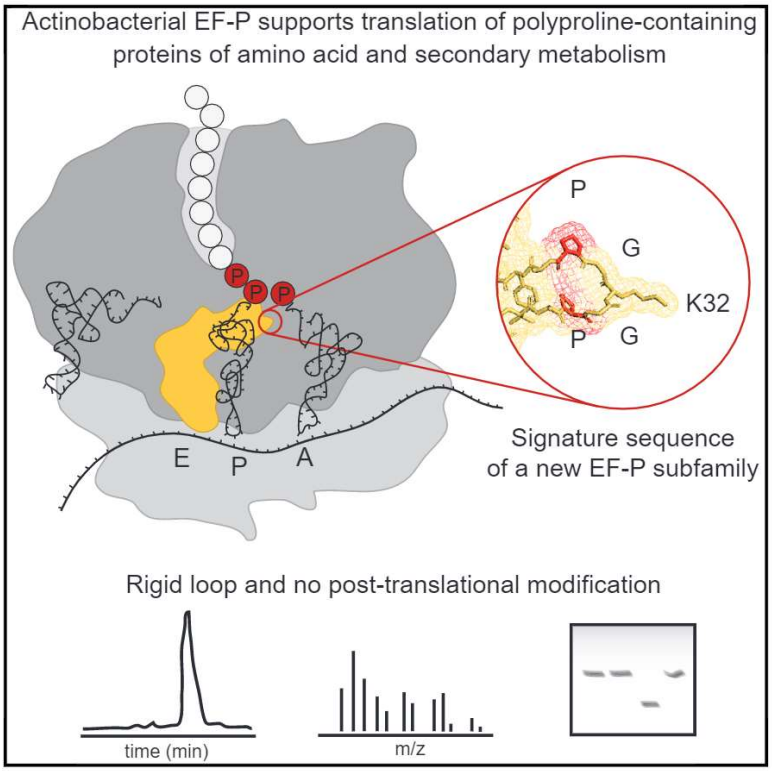
Article

Cell Reports

Structure and Function of an Elongation Factor P Subfamily in Actinobacteria

Graphical Abstract



Authors

Bruno Pinheiro, Christopher M. Scheidler, Pavel Kielkowski, ..., Stephan A. Sieber, Sabine Schneider, Kirsten Jung

Correspondence

sabine.schneider@cup.lmu.de (S.S.), jung@lmu.de (K.J.)

In Brief

Pinheiro et al. report on a subfamily of EF-P required for the translation of a high number of polyproline-containing proteins in Actinobacteria. While the function and 3D structure of these EF-Ps are conserved, they are active without post-translational modification and have the palindromic loop signature sequence PGKGP.

Highlights

- Unmodified EF-P is active in *Corynebacterium*, *Streptomyces*, and *Mycobacterium*
- The crystal structure of *C. glutamicum* EF-P was solved with 2.2-Å resolution
- The essential lysine32 is embedded in a rigid palindromic loop
- EF-P-dependent proteins are involved in amino acid and secondary metabolism

Pinheiro et al., 2020, Cell Reports *30*, 4332–4342
March 31, 2020
https://doi.org/10.1016/j.celrep.2020.03.009

CellPress

OPEN ACCESS
CellPress

Cell Reports
Article

Structure and Function of an Elongation Factor P Subfamily in Actinobacteria

Bruno Pinheiro,[1] Christopher M. Scheidler,[2] Pavel Kielkowski,[3] Marina Schmid,[1] Ignasi Forné,[4] Suhui Ye,[5,8] Norbert Reiling,[6,7] Eriko Takano,[5] Axel Imhof,[4] Stephan A. Sieber,[3] Sabine Schneider,[2,*] and Kirsten Jung[1,9,*]

[1]Department of Biology I, Microbiology, Ludwig-Maximilians-Universität München, Martinsried, Germany
[2]Department of Chemistry, Ludwig-Maximilians-Universität München, Munich, Germany
[3]Department of Chemistry, Technische Universität München, Garching, Germany
[4]Biomedical Center Munich, Ludwig-Maximilians-Universität München, Martinsried, Germany
[5]Manchester Institute of Biotechnology, Department of Chemistry, School of Natural Sciences, Faculty of Science and Engineering, University of Manchester, Manchester, UK
[6]RG Microbial Interface Biology, Research Center Borstel, Leibniz Lung Center, Borstel, Germany
[7]German Center for Infection Research (DZIF), Partner Site Hamburg-Lübeck-Borstel-Riems, Borstel, Germany
[8]Present address: Research Group BIONUC (Biotechnology of Nutraceuticals and Bioactive Compounds), Departamento de Biología Funcional, Área de Microbiología, Universidad de Oviedo, Avenida Julián Clavería s/n, 33006, Oviedo, Principality of Asturias, Spain
[9]Lead Contact
*Correspondence: sabine.schneider@cup.lmu.de (S.S.), jung@lmu.de (K.J.)
https://doi.org/10.1016/j.celrep.2020.03.009

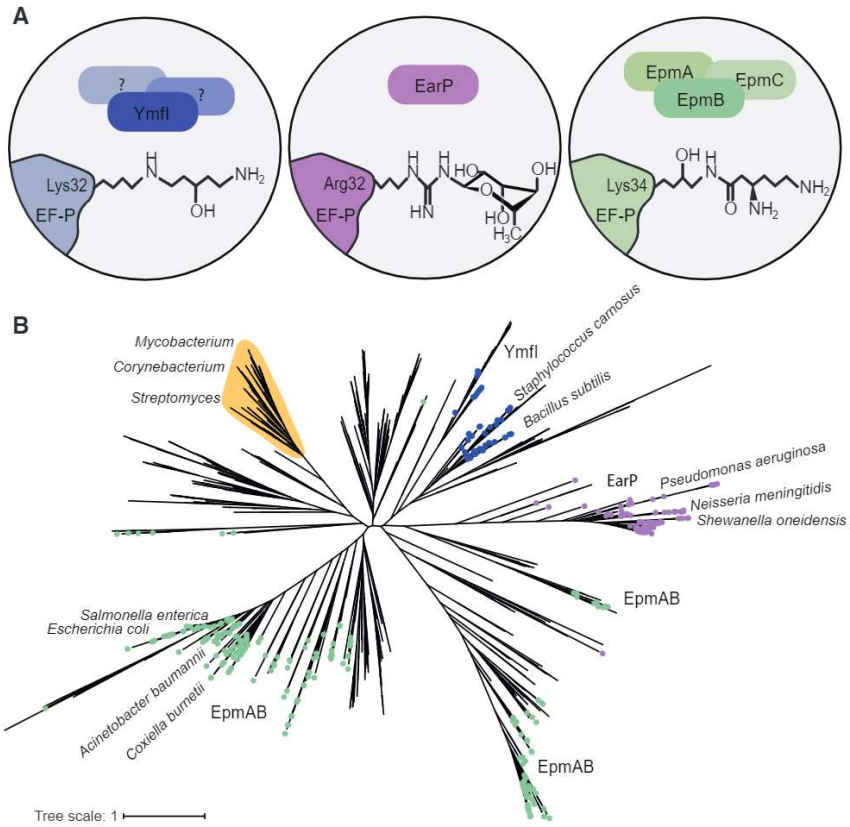
SUMMARY

Translation of consecutive proline motifs causes ribosome stalling and requires rescue via the action of a specific translation elongation factor, EF-P in bacteria and archaeal/eukaryotic a/eIF5A. In Eukarya, Archaea, and all bacteria investigated so far, the functionality of this translation elongation factor depends on specific and rather unusual post-translational modifications. The phylum Actinobacteria, which includes the genera *Corynebacterium*, *Mycobacterium*, and *Streptomyces*, is of both medical and economic significance. Here, we report that EF-P is required in these bacteria in particular for the translation of proteins involved in amino acid and secondary metabolite production. Notably, EF-P of Actinobacteria species does not need any post-translational modification for activation. While the function and overall 3D structure of this EF-P type is conserved, the loop containing the conserved lysine is flanked by two essential prolines that rigidify it. Actinobacteria's EF-P represents a unique subfamily that works without any modification.

INTRODUCTION

The incorporation of proline induces sharp turns and perturbs secondary structure in proteins, due to the unique conformational rigidity of the pyrrolidine ring (Doerfel et al., 2015; Vanhoof et al., 1995; Venkatachalam and Ramachandran, 1969). Proline-rich regions, including motifs containing consecutive prolines (XPPX), are often found in sites of protein-protein and protein-nucleic acid interaction (Adzhubei et al., 2013; Qi et al., 2018; Vanhoof et al., 1995). XPPX motifs are important in enzymes but might also play a regulatory role in copy number adjustment, as well as in protein folding and membrane insertion (Motz and Jung, 2018; Qi et al., 2018; Starosta et al., 2014b; Ude et al., 2013). The incorporation of polyproline motifs into proteins is challenging for ribosomes, and often leads to peptidyl-tRNA destabilization and detachment of the nascent chain, culminating in translational stalling and decreased protein levels (Doerfel et al., 2013; Huter et al., 2017; Peil et al., 2013; Ude et al., 2013). To overcome this translational burden, bacteria have evolved elongation factor P (EF-P) (Doerfel et al., 2013; Ude et al., 2013). Evidence that EF-P facilitates peptide bond formation when the ribosome encounters consecutive prolines is based on structural studies (Blaha et al., 2009; Huter et al., 2017), various *in vivo* translation-based reporter systems (Rajkovic et al., 2016; Ude et al., 2013), biochemical kinetic experiments (Doerfel et al., 2013), and ribosome profiling, which delineated a hierarchy of pausing motifs (Elgamal et al., 2014; Woolstenhulme et al., 2015). This protein, like its eukaryotic and archaeal orthologs eIF5A/aIF5A, binds to the ribosome between the P- and E-sites, projecting a conserved residue toward the ribosomal peptidyl-transferase center (PTC) and stimulating formation of the Pro-Pro bond (Doerfel et al., 2013; Gutierrez et al., 2013; Melnikov et al., 2016; Ude et al., 2013; Vanhoof et al., 1995). Binding of EF-P to the ribosome is required to stabilize the P-site tRNA, thereby enforcing a conformation of the polyproline-containing nascent chain that provides a favorable substrate geometry for peptide-bond formation (Huter et al., 2017).

In all organisms tested thus far, EF-P and its eIF5A/aIF5A orthologs require for their activation a post-translational modification (PTM) at a conserved amino acid residue located at the tip of a loop (Cooper et al., 1983; Navarre et al., 2010; Park et al., 1981; Peil et al., 2012; Yanagisawa et al., 2010). This modification extends the side chain of a conserved lysine or arginine (Huter et al., 2017). Whereas eukaryotes and archaea evolved a unique PTM, the hypusinyl side chain, to activate eIF5A/aIF5A (Cooper et al., 1983; Park et al., 1981), bacteria use diverse and unusual substrates and enzymes to activate their EF-Ps (Figure 1A).



