eukaryotic ribosome. *J. Mol. Biol.* 428, 3570–3576. doi: 10.1016/j.jmb.2016.05.011

Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Miller, J. H. (1992). *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia Coli and Related Bacteria*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

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. doi: 10.1016/j.bbrc.2007.12.072

Navarre, W. W., Zou, S. B., Roy, H., Xie, J. L., Savchenko, A., Singer, A., et al. (2010). PoxA, YjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. *Mol. Cell* 39, 209–221. doi: 10.1016/j.molcel.2010.06.021

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Niklasson, M., Otten, R., Ahlner, A., Andresen, C., Schlagnitweit, J., Petzold, K., et al. (2017). Comprehensive analysis of NMR data using advanced line shape fitting. *J. Biomol. NMR* 69, 93–99. doi: 10.1007/s10858-017-0141-6

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., Mcveigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733–D745. doi: 10.1093/nar/gkv1189

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. doi: 10.1073/pnas.0809211106

Peil, L., Starosta, A. L., Lassak, J., Atkinson, G. C., Virumae, K., Spitzer, M., et al. (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. doi: 10.1073/pnas.1310642110

Peil, L., Starosta, A. L., Virumae, K., Atkinson, G. C., Tenson, T., Remme, J., et al. (2012). Lys34 of translation elongation factor EF-P is hydroxylated by YfcM. *Nat. Chem. Biol.* 8, 695–697. doi: 10.1038/nchembio.1001

Pelechano, V., and Alepuz, P. (2017). eIF5A facilitates translation termination globally and promotes the elongation of many non polyproline-specific tripeptide sequences. *Nucleic Acids Res.* 45, 7326–7338. doi: 10.1093/nar/gkx479

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. doi: 10.1002/jcc.20084

Prunetti, L., Graf, M., Blaby, I. K., Peil, L., Makkay, A. M., Starosta, A. L., et al. (2016). Deciphering the translation initiation factor 5A modification pathway in halophilic archaea. *Archaea* 2016:7316725. doi: 10.1155/2016/7316725

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. doi: 10.1371/journal.pcbi.1005987

Rajkovic, A., Erickson, S., Witzky, A., Branson, O. E., Seo, J., Gafken, P. R., et al. (2015). Cyclic rhamnosylated elongation factor P establishes antibiotic resistance in *Pseudomonas aeruginosa*. *mBio* 6:e00823. doi: 10.1128/mBio.00823-15

Rajkovic, A., Hummels, K. R., Witzky, A., Erickson, S., Gafken, P. R., Whitelegge, J. P., et al. (2016). Translation control of swarming proficiency in *Bacillus subtilis* by 5-amino-pentanolylated elongation factor P. *J. Biol. Chem.* 291, 10976–10985. doi: 10.1074/jbc.M115.712091

Rajkovic, A., and Ibba, M. (2017). Elongation factor P and the control of translation elongation. *Annu Rev. Microbiol.* 71, 117–131. doi: 10.1146/annurev-micro-090816-093629

Revell, L. J. (2012).). phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3, 217–223. doi: 10.1111/j.2041-210X.2011.00169.x

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. doi: 10.1038/nprot.2010.5

Roy, H., Zou, S. B., Bullwinkle, T. J., Wolfe, B. S., Gilreath, M. S., Forsyth, C. J., et al. (2011). The tRNA synthetase paralog PoxA modifies elongation factor-P with (*R*)-β-lysine. *Nat. Chem. Biol.* 7, 667–669. doi: 10.1038/nchembio.632

Saini, P., Eyler, D. E., Green, R., and Dever, T. E. (2009). Hypusine-containing protein eIF5A promotes translation elongation. *Nature* 459, 118–121. doi: 10.1038/nature08034

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. doi: 10.1038/nbt.1568

Sattler, M., Schleucher, J., and Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Mag. Res. Spectrosc.* 34, 93–158. doi: 10.1016/S0079-6565(98)00025-9

Schlundt, A., Buchner, S., Janowski, R., Heydenreich, T., Heermann, R., Lassak, J., et al. (2017). Structure-function analysis of the DNA-binding domain of a transmembrane transcriptional activator. *Sci. Rep.* 7:1051. doi: 10.1038/s41598-017-01031-9

Schmidt, C., Becker, T., Heuer, A., Braunger, K., Shanmuganathan, V., Pech, M., et al. (2016). Structure of the hypusinylated eukaryotic translation factor eIF-5A bound to the ribosome. *Nucleic Acids Res.* 44, 1944–1951. doi: 10.1093/nar/gkv1517

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. doi: 10.1038/nmeth.2089

Schuller, A. P., Wu, C. C., Dever, T. E., Buskirk, A. R., and Green, R. (2017). eIF5A functions globally in translation elongation and termination. *Mol. Cell* 66, 194.e–205.e. doi: 10.1016/j.molcel.2017.03.003

Sengoku, T., Suzuki, T., Dohmae, N., Watanabe, C., Honma, T., Hikida, Y., et al. (2018). Structural basis of protein arginine rhamnosylation by glycosyltransferase EarP. *Nat. Chem. Biol.* 14, 368–374. doi: 10.1038/s41589-018-0002-y

Starosta, A. L., Lassak, J., Peil, L., Atkinson, G. C., Virumae, K., Tenson, T., et al. (2014a). Translational stalling at polyproline stretches is modulated by the sequence context upstream of the stall site. *Nucleic Acids Res.* 42, 10711–10719. doi: 10.1093/nar/gku768

Starosta, A. L., Lassak, J., Peil, L., Atkinson, G. C., Woolstenhulme, C. J., Virumae, K., et al. (2014b). A conserved proline triplet in Val-tRNA Synthetase and the origin of elongation factor P. *Cell Rep.* 9, 476–483. doi: 10.1016/j.celrep.2014.09.008

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. doi: 10.1074/jbc.M109.039040

Tetsch, L., Koller, C., Haneburger, I., and Jung, K. (2008). The membrane-integrated transcriptional activator CadC of *Escherichia coli* senses lysine indirectly via the interaction with the lysine permease LysP. *Mol. Microbiol.* 67, 570–583. doi: 10.1111/j.1365-2958.2007.06070.x

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. doi: 10.1126/science.1228985

Ude, S. C. M. (2013). *The Role of Elongation Factor EF-P in Translation and in Copy Number Control of the Transcriptional Regulator CadC in Escherichia coli*. Ph.D. dissertation, LMU München, München.

UniProt Consortium (2018). UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 46:269. doi: 10.1093/nar/gky092

Volkwein, W., Maier, C., Krafczyk, R., Jung, K., and Lassak, J. (2017). A versatile toolbox for the control of protein levels using Nε-acetyl-ʟ-lysine dependent amber suppression. *ACS Synth. Biol.* 6, 1892–1902. doi: 10.1021/acssynbio.7b00048

Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., et al. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59, 687–696. doi: 10.1002/prot.20449

Wagih, O. (2017). ggseqlogo: a versatile R package for drawing sequence logos. *Bioinformatics* 33, 3645–3647. doi: 10.1093/bioinformatics/btx469

Witzky, A., Hummels, K. R., Tollerson, R. II, Rajkovic, A., Jones, L. A., Kearns, D. B., et al. (2018). EF-P posttranslational modification has variable impact on polyproline translation in *Bacillus subtilis*. *mBio* 9:e00306-18. doi: 10.1128/mBio.00306-18

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. doi: 10.1074/jbc.M805316200

Wolff, E. C., Park, M. H., and Folk, J. E. (1990). Cleavage of spermidine as the first step in deoxyhypusine synthesis. The role of NAD. *J. Biol. Chem.* 265, 4793–4799.

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. doi: 10.1016/j.celrep.2015.03.014

Woolstenhulme, C. J., Parajuli, S., Healey, D. W., Valverde, D. P., Petersen, E. N., Starosta, A. L., et al. (2013). Nascent peptides that block protein synthesis in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 110, E878–E887. doi: 10.1073/pnas.1219536110

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. doi: 10.1038/nsmb.1889

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. doi: 10.1371/journal.pone.0147907

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* 12, 7–8. doi: 10.1038/nmeth.3213

Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T.-Y. (2017). ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* 8, 28–36. doi: 10.1111/2041-210x.12628

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9:40. doi: 10.1186/1471-2105-9-40
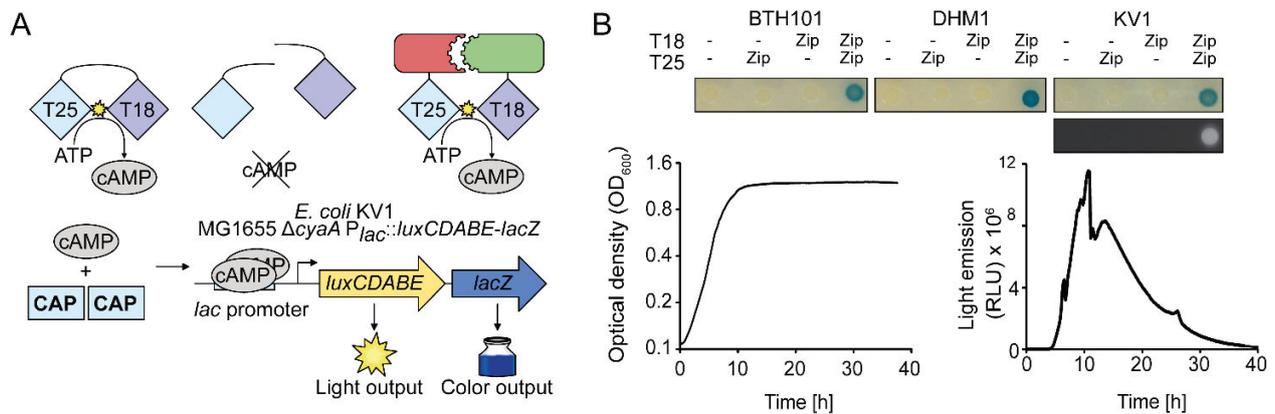
4.1 Supplementary Material

61

Bioinformatic related supplementary datasets S1, S2 and S3 are available online under:

https://www.frontiersin.org/articles/10.3389/fmicb.2019.01148/full#supplementary-material

