

---

# The Translation Elongation Factor P in Actinobacteria

---

Dissertation der Fakultät für Biologie  
der Ludwig-Maximilians-Universität München



vorgelegt von

**Bruno Pinheiro Damasceno Florentino**  
aus Belo Horizonte, Brasilien

München, Mai 2020

Gutachter:

1. Prof. Dr. Kirsten Jung
2. Prof. Dr. Jörg Nickelsen

Datum der Abgabe: 07.05.2020

Datum der mündlichen Prüfung: 06.07.2020

## **Eidesstattliche Erklahrung**

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbststandig und ohne unerlaubte Hilfe angefertigt wurde. Des Weiteren erklare ich, dass ich nicht anderweitig ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprufung zu unterziehen. Die folgende Dissertation liegt weder ganz, noch in wesentlichen Teilen einer anderen Prufungskommission vor.

Bruno Pinheiro

Munchen, 07.05.2020

## **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.

Bruno Pinheiro

Munich, 07.05.2020

# Table of Contents

<b>Eidesstattliche Erklärung</b> .....	<b>III</b>
<b>Statutory Declaration</b> .....	<b>III</b>
<b>Table of Contents</b> .....	<b>IV</b>
<b>Nomenclature</b> .....	<b>VI</b>
<b>Abbreviations</b> .....	<b>VII</b>
<b>Publications and Manuscripts Presented in this Thesis</b> .....	<b>IX</b>
<b>Contributions to Publications and Manuscripts presented in this Thesis</b> .....	<b>X</b>
<b>Summary</b> .....	<b>XII</b>
<b>Zusammenfassung</b> .....	<b>XIII</b>
<b>1 Introduction</b> .....	<b>1</b>
1.1 Actinobacteria and their relevance .....	1
1.1.1 <i>Corynebacterium glutamicum</i> .....	1
1.1.2 <i>Streptomyces coelicolor</i> .....	2
1.2 Translation in bacteria .....	2
1.3 Ribosome stalling .....	4
1.4 Translation elongation factor P .....	6
1.4.1 Function .....	6
1.4.2 General structure and interactions with the ribosome .....	7
1.4.3 Activation by post-translational modification.....	9
1.4.4 Mechanism of action.....	10
1.5 EF-P in Actinobacteria .....	12
1.6 Scope of the thesis .....	12
<b>2 Results</b> .....	<b>13</b>
2.1 Global effect of EF-P in <i>C. glutamicum</i> .....	13
2.2 Design of a reporter strain .....	16
2.3 A positive charge at position 32 is essential for EF-P activity .....	17
2.4 Development of a protocol for native EF-P purification .....	18
2.5 EF-P in <i>C. glutamicum</i> is not post-translationally modified .....	21
2.6 Crystal structure of <i>C. glutamicum</i> EF-P .....	22
2.7 P30 and P34 are essential for EF-P activity .....	25
2.8 A novel subfamily of EF-P in Actinobacteria.....	28
2.9 EF-P is required for fast growth of <i>C. glutamicum</i> on glucose .....	30
2.10 EII <sup>Glc</sup> translation requires EF-P .....	34
2.11 EF-P dependency of GntR2 does not affect transcription of <i>ptsG</i> .....	37
2.12 The polyproline motif is essential for EII <sup>Glc</sup> activity .....	38
2.13 Polyprolines are widespread in EIIC subunits .....	39
2.14 EF-P accumulates in Actinobacteria during early stationary phase .....	41
2.15 EF-P is not required for growth in <i>S. coelicolor</i> .....	42
2.16 Polyproline motifs are underrepresented in the <i>S. coelicolor</i> core proteome.....	44
2.17 Proteins required for secondary metabolite production contain polyprolines .....	45
2.18 EF-P is essential for antibiotic production in <i>S. coelicolor</i> .....	48
<b>3 Discussion</b> .....	<b>50</b>
3.1 Outlook.....	56

## Table of Contents

---

<b>4</b>	<b>Materials and Methods</b> .....	<b>57</b>
4.1	Construction of the phylogenetic trees .....	57
4.2	Polyproline motifs stalling strength and frequency in proteomes .....	57
4.3	Nucleotides, plasmids and bacterial strains construction .....	58
4.4	Growth conditions .....	64
4.5	Proteomic analysis.....	64
4.6	Protein cluster analysis .....	65
4.7	<i>E. coli</i> EF-P reporter strains.....	65
4.8	<i>C. glutamicum</i> EF-P reporter strain .....	65
4.9	Single-cell fluorescence microscopy and quantitative analysis .....	65
4.10	Western blot analysis.....	66
4.11	Quantitative Western blot .....	66
4.12	Purification of endogenous and recombinant EF-P .....	67
4.13	Mass spectrometry of intact proteins .....	68
4.14	Sample preparation for MS-based proteomics .....	68
4.15	MS measurement and analysis .....	69
4.16	Isoelectric focusing .....	70
4.17	Protein crystallisation and structure determination.....	70
4.18	RT-qPCR analysis .....	70
4.19	<i>In vitro</i> translation of EII <sup>Glc</sup> .....	71
4.20	D-glucose-6- <sup>14</sup> C uptake assay.....	71
4.21	Protein overproduction in <i>C. glutamicum</i> .....	72
4.22	EIIC <sup>Glc</sup> tertiary structure prediction.....	72
4.23	Statistical analysis .....	72
4.24	Crystallographic and proteomic data availability .....	72
	<b>Acknowledgements</b> .....	<b>73</b>
<b>5</b>	<b>References</b> .....	<b>74</b>

## Nomenclature

Amino acid sequences are numbered in a way that the first methionine/valine of the wild type protein is designated “1” independently of the N-terminal affinity or fluorescence tag (if existent). Affinity or fluorescence tags are marked in genes and proteins corresponding to their position, e.g. N-terminal mNG-EII<sup>Glc</sup> or C-terminal *efp*-6His. Amino acid substitutions in proteins are designated in one-letter code, followed by the respective amino acid position in the wild-type protein, and another one-letter code for the amino acid introduced by site-directed mutagenesis (e.g. EF-P K32A). Gene deletions are marked by “ $\Delta$ ”.

## Abbreviations

aa	Amino acid
aa-tRNA	Aminoacyl-tRNA
ACBDRib	Ribose specific ATP-binding cassette transporter for D-ribose import
ACT	Actinorhodin
aIF5A	Archaea translation initiation factor 5A
ArfA	Alternative ribosome-rescue factor A
ArfB	Alternative ribosome-rescue factor B
BHI	Brain-Heart Infusion
CGXII	Corynebacterium glutamicum minimal medium XII
CPK	Coelimycin P1
C $\alpha$	Carbon alpha
DNA	Deoxyribonucleic acid
DW	Dry weight
EF	Elongation factor
EF-P	Translation Elongation factor P
eGFP	Enhanced green fluorescent protein
EI	Enzyme I (encoded by <i>ptsI</i> )
eIF5A	Eukarya translation initiation factor 5A
eIF5Ad	Eukarya translation initiation factor 5A - depleted
EIIC <sup>Glc</sup>	Glucose specific permease subunit
EII <sup>Fru</sup>	Fructose-specific phosphotransferase system (encoded by <i>ptsF</i> )
EII <sup>Glc</sup>	Glucose-specific phosphotransferase system (encoded by <i>ptsG</i> )
EII <sup>Scr</sup>	Sucrose-specific phosphotransferase system (encoded by <i>ptsS</i> )
F1P	Fructose-1-phosphate
F6P	Fructose-6-phosphate
FBP	Fructose-1,6-biphosphate
Fbp	Fructose-1,6-biphosphatase
Fru	Fructose
G6P	Glucose-6-phosphate
Glk	Glucokinase
Glu	Glucose
Gnt	Gluconate
Gnt6P	6-phosphogluconate
GntK	Gluconate kinase
GntP	Gluconate permease
GTP	Guanosine-5'-triphosphate
HPr	Heat-stable phosphocarrier protein (encoded by <i>ptsH</i> )
IEC	Ion-exchange chromatography
IEF	Isoelectric focusing
IF	Initiation factor
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
mRNA	Messenger RNA
MS	Mass spectrometry

## Abbreviations

---

MU	Miller units
NC	Nascent peptide-chain
ORF	Open reading frame
PCR	Polymerase chain reaction
PDB	Protein data bank
PEP	Phosphoenolpyruvate
PfkA	6-phosphofructokinase
PfkB	Fructose-1-phosphate kinase
Pgi	Phosphoglucose isomerase
PpgK	Polyphosphate dependent glucokinase
PPP	Proline triplet
PTC	Peptidyl-transferase center
PTM	Post-translational modification
PTS	Phosphoenolpyruvate-dependent sugar phosphotransferase system
R5P	Ribose-5-phosphate
RBS	Ribosome binding site
RbsK1	Ribokinase 1
RbsK2	Ribokinase 2
REC	Recombinant
RED	Undecylprodigiosin
RF	Recycling factor
Rib	Ribose
RNA	Ribonucleic acid
RNAP	RNA polymerase
RRF	ribosome recycling factor
rRNA	Ribosomal RNA
RT-qPCR	Quantitative reverse transcription PCR
S6P	Sucrose-6-phosphate
ScrB	Sucrose-6-phosphate hydrolase
SEC	Size-exclusion chromatography
Suc	Sucrose
TCE	Trichloroethanol
tRNA	Transfer RNA
tRNA <sup>fMet</sup>	Initiator tRNA
ValS	Val-tRNA synthetase
WT	Wild type
XPPX	Polyproline motif

## Publications and Manuscripts Presented in this Thesis

### Chapter 1.1, 1.4, 2.1 – 2.3, 2.5 – 2.8:

**Pinheiro, B.**, Scheidler, C.M., Kielkowski, P., Schmid, M., Forne, I., Ye, S., Reiling, N., Takano, E., Imhof, A., Sieber, S.A., Schneider, S., Jung, K. (2020). Structure and Function of an Elongation Factor P Subfamily in Actinobacteria. Cell Reports 30, 4332-4342 e4335.

### Chapter 2.9 – 2.12:

**Pinheiro, B.**, Petrov, D. P., Martins, G. B., Bramkamp, M., Jung, K. (2020). Elongation factor P is required for EII<sup>Glc</sup> translation in *Corynebacterium glutamicum* due to an essential polyproline motif (working title). Molecular Microbiology, *Under review*.

### Chapter 2.14 – 2.18:

**Pinheiro, B.\***, Bilik, O.\*, Ye, S.\*, Takano, E., Jung, K. Influence of the Elongation Factor P in Actinobacteria amino acid and secondary metabolite production (working title). *In Preparation*.

\* Equal contributions.

## Contributions to Publications and Manuscripts presented in this Thesis

### Chapter 1.4 and 1.4

Parts of these chapters are already published in (Pinheiro et al., 2020).

### Chapter 2.1 – 2.3 and 2.5 – 2.8:

Bruno Pinheiro, Sabine Schneider and Kirsten Jung designed the experiments. Bruno Pinheiro constructed all *C. glutamicum* strains and plasmids. Bioinformatics, reporter assay for EF-P activity, microscopy imaging, fluorescence quantification and Isoelectric focusing of *C. glutamicum* proteins were done by Bruno Pinheiro. Bacterial strains cultivation, protein expression and purification were carried out by Bruno Pinheiro and further analyzed by Ignasi Forné (proteomics), Pavel Kielkowski (intact and digested MS) and Christopher Schleider (crystallization). *Streptomyces coelicolor* A3(2) *efp*-6HIS was constructed by Suhui Ye and Eriko Takano. Protein expression, purification and IEF gels of EF-P proteins from *Streptomyces* and *Mycobacterium* were performed by Marina Schmid in supervision of Bruno Pinheiro. *Mycobacterium tuberculosis* growth and inactivation was done by Norbert Reiling. Bruno Pinheiro, Sabine Schneider and Kirsten Jung wrote the manuscript (Pinheiro et al., 2020).

### Chapter 2.9 – 2.12:

Bruno Pinheiro, Dimitar Petrov, Marc Bramkamp and Kirsten Jung designed the experiments, Bruno Pinheiro constructed all strains and plasmids except by *pk19-mNG-ptsG*,  $\Delta ptsG$  and  $\Delta ptsG \Delta iolT1 \Delta iolT2$  strains. All experimental assays were carried out by Bruno Pinheiro except by radiolabeled glucose transport measurements done by Dimitar Petrov. Bruno Pinheiro and Kirsten Jung wrote the manuscript.

### Chapter 2.14 – 2.18:

Bruno Pinheiro, Suhui Ye, Eriko Takano and Kirsten Jung conceptualized the project. *C. glutamicum* growth, EF-P protein and mRNA quantification; bioinformatics experiments and *S. coelicolor* EF-P protein purification and quantification were done by Bruno Pinheiro. *S. coelicolor* growth and secondary metabolite production were quantified by Suhui Ye and Oksana Bilik. Figures were designed and manuscript was written by Bruno Pinheiro and Kirsten Jung.

We hereby confirm the above statements:

Bruno Pinheiro

Kirsten Jung

### Summary

The incorporation of consecutive prolines into a nascent peptide chain triggers ribosome stalling and requires rescue via a specific, universally conserved, translation elongation factor – bacterial EF-P 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. Actinobacteria, which includes the genera *Corynebacterium*, *Streptomyces* and *Mycobacterium* is of both medical and economical significance. This work focuses on the EF-P proteins from the most prominent Actinobacteria. Structure, mechanism of action, expression and specific physiological effects in the cells are explored.

Proteomic analysis of *C. glutamicum* revealed that EF-P is required for the translation of proteins involved in amino acid and secondary metabolite production. Analysis of the protein functionality, molecular mass, charge and X-ray crystal structure showed evidence that Actinobacterial EF-P does not require any post-translational modification for activation. While the function and overall 3D structure of this EF-P type is conserved, the  $\beta 2\Omega\beta 3$  loop containing the conserved lysine32 is flanked by two essential prolines (P30 and P34) that rigidify it. Amino acid substitutions that increased the flexibility of this loop completely inhibit EF-P function. This new EF-P subfamily is found in 11% of all bacteria.

The genome of *C. glutamicum* encodes 1,468 polyproline motifs, many of which are found in transcriptional regulators and carbohydrate transporters. The requirement of EF-P for translation of carbohydrate transporters can be a bottleneck in the performance of industrial strains. The EII<sup>Glc</sup> - glucose-specific permease of the phosphoenolpyruvate (PEP): sugar phosphotransferase system (PTS) - encoded by *ptsG*, is the most important glucose uptake system of *C. glutamicum*. Due to the presence of a polyproline, production of EII<sup>Glc</sup> is strongly reduced in the  $\Delta efp$  mutant strain, resulting in a lower glucose uptake rate and growth defect. The polyproline motif is essential for functionality and could not be replaced by any other amino acid sequence in an unbiased random mutagenesis. GntR2, a transcriptional activator of *ptsG*, is also dependent on EF-P. However, its reduced copy number can be compensated for by other regulators.

Analysis of *C. glutamicum* and *S. coelicolor* cell lysates revealed that EF-P is upregulated in the early stationary phase, a time point in which Actinobacteria produce and secrete large amounts of amino acids and other secondary metabolites. Evolutionary analysis of polyproline motifs in the proteome of *S. coelicolor* provided evidence of evolutionary selection against translational stalling in the core proteome, but not in the enzymes responsible for secondary metabolite production. Proteins essential for antibiotics production contain up to 15 polyproline motifs and without EF-P, production of coelimycin P1, undecylprodigiosin and actinorhodin were only a fraction of the parental strain.

This thesis provides evidence that EF-P is active without post-translational modifications in the most prominent Actinobacteria. Moreover, it acts as a bottleneck in the translation of carbohydrate transporters and enzymes of secondary metabolism.

# Zusammenfassung

Der Einbau konsekutiver Proline in eine entstehende Peptidkette führt zu einem ungewollten Stopp der Translation am Ribosom. Dieser Arrest kann aufgehoben werden durch einen spezifischen, universell konservierten Translations-Elongationsfaktor - bakterielles EF-P und archäales/eukaryotisches a/eIF5A. Bei Eukaryoten, Archaeen und allen bisher untersuchten Bakterien hängt die Funktionalität des Translations-Elongationsfaktors von spezifischen und eher ungewöhnlichen post-translationalen Modifikationen ab. Actinobacteria, zu denen die Gattungen *Corynebacterium*, *Streptomyces* und *Mycobacterium* gehören, sind sowohl von medizinischer als auch von wirtschaftlicher Bedeutung. Diese Arbeit fokussiert sich auf die EF-P-Proteine der prominentesten Actinobacteria, untersucht werden ihre Struktur, ihren Funktionsmechanismus, Expression und spezifische physiologische Effekte.

Die Proteomanalyse von *C. glutamicum* ergab, dass EF-P für die Translation von Proteinen erforderlich ist, welche an der Produktion von Aminosäuren und Sekundärmetaboliten beteiligt sind. Die Eigenschaften der posttranslationalen Modifikation von EF-P wurde in *Corynebacterium glutamicum*, *Streptomyces coelicolor* und *Mycobacterium tuberculosis* untersucht. Die Analyse der Proteinfunktionalität, der molekularen Masse, der Ladung und der Röntgenkristallstruktur führte zu Hinweisen, dass in Actinobakterien EF-P keine posttranslationale Modifikation zur Aktivierung benötigt. Während die Funktion und die gesamte 3D-Struktur dieses EF-P-Typs konserviert ist, wird die  $\beta 2\Omega\beta 3$  Schleife, die das konservierte Lysin<sup>32</sup> enthält, von zwei essentiellen Prolinen (P30 und P34) flankiert, die deren Flexibilität einschränken. Aminosäuresubstitutionen, die die Schleife flexibler machen, hemmen die EF-P-Funktion vollständig. Diese neue EF-P-Unterfamilie ist in 11 % aller Bakterien zu finden.

Das Genom von *C. glutamicum* kodiert für 1.468 Polyprolin-Motive, von denen viele in Transkriptionsregulatoren und Kohlenhydrattransportern zu finden sind. Die Notwendigkeit von EF-P für die Translation von Kohlenhydrattransportern kann ein Engpass in der Proteinproduktion sein, der die Leistung industrieller Stämme beeinträchtigt. EII<sup>Glc</sup> – die Glukose-spezifische Permease des Phosphoenolpyruvats (PEP): Zucker-Phosphotransferase-System (PTS) – kodiert durch *ptsG*, ist das wichtigste Glukoseaufnahmesystem von *C. glutamicum*. Da ein Polyprolin vorhanden ist, ist die Produktion von EII<sup>Glc</sup> im *efp* Deletionsstamm stark reduziert, was zu einer geringeren Glukoseaufnahmerate und einem Wachstumsdefekt führt. Das Polyprolin-Motiv ist für die Funktionalität essentiell und konnte in einer „unbiased random mutagenesis“ nicht durch eine andere Aminosäuresequenz ersetzt werden. GntR2, ein Transkriptionsaktivator von *ptsG*, ist ebenfalls EF-P-abhängig. Eine reduzierte Kopienzahl von GntR2 kann jedoch durch andere Regulatoren kompensiert werden.

Die Analyse der Zelllysate von *C. glutamicum* und *S. coelicolor* zeigte, dass die Menge an EF-P in der frühen stationären Phase hochreguliert wird, einem Zeitpunkt, zu dem Actinobakterien bekanntlich große Mengen an Aminosäuren und anderen Sekundärmetaboliten produzieren und sezernieren. Die evolutionäre Analyse von Polyprolin-Motiven im Proteom von *S. coelicolor* ergab Hinweise auf eine evolutionäre Selektion gegen Arrestmotive im Kernproteom, nicht aber in den Enzymen, die für die Produktion von Sekundärmetaboliten verantwortlich sind. Proteinkomplexe, die für die Antibiotika-Produktion essentiell sind, enthalten bis zu 15 Polyprolin-Motive. In einer *efp* Deletionsmutante war die Produktion von Coelimycin P1, Undecylprodigiosin und Actinorhodin äußerst stark reduziert.

Diese Arbeit liefert Hinweise darauf, dass EF-P ohne posttranslationale Modifikationen in den bekanntesten Actinobacteria aktiv ist. Außerdem bildet EF-P einen Engpass bei der Translation von Kohlenhydrattransportern und Enzymen des Sekundärstoffwechsels.

# 1 Introduction

## 1.1 Actinobacteria and their relevance

Actinobacteria constitutes one of the largest phylum among bacteria (Bergey et al., 2012). They are facultative or obligate anaerobes with a rigid cell wall containing muramic acid. The DNA C+G content can be over 70 mol% in some species (Bergey et al., 2012). Evolutionary, the divergence of Actinobacteria from other bacteria is ancient, making it impossible to identify the phylogenetically closest bacterial group (Lewin et al., 2016; Servin et al., 2008; Ventura et al., 2007). Furthermore, members of the Actinobacteria have adopted different lifestyles. They are widely distributed in aquatic and terrestrial habitats (*Streptomyces*), some are pathogens (e.g., *Corynebacterium*, *Mycobacterium*, *Gardnerella*), plant commensals (*Leifsonia*), gastrointestinal commensals (*Bifidobacterium*), etc. (Bergey et al., 2012). From morphology to molecular markers, the members of the phylum Actinobacteria are so diverse that do not have any unique characteristic as common shared to all species assigned to this phylum (Gao and Gupta, 2005, 2012).

Actinobacteria is the most economically significant as well as biotechnologically valuable phylum in the domain Bacteria. Explored in industry for naturally producing over 16.000 bioactive compounds and potential for unnumbered more, including amino acids and its derivatives, antibiotics, cytostatic, fungicides, modulators of immune responses, siderophores, odor compounds and many more (Bergey et al., 2012; Lewin et al., 2016; Manteca, 2019). *Corynebacterium glutamicum* and *Streptomyces coelicolor* are among the most prominent species of this phyla due to the substantial production of secondary metabolites and genetically availability for strain engineering.

### 1.1.1 *Corynebacterium glutamicum*

Originally isolated from sewage in a screening for organisms secreting L-glutamate (Udaka, 1960), *C. glutamicum* is, now a days, largely used in industry as a work-horse in the production of L-glutamate, L-lysine and derived amino acids (Becker et al., 2018; Eggeling and Bott, 2015; Heider and Wendisch, 2015; Sanchez et al., 2017). Furthermore, it is been successfully engineered to produce plant-derived aromatic compounds (Kallscheuer et al., 2017; Kallscheuer et al., 2016), diamines (Meiswinkel et al., 2013; Peters-Wendisch et al., 2014), carotenoids (Henke et al., 2018; Peters-Wendisch et al., 2014) and biofuels (Jojima et al., 2015; Siebert and Wendisch, 2015; Xiao et al., 2016) just to name a few. *C. glutamicum*'s biosafety, genetically availability and fast growth allows this bacterium to be a model for studies in Actinobacterial membrane synthesis and cell division, pathogenicity of *Corynebacteriales* to humans and animals, metabolic engineering and many others.

In contrast to most bacteria, *C. glutamicum* simultaneously metabolizes different carbon sources (Moon et al., 2007; Wendisch et al., 2000). Industrial fermentation in molasses and starch hydrolysates results in growth without diauxic behavior. Moreover, cultivation in simple carbohydrates mixtures or simply glucose, cheapen the fermentation process as well as further facilitates purification of the metabolites secreted (Becker et al., 2018; Kallscheuer et al., 2019; Kogure and Inui, 2018; Wendisch et al., 2016a; Wendisch et al., 2016b). For this reason, when engineering *C. glutamicum* strains towards large-scale metabolite production, carbohydrate transporters such as EII<sup>Glc</sup> and its transcriptional regulators are the main targets (Krause et al., 2010; Lindner et al., 2013; Xu et al., 2016).

### 1.1.2 *Streptomyces coelicolor*

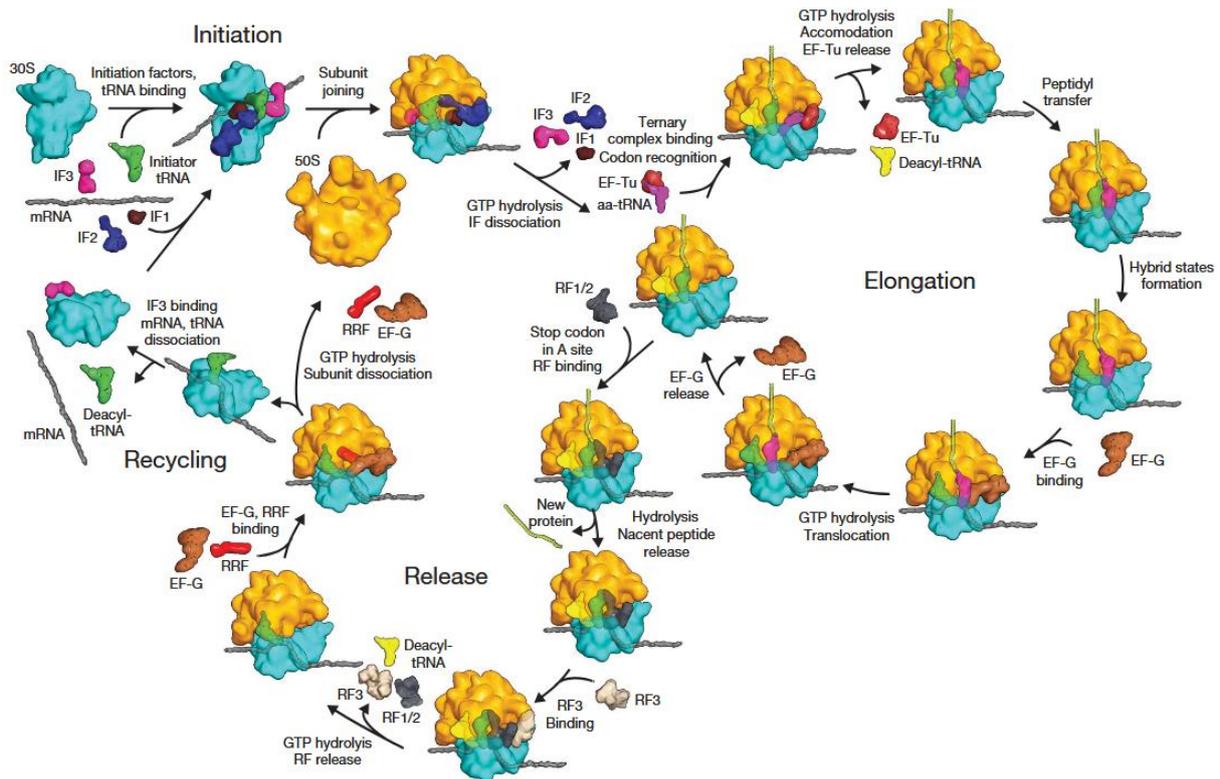
*Streptomyces* species have been the most important source of bioactive compounds that society is routinely using in clinic and agriculture. They are responsible for native production of more than 12,400 compounds of all sorts: antimicrobials, insecticides, immunosuppressant, antitumoral, just to name a few (Bergey et al., 2012; Manteca, 2019). They have a complex life cycle alternating between vegetative mycelium-like growth, aerial mycelium bearing chains of arthrospores and the arthrospores themselves (Bergey et al., 2012). In *S. coelicolor*, as well as in other *Streptomyces* and some other actinobacterial species, the early stage of the stationary phase is accompanied by a major metabolic switch resulting in activation of secondary metabolism, including production and secretion of many antibiotics and, further on, sporulation (Alam et al., 2010; Nieselt et al., 2010; Strauch et al., 1991).

*S. coelicolor* laboratory strain A(2), produce diverse antibiotics including coelimirin P1 (CPK), undecylprodigiosin (RED) and actinorhodin (ACT) (van Keulen and Dyson, 2014). Moreover, *S. coelicolor* A(2) genome codes for more than 20 secondary metabolite gene clusters (Bentley et al., 2002). Complex life cycle and secondary metabolism outline this bacterium as a model organism for cell division and differentiation, secondary metabolite production and antibiotics resistance.

## 1.2 Translation in bacteria

All bacterial cell activities such as growth and production of secondary metabolites, relies ultimately on the biosynthesis of several enzymes, transporters and structural proteins through the process of translation. In this process, a coding mRNA sequence is translated into the amino acid sequence of a protein by a complex, dynamic and highly conserved mechanism catalyzed by the ribosomal machinery. In bacteria, the ribosome consists of a large 50S and a small 30S subunit. The large subunit comprises the 23S RNA, 5S RNA and about 30 proteins while the small subunit, the 16S RNA and about 20 proteins (Alberts et al., 2002). This process of protein biosynthesis can be roughly divided into three stages: initiation, elongation and termination (further subdivided into release and ribosomal recycling) (Figure 1.1) (Alberts et al., 2002; Rodnina, 2018; Schmeing and Ramakrishnan, 2009).

## Introduction



**Figure 1.1** – Simplified translation mechanism in bacteria divided into four steps: initiation, elongation, release and recycling. IF *initiation factor*, EF *elongation factor*, RF *recycling factor*, GTP *guanosine-5'-triphosphate*, aa-tRNA *aminoacyl-tRNA*, RRF *ribosome recycling factor*. Adapted from (Schmeing and Ramakrishnan, 2009).

Due to the lack of nucleic compartmentalization in prokaryotes, translation initiate as soon as the ribosome binding site (RBS) of the leading mRNA is transcribed by the RNA polymerase (RNAP), permitting the assembly of the ribosomal machinery onto the nascent mRNA (Burmam et al., 2010; Proshkin et al., 2010). Therefore, no mRNA quality control such as capping, polyadenylation or splicing – common characteristics described for eukaryotes – are described in bacteria (Doma and Parker, 2007). Briefly, the 30S ribosomal subunit is positioned in close vicinity to the mRNA start codon by base pairing a close-complementary sequence of the 16S and the Shine- Dalgarno sequence in the mRNA. The initiation factors IF3 is also involved in this binding. Further association of IF1, IF2 and initiator tRNA ( $tRNA^{fMet}$ ) form the 30S initiation complex. IF2 further promotes joining of the 50S subunit to form the 70S initiation complex. The  $tRNA^{fMet}$  moves to the peptidyl transferase center (PTC) and the elongation starts (Alberts et al., 2002; Ramakrishnan, 2002; Rodnina, 2018; Schmeing and Ramakrishnan, 2009).

Elongation is the step in which the peptide chain is synthesized according to the mRNA codon sequence. For that, the ternary complexes, consisting of an aminoacyl-tRNA (aa-tRNA), elongation factor EF-Tu and GTP diffuses into the 70S aminoacyl-tRNA binding site (A-site). If the t-RNA anticodon perfectly pairs with the mRNA codon in the decoding center of the A-site, GTP is hydrolyzed favoring the peptidyl transferase reaction from the t-RNA in the peptidyl-tRNA binding site (P-site) to the aa-tRNA in the A-site, releasing the EF-Tu. EF-G catalyzes the translocation of the ribosomal complex by a distance of

one codon along the mRNA, moving the deacylated t-RNA to the ribosomal exit site (E-site). Release of EF-G allows the next elongation cycle (Alberts et al., 2002; Ramakrishnan, 2002; Rodnina, 2018; Schmeing and Ramakrishnan, 2009).

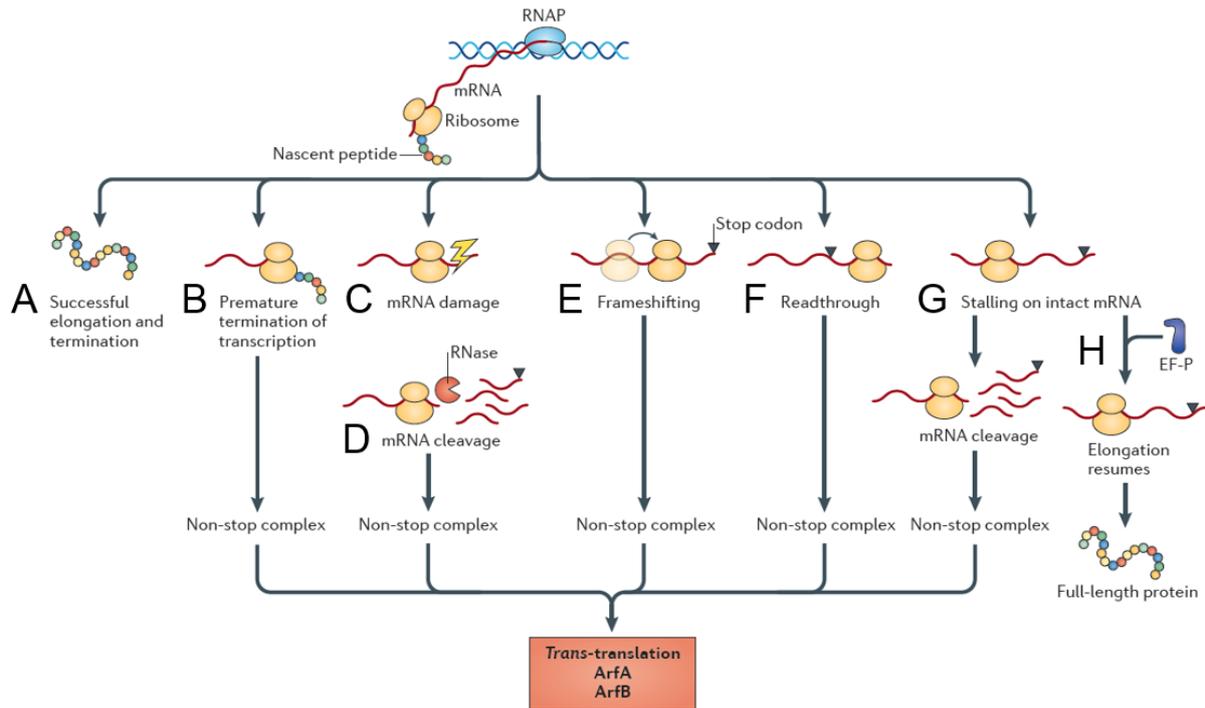
The elongation is the longest and most energetically-costly step of translation as well as requires around 50% of the total energy produced by a growing bacterial cell (Russell and Cook, 1995). Nevertheless, several events during this process might cause protein elongation to slow or stop. Usage of rare codons or Shine-Dalgarno-like sequences within coding regions reduce ribosome speed (Li et al., 2012; Mitarai et al., 2008; Sorensen and Pedersen, 1991). Other factors such as mRNA integrity, RNAP and ribosome fidelity and sequence of the translating peptide trigger ribosome stalling and rescue by one of the several ribosomal rescue systems (Keiler, 2015; Peil et al., 2013; Ude et al., 2013) (see below).

When the stop codon is reached by the translating ribosomal 70S complex and enters the decoding center of the ribosome, it is recognized by one of the release factors (RF1 or RF2) catalyzing the hydrolysis of the peptidyl-tRNA bond and release of the newly synthesized peptide. This characterizes the termination stage. Ribosome recycling factor and EF-G promotes dissociation of the ribosomal subunits, tRNA and mRNA, completing the translational cycle (Alberts et al., 2002; Ramakrishnan, 2002; Rodnina, 2018; Schmeing and Ramakrishnan, 2009).

### 1.3 Ribosome stalling

Many perturbations during translational elongation can trigger ribosome stalling (Figure 1.2) (Keiler, 2015; Rodnina, 2018). Premature termination of mRNA transcription by RNAP lead to synthesis of truncated mRNA coding for incomplete proteins and without stop codon (Figure 1.2B). Damage of the mRNA by external factors (Figure 1.2C), RNase activity (Figure 1.2D), shifts in the open-reading frame caused by slippery codons or readthrough (Figure 1.2E and F) results in ribosome stall due to the absence of codon in the decoding center (non-stop complex) (Keiler, 2015; Rodnina, 2018). In those cases, the ribosomal rescue trans-translation system or the alternative ribosome-rescue factor A (ArfA) or B (ArfB) are recruited and allow ribosome recycling (Figure E) (Keiler, 2015; Rodnina, 2018).

Ribosomes can stall, as well, on intact mRNAs for diverse reasons (Figure 1.2G). Shine-Dalgarno-like sequences within coding regions can anneal with the 16S rRNA interrupting ribosomal translocation (Li et al., 2012). Amino acid starvation triggers accumulation of non-charged tRNAs in the ribosomal A-site with no substrate for elongation (Loveland et al., 2016; Richter, 1976; Wendrich et al., 2002). Translation of peptides containing consecutive proline codons trigger stalling due to a sterically clash between the nascent peptide and the ribosome exit tunnel, impeding incorporation of downstream amino acids (Huter et al., 2017).