C

Actinobacteria strain	Proteome ID	Proteins encoded	Number of XPPX motifs			Frequency Motifs/Proteins
			Weak	Moderate	Strong	
Corynebacterium glutamicum ATCC13032	UP000000582	3093	751	347	380	0.48
Corynebacterium diphtheriae NCTC13129	UP000002198	2265	595	298	303	0.53
Mycobacterium smegmatis MC2 155	UP000000757	6602	2491	1292	2555	0.96
Mycobacterium leprae TN	UP000000806	1603	622	298	556	0.92
Mycobacterium tuberculosis H37Rv	UP000001584	3993	1826	996	1854	1.17
Streptomyces coelicolor A3(2)	UP000001973	8038	3508	1807	3406	1.08
Bifidobacterium bifidum S17	UP000006869	1783	412	218	205	0.47
Bifidobacterium longum NCC2705	UP000000439	1725	382	225	221	0.48
Bifidobacterium animalis subsp. lactis AD011	UP000002456	1525	345	186	166	0.46
Gardnerella vaginalis ATCC14019	UP000001453	1365	231	126	99	0.33
Mobiluncus curtisii ATCC43063	UP000006742	1904	636	348	381	0.72
Nocardia farcinica IFM10152	UP000006820	5941	2583	1365	2333	1.06

Figure 1. Phylogenetic Tree of the Bacterial EF-Ps and the Known PTMs

(A) Known bacterial EF-P subfamilies and their various post-translational modifications (PTMs).

(B) Phylogenic tree of bacterial EF-Ps. Dots highlight the distribution of enzymes required for the modification of EF-Ps: note the co-occurrence of EpmA and EpmB (green), EarP (purple), and YmfI (blue). The Actinobacteria EF-P subfamily is marked in yellow. Bacterial species in which some features of EF-P function were previously studied are indicated.

(C) Total number and frequency of polyproline motifs (XPPX) classified according to their stalling strengths in different species of Actinobacteria.

A phylogenetic tree of EF-P was constructed, and the various types of EF-P modifications were inferred from the PTM enzymes encoded by each organism (Figure 1B; Table S1). γ-Proteobacteria, including *Escherichia coli*, *Salmonella enterica*, and 29% of all other reference genomes encode the modification enzymes EF-P-(*R*)- β-lysine ligase EpmA (Roy et al., 2011; Yanagisawa et al., 2010) and L-lysine 2,3-aminomutase EpmB (Yanagisawa et al., 2010) to β-lysinylate EF-P. Only 13% of these bacteria co-encode the EF-P hydroxylase EpmC responsible for the last step in this modification. However, the hydroxyl group added by this enzyme was shown to have a negligible effect on EF-P activity in *E. coli* (Peil et al., 2012, 2013). Another EF-P subfamily is found in β-proteobacteria and some γ-proteobacteria (9% of all bacteria), in which lysine is replaced by arginine and the latter undergoes rhamnosylation catalyzed by the arginine rhamnosyltransferase EarP (Lassak et al., 2015). *Bacillus subtilis*, *Staphylococcus carnosus*, and other members of the Firmicutes encode the EF-P modification enzyme YmfI (4.5% of all bacteria). This enzyme catalyzes the reduction of EF-P-5-aminopentanone to EF-P-5-aminopentanol, the last step in the modification of EF-P in *B. subtilis* (Hummels et al., 2017). The complete pathway of this modification is still unclear. Thus far, the cognate PTMs were found to be essential for EF-P activity, and deletion mutants lacking one of the modification enzymes showed the same phenotypes as *efp* mutants, namely downregulation of proteins containing XPPX motifs (Peil et al., 2013; Qi et al., 2018; Starosta et al., 2014a; Witzky et al., 2018; Woolstenhulme et al., 2015), decreased fitness (Tollerson et al., 2018; Ude et al., 2013), decreased pathogenicity (Klee et al., 2018; Lassak et al., 2015; Navarre et al., 2010), or cell death (Yanagisawa et al., 2016).

The remaining 57% of bacterial genomes do not encode any known EF-P modification enzyme. Among them, the cluster formed by Actinobacteria EF-P represents 11% of all bacterial genomes (Figure 1B). This cluster comprises EF-P sequences with an average sequence identity of 68.1%, which implies that they share the same EF-P activation pathway.

Actinobacteria constitutes one of the largest phyla among bacteria and comprises Gram-positive bacteria with a high G+C DNA content. The divergence of Actinobacteria from other bacteria is ancient, making it impossible to identify the phylogenetically closest bacterial group. Furthermore, members of the Actinobacteria have adopted different lifestyles. Some are pathogens (e.g., *Corynebacterium, Mycobacterium*), others soil inhabitants (*Streptomyces*), plant commensals (*Leifsonia*), or gastrointestinal commensals (*Bifidobacterium*) (Ventura et al., 2007). *Mycobacterium*, *Streptomyces*, and *Corynebacterium*

are of medical and industrial importance. Here, we examine various species of Actinobacteria, specifically *C. glutamicum*, *S. coelicolor*, and *M. tuberculosis*, and report a new EF-P subfamily that differs from all others in that it alleviates ribosome stalling at polyproline motifs without needing any activating PTM.

RESULTS

Bioinformatics and Functional Importance of EF-P in Actinobacteria

EF-P is not only important for the translation of polyproline-containing proteins, but also for synthesis of specific subsets of proteins containing diprolyl motifs (X/PP/X) (Peil et al., 2013; Qi et al., 2018). There is a distinct hierarchy of stalling motifs, ranging from strong stallers, such as PPP, D/PP/D, PPW, APP, G/PP/G, and PPN, to weak stallers, such as CPP, L/PP/L, and HPP (Elgamal et al., 2014; Hersch et al., 2013; Peil et al., 2013; Starosta et al., 2014a; Woolstenhulme et al., 2015).

Strength and number of XPPX motifs in a proteome indicate how important EF-P activity is for the individual bacterial species. Previous studies have investigated the PTM status of EF-P primarily in three species: *E. coli*, *S. oneidensis*, and *B. subtilis*. Of these, *E. coli* is the one with most XPPX motifs. In total, 2,101 XPPX motifs are present in its proteome, corresponding to 0.49 per protein encoded (Qi et al., 2018). *S. oneidensis* and *B. subtilis* encode 0.38 and 0.28 motifs per protein, respectively (Table S2). We determined the numbers of XPPX motifs in the most relevant Actinobacteria and identified strikingly large numbers of them (Figure 1C). In *M. tuberculosis*, *S. coelicolor*, and *Nocardia farcinica*, the number of polyproline motifs actually exceeds the number of encoded proteins (Figure 1C).

Besides the control by EF-P, the relative amount of an XPPX-containing protein synthesized depends on the rate of translational initiation and the location of the stalling motif(s) it contains (Ude et al., 2013; Woolstenhulme et al., 2015). In addition, transcriptional regulators may contain XPPX motifs, so that levels of proteins that lack stalling motifs could still be indirectly controlled by EF-P. Therefore, to assess the global effect of EF-P on the proteome of a well-investigated member of Actinobacteria, we constructed a Δ*efp* mutant of *C. glutamicum* ATCC13032 and compared its proteome to that of the parental strain. The analysis of four independent replicates covered 1,604 out of 3,093 proteins described for the reference strain. In all, 222 proteins were downregulated (p value < 0.05; Table S3). As expected, the most markedly downregulated proteins contain XPPX motifs in their sequences (Figure 2A). The overall intensity of proteins containing polyproline motifs was reduced as well (Figure 2B, in blue), and proteins containing strong stalling XPPX motifs have the lowest relative intensity among them (indicated in yellow in Figure 2B).

Functional classification of the downregulated proteins revealed seven groups of proteins that were significantly overrepresented (Figure 2C). Among them are enzymes associated with the biosynthesis of amino acids, antibiotics, and other secondary metabolites. This underlines the role of EF-P in the production of these compounds in Actinobacteria.

The evolution of the EF-P and eIF5A/aIF5A ribosome rescue system is linked to an invariant proline triplet motif located in the active site of Valine-tRNA synthetase (ValS), which is found in all domains of life (Starosta et al., 2014b). In *C. glutamicum*, as well as in *E. coli*, ValS is one of the proteins most drastically downregulated when *efp* is deleted. Cgl1117 is a predicted glycosyltransferase that is homologous to the *E. coli* glycogen synthase GlgA and includes two strong XPPX motifs. It was found to be the most strongly downregulated protein in our proteomic analysis with a 15.4-fold reduction in its steady-state level relative to the *efp*$^+$ strain (Figure 2A; Table S3). This result of the mass spectrometry (MS) analysis was confirmed by testing the level of the chromosomally encoded Cgl1117-eGFP fusion protein in the *efp* mutant (Figure 2D). Overall, the impact in the downregulation of proteins containing XPPX motifs observed here is fully compatible with previous proteomic analyses of *E. coli* and other bacterial species lacking *efp* or one of its post-translational modification enzymes (Hersch et al., 2013; Peil et al., 2013).

EF-P Activity Depends on a Positively Charged Amino Acid at Position 32

The effect of EF-P on the *C. glutamicum* proteome led us to investigate the nature of the post-translational modification that might activate this elongation factor. We generated two reporter genes coding for the enhanced green fluorescent protein (eGFP), one with the strong stalling motif RPPP and another with the non-stalling sequence RPAP upstream of the eGFP sequence. Both genes were integrated chromosomally under the control of a constitutive promoter, and their expression was tested in *efp*$^+$ and *efp*$^-$ *C. glutamicum* cells. While the RPAP-eGFP variant was synthesized efficiently in both strains, RPPP-eGFP was barely detectable in the *efp*$^-$ mutant, revealing its strong dependence on EF-P, and the fluorescence intensity of these cells was only slightly higher than the background level (Figures 3A and 3B).

EF-P in *C. glutamicum* has a lysine (K32) at the tip of the loop, which in other organisms is known to undergo post-translational modification (see above). In order to determine whether K32 is important for EF-P activity, we constructed several EF-P variants to alter the properties of the amino acid at this position. Strains expressing various substitutions at this position (K32A, K32E, K32M, K32Q) were virtually unable to support RPPP-eGFP translation. Surprisingly, the EF-P^{K32R} variant enhanced RPPP-eGFP production to almost the same degree as the wild type (Figure 3C). Replacement of the lysine by arginine preserves the positively charged side chain of the amino acid, but nevertheless abolishes post-translational modification. While the EF-P^{K32R} variant is completely inactive in *E. coli* (Figure S1A), the unmodified EF-P^{K32R} variant retains substantial activity (e.g., swarming phenotype) in *Bacillus subtilis* (Hummels et al., 2017).