# Supplementary Material

1 Supplementary Figures and Tables

1.1 Supplementary Figures

Figure S1. (A) Molecular principle of the Euromedex bacterial two-hybrid system with *E. coli* KV1 (adjusted from the Euromedex KIT Manual): The *Bordetella pertussis* adenylate cyclase CyaA apoprotein catalyzes the formation of cyclic AMP from ATP. Splitting the enzyme into two parts – T18 and T25 – renders CyaA inactive even upon co-expression. Fusing T18 and T25 to interacting proteins brings the fragments into close proximity and thus allows reconstitution of the Apo-CyaA. In *E. coli* KV1, the cyclic AMP dependent lac promoter precedes a translational fusion of the *lux*-operon and *lacZ*, allowing the indirect measurement of protein-protein interactions by light emission and colorimetric detection. **(B)** Proof of principle of *E. coli* KV1 as an *in vivo* reporter for protein-protein interaction using the self-interacting leucine zipper of GCN4. The colorimetric detection in *E. coli* KV1 cells was assessed semi quantitatively based on the formation of blue colonies on LB (Miller)-plates containing 40 µg/mL X-Gal. *E. coli* KV1 light emission was measured in a time course experiment recording relative luminescence (RLU) and optical density at 600 nm (OD600) in time intervals of 10 minutes.
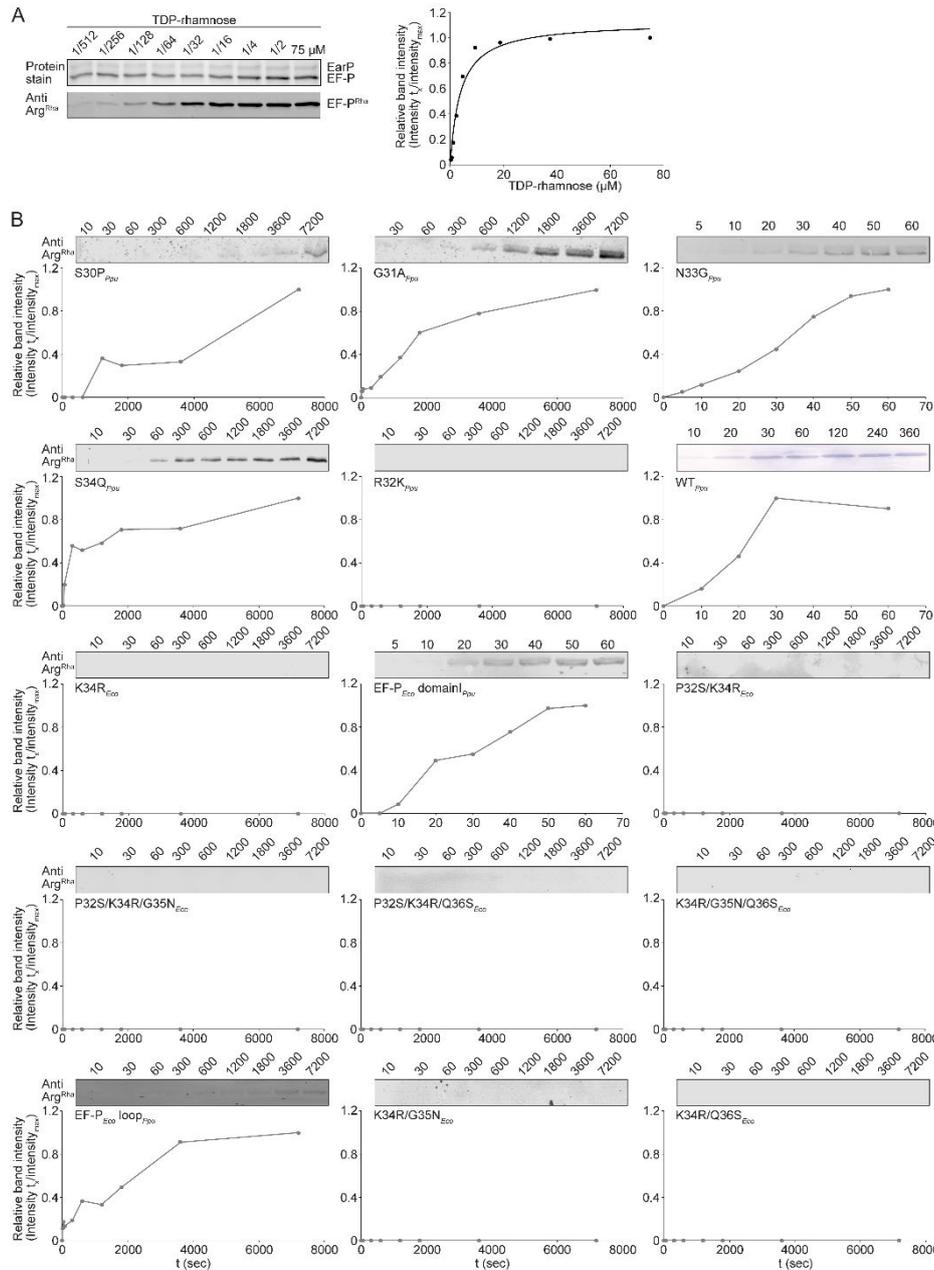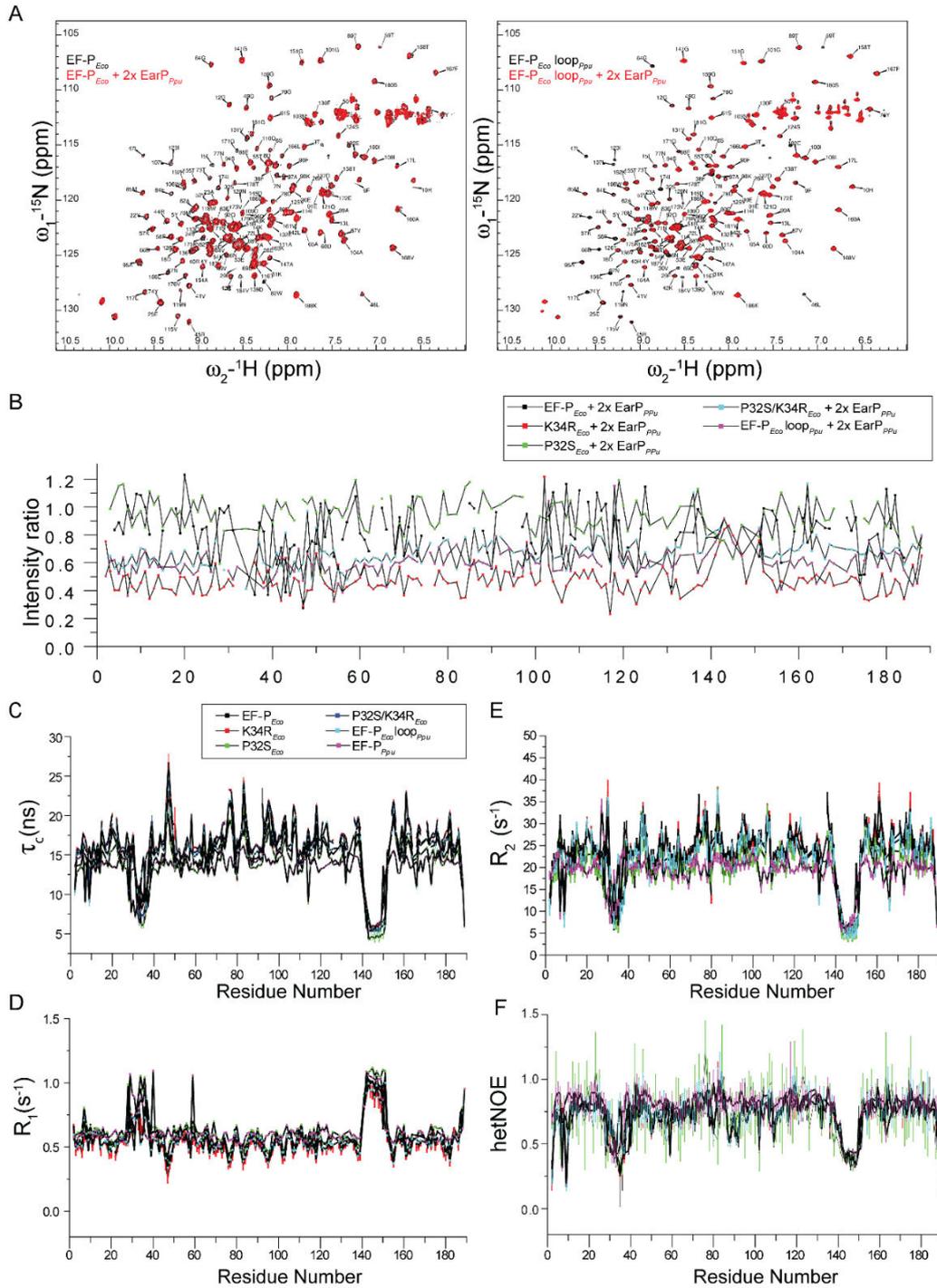
Figure S2. (A) Left: Analysis of EarP$_{Ppu}$ kinetic parameters. 0.5 µg of EF-P$_{Ppu}$ and 0.05 µg of EarP$_{Ppu}$ were subjected to SDS-PAGE after *in vitro* rhamnosylation (see Material and Methods) for 20 seconds at varying TDP-rhamnose concentrations. Proteins were transferred to nitrocellulose membrane by horizontal Western blotting. Rhamnosylated EF-P$_{Ppu}$ was detected using 0.25 µg/ml of Anti-Arg^{Rha}. Right: TDP-rhamnose saturation curve of EarP$_{Ppu}$. Band intensities on nitrocellulose membrane were quantified using ImageJ and relative band intensities were plotted against TDP-rhamnose concentration. **(B)** Timecourse analysis of various EF-P$_{Ppu}$ and EF-P$_{Eco}$ variants. 0.5 µg of EF-P$_{Ppu}$ and 0.05 µg of EarP$_{Ppu}$ were subjected to SDS-PAGE after *in vitro* rhamnosylation at a TDP-rhamnose concentration of 50 µM for varying timespans. Band intensities on nitrocellulose membrane were quantified using ImageJ and relative band intensities were plotted against time.
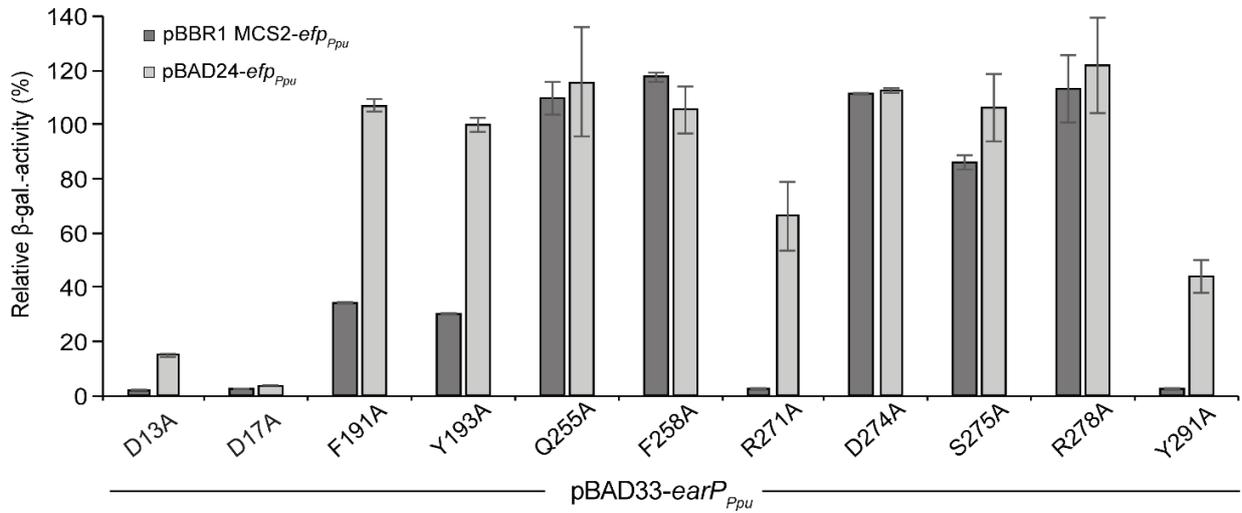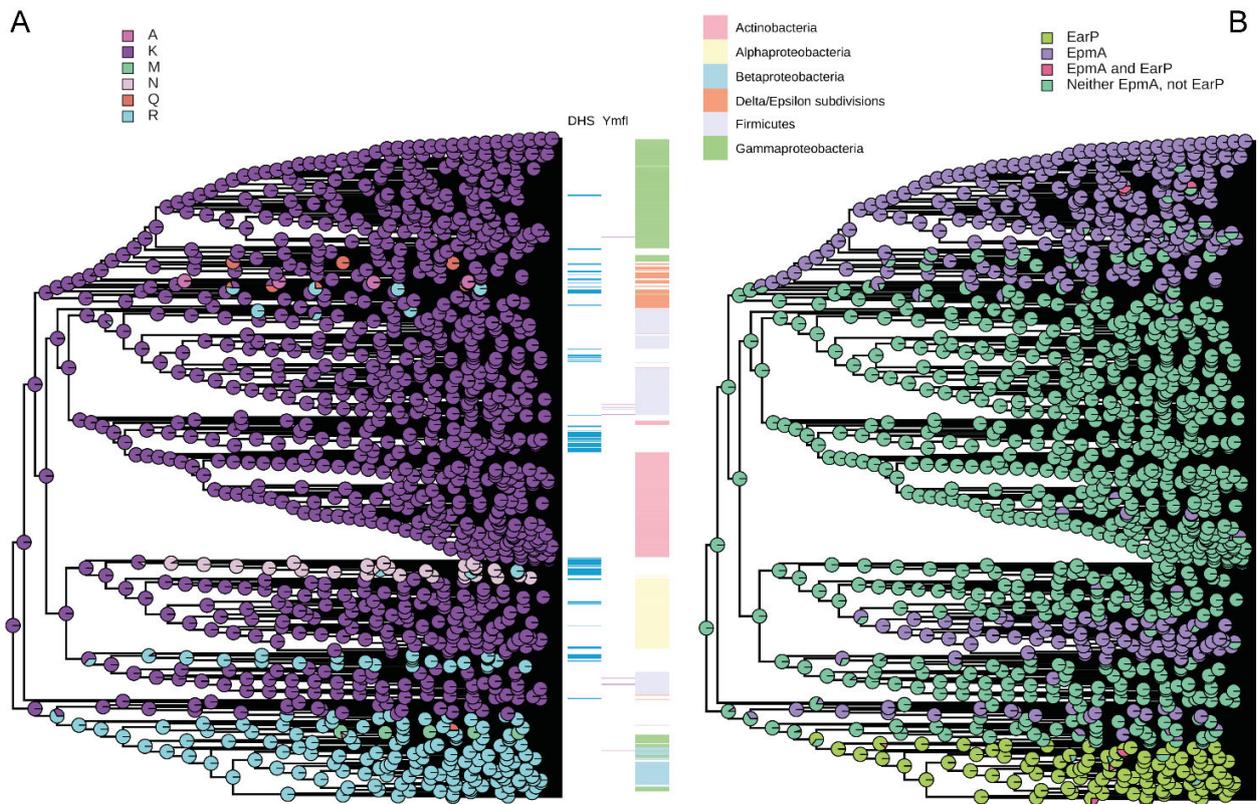
Figure S3. β-galactosidase activity in *E. coli* MG-CR-*efp*-Kan^S upon expression (dark grey) or co-overexpression (light grey) of EF-P$_{Ppu}$ and EarP$_{Ppu}$ and single-amino-acid substitution variants. Cells were incubated under *cadBA*-inducing conditions (LB, pH 5.8) at 30°C o/n.
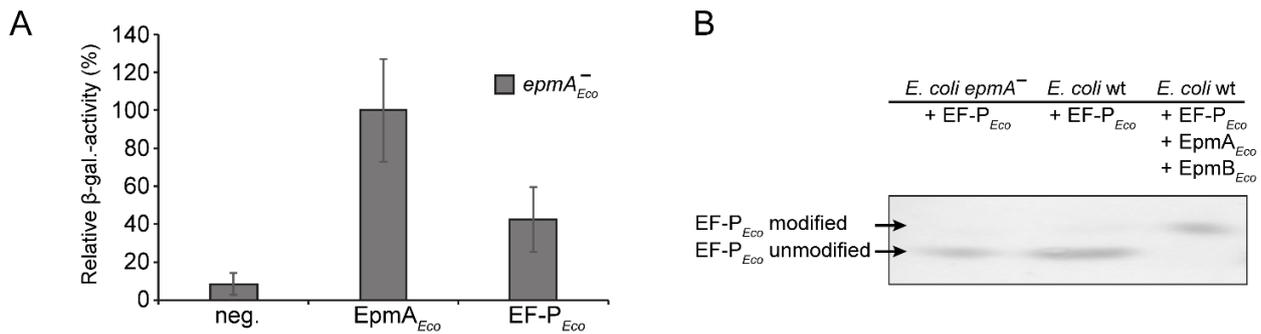
Figure S4. (A) β-galactosidase activity in the Δ*epmA* reporter strain (*E. coli* MG-CL-12-yjeA) harboring either a plasmid borne copy of *epmA* (pBAD33-*epmA*) or *efp* (pBAD24-*efp*$_{Eco}$). **(B)** Isoelectric focusing of *E. coli* EF-P, either overproduced in *E. coli epmA*$^-$ (BW25113-epmA), in which *epmA* was chromosomally deleted, or in *E. coli* wild type cells (BW25113). Furthermore, EF-P was overproduced in combination with its PTM proteins EpmA and EpmB in *E. coli* wild type (BW25113). Production of EF-P was verified by Western blot analysis.

Figure S5. (A) ^{1}H-^{15}N HSQC titrations of EF-P$_{Eco}$ and EF-P$_{Eco}$ loop$_{Ppu}$ with EarP$_{Ppu}$ is shown along with the backbone assignments. Both proteins show a decrease in intensity upon titration with EarP$_{Ppu}$ indicating their interaction with EarP$_{Ppu}$. **(B)** Intensity ratio of all the EF-P$_{Eco}$ mutants on titration with EarP$_{Ppu}$ is shown. **(C)** The correlation time (τ_c) for EF-P$_{Eco}$ and its variants along with **(D)** R1, **(E)** R2 rates and **(F)** hetNOE are shown.

Figure S6. (A) Phylogenetic trees of the EF-P KOW-like N-domain I, with reconstructed state of the 34th position (A, left), presence of known modification systems and taxonomy (A, right colored bars). **(B)** Phylogenetic reconstruction of emergence of EF-P modification systems.

1.2 Supplementary Tables

Table S1. Strains used in this study

Strain	Genotype	Reference
DH5α*λpir*	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 ΔlacZYA-argF* U169 ɸ80d*lacZ*ΔM15 λpir	(Macinga et al., 1995)
LMG194	F⁻ Δ*lacX74 galE galK thi rpsL ΔphoA* (PvuII) Δ*ara714 leu*::Tn*10*	(Guzman et al., 1995)
BL21(DE3)	F⁻ *ompT gal dcm lon hsdS_B*(rB⁻ mB⁻) λ(DE3)	(Studier and Moffatt, 1986)
DHM1	F⁻ *cya-854 recA1 endA1 gyrA96* (*NalR*) *thi1 hsdR17 spoT1 rfbD1 glnV44*(*AS*)	(Karimova et al., 2005)
BTH101	F⁻ *cya-99 araD139 galE15 galK16 rpsL1* (*StrR*) *hsdR2 mcrA1 mcrB1* additional *relA1* mutation reported by (Battesti and Bouveret, 2012)	Euromedex
JW4107	Δ*(araD-araB)567*, Δ*lacZ4787*(::rrnB-3), λ⁻, *rph-1*, Δ*(rhaD-rhaB)568*, Δ*efp-772::kan, hsdR514*	(Baba et al., 2006)
KV1	MG1655 *rpsL150* Δ*cyaA* P*lac*::*luxCDABE-lacZ*	This study
LF1	MG1655 *rpsL150* P*lac*::*rpsL-neo-kan*::*lacZ^{Δ1–100} bp;Kan^R Strp^S*	(Fried et al., 2012)
MG-CR-*efp*	MG1655 Δ*lacZ::tet rpsL150 efp::npt ΔcadBA* P*cadBA*::*lacZ*	(Lassak et al., 2015)
MG-CR-*efp*-KanS	MG1655 Δ*lacZ::tet rpsL150 Δefp ΔcadBA* P*cadBA*::*lacZ*	This study
MG-CR-*efp*-*epmA*-KanR	MG1655 Δ*lacZ::tet rpsL150 Δefp ΔcadBA epmA::npt* P*cadBA*::*lacZ*	This study
MG-CL-12-*yjeA*	MG1655 Δ*lacZ::tet rpsL150 yjeA::npt ΔcadBA cadBA*::*lacZ*	(Ude, 2013)
BW25113	Δ*(araD-araB)567*, Δ*lacZ4787*(::rrnB-3), λ-, rph-1, Δ*(rhaD-rhaB)568*, *hsdR514*	(Datsenko and Wanner, 2000)
BW25113-*epmA*	Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB)568, hsdR514, ΔepmA	This study
JW4116	F⁻, Δ*(araD-araB)567*, Δ*lacZ4787*(::rrnB-3), λ⁻, *rph-1*, Δ*(rhaD-rhaB)568*, Δ*poxA782::kan*, *hsdR514*	(Baba et al., 2006)