**Figure 1.2** – Causes of ribosome stalling in bacteria and ribosome rescue. **(A)** Most ribosomes elongate and terminate peptides successfully, producing a full-length protein. The absence of stop codons in mRNAs caused by **(B)** premature termination of the transcription, **(C)** mRNA damage by external factors or **(D)** RNase activity leads to the formation of a non-stop complex and usually rescued by trans-translation or by alternative pathways mediated by alternative ribosome-rescue factor A (ArfA) or B (ArfB). **(E)** Frameshifting and **(F)** readthrough of the stop codon can also produce non-stop complexes. **(G)** Ribosomes can stall at intact mRNA (details in the text), leading to premature termination or **(H)** rescue by EF-P. RNAP *RNA polymerase*. Adapted from (Keiler, 2015).

Proline is the only proteogenic amino acid that possesses a pyrrolidine ring including the C<sub>α</sub> and the amino group. This unusual connection limits the conformation of the amino acid, restricting flexibility and lowering the reactivity for the peptide bond formation (Pavlov et al., 2009; Schimmel and Flory, 1968; Venkatachalam and Ramachandran, 1969). Moreover, when consecutive prolines are incorporated in the nascent peptide, they adopt an unusual conformation in the ribosome exit tunnel and due to steric constraints, prevent the peptidyl transferase reaction from the peptidyl-tRNA<sup>Pro</sup> in the P-site to the Pro-tRNA in the A-site (Huter et al., 2017). Ribosomal translational speed remarkably drops during translation of polyproline motifs (XPPX), triggering ribosome stalling, translation interruption and release of truncated proteins (Doerfel et al., 2013; Peil et al., 2013; Ude et al., 2013). Ribosomes stalled at polyproline stretches are rescued by a specific release mechanism mediated by the translation elongation factor P (EF-P) (Figure 1.2H) (Huter et al., 2017; Ude et al., 2013).

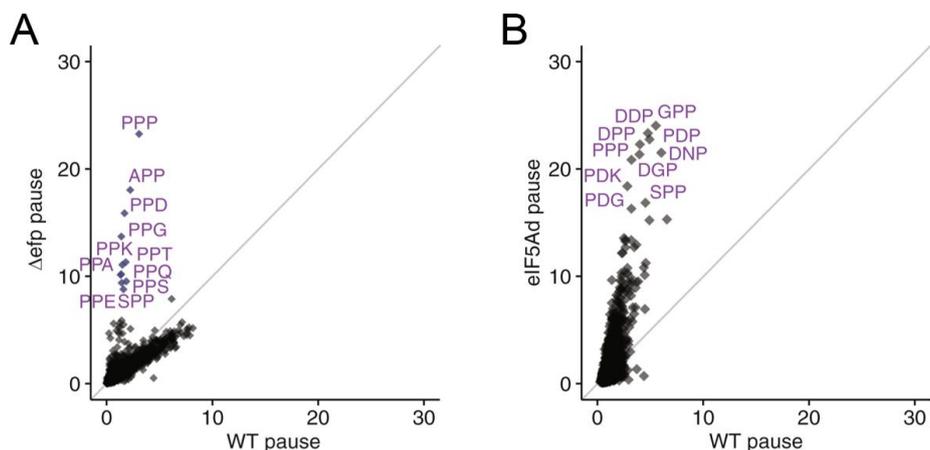
### 1.4 Translation elongation factor P

#### 1.4.1 Function

EF-P is a universally conserved translation elongation factor orthologue to the eukaryotic eIF5A and archaea aIF5A (Kyripides and Woese, 1998; Lassak et al., 2016; Rossi et al., 2014). It binds to stalled ribosomes favoring Pro-Pro bound formation and production of full-length proteins containing polyproline motifs (Doerfel et al., 2013; Ude et al., 2013). EF-P activity in maintenance of proline codon fidelity is also shown (Gamper et al., 2015).

Research on the EF-P function date from 1975 when Glick and Ganoza described positive effects of EF-P during translation initiation and elongation (Ganoza and Aoki, 2000; Glick et al., 1979; Glick and Ganoza, 1975). EF-P was formally mistaken as an essential protein for *Escherichia coli*, when it was firstly crystalized (Aoki et al., 1997). Recent studies show evidence that EF-P binds to ribosomes during most of elongation cycles (Mohapatra et al., 2017), has some effect in maintenance of proline and other codons fidelity probably through alleviating ribosome stalling (Gamper et al., 2015; Naganathan et al., 2015; Smith et al., 2019). However, it has a remarkable effect in proteostasis of di- and polyproline-containing proteins, without major effects on translation of other motifs (Figure 1.3A). Evidence that EF-P facilitates translation of XPPX motifs is based in many 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 (Elgamal et al., 2014; Woolstenhulme et al., 2015). To the contrary, the eukaryotic counterpart eIF5A is reported to have a more general positive effect in the translation of many amino acid motifs rather than only polyprolines (Figure 1.3B) (Schuller et al., 2017), and it is essential for eukaryotes viability (Dever et al., 2014; Gabel et al., 2013).

The stalling effect of XPPX motifs on ribosomes and rescue via EF-P raised the question why evolution favored the emergence of such specialized translational elongation factor, rather than the select against XPPX motifs in bacterial proteomes (Starosta et al., 2014b). A proline-triplet motif is found in the catalytic pocket of the Val-tRNA synthetase ValS. As this motif is conserved among all kingdoms and essential for ValS function (Starosta et al., 2014b), it was then proposed that EF-P was first selected to facilitate the translation of ValS, and later involved in the translation of other XPPX-containing proteins in each bacteria (Starosta et al., 2014b).



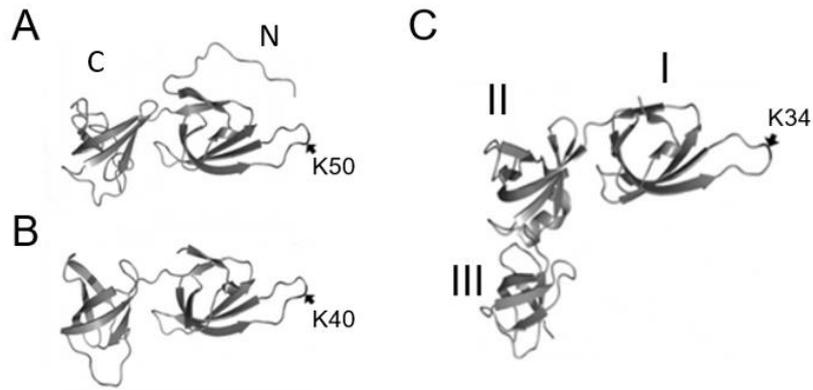
**Figure 1.3** - Effect of EF-P and the eukaryotic counterpart eIF5A in their respective host proteomes by ribosome profiling. **(A)** Comparison of pausing motifs in *E. coli* WT and  $\Delta efp$  cells. According to Woolstenhulme et al., 2015 dataset. **(B)** Comparison of tripeptide pausing in *S. cerevisiae* WT and eIF5A-depleted (eIF5Ad) cells. Each dot represents one tripeptide motif. The diagonal line indicates the distribution expected for no enrichment. Adapted from (Schuller et al., 2017).

Besides the catalytic function in ValS, polyproline motifs are secondary structure disrupters, inducing turns and kinks that favor protein interaction with biomolecules, other proteins and nucleic acids (Adzhubei et al., 2013; MacArthur and Thornton, 1991; Motz and Jung, 2018; Vanhoof et al., 1995). Characteristics that might explain why many bacteria encode for a large number of XPPX motifs in their proteomes (Qi et al., 2018; Starosta et al., 2014b).

Biochemical assays and ribosome profiling of *efp* deletion mutants defined a hierarchy of pausing motifs (Peil et al., 2013; Starosta et al., 2014a; Woolstenhulme et al., 2015). Depending on the amino acid context upstream (-1 and -2) and downstream (+1) of a polyproline motif, it can trigger ribosome stalling more frequently (Peil et al., 2013; Starosta et al., 2014a; Woolstenhulme et al., 2015). Classification of all XPPX motifs as weak, moderate or strong ribosome stallers can predict how important the EF-P rescue system is for production of a given protein according to the amino acid sequence (Qi et al., 2018).

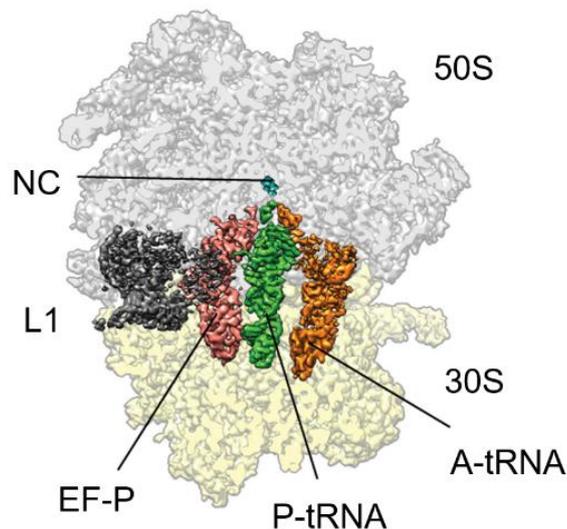
#### 1.4.2 General structure and interactions with the ribosome

The structure of EF-P consists of three  $\beta$ -barrel domains with an overall L-shape similar to a tRNA (Hanawa-Suetsugu et al., 2004). While EF-P domain I is topologically the same as the domain N of e/aiF5A, the domain II and III share the same topology of the e/aiF5A C domain, indicating that domain II and III arose in bacteria from duplication (Figure 1.4) (Hanawa-Suetsugu et al., 2004).



**Figure 1.4** – EF-P structure similarity between (A) Eukarya eIF5A from *Homo sapiens* (Uniprot Q6IS14), (B) Archaea aIF5A from *Sulfolobus acidocaldarius* (GenBank CAA44842) and (C) Bacteria EF-P from *E. coli* (GenBank, AP\_004648). Pointed by arrows, the conserved lysine found at a tip of the active  $\beta 2\Omega\beta 3$  loop. eIF5A and EF-P domains are indicated. Adapted from (Greganova et al., 2011).

Crystal and cryo-EM structures of the EF-P from *T. thermophilus* and *E. coli* in complex with the ribosome show that EF-P binds at the E-site, with its domain I located next to the acceptor stem of the P-site tRNA. The domain III of EF-P binds adjacent to the anticodon stem-loop of the P-site tRNA, while domain II interacts with the highly conserved ribosomal protein L1 (Blaha et al., 2009; Huter et al., 2017) (Figure 1.5). The ribosome-binding dynamics of EF-P and sucrose gradient centrifugation suggests that it binds to ribosomes during most elongation cycles, however binding is transient (Mohapatra et al., 2017; Naganathan et al., 2015).



**Figure 1.5** – Cryo-EM reconstruction of PPP-stalled ribosome complexes with EF-P. Ribosome subunit 30S (yellow) and 50S (gray), A-tRNA *A-site tRNA*, P-tRNA *P-site tRNA*, L1 *ribosomal protein L1* and NC *nascent peptide-chain*. Adapted from (Huter et al., 2017).

The active site of EF-P locates in a loop region connecting two  $\beta$ -strands  $\beta 2$  and  $\beta 3$  in domain I ( $\beta 2\Omega\beta 3$  loop). It interacts with the acceptor arm of the P-site tRNA and protrudes a conserved positive charged amino acid residue at position 32 (either lysine or arginine) into the catalytic pocket of the peptidyl-transferase reaction of the ribosome (Huter et al., 2017). The activity of EF-P in all bacteria investigated thus far, as well as in its eukaryotic and archaeal orthologs, is dependent of an unusual post-translational modification (PTM) that extends the side-chain of the amino acid at position 32 and allows increased interaction with the CCA end of the P-site tRNA (Hummels and Kearns, 2020; Huter et al., 2017; Lassak et al., 2016).

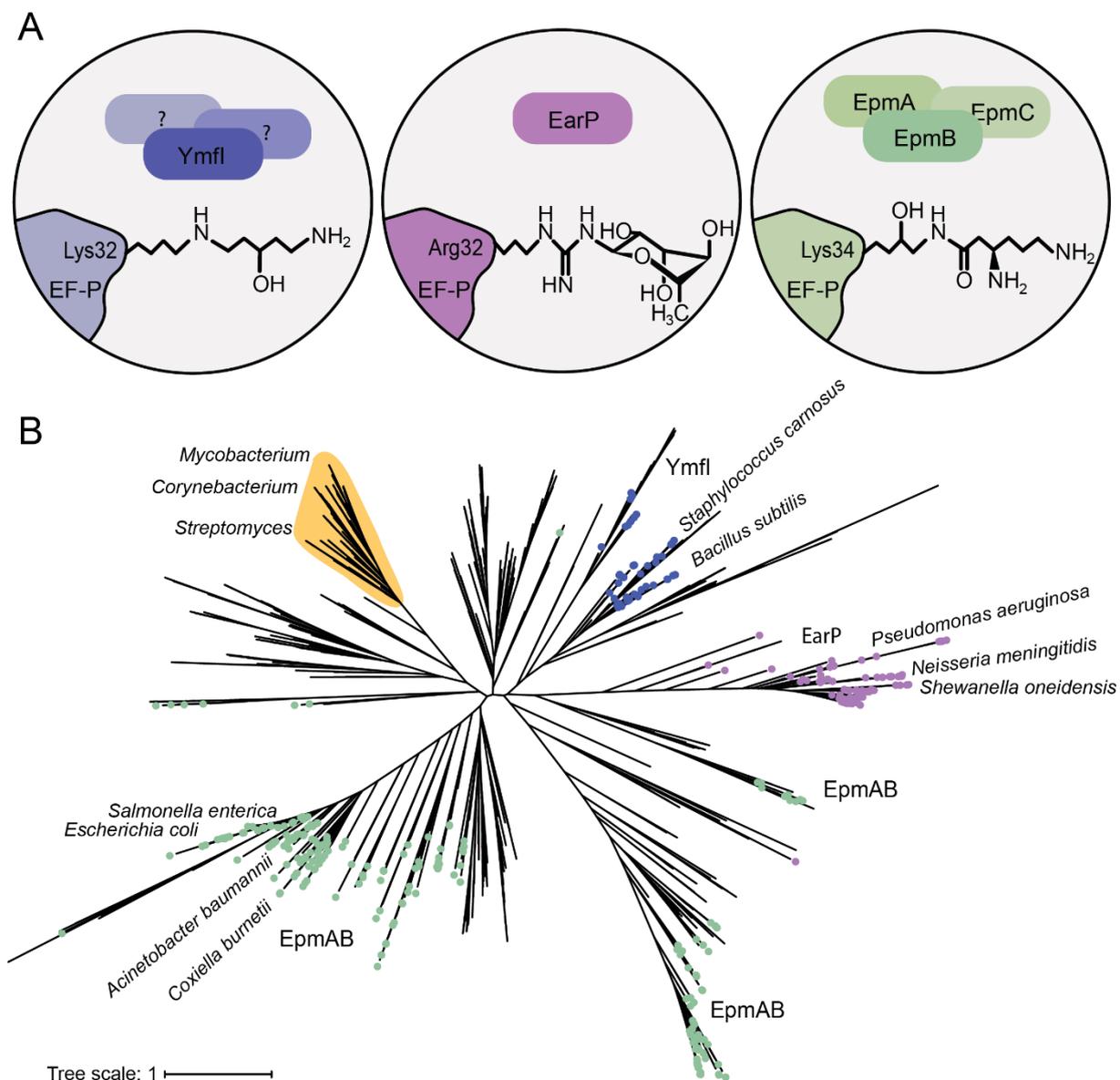
### 1.4.3 Activation by post-translational modification

While both e/alf5A are modified by a unique post-translational modification – hypusinylation of lysine 40/50, bacteria have evolved several ways to modify EF-P (Figure 1.6).  $\gamma$ -Proteobacteria, including *Escherichia coli*, *Salmonella enterica* and 29% of all other reference genomes encode the modification enzymes EF-P-(R)- $\beta$ -lysine ligase EpmA (Roy et al., 2011; Yanagisawa et al., 2010) and L-lysine 2,3-aminomutase EpmB (Yanagisawa et al., 2010) to  $\beta$ -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., 2013; Peil et al., 2012).

Another EF-P subfamily is found in  $\beta$ -proteobacteria and some  $\gamma$ -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. For the other bacteria with no homologous for any of those modification enzymes, the EF-P modification status is, so far, not known.

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). According to the affected proteins and pathways.



**Figure 1.6** – 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)** Phylogenetic 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 Ymfl (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 in bold type.

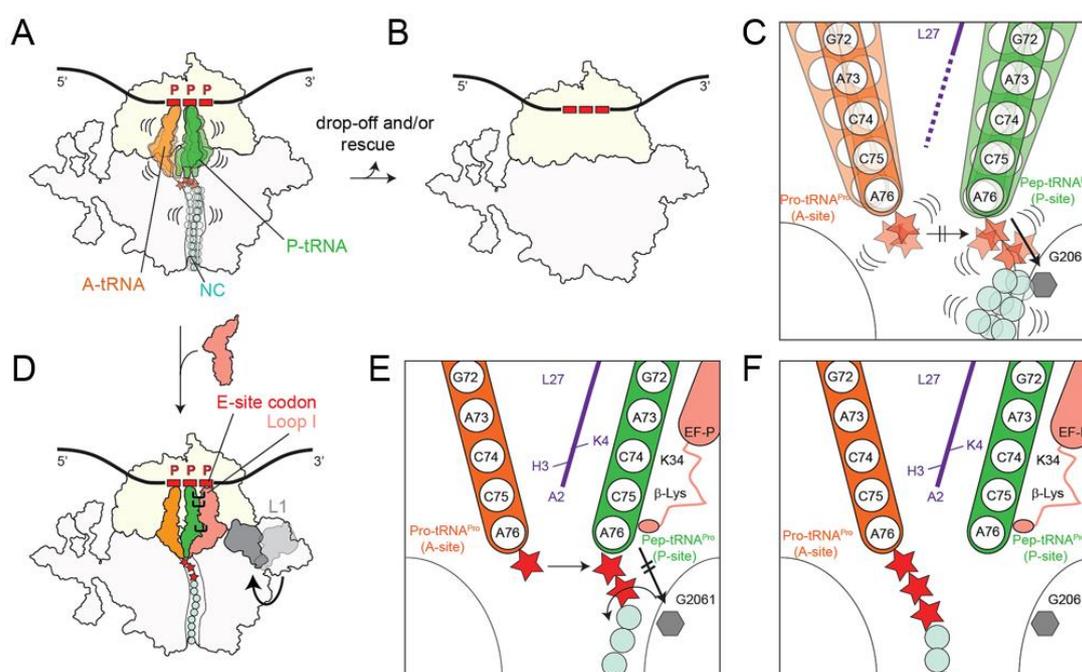
#### 1.4.4 Mechanism of action

The mechanism of ribosome rescue by EF-P is inferred from structural studies – crystal and cryo-EM structures of ribosome in complex with modified and unmodified EF-P or during translation of XPPX motifs (Blaha et al., 2009; Huter et al., 2017) – and biochemical studies on amino acid residues essential

for EF-P activity (Rajkovic et al., 2016; Ude et al., 2013; Volkwein et al., 2019). However, up to date, a full range of motion of this rescue mechanism is still missing.

Translation of a XPPX motif requires a peptidyl-transferase reaction from the proline containing peptidyl-tRNA<sup>Pro</sup> in the P-site, to the Pro-tRNA in the A-site (Figure 1.7A). If no functional EF-P is available, it leads to peptidyl-tRNA drop off and release of truncated proteins (Figure 1.7B). Due to amino acid proline structure, this transferase reaction causes a steric clash between the nascent peptide and the ribosomal 50S unit tunnel wall, specially residue G2061, impeding the accommodation of the A-site tRNA (Figure 1.7C) (Huter et al., 2017; Ude et al., 2013).

If EF-P is available, it binds to the stalled ribosomes in the E-site region and stabilizes the peptidyl-tRNA. This binding is facilitated by contacts with the ribosomal L1 protein (Figure 1.7D). Interactions of the  $\beta 2\Omega\beta 3$  loop with the acceptor arm, and the modified K32 with the CCA end of the P-site tRNA<sup>Pro</sup> force the prolines to adopt an alternative conformation that passes into the ribosomal exit tunnel (Figure 1.7E). Optimal geometry is achieved and peptide bound formation can occur (Figure 1.7F) (Huter et al., 2017).



**Figure 1.7** – Proposed EF-P mechanism of action on polyproline-stalled ribosomes. In details: ribosomes stall during translation of proteins containing polyproline motifs leading to (A) destabilization of the peptidyl-tRNA in the P site (B) Culminating in drop off due to (B) Steric clash between the prolines in the nascent peptide and the residue G2061 of the ribosomal 50S-subunit tunnel wall impeding accommodation of the A-site tRNA and further peptide bound formation (D) EF-P binds to ribosomes in the E-site region and stabilizes the peptidyl-tRNA. This binding is facilitated by contacts with the ribosomal L1 protein, P-site tRNA and E-site codon. (E) Interactions of the  $\beta 2\Omega\beta 3$  loop and the modified K34  $\epsilon(R)$ - $\beta$ -lysyl-hydroxylysine with the CCA end of the P-site tRNA<sup>Pro</sup> forces the prolines to adopt an alternative conformation that passes into the ribosomal exit tunnel. (F) Optimal geometry is achieved, and peptide bound formation can occur. Adapted from (Huter et al., 2017).

The proposed EF-P mechanism, however, fails to answer questions raised by some experimental datasets. For example, how EF-P acts in the maintenance of codon fidelity and what is its connection with ribosomal protein L9 (Gamper et al., 2015; Naganathan et al., 2015); What are the interaction residues between the P-site tRNA and the distinct PTM-molecules that activate EF-P? If a PTM is essential, how can overexpressing unmodified EF-P be enough to promote translation of proteins containing XPPX motifs? (Volkwein et al., 2019); How can the amino acid substitution K32R not cause the same phenotypes observed in the mutant K32A, once both inhibit PTM? (Hummels et al., 2017); Is it possible to have a fully active EF-P that does not require PTM? Overall, we are still in the very beginning of fully understanding all aspects of EF-P in bacteria.

### 1.5 EF-P in Actinobacteria

Not much is known thus far about EF-P proteins in Actinobacteria. Previous studies on *Mycobacterium tuberculosis* strains linked upregulation of *efp* as part of an antibiotic resistance mechanism against ofloxacin (Lata et al., 2015); Downregulation was identified under nutrient limitation (Betts et al., 2002); Saturating transposon mutagenesis studies suggested that *efp* is an essential gene for this bacterium (DeJesus et al., 2017; Sasseti et al., 2003).

Regardless of the relevance of Actinobacteria in human health and biotechnology, the nature of the post-translational modification that activates EF-P as well as global effects of this elongation factor in any member of this phylum is not yet investigated.

### 1.6 Scope of the thesis

EF-P has a crucial role on the translation of proteins containing XPPX motifs. The function of EF-P relies on an unusual post-translational modification at the conserved amino acid residue K32. The main scope of this thesis is to unravel the post-translational modification of EF-P, as well as the mechanism of action in members of the phyla Actinobacteria. Specific effects of EF-P on proteome homeostasis and physiology should also be investigated, as Actinobacteria genomes encode one of the highest frequencies of polyproline motifs among bacteria.

## 2 Results

The amino acid sequences of Actinobacterial EF-P proteins share high identities, clustering them apart of other EF-P sequences in the protein phylogenetic tree (Introduction, Figure 1.6). Thus far, besides the medical and economic importance of this group, neither the post-translational modification status nor the physiological relevance of EF-P in this bacterial phylum is known. By using *C. glutamicum* and *S. coelicolor* as a model organism for EF-P structure, function and actinobacterial physiology, we answer those questions and open a new research pathway towards biotechnological applications.

### 2.1 Global effect of EF-P in *C. glutamicum*

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* (Roy et al., 2011; Yanagisawa et al., 2010), *S. oneidensis* (Lassak et al., 2015) and *B. subtilis* (Rajkovic et al., 2016). Of these, *E. coli* is the one with most XPPX motifs. In total, 2101 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 2.1).

**Table 2.1** – Bacterial strains for which the EF-P PTM-enzymes are studied in detail. The polyproline motifs (XPPX) were classified according to their stalling strengths (Qi et al., 2018). A ratio (motifs/proteins) was calculated by dividing the total number of XPPX motifs by the total number of proteins encoded by each strain.

Strain	Proteome ID	Proteins encoded	Number of XPPX motifs			Motifs/Proteins
			Weak	Moderate	Strong	
<i>Escherichia coli</i> K12 MG1655	UP000000625	4291	1169	447	485	<b>0.49</b>
<i>Shewanella oneidensis</i> MR-1	UP000008186	4068	869	371	313	<b>0.38</b>
<i>Bacillus subtilis</i> 168	UP000001570	4260	605	280	329	<b>0.28</b>

The proteome of the most prominent Actinobacteria was downloaded and had the number of XPPX motifs determined, revealing a strikingly large number of them (Table 2.2). In *M. tuberculosis*, *S. coelicolor* and *Nocardia farcinica* the number of polyproline motifs exceeds the number of encoded proteins (Table 2.2).

## Results

**Table 2.2** – Total number and frequency of polyproline motifs (XPPX) classified according to their stalling strengths in different species of Actinobacteria.

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

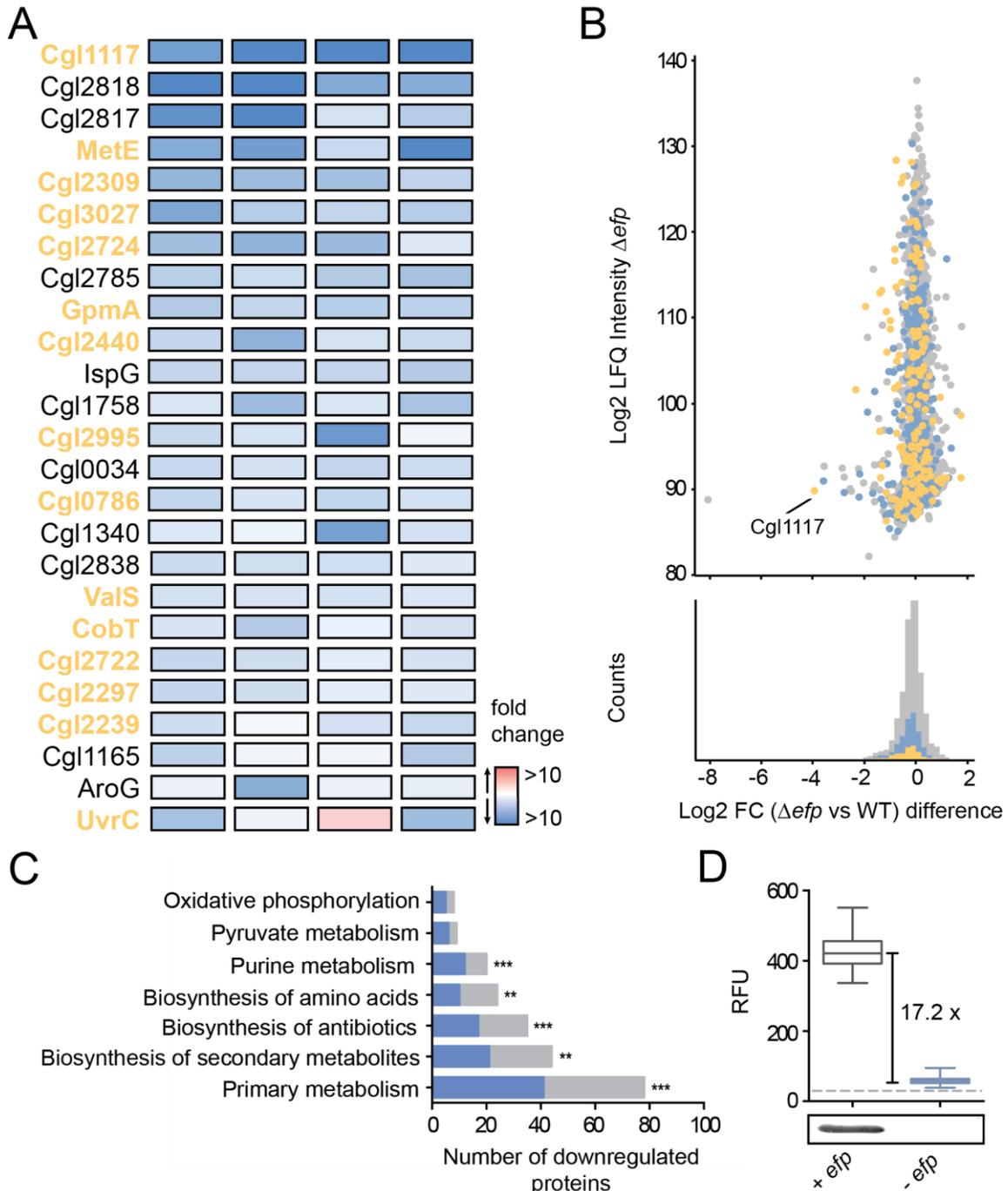
Besides the control by EF-P, the relative amount of an XPPX-containing protein synthesized depends on the rate of translational initiation and the location of the stalling motif(s) it contains (Ude et al., 2013; Woolstenhulme et al., 2015). In addition, transcriptional regulators may contain XPPX motifs, 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, a *efp* deleted mutant of *C. glutamicum* ATCC13032 was constructed and had its proteome compared to that of the parental strain. The analysis of four independent replicates covered 1604 out of 3093 proteins described for the reference strain. In all, 222 proteins were downregulated ( $P$  value <0.05). As expected, the most markedly downregulated proteins contain XPPX motifs in their sequences (Figure 2.1A, yellow marked proteins). The overall intensity of proteins containing polyproline motifs was reduced as well (Figure 2.1B, in blue), and proteins containing strong stalling XPPX motifs have the lowest relative intensity among them (indicated in yellow in Figure 2.1B).

Functional classification of the downregulated proteins revealed seven groups of proteins that were significantly overrepresented (Figure 2.1C). 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 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*<sup>+</sup> strain (Figure 2.1A). This result of the MS analysis was confirmed by testing the level of the chromosomally

## Results

encoded Cgl1117-eGFP fusion protein in the *efp* mutant (Figure 2.1D). 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.



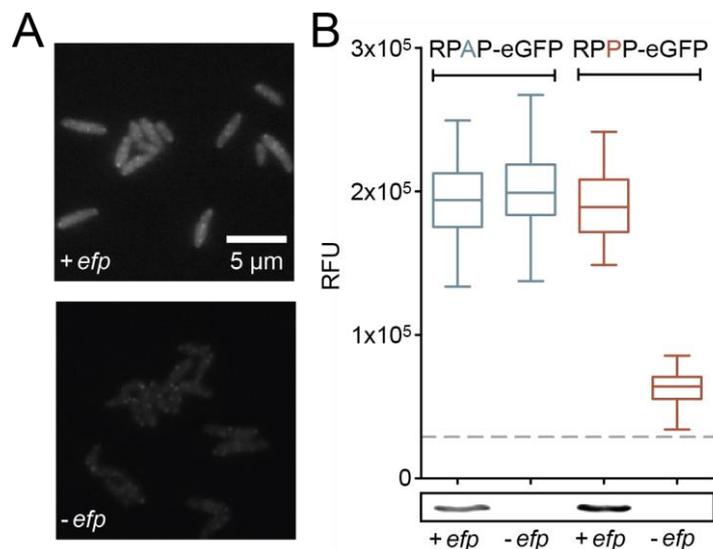
**Figure 2.1** – Proteomic analysis of *C. glutamicum*  $\Delta efp$  strain. **(A)** Heat map representation of the 25 most severely down-regulated proteins in the  $\Delta efp$  mutant relative to the *C. glutamicum* wild type. The fold changes in the four independent replicates are represented by the color gradient. Proteins marked in yellow contain polyproline motifs. EF-P itself as the most markedly downregulated protein is not shown on the heat map. **(B)** Scatter-plot of protein fold-change ratio (log<sub>2</sub>-transformed) relative to the summed

## Results

protein intensity of all biological replicates of the  $\Delta efp$  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-*efp* cells in *efp*<sup>+</sup> and *efp*<sup>-</sup> strains. Dashed lines mark the background fluorescence. Production of EF-P was confirmed by Western blot analysis.

### 2.2 Design of a reporter strain

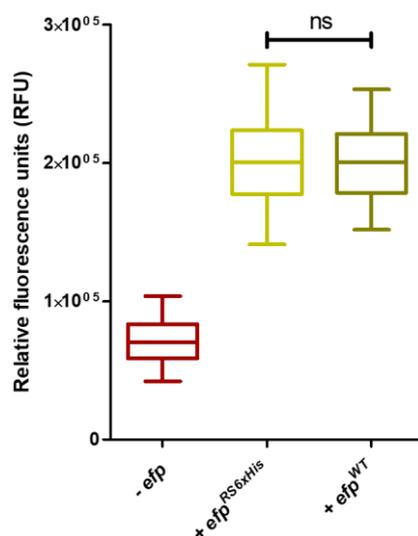
To study molecular characteristics of the EF-P from several Actinobacteria, two reporter genes were generated coding for the enhanced green fluorescent protein (eGFP), one with the strong stalling motif RPPP and another with the non-stalling sequence RPAP upstream of the eGFP sequence. Both genes were integrated chromosomally under the control of a constitutive promoter, and their expression was tested in *efp*<sup>+</sup> and *efp*<sup>-</sup> *C. glutamicum* cells. As expected, while the RPAP-eGFP variant was synthesized efficiently in both strains, RPPP-eGFP was barely detectable in the *efp*<sup>-</sup> mutant, revealing its strong dependence on EF-P, and the fluorescence intensity of these cells was only slightly higher than the background level (Figure 2.2).



**Figure 2.2** – EF-P enhances the translation of proteins containing polyproline motifs. **(A)** Single-cell fluorescence microscopy of *C. glutamicum* reporters producing RPPP-eGFP in *efp*<sup>+</sup> and *efp*<sup>-</sup> 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*<sup>+</sup> and *efp*<sup>-</sup> strains. The dashed line marks the background level of fluorescence. Western blots reveal levels of expression of *C. glutamicum* EF-P.

## Results

The activity of the his-tagged EF-P variant (connected by a short Arg-Ser linker,  $\text{efp}^{\text{RS6xHis}}$ ) was measured using this reporter strain and compared with the activity of the untagged variant ( $\text{efp}^{\text{WT}}$ ) sequence (Figure 2.3). No significant difference in EF-P activity was detected. Therefore, all complementation assays and further Western blot detections of this protein could be done using his-tagged variants.



**Figure 2.3** – Activity of tagged EF-P<sub>Cg</sub> ( $\text{efp}^{\text{RS6xHis}}$ ) and untagged ( $\text{efp}^{\text{WT}}$ ) was compared using *C. glutamicum* reporters as described in Figure 3. Boxplots represent the distribution of the relative fluorescence of 300 cells expressing RPPP-eGFP measured with an exposure time of 500-ms. Statistical significance via two-tailed Student's t-test. ns *non-significant*.

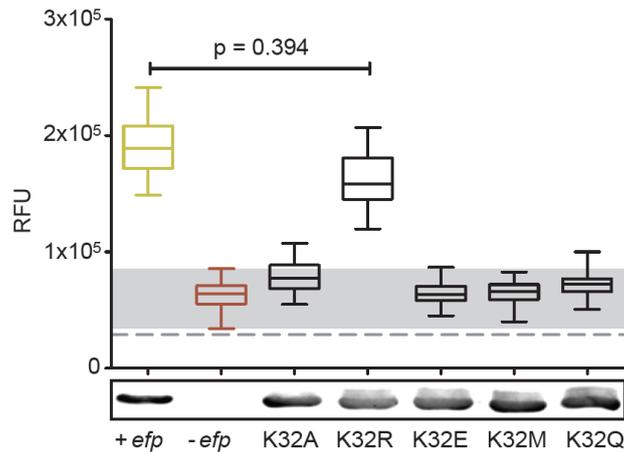
### 2.3 A positive charge at position 32 is essential for EF-P activity

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. In order to determine if K32 is important for EF-P activity, we constructed several EF-P variants to alter the properties of the amino acid at this position. After that, the activity of each variant was tested in the reporter strain described above.