EF-P in *C. glutamicum* Is Not Post-translationally Modified

The high activity of the EF-P^{K32R} variant prompted us to ask whether *C. glutamicum* EF-P (EF-P$_{Cg}$) actually requires post-translational modification for activation. To facilitate purification

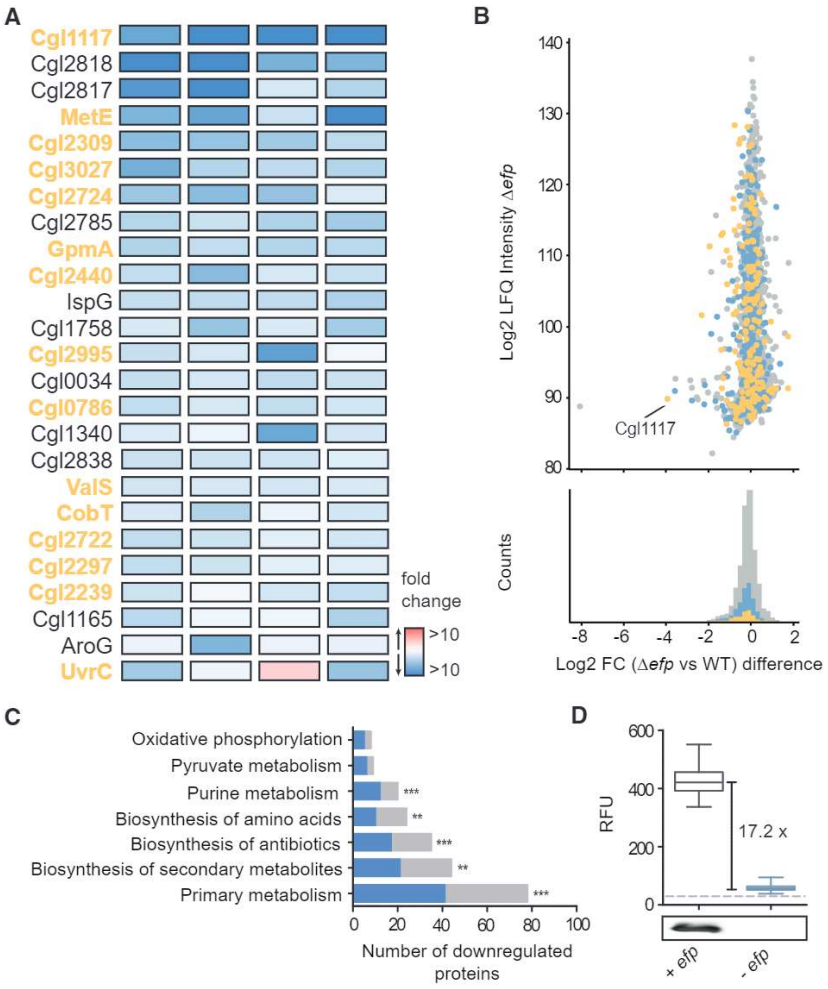


Figure 2. Proteomic Analysis of *C. glutamicum* Δ*efp* Strain

(A) Heatmap representation of the 25 most severely downregulated proteins in the Δ*efp* mutant relative to the *C. glutamicum* wild type. The fold changes in the four independent replicates are represented by the color gradient in red (upregulation) and blue (downregulation). Proteins marked in yellow contain polyproline motifs. EF-P itself as the most markedly downregulated protein is not shown on the heatmap but is highlighted on Table S3.

(B) Scatter-plot of protein fold-change ratio (log_2-transformed) relative to the summed protein intensity of all biological replicates of the Δ*efp* strain, including density plots showing the distribution of proteins with strong stalling XPPX motifs (yellow) relative to all XPPX-containing proteins (blue) and all proteins (gray).

(C) Clustering of downregulated proteins according to the Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Database. The blue segments represent the numbers of proteins containing XPPX motifs. Statistical significance was addressed with Fisher exact test, non-marked columns $p < 0.05$; **$p < 0.01$; ***$p < 0.0001$.

(D) Distribution of relative fluorescence signals from 300 *C. glutamicum* cgl1117-*egfp* cells in *efp*$^+$ and *efp*$^-$ strains. The dashed lines mark the background fluorescence. Production of EF-P was confirmed by western blot analysis.

of EF-P_{Cg}, we expressed a C-terminally 6×His-tagged EF-P variant (connected to the protein by a short Arg-Ser linker) from the endogenous locus under the control of the native promoter. The tag had no effect on EF-P activity (Figure S1B). This strategy allowed us to express and isolate native levels of EF-P_{Cg}, thus circumventing possible problems caused by low activity of PTM enzymes or substrate limitation triggered by protein overexpression.

A post-translational modification alters the mass and frequently the charge of a given protein (Deribe et al., 2010; Lakemeyer et al., 2019). To investigate the nature of PTM of *C. glutamicum*'s EF-P, we purified the His-tagged variant. In addition, we heterologously expressed this same gene in the Δ*efp* Δ*epmA E. coli* mutant, expecting that the resulting product would not be modified in this bacterium. Both proteins were then analyzed by mass spectrometry. The intact protein mass matched the calculated mass, irrespective of whether EF-P_{Cg} was isolated from *C. glutamicum* (Figure 4A, from mid-log growth phase; see Figure S2A for stationary growth phase) or from the transformed *E. coli* mutant (Figure S2). Liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis of the trypsin-digested endogenous EF-P confirmed that none of the peptides was modified (Figure 4B). Moreover, there was no difference in isoelectric point between endogenous and heterologous EF-P_{Cg} (Figure 4C). We therefore conclude that neither the endogenous nor the recombinantly produced protein undergoes PTM.

Subsequently, EF-P variants in which the conserved lysine at position 32 was replaced by alanine (K32A) or arginine (K32R) were also analyzed by mass spectrometry and isoelectric focusing (Figures 4C and S2). These variants exhibited only the expected shifts in mass and isoelectric point (IEP).

Crystal Structure of Actinobacteria EF-P

In order to gain insights into the structural configuration of this Actinobacteria EF-P, we isolated and crystallized the endogenous protein from *C. glutamicum* ATCC13032 and determined its X-ray crystal structure to 2.2-Å resolution (for data processing and structure refinement statistics, see Table S4). EF-P_{Cg} shares its overall folding topology with the previously reported bacterial EF-P structures and consists of three β-barrel domains (domains I, II, and III) with an overall L shape reminiscent of a tRNA (Hanawa-Suetsugu et al., 2004). The EF-P structure can be superimposed on homologous EF-P structures with a root-mean-square deviation (RMSD) ranging from 1.9 to 3 Å (Figure 5A). The loop connecting β2 and β3 in the N-terminal domain I, which in other EF-P proteins carries the post-translational modification on either a lysine or arginine residue, is fully defined in the electron-density map. The side chain of

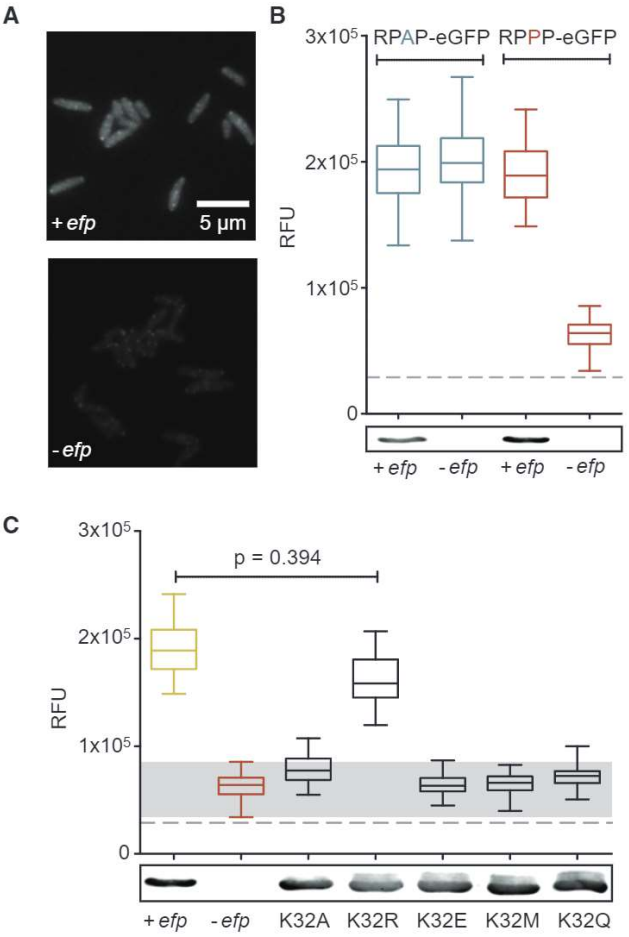


Figure 3. EF-P Enhances the Translation of Proteins Containing Polyproline Motifs

(A) Single-cell fluorescence microscopy of *C. glutamicum* reporters producing RPPP-eGFP in *efp*$^+$ and *efp*$^-$ strains. Pictures were taken with a 500-ms exposure time.

(B) Distribution of relative eGFP fluorescence signals obtained from a minimum of 300 *C. glutamicum* cells expressing RPAP-eGFP and RPPP-eGFP, respectively, in *efp*$^+$ and *efp*$^-$ strains. The dashed line marks the background level of fluorescence. Western blots reveal levels of expression of *C. glutamicum* EF-P.

(C) Synthesis of RPPP-eGFP in *C. glutamicum* strains expressing EF-P variants with the indicated amino acid replacements at position 32. Relative fluorescence units (RFUs) are shown. The dashed line corresponds to background fluorescence. The area shown in gray marks the range of fluorescence measured in the negative control (−*efp*). Strain expressing C. glutamicum EF-P (+ efp) and negative control as shown in (B).

Lys32 is not discernible and must therefore be flexible (Figure 5B).

Crystal and cryo-EM structures of the *T. thermophilus* and *E. coli* EF-Ps in complex with the ribosome show that EF-P binds at the ribosomal E-site, with its domain I located next to the acceptor stem of the P-site tRNA. The loop connecting β2 to β3 interacts with the acceptor arm of P-site tRNA. Domain III of EF-P binds adjacent to the anticodon stem-loop of the P-site tRNA, while domain II of EF-P interacts with the highly conserved ribosomal protein L1 (Blaha et al., 2009; Huter et al., 2017). Thus, EF-P most likely acts indirectly on peptide-bond formation either by stabilizing the initiator tRNA or by enforcing an alternative conformation of the nascent polyproline-containing chain to provide a favorable substrate geometry (Blaha et al., 2009; Huter et al., 2017). In the *E. coli* EF-P-ribosome complex, the modified Lys34, an ε(R)-β-lysylhydroxylysinyl residue, directly interacts with the CCA end of the P-site tRNA (Huter et al., 2017). Alignment of 150 different bacterial EF-P protein sequences reveals strong conservation of domains I and III, which interact with the tRNA and rRNA. Overall, domain III appears to be less conserved. However, the region that makes contact with the ribosomal L1 protein in the *E. coli* EF-P-ribosome complex also shows sequence conservation (Figure 5).

P30 and P34 Are Essential for EF-P Activity in Actinobacteria

Although the sequence of domain I in EF-P is highly conserved across the bacterial kingdom, a striking difference, a loop with the palindromic PGKGP motif, was identified in Actinobacteria EF-Ps. P30 and P34 flank this palindromic EF-P loop sequence that is found in 11% of all bacteria including all EF-P sequences from *Corynebacterium*, *Streptomyces*, *Mycobacterium*, *Bifidobacterium*, *Gardnerella*, *Mobiluncus*, and *Norcardia* (Figures 6B, S4, and S5).

P30 was previously identified as an invariant proline in all EF-Ps with a lysine at the tip of the loop in a phylogenetic analysis, and other amino acids replace proline when arginine is at the tip (Volkwein et al., 2019) (Figure S3E). The exchange of P30 for alanine or glutamine in EF-P of *C. glutamicum* led to non-functional variants (Figure 6A).