Kan^R: kanamycin resistance, Srep^S: streptomycin sensitive

Table S2. Plasmids used in this study

Plasmid	Features	Reference
Plasmids for strain construction		
pRED/ET® Amp	λ-RED recombinase in pBAD24; Amp^R	GeneBridges, Germany
FRT-PGK-gb2-neo-FRT template DNA	PCR-template (plasmid DNA) for generating a FRT-flanked PGK-gb2-neo cassette, Kan^R	GeneBridges, Germany
709-FLPe, amp	Plasmid for removal of FRT flanked resistance cassette, Supplier ID: A106	GeneBridges, Germany
pBAD/HisA-Lux	Contains the *luxCDABE* operon from *Photorhabdus luminescens*	(Volkwein et al., 2017)
Plasmids for mutational analysis of the *E. coli* loop and for overproduction		
pBAD24	Amp^R-cassette, pBBR322 origin, *araC* coding sequence, *ara* operator	(Guzman et al., 1995)
pBAD24-*efp$_{Eco}$*	C-terminal His$_6$-tagged *E. coli efp* amplified from pBAD33-*efp$_{E.c.}$*-His6 (Lassak et al., 2015) using P1/P2	this study
pBAD24-*efp$_{Eco}$* P32G	C-terminal His$_6$-tagged *E. coli efp* substitution variant P32G. Overlap PCR fragment was amplified from pBAD24-*efp$_{Eco}$* using P1/P16 and P2/P15	this study
pBAD24-*efp$_{Eco}$* P32S	C-terminal His$_6$-tagged *E. coli efp* substitution variant P32S. Overlap PCR fragment was amplified from pBAD24-*efp$_{Eco}$* using P1/P18 and P2/P17	this study
pBAD24-*efp$_{Eco}$* G33A	C-terminal His$_6$-tagged *E. coli efp* substitution variant G33A. Overlap PCR fragment was amplified from pBAD24-*efp$_{Eco}$* using P1/P20 and P2/P19	this study
pBAD24-*efp$_{Eco}$* G33S	C-terminal His$_6$-tagged *E. coli efp* substitution variant G33S. Overlap PCR fragment was amplified from pBAD24-*efp$_{Eco}$* using P1/P22 and P2/P21	this study
pBAD24-*efp$_{Eco}$* K34A	C-terminal His$_6$-tagged *E. coli efp* substitution variant K34A. Amplified from pBAD33-*efp$_{E.c.}$*-His6-K34A (Lassak et al., 2015) using P1/P2	this study
pBAD24-*efp$_{Eco}$* K34M	C-terminal His$_6$-tagged *E. coli efp* substitution variant K34M. Overlap PCR fragment was amplified from pBAD24-efpEco using P1/P24 and P2/P23	this study

pBAD24-*efp*$_{Eco}$ K34N	C-terminal His$_6$-tagged *E. coli efp* substitution variant K34N. Overlap PCR fragment was amplified from pBAD24-*efp*$_{Eco}$ using P1/P26 and P2/P25	this study
pBAD24-*efp*$_{Eco}$ K34Q	C-terminal His$_6$-tagged *E. coli efp* substitution variant K34Q. Overlap PCR fragment was amplified from pBAD24-*efp*$_{Eco}$ using P1/P28 and P2/P27	this study
pBAD24-*efp*$_{Eco}$ K34R	C-terminal His$_6$-tagged *E. coli efp* substitution variant K34R. Amplified from pBAD33-*efp*$_{E.c.}$-His6-K34R (Lassak et al., 2015) using P1/P2	this study
pBAD24-*efp*$_{Eco}$ G35N	C-terminal His$_6$-tagged *E. coli efp* substitution variant G35N. Overlap PCR fragment was amplified from pBAD24-*efp*$_{Eco}$ using P1/P30 and P2/P29	this study
pBAD24-*efp*$_{Eco}$ Q36S	C-terminal His$_6$-tagged *E. coli efp* substitution variant Q36S. Overlap PCR fragment was amplified from pBAD24-*efp*$_{Eco}$ using P1/P32 and P2/P31	this study
Plasmids for mutational analysis of the *P. putida* loop and for overproduction		
pBAD24-*efp*$_{Ppu}$	C-terminal His$_6$-tagged *P. putida efp* amplified from *Pseudomonas putida* KT2440 using P1/P2	this study
pBAD24-*efp*$_{Ppu}$ K29A	C-terminal His$_6$-tagged *efp P. putida* substitution variant K29A. Overlap PCR fragment was amplified from pBAD24-*efp*$_{Ppu}$ using P1/P36 and P2/P35	this study
pBAD24-*efp*$_{Ppu}$ K29R	C-terminal His$_6$-tagged *P. putida efp* substitution variant K29R. Overlap PCR fragment was amplified from pBAD24-*efp*$_{Ppu}$ using P1/P37 and P2/P35	this study
pBAD24-*efp*$_{Ppu}$ S30A	C-terminal His$_6$-tagged *P. putida efp* substitution variant S30A. Overlap PCR fragment was amplified from pBAD24-*efp*$_{Ppu}$ using P1/P38 and P2/P35	this study
pBAD24-*efp*$_{Ppu}$ S30G	C-terminal His$_6$-tagged *P. putida efp* substitution variant S30G. Overlap PCR fragment was amplified from pBAD24-*efp*$_{Ppu}$ using P1/P39 and P2/P35	this study
pBAD24-*efp*$_{Ppu}$ S30P	C-terminal His$_6$-tagged *P. putida efp* substitution variant S30P. Overlap PCR fragment was amplified from pBAD24-*efp*$_{Ppu}$ using P1/P40 and P2/P35	this study

pBAD24-*efp_Ppu* G31A	C-terminal His_6-tagged *P. putida efp* substitution variant G31A. Overlap PCR fragment was amplified from pBAD24-*efp_Ppu* using P1/P41 and P2/P35	this study
pBAD24-*efp_Ppu* G31S	C-terminal His_6-tagged *P. putida efp* substitution variant G31S. Overlap PCR fragment was amplified from pBAD24-*efp_Ppu* using P1/P42 and P2/P35	this study
pBAD24-*efp_Ppu* R32K	C-terminal His_6-tagged *P. putida efp* substitution variant R32K. Overlap PCR fragment was amplified from pBAD24-*efp_Ppu* using P1/P44 and P2/P43	this study
pBAD24-*efp_Ppu* N33D	C-terminal His_6-tagged *P. putida efp* substitution variant N33D. Overlap PCR fragment was amplified from pBAD24-*efp_Ppu* using P1/P46 and P2/P45	this study
pBAD24-*efp_Ppu* N33G	C-terminal His_6-tagged *P. putida efp* substitution variant N33G. Overlap PCR fragment was amplified from pBAD24-*efp_Ppu* using P1/P47 and P2/P45	this study
pBAD24-*efp_Ppu* S34A	C-terminal His_6-tagged *P. putida efp* substitution variant S34A. Overlap PCR fragment was amplified from pBAD24-*efp_Ppu* using P1/P48 and P2/P45	this study
pBAD24-*efp_Ppu* S34Q	C-terminal His_6-tagged *P. putida efp* substitution variant S34Q. Overlap PCR fragment was amplified from pBAD24-*efp_Ppu* using P1/P49 and P2/P45	this study
pBAD24-*efp_Ppu* A35S	C-terminal His_6-tagged *P. putida efp* substitution variant A35S. Overlap PCR fragment was amplified from pBAD24-*efp_Ppu* using P1/P50 and P2/P45	this study
pBAD33	Cm^R-cassette, p15A origin, *araC* coding sequence, *ara* operator	(Guzman et al., 1995)
pBAD33 PP1857-His6	C-terminal His6-Tag *earP* version from *P. putida* KT2440	(Krafczyk et al., 2017)

Plasmids for cross modification/actication and overproduction		
pBBRMCS2	pBBR origin of replication, *oriT*, KanR	(Kovach et al., 1995)
pBAD24-*efp$_{Eco}$* domainI-*efp$_{Ppu}$*	C-terminal His$_6$-tagged *E. coli efp* where the first 65 amino acids (domainI) were substituted by the first 65 amino acids from *P. putida efp*. Overlap PCR fragment was amplified from pBAD24-*efp$_{Ppu}$* using P1/P52 and pBAD24-*efp$_{Eco}$* P2/P51	this study
pBAD24-*efp$_{Eco}$* loop-*efp$_{Ppu}$*	C-terminal His$_6$-tagged *E. coli efp* P32S K34R G35N Q36S multiple amino acid substitution variant, corresponding to EF-P of *E. coli* carrying the acceptor loop of EF-P from *P. putida*. Overlap PCR fragment was amplified from pBAD24-*efp$_{Eco}$* using P1/P54 and P2/P53	this study
pBAD24-*efp$_{Eco}$* P32S K34R	C-terminal His$_6$-tagged *E. coli efp* substitution variant P32S K34R. Overlap PCR fragment was amplified from pBAD24-*efp$_{Eco}$* K34R using P1/P56 and P2/P55	this study
pBAD24-*efp$_{Eco}$* K34R G35N	C-terminal His$_6$-tagged *E. coli efp* substitution variant K34R G35N. Overlap PCR fragment was amplified from pBAD24-*efp$_{Eco}$* K34R using P1/P58 and P2/P57	this study
pBAD24-*efp$_{Eco}$* K34R Q36S	C-terminal His$_6$-tagged *E. coli efp* substitution variant K34R Q36S. Overlap PCR fragment was amplified from pBAD24-*efp$_{Eco}$* K34R using P1/P60 and P2/P59	this study
pBAD24-*efp$_{Eco}$* P32S K34R G35N	C-terminal His$_6$-tagged *E. coli efp* substitution variant P32S K34R G35N. Overlap PCR fragment was amplified from pBAD24-*efp$_{Eco}$* P32S K34R using P1/P62 and P2/P61	this study
pBAD24 -*efp$_{Eco}$* P32S K34R Q36S	C-terminal His$_6$-tagged *E. coli efp* substitution variant P32S K34R Q36S. Overlap PCR fragment was amplified from pBAD24-*efp$_{Eco}$* P32S K34R using P1/P64 and P2/P63	this study
pBAD24-*efp$_{Eco}$* K34R G35N Q36S	C-terminal His$_6$-tagged *E. coli efp* substitution variant K34R G35N Q36S. Overlap PCR fragment was amplified from pBAD24-*efp$_{Eco}$* K34R using P1/P66 and P2/P65	this study

pBAD33-*earP$_{So}$*	*earP* from *Shewanella oneidensis* MR-1 amplified from pBAD24-*earP$_{S.o.}$*-His$_6$ (Lassak et al., 2015) using P1/P67	this study
Plasmids for XPPX assay		
p3LC-TL30-APP	p3LC-TL30 + sequence encoding Ala-Pro-Pro	(Peil et al., 2013)
p3LC-TL30-PPD	p3LC-TL30 + sequence encoding Pro-Pro-Asp	(Peil et al., 2013)
p3LC-TL30-PPP	p3LC-TL30 + sequence encoding Pro-Pro-Pro	(Ude et al., 2013)
p3LC-TL30-DPP	p3LC-TL30 + sequence encoding Asp-Pro-Pro	(Peil et al., 2013)
p3LC-TL30-PPG	p3LC-TL30 + sequence encoding Pro-Pro-Gly	(Peil et al., 2013)
p3LC-TL30-PPN	p3LC-TL30 + sequence encoding Pro-Pro-Asn	(Peil et al., 2013)
Plasmids used in the reporter strain MG-CL-12-yjeA		
pBAD33-*epmA*	*epmA* from *E. coli*, amplified using P72/P75	this study
Plasmids for protein overproduction (NMR)		
pET SUMO	pBR322 origin, *lacI*, T7 *lac* promoter, N-terminal His$_6$ tag, SUMO coding sequence, Kan^R, Supplier ID: K300-01	Invitrogen
pET SUMO-*efp$_{Eco}$*	C-terminal genetic fusion of *efp* from *E. coli* to His$_6$-SUMO-tag. Amplified from pBAD24-*efp$_{Eco}$* using P68/P69	this study
pET SUMO-*efp$_{Ppu}$*	C-terminal genetic fusion of *efp* from *P. putida* KT2440 to His$_6$-SUMO-tag	(Krafczyk et al., 2017)
pET SUMO-*efp$_{Eco}$* P32S	C-terminal genetic fusion of *efp* from *E. coli* substitution variant P32S to His$_6$-SUMO-tag. Overlap PCR fragment was amplified from pET SUMO-*efp$_{Eco}$* using P5/P18 and P6/P17	this study
pET SUMO-*efp$_{Eco}$* K34R	C-terminal genetic fusion of *efp* from *E. coli* substitution variant K34R to His$_6$-SUMO-tag. Overlap PCR fragment was amplified from pET SUMO-*efp$_{Eco}$* using P5/P71 and P6/P70	this study
pET SUMO-*efp$_{Eco}$* P32S K34R	C-terminal genetic fusion of *efp* from *E. coli* substitution variant P32S K34R to His$_6$-SUMO-tag. Overlap PCR fragment was amplified from pET SUMO-*efp$_{Eco}$* using P5/P56 and P6/P55	this study
pET SUMO-*efp$_{Eco}$* loop$_{Ppu}$	C-terminal genetic fusion of *efp* from *E. coli* substitution variant P32S K34R G35N Q36S to His$_6$-SUMO-tag. Overlap PCR fragment was amplified from pET SUMO-*efp$_{Eco}$* using P5/P54 and P6/P53	this study

Plasmids for bacterial two-hybrid (BTH)		
pUT18-*zip*	N-terminal genetic fusion of the leucine zipper from GCN4 to the T18 fragment of CyaA	Euromedex
pKT25-*zip*	C-terminal genetic fusion of the leucine zipper from GCN4 to the T25 fragment of CyaA	Euromedex
pUT18C-PP1858	C-terminal genetic fusion of *efp* from *Pseudomonas putida t*o the T18 fragment of CyaA	(Krafczyk et al., 2017)
pKT25-PP1857	C-terminal genetic fusion of *earP* from *Pseudomonas putida* to the T25 fragment of CyaA	(Krafczyk et al., 2017)
pUT18C-*efp_{Eco}*	C-terminal genetic fusion of *efp* from *Escherichia coli* to the T18 fragment of CyaA. PCR fragment was amplified from pBAD24-*efp_{Eco}* using P80/P81	this study
pUT18C-*efp_{Eco}* K34R	C-terminal genetic fusion of *efp* from *Escherichia coli* to the T18 fragment of CyaA. K34R single amino acid exchange variant. PCR fragment was amplified from pBAD24-*efp_{Eco}* K34R using P80/P81	this study
pUT18C-*efp_{Eco}* P32S K34R	C-terminal genetic fusion of *efp* from *Escherichia coli* to the T18 fragment of CyaA. P32S K34R double amino acid exchange variant. PCR fragment was amplified from pBAD24-*efp_{Eco}* P32S K34R using P80/P81	this study
pUT18C-*efp_{Eco}* loop_{Ppu}	C-terminal genetic fusion of *efp* from *E. coli* to the T18 fragment of CyaA. P32S K34R G35N Q36S multiple amino acid exchange variant, corresponding to EF-P of *E. coli* carrying the acceptor loop of EF-P from *P. putida.* PCR fragment was amplified from pBAD24-*efp_{Eco}* loop_{Ppu} using P80/P81	this study
Plasmids for isoelectric focusing		
pBAD33-*efp*-His_6	C-terminal His_6-Tag *efp* version from *E. coli* into pBAD33	(Lassak et al., 2015)
pBAD33-*efp*-His_6-*epmAB*	C-terminal His_6-tagged *efp* and *epmAB from E. coli.* Overlap PCR fragment was amplified using P73/P76, P72/P75 and P74/P77	this study

Plasmids for EarP mutant test		
pBBR1-MCS2 NP_SO_PP1858-His6	C-terminal His6-Tag *efp* from *Pseudomonas putida* under control of the *Shewanella oneidensis* native *efp* promoter in pBBR1-MCS2	(Krafczyk et al., 2017)
pBAD33 PP1857 D13A-His6	C-terminal His$_6$-Tag *earP* from *Pseudomonas putida* in pBAD33. Mutant version; Asp13 exchanged to Ala	(Krafczyk et al., 2017)
pBAD33 PP1857 D17A-His6	C-terminal His$_6$-Tag *earP* from *Pseudomonas putida* in pBAD33. Mutant version; Asp17 exchanged to Ala	(Krafczyk et al., 2017)
pBAD33 PP1857 F191A-His6	C-terminal His$_6$-Tag *earP* from *Pseudomonas putida* in pBAD33. Mutant version; Phe191 exchanged to Ala	(Krafczyk et al., 2017)
pBAD33 PP1857 Y193A-His6	C-terminal His$_6$-Tag *earP* from *Pseudomonas putida* in pBAD33. Mutant version; Tyr193 exchanged to Ala	(Krafczyk et al., 2017)
pBAD33 PP1857 Q255A-His6	C-terminal His$_6$-Tag *earP* from *Pseudomonas putida* in pBAD33. Mutant version; Gln255 exchanged to Ala	(Krafczyk et al., 2017)
pBAD33 PP1857 F258A-His6	C-terminal His$_6$-Tag *earP* from *Pseudomonas putida* in pBAD33. Mutant version; Phe258 exchanged to Ala	(Krafczyk et al., 2017)
pBAD33 PP1857 R271A-His6	C-terminal His$_6$-Tag *earP* from *Pseudomonas putida* in pBAD33. Mutant version; Arg271 exchanged to Ala	(Krafczyk et al., 2017)
pBAD33 PP1857 D274A-His6	C-terminal His$_6$-Tag *earP* from *Pseudomonas putida* in pBAD33. Mutant version; Asp274 exchanged to Ala	(Krafczyk et al., 2017)
pBAD33 PP1857 S275A-His6	C-terminal His$_6$-Tag *earP* from *Pseudomonas putida* in pBAD33. Mutant version; Ser275 exchanged to Ala	(Krafczyk et al., 2017)
pBAD33 PP1857 R278A-His6	C-terminal His$_6$-Tag *earP* from *Pseudomonas putida* in pBAD33. Mutant version; Arg278 exchanged to Ala	(Krafczyk et al., 2017)
pBAD33 PP1857 Y291A-His6	C-terminal His$_6$-Tag *earP* from *Pseudomonas putida* in pBAD33. Mutant version; Tyr291 exchanged to Ala	(Krafczyk et al., 2017)

Amp^R, Cm^R, Kan^R: ampicillin, chloramphenicol, kanamycin resistance.