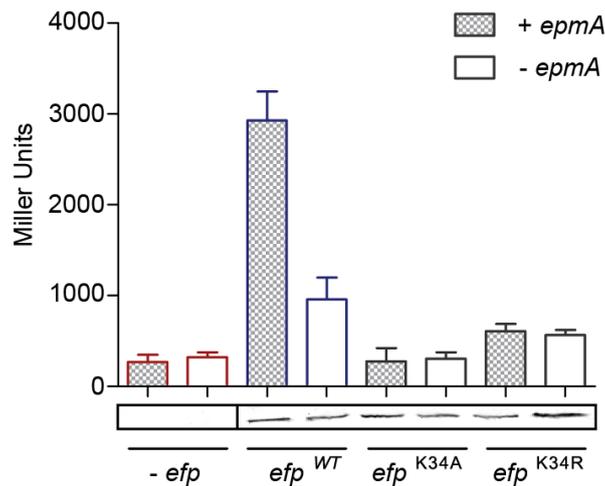
Strains expressing various the substitutions K32A, K32E, K32M and K32Q were virtually unable to support RPPP-eGFP translation. Surprisingly, the EF-P<sup>K32R</sup> variant enhanced RPPP-eGFP production to almost the same degree as the wild type (Figure 2.4).

Replacement of the lysine by arginine preserves the positively charged side chain of the amino acid, but nevertheless abolishes post-translational modification. Such amino acid substitution completely inactivate EF-P of *E. coli* (Figure 2.5) and *S. oneidensis* (Lassak et al., 2015) abolishing the attachment of a modification moiety. In *B. subtilis*, the substitution K32R abolishes post-translational modification but do not bring deleterious effect on swarming, observed in the  $\Delta\text{efp}$  strain (Hummels et al., 2017).

## Results



**Figure 2.4** – Synthesis of RPPP-eGFP in *C. glutamicum* strains expressing EF-P variants with the indicated amino acid replacements at position 32. Relative fluorescence units (RFU) are shown. The dashed line corresponds to background fluorescence. The area shown in grey marks the range of fluorescence measured in the negative control (-efp).



**Figure 2.5.** Activity of *E. coli* EF-P proteins. Amino acid replacements at the conserved lysine34 abolish EF-P activity in *E. coli* in presence or absence of the PTM-enzyme EpmA. *E. coli* reporter strain for EF-P activity as previously described (Ude et al., 2013). Mean and standard deviation were calculated from four independent replicates.

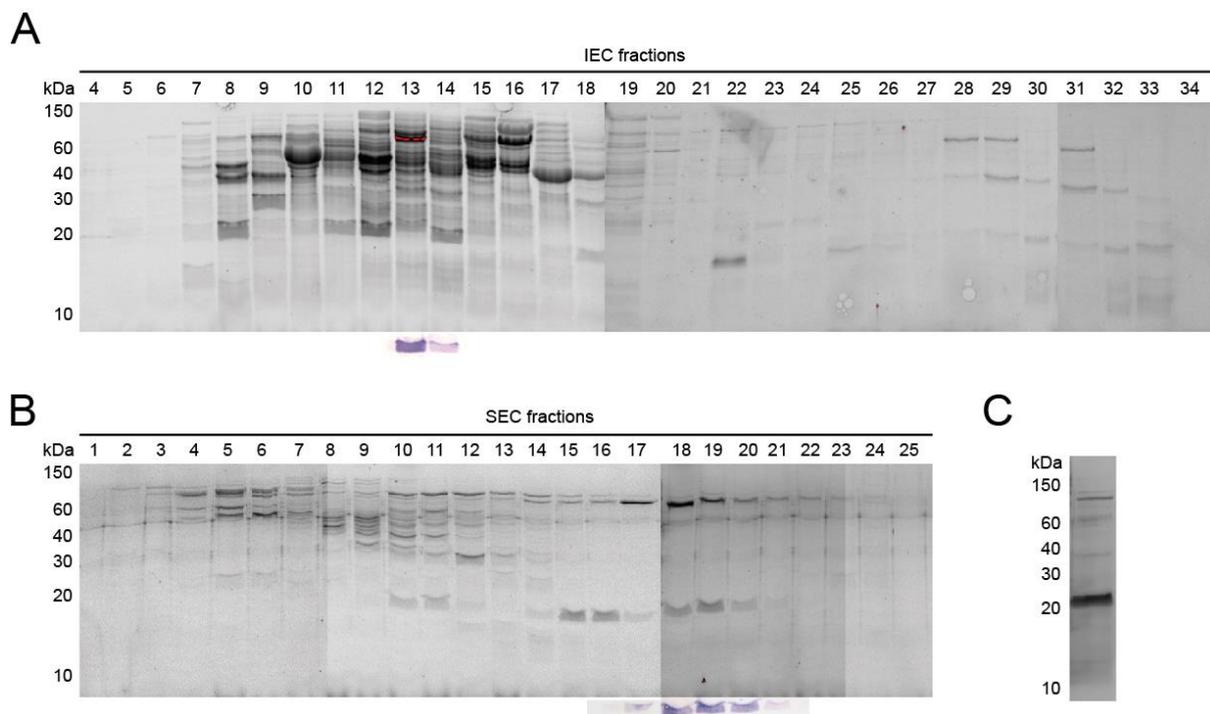
## 2.4 Development of a protocol for native EF-P purification

The high activity of the EF-P<sup>K32R</sup> variant prompted the question whether *C. glutamicum* EF-P (EF-P<sub>Cg</sub>) actually requires post-translational modification for activation. For this, it was essential that the native protein was purified and used for analytical steps, thus circumventing problems caused by low activity of PTM enzymes or substrate limitation triggered by protein overexpression.

## Results

The initial idea was to create a simple protocol to purify native EF-P from different Actinobacteria following a series of chromatographies in which EF-P proteins from different species would be enriched and eluted always in the same fraction. The native EF-P purification protocol set up involved the cultivation of cells until mid-exponential growth phase, suspension in lysis-buffer [Triethanolamine 20mM] to a final concentration of 0.2g/mL, cell lysis and EF-P enrichment by ion-exchange chromatography (IEC) [Mono-Q 8mL, GE Healthcare. Conditions: 4mL min<sup>-1</sup>, 0 – 1M KCl, total 40 fractions of 2mL], followed by size-exclusion chromatography (SEC) [Superdex 200, GE Healthcare. Conditions: 1mL min<sup>-1</sup>, 0.5 mL fractions] of the EF-P-richest fraction.

To validate this protocol, the native EF-P from three different bacterial species was purified: *E. coli* (EF-P<sub>Ec</sub>, Figure 2.6), *S. oneidensis* and *B. subtilis* (Table 2.3). The selection of those bacteria was due to the availability of antibodies that could specifically bind to the EF-P proteins from those species and allow detection via Western blot. The protein purifications were carried out resulting in EF-P enriched fractions with purity up to 95% (Figure 2.6C, Table 2.3).



**Figure 2.6** – Purification of native EF-P<sub>Ec</sub> using the method described here. **(A)** Trichloroethanol (TCE) stained SDS-polyacrylamide gel and Western blot detection of EF-P<sub>Ec</sub> of protein fractions obtained by IEC. **(B)** TCE stained SDS-polyacrylamide gel and Western blot of the protein fractions obtained after SEC of the fraction number 13 obtained in the previous step. **(C)** Sensitive silver staining of an SDS-page after elution of 10 $\mu$ L of the SEC-fraction number 19.

Although very efficient for native EF-P purification, the chromatographic method showed unpredictability in the EF-P-enriched fractions between bacterial species, been impossible to tell which exact fraction

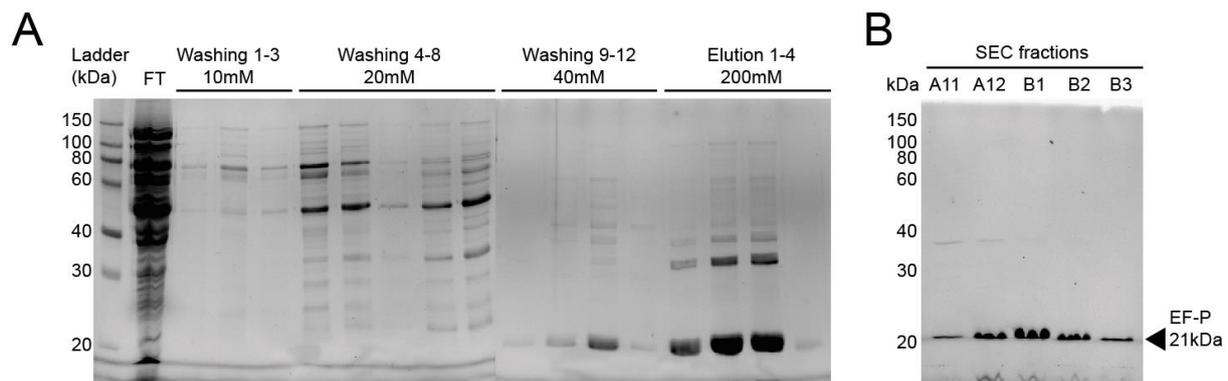
## Results

EF-P from different bacteria will be eluted (Table 2.3). To isolate the native EF-P from Actinobacteria using this method, it was necessary that an antibody suitable for detection was available. Our antibody collection against native EF-P from *E. coli* (1:10,000), *S. oneidensis* (1:5,000), *B. subtilis* (1:1,000) and *Thermus thermophilus* (1:1,000) were unable to detect native EF-P<sub>Cg</sub>.

**Table 2.3** – Native EF-P from different bacteria could be purified using the method described here.

Lysate	Inicial mass	IEC fraction	SEC fractions	Concentration (in 0.5 mL)	Purity
<i>E. coli</i> K12 MG1655	400mg	13-14	17-21	3.1 mg mL <sup>-1</sup>	95%
<i>S. oneidensis</i> MR-1	400mg	17-19	9-15	2.2 mg mL <sup>-1</sup>	90%
<i>B. subtilis</i> 168	400mg	15	13-16	1.5 mg mL <sup>-1</sup>	92%

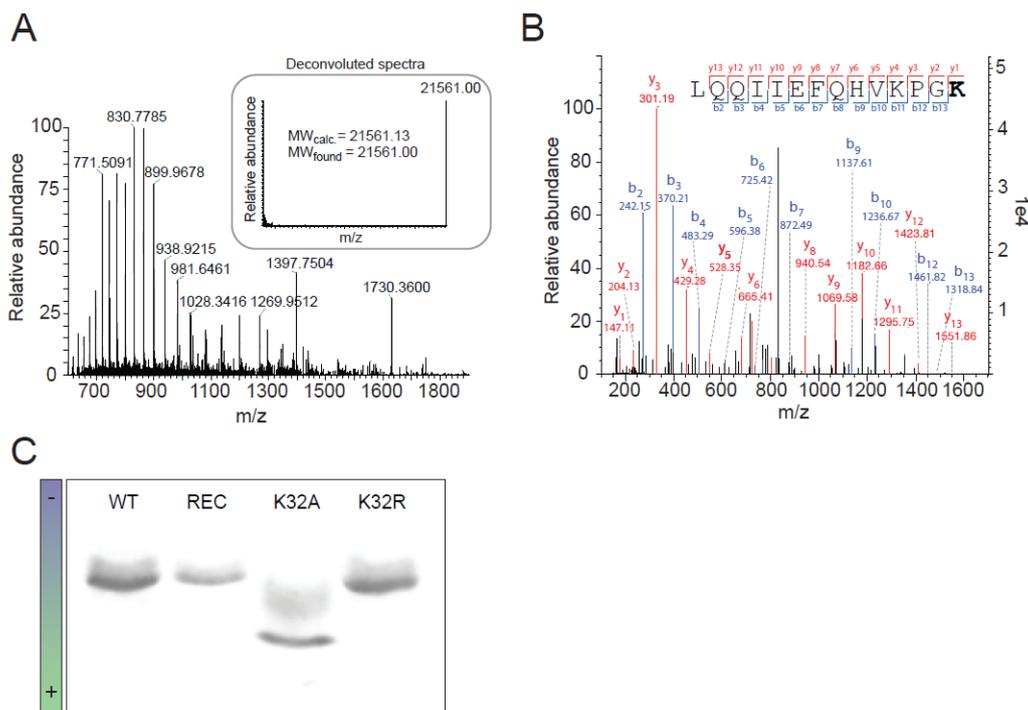
A histidine-tag (6xHis) is the most common amino acid sequences fused to proteins to assist purification and Western blot detection. It was shown in Figure 2.3 that EF-P<sub>Cg</sub> activity is not affected by the incorporation of a linker and a hexa-Histidine motif to the protein C-terminal. To further continue our investigation on the post-translational status of the EF-P from *C. glutamicum*, we incorporated this amino acid sequence to the chromosomally encoded copy of *efp*. This strategy allowed the expression and purification of native levels of EF-P<sub>Cg</sub>, circumventing problems caused by protein overexpression. One-step purification by Ni-NTA Agarose chromatography resulted in EF-P fractions of 90% purity (Figure 2.7A), after second-step SEC of fraction number 2, it could be increased to over 99% (Figure 2.7B).



**Figure 2.7** – Purification of native produced EF-PCg. **(A)** TCE-stained SDS-polyacrylamide gel of Ni-NTA chromatography fractions. Concentrations of imidazole used in the washing and elution buffers are shown. **(B)** SDS-polyacrylamide gel of protein fractions after SEC of the fraction E2 obtained in the previous step. 10µL of each fraction were loaded in each lane (0.6 – 10µg of total protein). FT flow through.

## 2.5 EF-P in *C. glutamicum* is not post-translationally modified

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, the native encoded His-tagged variant was purified (Figure 2.7). In addition, this same gene was heterologously expressed in the  $\Delta efp \Delta epmA$  *E. coli* mutant, expecting that the resulting product would not be modified in this bacterium. Both proteins were then analyzed by mass spectrometry. The total protein mass matched the calculated mass, irrespective of whether EF-P<sub>Cg</sub> was isolated from *C. glutamicum* (Figure 2.8A, Table 2.4) or from the transformed *E. coli* mutant (Table 2.4). LC-MS/MS analysis of the trypsin-digested endogenous EF-P confirmed that none of the peptides were modified (Figure 2.8B). Moreover, there was no difference in isoelectric point between endogenous and heterologous EF-P<sub>Cg</sub> (Figure 2.8C). We therefore conclude that neither the endogenous nor the recombinantly produced protein undergoes PTM.



**Figure 2.8** – EF-P is not post-translationally modified in *C. glutamicum*. **(A)** Mass spectra of the intact endogenous EF-P<sub>Cg</sub> 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*  $\Delta efp \Delta epmA$  (REC), and after replacement of lysine 32 by alanine (K32A) or arginine (K32R).

## Results

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 (Figure 2.8C, Table 2.4). These variants exhibited only the expected shifts in mass and IEP.

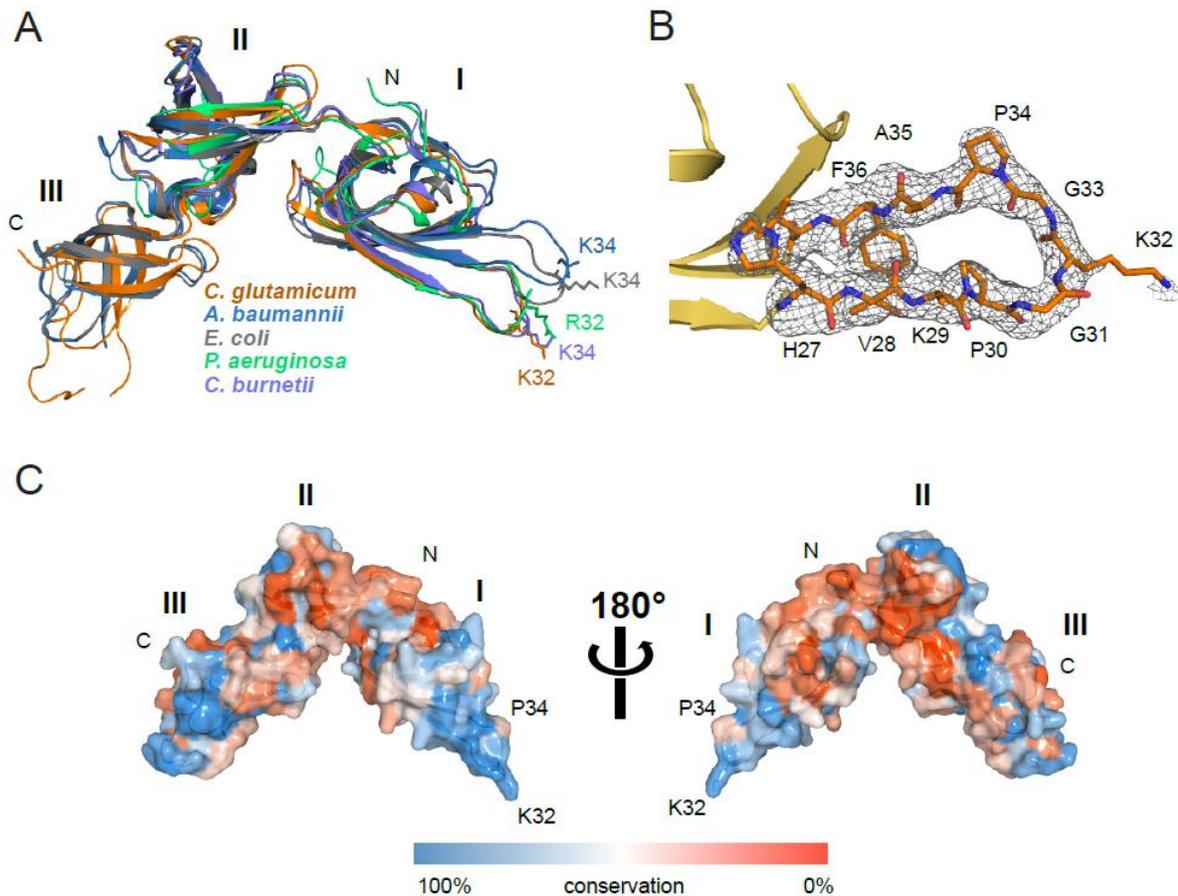
**Table 2.4** – Predicted calculated mass of the different EF-P variants and mass measured by mass spectra of the purified intact EF-P proteins listed.

EF-P variant	Predicted mass (Da)	Measured mass (Da)
EF-P <sub>Cg</sub> exponential	21575.15	21574.04
EF-P <sub>Cg</sub> stationary	21575.15	21574.04
EF-P <sub>Cg</sub> recombinant	21575.15	21576.03
EF-P <sup>K32A</sup>	21518.06	21517.99
EF-P <sup>K32R</sup>	21603.17	21603.05

### 2.6 Crystal structure of *C. glutamicum* EF-P

In order to gain insights into the structural configuration of this unmodified Actinobacteria EF-P, the endogenous protein from *C. glutamicum* ATCC13032 was isolated and had its X-ray crystal structure determined to 2.2 Å resolution (Table 2.5). EF-P<sub>Cg</sub> shares its overall folding topology with the previously reported bacterial EF-P structures and could be superimposed on homologous EF-P structures with an r.m.s.d. ranging from 1.9 to 3 Å (Figure 2.9A). The  $\beta 2\Omega\beta 3$  loop 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 sidechain of K32 is not discernible and must therefore be flexible (Figure 2.9B).

Alignment of 150 different bacterial EF-P protein sequences reveal 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 contacts the ribosomal L1 protein in the *E. coli* EF-P-ribosome complex also shows sequence conservation (Figure 2.9C).



**Figure 2.9** – Crystal structure of *C. glutamicum* EF-P. **(A)** Structural superposition and ribbon representations of available EF-P structures. *C. glutamicum* (PDB code 6S8Z, orange, this work), *Acinetobacter baumannii* (PDB code: 5J3B, blue), *E. coli* (PDB code: 3A5Z, grey), *Pseudomonas aeruginosa* (PDB code: 3OYY, green) and *Coxiella burnetii* (PDB code: 3TRE, purple) are shown. The Arg or Lys residue at the tip of the  $\beta 2\Omega\beta 3$  loop in the N-domain is shown as a stick model. **(B)** Simulated-annealing omit electron density map contoured at  $2.5 \sigma$  of the connecting loop between  $\beta 2$  and  $\beta 3$  in the domain I 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). Homologues 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) data base (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

## Results

---

**Table 2.5** Data collection and refinement statistics.

	EF-P <sub>Cg</sub>
<b>Wavelength (Å)</b>	0.9677
<b>Resolution range</b>	38.6 - 2.2 (2.26 - 2.2)
<b>Space group</b>	P2 <sub>1</sub>
<b>Unit cell</b>	51.2 38.1 55.3 90 93.1 90
<b>Total reflections</b>	52,547 (3,966)
<b>Unique reflections</b>	10,942 (799)
<b>Multiplicity</b>	4.8 (5.0)
<b>Completeness (%)</b>	98.7 (98.9)
<b>Mean I/sigma(I)</b>	5.1 (1.5)
<b>Wilson B-factor</b>	49.7
<b>R-merge</b>	0.129 (1.12)
<b>R-pim</b>	0.065 (1.158)
<b>CC1/2</b>	0.996 (0.629)
<b>Reflections used in refinement</b>	10,928 (1,080)
<b>Reflections used for R-free</b>	546 (54)
<b>R-work</b>	0.206 (0.356)
<b>R-free</b>	0.265 (0.404)
<b>CC(work)</b>	0.957 (0.519)
<b>CC(free)</b>	0.917 (0.211)
<b>Number of non-hydrogen atoms</b>	1,489
<b>Macromolecules</b>	1,423
<b>Ligands</b>	17
<b>Solvent</b>	49
<b>Protein residues</b>	185
<b>RMS (bonds)</b>	0.012
<b>RMS (angles)</b>	1.4
<b>Ramachandran favored (%)</b>	96.7
<b>Ramachandran allowed (%)</b>	2.76
<b>Ramachandran outliers (%)</b>	0.55
<b>Rotamer outliers (%)</b>	3.3
<b>Clashscore</b>	10.2
<b>Average B-factor</b>	57.4
<b>Macromolecules</b>	57.5
<b>Ligands</b>	66.9
<b>Solvent</b>	51.5

## 2.7 P30 and P34 are essential for EF-P activity

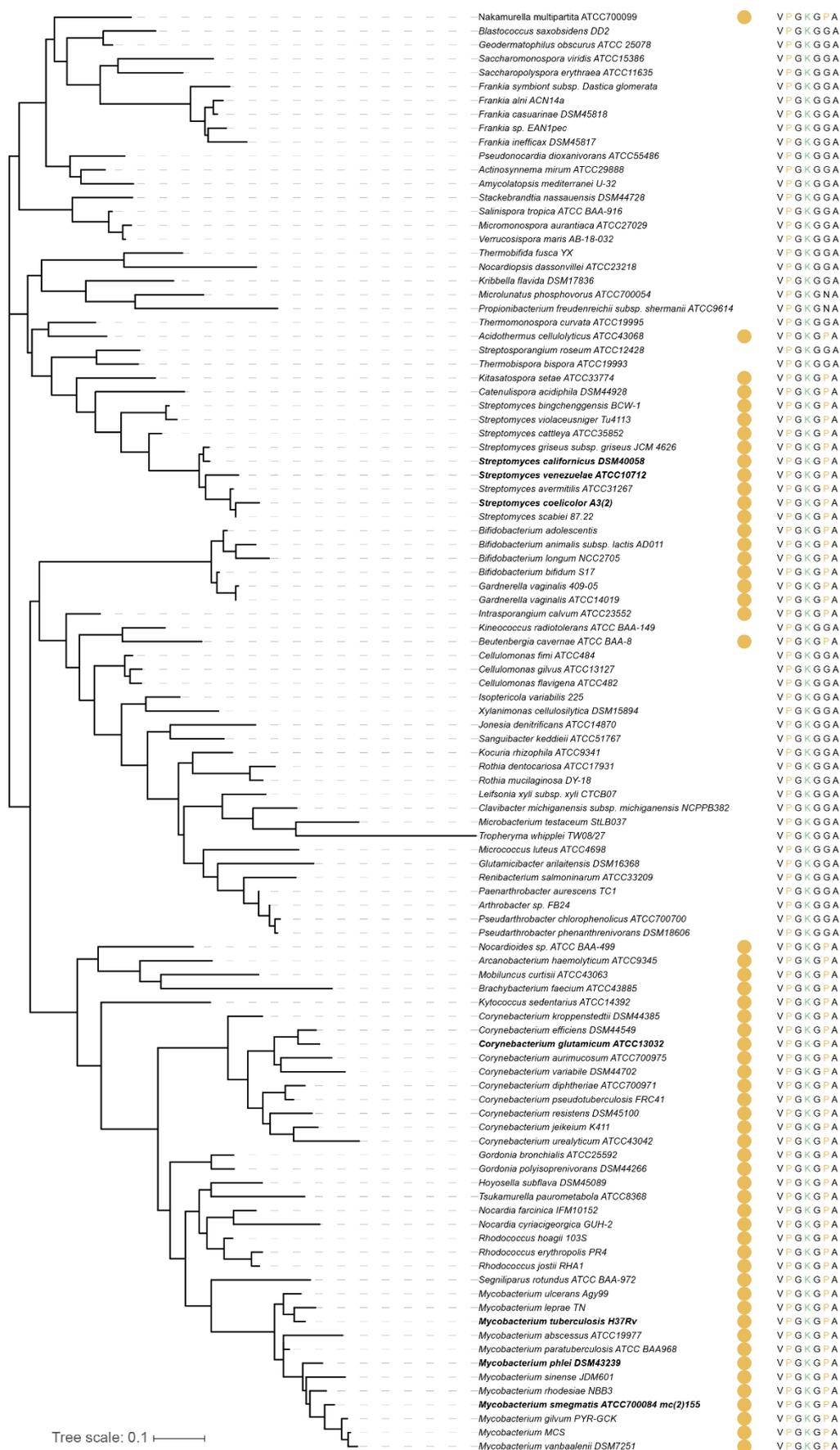
Although the sequence of domain I in EF-P is highly conserved across the bacterial kingdom, a striking difference, a loop with the palindromic PGKGP motif, was identified in Actinobacteria EF-Ps. P30 and P34 flank this palindromic EF-P loop sequence that is found in 11% of all bacteria including all EF-P sequences from *Corynebacterium*, *Streptomyces*, *Mycobacterium*, *Bifidobacterium*, *Gardnerella*, *Mobiluncus* and *Norcardia* (Figure 2.10).

P30 was previously identified as an invariant proline in all EF-Ps with a lysine at the tip of the  $\beta 2\Omega\beta 3$  in a phylogenetic analysis, and other amino acids replace proline when arginine is at the tip (Volkwein et al., 2019). The exchange of P30 for alanine or glutamine in EF-P of *C. glutamicum* led to non-functional variants (Figure 2.11).

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 2.12). Some actinobacterial EF-P sequences have glycine or asparagine at position 34 (Figures 2.10 and 2.12B and D). To determine the significance of P34 for *C. glutamicum* EF-P, this amino acid was replaced with alanine (*B. subtilis* EF-P has A34), glutamine (*E. coli* EF-P has Q34), glycine and asparagine and the EF-P activity of the resulting variants were quantified (Figure 2.11). All resulting EF-P variants were not able to support RPPP-eGFP translation in *C. glutamicum* (Figure 2.11).

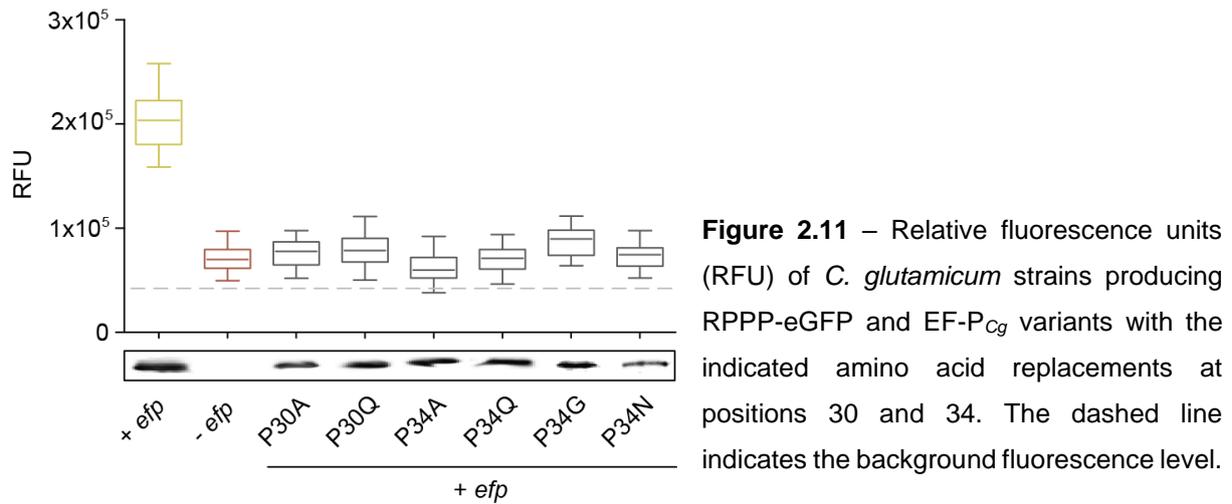
The two prolines in the consensus sequence most likely rigidify the  $\beta 2\Omega\beta 3$  loop, which in turn could enable these EF-Ps to stabilise the acceptor arm of the tRNA and thus allow translation of polyproline motifs without post-translational modification. Apart from this motif, no other sequence signature was identified in the structural and amino acid sequence analyses of unmodified EF-Ps (Figure 2.12).

## Results



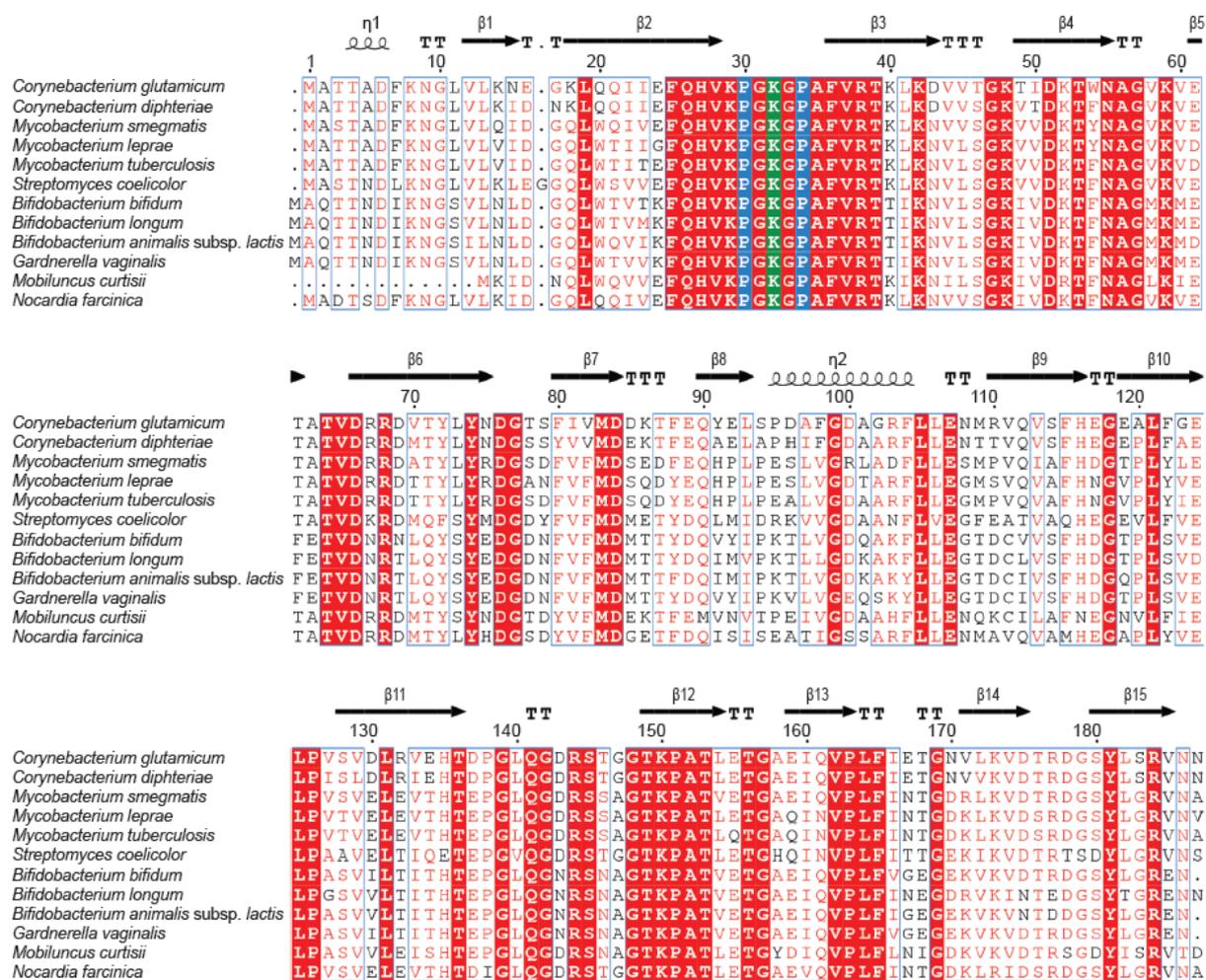
**Figure 2.10.** Actinobacteria branch of the EF-P phylogenetic tree. Yellow dots represent the species that have the palindromic  $\beta 2\Omega\beta 3$  sequence Pro-Gly-Lys-Gly-Pro.

## Results

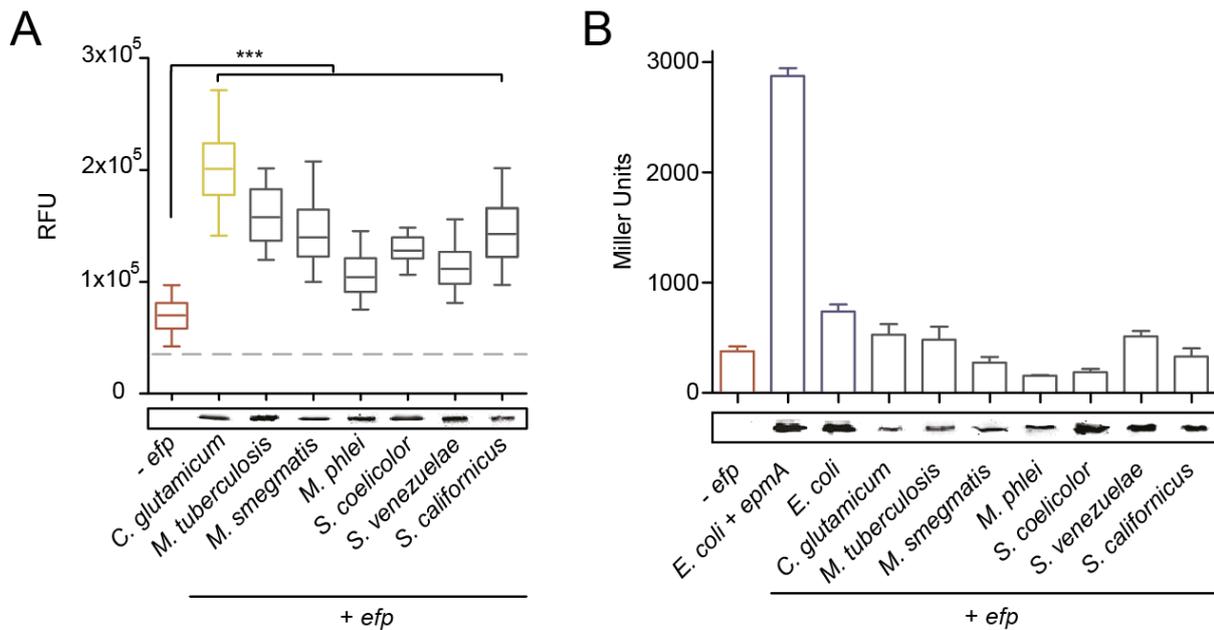


## 2.8 A novel subfamily of EF-P in Actinobacteria

EF-P protein sequences are highly conserved among the most relevant Actinobacteria (Figure 2.13). The same reporter strain was used to test whether EF-P proteins from other members of the Actinobacteria are able to complement the *C. glutamicum* *efp* deletion mutant. EF-Ps from *Mycobacterium tuberculosis*, *M. smegmatis*, *M. phlei*, *Streptomyces coelicolor*, *S. venezuelae* and *S. californicus*, significantly enhanced translation of RPPP-eGFP in *C. glutamicum* (Figure 2.14A). EF-P production was confirmed by Western blot analysis. In contrast, EF-Ps from *Actinobacteria* were unable to complement a  $\Delta efp \Delta epmA$  *E. coli* mutant (Figure 2.14B), which supports the hypothesis that actinobacterial EF-Ps constitute a separate subfamily of these elongation factors.