The second proline of the palindromic motif is only present in the subgroup of Actinobacteria described here. Glutamine and alanine are frequently found at this position in EF-P proteins in bacterial species encoding genes of known EF-P modification enzymes (Figure S3E). Some actinobacterial EF-P sequences have glycine or asparagine at position 34 (Figures S3D and S4). To determine the significance of P34 for *C. glutamicum* EF-P, we replaced this amino acid with alanine (*B. subtilis* EF-P has A34), glutamine (*E. coli* EF-P has Q34), glycine, and asparagine, and quantified the EF-P activity of the resulting variants (Figure 6A). All resulting EF-P variants were not able to support RPPP-eGFP translation in *C. glutamicum* (Figure 6A).

The two prolines in the consensus sequence most likely rigidify the loop, which in turn could enable these EF-Ps to stabilize the acceptor arm of the tRNA and thus allow translation of polyproline motifs without post-translational modification. Apart from this motif, no other sequence signature was identified in our structural and amino acid sequence analyses of unmodified EF-Ps (Figures 5 and S4). The palindromic loop sequence PGKGP is an essential feature of the new EF-P subfamily described here.

A Novel Subfamily of EF-P in Actinobacteria

EF-P protein sequences are highly conserved among the most relevant Actinobacteria (Figure 6B). We used the same

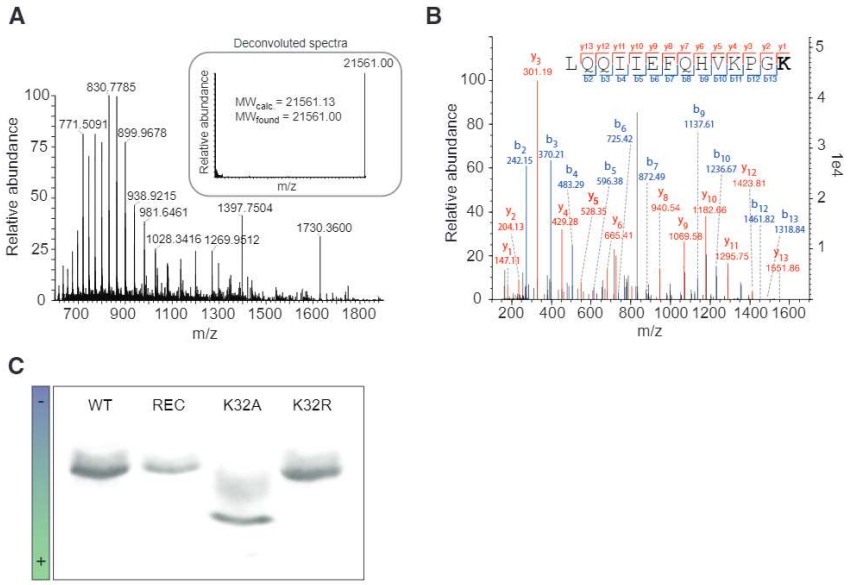


Figure 4. EF-P Is Not Post-translationally Modified in *C. glutamicum*

(A) Mass spectra of the intact endogenous EF-P*Cg* protein purified from cells in mid-log growth phase reveal no additional mass, and therefore no PTM of the protein.

(B) LC-MS/MS analysis of the trypsin-digested endogenous EF-P indicates that the peptide containing the conserved lysine 32 (bold) is unmodified (y- and b-fragments resulting from MS/MS sequencing are colored in red and blue, respectively).

(C) Western blot of an isoelectric focusing gel loaded with *C. glutamicum* EF-P isolated from its native host (WT), after heterologous expression in *E. coli* Δ*efp* Δ*epmA* (REC), and after replacement of lysine 32 by alanine (K32A) or arginine (K32R).

reporter setup described in Figure 3 to test whether EF-P proteins from other members of the Actinobacteria are able to complement the Δ*efp C. glutamicum* mutant. EF-Ps from *Mycobacterium smegmatis*, *M. phlei*, *M. tuberculosis*, *Streptomyces coelicolor*, *S. venezuelae*, and *S. californicus*, significantly enhanced translation of RPPP-eGFP in *C. glutamicum* (Figure 6C). EF-P production was confirmed by western blot analysis (Figure 6C). In contrast, EF-Ps from *Actinobacteria* were unable to complement a Δ*efp* Δ*epmA E. coli* mutant (Figure S5), which supports the hypothesis that actinobacterial EF-Ps constitute a separate subfamily of these elongation factors.

The endogenous *efp* genes in *S. coelicolor* and *M. smegmatis* were tagged as described for *C. glutamicum*. Endogenous His-tagged EF-P$_{Sc}$ and EF-P$_{Ms}$ proteins were then purified and compared to a recombinant variant produced in the Δ*efp* Δ*epmA E. coli* strain. In both cases, the intact protein mass measured by mass spectrometry matched the calculated for unmodified EF-P (Figures 7A and 7B). After digestion with chymotrypsin, LC-MS/MS analysis found all peptides comprising the loop region to be unmodified (Figures S6A–S6F). Isoelectric focusing showed no changes in protein charge between endogenous and heterologously produced versions of the proteins (Figures 7C and 7D). Last, unmodified EF-P could also be identified in lysates of *Mycobacterium tuberculosis* H37Rv (Figure 7E) and *M. smegmatis* (Figure S6G). Our results therefore confirm

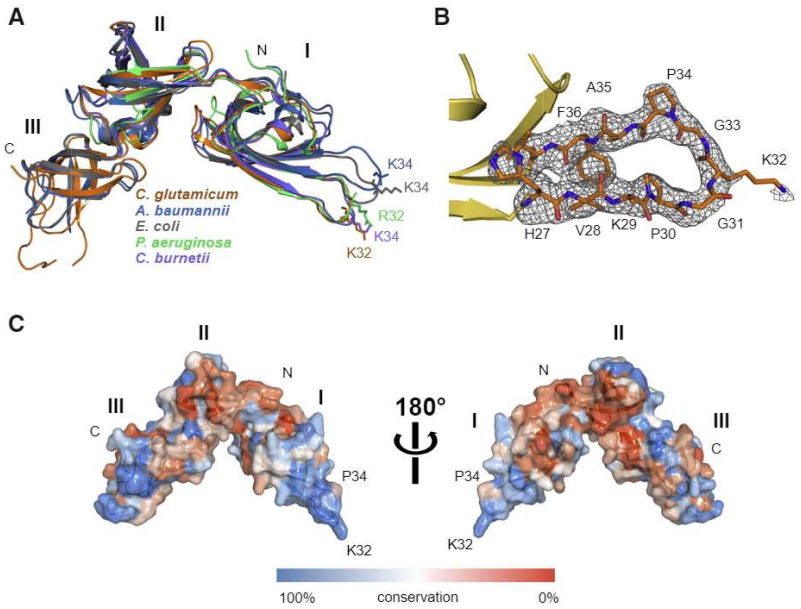


Figure 5. Crystal Structure of *C. glutamicum* EF-P

(A) Structural superposition and ribbon representations of available EF-P structures. *C. glutamicum* (PDB: 6S8Z, orange, this work), *Acinetobacter baumannii* (PDB: 5J3B, blue), *E. coli* (PDB: 3A5Z, gray), *Pseudomonas aeruginosa* (PDB: 3OYY, green), and *Coxiella burnetii* (PDB: 3TRE, purple) are shown. The Arg or Lys residue at the tip of the loop between β2 and β3 of the N-domain is shown as a stick model.

(B) Simulated-annealing omit electron density map contoured at 2.5σ of the connecting loop between β3 and β4 in the N-terminal domain of *C. glutamicum* EF-P.

(C) Sequence conservation across EF-P proteins mapped onto the surface of *C. glutamicum* EF-P using CONSURF (blue, conserved; orange, not conserved). Homologs were identified by BLAST search (Altschul et al., 1990) using the EF-P protein sequence against the translated nucleic acid sequences in the National Center of Biotechnological Information (NCBI) database (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

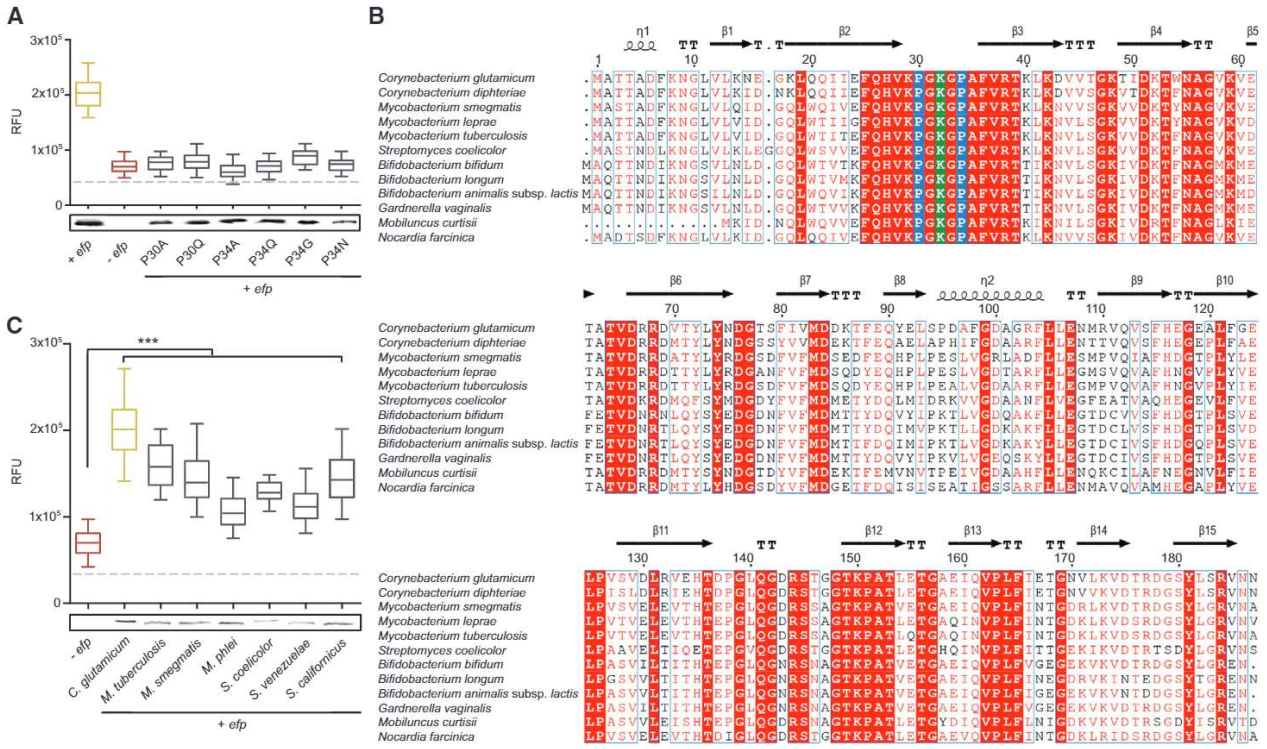


Figure 6. EF-P Conserved Prolines at Position 30 and 34 Are Essential for EF-P Activity

(A) Relative fluorescence units (RFUs) of *C. glutamicum* strains producing RPPP-eGFP and EF-P*Cg* variants with the indicated amino acid replacements at positions 30 and 34. The dashed line indicates the background fluorescence level.