Table S3. Primers used in this study

Identifier	Oligonucleotide	Sequence (5' - 3')	Restriction site	Reference
Primers for sequencing and cloning				
P1	Seq33 fw	GGC GTC CAC ACT TTG CTA TGC		(Lassak et al., 2010)
P2	pBAD HisA rev	CAG TTC CCT ACT CTC GCA TG		(Lassak et al., 2010)
P3	*epmA* chk fw	TAG GTA CAA CAG TAT AGT CTG ATG GAT AA		this study
P4	*epmA* chk rev	TGA GGC ATG AAA CCA TCC TTC ATT TC		this study
P5	T7 Prom	TAA TAC GAC TCA CTA TAG G		
P6	T7 Term	TAT GCT AGT TAT TGC TCA G		
Primers for strain construction of KV1				
P7	*lacI*-583-fw	GTC TGC GTC TGG CTG GCT GGC ATA		(Fried et al., 2012)
P8	*luxC*-OL-rev	TAG TGC CCA TAG CTG TTT CCT GTG TGA AAT TGT TAT CC		this study
P9	*luxC*-OL-fw	GGA AAC AGC TAT GGG CAC TAA AAA ATT TCA TTC AT TAT TAA CGG		this study
P10	*luxE*-OL-sRBS-*lacZ*-rev	AAT GTA CCT CCT TAC TTT ATT TAT TGT ATT TGT TTA GCT ATC AAA CGC TTC GGT TAA GCT C		this study
P11	OL-sRBS-*lacZ*-fw	ACA AAT ACA ATA AAT AAA GTA AGG AGG TAC ATT ATG ACC ATG ATT ACG GAT TCA CTG GCC G		this study
P12	*lacZ* 500bp anti	CGA CTG TCC TGG CCG TAA CCG ACC		(Fried et al., 2012)
P13	delta *cyaA* fw	GTT GGC GGA ATC ACA GTC ATG ACG GGT AGC AAA TCA GGC GAT ACG TCT TGA ATT AAC CCT CAC TAA AGG GCG		this study

P14	delta *cyaA* rev	TCC GCT AAG ATT GCA TGC CGG ATA AGC CTC GCT TTC CGG CAC GTT CAT CAT AAT ACG ACT CAC TAT AGG GCT C		this study
Primers *E. coli* loop mutation and overproduction constructs				
P15	*efp Eco* P32G fw	GTA AAA GGC GGT AAA GGC CAG G		this study
P16	*efp Eco* P32G rev	CCT GGC CTT TAC CGC CTT TTA C		this study
P17	*efp Eco* P32S fw	CGT AAA ATC GGG TAA AGG CCA GG		this study
P18	*efp Eco* P32S rev	CCT GGC CTT TAC CCG ATT TTA CG		this study
P19	*efp Eco* G33A fw	TAA AAC CGG CGA AAG GCC AGG		this study
P20	*efp Eco* G33A rev	CCT GGC CTT TCG CCG GTT TTA		this study
P21	*efp Eco* G33S fw	GTA AAA CCG TCG AAA GGC CAG G		this study
P22	*efp Eco* G33S rev	CCT GGC CTT TCG ACG GTT TTA C		this study
P23	*efp Eco* K34M fw	GTA AAA CCG GGT ATG GGC CAG GCA TTT		this study
P24	*efp Eco* K34M rev	AAA TGC CTG GCC CAT ACC CGG TTT TAC		this study
P25	*efp Eco* K34N fw	GTA AAA CCG GGT AAC GGC CAG GCA TTT		this study
P26	*efp Eco* K34N rev	AAA TGC CTG GCC GTT ACC CGG TTT TAC		this study
P27	*efp Eco* K34Q fw	GTA AAA CCG GGT CAG GGC CAG GCA TTT		this study
P28	*efp Eco* K34Q rev	AAA TGC CTG GCC CTG ACC CGG TTT TAC		this study
P29	*efp Eco* G35N fw	CCG GGT AAA AAC CAG GCA TTT GC		this study
P30	*efp Eco* G35N rev	GCA AAT GCC TGG TTT TTA CCC GG		this study
P31	*efp Eco* Q36S fw	GGG TAA AGG CAG CGC ATT TGC		this study

P32	*efp Eco* Q36S rev	GCA AAT GCG CTG CCT TTA CCC		this study
Primers *P. putida* loop mutation and overproduction constructs				
P33	NheI-NRBS-PP_1858-fw	GCA CTA GCT AGC CGC GGC CTC GAT TTT TAT AAA TCC	*Nhe*I	this study
P34	XbaI-PP_1858-GS-His6-rev	CGT CTA GAT TAG TGA TGG TGA TGG TGA TGC GAG CCC TTC TTG GAG CGG CCT TTG AA	*Xba*I	this study
P35	*efp Ppu* 29 30 31 OL fw	CGT AAC AGC GCG ATC ATG AAG ACC		this study
P36	*efp Ppu* K29A OL rev	CAT GAT CGC GCT GTT ACG GCC CGA CGC GGT GAA CTC AGC		this study
P37	*efp Ppu* K29R OL rev	CAT GAT CGC GCT GTT ACG GCC CGA GCG GGT GAA CTC AGC		this study
P38	*efp Ppu* S30A OL rev	CAT GAT CGC GCT GTT ACG GCC CGC CTT GGT GAA CTC AGC		this study
P39	*efp Ppu* S30G OL rev	CAT GAT CGC GCT GTT ACG GCC GCC CTT GGT GAA CTC AGC		this study
P40	*efp Ppu* S30P OL rev	CAT GAT CGC GCT GTT ACG GCC CGG CTT GGT GAA CTC AGC		this study
P41	*efp Ppu* G31A OL rev	CAT GAT CGC GCT GTT ACG CGC CGA CTT GGT GAA CTC		this study
P42	*efp Ppu* G31S OL rev	CAT GAT CGC GCT GTT ACG GCT CGA CTT GGT GAA CTC		this study
P43	*efp Ppu* R32K fw	ACC AAG TCG GGC AAG AAC AGC GCG ATC		this study
P44	*efp Ppu* R32K rev	GAT CGC GCT GTT CTT GCC CGA CTT GGT		this study
P45	*efp Ppu* 33 34 35 OL fw	ATC ATG AAG ACC AAG CTG AAG AAC CTG		this study
P46	*efp Ppu* N33D OL rev	CTT CAG CTT GGT CTT CAT GAT CGC GCT ATC ACG GCC CGA CTT		this study

P47	*efp Ppu* N33G OL rev	CTT CAG CTT GGT CTT CAT GAT CGC GCT GCC ACG GCC CGA CTT		this study
P48	*efp Ppu* S34A OL rev	CTT CAG CTT GGT CTT CAT GAT CGC CGC GTT ACG GCC CGA		this study
P49	*efp Ppu* S34Q OL rev	CTT CAG CTT GGT CTT CAT GAT CGC CTG GTT ACG GCC CGA		this study
P50	*efp Ppu* A35S OL rev	CTT CAG CTT GGT CTT CAT GAT GCT GCT GTT ACG GCC		this study
Primers cross modification/activation and overproduction				
P51	*efp Eco* domainI Ppu OL fw	AAG CTG GAC GAC GTG ATC CTG GAT ATG AAC CTG ACT TAC CTG		this study
P52	*efp Eco* domainI Ppu OL rev	CAG GTA AGT CAG GTT CAT ATC CAG GAT CAC GTC GTC CAG CTT		this study
P53	*efp Eco* loop *Ppu* OL fw	AAG TCG GGC CGT AAC AGC GCG TTT GCT CGC GTT AAA CTG CGT		this study
P54	*efp Eco* loop *Ppu* OL rev	CGC GCT GTT ACG GCC CGA CTT TAC GAA TTC ACT CGC TTC AAC		this study
P55	*efp Eco* P32S K34R OL fw	GAA TTC GTA AAA AGC GGT CGC GGC CAG GCA TTT		this study
P56	EF-P *Eco* P32S K34R OL rev	AAA TGC CTG GCC GCG ACC GCT TTT TAC GAA TTC		this study
P57	*efp Eco* K34R G35N OL fw	CCG GGT CGC AAC CAG GCA TTT GCT CG		this study
P58	*efp Eco* K34R G35N OL rev	CGA GCA AAT GCC TGG TTG CGA CCC GG		this study
P59	*efp Eco* K34R Q36S OL fw	CGC GGC AGC GCA TTT GCT CGC GTT A		this study
P60	*efp Eco* K34R Q36S OL rev	TAA CGC GAG CAA ATG CGC TGC CGC G		this study
P61	*efp Eco* P32S K34R G35N OL fw	AGC GGT CGC AAC CAG GCA TTT GCT CG		this study

P62	*efp Eco* P32S K34R G35N OL rev	CGA GCA AAT GCC TGG TTG CGA CCG CT		this study
P63	*efp Eco* P32S K34R Q36S OL fw	AGC GGT CGC GGC AGC GCA TTT GCT CG		this study
P64	*efp Eco* P32S K34R Q36S OL rev	CGA GCA AAT GCG CTG CCG CGA CCG CT		this study
P65	*efp Eco* K34R G35N Q36S OL fw	GGG TCG CAA CAG CGC ATT TG		this study
P66	*efp Eco* K34R G35N Q36S OL rev	CAA ATG CGC TGT TGC GAC CC		this study
P67	*earP So* rev	GCG GTA CCC GAT TTT CTA TTT CAG CGC AGC AT	*Kpn*I	this study
Primers pET SUMO constructs				
P68	*efp Eco*-SUMO-fw	ATG GCA ACG TAC TAT AGC AAC GAT TTT		this study
P69	*efp Eco*-SUMO-rev	TTA GTG ATG GTG ATG GTG ATG GCT		this study
P70	*efp Eco* K34R OL fw	GTA AAA CCG GGT CGC GGC CAG GCA TTT		this study
P71	*efp Eco* K34R OL rev	AAA TGC CTG GCC GCG ACC CGG TTT TAC		this study
Primers isoelectric focusing and pBAD33-*epmA*				
P72	*efp-yjeA*-OL-XbaI-fw	CGT GAA GTA ATC TAG ATT GTC AAA AAC TGG AGA TTT AAC TAT GAG C	*Xba*I	this study
P73	*efp-yjeA*-OL-XbaI-rev	TTT TTG ACA ATC TAG ATT ACT TCA CGC GAG AGA CGT ATT CA	*Xba*I	this study
P74	*yjeA-yjeK*-OL-PstI-fw	GCA TAA CTG CAG GGT AGC TAA GCC ACA AAA TGG CG	*Pst*I	this study
P75	*yjeA-yjeK*-OL-PstI-rev	GCT ACC CTG CAG TTA TGC CCG GTC AAC GCT AAA G	*Pst*I	this study
P76	SacI-RBS-*efp*-fw	GCG ATG AGC TCA ATT AAC AAA TTT CAG AGG GCC TTA TGG	*Sac*I	this study

P77	SphI-*yjeK*-rev	GCA TCG CAT GCT TAC TGC TGG CGT AGC TGG AG	*Sph*I	this study
Primer for strain construction of MG-CR-efp-epmA-KanR				
P78	*epmA* fw	CAC CGC TGT TTG ATT CCT GCG T		this study
P79	*epmA* rev	GCT ACA GAA TGG CGC TTA T CA CG		this study
Primer for for bacterial two-hybrid (BTH)				
P80	XbaI-*efp-Eco* FW	GTA TCG TCT AGA GGC AAC GTA CTA TAG CAA CGA TTT TCG TG	*Xba*I	this study
P81	XmaI-*efp-Eco* Rev	GTA TCG CCC GGG ACT TCA CGC GAG AGA CGT ATT CAC C	*Xma*I	this study

References

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., et al. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2**,** 2006 0008. doi:10.1038/msb4100050

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* 97**,** 6640-6645. doi:10.1073/pnas.120163297

Fried, L., Lassak, J., and Jung, K. (2012). A comprehensive toolbox for the rapid construction of *lacZ* fusion reporters. *J. Microbiol. Methods* 91**,** 537-543. doi:10.1016/j.mimet.2012.09.023

Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J. Bacteriol.* 177**,** 4121-4130.

Karimova, G., Dautin, N., and Ladant, D. (2005). Interaction network among *Escherichia coli* membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis. *J. Bacteriol.* 187**,** 2233-2243. doi:10.1128/JB.187.7.2233-2243.2005

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., 2nd, et al. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166**,** 175-176.

Krafczyk, R., Macosek, J., Jagtap, P.K.A., Gast, D., Wunder, S., Mitra, P., et al. (2017). Structural Basis for EarP-Mediated Arginine Glycosylation of Translation Elongation Factor EF-P. *mBio* 8**,** e01412-01417. doi:10.1128/mBio.01412-17

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. doi:10.1128/AEM.00512-10

Lassak, J., Keilhauer, E.C., Fürst, M., Wuichet, K., Gödeke, J., Starosta, A.L., et al. (2015). Arginine-rhamnosylation as new strategy to activate translation elongation factor P. *Nat. Chem. Biol.* 11**,** 266-270. doi:10.1038/nchembio.1751

Macinga, D.R., Parojcic, M.M., and Rather, P.N. (1995). Identification and analysis of *aarP*, a transcriptional activator of the 2'-N-acetyltransferase in *Providencia stuartii*. *J. Bacteriol.* 177**,** 3407-3413.

Peil, L., Starosta, A.L., Lassak, J., Atkinson, G.C., Virumae, K., Spitzer, M., et al. (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. doi:10.1073/pnas.1310642110

Studier, F.W., and Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189**,** 113-130.

Ude, S.C.M. (2013). *The role of elongation factor EF-P in translation and in copy number control of the transcriptional regulator CadC in Escherichia coli.* Dissertation, LMU München: Faculty of Biology.

Volkwein, W., Maier, C., Krafczyk, R., Jung, K., and Lassak, J. (2017). A versatile toolbox for the control of protein levels using N_{ϵ}-acetyl-L-lysine dependent amber suppression. *ACS Synth. Biol.* 6**,** 1892-1902. doi:10.1021/acssynbio.7b00048

5 Exceptionally Versatile – Arginine in Bacterial Post-translational Protein Modifications

Jürgen Lassak, Franziska Koller, Ralph Krafczyk and <u>Wolfram Volkwein</u> (2019). Exceptionally versatile – Arginine in bacterial post-translational protein modifications. Biol. Chem. 400, 1397-1427.

(Review Article)

Full-text article: https://doi.org/10.1515/hsz-2019-0182

6 Concluding Discussion and Outlook

Polyproline motifs are indispensable to endow proteins with important structural functionalities, and EF-P is needed to translate these motifs efficiently on the ribosome and thus present in all domains of life (Starosta et al., 2014b). In order to fulfill its molecular function, EF-P has to be post-translationally modified. Here different strategies have evolved which differ strongly in their chemical outcome. However, so far it remains unclear how these strategies have evolved and which requirements EF-P has to fulfil in order to be functional with chemical and structural distinct modifications.

The biochemical and bioinformatic data of this thesis now contribute substantially to the understanding of the evolution of the different PTM, and shows how the amino acid composition of the β3/β4 loop (β3Ωβ4) influences the functionality of the chemically divergent modifications. Furthermore, it determines prerequisites for rhamnosylation by the rhamnosyltransferase EarP, and based on this knowledge it demonstrates the possibility of switching the EarP acceptor substrate specificity (Chapter 4). Together this lays the foundation to evolve EarP into a general glycosyltransferase which would then allow the glycosylation of diverse proteins in the future.

In addition to that, a tool for the incorporation of diverse ncAAs into EF-P was established to further investigate the effect of chemical divergent PTMs on EF-Ps functionality. This tool allows to replace the natural PTM of EF-P with synthetic modifications (Chapter 2). Based on this approach, a toolbox was developed to control bacterial protein levels even of essential genes in diverse Gram-negative bacteria (Chapter 3). In addition, the here presented work has also begun to establish a screening system which will ultimately enable the identification of PylRS mutants capable of incorporation new, previously unknown ncAAs into proteins like EF-P (Chapter 2). Finally, this thesis also establishes and provides an improved bacterial two-hybrid (BTH) system for the detection of even transient protein-protein interaction in a time dependent manner by using luminescence as readout possibility (Chapter 4).

6.1 Strategy one: Translational activation of EF-P using the amber suppression system

By incorporating (*R*)-β-lysyl-lysine directly in a single translational step using the amber suppression system we aimed to synthetically activate *E. coli* EF-P. However, in the course of this work it turned out that the wt PylRS, which already accepts more than 20 lysine derivatives without active site engineering (Polycarpo et al., 2006; Wan et al., 2014; Crnkovic et al., 2016), does not accept (*R*)-β-lysyl-lysine as a substrate (Chapter 2). As a consequence, direct incorporation of (*R*)-β-lysyl-lysine was not possible. However, since the discovery of the amber suppression system and its use for the incorporation of ncAAs, also numerous mutants of the PylRS have been generated and described. Hence there is the possibility that one of these mutants accepts (*R*)-β-lysyl-lysine as a substrate. Considering this possibility, a cooperation with Kathrin Lang from the Technical University of Munich has been initiated. Her lab is in possession of several PylRS mutants described in literature and is currently screening

for potential candidates that accept (*R*)-β-lysyl-lysine as substrate. Furthermore, PylRS libraries will be generated in our lab in the future and screened using the FliC-based screening system developed in Chapter 2.