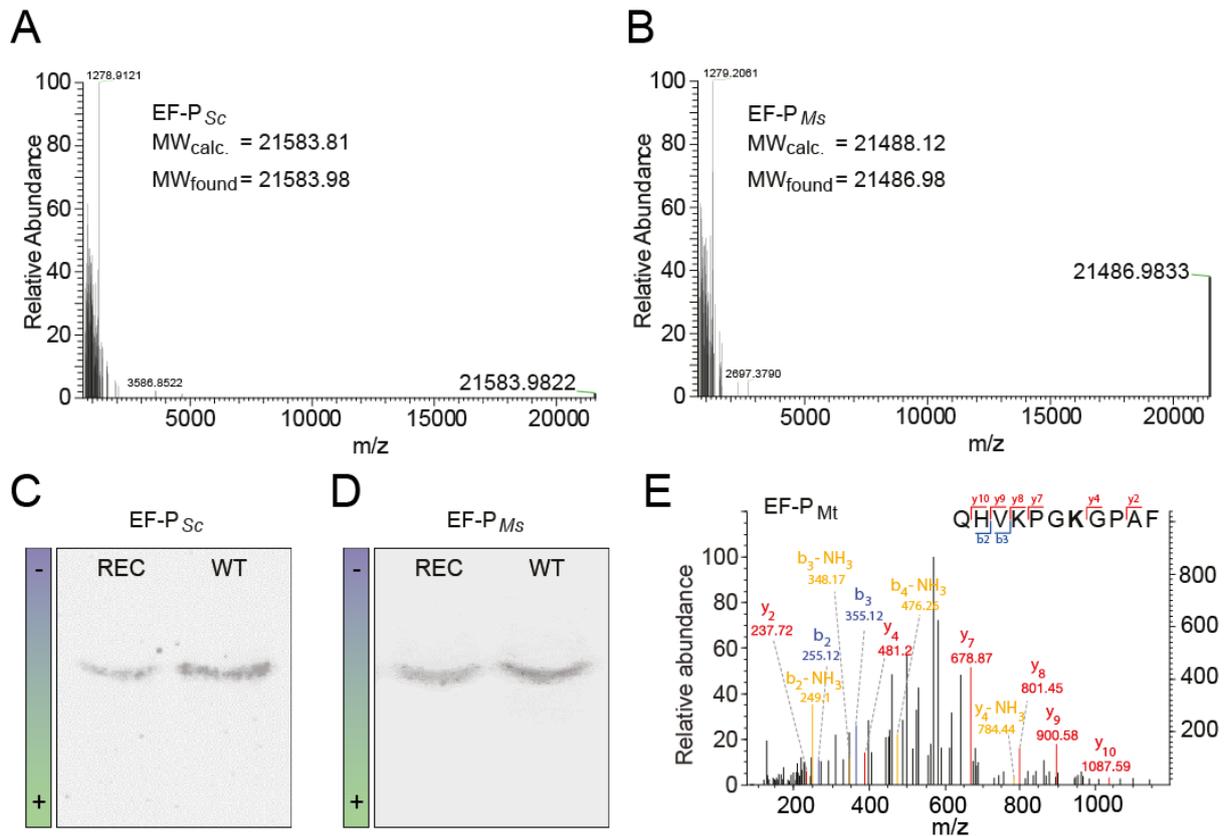


**Figure 2.13** – Alignment of EF-P amino acid sequences of different actinobacterial species. Identical amino acids are highlighted in red, conserved K32 in green, and P30 and P34 in blue.



**Figure 2.14** – A novel subfamily of EF-P in Actinobacteria. **(A)** 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, \*\*\*  $P < 0.0001$ . **(B)** Members of the new EF-P subfamily do not significantly enhance polyproline biosynthesis in *E. coli*  $\Delta efp \Delta epmA$ . Reporter assay as described previously (Ude et al., 2013), protocol as described in methods. Mean and standard deviation from four independent experiments. Heterologous protein production was confirmed by Western blot as shown.

The endogenous *efp* genes in *S. coelicolor* and *M. smegmatis* were tagged as described for *C. glutamicum*. Endogenous His-tagged EF-P<sub>Sc</sub> and EF-P<sub>Ms</sub> proteins were then purified and compared to a recombinant variant produced in the  $\Delta efp \Delta epmA$  *E. coli* strain. In both cases, the intact protein mass measured by mass spectrometry matched the calculated for unmodified EF-P (Figure 2.15A and B). Isoelectric focusing showed no changes in protein charge between endogenous and heterologously produced versions of the proteins (Figure 2.15C and D). Lastly, Unmodified EF-P could also be identified in lysates of *Mycobacterium tuberculosis* H37Rv (Figure 2.15E). Our results therefore confirm 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.

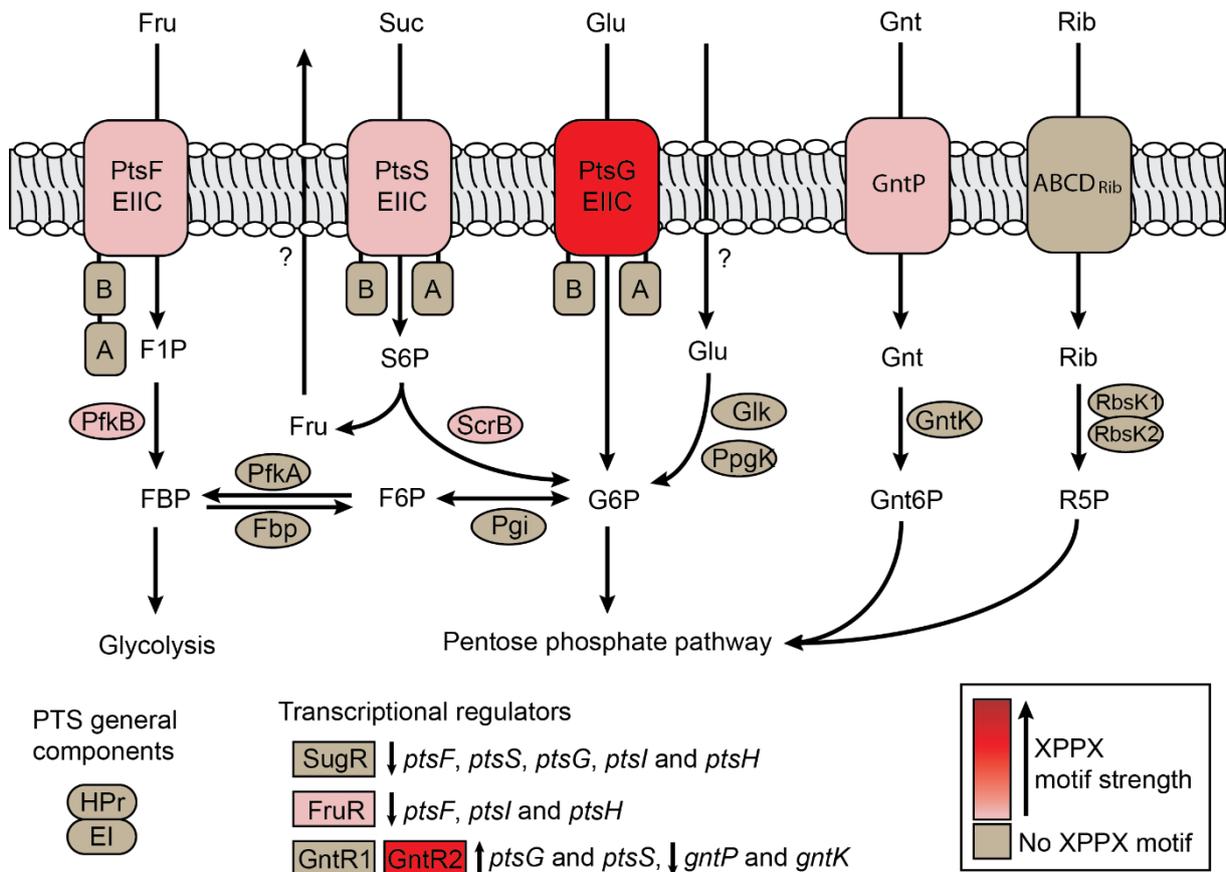


**Figure 2.15** – Different Actinobacteria do not PTM EF-P. **(A)** Deconvoluted MS spectra of intact endogenous EF-P from *S. coelicolor* (EF-P<sub>Sc</sub>) and **(B)** *M. smegmatis* (EF-P<sub>Ms</sub>) **(C)** IEF gel of recombinant (REC) and endogenous EF-P (WT) isolated from *S. coelicolor* and **(D)** *M. smegmatis*. **(E)** Annotated MS/MS spectra of K32 containing peptide of endogenous EF-P from *M. tuberculosis*.

## 2.9 EF-P is required for fast growth of *C. glutamicum* on glucose

Next, we were interested in the specific effects of this novel EF-P subfamily in relevant characteristics of *C. glutamicum*, such as carbohydrate transport and utilization. For that, the amino acid sequences of all known proteins responsible for carbohydrate uptake, metabolism and transcriptional regulation were analyzed (Figure 2.16). Among those sequences, the strongest motifs found were in the GntR2, a global transcriptional regulator that among others promotes *ptsG* transcription (Frunzke et al., 2008; Ikeda, 2012; Tanaka et al., 2014), and EII<sup>Glc</sup> itself the main glucose transporter in *C. glutamicum* (Figure 2.16, Table 2.6) (Martins et al., 2019; Parche et al., 2001; Pfeifer et al., 2017). The other motifs in EII<sup>Fru</sup>, PfkB, EII<sup>Scr</sup>, ScrB, GntP, and FruR are characterized as weak-stalling motifs and therefore, are less likely to stall ribosomes. Maltose can also be metabolized by *C. glutamicum* ATCC13032 however, the complete pathways of carbohydrate transport, phosphorylation and regulation are proposed but still unclear (Henrich et al., 2013; Kuhlmann et al., 2015).

## Results



**Figure 2.16** – Schematic depictions of the domain structures of PTS systems in *C. glutamicum*. Their substrates, metabolic contexts and transcriptional regulators. The level of shading indicates the relative strengths of the polyproline motifs they contain (listed in Table 1). Fru fructose, Suc sucrose, Glu glucose, Gnt gluconate, Rib ribose, PTS phosphoenolpyruvate-dependent sugar phosphotransferase system. EII<sup>Fru</sup> fructose-specific phosphotransferase system, EII<sup>Suc</sup> sucrose-specific phosphotransferase system, EII<sup>Glc</sup> glucose-specific phosphotransferase system (subunits A, B and C). GntP gluconate permease, ABCDRib ribose specific ATP-binding cassette transporter for D-ribose, F1P fructose-1-phosphate, FBP fructose-1,6-bisphosphate, F6P fructose-6-phosphate, S6P sucrose-6-phosphate, G6P glucose-6-phosphate, Gnt6P 6-phosphogluconate, R5P ribose-5-phosphate, PfkB fructose-1-phosphate kinase, PfkA 6-phosphofructokinase, Fbp fructose-1,6-bisphosphatase, ScrB sucrose-6-phosphate hydrolase, Pgi phosphoglucose isomerase, Glk glucokinase, PpgK polyphosphate dependent glucokinase, GntK gluconate kinase, RbsK1 ribokinase 1, RbsK2 ribokinase 2. HPr and EI, the heat-stable phosphocarrier protein and enzyme I, are general energy-coupling proteins of the phosphoenolpyruvate-dependent phospho-transferase systems (PTS). HPr and EI are encoded respectively by *ptsH* and *ptsI*.

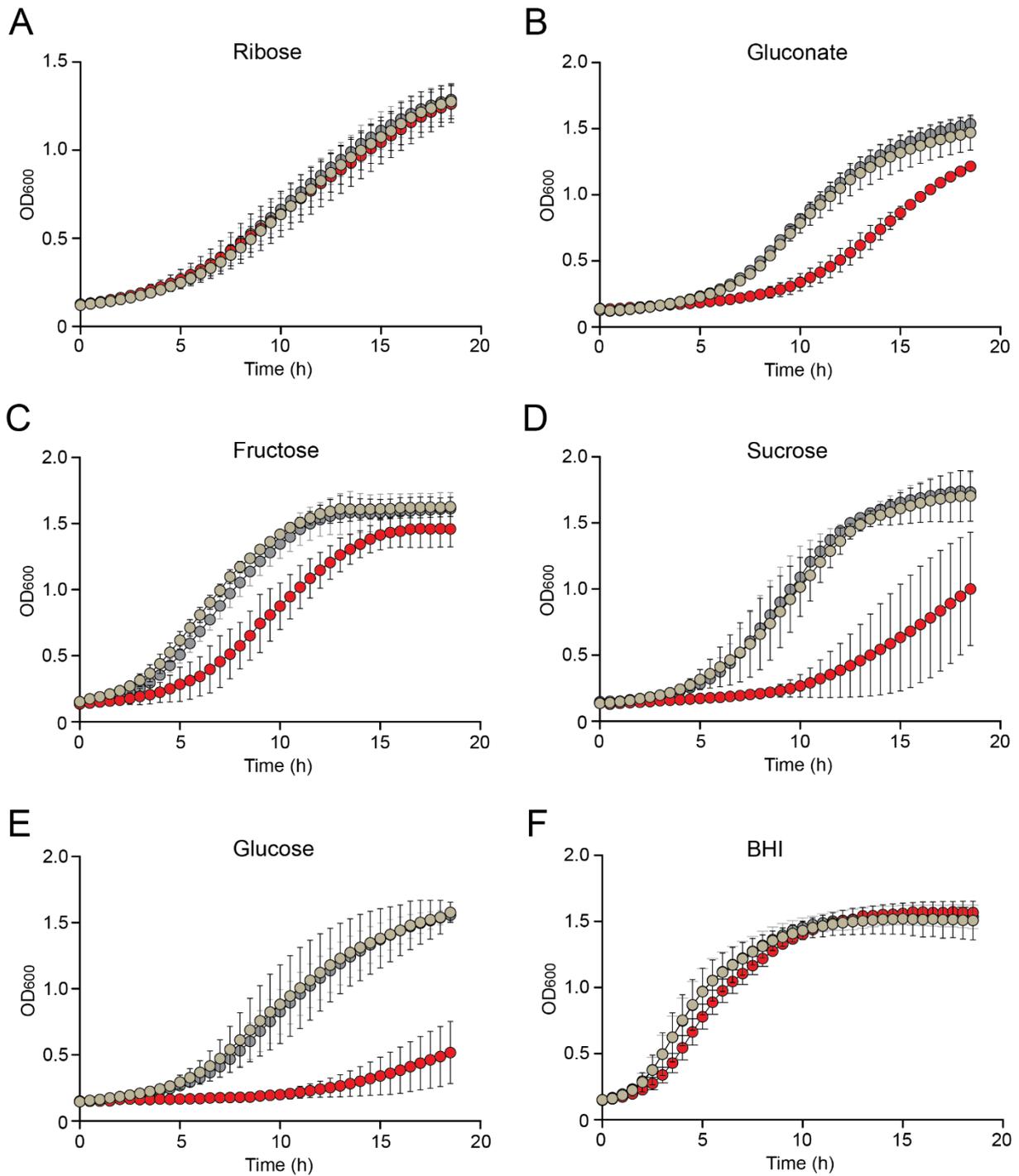
## Results

**Table 2.6** – Polyproline motifs found in proteins involved in carbohydrate transport and metabolism in *C. glutamicum* ATCC 13032. The amino acid context (positions -2, -1 and +1) of each motif is shown.

Name	Reference	Motif sequence	Polyproline position	Strength
EII <sup>Glc</sup>	Q46072	VFPPPL	235/236	Moderate
EII <sup>Fru</sup>	Q8NP80	MVPPI	479/480	Weak
EII <sup>Scr</sup>	Q8NMD6	SFPPI	328/329	Weak
GntP	Q79VC5	FVPPH	166/167	Weak
PfkB	Q8NP81	SLPPG	139/140	Weak
ScrB	Q8NMD5	VTPPQ	20/21	Weak
GntR2	Q8NPU3	MAPPI	187/188	Moderate
FruR	Q8NP82	TSPPR/GMPPE	67/68 and 79/80	Weak

To confirm the predictions, the *efp* deletion mutant was grown in minimum defined media supplemented with the different simple carbohydrates selected. Among those, it was observed the absence of growth defect in ribose (Figure 2.17A), supporting the lack of XPPX motif in any ribose-specific uptake and metabolism enzymes; Intermediate growth defect in gluconate, fructose and sucrose (Figure 2.17B-D), and that the *efp* mutant could hardly grow on glucose as the sole carbon source (Figure 2.17E). Moreover, the extend of the growth defect correlates with the strength of the polyproline motif observed in the respective transporter. Re-introducing the *efp* gene into the strain allowed complete growth recovery in all conditions (Figure 2.17, gray curves). As a positive control, the *efp* mutant was grown in rich medium BHI with no growth impairment (Figure 2.17F).

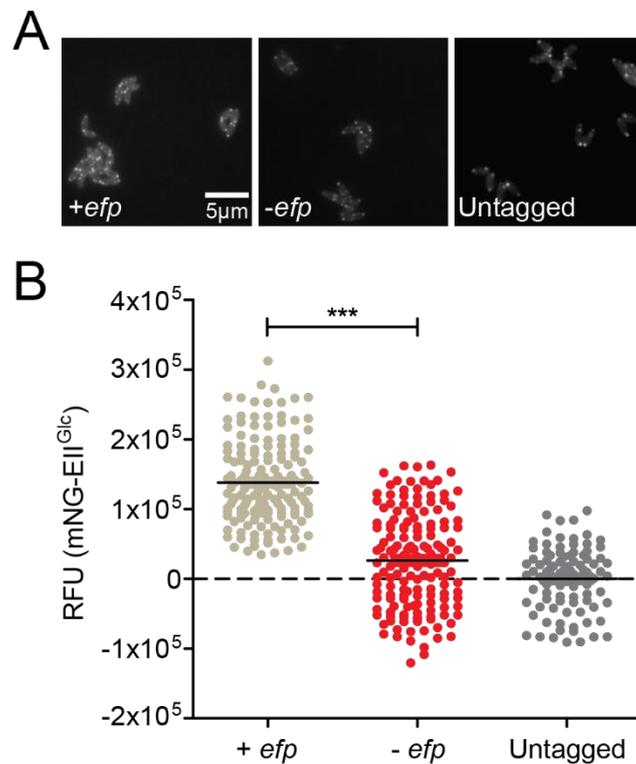
## Results



**Figure 2.17** – Typical growth curves of wild-type *C. glutamicum* ATCC 13032 (brown dots), the *efp* deletion mutant (red dots) and the complemented mutant with *efp in trans* (gray dots) grown on (A-E) minimal medium supplemented (at 2 % w/v) with the indicated carbon sources or on (F) complex BHI medium. Cells were grown in a microplate reader (30°C, 220rpm). Dots represent mean values, and bars depict the standard deviations of the mean of four independent replicates.

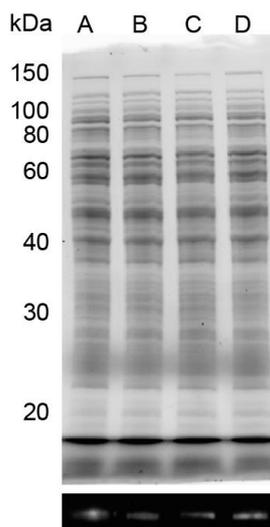
## 2.10 EII<sup>Glc</sup> translation requires EF-P

To investigate the effect of EF-P on the transporter production, a strain chromosomally coding the native EII<sup>Glc</sup> fused with mNeonGreen (N-terminal, as described previously) (Martins et al., 2019) was constructed and had the total fluorescence in the context of *+efp* and *-efp* analyzed. This fluorescent EII<sup>Glc</sup> variant was shown not to interfere neither with EII<sup>Glc</sup> copy number nor activity (Martins et al., 2019). Total fluorescence of the *-efp* cells was significantly lower than the *+efp* control (Figure 2.18A and B). Many  $\Delta efp$  cells showed the same fluorescence levels as the background given by the fluorescence of the untagged *C. glutamicum* ATCC13032 (Figure 2.18A and B), suggesting that part of the population have little or no EII<sup>Glc</sup> production. Furthermore, we tried to overexpress EF-P and measure EII<sup>Glc</sup> production, but the several strains constructed could not produce more EF-P proteins than the wild type (Figure 2.19).



**Figure 2.18** – EII<sup>Glc</sup> synthesis is dependent on EF-P. (A) Single-cell fluorescence microscopy of *C. glutamicum* cells expressing chromosomally encoded mNG-EII<sup>Glc</sup> in *efp*<sup>+</sup> and *efp*<sup>-</sup> strains. Cells of the untagged parental strain were used to determine background fluorescence. Exposure time in all cases was 2000-ms. (B) Distribution of relative mNeonGreen fluorescence intensities of a minimum of 300 *C. glutamicum* cells expressing mNG-EII<sup>Glc</sup> in either the *efp*<sup>+</sup> or the *efp*<sup>-</sup> strain. Black lines indicate the mean fluorescence. The dashed line represents the background fluorescence of the untagged strain. Statistical significance was assessed with the two-tailed t test. \*\*\*p<0.0001.

## Results

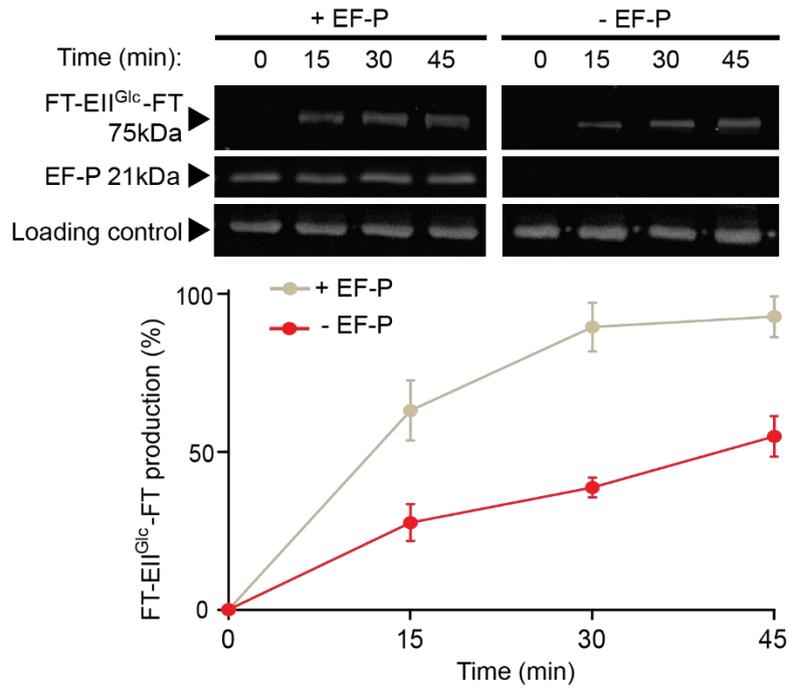


**Figure 2.19** – SDS-polyacrylamide gel of *C. glutamicum* cytosolic protein fractions and EF-P bands on Western blot membrane. Antibody against 6His (1:10.000) was used. **(A)** chromosomally integrated *efp*-6His in the native locus. **(B)**  $\Delta$ *efp* strain transformed with pEKEx2 with native Prom<sub>*efp*</sub> and *efp*-6His. **(C)**  $\Delta$ *efp* strain transformed with pEKEx2 with T7 promoter, native RBS<sub>*efp*</sub> and *efp*-6His. Cells were collected two hours after induction with 1mM IPTG. **(D)**  $\Delta$ *efp* strain transformed with pEKEx2 with T7 promoter, synthetic RBS<sub>*efp*</sub> and *efp*-6His. Two hours after induction with 1mM IPTG. 30 $\mu$ g of protein were loaded in each lane. EF-P detection by Western blot is shown below.

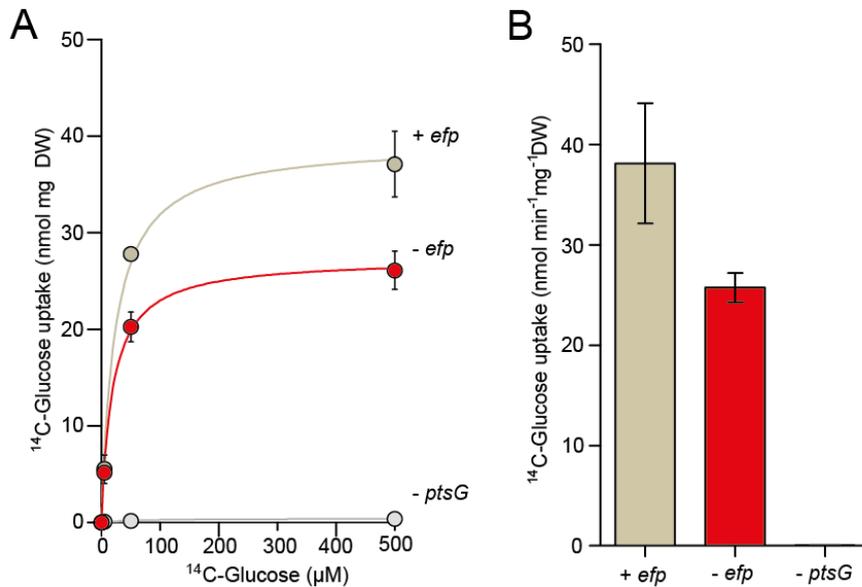
The effect of EF-P over EII<sup>Glc</sup> translation could also be demonstrated *in vitro*. For this, a reaction was set up containing all the substrates, ribosomes, tRNA, proteins and co-factors necessary for translation of EII<sup>Glc</sup>, varying only the presence or absence of functional EF-P. To detect the production of full-length EII<sup>Glc</sup> a reporter DNA sequence containing a double Flag-tagged EII<sup>Glc</sup> (FT-EII<sup>Glc</sup>-FT) was constructed to further quantify protein production by Western blot using anti-FlagTag antibodies. In the *in vitro* experiment, the addition of purified EF-P stimulates EII<sup>Glc</sup> translation confirmed by the 2.4 times faster accumulation of full-length EII<sup>Glc</sup> (Figure 2.20).

Next, it was questioned whether the growth defect of *C. glutamicum* in glucose would be caused by lower glucose uptake rate due to reduced EII<sup>Glc</sup> copy number. The kinetics for uptake of radiolabeled D-glucose-6-<sup>14</sup>C by *C. glutamicum* ATCC13032 were described previously (Lindner et al., 2011). EII<sup>Glc</sup> is characterized by a  $K_m$  of 14  $\mu$ M and a  $V_{max}$  of  $35 \pm 3$  nmol min<sup>-1</sup> mg<sup>-1</sup> DW. It was tested transport activities of wild type and the  $\Delta$ *efp* mutant at three different glucose concentrations (5, 50 and 500  $\mu$ M), whereby one concentration is below and two concentrations are above the  $K_m$  value. At all three concentrations the uptake rate of the  $\Delta$ *efp* mutant was significantly lower in comparison to the wild type. Using the highest concentration, the initial rate was determined to be  $38 \pm 4$  nmol min<sup>-1</sup> mg<sup>-1</sup> DW for the wild type and  $25 \pm 4$  nmol min<sup>-1</sup> mg<sup>-1</sup> DW for the  $\Delta$ *efp* mutant (Figure 2.21). As a negative control, it was measured the initial glucose uptake rate of the  $\Delta$ *ptsG* mutant, which was  $0.07 \pm 0.09$  nmol min<sup>-1</sup> mg<sup>-1</sup> DW (Figure 2.21).

## Results



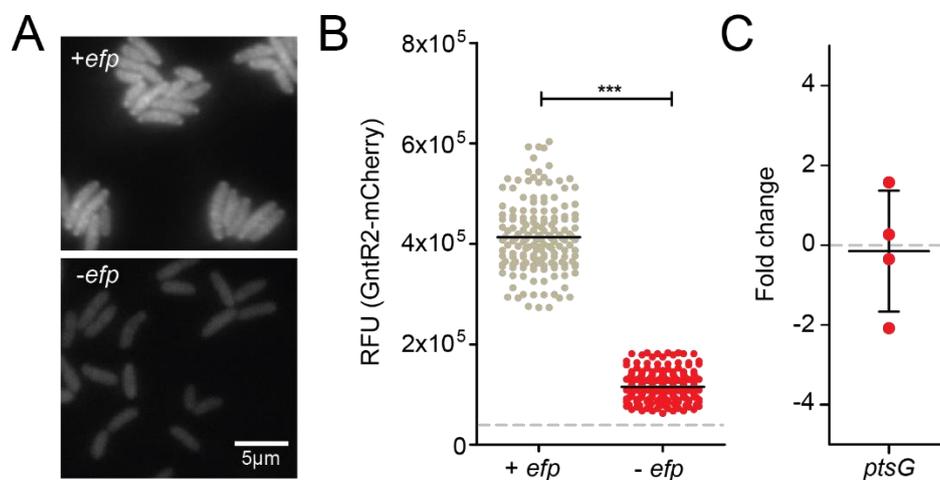
**Figure 2.20** – Quantification of *in vitro*-translated FT-EII<sup>Glc</sup>-FT synthesized in the presence and absence of purified EF-P. FT-EII<sup>Glc</sup>-FT was quantified on a Western blot, using primary monoclonal mouse anti-Flag antibodies (1:10.000) and the fluorescence-labeled secondary goat anti-mouse IgG antibody (1:20.000). Mean values and standard deviations of three independent experiments are shown.



**Figure 2.21** – Uptake of radiolabeled D-glucose-6-<sup>14</sup>C by *C. glutamicum efp*<sup>+</sup>, *efp*<sup>-</sup> and *ptsG*<sup>-</sup> strains. **(A)** Transport measurements as they fit to the Michaelis-Menten equation curve published previously (Lindner et al., 2011). **(B)** Vmax values for the wild type and  $\Delta efp$  strain. Values and standard deviations of five independent experiments are shown.

## 2.11 EF-P dependency of GntR2 does not affect transcription of *ptsG*

As discussed previously, EF-P might affect the transcription of certain genes by changing native levels of XPPX-containing transcriptional regulators. Among the network of transcriptional regulators of *ptsG* identified so far, GntR2 is the only one containing an XPPX motif in the sequence. GntR2 is a global transcriptional regulator of the GntR-type that simultaneously activates *ptsG* and *ptsS* expression, strongly represses *gntP* and *gntK* and weakly represses transcription of genes coding for enzymes of the pentose phosphate pathway (Frunzke et al., 2008; Tanaka et al., 2014). Due to the XPPX motif, it was asked how strongly EF-P contribute to GntR2 levels and consequently to *ptsG* transcription. To address this question, the native *gntR2* copy was chromosomally fused with a C-terminal fluorescent protein mCherry and had the fluorescence levels of GntR2-mCherry quantified in both *+efp* and *-efp* context (Figure 2.22A and B). GntR2-mCherry levels were 4.8 times lower in the  $\Delta efp$  strain, confirming EF-P dependency however, transcription of *ptsG* was not altered (Figure 2.22C). This effect could be due to *C. glutamicum* ATCC13032 encodes for two GntR-like regulators (GntR1 and GntR2) that are described to have redundant functions. In fact, phenotypic characterization of  $\Delta gntR1$  or  $\Delta gntR2$  strains shows wild-type phenotype and only growth defect in glucose when both genes are deleted (Frunzke et al., 2008).



**Figure 2.22** – Transcription of *ptsG* is not altered in the  $\Delta efp$  mutant. **(A)** Single-cell fluorescence microscopy of *C. glutamicum* cells expressing chromosomally encoded GntR2-mCherry in *efp*<sup>+</sup> and *efp*<sup>-</sup> strains. An 800-ms exposure time was used. **(B)** Distribution of mCherry fluorescence levels detected in samples containing a minimum of 300 cells of each *C. glutamicum* strain. The black lines indicate the mean fluorescence of these cells, and the gray dashed line represents the background fluorescence of the untagged strain. Statistical analysis was done by using two-tailed t test. \*\*\* $p < 0.0001$ . **(C)** Fold change (wild type/ $\Delta efp$ ) of *ptsG* mRNA levels. Cells were cultivated in BHI medium supplemented with 2% (w/v) glucose to OD<sub>600</sub> 2. Red dots represent the values of four biological replicates, and the mean and standard deviation are indicated.

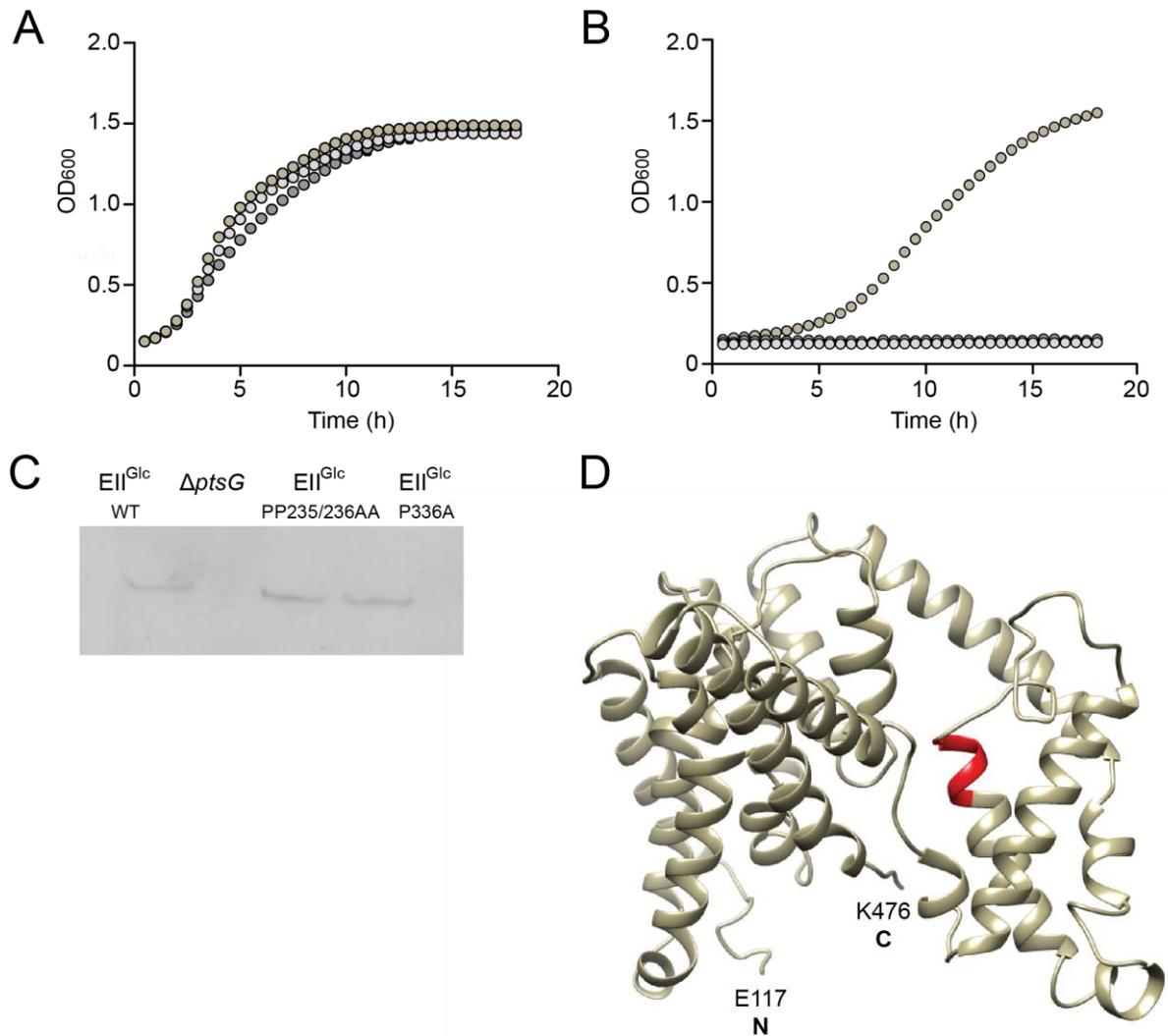
## 2.12 The polyproline motif is essential for EII<sup>Glc</sup> activity

Besides copy number control, a polyproline motif can be essential for a protein function (Motz and Jung, 2018). To get more details on further functions of the XPPX in EII<sup>Glc</sup>, the polyproline motif was exchanged by another amino acid sequence that do not require EF-P for translation, such as XAAX. The strain in which the native *ptsG* sequence was replaced by a synthetical variant coding for two consecutive alanine was successfully constructed (EII<sup>Glc</sup>-PP235/236-AA). Although this strain did not differ in growth against the wild type in rich medium BHI (Figure 2.23A), it was not capable to grow on glucose as the only carbon source (Figure 2.23B). Further exchange of only one of the prolines by alanine, still disrupting the polyproline motif (EII<sup>Glc</sup>-P236A), did not alter the phenotype (Figure 2.23B). Amino acid exchanges did not affect protein production as confirmed by Western blot (Figure 2.23C). In a next experimental setup, a *ptsG* library was constructed using mismatched pairs or primers in which the proline codons were exchanged by random amino acid codons. Several of the resulting clones were sequenced, confirming the diversity of sequences generated (Table 2.7). Nevertheless, introducing the EII<sup>Glc</sup>-variants in the  $\Delta ptsG \Delta ioT1 \Delta ioT2$  strain, originally unable to grow on glucose, only allowed growth when the wild-type sequence was inserted.