(B) Alignment of EF-P amino acid sequences of different actinobacterial species. Identical amino acids are highlighted in red, conserved K32 in green, and P30 and P34 in blue.

(C) Synthesis of RPPP-eGFP (expressed as RFU) in *C. glutamicum* strains expressing the *efp* genes from the actinobacterial species indicated below the bars. The dashed gray line corresponds to background fluorescence. Statistical significance was tested by unpaired, two-tailed t tests with 99% confidence intervals, ***$p < 0.0001$.

that an unmodified EF-P is active in the most relevant genera of Actinobacteria, including bacteria such as *S. coelicolor*, *M. smegmatis*, and *M. tuberculosis* whose proteomes are among those richest in XPPX motifs.

DISCUSSION

Actinobacteria such as *C. glutamicum* and *S. coelicolor* are important for the industrial production of amino acids, peptides, and other secondary metabolites. Proteomic analysis of the Δ*efp C. glutamicum* mutant reveals downregulation of 222 proteins, most of which contain XPPX motifs as expected. Moreover, this set is not a random sample of the proteome but is enriched in proteins involved in primary metabolism, and the biosynthesis of amino acids, antibiotics, and secondary metabolites. Bottlenecks in these pathways must be avoided during industrial-scale production, and EF-P might play an important role in this context, e.g., when upscaling production of metabolites by these organisms. Comparable proteome analyses of other EF-P/eIF-5A-depleted organisms revealed a predominant importance of EF-P for transcription and translation factors in *E. coli* (Starosta et al., 2014a), and for endoplasmic reticulum stress, and protein folding in HeLa cells (Mandal et al., 2016). These data are in accordance with the idea that EF-P and a/eIF-5A originally evolved to facilitate translation of ValS, the Val-tRNA synthetase with an invariant proline triplet in all kingdoms (Starosta et al., 2014b), but then independently adapted to the specific needs of groups of Bacteria, Archaea, and Eukarya.

Bacteria have evolved various modification mechanisms to activate EF-P. Synthesis and attachment of the specific post-translational modifications generate a functional elongation factor, but at a fitness cost. Here, we report on a novel EF-P subfamily in Actinobacteria that does not require any PTM for activation. The structure of this EF-P subclass was obtained by crystallizing the endogenously produced EF-P_{Cg}. The three β-barrel domains, which together form an L shape, retain all the residues required for interactions with rRNA, tRNAs, and the L1 protein, and cannot be distinguished from other homologs in terms of its overall structure. The functionally important Lys32 at the tip of the β-hairpin is encompassed by two prolines at positions 30 and 34. In subsequent studies, we confirmed that these two prolines are essential for EF-P function in *C. glutamicum*. Replacement of either of them

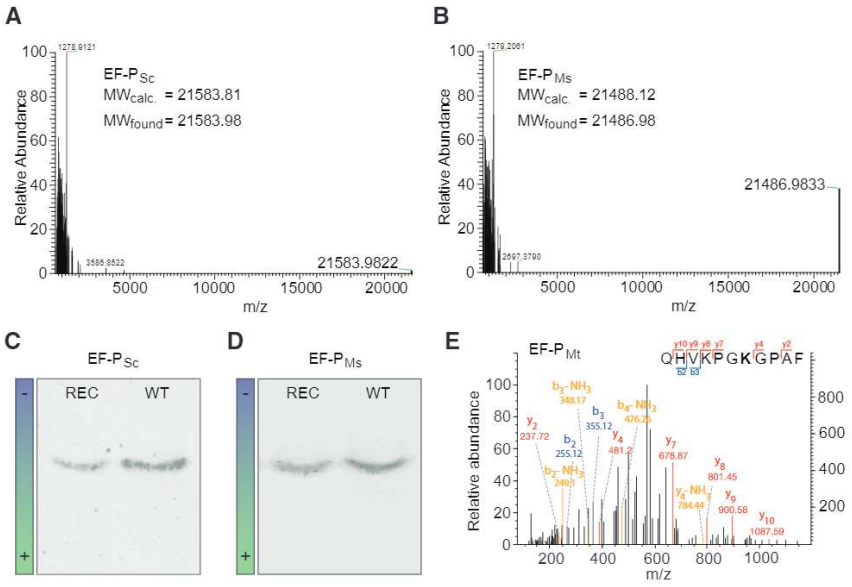


Figure 7. Different Actinobacteria Do Not Post-translationally Modify EF-P

(A–D) Deconvoluted MS spectra of intact endogenous EF-P from *S. coelicolor* (EF-P*Sc*) (A) and *M. smegmatis* (EF-P*Ms*) (B). Isoelectric focusing (IEF) gel of recombinant (REC) and endogenous EF-P (WT) isolated from *S. coelicolor* (C) and *M. smegmatis* (D).

(E) Annotated MS/MS spectra of K32 containing peptide of endogenous EF-P from *M. tuberculosis*.

by alanine, asparagine, glycine, or glutamine (amino acids, which are found in other EF-P subfamilies) results in an inactive EF-P. Due to the high amino acid sequence and structure conservation of EF-Ps, no other characteristic motif than the palindromic loop sequence PGKGP could be identified as unique for this new EF-P subfamily. The pyrrolidine ring of the proline constrains the main chain dihedral angles of the peptide chain and restricts the conformations of the neighboring residues (MacArthur and Thornton, 1991; Schimmel and Flory, 1968). This leads to a decrease in the conformational entropy and thus increase in rigidity of the β-hairpin. For the modified Lys 34, an ε(R)-β-lysylhydroxylysinyl residue, it was shown that it directly interacts with the backbone of the CCA end of the P-site tRNA (Huter et al., 2017). Here, it was argued that this not only stabilizes the P-site tRNA but also forces the prolines to adopt an alternative conformation and thus optimal geometry between the nascent chain and the aminoacyl-tRNA for peptide bond formation (Huter et al., 2017).

Mechanistically, we could not exclude the hypothesis that the conserved and unmodified Lys32 is able to reach further into the ribosomal PTC in comparison to the other modified EF-Ps. However, due to the overall structure identity and amino acid sequence conservation around the loop region (Figures 5 and S4), we suggest that the β-hairpin rigidified by the palindromic sequence Pro-Gly-Lys-Gly-Pro of this new subfamily of EF-Ps is able to position the protruding Lys32 and stabilizes the acceptor arm of the tRNA and facilitates XPPX translation in the absence of a PTM. This hypothesis is supported by several amino acid exchanges that did not shorten the loop but increased flexibility resulting in inactivation of EF-P (Figure 6A). This loop sequence is particularly found in EF-Ps of the Actinobacteria class (Figure S5), and it is definitely not present in EF-Ps of bacteria that possess EF-P modification enzymes (Figure S4E). Sequences rich in Pro-Gly pairs are commonly found in structural proteins, such as collagen, and confer rigidity and structural stability on these proteins (Bella, 2016; Shoulders and Raines, 2009). As we have shown above, the consensus loop sequence is not only essential for the functionality of this EF-P subclass, it also enabled us to determine the structure of this protein with unprecedented resolution.

Actinobacteria is a large and diverse phylogenetic group, and we have identified the unmodified EF-P as the sole form in its most prominent representatives, including species of *Corynebacterium*, *Streptomyces*, and *Mycobacterium*. Moreover, using prolines P30 and P34 as markers, we find this rigid loop in 11% of all bacteria, which suggests that EF-P in other bacteria also does not require a PTM for activation (Figures S4 and S7). Specifically, these encompass other genera of the Actinobacteria, such as *Bifidobacterium*, *Gardnerella*, and *Mobiluncus*, but also some *Flavobacterium* species belonging to the phylum Bacteroidetes and representatives of the Proteobacteria such as *Campylobacter lari* and *C. hominis*. Importantly, none of these bacteria contains any known EF-P modification enzyme.

Lys32 is a conserved residue at the tip of the loop in EF-Ps of Archaea, Eukarya, and most Bacteria, and it is the site of the activating PTM. In contrast to other homologs, the Lys32 of the actinobacterial subfamily of EF-Ps could be replaced by arginine without major loss of function. Although this confirms that Actinobacteria EF-P is active without any PTM, a positively charged amino acid side chain at position 32 is still essential for activity. Similarly, in *B. subtilis* the unmodified EF-P^{K32R}, but not the EF-P^{K32A} variant, has activity as it can restore the swarming phenotype (Hummels et al., 2017). These data reveal that EF-P with a positively charged amino acid at the tip of the loop retains some activity in distinct bacteria.

The fact that some bacteria have evolved a functional EF-P that does not rely on any PTM for its activity is interesting from both mechanistic and energetic points of view. Unicellular and free-living organisms, which often face nutrient limitation and stress conditions, require fast translational responses. Moreover, the burden of secondary metabolite production involving the translation of many polyproline-containing proteins may well be reduced by an EF-P that avoids the energetic and metabolic costs associated with an essential activation step.

STAR★METHODS

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

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Bioinformatics
  - Nucleotides, plasmids, and bacterial strains construction
  - Proteomic analysis
  - Protein cluster analysis
  - *E. coli* EF-P reporter strains
  - *C. glutamicum* EF-P reporter strains, single-cell fluorescence microscopy and quantitative analysis
  - Western blot analysis
  - Purification of endogenous and recombinant EF-P
  - Mass spectrometry of intact proteins
  - Sample preparation for MS-based proteomics to identify putative PTMs on peptide level
  - Sample preparation for MS-based proteomics to detect unmodified EF-P peptides in cell lysates
  - MS measurement and analysis of putatively modified peptides
  - Isoelectric focusing
  - Protein crystallization and structure determination
  - Alignment and Sequence logo
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2020.03.009.

ACKNOWLEDGMENTS

We thank Dr. Marc Bramkamp for providing the wild-type strain of *C. glutamicum* and the plasmids pK19mobSacB and pEKEx2, and Dr. H. Zhao for the plasmid pCM4.4. We thank the European Synchrotron Radiation Facility (ESRF) for beamtime, and the staff of beamline ID30-A3 for setting up the beamline for data collection. This work was financially supported by Deutsche Forschungsgemeinschaft (JU270/17-1 to K.J. and GRK2062 to K.J. and S.S.). S.Y. and E.T. received funding from the European Union's Horizon 2020 Research and Innovation Programme, Grant Agreement No. 720793, TOPCAPI—Thoroughly Optimised Production Chassis for Advanced Pharmaceutical Ingredients.

AUTHOR CONTRIBUTIONS

Conceptualization, B.P., S.S., K.J.; Methodology, B.P., C.M.S., P.K., I.F., S.Y., N.R., E.T., A.I., S.A.S., S.S.S., and K.J.; Investigation, B.P., C.M.S., P.K., M.S., I.F., and S.Y.; Writing – Original Draft, B.P., P.K., S.S., and K.J.; Writing – Review & Editing, B.P., P.K., I.F., S.Y., S.S., K.J.; Funding Acquisition and Resources, E.T., A.I., S.A.S., S.S., and K.J.; Supervision, K.J.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 20, 2019
Revised: February 6, 2020
Accepted: March 2, 2020
Published: March 31, 2020

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.