These PylRS libraries will be generated based on rational design since crystal structures of wild-type PylRS (Kavran et al., 2007; Yanagisawa et al., 2008a) and its mutants in complex with numerous substrates are nowadays available (Yanagisawa et al., 2008b; Takimoto et al., 2011; Schneider et al., 2013; Yanagisawa et al., 2013; Flugel et al., 2014; Guo et al., 2014; Schmidt et al., 2014; Englert et al., 2015; Lee et al., 2016; Yanagisawa et al., 2019). This information allows *in silico* docking of new ncAAs into the catalytic fragment (residues 185-454 (Kavran et al., 2007) according to the *M. mazei* numbering) of the PylRS. These *in silico* experiments are essential to generate focused libraries, since the catalytic fragment contains up to 30 active site residues (Baumann et al., 2019), a number which cannot be covered by conventional library design. Furthermore, existing information on PylRS mutants can be integrated in the library design. For example, it is known that the replacement of tyrosine at position 384 with phenylalanine (Y384F) increases *in vivo* suppression activity and *in vitro* aminoacylation (Yanagisawa et al., 2008b). This effect can even be enhanced for certain ncAAs by replacing tyrosine at position 306 with alanine (Y306A). This allows the PylRS binding pocket to accommodate even large lysine derivatives and by that further broadens the substrate spectrum (Yanagisawa et al., 2019).

Regarding the development of the FliC based screening system, it should be mentioned that it is based on a totally different screening principle than the already existing ones. The development was driven by the following two aspects:

- First: Screening systems based on fluorescence reporters such as the green fluorescent protein (GFP) identify the desired cells/mutants by fluorescence-activated cell sorting (Santoro et al., 2002; Kuhn et al., 2010; Reichert et al., 2015). This technology is not available in every lab, elaborate and expensive.

- Second: Other available screening systems different from GFP based ones are usually based on antibiotic resistances such as the chloramphenicol acetyltransferase (CAT) (Pastrnak et al., 2000; Wang et al., 2001) for positive selection, or on toxic gene products such as barnase (Wang et al., 2001; Wang and Schultz, 2001), CcdB (Umehara et al., 2012) or TolC (Amiram et al., 2015) for negative selection. We and others (Umehara et al., 2012) assume that screening systems which are based on different principles, lead in case of successful selection to a different pool of positive variants. Therefore, it seems reasonable to establish a screening system based on the FliC mediated swimming ability of *E. coli*.

One might also wonder why it is even possible to introduce the orthogonal amber suppression system into *E. coli* cells, because the system not only incorporates the ncAA in the gene product of interest at the previously introduced amber position, but also in chromosomally encoded gene products bearing native amber stop codons. The rationale behind this is the fact, that the amber stop codon (UAG) is the least used of

the three stop codons and rarely terminates essential genes (Korkmaz et al., 2014), and if so, is followed either by an ochre (UAA) or an opal (UGA) stop codon at a relatively short distance downstream (Johnson et al., 2012). In *E. coli* MG1655, for example, 321 genes are terminated by the amber stop codon of which just 7 are essential (Lajoie et al., 2013). However, it has been shown that the conversion of these 321 amber stop codons to ochre stop codons has the advantage of creating a new sense codon exclusively for the ncAA of interest, resulting in improved properties for the incorporation of the ncAA by the amber suppression system (Lajoie et al., 2013). For this reason, the *E. coli* strain C321.ΔA.exp, in which all 321 amber stop codons are converted to ochre stop codons, was chosen in this work as the chassis to develop a screening system based on FliC.

Should one of the described strategies lead to the identification of a PylRS mutant capable of incorporating (*R*)-β-lysyl-lysine into EF-P from *E. coli*, the next logical step would be to ask which of the chemical groups or structural properties of (*R*)-β-lysyl-lysine (Figure 11) are essential for its functionality. To determine this, the potential of chemical substrate synthesis could be used to, for example, shorten the chain length of (*R*)-β-lysyl-lysine or add new functional groups to it.
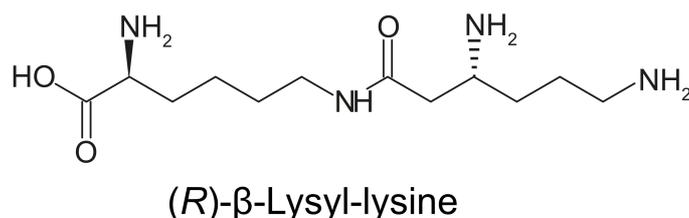

(*R*)-β-Lysyl-lysine

Figure 11: Chemical structure of (*R*)-β-lysyl-lysine.

Moreover, the translational incorporation of (*R*)-β-lysyl-lysine or derivatives of it in *E. coli* EF-P would not only allow to draw conclusions on functional aspects. It would also allow to produce synthetic constitutive active EF-P variants, possibly with even enhanced activity. The timescale EF-P needs to find and bind to diprolyl/polyproline stalled ribosomes in the cell was determined to be about 16 milliseconds (Mohapatra et al., 2017). Increased levels of these synthetic constitutive active EF-Ps could on the one hand reduce this time and on the other hand assist to produce diprolyl/polyproline containing proteins more efficiently.

6.1.1 Alternative strategy to activate EF-P synthetically with different ncAAs

An alternative strategy to equip EF-P with different PTMs could be offered by EpmA. EpmA is a bacterial class II lysyl-tRNA synthetase homolog lacking the tRNA anticodon binding domain (Ambrogelly et al., 2010; Bailly and de Crecy-Lagard, 2010; Navarre et al., 2010; Yanagisawa et al., 2010). Hence it specifically modifies EF-P post-translationally with (*R*)-β-lysine and no endogenous tRNAs. Interestingly, it has been shown that EpmA can also activate lysine analogs *in vitro* (Figure 12A) (Ambrogelly et al., 2010; Roy et al., 2011). A property which could be further broadened by rational design of the active site of EpmA, based on the crystal structure of EpmA as well as the co-crystal structure of EpmA in complex with EF-P (Yanagisawa et al., 2010).

The possibility that EpmA can attach the activated lysine analogs to K34 of EF-P from *E. coli* in a subsequent step seems tempting, but remains a speculation so far (Figure 12B). Again, and since the interaction sites between EF-P and EpmA are known (Yanagisawa et al., 2010; Katz et al., 2014), rational design could enable the EpmA mediated transfer of lysine analogs to K34 of EF-P from *E. coli*. As a result, EF-P could be equipped with so far unknown PTMs (Figure 12C) and their influence on EF-P functionality could be investigated.
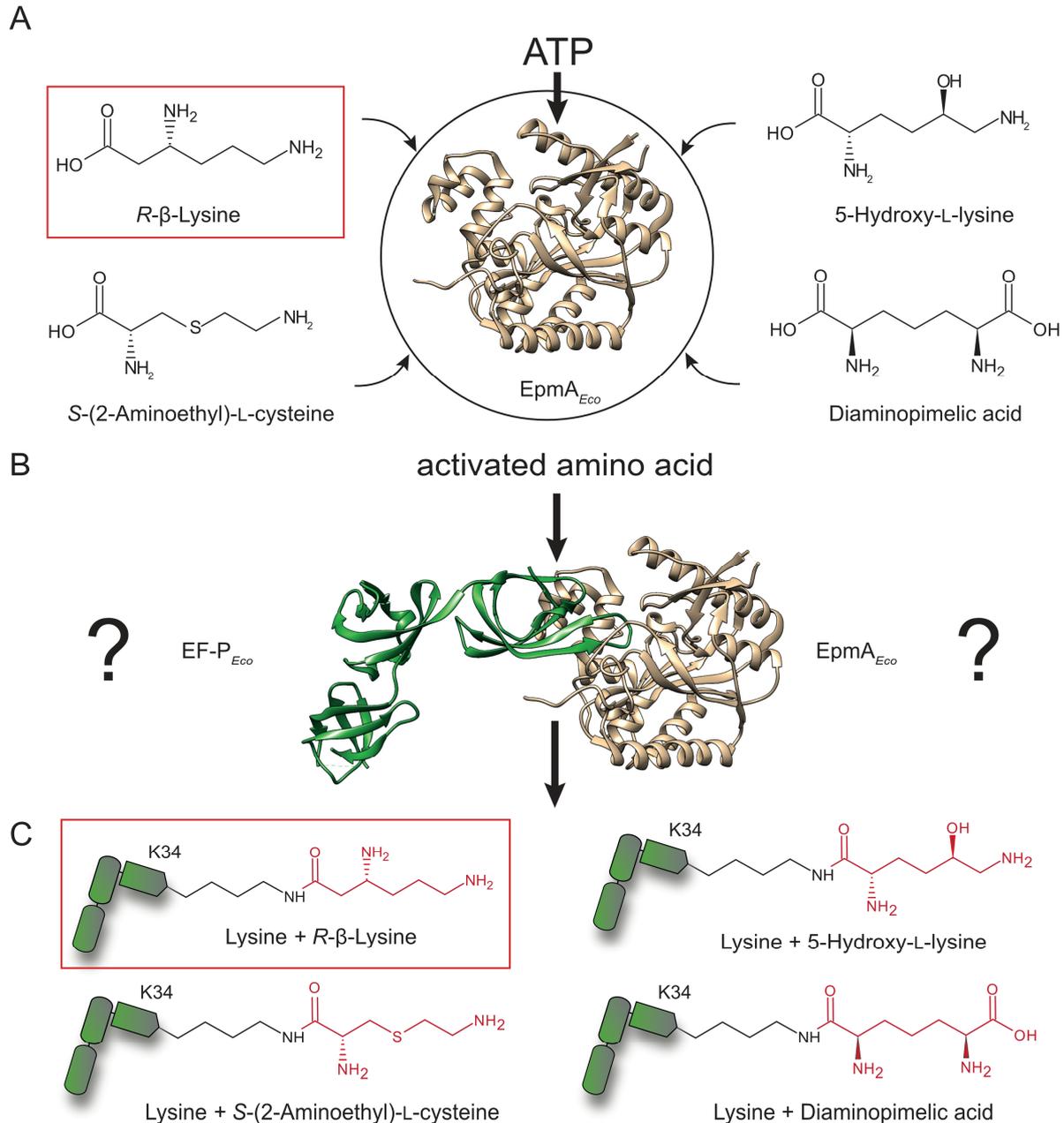
Figure 12: EpmA mediated activation of L-lysine analogs. (A) Examples of lysine analogs which can be activated *in vitro* by EpmA (Ambrogelly et al., 2010; Roy et al., 2011). Red framed the natural substrate for EpmA. **(B)** Crystal structure of EF-P from *E. coli* in complex with EpmA from *E. coli* (PDB: 3A5Z) (Sumida et al., 2010). For visualization, UCSF Chimera was used (Pettersen et al., 2004). Question marks indicate that it is unknown if the activated lysine derivatives from (A) can be transferred onto EF-P or not. **(C)** Resulting post-translational modifications of EF-P if the transfer of the activated amino acids depicted in (A) is possible. Red framed is the natural PTM for EF-P from *E. coli*.

6.2 The amber suppression system as a tool to regulate bacterial protein levels

During the work on the FliC screening system described in Chapter 2, the idea was born to use the amber suppression system not only for the incorporation of the ncAA of interest into proteins, but also to control bacterial protein levels with it. For this purpose, we generated the in Chapter 3 described toolbox based on the amber suppression system which is applicable not only in *E. coli* but also in other Gram-negative bacteria.

During the development, the first decision to make was, which of the numerous substrates available for the amber suppression system is the most suitable for our purpose. In the end, N^ε-acetyl-lysine was chosen for the following reasons, which made this substrate particularly suitable: It is commercially available, cheap and enables the production of unaltered proteins when N^ε-acetyl-lysine is incorporated into a protein instead of a cognate lysine. This is based on the fact that in bacteria the acetyl group is frequently removed by deacetylases (Blander and Guarente, 2004; Yang and Seto, 2008; Zheng et al., 2018) (Figure 13).
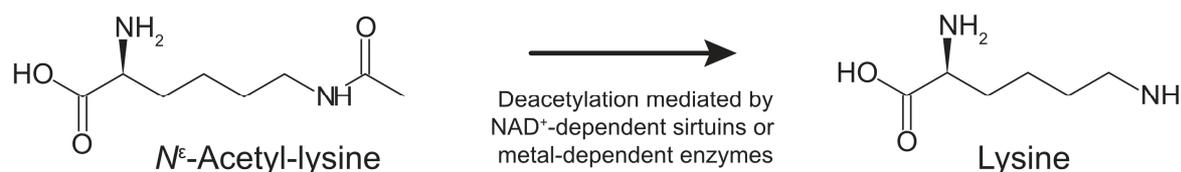
Figure 13: Deacetylation of N^ε-acetyl-lysine in bacteria. (Left) Chemical structure of N^ε-acetyl-lysine which is in bacteria frequently deacetylated, either by sirtuins or metal-dependent enzymes, resulting in lysine **(right).**

Furthermore, we decided to introduce a transcriptional regulation level in addition to the translational regulation level provided by the amber suppression system. This coupling of transcriptional and translational regulation is of particular importance if the system is applied to essential genes. The reason for this is that a two-level control makes it much harder for bacteria to bypass it by introducing spontaneous mutations, in comparison to a one level system. Furthermore, the work presented here shows that protein production at the translational level can be regulated not only by the amber suppression system, but also by the codon context in which the amber stop codon itself is embedded. Finally, it shows that only the described combination of transcriptional and translational regulation allows a complete shutdown of protein production while still being gradually inducible if required, and that the system can be transferred to other Gram-negative bacteria like *S. enterica* and *Vibrio cholerae*.

6.2.1 Strategies to improve the uptake of ncAAs into the cell

The exact mechanism by which the numerous available ncAAs (especially polar ones) for the amber suppression system are taken up by the cell has not been clarified yet. Hence, the idea developed to further improve our toolbox by optimizing the uptake mechanism for N^ε-acetyl-lysine into the cell. Usually the medium is supplemented with N^ε-acetyl-lysine in the millimolar range (Neumann et al., 2008; Umehara et al., 2012; Volkwein et al., 2017). Should it be possible to shift this into the nanomolar range, this

would reduce the substrate costs at least by a factor of 10. To accomplish this goal, a master's thesis was supervised within the scope of this work in which the mechanism for N^ε-acetyl-lysine uptake was started to be investigated using the following approach:

With the help of the database "Transport DB" (Ren et al., 2006), *E. coli* membrane proteins were identified which mediate the uptake of amino acids/peptides or whose substrate specificity is not known. Then, the corresponding deletion mutants were co-transformed with both, the in Chapter 3 described *luxCDABE* reporter system in which an amber stop codon is inserted into the *luxC* reductase subunit and the amber suppression system for N^ε-acetyl-lysine incorporation. As a result, the reductase subunit LuxC is produced only in the presence of N^ε-acetyl-lysine, leading to a luminescence signal (Figure 14A). In contrast, a lack of luminescence indicates that N^ε-acetyl-lysine is not taken up into the cell and that the deleted gene possibly encodes for a putative N^ε-acetyl-lysine transporter (Figure 14B).
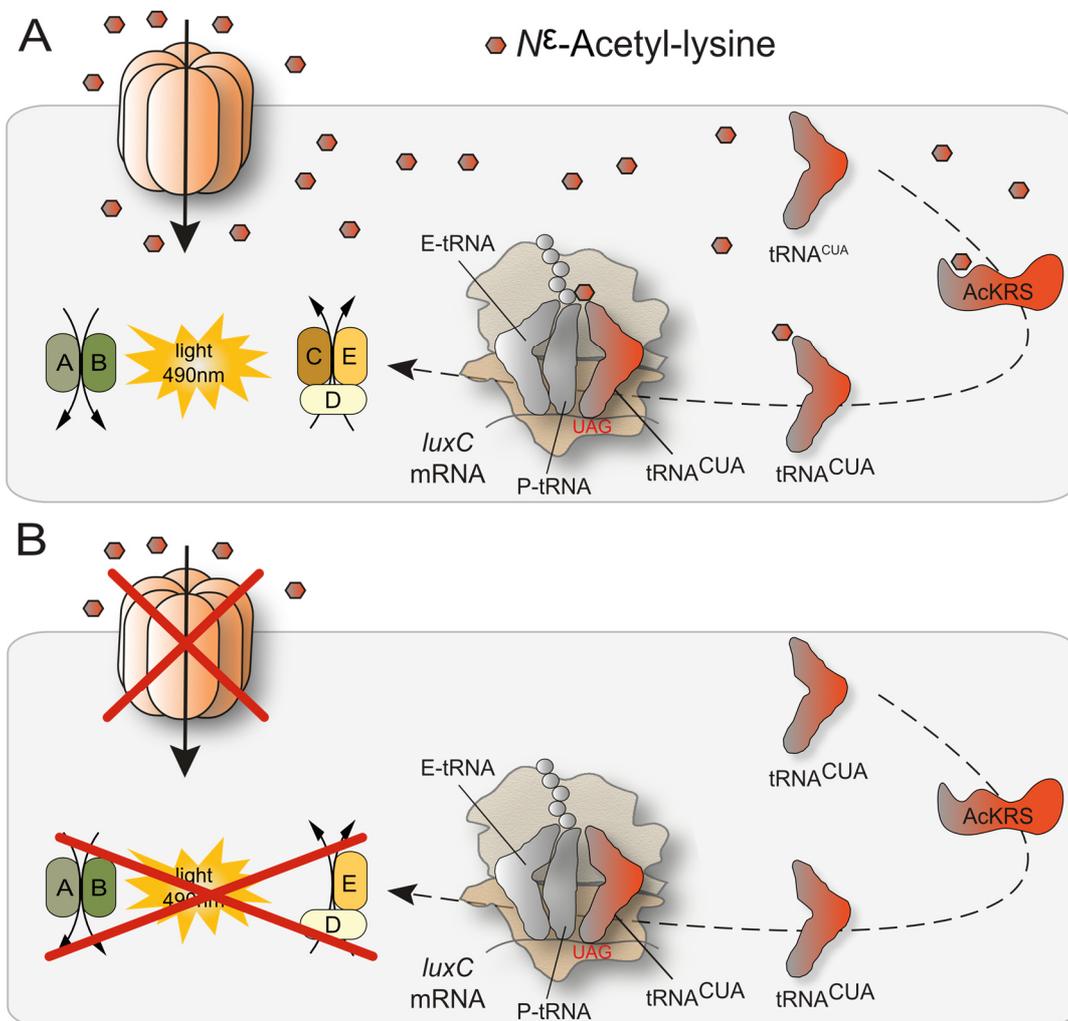
Figure 14: Schematic illustration of the screening system for the identification of N^ε-acetyl-lysine transporters in *E. coli*. (A) In the presence of the corresponding transporter, N^ε-acetyl-lysine is taken up into the cell which contains on the one hand the acetyl lysyl-tRNA synthetase (AcKRS)/tRNA[CUA] pair and on the other hand the *luxCDABE* reporter system in which an amber stop codon is integrated into the *luxC* subunit. Consequently, the presence of N^ε-acetyl-lysine leads to the production of LuxC, which finally results in luminescence. **(B)** The transporter for N^ε-acetyl-lysine is deleted, therefore LuxC cannot be produced since N^ε-acetyl-lysine is absent in the cells and as a result, cells are not emitting light.

Worth mentioning here is that that this approach can also be applied to ncAAs different than N^ε-acetyl-lysine. This is especially of interest for ncAAs which are not commercially available and have to be specifically synthesized. A procedure which is expensive and, also important, time consuming. Here, the reduction of the required substrate quantity would assist the applicability of these substrates in research and application.

The use of the described screening system has already led to the identification of several genes that may be involved in the uptake of N^ε-acetyl-lysine. In order to make a reliable statement whether the identified membrane proteins are involved in the uptake, and if so, to what extent it is necessary to perform complementation experiments with the identified transport proteins in the corresponding deletion mutant background. If the complementation is successful, it would be interesting to see whether overproduction of the transporter has a positive effect on the N^ε-acetyl-lysine import and to what extent the concentration can be reduced while still having a functional system to control bacterial protein levels. Furthermore, it would be interesting to see if other ncAAs besides N^ε-acetyl-lysine are taken up into the cell by corresponding transporters.

Additionally, and in the context of the master's thesis already mentioned, it was also possible to isolate bacteria from soil, which were able to grow in the presence of N^ε-acetyl-lysine as a singular carbon source. This indicates the presence of a high affinity transporter for N^ε-acetyl-lysine. The successfully isolated bacteria belonging to the genera of *Enterobacter*, *Pseudomonas* and *Xanthomonadaceae*. Interestingly, laboratory strains from *P. putida* and *Klebsiella aerogenes* (formerly named *Enterobacter aerogenes*) were not able to grow with N^ε-acetyl-lysine as single carbon source. This fact indicates that either the isolated strains from soil evolved specific transporters to adapt to the environment, or that the laboratory strains lost their ability to do so during ongoing evolution in the lab. However, further research has to be conducted to finally identify the membrane proteins responsible for the transport of N^ε-acetyl-lysine in the identified bacteria.

6.2.2 The N^ε-acetyl-lysine metabolism

Another yet unanswered question is, since the just mentioned bacteria can grow with N^ε-acetyl-lysine as only carbon source, how N^ε-acetyl-lysine is delivered to the cell metabolism. Here, we assume that N^ε-acetyl-lysine is deacetylated by a yet unknown deacetylase, and the resulting lysine is fed into the lysine metabolism of the cell. In this case, two scenarios are possible: Either N^ε-acetyl-lysine is taken up by a transporter and is then deacetylated inside of the cell (Figure 15A) or N^ε-acetyl-lysine is deacetylated outside of the cell and the resulting lysine is transported in the cell where it is fed directly into the lysine metabolism (Figure 15B). To elucidate this, further research is needed in this area to identify the responsible enzyme and clarify its mechanism of action.
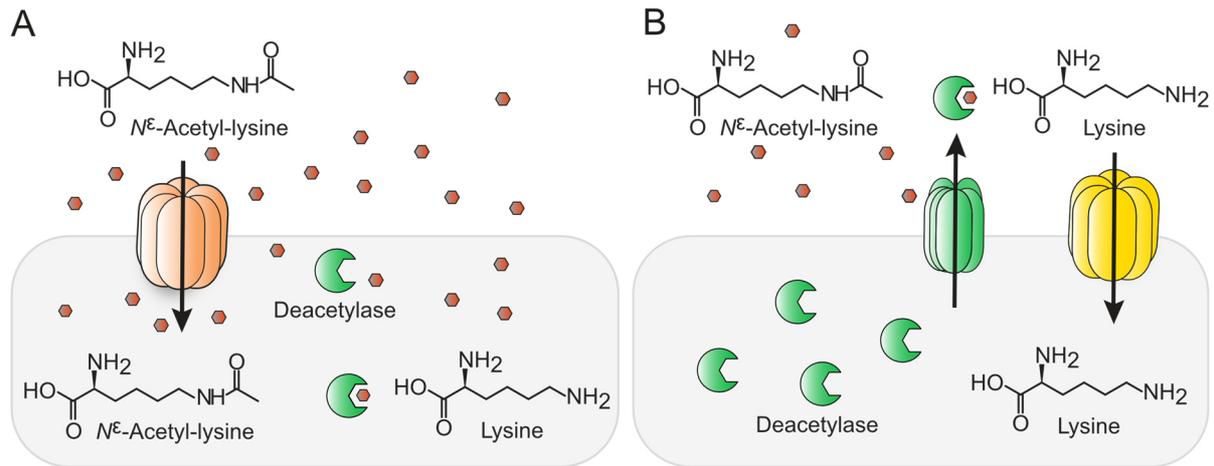
Figure 15: Possibilities of *N*^ε-acetyl-lysine metabolism in bacteria. (A) In this scenario, the transport of *N*^ε-acetyl-lysine into the cell takes place first and in the following step *N*^ε-acetyl-lysine is deacetylated and thus accessible for the lysine metabolism. **(B)** Another possibility is that *N*^ε-acetyl-lysine is degraded first outside of the cell by an exo-deacetylase, and then taken up into the cell as lysine, where it is then fed into the lysine metabolism.

6.3   Strategy two: Post-translational activation of EF-P by non-cognate rhamnosylation

In addition to the efforts to synthetically activate EF-P by translational means (Chapter 2 and 3), a second strategy was pursued in parallel. This strategy took advantage of the fact that despite large sequence diversity, all EF-Ps are structurally similar (Yanagisawa et al., 2010; Choi and Choe, 2011). This similarity in turn led to the hypothesis that cross-activation of EF-P from *E. coli* with a non-cognate modification system might be possible.

6.3.1 Synthetic rhamnosylation of EF-P

As interaction is the prerequisite for cross-activation, an *E. coli* strain that is compatible with the commercially available BTH system was developed in order to determine if there is interaction between *E. coli* EF-P and *P. putida* EarP. The advantage of this strain lies in the fact that it offers, compared to the commercial BTH strains, an additional luminescence readout possibility. This possibility in turn allows continuous monitoring of the luminescence development over a long period of time. Hence, the new strain also enables the detection of even transient protein-protein interactions like we expected for *E. coli* EF-P with its non-cognate modification system EarP from *P. putida*. By using this advantage and subsequent biochemical analyses, we could show that *E. coli* EF-P and *P. putida* EarP do cross-interact, and furthermore, that for rhamnosylation of *E. coli* EF-P by EarP a single amino acid exchange of lysine to arginine on the tip of the β3Ωβ4 is sufficient (Chapter 4).

Despite the fact that this *E. coli* EF-P variant is rhamnosylated but not functional, this result is still of particular interest. It shows that EarP can rhamnosylate a protein which shares only about 30% sequence identity compared to the natural substrate (EF-P from *P. putida*) in the EF-P KOW-like N-domain I that mediates interaction with EarP (Krafczyk et al., 2017; Sengoku et al., 2018). Furthermore, this fact provides first evidence that it could be possible to evolve EarP into a general glycosynthase. In order

to accomplish this, however, the next steps would be to identify A) the minimum recognition motif, and B) the extent to which the overall structure of EF-P affects the recognition by EarP. With this information in hand it should be possible to specifically rhamnosylate proteins other than EF-P, enabling to change the carbochemistry of for example therapeutic proteins and by that providing them with new, improved properties (Sinclair and Elliott, 2005). Here, tools like the first anti-rhamnosyl arginine specific antibody (Li et al., 2016) will be useful to support this development.

6.3.2 Glycoengineering of EarP

Since detailed information on the structure of EarP in complex with TDP-β-L-rhamnose as well as a thorough biochemical characterization of the rhamnosyltransferase reaction are available (Krafczyk et al., 2017; Sengoku et al., 2018; He et al., 2019), it should be possible to rationally engineer the highly conserved TDP-β-L-rhamnose binding pocket in the C-domain of EarP. This could be used to improve the rhamnosyltransferase reaction on the one hand or to broaden the substrate spectrum from EarP to nucleotide sugars other than TDP-β-L-rhamnose on the other hand. EarP, equipped with these characteristics, could help to improve the yield of heterologously produced glycoproteins and introduce new carbochemistry into proteins.

Moreover, the detailed information about EarP residues involved in binding of TDP-β-L-rhamnose could also be used to specifically design EarP-targeted inhibitors. Since it is known that the post-translational modification of EF-P with TDP-β-L-rhamnose is critical for pathogenicity of certain bacteria (Lassak et al., 2015; Rajkovic et al., 2015; Yanagisawa et al., 2016; Tollerson et al., 2018), such small molecule inhibitors could potentially be used as a new class of antibiotics.

6.3.3 Synthetic activation of EF-P

Having shown that *E. coli* EF-P can be rhamnosylated by its non-cognate modification system EarP*,* we were curious to see if it is also possible to generate an *E. coli* EF-P variant which is not just rhamnosylated but also functional, and by that capable of resolving ribosome stalling caused by polyproline stretches. To investigate this, we analyzed in a first step the amino acids composition of the seven amino acids long loop between β3Ωβ4 of *E. coli* and *P. putida* EF-P by biochemical and bioinformatical means. The result was that A: the composition of this loop is essential for EF-Ps ability to rescue polyproline arrested ribosomes, and B: that the amino acid two positions upstream of the modification site is crucial for the functionality of the different chemical modifications. With this knowledge we were able to generate *E. coli* EF-P mutants which are both, rhamnosylated and functional, and showed that a minimal change of only two amino acids in the *E. coli* EF-P loop according to the *P. putida* loop are sufficient to achieve this (Figure 16A and B).
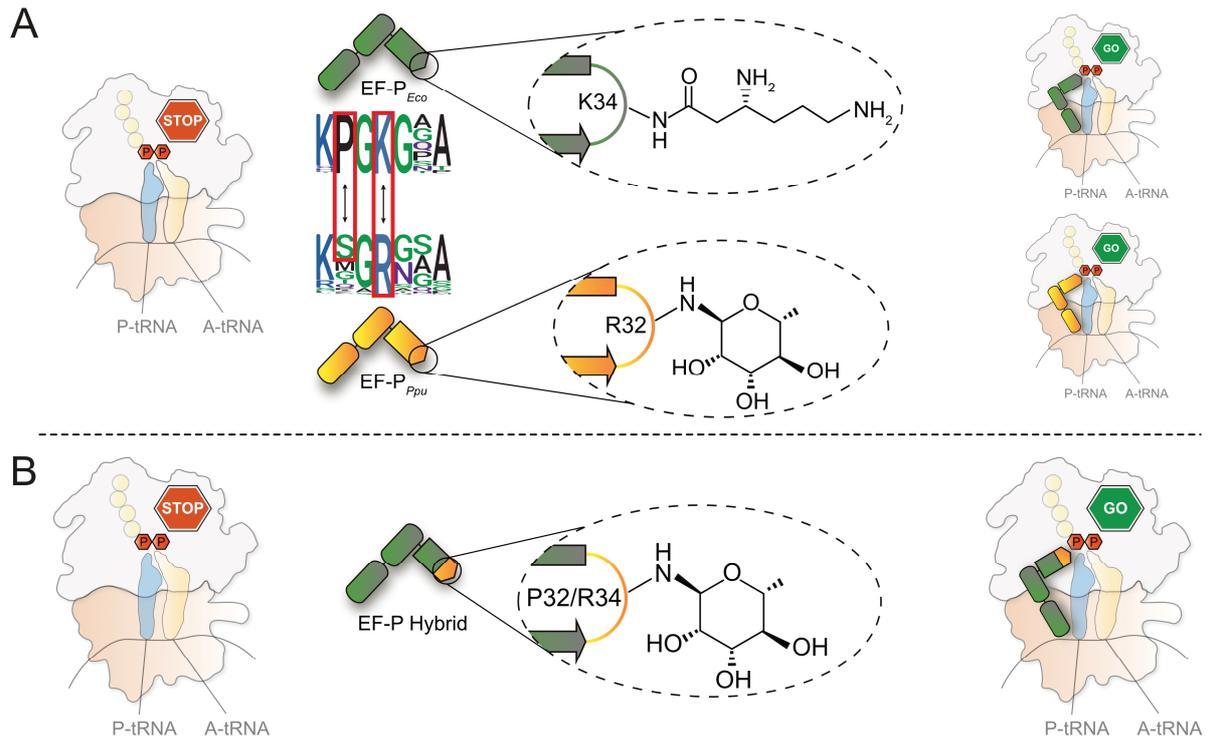
Figure 16: Overview of the different activation strategies for EF-P. (A) Natural activation of EF-P. (Left) Polyproline mediated ribosome stalling. (Middle) Depicted are EF-Ps from *E. coli* and *P. putida* in its activated/post-translationally modified form. The chemical structures of the PTMs are enlarged. Weblogos of the seven amino acids long loop between beta-strands β3/β4 (β3Ωβ4) are shown for lysine and arginine type EF-P. Red boxes in the weblogos indicate amino acids enabling, if switched, the synthetic activation of EF-P$_{Eco}$. (Right) Resolving of the polyproline mediated ribosome stalling by activated (post-translationally modified) EF-Ps from *E coli* and *P putida*. **(B)** Synthetic activation of EF-P$_{Eco}$. Here, two amino acids are changed in the β3Ωβ4 of *E. coli* EF-P according to the weblogo in **(A)**, allowing the switch of the modification strategy from lysylation to rhamnosylation while maintaining EF-Ps functionality to resolve polyproline mediated ribosome stalling.

Interestingly, the second amino acid to change (besides the amino acid of the modification site itself, K34R) is proline, located two amino acids upstream of the modification site, which when exchanged leads to rhamnosylated and additionally functional *E. coli* EF-P. This result led to the hypothesis that the rigidity of proline influences the dynamics of the loop of EF-P and thereby its functionality. However, this could not be proven experimentally with the in Chapter 4 described methods and the exact molecular mechanism by which proline influences functionality therefore remains unclear.

After having shown in Chapter 4 that the rhamnosylation of *E. coli* EF-P by its non-cognate modification system is possible and that the exchange of only two amino acids within the loop region establishes functionality, it would be interesting to see if this approach also works vice versa, resulting in lysylated *P. putida* EF-P. Similar to the work that was presented here, this would help to identify the minimal recognition motif of *E. coli* EpmA and thus lay the foundation for the generation of a general lysylase. Furthermore, the generation of a specific antibody against (*R*)-β-lysyl-lysine would allow the identification of further (*R*)-β-lysylated proteins in divers species.