The XPPX motif is located in the EIIC domain of the protein, responsible for the translocation of the carbohydrate (Figure 2.16). Although the members of the EIIC glucose superfamily have low sequence identity, they are reported to have identical topologies (McCoy et al., 2016; Nguyen et al., 2006). The EIIC component of the maltose specific PTS component from *B. cereus* had its structure previously solved and shares 18% sequence identity with *C. glutamicum* EII<sup>Glc</sup> (McCoy et al., 2016). Using *B. cereus* MalT to predict the structure of *C. glutamicum* EII<sup>Glc</sup>, the XPPX motif was placed in a transmembrane domain involved in the dimerization of the PTS component (Figure 2.23D) and might play a vital role in functionality.

**Table 2.7** - EII<sup>Glc</sup> motif sequences generated using a mismatched pair of primers exchanging one of the prolines of the by a random codon.

VFXPL	VFPXL
VFIPL	VFPFL
VFRPL	VFPLL
VFTPL	VFPVL
VFYPL	VFPHL
VFEPL	VFPQL
VFKPL	VFPKL
VFQPL	VFPDL
VFPPPL	VFPDL
	VFP*L (stop codon)



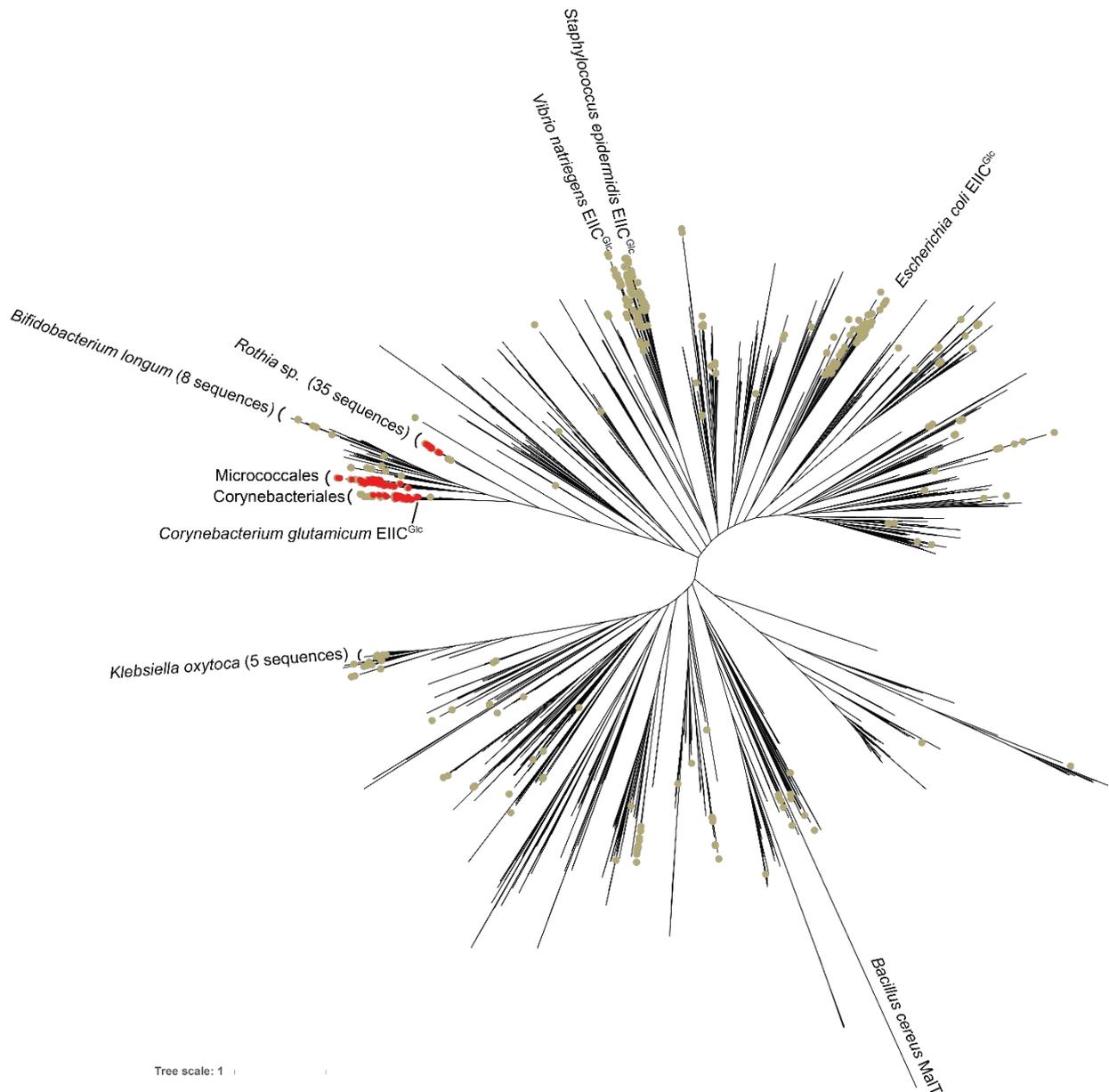
**Figure 2.23** – The polyproline motif is essential for EII<sup>Glc</sup> function. **(A)** Growth curves of wild-type *C. glutamicum* (brown), *C. glutamicum* mNG-EII<sup>Glc</sup> PP235/236AA (light gray) and mNG-EII<sup>Glc</sup> P236A (dark gray) in complex BHI medium. **(B)** Growth curves of the same strains as in **(A)** in minimal medium with glucose as sole carbon source (2% w/v). **(C)** Western blot analysis of mNG-EII<sup>Glc</sup> in whole cell lysates of the same *C. glutamicum* strains as in **(A)**. The  $\Delta ptsG$  mutant was included as negative control. **(D)** 3D structure of *C. glutamicum* EII<sup>Glc</sup> as predicted by Phyre2 (Kelley et al., 2015). Subunit C of EII<sup>Glc</sup> is shown from the side, and the position of the VFPPPL stalling motif is marked in red.

### 2.13 Polyprolines are widespread in EIIC subunits

PTS systems are comprised of a sugar-specific permease EIIC, as well as two other domains responsible for the phosphorylation of the incoming sugar EIIA and EIIB. These components might or not be part of one single protein, as well as can be structured in different orders depending on the organism. The sequence of *C. glutamicum* EII<sup>Glc</sup> was used to identify and collect similar amino acid

## Results

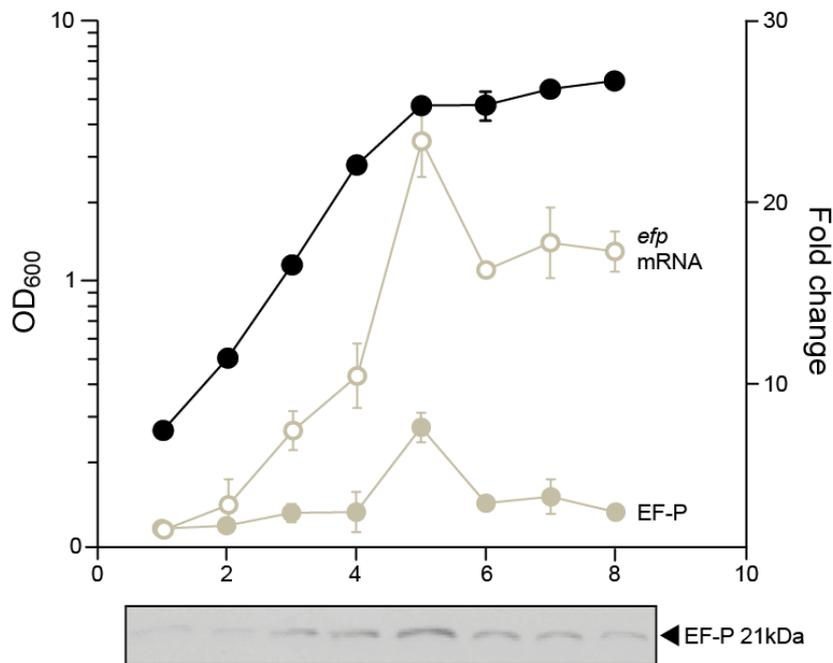
sequences. From the 4219 non-redundant sequences collected, 14.6% have XPPX motifs, including permeases sequences from *E. coli*, *Vibrio natriegens*, *Staphylococcus epidermidis*, *Klebsiella oxytoca* and might, as well, have protein production affected by EF-P. The amino acids up- and downstream the polyproline motifs highly differ among the sequences, and corroborates to the low amino acid sequence conservation among the EIIC of the glucose superfamily. The same motif found in *C. glutamicum* (VFPPL) is present in 129 sequences analyzed – mostly belonging to species from the order of Corynebacteriales, and closely Micrococcales and genus *Rothia* (Figure 2.24).



**Figure 2.24** – Phylogenetic tree of EIIC permeases from the glucose superfamily, sequences containing VFPPL motif (red dots) or general EF-P-dependent XPPX motifs (light brown) are highlighted. Sequence of *B. cereus* MaIT was added for comparison.

## 2.14 EF-P accumulates in Actinobacteria during early stationary phase

EF-P was not required for exponential growth of *C. glutamicum* in rich medium BHI (Figure 2.17E) raising the question whether EF-P is produced in this growth stage. *C. glutamicum* was grown, and samples were collected to simultaneously measure transcription and translation of EF-P at the different time points and growth stages. In *E. coli*, EF-P is more abundant in the exponential growth phase with around 0.1 copies per ribosome (An et al., 1980). To the contrary, *C. glutamicum* EF-P accumulates within the cell reaching the maximum level at the beginning of the stationary phase (Figure 2.25). This interesting analysis not only support the lack of growth impairment in rich medium, but outlines a time point which ribosomal rescue by EF-P might have the most relevance in Actinobacteria: the early stationary phase.



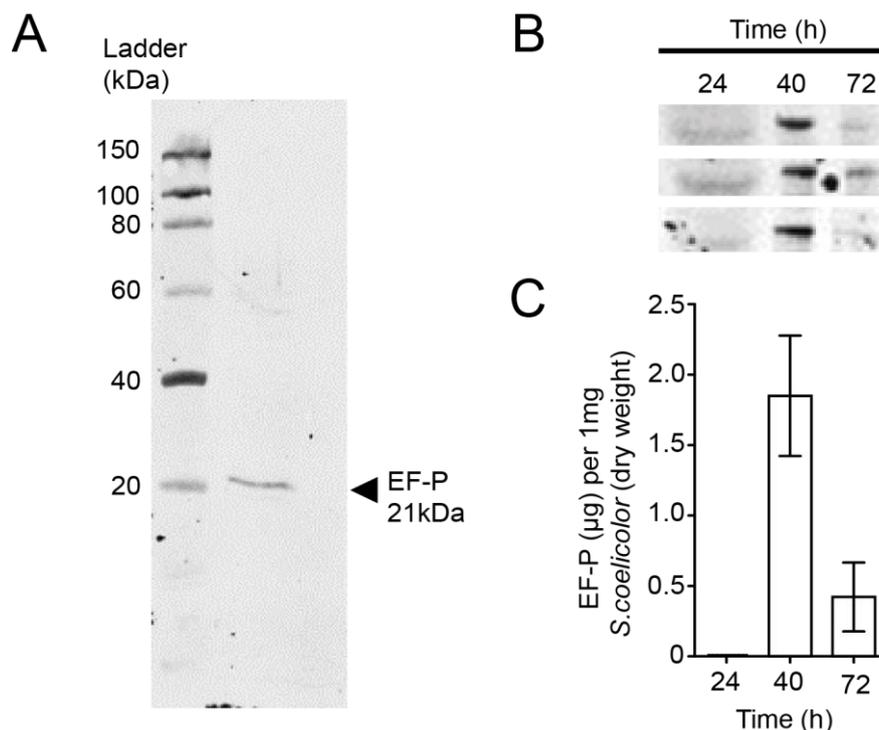
**Figure 2.25** – EF-P quantification over *C. glutamicum* growth: *efp* expression and EF-P production are enhanced in the beginning of stationary phase as shown. OD<sub>600</sub> - black dots, mRNA fold change - open brown circles, protein fold change – brown dots. Standard deviation bars were calculated after three independent replicates.

The early stage of the stationary phase in Actinobacteria such as *C. glutamicum* and, most dramatically, in *Streptomyces* species is accompanied by a major metabolic switch resulting in a strong activation of secondary metabolism, including production and secretion of several amino acids and antibiotics (Amara et al., 2018; Kolter et al., 1993; Nieselt et al., 2010).

To confirm if the accumulation of EF-P in the early stationary phase is a common feature in Actinobacteria, cultures of *S. coelicolor* were grown and samples for EF-P protein quantification were

## Results

collected at three time points: exponential growth, early stationary and late stationary phase. For native EF-P protein quantification we firstly recombinantly produced EF-P<sub>Sc</sub> and screened for a specific antibody suitable for Western blot detection. Antibodies against *B. subtilis* EF-P (1:1,000) were able to specifically detect EF-P<sub>Sc</sub> in *E. coli* cell lysates (Figure 2.26A). Quantitative Western blot of the *S. coelicolor* revealed increases in EF-P production in the early stage of the stationary phase as observed previously for *C. glutamicum* however, in a more dramatic fold change, of around 250-fold (Figure 2.26B and C).



**Figure 2.26** – Quantification of EF-P proteins in *S. coelicolor* cell lysates. **(A)** Western blot membrane of SDS-polyacrylamide gel from *E. coli* recombinantly producing EF-P<sub>Sc</sub>, primary rabbit-antibody against EF-P<sub>Bs</sub> (1:1,000) was used. **(B)** EF-P bands of Western blot membrane from *S. coelicolor* cell lysates from different growth stages: exponential growth (24h), early stationary phase (40h) and late stationary (72h). 100µg of total proteins were loaded in the lanes 24h and 15 µg in lanes 40h and 72h. **(C)** EF-P proteins in the three stages of *S. coelicolor* growth by quantitative Western blot.

### 2.15 EF-P is not required for growth in *S. coelicolor*

The public available reference proteomes of several *Streptomyces* species were analyzed regarding number and the staling strength of the polyprolines encoded (Table 2.8). Almost all *Streptomyces* species code for more XPPX motifs than the total number of proteins. Among them, *S. clavuligerus* is the specie with the highest frequency of XPPX motifs (1.48). This specie is described to produce more than twenty different secondary metabolites, including many beta-lactam antibiotics such as clavulanic

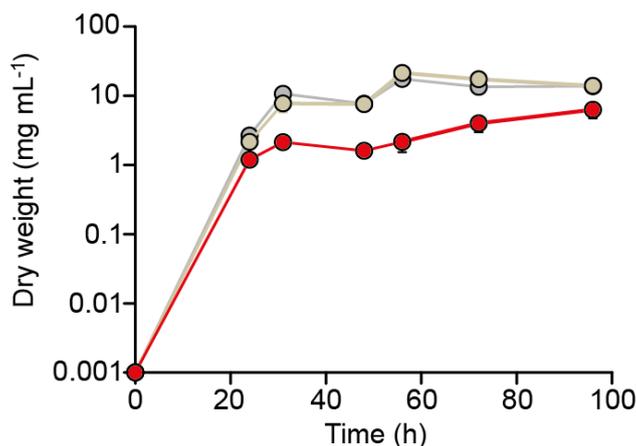
## Results

acid (Song et al., 2010). The model organisms *S. coelicolor* A(3) and *S. venezuelae* ATCC 10712 also encode for a high frequency of XPPX motifs, respectively, 1.08 and 1.17 XPPX motifs per protein encoded.

**Table 2.8** – *Streptomyces* species encode for a high number of XPPX containing proteins. Strength of the stalling motif as defined previously (Qi et al., 2018).

Streptomyces strain	Proteome ID	Proteins encoded	Number of XPPX motifs			Frequency Motifs/Proteins
			Weak	Moderate	Strong	
<i>S. coelicolor</i> A3(2)	UP000001973	8038	3508	1807	3406	<b>1.08</b>
<i>S. viridosporus</i> ATCC 14672	UP000003824	7865	3481	1790	3332	<b>1.09</b>
<i>S. peucetius</i> subsp. <i>caesioides</i> ATCC 27952	UP000230011	6389	2693	1435	2454	<b>1.03</b>
<i>S. albus</i> ATCC 21838	UP000031523	7322	3494	1901	3543	<b>1.22</b>
<i>S. natalensis</i> ATCC 27448	UP000032458	6559	2834	1472	2494	<b>1.04</b>
<i>S. avermitilis</i> ATCC 31267	UP000000428	7672	3543	1835	3127	<b>1.11</b>
<i>S. pristinaespiralis</i> ATCC 25486	UP000002805	6866	2829	1476	2585	<b>1.00</b>
<i>S. pratensis</i> ATCC 33331	UP000002066	6558	2793	1453	2640	<b>1.05</b>
<i>S. venezuelae</i> ATCC 10712	UP000006854	7451	3504	1807	3384	<b>1.17</b>
<i>S. cattleya</i> ATCC 35852	UP000007842	7537	3764	1912	3819	<b>1.26</b>
<i>S. clavuligerus</i> ATCC 27064	UP000002357	7290	4129	2178	4503	<b>1.48</b>
<i>S. himastatinicus</i> ATCC 53653	UP000003963	9177	4272	2146	4141	<b>1.15</b>
<i>S. ambofaciens</i> ATCC 23877	UP000061018	7568	3248	1748	3121	<b>1.07</b>
<i>S. rubellomurinus</i> ATCC 31215	UP000033699	6267	2601	1293	2090	<b>0.95</b>
<i>S. cinnamomeus</i> ATCC 21532	UP000222531	5593	2502	1412	2509	<b>1.15</b>
<i>S. nodosus</i> ATCC 14899	UP000031526	5863	2406	1229	2064	<b>0.97</b>
<i>S. armeniacus</i> ATCC 15676	UP000254425	6639	3077	1624	3489	<b>1.23</b>
<i>S. silvensis</i> ATCC 53525	UP000054804	7592	3352	1714	3340	<b>1.11</b>
<i>S. roseus</i> ATCC 31245	UP000035932	6493	2681	1470	2538	<b>1.03</b>
<i>S. actuosus</i> ATCC 25421	UP000247634	7068	3273	1645	2946	<b>1.11</b>
<i>S. showdoensis</i> ATCC 15227	UP000265325	6933	2862	1589	2721	<b>1.03</b>
<i>S. laurentii</i> ATCC 31255	UP000217676	7440	3371	1743	3048	<b>1.10</b>
<i>S. tateyamensis</i> ATCC 21389	UP000248039	6427	2756	1583	2369	<b>1.04</b>
<i>S. griseochromogenes</i> ATCC 14511	UP000092659	9005	3835	1987	3400	<b>1.02</b>
<i>S. xinghaiensis</i> ATCC 19609	UP000028058	6234	3058	1635	3510	<b>1.32</b>
<i>S. vitaminophilus</i> ATCC 31673	UP000050867	5132	2465	1319	2556	<b>1.24</b>
<i>S. setae</i> ATCC 33774	UP000007076	7443	3867	2031	3689	<b>1.29</b>
<i>S. aureofaciens</i> ATCC 10762	UP000037395	7620	3361	1729	2743	<b>1.03</b>

A *efp* deletion mutant of *S. coelicolor* was constructed and had its growth curve compared to the parental strain. As described for *C. glutamicum*, the exponential growth of *S. coelicolor* in rich medium was not affected by the deletion of *efp* (Figure 2.27).

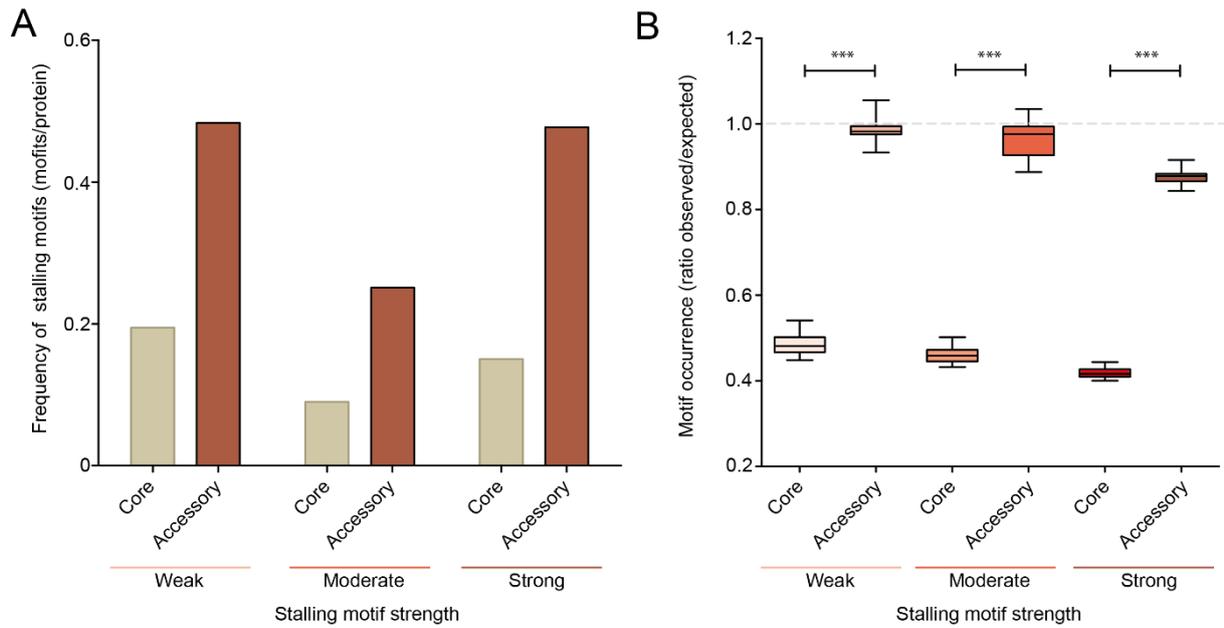


**Figure 2.27** – Growth curve of *S. coelicolor* A3(2) wild type (light brown), *efp* deletion mutant (red) and complemented strain (gray). Cells were grown in chemically defined medium. Dots represent the mean and bars the standard deviation of three independent replicates. Data obtained in collaboration with Prof. Eriko Takano and Dr. Suhui Ye (University of Manchester).

## 2.16 Polyproline motifs are underrepresented in the *S. coelicolor* core proteome

*S. coelicolor* do not require EF-P rescue activity during exponential growth. Next, it was asked whether proteins from secondary metabolism requires EF-P for translation. For that, the proteome of *S. coelicolor* A3(2) was divided into two groups: core proteome – proteins with homologs found in all *Streptomyces* species – and accessory proteome – proteins which are not present in all *Streptomyces* species sequenced thus far. With this strategy, the proteins more likely to be important or essential for primary metabolism and growth were grouped as belonging to the core, and the proteins responsible for regulation, production, and secretion of secondary metabolites as accessory proteins. Proportionally less proteins in the core proteome encode for XPPX motifs when compared to the accessory proteome, regardless of the motif stalling strength (Figure 2.28A). Subsequently, the amino acid sequences of those proteins were shuffled several times and the number of XPPX motifs in them were quantified. This strategy reveals if the number of polyprolines in each protein group are randomly encoded, or if there is some sort of selection involved to maintain or neglect such motifs. It was not found significant difference between observed and expected number of XPPX motifs in the accessory proteome, however, the XPPX frequency in the core proteome turned out much lower than expected (Figure 2.28B). This analysis supports that there is a strong selection pressure against XPPX motifs in the core proteome, while the accessory proteome do not undergo such counter selection. As comparison, similar analysis done in *E. coli* proteomes found such selection pressure against XPPX motifs in both protein groups (Qi et al., 2018).

## Results



**Figure 2.28** - Polyproline motifs encoded by *S. coelicolor* A3(2) genome. **(A)** Percentage of proteins containing XPPX motifs in the accessory and core proteome of *S. coelicolor* divided by stalling strength. **(B)** Ratio observed versus expected of XPPX motifs in those proteins. Statistical significance was addressed with the two-tailed t test. \*\*\* $p < 0.0001$ .

### 2.17 Proteins required for secondary metabolite production contain polyprolines

Three antibiotics produced by *S. coelicolor* are used as model for studies on secondary metabolite production in Actinobacteria. They are the coelimycin P1 (CPK), undecylprodigiosin (RED) and actinorhodin (ACT). The pathways of biosynthesis of each metabolite was analyzed to predict how EF-P would affect their production. Several XPPX motifs, including motifs with strong EF-P dependency, are present in all stages of production of the secondary metabolites (Table 2.9 for CPK, 2.10 for RED and 2.11 for ACT). Many enzymes contain more than one motif in the sequence and this number could be up to 15 motifs, as it was observed for the acyl-synthetase encoded by SCO5892 essential for RED production.

## Results

**Table 2.9** – Gene ID, Protein ID, function, XPPX motif and strength of enzymes and transporters predicted to be involved in Coelimycin P1 (CPK) production and secretion (Amara et al., 2018).

Gene ID	Protein ID	Function	XPPX motifs	Strenght
SCO6265	Q7AKF1	Butyrolactone-responsive repressor protein	QEPPQ	Moderate
SCO6266	Q7AKF0	Butenolide synthase involved in SCB1 biosynthesis	-	
SCO6267	Q9RKS7	Reductase involved in SCB1 biosynthesis	-	
SCO6268	Q9RKS6	Two-component system histidine kinase	VLPLL	Weak
			TRPPG	Strong
SCO6269	Q9RKS5	$\alpha$ -Ketoacid-dependent ferredoxin reductase $\beta$ -subunit	-	
SCO6270	Q9RKS4	$\alpha$ -Ketoacid-dependent ferredoxin reductase $\alpha$ -subunit	AFPPG	Moderate
SCO6271	Q9EWW4	Acyl-CoA carboxylase $\alpha$ -subunit (biotinylated)	GPPPH	Strong
SCO6272	Q9EX55	Secreted flavin-dependent epoxidase/dehydrogenase	GAPPA	Strong
SCO6273	Q9EX54	Polyketide synthase module 5	EPPPA	Moderate
			QWPPA	Moderate
			RLPPT	Weak
			FYPPI	Weak
SCO6274	Q9EX53	Polyketide synthase modules 3 and 4	SAPPV	Weak
			VWPPT	Moderate
			YAPPV	Strong
SCO6275	Q8CJN6	Polyketide synthase loading module, and modules 1 and 2	EVPPG	Strong
			GEPLL	Moderate
			GLPPV	Weak
			APPPS	Strong
			AAPPA	Strong
			VWPPE	Strong
			VRPPG	Strong
			FRPPR	Weak
			GEPLL	Moderate
			PWPPA	Moderate
			YAPPV	Strong
SCO6276	Q93S13	Secreted flavin-dependent epoxidase/dehydrogenase	DHPPT	Moderate
SCO6277	Q93S12	Isomerase	-	
SCO6278	Q93S11	Transmembrane efflux protein	AAPPE	Strong
SCO6279	Q93S10	Pyridoxal-dependent aminotransferase	TVPPR	Weak
SCO6280	Q93S09	SARP-family transcriptional activator	QPPPA	Strong
			PAPPG	Strong
SCO6281	Q93S08	Secreted flavin-dependent epoxidase/dehydrogenase	-	
SCO6282	Q93S07	Nicotinamide-dependent dehydrogenase	-	
SCO6283	Q93S06	unknown function	-	
SCO6284	Q93S05	Acyl-CoA carboxylase $\beta$ -subunit	ELPPA	Weak
SCO6285	Q93S04	Hypothetical protein (unknown function)	-	
SCO6286	Q93S03	Butyrolactone-responsive repressor protein	-	
SCO6287	Q9LAS9	Type II thioesterase	REPLL	Moderate
			VMPLL	Weak
			PGPLL	Strong
SCO6288	Q93RY4	SARP-family transcriptional activator	DDPPR	Strong
			DPPPG	Strong

## Results

**Table 2.10** – Gene ID, Protein ID, function, XPPX motif and strength of enzymes and transporters predicted to be involved in Undecylprodigiosin (RED) production and secretion (Amara et al., 2018).

Gene ID	Protein ID	Function	XPPX motifs	Strenght
SCO5878	O54142	pyrrole-2-carboxyl intermediate transferase	GIPPA	Moderate
			TVPPV	Weak
			QPPPG	Strong
			RIPPH	Weak
			TLPPS	Weak
SCO5879	REDW	L-prolyl-RedO oxidase	-	
SCO5886	O54149	3-oxoacyl-[acyl-carrier-protein] synthase (n-C8:0)	TVPPT	Moderate
SCO5888	FABH3	beta-ketoacyl-ACP synthase (2) (undecylprodigiosin ACP)	GGPPR	Strong
SCO5889	O54152	L-prolyl-adenylate-PCP transaminoacylase	-	
SCO5890	O54153	L-serine acyltransferase/alpha-oxoamine synthase	-	
SCO5891	REDM	L-prolyl-adenylate synthase	LLPPG	Weak
			DAPPR	Strong
			LHPPG	Strong
			AYPPD	Strong
			WLPPY	Weak
			ADPPV	Strong
			EVPLL	Weak
			PAPPA	Strong
			FQPPA	Moderate
			EHPPA	Moderate
			AAPPA	Strong
			AAPPV	Strong
			PAPPH	Strong
			GYPPH	Weak
PYPPQ	Moderate			
SCO5892	O54155	acyl-[acyl-carrier-protein] synthetase (n-C12:0)	FPPPA	Strong
			PAPPS	Strong
SCO5893	O54156	2-undecyl-4-pyrrolinone reductase/dehydratase	-	
SCO5895	O54158	4-hydroxy-2_2'-bipyrrole-5-methanol methyltransferase	-	
SCO5896	Q8CJQ2	undecylprodigiosin synthase	PVPPA	Moderate
			PEPPL	Moderate

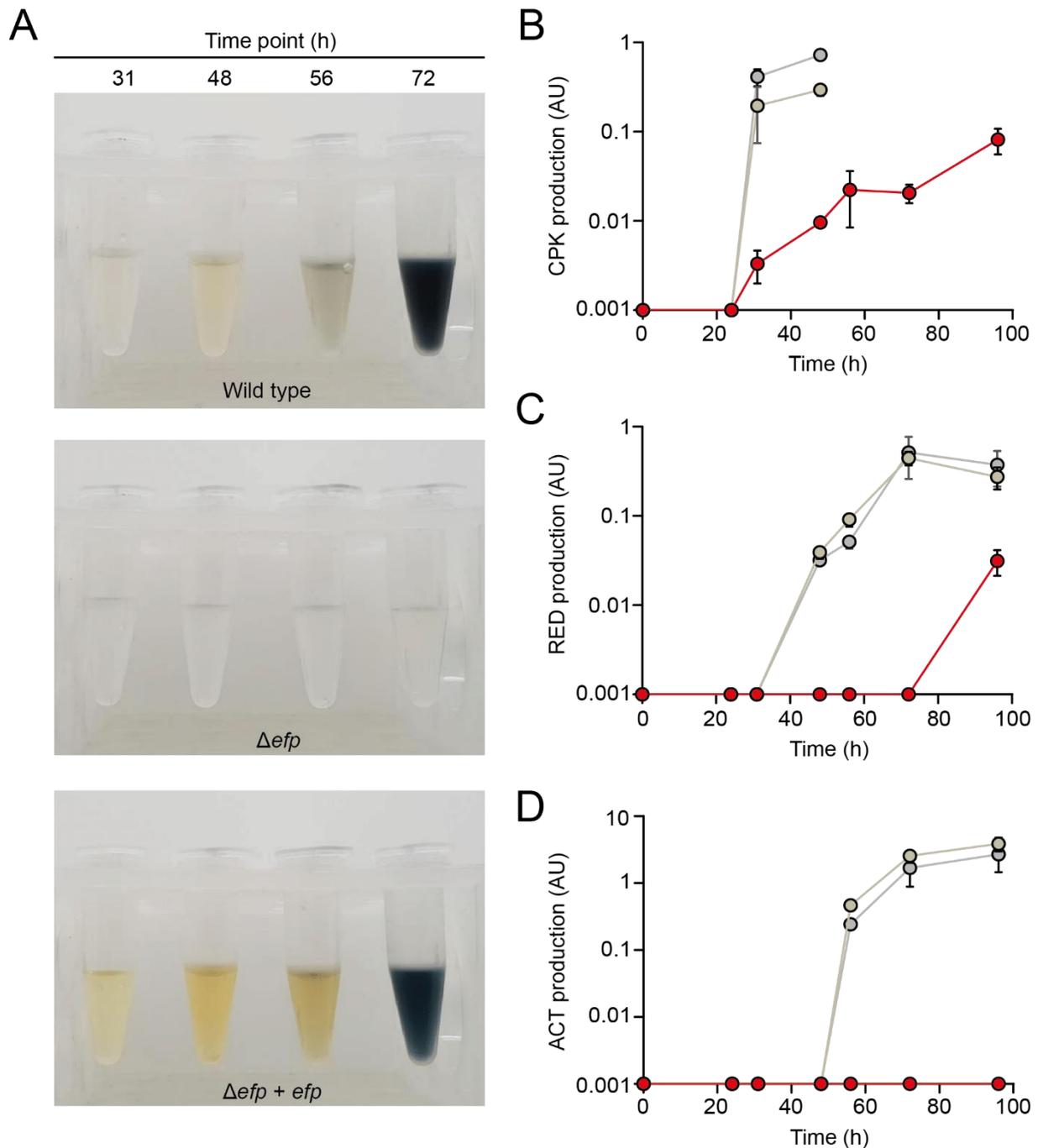
## Results

**Table 2.11** – Gene ID, Protein ID, function, XPPX motif and strength of enzymes and transporters predicted to be involved in Actinorhodin (ACT) production and secretion (Amara et al., 2018).