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. *68*, 352–367.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Bella, J. (2016). Collagen structure: new tricks from a very old dog. Biochem. J. *473*, 1001–1025.

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.

Cooper, H.L., Park, M.H., Folk, J.E., Safer, B., and Braverman, R. (1983). Identification of the hypusine-containing protein hy+ as translation initiation factor eIF-4D. Proc. Natl. Acad. Sci. USA *80*, 1854–1857.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res. *14*, 1188–1190.

Deribe, Y.L., Pawson, T., and Dikic, I. (2010). Post-translational modifications in signal integration. Nat. Struct. Mol. Biol. *17*, 666–672.

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.

Doerfel, L.K., Wohlgemuth, I., Kubyshkin, V., Starosta, A.L., Wilson, D.N., Budisa, N., and Rodnina, M.V. (2015). Entropic contribution of elongation factor P to proline positioning at the catalytic center of the ribosome. J. Am. Chem. Soc. *137*, 12997–13006.

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

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

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

Hersch, S.J., Wang, M., Zou, S.B., Moon, K.M., Foster, L.J., Ibba, M., and Navarre, W.W. (2013). Divergent protein motifs direct elongation factor P-mediated translational regulation in *Salmonella enterica* and *Escherichia coli*. MBio *4*, e00180-e13.

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.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Hummels, K.R., Witzky, A., Rajkovic, A., Tollerson, R., 2nd, Jones, L.A., Ibba, M., and Kearns, D.B. (2017). Carbonyl reduction by YmfI in *Bacillus subtilis* prevents accumulation of an inhibitory EF-P modification state. Mol. Microbiol. *106*, 236–251.

Huter, P., Arenz, S., Bock, L.V., Graf, M., Frister, J.O., Heuer, A., Peil, L., Starosta, A.L., Wohlgemuth, I., Peske, F., et al. (2017).). Structural basis for polyproline-mediated ribosome stalling and rescue by the translation elongation factor EF-P. Mol. Cell *68*, 515–527.e6.

Kabsch, W. (2010a). Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D Biol. Crystallogr. *66*, 133–144.

Kabsch, W. (2010b). Xds. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. *28*, 27–30.

Katoh, K., Rozewicki, J., and Yamada, K.D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinform. *20*, 1160–1166.

Keilberg, D., Wuichet, K., Drescher, F., and Søgaard-Andersen, L. (2012). A response regulator interfaces between the Frz chemosensory system and the MglA/MglB GTPase/GAP module to regulate polarity in Myxococcus xanthus. PLoS Genet. *8*, e1002951.

Kieser, T.B.M., Buttner, M.J., Chater, K.F., and Hopwood, D.A. (2000). Practical *Streptomyces* Genetics (John Innes Foundation).

Klee, S.M., Mostafa, I., Chen, S., Dufresne, C., Lehman, B.L., Sinn, J.P., Peter, K.A., and McNellis, T.W. (2018). An *Erwinia amylovora* yjeK mutant exhibits reduced virulence, increased chemical sensitivity and numerous environmentally dependent proteomic alterations. Mol. Plant Pathol. *19*, 1667–1678.

Lakemeyer, M., Bertosin, E., Möller, F., Balogh, D., Strasser, R., Dietz, H., and Sieber, S.A. (2019). Tailored peptide phenyl esters block ClpXP proteolysis by an unusual breakdown into a heptamer-hexamer assembly. Angew. Chem. Int. Ed. Engl. *58*, 7127–7132.

Larsen, M.H., Biermann, K., Tandberg, S., Hsu, T., and Jacobs, W.R., Jr. (2007). Genetic manipulation of Mycobacterium tuberculosis. Curr. Protoc. Microbiol. *Chapter 10*, Unit 10A.12.

Lassak, J., Keilhauer, E.C., Fürst, M., Wuichet, K., Gödeke, J., Starosta, A.L., Chen, J.M., Søgaard-Andersen, L., Rohr, J., Wilson, D.N., et al. (2015). Arginine-rhamnosylation as new strategy to activate translation elongation factor P. Nat. Chem. Biol. *11*, 266–270.

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. *44* (*W1*), W242–W245.

Luft, J.R., and DeTitta, G.T. (1999). A method to produce microseed stock for use in the crystallization of biological macromolecules. Acta Crystallogr. D Biol. Crystallogr. *55*, 988–993.

Ly, M.A., Liew, E.F., Le, N.B., and Coleman, N.V. (2011). Construction and evaluation of pMycoFos, a fosmid shuttle vector for Mycobacterium spp. with inducible gene expression and copy number control. J. Microbiol. Methods *86*, 320–326.

MacArthur, M.W., and Thornton, J.M. (1991). Influence of proline residues on protein conformation. J. Mol. Biol. *218*, 397–412.

MacNeil, D.J., Gewain, K.M., Ruby, C.L., Dezeny, G., Gibbons, P.H., and MacNeil, T. (1992). Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. Gene *111*, 61–68.

Mandal, A., Mandal, S., and Park, M.H. (2016). Global quantitative proteomics reveal up-regulation of endoplasmic reticulum stress response proteins upon depletion of eIF5A in HeLa cells. Sci. Rep. *6*, 25795.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Cryst. *40*, 658–674.

Melnikov, S., Mailliot, J., Shin, B.S., Rigger, L., Yusupova, G., Micura, R., Dever, T.E., and Yusupov, M. (2016). Crystal structure of hypusine-containing translation factor eif5a bound to a rotated eukaryotic ribosome. J. Mol. Biol. *428*, 3570–3576.

Motz, M., and Jung, K. (2018). The role of polyproline motifs in the histidine kinase EnvZ. PLoS ONE *13*, e0199782.

Navarre, W.W., Zou, S.B., Roy, H., Xie, J.L., Savchenko, A., Singer, A., Edvokimova, E., Prost, L.R., Kumar, R., Ibba, M., and Fang, F.C. (2010). PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. Mol. Cell *39*, 209–221.

Paget, M.S., Chamberlin, L., Atrih, A., Foster, S.J., and Buttner, M.J. (1999). Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). J. Bacteriol. *181*, 204–211.

Park, M.H., Cooper, H.L., and Folk, J.E. (1981). Identification of hypusine, an unusual amino acid, in a protein from human lymphocytes and of spermidine as its biosynthetic precursor. Proc. Natl. Acad. Sci. USA *78*, 2869–2873.

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

Peil, L., Starosta, A.L., Lassak, J., Atkinson, G.C., Virumäe, K., Spitzer, M., Tenson, T., Jung, K., Remme, J., and Wilson, D.N. (2013). Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. Proc. Natl. Acad. Sci. USA *110*, 15265–15270.

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. *47* (*D1*), D442–D450.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS ONE *5*, e9490.

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.

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

Reiling, N., Homolka, S., Walter, K., Brandenburg, J., Niwinski, L., Ernst, M., Herzmann, C., Lange, C., Diel, R., Ehlers, S., and Niemann, S. (2013). Clade-specific virulence patterns of *Mycobacterium tuberculosis* complex strains in human primary macrophages and aerogenically infected mice. MBio *4*, e00250-13.

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. *42*, W320–W324.

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

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., and Pühler, A. (1994). Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene *145*, 69–73.

Schimmel, P.R., and Flory, P.J. (1968). Conformational energies and configurational statistics of copolypeptides containing L-proline. J. Mol. Biol. *34*, 105–120.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schrödinger, L.L.C. (2015). The PyMOL Molecular Graphics System, Version 1.8 (Schrödinger, L.L.C.).

Shoulders, M.D., and Raines, R.T. (2009). Collagen structure and stability. Annu. Rev. Biochem. *78*, 929–958.

Starosta, A.L., Lassak, J., Peil, L., Atkinson, G.C., Virumäe, K., Tenson, T., Remme, J., Jung, K., and Wilson, D.N. (2014a). Translational stalling at polyproline stretches is modulated by the sequence context upstream of the stall site. Nucleic Acids Res. *42*, 10711–10719.

Starosta, A.L., Lassak, J., Peil, L., Atkinson, G.C., Woolstenhulme, C.J., Virumäe, K., Buskirk, A., Tenson, T., Remme, J., Jung, K., and Wilson, D.N. (2014b). A conserved proline triplet in Val-tRNA synthetase and the origin of elongation factor P. Cell Rep. *9*, 476–483.

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.

Tollerson, R., 2nd, Witzky, A., and Ibba, M. (2018). Elongation factor P is required to maintain proteome homeostasis at high growth rate. Proc. Natl. Acad. Sci. USA *115*, 11072–11077.

Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat. Protoc. *11*, 2301–2319.

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.

UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. *47* (*D1*), D506–D515.

Vanhoof, G., Goossens, F., De Meester, I., Hendriks, D., and Scharpé, S. (1995). Proline motifs in peptides and their biological processing. FASEB J. *9*, 736–744.

Venkatachalam, C.M., and Ramachandran, G.N. (1969). Conformation of polypeptide chains. Annu. Rev. Biochem. *38*, 45–82.

Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G.F., Chater, K.F., and van Sinderen, D. (2007). Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. Microbiol. Mol. Biol. Rev. *71*, 495–548.

Volkwein, W., Krafczyk, R., Jagtap, P.K.A., Parr, M., Mankina, E., Macošek, J., Guo, Z., Fürst, M.J.L.J., Pfab, M., Frishman, D., et al. (2019). Switching the post-translational modification of translation elongation factor EF-P. Front. Microbiol. *10*, 1148.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. *67*, 235–242.

Witzky, A., Hummels, K.R., Tollerson, R., 2nd, Rajkovic, A., Jones, L.A., Kearns, D.B., and Ibba, M. (2018). EF-P posttranslational modification has variable impact on polyproline translation in *Bacillus subtilis*. MBio *9*, e00306-18.

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.

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.

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.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit monoclonal anti 6xHis antibodies	Abcam	ab200537
Goat anti-Rabbit IgG	Abcam	ab216773
Bacterial and Virus Strains		
See Table S7		N/A
Deposited Data		
Crystallographic data of EF-P structures	Protein Data Bank	PDB: 6S8Z
Mass spectrometry proteomics data	ProteomeXchange	ProteomeXchange: PXD014742
Oligonucleotides		
See Table S5		N/A
Recombinant DNA		
See Table S6		N/A
Software and Algorithms		
MAFF software package version 7.409	Katoh et al., 2019	https://mafft.cbrc.jp/alignment/software/
FastTree version 2.1	Price et al., 2010	http://www.microbesonline.org/fasttree/
iTOL version 3	Letunic and Bork, 2016	https://itol.embl.de/
Fiji version 1.52i	Schindelin et al., 2012	https://imagej.net/Fiji
MaxQuant version 1.5.2.8	Tyanova et al., 2016	https://www.maxquant.org/
Database for annotation, visualization and integrated discovery (DAVID) version 6.7	Huang et al., 2009	https://david-d.ncifcrf.gov/
Xcalibur Software version 3.0sp2	Thermo Scientific	N/A
Phenix	Afonine et al., 2012	http://www.phenix-online.org/
PyMOL	Schrödinger, 2015	https://pymol.org/2/
ESPript 3.0	Robert and Gouet, 2014	http://espript.ibcp.fr
WebLogo3	Crooks et al., 2004	http://weblogo.threeplusone.com/
GraphPad Prism version 8.3.1 for Windows	GraphPad Software, San Diego, California USA	https://www.graphpad.com/

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Kirsten Jung (jung@lmu.de). All unique reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All the bacterial strains used in this study are listed in Table S7. *E. coli* was grown in lysogeny broth (LB) under aerobic conditions at 37°C. *C. glutamicum* was routinely grown in Brain-Heart Infusion (BHI) broth at 30°C unless indicated otherwise. *S. coelicolor* was cultivated aerobically in SFM medium or tryptic soy broth (TSB) medium at 30°C. *M. smegmatis* was grown in BHI medium at 37°C. *M. tuberculosis* was cultivated in Middlebrook 7H9 medium supplemented with 10% oleic acid-albumin-dextrose-catalase, 0.05% tween 80 and 0.2% glycerol, incubated aerobically at 37°C and carried out under biosafety level 3 conditions (Reiling et al., 2013). When necessary, antibiotics were used in the following concentrations: apramycin 50 μg/mL; chloramphenicol 10 μg/mL (*S. coelicolor*) or 34 μg/mL (*E. coli*); kanamycin 10 μg/mL (*M. smegmatis*), 25 μg/mL (*C. glutamicum*, *E. coli*) or 50 μg/mL (*S. coelicolor*).