Another interesting issue arising from the results achieved so far in Chapter 4, which has not yet been addressed, is why different PTMs have evolved. The data presented here and elsewhere (Lassak et al., 2015) clearly show that rhamnosylated heterologously produced EF-P is also functional in *E. coli*, but the rationale behind this evolution of a chemically totally different PTM is so far unclear and can only be speculative at this point. One possible explanation might be that the different chemical modifications offer advantages for a particular set of diprolyl motifs (XPPX) embedded in key proteins of the bacteria affected. These advantages could be, on the one hand, that the protein production is slowed down to provide time for necessary folding processes, or that the production is accelerated to produce more protein in less time.

In this context, another interesting aspect emerges from the bioinformatic data presented in Chapter 4. The data indicate that rhamnosylated EF-P most likely occurred evolutionarily after (*R*)-β-lysylated EF-P, and even more importantly that rhamnosylated EF-P seems to have spread into other phyla, in some cases replacing an already present (*R*)-β-lysylated EF-P variant. On the contrary, replacement of rhamnosylated EF-P by (*R*)-β-lysylated EF-P variants was not observed. This indicates that it might be possible that rhamnosylated EF-P has an extended functionality beyond resolving polyproline mediated ribosome stalling. Extended functionality has already been reported for the eukaryotic counterpart of EF-P (Pelechano and Alepuz, 2017; Schuller et al., 2017) but experimental evidence for rhamnosylated EF-P in prokaryotes is lacking and needs further investigation. Furthermore, the available bioinformatic data show that not all EF-Ps carry an arginine or lysine at the tip of the β3Ωβ4 loop (Chapter 4), which raises the question for alternative modification strategies in addition to the already known ones.

However, it should also be taken into consideration that bacteria might have adapted their genome in order to reduce EF-P dependence to an amount were basal functionality of unmodified EF-P (Volkwein et al., 2019) is sufficient, hence making an additional PTM obsolete. The same might also apply for lysine and arginine EF-Ps were modification pathways could not be identified experimentally and bioinformatically.

Danksagung

Abschließend möchte ich mich noch bei folgenden Personen herzlich bedanken die zur Entstehung dieser Arbeit, sei es wissentlich oder unwissentlich, maßgeblich beigetragen haben:

Mein besonderer Dank gilt hier in erster Linie meinem Betreuer Prof. Dr. Jürgen Lassak für die herausragende Betreuung, für das entgegengebrachte Vertrauen in mich und das stets offene und freundliche Arbeitsklima in dem ich auch dank seiner Hilfe meine Doktorarbeit anfertigen durfte.

Ebenso gilt aber mein besonderer Dank auch Prof. Dr. Kirsten Jung, die mich in Ihre Arbeitsgruppe aufnahm und mich ebenfalls jahrelang unterstützte. Darüber hinaus möchte ich mich bei allen Mitgliedern meiner Prüfungskommission und im Besonderen bei Herrn Prof. Dr. Andreas Klingl bedanken, der sich als Zweitgutachter zur Verfügung gestellt hat. Des Weitern möchte ich mich bei unseren Kooperationspartnern Herrn Prof. Dr. Thomas Carell und Herrn Prof. Dr. Arne Skerra für die lehrreiche Zusammenarbeit bedanken. Mein Dank gilt zudem allen Mitgliedern der Graduiertenschule GRK2062 „Molecular Principles of Synthetic Biology" und insbesondere unserer Koordinatorin Beate Hafner für die hervorragende Organisation der Graduiertenschule.

Frisch zum Masterstudium in München angekommen hatte ich den ersten Kontakt mit der Arbeitsgruppe durch ein Forschungspraktikum, betreut durch Stefan Behr, welcher mich den anderen Mitgliedern zum damaligen Zeitpunkt vorstellte: Luitpold Fried, Sophie Buchner, Martina Rauschmeier, Yang Wang, Hannah Schramke, Laure Plener, Frank Landgraf, Ingrid Weitl, Korinna Burdack, Susanne Bracher, Felix Becker, Sabine Peschek, Michelle Eder, Tatijana Anikanova, Magdalena Motz, Sophie Brameyer und Matthias Reiger. Jeder von euch hat ein kleines Stück dazu beigetragen die Arbeitsgruppe zu dem zu machen was sie zum damaligen Zeitpunkt war und dafür bin ich euch allen dankbar. Während meiner Doktorarbeit habe ich es auch sehr zu schätzen gelernt, dass wir mit Grazyna Wlodarska-Lauer eine Weltklasse Sekretärin bekommen haben, die die Gabe besitz einem in einfachen Worten zu erklären was die Verwaltung eigentlich von einem will und was der ganze Zirkus eigentlich soll.

Eine besondere Beziehung hat jeder Doktorand natürlich auch zu seinen Büropartnern, so auch ich. So durfte ich dank Cláudia Vilhena und Ivica Krištofičová mit eigenen Augen miterleben wie vor meiner Nase ein Schokoladen und Kaffeeimperium sondergleichen aufgebaut wurde, und von denen selbst Drogendealer noch einiges im Bezug auf Kundenbindung und innovatives Marketing hätten lernen können. Ralph Krafczyk und Franziska Koller (nennt irgendjemand dich eigentlich Franziska, Franzi?) waren aufgrund ihres Naturells die wohl angenehmsten Büropartner die man sich vorstellen kann. Von Ralph habe ich außerdem gelernt, dass es durchaus möglich ist mehr als 1000 Worte in einer Minute zu sprechen und dass

Nudeln kombiniert mit Kartoffeln sehr wohl eine vollständige Mahlzeit sein können (natürlich aber nur wenn das ganze ordentlich gesalzen ist). Franzi zeigte einem die kleinen Dinge des Lebens zu schätzen und dass man nicht alles gleich wegschmeißen muss. Claudia, Ivica, Ralph und Franzi: Danke, dass ich mit euch immer was zu lachen hatte, auch wenn's mal Scheiße lief.

Natürlich sind über die Jahre in der Arbeitsgruppe auch neue Personen hinzugekommen welche ich ebenfalls sehr zu schätzen gelernt habe: Simone Eckstein, Miriam Pfab, Jannis Brehm, Florian Fabiani, Ana Gasperotti, Bruno Pinheiro, Larissa Wirtz und Dimitar Petrov, euch allen wünsche ich das Beste für die Zukunft. Gleiches gilt für meine/unsere Studenten Benedikt Frederik Camille Graf von Armansperg, Elena Fajardo Ruiz, Alina Sieber, Anna Haslberger, Laura Lindenthal und Susanne Kurz.

Der größte Dank gehört jedoch meiner ganzen Familie, meinen Geschwistern, Neffen und Nichten und insbesondere meinen Eltern die mich einfach immer das machen ließen was ich für richtig hielt. Nicola und Chris, ihr habt euch bestimmt schon gewundert warum ihr bis jetzt nicht aufgetaucht seid. Der Grund war das es sich besser anfühlte euch in diesem Abschnitt unterzubringen. Danke für die schöne zusammen verbrachte Zeit. Gleiches gilt auch für meinen Freundeskreis aus Eisenharz, meinem Heimatdorf, in dem ich aufwachsen durfte und der mich schon mein ganzes Leben lang begleitet und dies auch hoffentlich noch bis zu dessen Ende tun wird.

So, des wars,

Wolfram

References

AbouElfetouh, A., Kuhn, M.L., Hu, L.I., Scholle, M.D., Sorensen, D.J., Sahu, A.K., Becher, D., Antelmann, H., Mrksich, M., Anderson, W.F., Gibson, B.W., Schilling, B., Wolfe, A.J., 2015. The *E. coli* sirtuin CobB shows no preference for enzymatic and nonenzymatic lysine acetylation substrate sites. Microbiologyopen 4, 66-83.

Abraham, A.K., Pihl, A., 1980. Variable rate of polypeptide chain elongation *in vitro*. Effect of spermidine. Eur. J. Biochem. 106, 257-262.

Ambrogelly, A., O'Donoghue, P., Soll, D., Moses, S., 2010. A bacterial ortholog of class II lysyl-tRNA synthetase activates lysine. FEBS Lett. 584, 3055-3060.

Amiram, M., Haimovich, A.D., Fan, C., Wang, Y.S., Aerni, H.R., Ntai, I., Moonan, D.W., Ma, N.J., Rovner, A.J., Hong, S.H., Kelleher, N.L., Goodman, A.L., Jewett, M.C., Soll, D., Rinehart, J., Isaacs, F.J., 2015. Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. Nat. Biotechnol. 33, 1272-1279.

Aoki, H., Adams, S.L., Chung, D.G., Yaguchi, M., Chuang, S.E., Ganoza, M.C., 1991. Cloning, sequencing and overexpression of the gene for prokaryotic factor EF-P involved in peptide bond synthesis. Nucleic Acids Res. 19, 6215-6220.

Aoki, H., Dekany, K., Adams, S.L., Ganoza, M.C., 1997. The gene encoding the elongation factor P protein is essential for viability and is required for protein synthesis. J. Biol. Chem. 272, 32254-32259.

Aoki, H., Xu, J., Emili, A., Chosay, J.G., Golshani, A., Ganoza, M.C., 2008. Interactions of elongation factor EF-P with the *Escherichia coli* ribosome. FEBS J. 275, 671-681.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., Mori, H., 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2, 2006.0008.

Bailly, M., 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.

Balibar, C.J., Iwanowicz, D., Dean, C.R., 2013. Elongation factor P is dispensable in *Escherichia coli* and *Pseudomonas aeruginosa*. Curr. Microbiol. 67, 293-299.

Baumann, T., Hauf, M., Richter, F., Albers, S., Moglich, A., Ignatova, Z., Budisa, N., 2019. Computational Aminoacyl-tRNA Synthetase Library Design for Photocaged Tyrosine. Int. J. Mol. Sci. 20.

Behshad, E., Ruzicka, F.J., Mansoorabadi, S.O., Chen, D., Reed, G.H., 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.

References

Berg, H.C., 2003. The rotary motor of bacterial flagella. Annu. Rev. Biochem. 72, 19-54.

Bertani, G., 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. J. Bacteriol. 62, 293-300.

Blaha, G., Stanley, R.E., Steitz, T.A., 2009. Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. Science 325, 966-970.

Blander, G., Guarente, L., 2004. The Sir2 family of protein deacetylases. Annu. Rev. Biochem. 73, 417-435.

Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., Shao, Y., 1997. The complete genome sequence of *Escherichia coli* K-12. Science 277, 1453-1462.

Chaney, W.G., Morris, A.J., 1978. Nonuniform size distribution of nascent peptides: the role of messenger RNA. Arch. Biochem. Biophys. 191, 734-741.

Chevance, F.F., Le Guyon, S., Hughes, K.T., 2014. The effects of codon context on *in vivo* translation speed. PLoS Genet. 10, e1004392.

Choi, S., Choe, J., 2011. Crystal structure of elongation factor P from *Pseudomonas aeruginosa* at 1.75 Å resolution. Proteins 79, 1688-1693.

Crnkovic, A., Suzuki, T., Soll, D., Reynolds, N.M., 2016. Pyrrolysyl-tRNA synthetase, an aminoacyl-tRNA synthetase for genetic code expansion. Croat. Chem. Actas 89, 163-174.

Cymer, F., Hedman, R., Ismail, N., von Heijne, G., 2015. Exploration of the arrest peptide sequence space reveals arrest-enhanced variants. J. Biol. Chem. 290, 10208-10215.

Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., Rodnina, M.V., 2013. EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science 339, 85-88.

Doerfel, L.K., Wohlgemuth, I., Kubyshkin, V., Starosta, A.L., Wilson, D.N., Budisa, N., 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.

Dumas, A., Lercher, L., Spicer, C.D., Davis, B.G., 2015. Designing logical codon reassignment - Expanding the chemistry in biology. Chem. Sci. 6, 50-69.

Elgamal, S., Katz, A., Hersch, S.J., Newsom, D., White, P., Navarre, W.W., Ibba, M., 2014. EF-P dependent pauses integrate proximal and distal signals during translation. PLoS Genet. 10, e1004553.

Englert, M., Nakamura, A., Wang, Y.S., Eiler, D., Soll, D., Guo, L.T., 2015. Probing the active site tryptophan of *Staphylococcus aureus* thioredoxin with an analog. Nucleic Acids Res. 43, 11061-11067.

References

Flugel, V., Vrabel, M., Schneider, S., 2014. Structural basis for the site-specific incorporation of lysine derivatives into proteins. PLoS One 9, e96198.

Gattner, M.J., Vrabel, M., Carell, T., 2013. Synthesis of ε-*N*-propionyl-, ε-*N*-butyryl-, and ε-*N*-crotonyl-lysine containing histone H3 using the pyrrolysine system. Chem. Commun. (Camb) 49, 379-381.

Glick, B.R., Chladek, S., 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.

Glick, B.R., Ganoza, M.C., 1975. Identification of a soluble protein that stimulates peptide bond synthesis. Proc. Natl. Acad. Sci. U. S. A. 72, 4257-4260.

Guo, L.T., Wang, Y.S., Nakamura, A., Eiler, D., Kavran, J.M., Wong, M., Kiessling, L.L., Steitz, T.A., O'Donoghue, P., Soll, D., 2014. Polyspecific pyrrolysyl-tRNA synthetases from directed evolution. Proc. Natl. Acad. Sci. U. S. A. 111, 16724-16729.

Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., Kim, J.R., Saini, P., Buskirk, A.R., Dever, T.E., 2013. eIF5A promotes translation of polyproline motifs. Mol. Cell 51, 35-45.

Hanawa-Suetsugu, K., Sekine, S., Sakai, H., Hori-Takemoto, C., Terada, T., Unzai, S., Tame, J.R., Kuramitsu, S., Shirouzu, M., Yokoyama, S., 2004. Crystal structure of elongation factor P from *Thermus thermophilus* HB8. Proc. Natl. Acad. Sci. U. S. A, 101, 9595-9600.

He, C., Liu, N., Li, F., Jia, X., Peng, H., Liu, Y., Xiao, Y., 2019. Complex structure of *Pseudomonas aeruginosa* arginine rhamnosyltransferase EarP with its acceptor elongation factor P. J. Bacteriol.

Hersch, S.J., Elgamal, S., Katz, A., Ibba, M., Navarre, W.W., 2014. Translation initiation rate determines the impact of ribosome stalling on bacterial protein synthesis. J. Biol. Chem. 289, 28160-28171.

Hersch, S.J., Wang, M., Zou, S.B., Moon, K.M., Foster, L.J., Ibba, M., Navarre, W.W., 2013. Divergent protein motifs direct elongation factor P-mediated translational regulation in *Salmonella enterica* and *Escherichia coli*. mBio 4, e00180-00113.

Hummels, K.R., Witzky, A., Rajkovic, A., Tollerson, R., 2nd, Jones, L.A., Ibba, M., 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.

Huter, P., Arenz, S., Bock, L.V., Graf, M., Frister, J.O., Heuer, A., Peil, L., Starosta, A.L., Wohlgemuth, I., Peske, F., Novacek, J., Berninghausen, O., Grubmüller, H., Tenson, T., Beckmann, R., Rodnina, M.V., Vaiana, A.C., Wilson, D.N., 2017. Structural Basis for Polyproline-Mediated Ribosome Stalling and Rescue by the Translation Elongation Factor EF-P. Mol. Cell 68, 515-527 e516.

Iannino, F., Ugalde, J.E., Inon de Iannino, N., 2012. *Brucella abortus efp* gene is required for an efficient internalization in HeLa cells. Microb. Pathog. 52, 31-40.

References

Johansson, M., Ieong, K.W., Trobro, S., Strazewski, P., Aqvist, J., Pavlov, M.Y., Ehrenberg, M., 2011. pH-sensitivity of the ribosomal peptidyl transfer reaction dependent on the identity of the A-site aminoacyl-tRNA. Proc. Natl. Acad. Sci. U. S. A. 108, 79-84.

Johnson, D.B.F., Wang, C., Xu, J.F., Schultz, M.D., Schmitz, R.J., Ecker, J.R., Wang, L., 2012. Release factor one is nonessential in *Escherichia coli*. ACS Chem. Biol. 7, 1337-1344.

Jovine, L., Djordjevic, S., Rhodes, D., 2000. The crystal structure of yeast phenylalanine tRNA at 2.0 Å resolution: cleavage by Mg^{2+} in 15-year old crystals. J. Mol. Biol. 301, 401-414.

Katz, A., Solden, L., Zou, S.B., Navarre, W.W., Ibba, M., 2014. Molecular evolution of protein-RNA mimicry as a mechanism for translational control. Nucleic Acids Res. 42, 3261-3271.

Kavran, J.M., Gundllapalli, S., O'Donoghue, P., Englert, M., Soll, D., Steitz, T.A., 2007. Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. Proc. Natl. Acad. Sci. U. S. A. 104, 11268-11273.