Gene ID	Protein ID	Function	XPPX motifs	Strenght
SCO5023	Q9KY79	actVI oxidoreductase	WSPPA	Moderate
			SFPPG	Moderate
			DPPPA	Strong
			DAPPC	Strong
SCO5071	Q7AKI2	actVI dehydratase	-	
SCO5072	Q7AKI1	hydroxylacyl-CoA dehydrogenase	FNPPH	Weak
SCO5074	Q9ADD7	actVI dehydratase	-	
SCO5075	Q8CJS3	actVI oxidoreductase	VKPPE	Strong
SCO5080	Q7AKH7	actVA-5/actVB monooxygenase	-	
SCO5081	Q7AKH6	actVA-5/actVB monooxygenase	MGPPK	Moderate
SCO5083	ACT22	actinorhodin transport via facilitated transport	TAPPS	Strong
			GPPPY	Strong
			MFPPK	Moderate
			RLPPK	Weak
SCO5084	MMPLA	actinorhodin transport via facilitated transport	VAPPQ	Strong
SCO5086	ACT3	actIII ketoacyl reductase	-	
SCO5087	KASA	polyketide chain elongation 1	FEPPS	Strong
			VVPPT	Moderate
SCO5088	KASB	polyketide chain elongation 1	-	
SCO5089	ACPX	Malonyl-CoA-ACP transacylase (actinorhodin ACP)	-	
SCO5090	CYPK	actVII aromatase	TQPPI	Weak
			RQPPE	Strong
			LVPPA	Moderate
SCO5091	Q93IZ0	actIV cyclase	LRPPN	Strong
SCO5092	DIM6	actVA-5/actVB monooxygenase	MEPPL	Moderate

### 2.18 EF-P is essential for antibiotic production in *S. coelicolor*

The increase of EF-P copy number in early exponential growth phase and the high frequency of XPPX motifs found in proteins responsible for secondary metabolite production, support the relevance of EF-P for the translation of XPPX-rich enzymes of secondary metabolism and ultimately, antibiotics production. In the next experimental set up, the native production of the antibiotics CPK, RED and ACT by *S. coelicolor*  $\Delta efp$  was quantified and compared to the production of the parental strain. As control, the mutant strain was complemented by reintroducing *efp* into the native chromosomal locus. The production of antibiotics by the *efp* deletion strain was virtually abolished (Figure 2.29), while the complemented strain recovered the production capacity of the wild type (Figure 2.29, gray curves).



**Figure 2.29** - Quantification of coelimycin P1 (CPK), undecylprodigiosin (RED) and actinorhodin (ACT) production by *S. coelicolor* A3(2) wild-type, *efp* deletion mutant and complemented strain. **(A)** Supernatant of wild-type strain,  $\Delta ef p$  deletion mutant and complemented strain grown in chemically defined medium and collected at the time points indicated. Quantification of **(B)** CPK, **(C)** RED and **(D)** ACT production by analysis of the supernatant of different strains. Wild type (light brown),  $\Delta ef p$  (red) and complemented strain (gray). Dots represent the mean and bars the standard deviation of three independent replicates. Data obtained in collaboration with Prof. Eriko Takano and Dr. Suhui Ye (University of Manchester).

### 3 Discussion

Actinobacteria such as *C. glutamium* are important workhorses for the industrial production of amino acids, peptides and other secondary metabolites. Proteomic analysis of the *C. glutamicum*  $\Delta efp$  mutant reveals downregulation of 222 proteins, most of which containing a XPPX motif. Moreover, this set is not a random sample of the proteome but is enriched in proteins involved in 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, for example, 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 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.

A protocol for the isolation of native EF-P proteins from bacterial lysates was set up in order to study the post-translational modification status of EF-P in diverse Actinobacteria. This protocol consisted of two sequential chromatographic steps that resulted in EF-P purity of around 95%. Although the protein enrichment was probably enough for the first analytical measurements, limitations to detect the EF-P<sub>Cg</sub>-enriched fractions after the chromatographies forced the migration to another purification method with an even better outcome – fusion of a histidine-tag to the native chromosomally encoded *efp* and further affinity chromatography purification. This strategy was successfully used in *C. glutamicum* and other Actinobacteria described here, and only possible due to the recent emergency of molecular tools suitable for genomic engineering in Actinobacteria.

Bacteria have evolved many 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, a novel EF-P subfamily in Actinobacteria that does not require any PTM for activation is described. The structure of this EF-P subclass was obtained by crystallizing the endogenously produced EF-P<sub>Cg</sub>. The three  $\beta$ -barrel domains, which together form a 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 (Figure 2.9).

The functionally important Lys32 at the tip of the  $\beta$ -hairpin is encompassed by two prolines at position 30 and 34. Replacement of either of them by alanine, asparagine, glycine or glutamine (amino acids, which are found in other EF-P subfamilies) results in an inactive EF-P. Confirming that the two proline residues are essential for EF-P<sub>Cg</sub> activity. Due to the high amino acid sequence and structure conservation of EF-Ps, no other characteristic motif rather than the palindromic loop sequence PGKGP could be identified as unique for this new EF-P subfamily. Here, it is argued that this not only stabilizes the P-site tRNA but also forces the prolines to adopt an alternative conformation and thus optimal

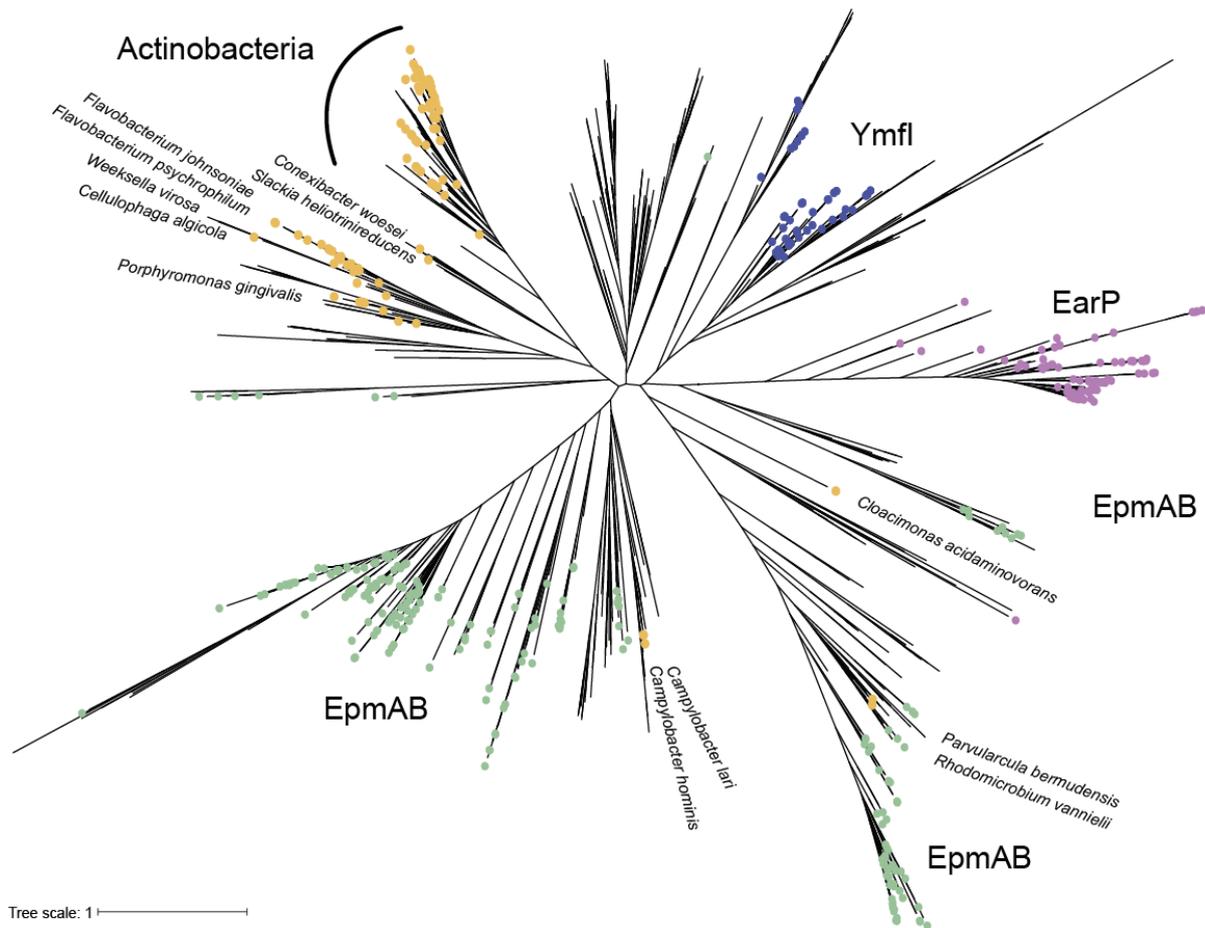
geometry between the nascent chain and the aminoacyl-tRNA for peptide bond formation (Huter et al., 2017).

Mechanistically, we cannot 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, we suggest that the  $\beta$ -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 facilitate 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 2.11). This loop sequence is particularly found in EF-Ps of the Actinobacteria class (Figure 2.12), and it is definitely not present in EF-Ps of bacteria that possess EF-P modification enzymes.

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 it was shown above, the consensus loop sequence is not only essential for the functionality of this EF-P subclass, it also enabled the structure determination of the  $\beta 2\Omega\beta 3$  loop region with unprecedented resolution.

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 (Hummels and Kearns, 2020; Lassak et al., 2016). 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 sidechain at position 32 is still essential for activity. Similarly, in *B. subtilis* the unmodified EF-PK<sup>32R</sup>, but not the EF-PK<sup>32A</sup> 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.

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  $\beta 2\Omega\beta 3$  loop in 11% of all bacteria, which suggests that EF-P in other bacteria also does not require a PTM for activation (Figure 3.1). 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.



**Figure 3.1.** Proline at position 34 of EF-P (represented by yellow dots) is only found in bacterial species that lack homologs for EF-P modification enzymes, which are shown as: blue (YmfI), purple (EarP) and green dots (co-occurrence of EpmA and EpmB).

EII<sup>Glc</sup> is the major glucose transporter in *C. glutamicum*. Induction of carbohydrate transporters expression is one of the most popular strategies to improve bacterial amino acid and secondary metabolite production (Krause et al., 2010; Lindner et al., 2013; Xu et al., 2016). By theoretical predictions and growth analysis of the *C. glutamicum*  $\Delta efp$ , it was noticed that EF-P might play an important role during translation of carbohydrate transporters. Without changes in *ptsG* transcriptional levels, protein levels were significantly reduced when EF-P is not present. The direct effect of EF-P on EII<sup>Glc</sup> translation could also be confirmed *in vitro* by reduced translational rate. *In vivo*, *efp* deletion resulted in EII<sup>Glc</sup> low copy number, low glucose uptake rate, and impaired growth.

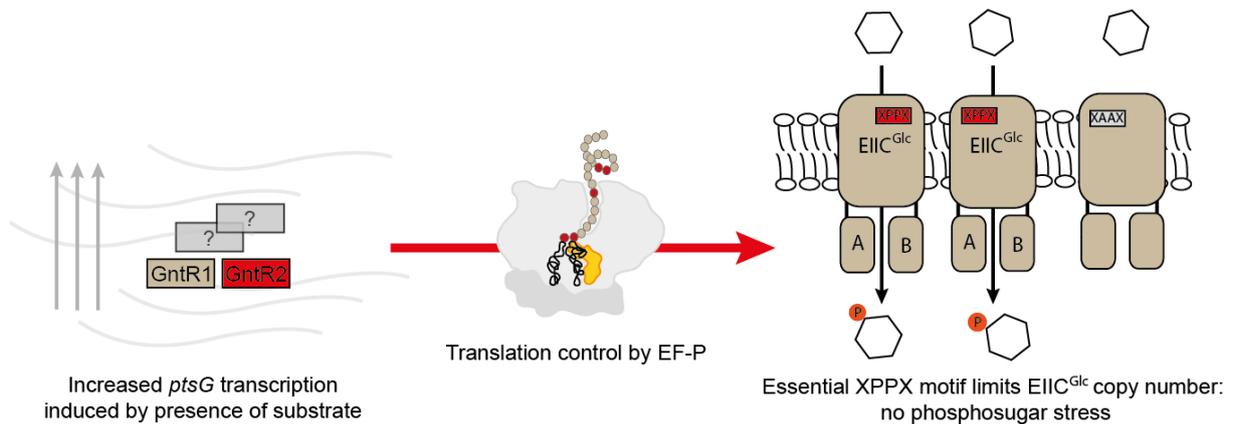
The growth impairment of the *efp* deletion mutant on different carbohydrates correlates to the stalling strength of the XPPX located in each specific carbohydrate transporter rather than in their transcriptional regulators. GntR2 is a global transcriptional regulator of sugar utilization and have levels influenced by EF-P. However, this influence in copy number was not propagated to *ptsG* transcription probably due to the redundant function with other regulators present in the *C. glutamicum* ATCC13032 strain (Frunzke et al., 2008).

Numerous studies have shown that logarithmic increases in *ptsG* transcription can be easily achieved by inducible promoters however, the increase in mRNA levels is only followed by a minor increase in glucose uptake rate (Krause et al., 2010; Lindner et al., 2013; Wang et al., 2014; Wang et al., 2018). Here, we argue that besides the complex network of *ptsG* transcriptional regulators, a translational control by EF-P play a significant role in protein copy number adjustment due to an essential XPPX motif. This translational regulation might prevent an unlimited increase of EII<sup>Glc</sup> molecules to protect *C. glutamicum* from the so-called phosphosugar stress. In *E. coli*, protection against phosphosugar stress is accomplished by a complex regulatory network comprising small RNA-initiated inhibition of *ptsG* translation, and Hfq-dependent *ptsG* mRNA degradation by RNase E (Maki et al., 2008; Morita et al., 2005). *C. glutamicum* does not possess an Hfq homologue (Kalinowski et al., 2003) and therefore, might use the polyproline-dependent stalling regulation instead.

Polyproline motifs are frequently found in protein-protein interaction sites. In EII<sup>Glc</sup>, the EF-P dependent polyproline motif is predicted to be located in a transmembrane domain of the subunit EIIC involved in dimer formation. Avoiding a bottleneck in EII<sup>Glc</sup> production by replacing one or the two prolines of this motif completely inhibited growth with glucose as sole carbon source. An unbiased semi-random mutagenesis approach further confirmed that the consecutive prolines are essential for functionality of EII<sup>Glc</sup>. Previous studies on the role of polyproline motifs in membrane-integrated proteins focused on the *E. coli* acid stress receptor CadC (Ude et al., 2013) and the osmosensor EnvZ (Motz and Jung, 2018). CadC has a strong polyproline motif that fine-tunes its copy number (Ude et al., 2013). A CadC variant in which the polyproline motif is replaced with polyalanine is characterized by 3-fold higher copy number and a less sensitive stress response. On the other hand, the polyproline motif in EnvZ did not affect receptor copy number, but was found to be essential for dimerization and interaction with the modulator MzrA (Motz and Jung, 2018). Here, it is proposed that the polyproline motif in *C. glutamicum* EII<sup>Glc</sup> has a dual function: it is essential for its functionality, but also fine-tunes the copy of this transporter (Figure 3.2).

A polyproline motif can be found in the permease subunits of diverse PTS sequences (Figure 2.24). Not only in sequences retrieved from Actinobacteria, but as well in Firmicutes (eg. *Staphylococcus*) and Proteobacteria (eg. *Escherichia coli*, *Klebsiella*), suggesting that bacteria from different phyla might address EII<sup>Glc</sup> function, copy-number fine tuning or both by the mechanism described here.

The proteome analysis of *C. glutamicum*  $\Delta$ *efp* highlighted the importance of this elongation factor for the translation of proteins involved in production of secondary metabolites. Here, it was further identified that EF-P proteins are eight times more abundant in *C. glutamicum* and 250 times more abundant in *S. coelicolor* cells in the early stationary phase. Generally, in Actinobacteria, this period is accompanied by a major metabolic switch, which induces amino acid, antibiotics and other secondary metabolite production.

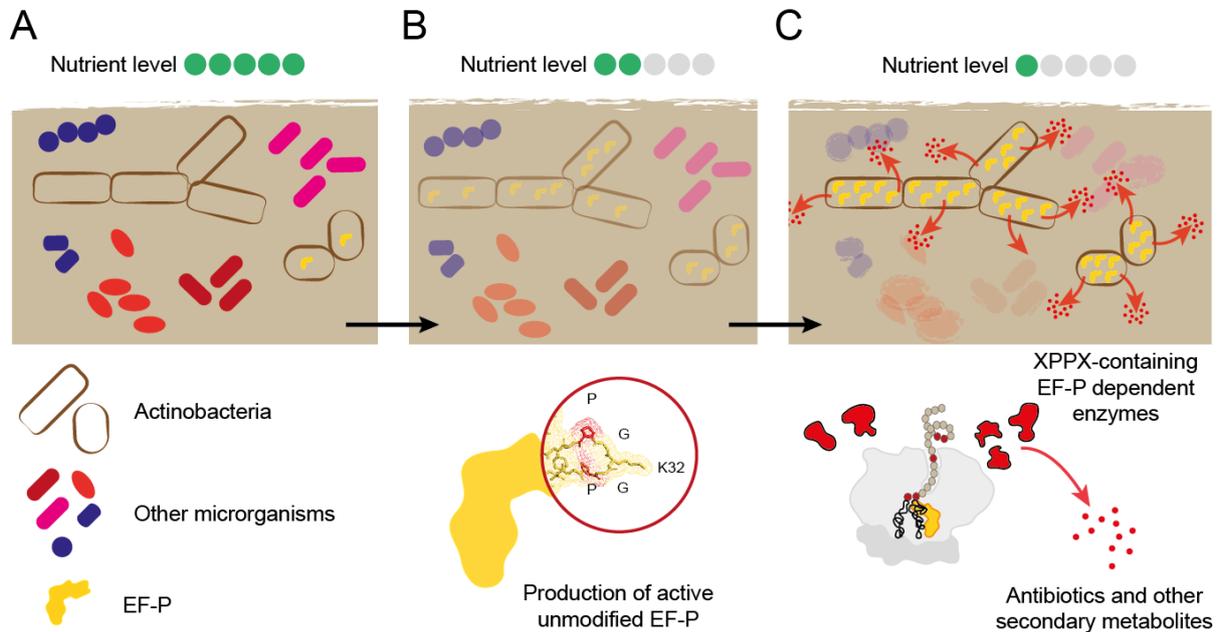


**Figure 3.2** – In *C. glutamicum*,  $EII^{Glc}$  copy-number fine tuning and functionality is addressed by an EF-P dependent, polyproline motif. Increases in *ptsG* transcription can be obtained by the presence of the substrate glucose or overexpression of one of its many transcriptional activators. Production of  $EII^{Glc}$ , however, is under EF-P control. This would be one mechanism to avoid an excess of transporters and the so-called phosphosugar stress. The polyproline motif is essential for  $EII^{Glc}$  activity, therefore, an eventual loss of the XPPX motif and production independently of EF-P would lead to unfunctional transporters.

*S. coelicolor* is one Actinobacteria known for the native production of secondary metabolites including several antibiotics (van Keulen and Dyson, 2014). Its genome codes for proteins containing one of the highest XPPX motif frequency among bacteria - 1.08 motifs per protein encoded. The hypothesis that EF-P is essential for antibiotics production in *S. coelicolor* is supported by *in silico* analysis - polyproline motifs are significantly less frequent in the proteins of the primary metabolism and abundant in enzymes responsible for antibiotics production; and *in vivo* experiments – lack of growth defect and virtually no antibiotic production by the  $\Delta efp$  strain.

Antibiotics synthesis are energetically expensive for the bacterial cell, requiring maintenance of tens of kilobases of DNA (around 20-30 enzymes) whereas only hundreds of bases are necessary for most resistance mechanisms (Hoskisson and Fernandez-Martinez, 2018; Nett et al., 2009; Traxler and Kolter, 2015; van Keulen and Dyson, 2014). Therefore, production in the right amounts at the right time is crucial to avoid emergence of resistant competitors. Antibiotics production pathways are under tight control of several transcriptional regulators (Bibb, 2005; Wei et al., 2018). In *S. coelicolor*, each antibiotic production pathway is controlled by one specific and essential transcriptional activator as well as under control of higher-level global regulators (Wei et al., 2018). Deletion or inactivation of the transcription activator not always leads to a complete inhibition of the corresponding antibiotic production (Fernandez-Moreno et al., 1991; Takano et al., 1992; Takano et al., 2005; Uguru et al., 2005; Wei et al., 2018). Actinorhodin (ACT) had production levels undetectable in the  $\Delta efp$  mutant strain (Figure 2.39D) while the deletion of the specific transcriptional activator of ACT pathway is described to lead only to a reduced production (Fernandez-Moreno et al., 1991; Uguru et al., 2005). This supports that EF-P plays a tight second-level control at the translation level of such enzymes, circumventing leaky production. When nutrients start to become limited, antibiotic production are a great advantage against competitor bacteria

towards nutrient scavenging, survival and spore formation. Here, we argue that up-regulating EF-P at this time point would permit translation of secondary metabolism enzymes circumventing earlier and unnecessary production of those metabolites (Figure 3.3).



**Figure 3.3** – Physiological role of EF-P in Actinobacteria. **(A)** Actinobacteria colonize diverse environments in combination with other microorganisms. EF-P is not necessary for Actinobacteria growth in nutrient rich conditions, and production is kept at minimal levels. **(B)** Actinobacterial EF-P do not require post-translational modifications for activation, it is optimized to be produced in nutrient limiting conditions. **(C)** Close to starvation, EF-P is abundant and enzymes rich in polyprolines can be translated. They are responsible for secondary metabolite production and function as an advantage for Actinobacteria during colonization and survival in diverse environments.

It is not clear why Actinobacteria evolved an EF-P that dispenses post-translational modifications. Hypothesis that bacteria might use activation and de-activation of EF-P by adding or erasing a PTM was raised but still lack experimental evidences (Hummels and Kearns, 2020). Modulating EF-P levels, however, is the strategy that Actinobacteria apply to regulate, at the translational level, production of proteins containing XPPX motifs. By lacking further activation mechanism, Actinobacteria is prompt to faster translational responses at a low energetic cost. A very interesting strategy used by both free-living and pathogenic bacteria, to survive in a continuously changing environment while directing energy for production of a wide range of secondary metabolites. Altogether, we present a new EF-P subfamily that regulates translation of specific pathways and carry great biotechnological potential.

### 3.1 Outlook

In this thesis, a novel subfamily of EF-P proteins was characterized regarding structure, mechanism of action and specific functions in the model organism *C. glutamicum* and *S. coelicolor*. Interesting questions arose from this study and can be subject of future research.

Actinobacteria, including the ones amongst the bacteria with the largest genomes, do not need a PTM to activate EF-P while other bacteria invest limited resources in the production of intricate substrates and modification enzymes. It would be interesting to introduce this new EF-P subclass into bacteria that usually modify EF-P to investigate the evolutionary advantages of it. The redirection of energetic resources into biomass or metabolite production could also be investigated.

In the same direction, increases in biomass or amino acid production by *C. glutamicum* could be tested by expressing a glucose transporter from another close-related organism. A transporter that, ideally, is functional without a polyproline motif. Moreover, the presence of a XPPX in some transporters but not in others suggests that EF-P might play a more general role in *C. glutamicum* metabolism regulation; this effect could also be investigated in more detail.

Actinobacteria are organisms that deal with many environmental stresses. It is described here that EF-P is upregulated when nutrients become limiting. It would be interesting to find its transcriptional regulator(s), perhaps a sigma factor, responsible for the stress-sensing and promoting *efp* transcription. Other stresses that might upregulate EF-P should also be investigated, especially in species of *Corynebacterium*, *Streptomyces* and *Mycobacterium* due to their relevance in biotechnology and human health.

Furthermore, designing of *C. glutamicum* and *S. coelicolor* constitutively over-expressing EF-P would be a very interesting starting point for the construction of industrial strains with increased efficiency towards production of recombinant proteins or secondary metabolites.

## 4 Materials and Methods

### 4.1 Construction of the phylogenetic trees

For the phylogenetic tree of EF-P proteins, a set of fully sequenced prokaryotic genomes identified by (Keilberg et al., 2012; 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 ymfl adh\_short or adh\_shortC2 domains as described 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 I-ins-I algorithm implemented in the MAFF software package version 7.409 (Kato et al., 2017). 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).

The phylogenetic tree of EIIC subunits from the glucose superfamily, the amino acid sequence of the EIIC domain from *C. glutamicum* EII<sup>Glc</sup> (Uniprot reference Q46072), corresponding to the amino acids from position 117 to 476 was downloaded and the 20.000 most similar sequences identified by Basic local alignment search tool (BLAST), excluding models and uncultured/environmental samples. Dataset was reined by deleting partial sequences, hypothetical proteins, sequences that do not contain domain IIC and identical sequences. Data was retrieved against Uniprot annotated database resulting in 4.218 non-redundant sequences. Sequence of *Bacillus cereus* MaIT was added for comparison. Protein sequences alignment was performed with MAFFT FFT-NS-2 method. Maximum likelihood protein trees were constructed with RAxML-HPC v.8 (Stamatakis, 2014). Tree display was done using iTOL v3.

### 4.2 Polyproline motifs stalling strength and frequency in proteomes

To search for polyproline motifs, reference proteomes for each selected strain were downloaded from UniProt (UniProt, 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). Core and accessory proteomes were identified as described previously (Qi et al., 2018) against all reference *Streptomyces* proteomes available in 15.07.2018. Proteomes were shuffled using the program shuffleseq from EMBOSS package (Rice et al., 2000), for graphical representation and statistical analysis, a minimum number of 100 shuffled sequences were used.

### 4.3 Nucleotides, plasmids and bacterial strains construction

Strains, plasmids and primers used in this study are listed in Tables 4.1 – 4.3. 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 T, 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* DH5 $\alpha$  (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 (Schafer 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 intergeneric conjugation of pTE1213 into wild type *S. coelicolor* M145 and subsequent apramycin selection, followed by a second intergeneric conjugation with pTE1208 and counterselection on kanamycin, and finally two consecutive passages on Soya Flour Mannitol (SFM) without antibiotics. Plasmids pCM4.4 (kindly provided by H. Zhao, unpublished) and pCMU-4K (S. Ye, unpublished) were used for *S. coelicolor* genome editing by CRISPR-Cas9 and for plasmid curing, respectively.

## Materials and Methods

**Table 4.1** – Strains used in this thesis

<i>E. coli</i>	Relevant characteristics	Source
DH5α	<i>fhuA2 lacΔU169 phoA glnV44 Φ80' lacZΔM15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17</i>	Promega
MG1655 <i>PcadBA::lacZ ΔcadBA Δefp</i>	K-12 reference strain MG1655 with a translational fusion of <i>PcadBA</i> and <i>lacZ</i> , deleted for <i>cadB</i> , <i>cadA</i> and <i>efp</i>	Ude et al., 2013
MG1655 <i>PcadBA::lacZ ΔcadBA Δefp ΔepmA</i>	K-12 reference strain MG1655 with a translational fusion of <i>PcadBA</i> and <i>lacZ</i> , deleted for <i>cadB</i> , <i>cadA</i> , <i>efp</i> and the modification enzyme <i>EpmA</i>	Ude et al., 2013
ET12567 [pUZ8002]	<i>dam dcm hsdM hsdS hsdR cat tet</i> , carrying plasmid pUZ8002	E. Takano
<i>C. glutamicum</i>	Relevant characteristics	Source
ATCC13032	Biotin-auxotrophic wild type	M. Bramkamp
ATCC13032 <i>Δefp</i>	ATCC13032 with in-frame deletion of <i>efp</i>	this study
ATCC13032 <i>efp</i> -histag	ATCC13032 native <i>efp</i> with translational fusion to 6xhis	this study
ATCC13032 <i>Δefp ::PdnaK-RPAP-egfp</i>	ATCC13032 <i>Δefp</i> with RPAP- <i>egfp</i> under control of Pdnak integrated into intergenic region cg1824-cg1825	this study
ATCC13032 <i>Δefp ::PdnaK-RPPP-egfp</i>	ATCC13032 <i>Δefp</i> with RPPP- <i>egfp</i> under control of Pdnak integrated into intergenic region cg1824-cg1825	this study
ATCC13032 <i>efp</i> -histag <i>Cgl1117-egfp</i>	ATCC13032, native <i>efp</i> with translational fusion to 6xhis and <i>cgl1117</i> with translational fusion to <i>egfp</i>	this study
ATCC13032 <i>Δefp Cgl1117-egfp</i>	ATCC13032 <i>Δefp</i> , native <i>Cgl1117</i> with translational fusion to <i>egfp</i>	this study
ATCC13032 <i>gntR2</i> -mCherry	ATCC13032 native <i>gntR2</i> (Cgl1718) with translational fusion to mCherry	this study
ATCC13032 <i>Δefp gntR2</i> -mCherry	ATCC13032 <i>Δefp</i> , native <i>gntR2</i> (Cgl1718) with translational fusion to mCherry	this study
ATCC13032 mNeonGreen- <i>ptsG</i>	ATCC13032 native <i>ptsG</i> (Cgl1360) with translational fusion to mNeonGreen	this study
ATCC13032 <i>Δefp mNeonGreen-ptsG</i>	ATCC13032 <i>Δefp</i> , native <i>ptsG</i> (Cgl1360) with translational fusion to mNeonGreen	this study
ATCC13032 mNeonGreen- <i>ptsG</i> PP235/236AA	ATCC13032 native <i>ptsG</i> (Cgl1360) with translational fusion to mNeonGreen and amino acid substitution PP235/236AA	this study
ATCC13032 mNeonGreen- <i>ptsG</i> P236A	ATCC13032 native <i>ptsG</i> (Cgl1360) with translational fusion to mNeonGreen and amino acid substitution P236A	this study
ATCC13032 <i>ΔptsG</i>	ATCC13032 with in-frame deletion of <i>ptsG</i>	M. Bramkamp
ATCC13032 <i>ΔptsG ΔioT1 ΔioT2</i>	ATCC13032 with in-frame deletion of <i>ptsG</i> , <i>ioT1</i> and <i>ioT2</i>	D. Petrov
<i>S. coelicolor</i>	Relevant characteristics	Source
M145	<i>SCP1- SCP2-</i> derivative from A3(2) wild type	E. Takano
M145 <i>Δefp</i>	M145 with in-frame deletion of <i>efp</i>	E. Takano
LW277	M145 native <i>efp</i> with translational fusion to 6xhis	this study
<i>M. smegmatis</i>	Relevant characteristics	Source
mc(2) 155	Derivative from mc(2) 154 wild type	ATCC
<i>M. tuberculosis</i>	Relevant characteristics	Source
H37Rv	Human-lung isolate	N. Reiling

## Materials and Methods

**Table 4.2** – Plasmids used in this thesis

Vector backbones	Feature and construction comments	Source
pBAD33 pEKEx2	Cam-cassette, p15A origin, araC coding sequence, ara operator Kan-cassette, E. coli-C. glutamicum shuttle vector for regulated gene expression. <i>Ptac lacI</i> , pBL1 <i>oriVC.g</i> , pUC18 <i>oriVE.c</i> .	Novagen (Merk Millipore) B. Eikmanns
pK19mobSacB pCM4.4	Kan-cassette, pK18 origin, <i>sacB lacZ</i> . For allelic exchange in <i>C. glutamicum</i> Apra-cassette, oriT, rep pSG5(ts), ori ColE1, ermE* <sup>p</sup> -sPcas9, synthetic guide RNA cassette. <i>E. coli-Streptomyces</i> shuttle vector for genetic recombination in <i>S.</i> <i>coelicolor</i> by CRISPR-Cas9	M. Bramkamp H. Zhao (unpublished)
pCMU-4K	Kan-cassette, Amp-cassette, oriT, rep pSG5(ts), ori ColE1, D4-ribo-Spcas9, synthetic guide RNA cassette. <i>E. coli-Streptomyces</i> shuttle vector for pCM4.4- derivative plasmid curing	S. Ye (unpublished)
pMycoFos	Kan-cassette, E.coli-Mycobacterium shuttle vector for gene expression.	Addgene; Ly et al., 2011
<b>Strain construction and crossosomal editing</b>		
pK19mobSacB-efp-histag	Translational fusion of native <i>efp</i> and a c-terminal 6x histag	this study
pK19mobSacB-efp-del	Contains up- and downstream regions of <i>efp</i> . For <i>efp</i> deletion	this study
pK19mobSacB-RPAP-egfp	<i>egfp</i> varinat contiang a RPAP motif under control of PdnaK	this study
pK19mobSacB-RPPP-egfp	<i>egfp</i> varinat contiang a RPPP motif under control of PdnaK	this study
pK19mobSacB-cgl117-egfp	Translational fusion of native <i>cgl117</i> and a c-terminal 6x histag	this study
pK19mobSacB-mNG-PtsG	Translational fusion of native <i>ptsG</i> and a n-terminal mNeonGreen	Martins et al. 2019
pK19mobSacB-gntR2-mCherry	Translational fusion of native <i>gntR2</i> and a c-terminal mCherry	this study
pK19mobSacB-ptsG-del	Contains up- and downstream regions of <i>ptsG</i> . For <i>ptsG</i> deletion	this study
pK19mobSacB-mNG-PtsG-PP/AA	Translational fusion of native <i>ptsG</i> and a n-terminal mNeonGreen and codon substitutions at positions 235/236 PP/AA	this study
pK19mobSacB-mNG-PtsG-P/A	Translational fusion of native <i>ptsG</i> and a n-terminal mNeonGreen and codon substitutions at position 236 P/A	this study
pTE1213	Translational fusion of <i>S. coelicolor</i> native <i>efp</i> and a c-terminal 6x histag	this study
pTE1208	Protospacer targeting Apra-cassette	this study
<b>Recombinant expression of EF-P in E. coli</b>		
pBAD33-Cg-efp-histag	<i>efp</i> from <i>C. glutamicum</i> fused with histag	this study
pBAD33-Scoe-efp-histag	<i>efp</i> from <i>S. coelicolor</i> fused with histag	this study
pBAD33-Ms-efp-histag	<i>efp</i> from <i>M. smegmatis</i> fused with histag	this study
pBAD33-Mp-efp-histag	<i>efp</i> from <i>M. phlei</i> fused with histag	this study
pBAD33-Sv-efp-histag	<i>efp</i> from <i>S. venezuelae</i> fused with histag	this study
pBAD33-Sc-efp-histag	<i>efp</i> from <i>S. californicus</i> fused with histag	this study
pBAD33-Mt-efp-histag	<i>efp</i> from <i>M. tuberculosis</i> fused with histag	this study
pBAD33-Ec-efp-histag	<i>efp</i> <i>E. coli</i> fused with histag	this study
pBAD33-Ec-efp-K34A	<i>efp</i> <i>E. coli</i> K34A fused with histag	this study
pBAD33-Ec-efp-K34R	<i>efp</i> <i>E. coli</i> K34R fused with histag	this study
<b>Expression of EF-P in C. glutamicum</b>		
pEKEx2-Cg-efp-histag	<i>efp</i> native copy fused with histag	this study
pEKEx2-Cg-efp	<i>efp</i> native copy including the native promotor	this study
pEKEx2-Cg-efp-K32A	<i>efp</i> K32A including the native promotor	this study
pEKEx2-Cg-efp-K32R	<i>efp</i> K32R including the native promotor	this study
pEKEx2-Cg-efp-K32M	<i>efp</i> K32M including the native promotor	this study
pEKEx2-Cg-efp-K32E	<i>efp</i> K32E including the native promotor	this study
pEKEx2-Cg-efp-K32Q	<i>efp</i> K32Q including the native promotor	this study
pEKEx2-Cg-efp-P30G	<i>efp</i> P30G including the native promotor	this study
pEKEx2-Cg-efp-P30A	<i>efp</i> P30A including the native promotor	this study
pEKEx2-Cg-efp-P34G	<i>efp</i> P34G including the native promotor	this study
pEKEx2-Cg-efp-P34A	<i>efp</i> P34A including the native promotor	this study
pEKEx2-Cg-efp-P34Q	<i>efp</i> P34Q including the native promotor	this study
pEKEx2-Cg-efp-P34N	<i>efp</i> P34N including the native promotor	this study
pEKEx2-Ms-efp	<i>efp</i> from <i>M. smegmatis</i> under control of <i>C. glutamicum</i> <i>efp</i> native promotor	this study
pEKEx2-Mp-efp	<i>efp</i> from <i>M. phlei</i> under control of <i>C. glutamicum</i> <i>efp</i> native promotor	this study
pEKEx2-Mt-efp	<i>efp</i> from <i>M. tuberculosis</i> under control of <i>C. glutamicum</i> <i>efp</i> native promotor	this study
pEKEx2-Scoe-efp	<i>efp</i> from <i>S. coelicolor</i> under control of <i>C. glutamicum</i> <i>efp</i> native promotor	this study
pEKEx2-Scal-efp	<i>efp</i> from <i>S. californicus</i> under control of <i>C. glutamicum</i> <i>efp</i> native promotor	this study
pEKEx2-Sv-efp	<i>efp</i> from <i>S. venezuelae</i> under control of <i>C. glutamicum</i> <i>efp</i> native promotor	this study
pEKEx2-Cg-efp-natRBS	<i>efp</i> native copy fused with 6xHis tag under control of Ptac promotor. Native RBS	this study
pEKEx2-Cg-efp-synRBS	<i>efp</i> native copy fused with 6xHis tag under control of Ptac promotor. RBS sequence AGG AGG TTT GGA ATG GTG	this study
<b>Expression of EF-P in M. smegmatis</b>		
pMycoFos-Ms-efp	<i>efp</i> from <i>M. smegmatis</i> under control of <i>M. smegmatis</i> <i>efp</i> native promotor	this study
<b>Expression of EIIGlc in C. glutamicum</b>		
pEKEx2-mNG-ptsG	Translational fusion of native <i>ptsG</i> and a n-terminal mNeonGreen	this study
pEKEx2-mNG-ptsG-P235X	Translational fusion of native <i>ptsG</i> and a n-terminal mNeonGreen and codon substitutions at positions 235 P/X	this study
pEKEx2-mNG-ptsG-P236X	Translational fusion of native <i>ptsG</i> and a n-terminal mNeonGreen and codon substitutions at positions 236 P/X	this study