METHOD DETAILS

Bioinformatics

A set of fully sequenced prokaryotic genomes identified by Keilberg et al. (2012) and Lassak et al. (2015) had efp sequence(s) and its known modification enzymes identified as following: efp - EF-P domain; epmA – tRNA-synt_2 without tRNA_anti-codon domain; epmB – Radical_SAM domain containing protein within a distance of 4 coding regions from efp; epmC – DUF462 domain; earP – DUF2331 domain and ymfI adh_short or adh_shortC2 domains as discribed previously (Hummels et al., 2017; Lassak et al., 2015). EF-P sequences were obtained by retrieving the data against Uniprot database resulting in a set of 937 sequences. Multiple sequence alignments were constructed using the l-ins-l algorithm implemented in the MAFF software package version 7.409 (Katoh et al., 2019). Phylogenetic trees were constructed using FastTree 2.1 with default settings (Price et al., 2010). Annotations and management of phylogenetic trees were done with iTOL version 3 (Letunic and Bork, 2016). To search for polyproline motifs, reference proteomes for each selected strain were downloaded from UniProt (UniProt Consortium, 2019). Motifs search was done using the program CLC Main Workbench v. 7.7.3 and characterized as weak, moderate or strong according to Qi et al. (2018).

Nucleotides, plasmids, and bacterial strains construction

Primers, plasmids and strains used in this study are listed in Tables S5–S7. Enzymes and kits were used according to the manufacturers' standard protocols. Genomic DNA from *C. glutamicum*, *E. coli*, *M. smegmatis*, *M. phlei*, *S. venezuelae*, and *S. californicus* was purified with Nucleospin Microbial DNA from Macherey-Nagel. gDNA from *S. coelicolor* was extracted following the salting-out procedure (Kieser et al., 2000). *M. tuberculosis* gDNA was purified from liquid cultures by chloroform/isoamyl alcohol extraction protocol after treatment with lysozyme, proteinase K and cetrimide (hexadecyltrimethylammonium bromide) (Larsen et al., 2007). DNA polymerases (Q5) and restriction endonucleases were purchased from New England Biolabs (NEB). DNA fragments were purified from agarose gels using the High-Yield PCR Cleanup and Gel Extraction kit from Süd-Laborbedarf Gauting. Plasmid purifications from liquid cultures were performed using the High-Yield Plasmid DNA Purification kit from the same source. All amino acid exchanges were constructed by two-step PCR using mismatched primer pairs (Ho et al., 1989). NEBuilder HiFi DNA Assembly Master Mix (NEB) was used for Gibson assembly. *E. coli* DH5α (Promega) and *E. coli* ET12567[pUZ802] (MacNeil et al., 1992; Paget et al., 1999) were used as hosts for plasmid construction and maintenance, and for intergeneric conjugation with *S. coelicolor*, respectively. *C. glutamicum* mutants were constructed using the pK19mobsacB recombinant vector with SacB counterselection (Schäfer et al., 1994). *S. coelicolor* LW277, harboring the native *efp* coding sequence with an in-frame translational fusion to the His-tag, was constructed by intergenic conjugation of pTE1213 into wild-type *S. coelicolor* M145 and subsequent apramycin selection, followed by a second intergenic conjugation with pTE1208 and counterselection on kanamycin, and finally two consecutive passages on Soya Flour Mannitol (SFM) without antibiotics. Plasmids pCM4.4 (E.T. and S.Y., kindly provided by H. Zhao, unpublished data) and pCMU-4K (S.Y., unpublished data) were used for *S. coelicolor* genome editing by CRISPR-Cas9 and for plasmid curing, respectively.

Proteomic analysis

Cells ($5x10^8$) were processed using the iST kit (PreOmics) following the manufacturer's instructions and resuspended to yield 0.8 mg/mL protein. For LC-MS/MS purposes, 5-μl aliquots of desalted peptides were injected into an Ultimate 3000 RSLCnano system (Thermo), separated in a 15-cm analytical column (75 μm ID home-packed with ReproSil-Pur C18-AQ 2.4 μm from Dr. Maisch) using a 120-min gradient from 5 to 60% acetonitrile in 0.1% (v/v) formic acid. The effluent from the HPLC was directly electrosprayed into a Q Exactive HF (Thermo) operated in data-dependent mode to automatically switch between full-scan MS and MS/MS acquisition. Survey full-scan MS spectra (from *m/z* 375–1600) were acquired with resolution 60,000 at *m/z* 400 (AGC target of $3x10^6$). The 10 most intense peptide ions with charge states between 2 and 5 were sequentially isolated to a target value of $1x10^5$, and fragmented at 27% normalized collision energy with resolution 15,000 at *m/z* 400. Typical mass spectrometric conditions were: spray voltage, 1.5 kV; no sheath or auxiliary gas flow; heated capillary temperature, 250°C; ion selection threshold, 33,000 counts. MaxQuant 1.5.2.8 (Tyanova et al., 2016) was used to identify proteins and quantify them by LFQ with the following parameters: Database, uniprot_3AUP000000582_Cglutamicum _15032017; MS tol, 10ppm; MS/MS tol, 20ppm; Peptide FDR, 0.1; Protein FDR, 0.01 Min.; peptide Length, 5; Variable modifications, Oxidation (M); Fixed modifications, Carbamidomethyl (C); Peptides for protein quantitation, razor and unique; Min. peptides, 1; Min. ratio count, 2. Identified proteins were considered as statistically significant with FDR = 0.05 and s0 = 1 (Two-sample test adjusted for multiple comparisons, Perseus). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD014742.

Protein cluster analysis

Downregulated proteins identified by MaxQuant were uploaded into the database for annotation, visualization and integrated discovery (DAVID) (Huang et al., 2009). Standard settings were used and significantly overrepresented pathways were identified in the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000) Pathway Database. Functional clusters with *P*-values below 0.05 were considered overrepresented.

E. coli EF-P reporter strains

E. coli P_{cadBA}::*lacZ* Δ*cadAB* Δ*efp* or *E. coli* P_{cadBA}::*lacZ* Δ*cadAB* Δ*efp* ΔepmA cells were transformed with pBAD33 expressing EF-P variants under the control of the P_{BAD}, and grown overnight with continuous shaking in 1.8-mL aliquots of LB buffered with phosphate to pH 5.8 [KH_2PO_4 91.5 mM, K_2HPO_4 8.5 mM] in tightly closed 2-mL centrifuge tubes to provide a microaerophilic atmosphere. Cells were harvested and β-galactosidase activity was determined as previously described (Tetsch et al., 2008; Ude et al., 2013).

C. glutamicum EF-P reporter strains, single-cell fluorescence microscopy and quantitative analysis

For quantification of EF-P activity in *C. glutamicum*, the strain ATCC13032 Δ*efp* P*dnaK*-RPPP-*egfp* was transformed with pEKEx2 containing a copy of *efp* under the control of the native EF-P_{Cg} promoter. A 50 μL aliquot of an overnight culture was inoculated into fresh BHI medium and incubated under vigorous shaking for 2 h at 30°C to stimulate exponential growth, followed by 1 h at 40°C to induce P_{dnaK} and 1 h at 30°C to allow for recovery and folding of eGFP. Aeration was provided during the whole incubation time. To measure eGFP fluorescence, cells were washed in ice-cold, phosphate-buffered saline (PBS) and fixed on an agarose pad [1% w/v in PBS] placed on a microscope slide with coverslip. Micrographs were taken on a Leica microscope DMI 6000B equipped with a Leica DFC 365Fx camera (Andor, 12bit). eGFP fluorescence was visualized using an excitation wavelength of 460 nm and a 512 nm emission filter with a 75-nm bandwidth. Fluorescence intensities of a minimum of 300 cells per *efp* transformant were collected and quantified using Fiji (Schindelin et al., 2012). Statistical analysis was done by using two-tailored t test. Quantification of Cgl1117 production was done by monitoring fluorescence of Cgl1117-eGFP. The strains *C. glutamicum* Δ*efp* and *C. glutamicum efp*-6His (which produces His-tagged EF-P) were grown in BHI medium to an OD_{600} of 2. Cells were collected by centrifugation, washed in ice-cold PBS, fixed and imaged under the microscope. eGFP fluorescence was quantified as described above.

Western blot analysis

SDS-polyacrylamide gels were used to fractionate proteins. When necessary, the gels were stained with Instant blue. Proteins were transferred to nitrocellulose membranes by the wet-transfer method. Endogenous EF-P containing an N-terminal 6xHis tag was detected with 1:10,000 dilutions of anti-6xHis antibodies (Abcam) in TBS supplemented with 3% (w/v) BSA. The fluorescence-labeled secondary antibody (IRDye 800CW, Abcam) was used at a concentration of 1:20,000. Membranes were washed in TBS-TT buffer [10 mM Tris/HCl pH 7.5, 150 mM NaCl, 0.05% (v/v) Tween 20, 0.2% (v/v) Triton 100] and scanned using the Odyssey CLx imaging system (LI-COR Biosciences).

Purification of endogenous and recombinant EF-P

For purification of the *C. glutamicum* EF-P, the strain *C. glutamicum efp*-6His was grown in BHI medium to an OD_{600} of 2 (exponential phase) or overnight (stationary growth phase). Endogenous EF-P was isolated from the *S. coelicolor efp*-6His strain after cultivation of cells for 48 h or 7 days. Endogenous EF-P was isolated from *M. smegmatis* pMycoFos (Ly et al., 2011) P_{efp}-*efp*-6His transformants grown at 37°C for 24 or 72 h.