Kearns, D.B., Chu, F., Rudner, R., Losick, R., 2004. Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility. Mol. Microbiol. 52, 357-369.

Kobayashi, K., Katz, A., Rajkovic, A., Ishii, R., Branson, O.E., Freitas, M.A., Ishitani, R., Ibba, M., Nureki, O., 2014a. The non-canonical hydroxylase structure of YfcM reveals a metal ion-coordination motif required for EF-P hydroxylation. Nucleic Acids Res. 42, 12295-12305.

Kobayashi, K., Suzuki, T., Dohmae, N., Ishitani, R., Nureki, O., 2014b. Crystallization and preliminary X-ray crystallographic analysis of YfcM: an important factor for EF-P hydroxylation. Acta Crystallogr. F Struct. Biol. Commun. 70, 1236-1239.

Korkmaz, G., Holm, M., Wiens, T., Sanyal, S., 2014. Comprehensive analysis of stop codon usage in bacteria and its correlation with release factor abundance. J. Biol. Chem. 289, 30334-30342.

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., 2nd, Peterson, K.M., 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene 166, 175-176.

Krafczyk, R., Macošek, J., Jagtap, P.K.A., Gast, D., Wunder, S., Mitra, P., Jha, A.K., Rohr, J., Hoffmann-Röder, A., Jung, K., Hennig, J., Lassak, J., 2017. Structural basis for EarP-mediated arginine glycosylation of translation elongation factor EF-P. mBio 8.

Kuhn, S.M., Rubini, M., Fuhrmann, M., Theobald, I., Skerra, A., 2010. Engineering of an orthogonal aminoacyl-tRNA synthetase for efficient incorporation of the non-natural amino acid O-methyl-L-tyrosine using fluorescence-based bacterial cell sorting. J. Mol. Biol. 404, 70-87.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.

Lajoie, M.J., Rovner, A.J., Goodman, D.B., Aerni, H.R., Haimovich, A.D., Kuznetsov, G., Mercer, J.A., Wang, H.H., Carr, P.A., Mosberg, J.A., Rohland, N., Schultz, P.G., Jacobson, J.M., Rinehart, J., Church, G.M., Isaacs, F.J., 2013. Genomically recoded organisms expand biological functions. Science 342, 357-360.

Lassak, J., Keilhauer, E.C., Fürst, M., Wuichet, K., Gödeke, J., Starosta, A.L., Chen, J.M., Sogaard-Andersen, L., Rohr, J., Wilson, D.N., Haussler, S., Mann, M., Jung, K., 2015. Arginine-rhamnosylation as new strategy to activate translation elongation factor P. Nat. Chem. Biol. 11, 266-270.

Lee, Y.J., Schmidt, M.J., Tharp, J.M., Weber, A., Koenig, A.L., Zheng, H., Gao, J., Waters, M.L., Summerer, D., Liu, W.R., 2016. Genetically encoded fluorophenylalanines enable insights into the recognition of lysine trimethylation by an epigenetic reader. Chem. Commun. (Camb) 52, 12606-12609.

Li, X., Krafczyk, R., Macosek, J., Li, Y.L., Zou, Y., Simon, B., Pan, X., Wu, Q.Y., Yan, F., Li, S., Hennig, J., Jung, K., Lassak, J., Hu, H.G., 2016. Resolving the alpha-glycosidic linkage of arginine-rhamnosylated translation elongation factor P triggers generation of the first Arg^{Rha} specific antibody. Chem. Sci. 7, 6995-7001.

Lizardi, P.M., Mahdavi, V., Shields, D., Candelas, G., 1979. Discontinuous translation of silk fibroin in a reticulocyte cell-free system and in intact silk gland cells. Proc. Natl. Acad. Sci. U. S. A. 76, 6211-6215.

Macinga, D.R., Parojcic, M.M., Rather, P.N., 1995. Identification and analysis of *aarP*, a transcriptional activator of the 2'-*N*-acetyltransferase in *Providencia stuartii*. J. Bacteriol. 177, 3407-3413.

Mandal, A., Mandal, S., Park, M.H., 2014. Genome-wide analyses and functional classification of proline repeat-rich proteins: potential role of eIF5A in eukaryotic evolution. PLoS One 9, e111800.

Miller, J.H., 1972. Experiments in molecular genetics. Cold Spring Harbor, N.Y.

Miller, J.H., 1992. A short course in bacterial genetics: A laboratory manual for *Escherichia coli* and related bacteria. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 456 pp.

Mohapatra, S., Choi, H., Ge, X., Sanyal, S., Weisshaar, J.C., 2017. Spatial Distribution and Ribosome-Binding Dynamics of EF-P in Live *Escherichia coli*. mBio 8.

Morgan, A.A., Rubenstein, E., 2013. Proline: the distribution, frequency, positioning, and common functional roles of proline and polyproline sequences in the human proteome. PLoS One 8, e53785.

Muto, H., Ito, K., 2008. Peptidyl-prolyl-tRNA at the ribosomal P-site reacts poorly with puromycin. Biochem. Biophys. Res. Commun. 366, 1043-1047.

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., 2010. PoxA, YjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. Mol. Cell 39, 209-221.

Neumann, H., Peak-Chew, S.Y., Chin, J.W., 2008. Genetically encoding N^ε-acetyllysine in recombinant proteins. Nat. Chem. Biol. 4, 232-234.

Ohashi, Y., Inaoka, T., Kasai, K., Ito, Y., Okamoto, S., Satsu, H., Tozawa, Y., Kawamura, F., Ochi, K., 2003. Expression profiling of translation-associated genes in sporulating *Bacillus subtilis* and consequence of sporulation by gene inactivation. Biosci. Biotechnol. Biochem. 67, 2245-2253.

Park, J.H., Johansson, H.E., Aoki, H., Huang, B.X., Kim, H.Y., Ganoza, M.C., 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.

Pastrnak, M., Magliery, T.J., Schultz, P.G., 2000. A new orthogonal suppressor tRNA/aminoacyl-tRNA synthetase pair for evolving an organism with an expanded genetic code. Helvetica Chimica Acta 83, 2277-2286.

Pavlov, M.Y., Watts, R.E., Tan, Z., Cornish, V.W., Ehrenberg, M., 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.

Peil, L., Starosta, A.L., Lassak, J., Atkinson, G.C., Virumae, K., Spitzer, M., Tenson, T., Jung, K., Remme, J., Wilson, D.N., 2013. Distinct X/PP/X 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.

Peil, L., Starosta, A.L., Virumae, K., Atkinson, G.C., Tenson, T., Remme, J., Wilson, D.N., 2012. Lys34 of translation elongation factor EF-P is hydroxylated by YfcM. Nat. Chem. Biol. 8, 695-697.

Pelechano, V., Alepuz, P., 2017. eIF5A facilitates translation termination globally and promotes the elongation of many non polyproline-specific tripeptide sequences. Nucleic Acids Res. 45, 7326-7338.

Peng, W.T., Banta, L.M., Charles, T.C., 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.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605-1612.

Polycarpo, C.R., Herring, S., Berube, A., Wood, J.L., Soll, D., Ambrogelly, A., 2006. Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase. FEBS Lett. 580, 6695-6700.

Protzel, A., Morris, A.J., 1974. Gel chromatographic analysis of nascent globin chains. Evidence of nonuniform size distribution. J. Biol. Chem. 249, 4594-4600.

Qi, F., Motz, M., Jung, K., Lassak, J., Frishman, D., 2018. Evolutionary analysis of polyproline motifs in *Escherichia coli* reveals their regulatory role in translation. PLoS Comput. Biol. 14, e1005987.

References

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., Darwin, A.J., Ibba, M., 2015. Cyclic rhamnosylated elongation factor P establishes antibiotic resistance in *Pseudomonas aeruginosa*. mBio 6, e00823.

Rajkovic, A., Hummels, K.R., Witzky, A., Erickson, S., Gafken, P.R., Whitelegge, J.P., Faull, K.F., Kearns, D.B., Ibba, M., 2016. Translation control of swarming proficiency in *Bacillus subtilis* by 5-amino-pentanolylated elongation factor P. J. Biol. Chem. 291, 10976-10985.

Randall, L.L., Josefsson, L.G., Hardy, S.J., 1980. Novel intermediates in the synthesis of maltose-binding protein in *Escherichia coli*. Eur. J. Biochem. 107, 375-379.

Reichert, A.J., Poxleitner, G., Dauner, M., Skerra, A., 2015. Optimisation of a system for the co-translational incorporation of a keto amino acid and its application to a tumour-specific Anticalin. Protein Eng. Des. Sel. 28, 553-565.

Ren, Q., Chen, K., Paulsen, I.T., 2006. TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. Nucleic Acids Res. 35, D274-D279.

Ringel, A.E., Roman, C., Wolberger, C., 2014. Alternate deacylating specificities of the archaeal sirtuins Sir2Af1 and Sir2Af2. Protein Sci. 23, 1686-1697.

Roy, H., Zou, S.B., Bullwinkle, T.J., Wolfe, B.S., Gilreath, M.S., Forsyth, C.J., Navarre, W.W., Ibba, M., 2011. The tRNA synthetase paralog PoxA modifies elongation factor-P with (*R*)-beta-lysine. Nat. Chem. Biol. 7, 667-669.

Santoro, S.W., Wang, L., Herberich, B., King, D.S., Schultz, P.G., 2002. An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. Nat. Biotechnol. 20, 1044-1048.

Sassetti, C.M., Boyd, D.H., Rubin, E.J., 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol. Microbiol. 48, 77-84.

Schmidt, M.J., Weber, A., Pott, M., Welte, W., Summerer, D., 2014. Structural basis of furan-amino acid recognition by a polyspecific aminoacyl-tRNA-synthetase and its genetic encoding in human cells. Chembiochem 15, 1755-1760.

Schneider, S., Gattner, M.J., Vrabel, M., Flugel, V., Lopez-Carrillo, V., Prill, S., Carell, T., 2013. Structural insights into incorporation of norbornene amino acids for click modification of proteins. Chembiochem 14, 2114-2118.

Schuller, A.P., Wu, C.C., Dever, T.E., Buskirk, A.R., Green, R., 2017. eIF5A functions globally in translation elongation and termination. Mol. Cell 66, 194-205.

Sengoku, T., Suzuki, T., Dohmae, N., Watanabe, C., Honma, T., Hikida, Y., Yamaguchi, Y., Takahashi, H., Yokoyama, S., Yanagisawa, T., 2018. Structural basis of protein arginine rhamnosylation by glycosyltransferase EarP. Nat. Chem. Biol. 14, 368-374.

Sinclair, A.M., Elliott, S., 2005. Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins. J. Pharm. Sci. 94, 1626-1635.

References

Starosta, A.L., Lassak, J., Peil, L., Atkinson, G.C., Virumae, K., Tenson, T., Remme, J., Jung, K., 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.

Sumida, T., Yanagisawa, T., Ishii, R., Yokoyama, S., 2010. Crystallization and preliminary X-ray crystallographic study of GenX, a lysyl-tRNA synthetase paralogue from *Escherichia coli,* in complex with translation elongation factor P. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 66, 1115-1118.

Takimoto, J.K., Dellas, N., Noel, J.P., Wang, L., 2011. Stereochemical basis for engineered pyrrolysyl-tRNA synthetase and the efficient *in vivo* incorporation of structurally divergent non-native amino acids. ACS Chem. Biol. 6, 733-743.

Tanner, D.R., Cariello, D.A., Woolstenhulme, C.J., Broadbent, M.A., Buskirk, A.R., 2009. Genetic identification of nascent peptides that induce ribosome stalling. J. Biol. Chem. 284, 34809-34818.

Tetsch, L., Koller, C., Haneburger, I., Jung, K., 2008. The membrane-integrated transcriptional activator CadC of *Escherichia coli* senses lysine indirectly via the interaction with the lysine permease LysP. Mol. Microbiol. 67, 570-583.

Tollerson, R., 2nd, Witzky, A., 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.

Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N., Jung, K., 2013. Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. Science 339, 82-85.

Umehara, T., Kim, J., Lee, S., Guo, L.T., Soll, D., Park, H.S., 2012. *N*-acetyl lysyl-tRNA synthetases evolved by a CcdB-based selection possess *N*-acetyl lysine specificity *in vitro* and *in vivo*. FEBS Lett. 586, 729-733.

Varenne, S., Buc, J., Lloubes, R., Lazdunski, C., 1984. Translation is a non-uniform process - effect of transfer-RNA availability on the rate of elongation of nascent polypeptide-chains. J. Mol. Biol. 180, 549-576.

Varenne, S., Cavard, D., Lazdunski, C., 1981. Biosynthesis and export of colicin A in *Citrobacter freundii* CA31. Eur. J. Biochem. 116, 615-620.

Varenne, S., Knibiehler, M., Cavard, D., Morlon, J., Lazdunski, C., 1982. Variable rate of polypeptide chain elongation for colicins A, E2 and E3. J. Mol. Biol. 159, 57-70.

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., Hennig, J., Jung, K., Lassak, J., 2019. Switching the post-translational modification of translation elongation factor EF-P. Front. Microbiol. 10.

Volkwein, W., Maier, C., Krafczyk, R., Jung, K., Lassak, J., 2017. A versatile toolbox for the control of protein levels using N^ε-acetyl-L-lysine dependent amber suppression. ACS Synth. Biol. 6, 1892-1902.

Wan, W., Tharp, J.M., Liu, W.R., 2014. Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. Biochim. Biophys. Acta 1844, 1059-1070.

Wang, L., Brock, A., Herberich, B., Schultz, P.G., 2001. Expanding the genetic code of *Escherichia coli*. Science 292, 498-500.

Wang, L., Schultz, P.G., 2001. A general approach for the generation of orthogonal tRNAs. Chemistry & biology 8, 883-890.

Witzky, A., Hummels, K.R., Tollerson, R., 2nd, Rajkovic, A., Jones, L.A., Kearns, D.B., Ibba, M., 2018. EF-P posttranslational modification has variable impact on polyproline translation in *Bacillus subtilis*. mBio 9.

Wohlgemuth, I., Brenner, S., Beringer, M., 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.

Woolstenhulme, C.J., Guydosh, N.R., Green, R., Buskirk, A.R., 2015. High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EF-P. Cell Rep. 11, 13-21.

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., 2013. Nascent peptides that block protein synthesis in bacteria. Proc. Natl. Acad. Sci. U. S. A. 110, E878-887.

Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., Yokoyama, S., 2008a. Crystallographic studies on multiple conformational states of active-site loops in pyrrolysyl-tRNA synthetase. J. Mol. Biol. 378, 634-652.

Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., Yokoyama, S., 2008b. Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N^{ε}-(*o*-azidobenzyloxycarbonyl) lysine for site-specific protein modification. Chemistry & biology 15, 1187-1197.

Yanagisawa, T., Kuratani, M., Seki, E., Hino, N., Sakamoto, K., Yokoyama, S., 2019. Structural basis for genetic-code expansion with bulky lysine derivatives by an engineered pyrrolysyl-tRNA synthetase. Cell Chem. Biol. 26, 936-949.

Yanagisawa, T., Sumida, T., Ishii, R., Takemoto, C., 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.

Yanagisawa, T., Sumida, T., Ishii, R., Yokoyama, S., 2013. A novel crystal form of pyrrolysyl-tRNA synthetase reveals the pre- and post-aminoacyl-tRNA synthesis conformational states of the adenylate and aminoacyl moieties and an asparagine residue in the catalytic site. Acta Crystallogr. D Biol. Crystallogr. 69, 5-15.

Yanagisawa, T., Takahashi, H., Suzuki, T., Masuda, A., Dohmae, N., Yokoyama, S., 2016. *Neisseria meningitidis* translation elongation factor P and its active-site arginine residue are essential for cell viability. PLoS One 11, e0147907.

References

Yang, X.J., Seto, E., 2008. The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. Nat. Rev. Mol. Cell. Biol. 9, 206-218.

Yanofsky, C., 1981. Attenuation in the control of expression of bacterial operons. Nature 289, 751-758.

Zheng, Y., Liu, Q., Shen, H., Yang, G., 2018. To increase the incorporation efficiency of genetically encoding N^ε-acetyllysine in recombinant protein. Protein Expr. Purif. 145, 59-63.

Zou, S.B., Hersch, S.J., Roy, H., Wiggers, J.B., Leung, A.S., Buranyi, S., Xie, J.L., Dare, K., Ibba, M., Navarre, W.W., 2012. Loss of elongation factor P disrupts bacterial outer membrane integrity. J. Bacteriol. 194, 413-425.

Zou, S.B., Roy, H., Ibba, M., Navarre, W.W., 2011. Elongation factor P mediates a novel post-transcriptional regulatory pathway critical for bacterial virulence. Virulence 2, 147-151.