## Materials and Methods

**Table 4.3** – Primers used in this thesis

Sequencing	Sequence	Restriction site
pBAD33-Fw	GGC GTC ACA CTT TGC TAT GC	
pBAD33-Rv	CAG TTC CCT ACT CTC GCA TG	
pEKEx2-Fw	TAT TCT GAA ATG AGC TGT TGA CAA TTA ATC	
pEKEx2-Rv	CAA ATT CTG TTT TAT CAG ACC GCT TCT GCG	
pK19mobSacB-Fw	AGT GAG CGC AAC GCA ATT AAT GTG A	
pK19mobSacB-Rv	CGC CAT TCG CCA TTC AGG CTG CGC A	
pMycFos-Fw	GAT CGG TTT TGC CGC AGT TGT TCT C	
PMycFos-Rv	TGA ATA TGG CTC ATA ACA CCC CTT G	
Gene deletions	Sequence	Restriction site
Cg-efpdel-1-Fw	ACG CCA AGC TTG TGC CAG AAC TGC TTC TGG GTT GGC GGC G	HindIII
Cg-efpdel-2-Rv	AAG GTG AGT AAG GAC CTC ATC GTG GTG TAA GAT TCT TAA AAC CTT TAA GAA TCA GCC AGA	
Cg-efpdel-3-Fw	TCT GGC TGA TTC TTA AAG GTT TTA AGA ATC TTA CAC CAC GAT GAG GTC CTT ACT CAC CTT	
Cg-efpdel-4-Rv	AGT GAA TTC GGC GCC TGT TGG GAG CAG AAC GCC CCA	EcoRI
Cg-ptsGdel-1-Fw	ATT AAG CTT TTT TGG CGG GCG CTT	HindIII
Cg-ptsGdel-2-Rv	AAA CCC GGG GTC AAA CCT TTC TAA ACG TAG GGT C	XmaI
Cg-ptsGdel-3-Fw	AAA CCC GGG CCT GGG ATC CAT GTT GCG	XmaI
Cg-ptsGdel-4-Rv	AAA GAA TTC CGG GCA CTT GAG AAG CGA TT	EcoRI
Insertion and Fusion	Sequence	Restriction site
Cg-efpHIS6-1-Fw	ACG CCA AGC TTG TGC CAG AAC TGC TTC TGG GTT GGC GGC G	HindIII
Cg-efpHIS6-2-Rv	TCC TAC CTC TCC CGC GTT AAC AAC AGA TCT CAT CAC CAT CAC TAA GAT TCT TAA AAC CTT TAA GAA TCA GCC	
Cg-efpHIS6-3-Fw	GGC TGA TTC TTA AAG GTT TTA AGA ATC TTA GTG ATG GTG ATG GTG ATG AGA TCT GTT GTT AAC GCG GGA GAT GTA GGA	
Cg-efpHIS6-4-Rv	CAG TGA ATT CGT GAT TTC AAG AAC GGT CTA GTT CTC AAG AA	EcoRI
Cg-PdnaK-egfp-1-WT	GTC GAC TCT AGA AGG AGC ACC GGA AAT CAT CAC CAA G	XbaI
Cg-PdnaK-egfp-2-WT	GCC ATA AAC TTT AGA AGG ACA CCC TTT AGT TGT TAA CGC GGG AGA GGT AG	
Cg-PdnaK-egfp-3-WT	CTA CCT CTC CCG CGT TAA CAA CTA AAG GGT GTC CTT CTA AAG TTT ATG GC	
Cg-PdnaK-egfp-1-deltaEFP	GTC GAC TCT AGA AAT CCA TTC GCC TGA CCA AAG ACA GCT TC	XbaI
Cg-PdnaK-egfp-2-deltaEFP	GCC ATA AAC TTT AGA AGG ACA CCC TTT ACA CCA CGA TGA GGT CCT TAC TC	
Cg-PdnaK-egfp-3-deltaEFP	GAG TAA GGA CCT CAT CGT GGT GTA AAG GGT GTC CTT CTA AAG TTT ATG GC	
Cg-PdnaK-egfp-4-NoMotif	CAG TGA AAA GTT CTT CTC CTT TAC GCA TAA TGG TGC CTC TTT TAG GTG AT	
Cg-PdnaK-egfp-5-NoMotif	ATC ACC TAA AAC AGG AGG CAC CAT TAT GCG TAA AGG AGA ACT TTT CAC TG	
Cg-PdnaK-egfp-4-RPAP	CAG TGA AAA GTT CTT CTC CTT TAC GCG GCG CCG GGC GCA TAA TGG TGC CTC CTG TTT TAG GTG AT	
Cg-PdnaK-egfp-5-RPAP	ATC ACC TAA AAC AGG AGG CAC CAT TAT GCG CCC GGC GCC GCG TAA AGG AGA AGA ACT TTT CAC TG	
Cg-PdnaK-egfp-4-R3P	CAG TGA AAA GTT CTT CTC CTT TAC GCG GCG GCG GGC GCA TAA TGG TGC CTC CTG TTT TAG GTG AT	
Cg-PdnaK-egfp-5-R3P	ATC ACC TAA AAC AGG AGG CAC CAT TAT GCG CCC GCC GCC GCG TAA AGG AGA AGA ACT TTT CAC TG	
Cg-PdnaK-egfp-6	CTG ATT CTT AAA GGT TTT AAG AAT CTT AAT GAT GAT GAT GAT GAG CC	
Cg-PdnaK-egfp-7	GGC TCA TCA TCA TCA TCA TTA AGA TTC TTA AAA CCT TTA AGA ATC AG	
Cg-PdnaK-egfp-8	GCC AGT GAA TTC TCT TCT GCG ACT TCG GAA ACC TCT A	EcoRI
Cg-cgl1117-egfp-1-Fw	TAG AGG ATC CAA TTG CAT CCG CGG AAC CAG TG	
Cg-cgl1117-egfp-2-Rv	TTC TTC TCC TTT ACG CGG CGG GCG CAT CAT CAA GGA TTT GTA CAC ATC AAT GGT CTG	BamHI
Cg-cgl1117-egfp-3-Fw	CAG ACC ATT GAT GTG TAC AAA TCC TTG ATG ATG CGC CCG CCG CGT AAA GGA GAA GAA	
Cg-cgl1117-egfp-4-Rv	AGC CGG GGA GAG GTT CCC CGG CTT TCG GTT TTA ATG ATG ATG ATG ATG AGC CAT ACC	
Cg-cgl1117-egfp-5-Fw	GGT ATG GCT CAT CAT CAT CAT CAT TAA AAC CGA AAG CCG GGG AAC CTC TCC CCG GCT	
Cg-cgl1117-egfp-6-Rv	CAG TGA ATT CGA CTC TAA TCC AGA GAG GCC GTT GA	EcoRI
Cg-GntR2mCherry-1-Fw	CAG GTC GAC GGG CAT CAC TGT TGA GGA AGG TAA T	Sall
Cg-GntR2mCherry-2-Rv	TAT TTT CTT CTC CTT TAC TAA CCA TTG AAC CTC CAC CTC CTG AAC CTC CAC CTC CTG AAC CTC CAC CTC CCG AGG TCA GTG CGT CGC GCA CCT CA	
Cg-GntR2mCherry-3-Fw	TGA GGT GCG CGA CGC ACT GAC CTC GGG AGG TGG AGG TTC AGG AGG TGG AGG TTC AGG AGG TGG AGG TTC AAT GGT TAG TAA AGG AGA AGA AAA TA	
Cg-GntR2mCherry-4-Rv	TGT GTA GTG ACT CGT TTA GTG GCA ATT AGT GAT GGT GAT GAT GAT	
Cg-GntR2mCherry-5-Fw	ATC TCA TCA CCA TCA CCA TCA CTA ATT GCC ACT AAA CGA GTC ACT ACA CA	
Cg-GntR2mCherry-6-Rv	CGG GGA TCC CAT TAG AAC TGA GGC TAG CGC TGT G	BamHI
Sc-efphis-prot-Fw	ACG CGG CCG TAT TGC GAG CAG CCA	
Sc-efphis-prot-Rv	AAA CTG GCT GCT CGC AAT ACG GCC	
Sc-Apra-prot-Fw	ACG CCC CAT CGA GTT CAT GGA CAC	
Sc-Apra-prot-Rv	AAA CGT GTC CAT GAA CTC GAT GGG	
Sc-pCM4.4-HA-site-UNS7-Fw	CAA GAC GCT GGC TCT GAC ATT TCC GCT ACT GAA CTA CTC GAC GCT CAG TGG AAC GAA AAC	
Sc-dLancI-II-1-UNS2-Rv	GCT TGG ATT CTG CGT TTG TTT CCG TCT ACG AAC TCC CAG CGG GAC GTG CTT GGC AAT CA	
Sc-Cas9-UNS2-Fw	GCT GGG AGT TCG TAG ACG GAA ACA AAC GCA GAA TCC AAG CCA TGC GCT CCA TCA AGA A	
Sc-pCM4.4-HA-site-UNS6-Rv	GTA TGT GAC CGT AGA GTA TTC TTA GGT GGC AGC GAA CGA GCA GAC CCC GTA GAA AAG A	
Sc-efphis-UNS6-Fw	CTC GTT CGC TGC CAC CTA AGA ATA CTC TAC GGT CAC ATA CAT ACG ACA AGA AGA CCC A	
Sc-efphis-Mut-Rv	AGT TAG TGA TGG TGA TGG TGA TGA GAT CTG CTG TTC ACC CGG CCG A	
Sc-efphis-Mut-Fw	AGA TCT CAT CAC CAT CAC CAT CAC TAA CTG TGG CTG CTC GCA ATA CG	
Sc-efphis-UNS7-Rv	CGA GTA GTT CAG TAG CGG AAA TGT CAG AGC CAG CGT CTT GGG AAG AGC GGG ATA ATT C	

## Materials and Methods

Sc-dLancl-II-1-bis-UNS1-Fw	CAT TAC TCG CAT CCA TTC TCA GGC TGT CTC GTC TCG TCT CGG CAG ACC CTC GGC TTT C	
Sc-dgly-GA-UNS2-Rv	GCT TGG ATT CTG CGT TTG TTT CCG TCT ACG AAC TCC CAG CCT AGG TTT CTG CAC TTT CTG	
Sc-dgly-GA-UNS2-Fw	GCT GGG AGT TCG TAG ACG GAA ACA AAC GCA GAA TCC AAG CGT GCA GGT CGA CGG ATC TT	
Sc-pCMU-4 UNS7-Rv	CGA GTA GTT CAG TAG CGG AAA TGT CAG AGC CAG CGT CTT GTG AGC GTC AGA CCC CGT A	
Sc-dLancl-II-4-UNS4-Rv	GAC TTT GCG TGT TGT CTT ACT ATT GCT GGC AGG AGG TCA GAC TTT TCG GGG AAA TGT G	
Sc-dLancl-II-5-bis-UNS4-Fw	CTG ACC TCC TGC CAG CAA TAG TAA GAC AAC ACG CAA AGT CAA CTC TAC ACA TCG AAT TCC T	
Sc-dLancl-II-5-UNS1-Rv	GAG ACG AGA CGA GAC AGC CTG AGA ATG GAT GCG AGT AAT GTA CGC CAA GCT TTC AGA AG	
<b>In vitro translation of EII<sup>91c</sup></b>	<b>Sequence</b>	<b>Restriction site</b>
T7-RBS-FlagTag-PtsG-Fw	GCG AAT TAA TAC GAC TCA CTA TAG GGC TTA AGT ATA AGG AGG AAA AAA TAT GGA CTA CAA GGA CGA CGA TGA CAA GGC GTC CAA ACT GAC GAC GAC ATC GC	
FlagTag-PtsG-Rv	AAA CCC CTC CGT TTA GAG AGG GGT TAT GCT AGT TAC TTG TCA TCG TCG TCC TTG TAG TCC TCG TTC TTG CCG TTG ACC TTG ATC	
<b>EF-P express. in <i>C. glutamicum</i></b>	<b>Sequence</b>	<b>Restriction site</b>
Cg-Pefp-efp-Fw	TAG AGG ATC CTC TAG GTG AGC TAA TCG GTC TGC GCG TTT T	BamHI
Cg-efp-HIS6-Rv	CAG TGA ATT CTT AGT GAT GGT GAT GGT GAT GAG ATC TTT AGT TGT TAA CGC GGG AGA GGT AGG AA	EcoRI
Cg-efp-K32A-Fw	AGCA CGT CAA GCC AGG CGC GGG CCC AGC ATT CGT GCG AAC C	
Cg-efp-K32A-Rv	GGT TCG CAC GAA TGC TGG GCC CGC GCC TGG CTT GAC GTG CT	
Cg-efp-K32R-Fw	GCA CGT CAA GCC AGG CCG CGG CCC AGC ATT CGT GCG AAC C	
Cg-efp-K32R-Rv	GGT TCG CAC GAA TGC TGG GCC CGC GCC TGG CTT GAC GTG C	
Cg-efp-K32M-Fw	CGA GTT CCA GCA CGT CAA GCC AGG CAT GGG CCC AGC ATT CGT GCG AAC CAA AC	
Cg-efp-K32M-Rv	GTT TGG TTC GCA CGA ATG CTG GGC CCA TGC CTG GCT TGA CGT GCT GGA ACT CG	
Cg-efp-K32E-Fw	CGA GTT CCA GCA CGT CAA GCC AGG CGA AGG CCC AGC ATT CGT GCG AAC CAA AC	
Cg-efp-K32E-Rv	GTT TGG TTC GCA CGA ATG CTG GGC CTT CGC CTG GCT TGA CGT GCT GGA ACT CG	
Cg-efp-K32Q-Fw	CGA GTT CCA GCA CGT CAA GCC AGG CCA GGG CCC AGC ATT CGT GCG AAC CAA AC	
Cg-efp-K32Q-Rv	GTT TGG TTC GCA CGA ATG CTG GGC CCT GGC CTG GCT TGA CGT GCT GGA ACT CG	
Cg-efp-P30Q-Fw	GAT CAT CGA GTT CCA GCA CGT CAA GCA GGG CAA GGG CCC AGC ATT CGT GCG AA	
Cg-efp-P30Q-Rv	TTC GCA CGA ATG CTG GGC CCT TGC CCT GCT TGA CGT GCT GGA ACT CGA TGA TC	
Cg-efp-P30A-Fw	GAT CAT CGA GTT CCA GCA CGT CAA GGC AGG CAA GGG CCC AGC ATT CGT GCG AA	
Cg-efp-P30A-Rv	TTC GCA CGA ATG CTG GGC CCT TGC CCT GCT TGA CGT GCT GGA ACT CGA TGA TC	
Cg-efp-P34Q-Fw	CCA GCA CGT CAA GCC AGG CAA GGG CCA GGC ATT CGT GCG AAC CAA ACT CAA GG	
Cg-efp-P34Q-Rv	CCT TGA GTT TGG TTC GCA CGA ATG CCT GGC CCT TGC CTG GCT TGA CGT GCT GG	
Cg-efp-P34A-Fw	CCA GCA CGT CAA GCC AGG CAA GGG CGC AGC ATT CGT GCG AAC CAA ACT CAA GG	
Cg-efp-P34A-Rv	CCT TGA GTT TGG TTC GCA CGA ATG CTG CGC CCT TGC CTG GCT TGA CGT GCT GG	
Cg-efp-P34G-Fw	CCA GCA CGT CAA GCC AGG CAA GGG CGC AGC ATT CGT GCG AAC CAA ACT CAA GG	
Cg-efp-P34G-Rv	CCT TGA GTT TGG TTC GCA CGA ATG CGC CGC CCT TGC CTG GCT TGA CGT GCT GG	
Cg-efp-P34N-Fw	CCA GCA CGT CAA GCC AGG CAA GGG CAA CGC ATT CGT GCG AAC CAA ACT CAA GG	
Cg-efp-P34N-Rv	CCT TGA GTT TGG TTC GCA CGA ATG CGT TGC CCT TGC CTG GCT TGA CGT GCT GG	
Cg-Pefp- <i>Ms-efp-Fw</i>	TAG AGG ATC CTC TAG GTG AGC TAA TCG GTC TGC GCG TTT TTC AAG CAC TTT GCG CAG ACC CCC ATC CAC GTA ATC CAC GAG GGA GAT CAC ATC CGT GGC AAC GAC CGC CGA CTT CAA GA	BamHI
<i>Ms-efp-HIS6-Rv</i>	TGC CGA CTT CAA GA CAG TGA ATT CTT AGT GAT GGT GAT GGT GAT GAG ATC TGG CAT TCA CCC GGC CCA GGT AGC TG	EcoRI
Cg-Pefp- <i>Mp-efp-Fw</i>	TAG AGG ATC CTC TAG GTG AGC TAA TCG GTC TGC GCG TTT TTC AAG CAC TTT GCG CAG ACC CCC ATC CAC GTA ATC CAC GAG GGA GAT CAC ATC CGT GGC AAC GAC CGC CGA CTT CAA GA	BamHI
<i>Mp-efp-HIS6-Rv</i>	CAG TGA ATT CTT AGT GAT GGT GAT GGT GAT GAG ATC TGG CAT TCA CCC GGC CCA GGT AGC TG	EcoRI
Cg-Pefp- <i>Scoe-efp-Fw</i>	TAG AGG ATC CTC TAG GTG AGC TAA TCG GTC TGC GCG TTT TTC AAG CAC TTT GCG CAG ACC CCC ATC CAC GTA ATC CAC GAG GGA GAT CAC ATC CGT GGC TTC CAC GAA CGA CCT CAA GA	BamHI
<i>Scoe-efp-HIS6-Rv</i>	CAG TGA ATT CTT AGT GAT GGT GAT GGT GAT GAG ATC TGC TGT TCA CCC GGC CGA GGT AGT CG	EcoRI
Cg-Pefp- <i>Scal-efp-Fw</i>	TAG AGG ATC CTC TAG GTG AGC TAA TCG GTC TGC GCG TTT TTC AAG CAC TTT GCG CAG ACC CCC ATC CAC GTA ATC CAC GAG GGA GAT CAC ATC CGA CCT CAA GAA CGG CAT GGT GCT CA	BamHI
<i>Scal-efp-HIS6-Rv</i>	CAG TGA ATT CTT AGT GAT GGT GAT GGT GAT GAG ATC TCT TCT CGC CGG TGG AGA TGA AGA GC	EcoRI
Cg-Pefp- <i>Sv-efp-Fw</i>	TAG AGG ATC CTC TAG GTG AGC TAA TCG GTC TGC GCG TTT TTC AAG CAC TTT GCG CAG ACC CCC ATC CAC GTA ATC CAC GAG GGA GAT CAC ATC CGT GGC TTC CAC GAA CGA CCT CAA GA	BamHI
<i>Sv-efp-HIS6-Rv</i>	CAG TGA ATT CTT AGT GAT GGT GAT GGT GAT GAG ATC TGC TGT TCA CCC GGC CGA GGT AGC TG	EcoRI
Cg-Pefp- <i>Mt-efp-Fw</i>	TAG AGG ATC CTC TAG GTG AGC TAA TCG GTC TGC GCG TTT TTC AAG CAC TTT GCG CAG ACC CCC ATC CAC GTA ATC CAC GAG GGA GAT CAC ATC CGT GGC GAC CAC TGC TGA CTT CAA GA	BamHI
<i>Mt-efp-HIS6-Rv</i>	CAG TGA ATT CTT AGT GAT GGT GAT GGT GAT GAG ATC TGG CGT TGA CCC GGC CCA GGT AGC TA	EcoRI
<b>EF-P expression in <i>E. coli</i></b>	<b>Sequence</b>	<b>Restriction site</b>
<i>Mt-efp-Fw</i>	TCC TCT AGA ATA TTA AAG AGG AGG ACA GCT ATG GCG ACC ACT GCT GAC TTC AAG AAC G	XbaI
<i>Mt-efp-Rv</i>	AGC CAA GCT TTT AGT GAT GGT GAT GGT GAT GAG ATC TGG CGT TGA CCC GGC CCA GGT AGC TA	HindIII
<i>Ms-efp-Fw</i>	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GGC ATC GAC TGC CGA CTT CAA GA	XbaI
<i>Ms-efp-Rv</i>	GCC AAG CTT TTA GTG ATG GTG ATG GTG ATG AGA TCT GGC ATT CAC CCG GCC CAG GTA G	HindIII
<i>Scoe-efp-Fw</i>	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TGT GGC TTC CAC GAA CGA CCT CAA GA	XbaI

## Materials and Methods

Scoe-efp-Rv	AGC CAA GCT TTT AGT GAT GGT GAT GGT GAT GAG ATC TGC TGT TCA CCC GGC CGA GGT AGT CG	HindIII
Mp-efp-Fw	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TGT GGC AAC GAG CAA CGA CAT CAA GA	XbaI
Mp-efp-Rv	GCC AAG CTT TCA GTG ATG GTG ATG GTG ATG AGA TCT GTC ATT GAC GCG GCC GAG GTA G	HindIII
Cg-efp-Fw	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GGC AAC TAC CGC TGA TTT CAA GAA CG	XbaI
Cg-efp-Rv	GCC AAG CTT TTA GTG ATG GTG ATG GTG ATG AGA TCT GTT GTT AAC GCG GGA GAG GTA GGA A	HindIII
Sv-efp-Fw	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GGC TTC CAC GAA CGA CCT CAA GA	XbaI
Sv-efp-Rv	AGC CAA GCT TTT AGT GAT GGT GAT GGT GAT GAG ATC TGC TGT TCA CCC GGC CGA GGT AGC TG	HindIII
Scal-efp-Fw	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GGA CCT CAA GAA CGG CAT GGT GC	XbaI
Scal-efp-Rv	GCC AAG CTT TTA GTG ATG GTG ATG GTG ATG AGA TCT CTT CTC GCC GGT GGA GAT GAA GAG C	HindIII
Ec-efp-Fw	ATC CTC TAG AAT ATT AAA GAG GAG GAC AGC TAT GGC AAC GTA CTA TAG CAA CGA TT	XbaI
Ec-efp-Rv	GCC AAG CTT TTA GTG ATG GTG ATG GTG ATG AGA TCT CTT CAC GCG AGA GAC GTA TTC A	HindIII
Ec-efp-K34A-Fw	AGC GAG TGA ATT CGT AAA ACC GGG TGC GGG CCA GGC ATT TGC TCG CGT TAA AC	
Ec-efp-K34A-Rv	GTT TAA CGC GAG CAA ATG CCT GGC CCG CAC CCG GTT TTA CGA ATT CAC TCG CT	
Ec-efp-K34R-Fw	AGC GAG TGA ATT CGT AAA ACC GGG TGC CGG CCA GGC ATT TGC TCG CGT TAA AC	
Ec-efp-K34R-Rv	GTT TAA CGC GAG CAA ATG CCT GGC CGC GAC CCG GTT TTA CGA ATT CAC TCG CT	
<b>EF-P express. in <i>M. smegmatis</i></b>	<b>Sequence</b>	<b>Restriction site</b>
pmycofos-ms-efp-Fw	GTA AGA ATT CGA CCG TGG AGC CTG GTG TGT ACC TG	EcoRI
pmycofos-ms-efp-Rv	ACA CGG ATC CTC AGT GAT GGT GAT GGT GAT GAG ATC TGG CAT TCA CCC GGC CCA GGT AGC TG	BamHI
<b>EII<sup>6c</sup> a. a. substitutions</b>	<b>Sequence</b>	<b>Restriction site</b>
PtsG-PP235-236AA-P1-Fw	GCA GGT CGA CTT TTG GCG GGC GCT TCG GCG AAA AT	Sall
PtsG-PP235-236AA-P2-Rv	GAA GTA GAA GAC CGA GCG CCA CAT GGA GTG CAG GAA TAC ATA AGT ATC AG	
PtsG-PP235-236AA-P3-Fw	CTG ATA CTT ATG TAT TCC TGC ACT CCA TGT GGC GCT CGG TCT TCT ACT TC	
PtsG-PP235-236AA-P4-Rv	AGT ACA GAC CAA TTG CTG CAA TCA GTG CTG CGA ATA CCT GTC CGG AGT AGT CAT TC	
PtsG-PP235-236AA-P5-Fw	GAA TGA CTA CTC CGG ACA GGT ATT CGC AGC ACT GAT TGC AGC AAT TGG TCT GTA CT	
PtsG-PP235-236AA-P6-Rv	TAG AGT CGA CTT ACT CGT TCT TGC CGT TGA CCT TG	XbaI
PtsG-P236A-P2-Rv	AAT GAC TAC TCC GGA CAG GTA TTC CCA GCA CTG ATT GCA GCA ATT GGT CTG TA	
PtsG-P236A-P3-Fw	TAC AGA CCA ATT GCT GCA ATC AGT GCT GGG AAT ACC TGT CCG GAG TAG TCA TT	
Cory-1stP-sub-Fw	GGT TCT GAA TGA CTA CTC CGG ACA GGT ATT CNN NCC GCT GAT TGC AGC AAT TGG TCT GTA CTG GG	
Cory-1stP-sub-Rv	CCC AGT ACA GAC CAA TTG CTG CAA TCA GCG GNN NGA ATA CCT GTC CGG AGT AGT CAT TCA GAA CC	
Cory-2nd-sub-Fw	GGT TCT GAA TGA CTA CTC CGG ACA GGT ATT CCC ANN NCT GAT TGC AGC AAT TGG TCT GTA CTG GG	
Cory-2nd-sub-Rv	CCC AGT ACA GAC CAA TTG CTG CAA TCA GNN NTG GGA ATA CCT GTC CGG AGT AGT CAT TCA GAA CC	
<b>EF-P overexpr. in <i>C. glutamicum</i></b>	<b>Sequence</b>	<b>Restriction site</b>
Cg-Pefp-efp-Fw	TAG AGG ATC CTC TAG GTG AGC TAA TCG GTC TGC GCG TTT T	BamHI
Cory-Cory-synRBS-efp-Fw	TAG AGG ATC CAG GAG GTT TGG AAT GGT GGC AAC TAC CGC TGA TTT CAA GAA CGG T	BamHI
Cory-nativeRBS-efp-Fw	TAG AGG ATC CAT CCA CGA GGG AGA TCA CAT CCA TGG TGG CAA CTA CCG CTG ATT TCA AGA ACG GT	BamHI
Cg-efp-HIS6-Rv	CAG TGA ATT CTT AGT GAT GGT GAT GGT GAT GAG ATC TTT AGT TGT TAA CGC GGG AGA GGT AGG AA	EcoRI
<b>RT-qPCR</b>	<b>Sequence</b>	<b>Restriction site</b>
Cory-RT-16S-Fw	ACC CTT GTC TTA TGT TGC CAG	
Cory-RT-16S-Rv	TGT ACC GAC CAT TGT AGC ATG	
Cory-RT-dnaE-2-Fw	TTC GTC ATC AGG AAC GGT GA	
Cory-RT-dnaE-2-Rv	GAT CGA TAC CCT GCC GTG TA	
Cory-RT-efp-Fw	CTC GAT GAA CAG TGG GAC CT	
Cory-RT-efp-Rv	GTT CAG GTC TCC TTC CAC GA	

### 4.4 Growth conditions

*E. coli* was grown in lysogeny broth (LB) under aerobic conditions at 37°C. *S. coelicolor* was cultivated aerobically in SFM medium, tryptic soy broth (TSB) medium or chemically defined medium supplemented with sodium glutamate (Nieselt et al., 2010) at 30°C. *M. smegmatis* was grown in BHI (brain-heat infusion) 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. BHI was used as complex media for *C. glutamicum* growth. Cells were cultivated in baffled flasks at 30°C in a rotatory shaker. Growth in chemical defined media CGXII supplemented with 2% of the specified carbon source were followed in 96 well plates incubated at 30°C and 220rpm. 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*).

### 4.5 Proteomic analysis

Cells ( $5 \times 10^8$ ) 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  $3 \times 10^6$ ). The 10 most intense peptide ions with charge states between 2 and 5 were sequentially isolated to a target value of  $1 \times 10^5$ , 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\_3AUP00000582\_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  $s_0=1$  (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.

#### 4.6 Protein cluster analysis

Downregulated proteins identified by MaxQuant were uploaded into the database for annotation, visualization and integrated discovery (DAVID) (Huang da 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.

#### 4.7 *E. coli* EF-P reporter strains

*E. coli*  $P_{cadBA}::lacZ\Delta cadAB\Delta efp$  or *E. coli*  $P_{cadBA}::lacZ\Delta cadAB\Delta efp$  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  $\beta$ -galactosidase activity was determined as previously described (Tetsch et al., 2008; Ude et al., 2013).

#### 4.8 *C. glutamicum* EF-P reporter strain

For quantification of EF-P activity in *C. glutamicum*, the strain ATCC13032  $\Delta efp P_{dnaK}$ -RPPP-*egfp* was transformed with pEKEx2 containing a copy of *efp* under the control of the native EF- $P_{Cg}$  promoter. A 50- $\mu$ L 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.

#### 4.9 Single-cell fluorescence microscopy and quantitative analysis

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. The exposure time was 500-ms unless described otherwise in figure legends. 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-tailed t-test. Quantification of Cgl1117 production was done by monitoring fluorescence of Cgl1117-eGFP. The

strains *C. glutamicum*  $\Delta efp$  and *C. glutamicum* *efp*-6His (which produces His-tagged EF-P) were grown in BHI medium to an OD<sub>600</sub> 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.

For quantification of mNG-EII<sup>Glc</sup> and GntR2-mCherry, cells were grown in rich BHI medium supplemented with 2% (w/v) glucose until OD<sub>600</sub> = 2. Cells were then collected, washed in ice-cold phosphate-buffered saline (PBS; pH 7.4), placed on an agarose pad (1% w/v agarose in PBS) and covered with a coverslip. Fluorescence images of cells producing mNG-EII<sup>Glc</sup> were taken on a Delta Vision Elite (GE Healthcare, Applied Precision) equipped with Insight SSI illumination, X4 laser module and a Cool Snap HQ2 CCD camera. Exposure times were limited to 2000 ms. Images of cells expressing GntR2-mCherry were taken on a Leica microscope DMI6000B equipped with a DFC365 Fx camera (Leica) and a 300-ms exposure time was used. Excitation and emission filters were selected as appropriate for the relevant fluorophore: 460/512 for mNeonGreen and 546/605 for mCherry. At least 300 cells per condition were analyzed. Digital images were analyzed using Fiji (Schindelin et al., 2012).

### 4.10 Western blot analysis

SDS-polyacrylamide gels were used to fractionate proteins. When necessary, the gels were stained with Instant blue. Wet-transfer method was used to transfer the proteins from SDS-polyacrylamide gels to nitrocellulose membranes (Amersham, GE Healthcare). Primary and secondary antibodies were diluted in TBS [10mM Tris/HCl pH 7.5, 150mM NaCl] supplemented with 3% (wt/vol) BSA as following dilutions: 1:1000 for monoclonal mouse anti-mNeonGreen antibodies (Chromotek, 32F6), 1:4,000 for polyclonal rabbit antibodies against *E. coli* EF-P (Eurogentec), 1:10,000 for monoclonal mouse anti-Flag antibodies (Sigma, A8592), 1:20,000 for rabbit anti-6xHis antibody. Secondary fluorescent antibodies dilutions were 1:20,000 for goat anti-mouse IgG antibodies (Abcam, ab216776) and 1:20,000 for goat anti-rabbit antibodies (Abcam, ab216773). Between primary and secondary antibodies membranes were washed in TBS-TT buffer [10mM Tris/HCl pH 7.5, 500mM NaCl, 0.05% (v/v) Tween 20, 0.2% (v/v) Triton 100]. Images were taken using the Odyssey CLx imaging system (LI-COR Biosciences).

### 4.11 Quantitative Western blot

*C. glutamicum* expressing endogenous copy of *efp*-6xHis and mNG-ptsG were grown and samples for collected after every hour of incubation. Cells were lysed as described above and normalized by protein concentration. Fractions were loaded on a SDS-polyacrylamide gel and transferred to a nitrocellulose membrane by Western blot. After imaging, the fold change in the protein level as calculated using Fiji. To determine the amount of EF-P per cell, the cytosolic fractions were loaded in a SDS-polyacrylamide gel and transferred to a nitrocellulose membrane as described among with a standard curve determined

by using purified *C. glutamicum* EF-P in serial increasing amounts (0.10, 0.20, 0.40, 0.80, 1.6, 3.2 and 6.4 $\mu$ g).

For *S. coelicolor* EF-P quantification, cell pellets were suspended in HEPES buffer to a final concentration of 10mg mg (DW)/mL. Bacterial clumps were disrupted by sonication (20% intensity, 3 cycles of 30 seconds 1s on - 1s off) and further lysed using high-pressure system (Constant Systems). Samples were loaded in a SDS-polyacrylamide gel along with a serial dilution (1:2) of purified EF-P<sub>Sc</sub> as standard and transferred to a nitrocellulose membrane (Amersham, GE Healthcare). Primary and secondary antibodies were diluted in TBS [10mM Tris/HCl pH 7.5, 150mM NaCl] supplemented with 3% (wt/vol) BSA as following dilutions: 1:1.000 rabbit anti EF-P<sub>Bs</sub> (Provided by Dr. Daniel Wilson) and 1:20.000 goat anti-rabbit (Abcam, ab216773). Fluorescence images were taken using Odyssey CLx imaging system (LI-COR Biosciences) and bands intensity quantified by Fiji software.

### 4.12 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<sub>600</sub> of 2 (exponential phase) or overnight (stationary 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<sub>*efp*</sub>-*efp*-6His transformants grown at 37°C for 24 or 72 h.

EF-P variants with amino acid replacements were produced in *C. glutamicum*  $\Delta$ *efp* cells grown in rich medium supplemented with the appropriate antibiotic overnight. Recombinant EF-P was expressed in *E. coli* MG1655  $\Delta$ *efp*  $\Delta$ *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<sup>2+</sup>-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.

The purification of modified EF-P from *E. coli*, used in the in vitro reaction was done by transforming the *E. coli* K12 MG1655 with pBAD33 encoding for EF-P-6xHisTag, *EpmA* and *EpmB* were grown until OD<sub>600</sub> 2, induced with 0.2% arabinose and incubated overnight at 18°C under constant aeration. Cell lysis and protein purification were done as described above.

Recombinant EF-P<sub>Sc</sub> were produced in *E. coli*  $\Delta$ *efp*  $\Delta$ *epmA* transformed with replicative plasmid pBAD33 EF-P<sub>Sc</sub>-6His. Cultivation was done in LB supplemented with chloramphenicol, aerobically at 37°C until

OD<sub>600</sub> of 0.5, when arabinose was added to a final concentration of 0.2% w/v and cultivation temperature changed to 18°C overnight. Purification was done by affinity chromatography using Ni-NTA (Qiagen). Protein was kept in HEPES buffer (25mM HEPES-KOH pH 8, 125mM NaCl, 25mM KCl 10% v/v glycerol) at 4°C until further use.