EF-P variants with amino acid replacements were produced in *C. glutamicum* Δ*efp* cells grown in rich medium supplemented with the appropriate antibiotic overnight. Recombinant EF-P was expressed in *E. coli* MG1655 Δ*efp* Δ*epmA* cells grown in LB medium supplemented with 0.2% (w/v) arabinose and antibiotic as needed. Cells were collected and resuspended in lysis buffer (25 mM HEPES pH 8, 125 mM NaCl, 25 mM KCl). Cells were lysed using the high-pressure system from Constant Systems, and the cytosolic fractions were obtained after ultracentrifugation. Fractions were kept on ice prior to further purification. Endogenous and recombinant proteins for MS analysis were purified using a Ni^{2+}-nitrilotriacetic acid (NTA) resin (QIAGEN), washed and eluted, respectively, in lysis buffer supplemented with 20 mM and 200 mM imidazole. For crystallization of the endogenous *C. glutamicum* EF-P, the eluate was further fractionated by size-exclusion chromatography on a Superdex 200 10/300GL column (GE Life Sciences). Fractions containing EF-P were combined, injected into a dialysis tube (SnakeSkin, ThermoFisher) and concentrated by incubation in lysis buffer saturated with sucrose.

Mass spectrometry of intact proteins

Samples were desalted and measured using a MassPREP On-Line Desalting Cartridge (Waters) on an Ultimate 3000 HPLC system (Dionex) coupled to a Finnigan LTQ-FT Ultra mass spectrometer (Thermo Scientific) with electrospray ionization (spray voltage 4.0 kV, tube lens 110 V, capillary voltage 48 V, sheath gas 60 arb, aux gas 10 arb, sweep gas 0.2 arb). Xcalibur Xtract Software (Thermo Scientific) was used for data analysis and deconvolution.

Sample preparation for MS-based proteomics to identify putative PTMs on peptide level

Purified protein (1 μg) was dissolved in 200 μl of X-buffer (7 M urea, 2 M thiourea in 20 mM HEPES buffer, pH 7.5) for trypsin digestion or in 200 μl of ABC buffer (25 mM ammonium bicarbonate) for digestion by chymotrypsin. Upon reduction with 1 mM DTT (0.2 μl of 1 M stock in ddH_2O) for 45 min at 25°C, proteins were alkylated using 5.5 mM IAA (2 μl of 550 mM stock in ddH_2O) for 30 min at 25°C and samples were quenched with 4 mM DTT (0.8 μl of 1 M stock) for 30 min at 25°C. Samples intended for trypsin digestion were first digested with LysC (1 μl of 0.5 μg/μl, Wako, MS grade) for 2 h at 25°C, then diluted with triethylammonium bicarbonate (TEAB) buffer (600 μl of 50 mM stock in ddH_2O) and digested with trypsin (1.5 μl of 0.5 μg/μl in 50 mM acetic acid, Promega, sequencing grade) for a further 16 h at 37°C. Samples for chymotrypsin digestion were supplemented with 2 μl of 1 M $CaCl_2$, then 1 μl of 0.5 μg/μl

chymotrypsin (Promega) was added and the mixture was incubated for 16 h at 25°C. Samples were then acidified with 1% (v/v) FA and desalted using SepPak® C18 cartridges (50 mg, Waters) with a vacuum manifold. The cartridges were first washed with ACN (2 × 1 ml) and equilibrated with 0.5% (v/v) FA (3 × 1 ml) prior to loading the samples. After washing with 0.5% (v/v) FA (3 × 1 ml), peptides were eluted with 80% (v/v) ACN containing 0.5% FA (2 × 0.25 ml) and freeze-dried using a Speedvac centrifuge. Samples were prepared for MS analysis by dissolving them in 30 μl of 1% (v/v) FA and filtering through 0.22-μm PVDF filters (Millipore).

Sample preparation for MS-based proteomics to detect unmodified EF-P peptides in cell lysates

M. tuberculosis H37Rv cultures were inactivated by incubating frozen 0.5-1g wet weight pallets in 5mL methanol under agitation at room temperature for 16 hours. Subsequently, methanol was removed by agitation at 37°C overnight. The remaining pellet was solubilized in 200μL lysis buffer for future proteomic analysis. Successful inactivation was validated by Mycobacteria Growth Indicator Tube analysis (MGIT) and culture on 7H10 agar supplemented with 10% bovine calf serum. *M. smegmatis* cultures were harvested and resuspended in lysis buffer. Cell disruption was achieved in two steps. First by transferring 300μL of the cell suspensions to a sterile 2mL screwcap centrifuge tubes, including 300μL of 0.1mm glass pearls and agitating at maximum speed for three rounds of 45 s using a sample homogenizer (FastPrep-24, MP Biomedicals). Second, we transferred the cell suspensions to a new tube and optimized the lysis using the high pressure system (Constant Systems) as described above. Mass spectrometry samples were prepared as described above using trypsin for *M. smegmatis* and chymotrypsin for *M. tuberculosis*.

MS measurement and analysis of putatively modified peptides

MS analysis was performed on a Q Exactive Plus instrument coupled to an Ultimate3000 Nano-HPLC via an Easy-Spray ion source (Thermo Scientific). Samples were loaded on a 2-cm PepMap RSLC C18 trap column (particles 3 μm, 100A, inner diameter 75 μm, Thermo Scientific) with 0.1% (v/v) TFA and separated on a 50 cm PepMap RSLC C18 column (particles 2 μm, 100A, inner diameter 75 μm, Thermo Scientific) held at a constant temperature of 50°C. The gradient was run from 5%–32% acetonitrile, 0.1% (v/v) FA over a period of 152 min (7 min 5%, 105 min to 22%, 10 min to 32%, 10 min to 90%, 10 min wash at 90%, 10 min equilibration at 5%) at a flow rate of 300 nL/min. For measurements of chemical-proteomic samples on the fusion instrument, survey scans (*m/z* 300-1,500) were acquired in the orbitrap with a resolution of 120,000 at *m/z* 200 and the maximum injection time set to 50 ms (target value 2e5). Most intense ions of charge states 2-7 were selected for fragmentation with high-energy collisional dissociation at a collision energy of 30%. The instrument was operated in top-speed mode and spectra were acquired in the ion trap with the maximum injection time set to 50 ms (target value 1e4). The option to inject ions for all available parallelizable times was enabled. Dynamic exclusion of sequenced peptides was set to 60 s. Real-time mass calibration was based on internally generated fluoranthene ions. Data were acquired using Xcalibur software version 3.0sp2 (Thermo Scientific).

For measurements of the endogenous EF-P from *M. smegmatis* and *M. tuberculosis* we used Orbitrap Fusion (Thermo Scientific) with Ionopticks 25 cm Aurora Series emitter column (25 cm x 75 μm ID, 1.6 μm C18), the flow was adjusted to 400 nL/min and 40°C was used for trap and main column. Fusion instrument survey scans (m/z 300-1,500) were acquired in the orbitrap with a resolution of 120,000 at m/z 200 and the maximum injection time set to 50 ms (target value 2e5). Most intense ions of charge states 2-7 were selected for fragmentation with high-energy collisional dissociation at a collision energy of 30%. The instrument was operated in top speed mode and spectra were acquired in the ion trap with the maximum injection time set to 50 ms (target value 1e4). The option to inject ions for all available parallelizable time was enabled. Dynamic exclusion of sequenced peptides was set to 60 s.

MS raw files were analyzed with MaxQuant software (version 1.5.3.8). MS/MS-based peptide identification was carried out using the Andromeda search engine with fasta files containing WT *efp* and *efp* point mutants. For recombinant EF-P from *E. coli*, the *E. coli* UniProtKB database was also used. The following parameter settings were employed: peptide and protein FDR, 1%; enzyme specificity, trypsin; minimal number of amino acids required for peptide identification, 7; variable modification, methionine oxidation; fixed modification, carbamidomethylation. At least one unique peptide was required for the identification of protein. All other parameters were used according to the default settings. For identification of putatively unknown modifications, search for dependent peptides was enabled. Potential contaminants and reverse hits were removed from the result lists.

Isoelectric focusing

Purified EF-P (1 μg) was loaded on a precast isoelectric focusing gel with a pH gradient range of 4-7 (SERVAGel, Serva). Prior to sample application, gels were prefocused at 100 V for 10 min. Samples were then focused for 1 h at 200 V, 1 h at 300 V and 30 min at 500 V. Proteins were transferred to a nitrocellulose membrane by the wet-transfer method and detected with anti-6xHis antibodies as described above.

Protein crystallization and structure determination

Prior to crystallization, EF-P was concentrated to 3 mg/ml using centrifugal filter devices (Amicon, Merck). Aggregates and debris were removed by centrifugation (16.000x*g* for 20 min). Diffraction-quality crystals of *C. glutamicum* EF-P were obtained by micro-seeding in 100 mM sodium acetate, 100 mM HEPES pH 7.5, 22% (w/v) PEG4000 at 4°C (0.3 μL protein, 0.2 μL precipitant, 0.1 μL seed stock). The seed stock was generated using the Seed Bead kit (Hampton Research) according to the manufacturer's protocol (Luft and DeTitta, 1999) and crystals of low diffraction quality, previously obtained under various PEG-based conditions in a high-throughput screening campaign. For cryoprotection, crystals were soaked in mother liquor supplemented with 30% (w/v)

ethylene glycol, flash-cooled and stored in liquid nitrogen. Data collection was carried out at the synchrotron beamline ID 30-A3 at the ESRF (European Synchrotron Radiation Facility, Grenoble, France). The data were processed with XDS (Kabsch, 2010b) and the structure solved by molecular replacement with PHASER (McCoy et al., 2007), using the coordinates of EF-P from *P. aeruginosa* (PDB code: 3OYY), which shares 33% sequence identity with the homolog from *C. glutamicum*, as search model, after truncation to the C-alpha carbon atoms (CHAINSAW, CCP4) (Kabsch, 2010a; McCoy et al., 2007; Winn et al., 2011). Model building was done with COOT and refinement of the coordinates was carried out with Phenix (Afonine et al., 2012). The structural figures were prepared using PyMOL 2.3 (Schrödinger, 2015). For data processing and structure refinement statistics, see Table S4.

Alignment and Sequence logo

Alignments of EF-P sequences from diverse Actinobacteria (Figure 6) and previously crystallized EF-P proteins (Figure S3A) were uploaded in combination with EF-P_{Cg} structure to the ESPript 3.0 platform (Robert and Gouet, 2014) for rendering analysis of secondary structure. Sequence logos of β2-loop-β3 EF-P sequences were generated by uploading the alignment of all EF-P sequences selected according to Figures S3B–S3E on WebLogo3 platform (Crooks et al., 2004).

QUANTIFICATION AND STATISTICAL ANALYSIS

Two-tailed t tests were performed using GraphPad Prism version 8.3.1 for Windows. Differences are considered significant when *p-value*s were less than 0.05. Mean and standard deviation are shown unless indicated otherwise in the figure legends. All data are representative of at least three different experiments. The number of bacterial cells analyzed under fluorescence microscopy are shown in the respective figure legends. Pathways significantly enriched in downregulated proteins were identified by Fisher Exact test using the online platform DAVID (Huang et al., 2009) as indicated above.

DATA AND CODE AVAILABILITY

Crystallographic data of EF-P structures have been deposited in the Protein Data Bank (https://www.ebi.ac.uk/pdbe/) under the PDB accession code PDB: 6S8Z. The mass spectrometry proteomics data are available via ProteomeXchange with identifier ProteomeXchange: PXD014742.