### 4.13 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.

### 4.14 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<sub>2</sub>O) for 45 min at 25°C, proteins were alkylated using 5.5 mM IAA (2 µl of 550 mM stock in ddH<sub>2</sub>O) 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<sub>2</sub>O) 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<sub>2</sub>, 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 x 1 ml) and equilibrated with 0.5% (v/v) FA (3 x 1 ml) prior to loading the samples. After washing with 0.5% (v/v) FA (3 x 1 ml), peptides were eluted with 80% (v/v) ACN containing 0.5% FA (2 x 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).

To detect unmodified EF-P peptides in cell lysates, *M. tuberculosis* H37Rv cultures were inactivated by incubating frozen 0.5-1g wet weight pellets 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 PBS for future proteomic analysis. Successful inactivation was validated by Mycobacteria Growth Indicator Tube (MGIT) analysis 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 300 $\mu$ L of the cell suspensions to a sterile 2mL screwcap centrifuge tubes, including 300 $\mu$ L of 0.1mm glass pearls and agitating at maximum speed for three rounds of 45 seconds 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*.

### 4.15 MS measurement and analysis

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  $\mu$ m, 100A, inner diameter 75  $\mu$ m, Thermo Scientific) with 0.1% (v/v) TFA and separated on a 50 cm PepMap RSLC C18 column (particles 2  $\mu$ m, 100A, inner diameter 75  $\mu$ m, 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).

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.

### 4.16 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.

### 4.17 Protein crystallisation and structure determination

Prior to crystallisation, EF-P was concentrated to 3 mg/ml using centrifugal filter devices (Amicon, Merck). Aggregates and debris were removed by centrifugation (16.000xg 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 homologue 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; Emsley et al., 2010). The structural figures were prepared using PyMOL 2.3 (Schrödinger, 2015). For data processing and structure refinement statistics, see Supplemental Table S1.

### 4.18 RT-qPCR analysis

The RNA used for reverse transcription qPCR was isolated using the phenol-chloroform-isoamyl alcohol (PCI) protocol (Russell and Sambrook, 2001) with modifications. 50 mg of bacterial pellet was washed in 1 mL of ice-cold AE buffer [20 mM sodium acetate brought to pH 5.2 with acetic acid, 1 mM EDTA pH 8.0] and resuspended in 500 µL of the same buffer. Then 500 µL of pre-warmed PCI for RNA extraction (Roth, X985) and 25 µL of 10% (w/v) SDS were added and the mixture was incubated for 5 min at 60°C under vigorous agitation. Samples were cooled for 2 h on ice and centrifuged for 1 h at 16,000g. The supernatant was transferred to phase-lock tubes (Quanta), and 1.0 volume of PCI and 0.1 volume of 3

M sodium acetate (pH 5.2) were added before centrifugation for 15 min. The supernatant was collected, mixed with 2.3 volumes of ethanol and placed in a -80°C freezer overnight. After centrifugation at 16,000g for 1 h, the supernatant was discarded and the pellet was washed twice with 70% (v/v) ethanol, dried and resuspended in 100 µL of RNase-free water. A 1-µg aliquot of the isolated total RNA was converted to cDNA with the iScript Advanced Script (Bio-Rad) according to the manufacturer's protocol. Samples were mixed with SsoAdvanced Univ SYBR Green Supermix (Bio-Rad), dispensed in triplicate onto a 96-well PCR plate (Bio-Rad) and subjected to qPCR in a Bio-Rad CFX real-time cycler. Evaluation of the data obtained was performed according to the  $\Delta\Delta C_t$  method, using 16S rRNA and *dnaE* amounts as internal references (Schmittgen and Livak, 2008).

### 4.19 *In vitro* translation of EII<sup>Glc</sup>

The PURExpress *In Vitro* Protein Synthesis Kit (NEB, E6800) was used for in-vitro translation of EII<sup>Glc</sup>, with the following modifications. To avoid protein aggregation, the reaction mix was supplemented with 1 mM arginine (pH 7.0) and 1 mM  $\beta$ -mercaptoethanol. Purified post-translationally modified *E. coli* EF-P was added at 1 µM concentration. The same amount of lysis buffer was added to the negative control. The tubes were incubated for 3 min at 37°C prior to the addition of DNA. To start the reaction, 200 ng of DNA coding for FT-EII<sup>Glc</sup>-FT was added. Samples were taken at 0, 15, 30 and 45 min, and the translation reaction was stopped by homogenizing the sample in a stop-solution containing 0.1 M kanamycin, 8 M urea, 400 mM arginine (pH 7.0) and 10 mM  $\beta$ -mercaptoethanol. Samples (2.5 µL) were then loaded onto a SDS-polyacrylamide gel (Laemmli, 1970) to fractionate the proteins by size, and further analyzed after Western blotting.

### 4.20 D-glucose-6-<sup>14</sup>C uptake assay

All uptake measurements were performed as described earlier with minor modifications (Lindner et al., 2011). Cells were grown in BHI medium supplemented with 2% (w/v) glucose to OD<sub>600</sub> 2, harvested by centrifugation, then washed three times in ice-cold CGXII medium (without carbon sources), resuspended to OD<sub>600</sub> = 2 and stored on ice until analyzed. Prior to measurement, 2-mL aliquots of cell culture were incubated for 3 min at 30°C in a water bath. The reaction was started by the addition of 5, 50 or 500 µM D-glucose-6-<sup>14</sup>C (59 mCi/nmol; Sigma, G9899). At 30s intervals, 200-µL samples were filtered through glass-fiber filters (Typ F, Millipore) and washed twice with 2.5 mL of 100 mM LiCl. The radioactivity in the samples was determined using scintillation fluid (MP Biomedicals) and a scintillation counter (PerkinElmer).

### 4.21 Protein overproduction in *C. glutamicum*

Cells transformed with inducible vectors were grown in BHI medium until OD<sub>600</sub> 2 and induced with 10mM IPTG (Isopropyl β-D-1-thiogalactopyranoside). Culture were kept at 30°C under continuous aeration for more 2 hours. Cells were suspended to an OD<sub>600</sub> 10 in lysis buffer [25 mM HEPES-KOH pH 8, 125 mM NaCl, 25 mM KCl] and disrupted by high-pressure system (Constant Systems). To confirm overproduction, cytosolic protein solution was obtained by ultracentrifugation and fractionated by size on a SDS-polyacrylamide gel. Proteins were transferred to a nitrocellulose membrane and detected via Western blot.

### 4.22 EII<sup>Glc</sup> tertiary structure prediction

The amino acid sequence of the EII domain of EII<sup>Glc</sup> was uploaded to the Phyre2 platform using standard parameters (Kelley et al., 2015). The structure was modelled with 100% confidence with a single highest scoring template, 5IWS (*Bacillus cereus* MalT). Output PDB files were visualized using UCSF Chimera v1.14 (Pettersen et al., 2004).

### 4.23 Statistical analysis

Two-tailed t tests were performed using GraphPad Prism version 8.3.1 for Windows. Differences are considered significant when *p-values* 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 da et al., 2009).

### 4.24 Crystallographic and proteomic data availability

Crystallographic data of EF-P structures have been deposited in the Protein Data Bank under the PDB accession code 6S8Z. The mass spectrometry proteomics data are available via ProteomeXchange with identifier PXD014742.

## Acknowledgements

In This session I would like to warmly acknowledge all the people who directly or indirectly, made a significant contribution to this work:

My special thanks go to Prof. Kirsten Jung for the uncomplicated admission to the group, the trust deposited in my work and the outstanding support and supervision throughout this doctoral thesis.

I would like to thank Prof. Jörg Nickelsen for been available as a second reviewer, as well as the other members this thesis examination committee Prof. Andreas Klingl, Prof. Christof Osman, Dr. Macarena Marín and Prof. Thomas Nägele.

Special acknowledgements to all the scientific collaborators for the time and resources invested in the experiments and manuscripts: Christopher Scheidler, Pavel Kielkowski, Marina Schmid, Ignasi Forné, Suhui Ye, Norbert Reiling, Eriko Takano, Axel Imhof, Stephan Sieber, Sabine Schneider, Dimitar Petrov, Gustavo Martins and Marc Bramkamp.

To all the members of the GRK2062 Molecular Principles of Synthetic Biology Graduate School, specially our coordinator Dr. Beate Hafner for the outstanding support.

Special thanks to the members of the 01.039 lab Ingrid Weigl and Korinna Burdack for always taking care of the equipment, reagents and the good music.

To the actual and former members of the coffee-popcorn-beer office: Elisabeth Hoyer, Julia Schwarz, Lingyun Guo, Dr. Magdalena Motz and Dr. Yang Wang. The time with you guys was incredible.

To the colleges and former colleges from all AGs that were always warm and very helpful, particularly: Tatiana Anikanova, Sabine Peschek, Hannah Schramke, Ivica Krištofičová, Florian Fabiani, Sophie Brameyer, Stephanie Göing, Miriam Pfab, Ralph Krafczyk, Wolfram Volkwein, Michelle Eder and Larissa Wirtz.

Nothing here would have been possible without the awesome people I met in Munich and that will always be in my heart. The ones that always have a word – and attitude – of support and understanding (unless the invitation for the “last beer” is refused). The ones that made enduring a PhD thesis a much easier task. The most special acknowledgements go to my “Munich family” Tobias Möller, Daniel Thiemicke, Nicola Lorenz and João Guilherme, I love you guys, thank you for the endless support and countless stories to tell. Marko Gogala, Claudia Vilhena and Ana Gasperotti thank you for so many good moments in and outside Munich. Dayane Araújo e Ana Antonello, for been always a good company, anywhere “myplace” is.

Muito obrigado à minha família pelo respeito e suporte incondicional em todos os aspectos da minha vida. E a todos os meus queridos em Belo Horizonte, que aqui seria impossível nomear, mas que fizeram-se muito presentes mesmo na com um oceano de distância.

Muito obrigado, também, às jardineiras Débora Barbosa, Fabiana Lima, Lorena Lana e Thiago Câmara. Se eu as conheço bem, abriram a tese de trás para frente só para constar se o nome delas estavam aqui.

Cheers, Prosit, Saúde!

Muito obrigado. Vielen Dank. Thank you. Merci beaucoup. Hvala vam. Muchas gracias. Takk fyrir. Grazie. 谢谢你.

## 5 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.
- Alam, M.T., Merlo, M.E., Consortium, S., Hodgson, D.A., Wellington, E.M., Takano, E., and Breitling, R. (2010). Metabolic modeling and analysis of the metabolic switch in *Streptomyces coelicolor*. *BMC Genomics* 11, 202.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). Molecular biology of the cell. In *Molecular biology of the cell*.
- 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.
- Amara, A., Takano, E., and Breitling, R. (2018). Development and validation of an updated computational model of *Streptomyces coelicolor* primary and secondary metabolism. *BMC Genomics* 19, 519.
- An, G., Glick, B.R., Friesen, J.D., and Ganoza, M.C. (1980). Identification and quantitation of elongation factor EF-P in *Escherichia coli* cell-free extracts. *Can J Biochem* 58, 1312-1314.
- Aoki, H., Dekany, K., Adams, S.L., and 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.
- Becker, J., Giesselmann, G., Hoffmann, S.L., and Wittmann, C. (2018). *Corynebacterium glutamicum* for sustainable bioproduction: from metabolic physiology to systems metabolic engineering. *Adv Biochem Eng Biotechnol* 162, 217-263.
- Bella, J. (2016). Collagen structure: new tricks from a very old dog. *Biochem J* 473, 1001-1025.
- Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., *et al.* (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417, 141-147.
- Bergey, D.H., Whitman, W.B., Goodfellow, M., Kämpfer, P., and Busse, H.-J. (2012). *Bergey's manual of systematic bacteriology*. Vol. 5, the Actinobacteria. (New York, Springer,) p. 1 online resource (2 vol.).
- Betts, J.C., Lukey, P.T., Robb, L.C., McAdam, R.A., and Duncan, K. (2002). Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol Microbiol* 43, 717-731.
- Bibb, M.J. (2005). Regulation of secondary metabolism in streptomycetes. *Curr Opin Microbiol* 8, 208-215.
- 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.
- Burmann, B.M., Schweimer, K., Luo, X., Wahl, M.C., Stitt, B.L., Gottesman, M.E., and Rosch, P. (2010). A NusE:NusG complex links transcription and translation. *Science* 328, 501-504.
- Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. *Genome Res* 14, 1188-1190.
- DeJesus, M.A., Gerrick, E.R., Xu, W., Park, S.W., Long, J.E., Boutte, C.C., Rubin, E.J., Schnappinger, D., Ehrt, S., Fortune, S.M., *et al.* (2017). Comprehensive essentiality analysis of the *Mycobacterium tuberculosis* genome via saturating transposon mutagenesis. *mBio* 8.
- Deribe, Y.L., Pawson, T., and Dikic, I. (2010). Post-translational modifications in signal integration. *Nat Struct Mol Biol* 17, 666-672.

## References

---

- Dever, T.E., Gutierrez, E., and Shin, B.S. (2014). The hypusine-containing translation factor eIF5A. *Crit Rev Biochem Mol Biol* 49, 413-425.
- 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.
- Doma, M.K., and Parker, R. (2007). RNA quality control in eukaryotes. *Cell* 131, 660-668.
- Eggeling, L., and Bott, M. (2015). A giant market and a powerful metabolism: L-lysine provided by *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol* 99, 3387-3394.
- 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.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66, 486-501.
- Fernandez-Moreno, M.A., Caballero, J.L., Hopwood, D.A., and Malpartida, F. (1991). The act cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* tRNA gene of *Streptomyces*. *Cell* 66, 769-780.
- Frunzke, J., Engels, V., Hasenbein, S., Gatgens, C., and Bott, M. (2008). Co-ordinated regulation of gluconate catabolism and glucose uptake in *Corynebacterium glutamicum* by two functionally equivalent transcriptional regulators, GntR1 and GntR2. *Mol Microbiol* 67, 305-322.
- Gabel, K., Schmitt, J., Schulz, S., Nather, D.J., and Soppa, J. (2013). A comprehensive analysis of the importance of translation initiation factors for *Haloferax volcanii* applying deletion and conditional depletion mutants. *PLoS One* 8, e77188.
- Gamper, H.B., Masuda, I., Frenkel-Morgenstern, M., and Hou, Y.M. (2015). Maintenance of protein synthesis reading frame by EF-P and m(1)G37-tRNA. *Nat Commun* 6, 7226.
- Ganoza, M.C., and Aoki, H. (2000). Peptide bond synthesis: function of the *efp* gene product. *Biol Chem* 381, 553-559.
- Gao, B., and Gupta, R.S. (2005). Conserved indels in protein sequences that are characteristic of the phylum Actinobacteria. *Int J Syst Evol Microbiol* 55, 2401-2412.
- Gao, B., and Gupta, R.S. (2012). Phylogenetic framework and molecular signatures for the main clades of the phylum Actinobacteria. *Microbiol Mol Biol Rev* 76, 66-112.
- Glick, B.R., Chladek, S., and Ganoza, M.C. (1979). Peptide bond formation stimulated by protein synthesis factor EF-P depends on the aminoacyl moiety of the acceptor. *Eur J Biochem* 97, 23-28.
- Glick, B.R., and Ganoza, M.C. (1975). Identification of a soluble protein that stimulates peptide bond synthesis. *Proc Natl Acad Sci U S A* 72, 4257-4260.
- Greganova, E., Altmann, M., and Butikofer, P. (2011). Unique modifications of translation elongation factors. *FEBS J* 278, 2613-2624.
- Hanawa-Suetsugu, K., Sekine, S., Sakai, H., Hori-Takemoto, C., Terada, T., Unzai, S., Tame, J.R., Kuramitsu, S., Shirouzu, M., and Yokoyama, S. (2004). Crystal structure of elongation factor P from *Thermus thermophilus* HB8. *Proc Natl Acad Sci U S A* 101, 9595-9600.
- Heider, S.A., and Wendisch, V.F. (2015). Engineering microbial cell factories: Metabolic engineering of *Corynebacterium glutamicum* with a focus on non-natural products. *Biotechnol J* 10, 1170-1184.
- Henke, N.A., Wiebe, D., Perez-Garcia, F., Peters-Wendisch, P., and Wendisch, V.F. (2018). Coproduction of cell-bound and secreted value-added compounds: Simultaneous production of carotenoids and amino acids by *Corynebacterium glutamicum*. *Bioresour Technol* 247, 744-752.
- Henrich, A., Kuhlmann, N., Eck, A.W., Kramer, R., and Seibold, G.M. (2013). Maltose uptake by the novel ABC transport system MusEFGK2I causes increased expression of *ptsG* in *Corynebacterium glutamicum*. *J Bacteriol* 195, 2573-2584.
- 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.
- Hoskisson, P.A., and Fernandez-Martinez, L.T. (2018). Regulation of specialised metabolites in Actinobacteria - expanding the paradigms. *Environ Microbiol Rep* 10, 231-238.

## References

---

- Huang da, 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., and Kearns, D.B. (2020). Translation Elongation Factor P (EF-P). *FEMS Microbiol Rev.*
- 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 e516.
- Ikeda, M. (2012). Sugar transport systems in *Corynebacterium glutamicum*: features and applications to strain development. *Appl Microbiol Biotechnol* 96, 1191-1200.
- Jojima, T., Noburyu, R., Sasaki, M., Tajima, T., Suda, M., Yukawa, H., and Inui, M. (2015). Metabolic engineering for improved production of ethanol by *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol* 99, 1165-1172.
- 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.
- Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B.J., Gaigalat, L., *et al.* (2003). The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J Biotechnol* 104, 5-25.
- Kallscheuer, N., Classen, T., Drepper, T., and Marienhagen, J. (2019). Production of plant metabolites with applications in the food industry using engineered microorganisms. *Curr Opin Biotechnol* 56, 7-17.
- Kallscheuer, N., Vogt, M., Bott, M., and Marienhagen, J. (2017). Functional expression of plant-derived O-methyltransferase, flavanone 3-hydroxylase, and flavonol synthase in *Corynebacterium glutamicum* for production of pterostilbene, kaempferol, and quercetin. *J Biotechnol* 258, 190-196.
- Kallscheuer, N., Vogt, M., Stenzel, A., Gatgens, J., Bott, M., and Marienhagen, J. (2016). Construction of a *Corynebacterium glutamicum* platform strain for the production of stilbenes and (2S)-flavanones. *Metab Eng* 38, 47-55.
- 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. (2017). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.*
- Keilberg, D., Wuichet, K., Drescher, F., and Sogaard-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.
- Keiler, K.C. (2015). Mechanisms of ribosome rescue in bacteria. *Nat Rev Microbiol* 13, 285-297.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10, 845-858.
- Kieser T, B.M., Buttner MJ, Chater KF, Hopwood DA (2000). *Practical Streptomyces Genetics* (Norwich, UK: 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.
- Kogure, T., and Inui, M. (2018). Recent advances in metabolic engineering of *Corynebacterium glutamicum* for bioproduction of value-added aromatic chemicals and natural products. *Appl Microbiol Biotechnol* 102, 8685-8705.
- Kolter, R., Siegele, D.A., and Tormo, A. (1993). The stationary phase of the bacterial life cycle. *Annu Rev Microbiol* 47, 855-874.

## References

---

- Krause, F.S., Henrich, A., Blombach, B., Kramer, R., Eikmanns, B.J., and Seibold, G.M. (2010). Increased glucose utilization in *Corynebacterium glutamicum* by use of maltose, and its application for the improvement of L-valine productivity. *Appl Environ Microbiol* **76**, 370-374.
- Kuhlmann, N., Petrov, D.P., Henrich, A.W., Lindner, S.N., Wendisch, V.F., and Seibold, G.M. (2015). Transcription of *malP* is subject to phosphotransferase system-dependent regulation in *Corynebacterium glutamicum*. *Microbiology* **161**, 1830-1843.
- Kyrpides, N.C., and Woese, C.R. (1998). Universally conserved translation initiation factors. *Proc Natl Acad Sci U S A* **95**, 224-228.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- Lakemeyer, M., Bertosin, E., Moller, 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., Furst, M., Wuichet, K., Godeke, J., Starosta, A.L., Chen, J.M., Sogaard-Andersen, L., Rohr, J., Wilson, D.N., *et al.* (2015). Arginine-rhamnosylation as new strategy to activate translation elongation factor P. *Nat Chem Biol* **11**, 266-270.
- 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.
- Lata, M., Sharma, D., Deo, N., Tiwari, P.K., Bisht, D., and Venkatesan, K. (2015). Proteomic analysis of ofloxacin-mono resistant *Mycobacterium tuberculosis* isolates. *J Proteomics* **127**, 114-121.
- 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**, W242-245.
- Lewin, G.R., Carlos, C., Chevrette, M.G., Horn, H.A., McDonald, B.R., Stankey, R.J., Fox, B.G., and Currie, C.R. (2016). Evolution and ecology of Actinobacteria and their bioenergy applications. *Annu Rev Microbiol* **70**, 235-254.
- Li, G.W., Oh, E., and Weissman, J.S. (2012). The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* **484**, 538-541.
- Lindner, S.N., Petrov, D.P., Haggmann, C.T., Henrich, A., Kramer, R., Eikmanns, B.J., Wendisch, V.F., and Seibold, G.M. (2013). Phosphotransferase system-mediated glucose uptake is repressed in phosphoglucosyltransferase-deficient *Corynebacterium glutamicum* strains. *Appl Environ Microbiol* **79**, 2588-2595.
- Lindner, S.N., Seibold, G.M., Henrich, A., Kramer, R., and Wendisch, V.F. (2011). Phosphotransferase system-independent glucose utilization in *Corynebacterium glutamicum* by inositol permeases and glucokinases. *Appl Environ Microbiol* **77**, 3571-3581.
- Loveland, A.B., Bah, E., Madireddy, R., Zhang, Y., Brilot, A.F., Grigorieff, N., and Korostelev, A.A. (2016). Ribosome\*RelA structures reveal the mechanism of stringent response activation. *Elife* **5**.
- 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.
- Maki, K., Uno, K., Morita, T., and Aiba, H. (2008). RNA, but not protein partners, is directly responsible for translational silencing by a bacterial Hfq-binding small RNA. *Proc Natl Acad Sci U S A* **105**, 10332-10337.

## References

---

- 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.
- Manteca, A.Y., P. (2019). *Antimicrobials, Antibiotic Resistance, Antibiofilm Strategies and Activity Methods* (IntechOpen).
- Martins, G.B., Giacomelli, G., Goldbeck, O., Seibold, G.M., and Bramkamp, M. (2019). Substrate-dependent cluster density dynamics of *Corynebacterium glutamicum* phosphotransferase system permeases. *Mol Microbiol* 111, 1335-1354.
- 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 Crystallogr* 40, 658-674.
- McCoy, J.G., Ren, Z., Stanevich, V., Lee, J., Mitra, S., Levin, E.J., Poget, S., Quick, M., Im, W., and Zhou, M. (2016). The structure of a sugar transporter of the glucose EIIC superfamily provides insight into the elevator mechanism of membrane transport. *Structure* 24, 956-964.
- Meiswinkel, T.M., Gopinath, V., Lindner, S.N., Nampoothiri, K.M., and Wendisch, V.F. (2013). Accelerated pentose utilization by *Corynebacterium glutamicum* for accelerated production of lysine, glutamate, ornithine and putrescine. *Microb Biotechnol* 6, 131-140.
- Mitarai, N., Sneppen, K., and Pedersen, S. (2008). Ribosome collisions and translation efficiency: optimization by codon usage and mRNA destabilization. *J Mol Biol* 382, 236-245.
- Mohapatra, S., Choi, H., Ge, X., Sanyal, S., and Weisshaar, J.C. (2017). Spatial distribution and ribosome-binding dynamics of EF-P in live *Escherichia coli*. *MBio* 8.
- Moon, M.W., Park, S.Y., Choi, S.K., and Lee, J.K. (2007). The phosphotransferase system of *Corynebacterium glutamicum*: features of sugar transport and carbon regulation. *J Mol Microbiol Biotechnol* 12, 43-50.
- Morita, T., Maki, K., and Aiba, H. (2005). RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev* 19, 2176-2186.
- Motz, M., and Jung, K. (2018). The role of polyproline motifs in the histidine kinase EnvZ. *PLoS One* 13, e0199782.
- Naganathan, A., Wood, M.P., and Moore, S.D. (2015). The large ribosomal subunit protein L9 enables the growth of EF-P deficient cells and enhances small subunit maturation. *PLoS One* 10, e0120060.
- Navarre, W.W., Zou, S.B., Roy, H., Xie, J.L., Savchenko, A., Singer, A., Edvokimova, E., Prost, L.R., Kumar, R., Ibba, M., *et al.* (2010). PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. *Mol Cell* 39, 209-221.
- Nett, M., Ikeda, H., and Moore, B.S. (2009). Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Nat Prod Rep* 26, 1362-1384.
- Nguyen, T.X., Yen, M.R., Barabote, R.D., and Saier, M.H., Jr. (2006). Topological predictions for integral membrane permeases of the phosphoenolpyruvate:sugar phosphotransferase system. *J Mol Microbiol Biotechnol* 11, 345-360.
- Nieselt, K., Battke, F., Herbig, A., Bruheim, P., Wentzel, A., Jakobsen, O.M., Sletta, H., Alam, M.T., Merlo, M.E., Moore, J., *et al.* (2010). The dynamic architecture of the metabolic switch in *Streptomyces coelicolor*. *BMC Genomics* 11, 10.
- 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.
- Parche, S., Burkovski, A., Sprenger, G.A., Weil, B., Kramer, R., and Titgemeyer, F. (2001). *Corynebacterium glutamicum*: a dissection of the PTS. *J Mol Microbiol Biotechnol* 3, 423-428.
- 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.
- Peil, L., Starosta, A.L., Lassak, J., Atkinson, G.C., Virumae, K., Spitzer, M., Tenson, T., Jung, K., Remme, J., and Wilson, D.N. (2013). Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. *Proc Natl Acad Sci U S A* 110, 15265-15270.

## References

---

- Peil, L., Starosta, A.L., Virumae, K., Atkinson, G.C., Tenson, T., Remme, J., and Wilson, D.N. (2012). Lys34 of translation elongation factor EF-P is hydroxylated by YfcM. *Nat Chem Biol* 8, 695-697.
- 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, D442-D450.
- Peters-Wendisch, P., Gotker, S., Heider, S.A., Komati Reddy, G., Nguyen, A.Q., Stansen, K.C., and Wendisch, V.F. (2014). Engineering biotin prototrophic *Corynebacterium glutamicum* strains for amino acid, diamine and carotenoid production. *J Biotechnol* 192 Pt B, 346-354.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25, 1605-1612.
- Pfeifer, E., Gatgens, C., Polen, T., and Frunzke, J. (2017). Adaptive laboratory evolution of *Corynebacterium glutamicum* towards higher growth rates on glucose minimal medium. *Sci Rep* 7, 16780.
- Pinheiro, B., Scheidler, C.M., Kielkowski, P., Schmid, M., Forne, I., Ye, S., Reiling, N., Takano, E., Imhof, A., Sieber, S.A., *et al.* (2020). Structure and function of an elongation factor P subfamily in Actinobacteria. *Cell Rep* 30, 4332-4342 e4335.
- Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 5, e9490.
- Proshkin, S., Rahmouni, A.R., Mironov, A., and Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* 328, 504-508.
- 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-pentanoylated elongation factor P. *J Biol Chem* 291, 10976-10985.
- Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. *Cell* 108, 557-572.
- Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* 16, 276-277.
- Richter, D. (1976). Stringent factor from *Escherichia coli* directs ribosomal binding and release of uncharged tRNA. *Proc Natl Acad Sci U S A* 73, 707-711.
- Rodnina, M.V. (2018). Translation in prokaryotes. *Cold Spring Harb Perspect Biol* 10.
- Rossi, D., Kuroshu, R., Zanelli, C.F., and Valentini, S.R. (2014). eIF5A and EF-P: two unique translation factors are now traveling the same road. *Wiley Interdiscip Rev RNA* 5, 209-222.
- Roy, H., Zou, S.B., Bullwinkle, T.J., Wolfe, B.S., Gilreath, M.S., Forsyth, C.J., Navarre, W.W., and Ibba, M. (2011). The tRNA synthetase paralog PoxA modifies elongation factor-P with (R)-beta-lysine. *Nat Chem Biol* 7, 667-669.
- Russell, D.W., and Sambrook, J. (2001). *Molecular cloning: a laboratory manual*, Vol 1 (Cold Spring Harbor Laboratory Cold Spring Harbor, NY).
- Russell, J.B., and Cook, G.M. (1995). Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiol Rev* 59, 48-62.
- Sanchez, S., Rodriguez-Sanoja, R., Ramos, A., and Demain, A.L. (2017). Our microbes not only produce antibiotics, they also overproduce amino acids. *J Antibiot (Tokyo)*.
- Sasseti, C.M., Boyd, D.H., and Rubin, E.J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* 48, 77-84.
- Schafer, A., Tauch, A., Jager, W., Kalinowski, J., Thierbach, G., and Puhler, 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.

## References

---

- 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.
- Schmeing, T.M., and Ramakrishnan, V. (2009). What recent ribosome structures have revealed about the mechanism of translation. *Nature* 461, 1234-1242.
- Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 3, 1101-1108.
- Schrödinger, L.L.C. (2015). The PyMOL Molecular Graphics System, Version 1.8.
- 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, 194-205 e195.
- Servin, J.A., Herbold, C.W., Skophammer, R.G., and Lake, J.A. (2008). Evidence excluding the root of the tree of life from the actinobacteria. *Mol Biol Evol* 25, 1-4.
- Shoulders, M.D., and Raines, R.T. (2009). Collagen structure and stability. *Annu Rev Biochem* 78, 929-958.
- Siebert, D., and Wendisch, V.F. (2015). Metabolic pathway engineering for production of 1,2-propanediol and 1-propanol by *Corynebacterium glutamicum*. *Biotechnol Biofuels* 8, 91.
- Smith, A.M., Costello, M.S., Kettring, A.H., Wingo, R.J., and Moore, S.D. (2019). Ribosome collisions alter frameshifting at translational reprogramming motifs in bacterial mRNAs. *Proc Natl Acad Sci U S A* 116, 21769-21779.
- Song, J.Y., Jeong, H., Yu, D.S., Fischbach, M.A., Park, H.S., Kim, J.J., Seo, J.S., Jensen, S.E., Oh, T.K., Lee, K.J., *et al.* (2010). Draft genome sequence of *Streptomyces clavuligerus* NRRL 3585, a producer of diverse secondary metabolites. *J Bacteriol* 192, 6317-6318.
- Sorensen, M.A., and Pedersen, S. (1991). Absolute in vivo translation rates of individual codons in *Escherichia coli*. The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate. *J Mol Biol* 222, 265-280.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312-1313.
- Starosta, A.L., Lassak, J., Peil, L., Atkinson, G.C., Virumae, 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., Virumae, K., Buskirk, A., Tenson, T., Remme, J., Jung, K., *et al.* (2014b). A conserved proline triplet in Val-tRNA synthetase and the origin of elongation factor P. *Cell Rep* 9, 476-483.
- Strauch, E., Takano, E., Baylis, H.A., and Bibb, M.J. (1991). The stringent response in *Streptomyces coelicolor* A3(2). *Mol Microbiol* 5, 289-298.
- Takano, E., Gramajo, H.C., Strauch, E., Andres, N., White, J., and Bibb, M.J. (1992). Transcriptional regulation of the redD transcriptional activator gene accounts for growth-phase-dependent production of the antibiotic undecylprodigiosin in *Streptomyces coelicolor* A3(2). *Mol Microbiol* 6, 2797-2804.
- Takano, E., Kinoshita, H., Mersinias, V., Bucca, G., Hotchkiss, G., Nihira, T., Smith, C.P., Bibb, M., Wohlleben, W., and Chater, K. (2005). A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide biosynthetic gene cluster of *Streptomyces coelicolor*. *Mol Microbiol* 56, 465-479.
- Tanaka, Y., Takemoto, N., Ito, T., Teramoto, H., Yukawa, H., and Inui, M. (2014). Genome-wide analysis of the role of global transcriptional regulator GntR1 in *Corynebacterium glutamicum*. *J Bacteriol* 196, 3249-3258.
- 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 U S A* 115, 11072-11077.

## References

---

- Traxler, M.F., and Kolter, R. (2015). Natural products in soil microbe interactions and evolution. *Nat Prod Rep* 32, 956-970.
- Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 11, 2301-2319.
- Udaka, S. (1960). Screening method for microorganisms accumulating metabolites and its use in the isolation of *Micrococcus glutamicus*. *J Bacteriol* 79, 754-755.
- 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.
- Uguru, G.C., Stephens, K.E., Stead, J.A., Towle, J.E., Baumberg, S., and McDowall, K.J. (2005). Transcriptional activation of the pathway-specific regulator of the actinorhodin biosynthetic genes in *Streptomyces coelicolor*. *Mol Microbiol* 58, 131-150.
- UniProt, C. (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47, D506-D515.
- van Keulen, G., and Dyson, P.J. (2014). Production of specialized metabolites by *Streptomyces coelicolor* A3(2). *Adv Appl Microbiol* 89, 217-266.
- Vanhoof, G., Goossens, F., De Meester, I., Hendriks, D., and Scharpe, 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., Macosek, J., Guo, Z., Furst, M., Pfab, M., Frishman, D., *et al.* (2019). Switching the post-translational modification of translation elongation factor EF-P. *Front Microbiol* 10, 1148.
- Wang, C., Cai, H., Zhou, Z., Zhang, K., Chen, Z., Chen, Y., Wan, H., and Ouyang, P. (2014). Investigation of *ptsG* gene in response to xylose utilization in *Corynebacterium glutamicum*. *J Ind Microbiol Biotechnol* 41, 1249-1258.
- Wang, Z., Liu, J., Chen, L., Zeng, A.P., Solem, C., and Jensen, P.R. (2018). Alterations in the transcription factors GntR1 and RamA enhance the growth and central metabolism of *Corynebacterium glutamicum*. *Metab Eng* 48, 1-12.
- Wei, J., He, L., and Niu, G. (2018). Regulation of antibiotic biosynthesis in actinomycetes: Perspectives and challenges. *Synth Syst Biotechnol* 3, 229-235.
- Wendisch, V.F., Brito, L.F., Gil Lopez, M., Hennig, G., Pfeifenschneider, J., Sgobba, E., and Veldmann, K.H. (2016a). The flexible feedstock concept in Industrial Biotechnology: Metabolic engineering of *Escherichia coli*, *Corynebacterium glutamicum*, *Pseudomonas*, *Bacillus* and yeast strains for access to alternative carbon sources. *J Biotechnol* 234, 139-157.
- Wendisch, V.F., de Graaf, A.A., Sahm, H., and Eikmanns, B.J. (2000). Quantitative determination of metabolic fluxes during cointilization of two carbon sources: comparative analyses with *Corynebacterium glutamicum* during growth on acetate and/or glucose. *J Bacteriol* 182, 3088-3096.
- Wendisch, V.F., Jorge, J.M.P., Perez-Garcia, F., and Sgobba, E. (2016b). Updates on industrial production of amino acids using *Corynebacterium glutamicum*. *World J Microbiol Biotechnol* 32, 105.
- Wendrich, T.M., Blaha, G., Wilson, D.N., Marahiel, M.A., and Nierhaus, K.H. (2002). Dissection of the mechanism for the stringent factor RelA. *Mol Cell* 10, 779-788.
- 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.
- 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 EF-P. *Cell Rep* 11, 13-21.

## References

---

- Xiao, S., Xu, J., Chen, X., Li, X., Zhang, Y., and Yuan, Z. (2016). 3-Methyl-1-butanol biosynthesis in an engineered *Corynebacterium glutamicum*. *Mol Biotechnol* 58, 311-318.
- Xu, J., Zhang, J., Liu, D., and Zhang, W. (2016). Increased glucose utilization and cell growth of *Corynebacterium glutamicum* by modifying the glucose-specific phosphotransferase system (PTS(Glc)) genes. *Can J Microbiol* 62, 983-992.
- 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.